EQUILIBRIUM STATES OF THE PRESSURE FUNCTION FOR PRODUCTS OF MATRICES

DE-JUN FENG AND ANTTI KÄENMÄKI

ABSTRACT. Let $\{M_i\}_{i=1}^{\ell}$ be a non-trivial family of $d \times d$ complex matrices, in the sense that for any $n \in \mathbb{N}$, there exists $i_1 \cdots i_n \in \{1, \ldots, \ell\}^n$ such that $M_{i_1} \cdots M_{i_n} \neq \mathbf{0}$. Let $P \colon (0, \infty) \to \mathbb{R}$ be the pressure function of $\{M_i\}_{i=1}^{\ell}$. We show that for each $q > 0$, there are at most d ergodic q-equilibrium states of P, and each of them satisfies certain Gibbs property.

1. INTRODUCTION AND RESULTS

In this paper, we study the thermodynamic formalism for matrix products. We will characterize the structure of equilibrium states of pressure functions, and also examine the Gibbs properties of such states. This work was first carried out in [11] in the case that the involved matrices are non-negative and satisfy a kind of irreducibility. Some applications were given in the multifractal analysis of the top Lyapunov exponents of matrix products [11, 6, 8] (see also [10]). In this paper, we will consider arbitrary complex matrices.

Let (Σ, σ) be the one-sided full shift over the alphabet $\{1, \ldots, \ell\}$ (cf. [1]) and let $\{M_i\}_{i=1}^{\ell}$ be a family of $d \times d$ complex matrices. For $q > 0$, we define

$$(1.1) \qquad P(q) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{J \in \Sigma_n} \|M_J\|^q,$$

where Σ_n is the collection of all words of length n over $\{1, \ldots, \ell\}$, $M_J = M_{j_1} \cdots M_{j_n}$ for $J = j_1 \cdots j_n$, and $\|\cdot\|$ is the standard matrix norm. By sub-additivity, the above limit exists and takes values in the set $\mathbb{R} \cup \{-\infty\}$. The function P is called the *pressure function* of $\{M_i\}_{i=1}^{\ell}$. It plays an important role in the multifractal analysis of Lyapunov exponents of matrices [11, 6, 8]. Moreover, it is closely related to the dimension theory of self-affine sets and measures [4, 15].

Denote the collection of all σ-invariant Borel probability measures on Σ by $\mathcal{M}_\sigma(\Sigma)$. Endow $\mathcal{M}_\sigma(\Sigma)$ with the weak-star topology. For $\mu \in \mathcal{M}_\sigma(\Sigma)$, we define

$$(1.2) \qquad M_*(\mu) = \lim_{n \to \infty} \frac{1}{n} \sum_{J \in \Sigma_n} \mu([J]) \log \|M_J\|,$$

where $[J]$ denotes the n-th cylinder $\{x = (x_i)_{i=1}^\infty \in \Sigma : x_1 \cdots x_n = J\}$ in Σ. The term $M_*(\mu)$ is called the *Lyapunov exponent of* $\{M_i\}_{i=1}^\ell$ *with respect to* μ. It also takes values in the set $\mathbb{R} \cup \{-\infty\}$. The following variational principle for P was proved in [3] in a more general sub-additive setting:

$$(1.3) \qquad P(q) = \sup\{qM_*(\mu) + h(\mu) : \ \mu \in \mathcal{M}_\sigma(\Sigma)\},$$

where $h(\mu)$ denotes the measure-theoretic entropy of μ with respect to σ (cf. [19]). We remark that (1.3) was proved earlier in [7, 15] when the matrices are non-negative or invertible, respectively. For given $q > 0$, let

$$(1.4) \qquad \mathcal{I}_q = \{\mu \in \mathcal{M}_\sigma(\Sigma) : \ P(q) = qM_*(\mu) + h(\mu)\}.$$

Each element μ in $\mathcal{I}_q$ is called a *q-equilibrium state of P*. Since both $M_*(\cdot)$ and $h(\cdot)$ are upper semi-continuous on $\mathcal{M}_\sigma(\Sigma)$, $\mathcal{I}_q$ is a non-empty closed convex subset of $\mathcal{M}_\sigma(\Sigma)$. In particular, $\mathcal{I}_q$ contains ergodic elements (each extreme point of $\mathcal{I}_q$ is an ergodic measure).

Our main purpose is to characterize the structure of $\mathcal{I}_q$. This question was partially raised from [16]. A complete characterization is given in Theorem 1.7. In the following, we shall present the setting and results. Proofs of the results are postponed until §2.

Definition 1.1. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. A family of $d \times d$ matrices $\{M_i\}_{i=1}^\ell$ with entries in $\mathbb{F}$ is said to be *irreducible over* $\mathbb{F}^d$ if there is no non-zero proper linear subspace V of $\mathbb{F}^d$ such that $M_i V \subseteq V$ for all $i \in \{1, \ldots, \ell\}$.

The above definition is adopted from [2, p. 48]. If $\{M_i\}_{i=1}^\ell$ is irreducible over $\mathbb{F}^d$, then there exist $D > 0$ and $k \in \mathbb{N}$ such that for any words $I, J \in \Sigma^* = \bigcup_{n=1}^\infty \{1, \ldots, \ell\}^n$, there exists a word K in $\bigcup_{n=1}^k \{1, \ldots, \ell\}^n$ such that

$$(1.5) \qquad \|M_{IKJ}\| \geq D\|M_I\|\|M_J\|.$$

For a proof, see [8, Proposition 2.8]. This property is crucial in the proof of the following proposition.

Proposition 1.2. *Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$, and $\{M_i\}_{i=1}^{\ell}$ a family of $d \times d$ matrices with entries in $\mathbb{F}$. If $\{M_i\}_{i=1}^{\ell}$ is irreducible over $\mathbb{F}^d$, then for each $q > 0$, P has a unique q-equilibrium state μ_q. Furthermore, μ_q has the following Gibbs property:*

(1.6) $$C^{-1}\exp(-nP(q))\|M_J\|^q \leq \mu_q([J]) \leq C\exp(-nP(q))\|M_J\|^q$$

for all $n \in \mathbb{N}$ and $J \in \Sigma_n$. Moreover, P is differentiable over $(0, \infty)$ and $P'(q) = M_*(\mu_q)$ for $q > 0$.

Remark 1.3. Proposition 1.2 is an analogue of Bowen's theory about the equilibrium state of Hölder continuous additive potentials (cf. [1]). See [18, 19] for backgrounds and more details about the classical thermodynamic formalism of additive potentials. Proposition 1.2 was first proved in [11] for non-negative matrices under a different irreducibility assumption (that is, there exists $r \in \mathbb{N}$ so that $\sum_{i=1}^{r}(M_1 + \cdots + M_\ell)^r$ is a strictly positive matrix). An extension was recently given in [9, Theorem 5.5] to certain sub-additive potentials.

Let us next consider the non-irreducibility case. Denote the $n \times m$ zero matrix by $\mathbf{0}_{n \times m}$.

Proposition 1.4. *Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$, and $\{M_i\}_{i=1}^{\ell}$ a family of $d \times d$ matrices with entries in $\mathbb{F}$. Then there exist an invertible $d \times d$ matrix T, $t \in \{1, \ldots, d\}$, and positive integers $d_1, \ldots, d_t$ with $d = d_1 + \cdots + d_t$ such that for every $i \in \{1, \ldots, \ell\}$ the product $T^{-1}M_iT$ is a partitioned matrix of the form*

(1.7) $$T^{-1}M_iT = \left(A_i^{(j,k)}\right)_{1 \leq j,k \leq t},$$

where $A_i^{(j,k)}$, $j, k \in \{1, \ldots, t\}$, satisfy the following two properties:

(i) *$A_i^{(j,k)}$ is a $d_j \times d_k$ matrix and $A_i^{(j,k)} = \mathbf{0}_{d_j \times d_k}$ when $j > k$.*
(ii) *For any $j \in \{1, \ldots, t\}$, either the family $\{A_i^{(j,j)}\}_{i=1}^{\ell}$ is irreducible over $\mathbb{F}^{d_j}$, or $A_i^{(j,j)} = \mathbf{0}_{d_j \times d_j}$ for all $i \in \{1, \ldots, \ell\}$.*

Considering the partition (1.7) in the above proposition, we set

$$\Lambda = \Lambda(\{M_i\}_{i=1}^{\ell}) = \{j \in \{1, \ldots, t\} : \{A_i^{(j,j)}\}_{i=1}^{\ell} \text{ is irreducible over } \mathbb{F}^{d_j}\}.$$

Remark 1.5. It is possible that $\Lambda = \emptyset$. For instance, this is the case for $\{M_i\}_{i=1}^{2}$, where

$$M_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

Anyhow, it holds that $\Lambda = \emptyset$ if and only if there is $k \in \mathbb{N}$ such that $M_{i_1} \cdots M_{i_n} = \mathbf{0}_{d \times d}$ for all $n \geq k$ and $i_1 \cdots i_n \in \{1, \ldots, \ell\}^n$. Observe first that $T^{-1} M_{i_1} \cdots M_{i_n} T$ is a partitioned matrix of the form $(B^{(j,k)})_{1 \leq j,k \leq t}$, where

$$
\begin{aligned}
B^{(j,k)} &= \sum_{1 \leq y_1, \ldots, y_{n-1} \leq t} A_{i_1}^{(j,y_1)} A_{i_2}^{(y_1,y_2)} \cdots A_{i_n}^{(y_{n-1},k)} \\
&= \sum_{j \leq y_1 \leq y_2 \leq \cdots \leq y_{n-1} \leq k} A_{i_1}^{(j,y_1)} A_{i_2}^{(y_1,y_2)} \cdots A_{i_n}^{(y_{n-1},k)}
\end{aligned}
\tag{1.8}
$$

is a $d_j \times d_k$ matrix. According to (ii) of Proposition 1.4, $\Lambda = \emptyset$ implies $A_i^{(j,j)} = \mathbf{0}_{d_j \times d_j}$ for all $i \in \{1, \ldots, \ell\}$ and $j \in \{1, \ldots, t\}$. Hence $M_{i_1} \cdots M_{i_n} = \mathbf{0}_{d \times d}$ for all $n > t$ by (1.8). To see the converse, assume contrarily that $\{A_i^{(j,j)}\}_{i=1}^{\ell}$ is irreducible over $\mathbb{C}^{d_j}$ for some $j \in \{1, \ldots, t\}$. It follows now from (1.5) that for every $n \in \mathbb{N}$ there exists a word $i_1 \cdots i_n$ such that $A_{i_1}^{(j,j)} \cdots A_{i_n}^{(j,j)} \neq \mathbf{0}_{d_j \times d_j}$ and, consequently, $M_{i_1} \cdots M_{i_n} \neq \mathbf{0}_{d \times d}$.

Definition 1.6. A family of $d \times d$ complex matrices $\{M_i\}_{i=1}^{\ell}$ is called *non-trivial* if $\Lambda \neq \emptyset$, or equivalently, for each $n \in \mathbb{N}$, there exists $I \in \{1, \ldots, \ell\}^n$ such that $M_I \neq \mathbf{0}_{d \times d}$.

In the following, we always assume that $\{M_i\}_{i=1}^{\ell}$ is non-trivial. If $j \in \Lambda$, then the pressure function of $\{A_i^{(j,j)}\}_{i=1}^{\ell}$ is denoted by P_j and the Lyapunov exponent of $\{A_i^{(j,j)}\}_{i=1}^{\ell}$ with respect to μ is denoted by $A_*^{(j)}(\mu)$. The following is the main result of our paper.

Theorem 1.7. *In the above general setting, it holds that*

(i) $M_*(\mu) = \max\{A_*^{(j)}(\mu) : j \in \Lambda\}$ *for each ergodic measure* $\mu \in \mathcal{M}_\sigma(\Sigma)$.

(ii) P *is a real-valued convex function on* $(0, \infty)$, *and* $P(q) = \max\{P_j(q) : j \in \Lambda\}$ *for all* $q > 0$.

(iii) *if* $q > 0$ *and* $\mu_{j,q}$, $j \in \Lambda$, *is the unique q-equilibrium state for* P_j, *then*

$$
\mathcal{I}_q = \mathrm{conv}\{\mu_{j,q} : P_j(q) = P(q)\},
$$

where $\mathrm{conv}(A)$ *is the convex hull of* A.

Remark 1.8. The equality in (i) of Theorem 1.7 may fail for non-ergodic measures of $\mathcal{M}_\sigma(\Sigma)$. For instance, consider $\{M_i\}_{i=1}^2$, where $M_1 = \mathrm{diag}(1, 2)$ and $M_2 = \mathrm{diag}(3, 2)$. Let $\mu_1 = \delta_{1\infty}$, $\mu_2 = \delta_{2\infty}$ (here δ_x denotes the Dirac measure at x), and $\mu = p\mu_1 + (1 - p)\mu_2$ for some $0 < p < 1$. It is easy to check that

$$
M_*(\mu_1) = \log 2, \quad A_*^{(1)}(\mu_1) = 0, \quad A_*^{(2)}(\mu_1) = \log 2
$$

and
$$M_*(\mu_2) = \log 3, \quad A_*^{(1)}(\mu_2) = \log 3, \quad A_*^{(2)}(\mu_2) = \log 2.$$
Since $M_*(\cdot)$, $A_*^{(1)}(\cdot)$, and $A_*^{(2)}(\cdot)$ are affine on $\mathcal{M}_\sigma(\Sigma)$, we have
$$M_*(\mu) = p \log 2 + (1-p) \log 3, \quad A_*^{(1)}(\mu) = (1-p) \log 3, \quad A_*^{(2)}(\mu) = \log 2,$$
and thus, $M_*(\mu) > \max\{A_*^{(i)}(\mu) : i \in \{1, 2\}\}$.

Remark 1.9. The pressure function for products of matrices has been studied in the literature under some stronger conditions. Let $\{M_i\}_{i=1}^\ell$ be a family of real invertible matrices. Assume that $\{M_i\}_{i=1}^\ell$ satisfies the strong irreducibility and contraction conditions (cf. [2, 13]). Guivarc'h and Le Page showed in [13, Theorem 8.8] that the pressure function P of $\{M_i\}_{i=1}^\ell$ corresponds to the logarithm of the spectral radius of certain Ruelle transfer operator and moreover, P is real analytic on $(0, \infty)$, and it can be extended to an analytic function on $\{z \in \mathbb{C} : \Re z > 0\}$. This strengthens an early result of Le Page [17].

2. Proofs of the results

This section is dedicated to the proof of Theorem 1.7. For the convenience of the reader we shall also present complete proofs for Propositions 1.2 and 1.4.

Proof of Proposition 1.2. Let $q > 0$. Define a sequence of probability measures $(\nu_{n,q})_{n \geq 1}$ on Σ so that
$$\nu_{n,q}([I]) = \frac{\|M_I\|^q}{\sum_{J \in \Sigma_n} \|M_J\|^q}$$
for all $I \in \Sigma_n$. Let ν_q be a limit point of the sequence $(\nu_{n,q})_{n \geq 1}$ in the weak topology. Furthermore, let μ_q be a limit point of the sequence
$$\left(\frac{1}{n} \sum_{j=0}^{n-1} \nu_q \circ \sigma^{-j} \right)_{n \geq 1}$$
in the weak topology. Using (1.5) and a proof essentially identical to that of [11, Theorem 3.2], we see that $\mu_q \in \mathcal{M}_\sigma(\Sigma)$ is ergodic and has the Gibbs property (1.6). Thus
$$qM_*(\mu_q) + h(\mu_q) \geq \lim_{n \to \infty} \frac{1}{n} \sum_{J \in \Sigma_n} \mu_q([J]) \log\left(C^{-1} \exp(nP(q)) \mu_q([J])\right)$$
$$- \lim_{n \to \infty} \frac{1}{n} \sum_{J \in \Sigma_n} \mu_q([J]) \log \mu_q([J]) = P(q).$$

Recalling (1.3), this implies $\mu_q \in \mathcal{I}_q$.

Applying (1.6) and the ergodicity of μ_q, and using an identical argument as in [1, proof of Theorem 1.22] (or using [16, Theorem 3.6]), we see that μ_q is the unique element in $\mathcal{I}_q$. According to this uniqueness, we have $P'(q) = M_*(\mu_q)$, which follows from the Ruelle-type derivative formula of pressures obtained in [7, Theorem 1.2]:

$$P'(q-) = \inf\{\mathcal{M}_*(\mu) : \mu \in \mathcal{I}_q\}, \quad P'(q+) = \sup\{\mathcal{M}_*(\mu) : \mu \in \mathcal{I}_q\}.$$

We remark that although [7, Theorem 1.2] only deals with non-negative matrices, the proof given there works for arbitrary matrices. Alternatively, to show that $P'(q) = M_*(\mu_q)$, we may apply (1.6) and the ergodicity of μ_q, and follow [14, proof of Theorem 2.1] (see also [16, Theorem 4.4]). $\qquad\square$

Proof of Proposition 1.4. We prove the proposition by induction on d. Clearly the proposition is true when $d = 1$. Assuming there exists an integer p so that the proposition is true for all $d \leq p$, we show below that it remains true for $d = p + 1$. Let $L(n, m)$ be the collection of all $n \times m$ matrices with entries in $\mathbb{F}$.

If $\{M_i\}_{i=1}^\ell$ is irreducible over $\mathbb{F}^d$, we simply take $t = 1$ and have nothing else to prove. We may thus assume that $\{M_i\}_{i=1}^\ell$ is reducible, that is, there exists a non-zero proper linear space V of $\mathbb{F}^d$ such that $M_i V \subset V$. If we let $v = \dim V$, then $1 \leq v$ and $d - v \leq d - 1 = p$. We choose an invertible linear map $T_1 : \mathbb{F}^d \to \mathbb{F}^d$ such that $T_1(\mathbb{F}^v \times \{0\}) = V$. Then for each $i \in \{1, \ldots, \ell\}$ there exist $E_i \in L(v, v)$, $B_i \in L(v, d - v)$, $D_i \in L(d - v, d - v)$ so that

$$T_1^{-1} M_i T_1 = \begin{pmatrix} E_i & B_i \\ \mathbf{0}_{(d-v)\times v} & D_i \end{pmatrix}.$$

Now by the induction assumption, there exist invertible matrices $T_2 \in L(v, v)$ and $T_3 \in L(d - v, d - v)$ such that $(T_2^{-1} E_i T_2)_{i=1}^\ell$ and $(T_3^{-1} D_i T_3)_{i=1}^\ell$ have the desired partitioned form for all $i \in \{1, \ldots, \ell\}$. It follows that

$$T_4 = T_1 \begin{pmatrix} T_2 & \mathbf{0}_{v\times(d-v)} \\ \mathbf{0}_{(d-v)\times v} & T_3 \end{pmatrix}$$

is an invertible $d \times d$ matrix and

$$T_4^{-1} M_i T_4 = \begin{pmatrix} T_2^{-1} E_i T_2 & T_2^{-1} B_i T_3 \\ \mathbf{0}_{(d-v)\times v} & T_3^{-1} D_i T_3 \end{pmatrix}$$

has the desired partitioned form for all $i \in \{1, \ldots, \ell\}$. $\qquad\square$

Before proving Theorem 1.7, we shall first prove the following auxiliary result.

Proposition 2.1. *Let $(X, \mathcal{F}, \mu)$ be a probability space and $T : X \to X$ an ergodic measure-preserving transformation. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of non-negative Borel measurable functions on X such that $\sup_{x \in X} f_1(x) < \infty$ and*

$$(2.1) \qquad\qquad f_{n+m}(x) \le f_m(x) f_n(T^m x)$$

for all $m, n \in \mathbb{N}$ and $x \in X$. If $\epsilon > 0$ and $\alpha = \lim_{n \to \infty} (1/n) \int \log f_n \, d\mu$, then the following claims hold:

   (i) *If $\alpha \ne -\infty$, then for μ-almost every $x \in X$, there exists a positive integer $n_0(x)$ such that*

$$(2.2) \qquad\qquad |\log f_n(T^m x) - n\alpha| \le (n + m)\epsilon$$

   for all $n \ge n_0(x)$ and $m \in \mathbb{N}$.
   (ii) *If $\alpha = -\infty$, then for any $N > 0$ and μ-almost every $x \in X$, there exists a positive integer $n_0(x)$ such that*

$$(2.3) \qquad\qquad \log f_n(T^m x) \le -Nn + (n + m)\epsilon$$

   for all $n \ge n_0(x)$ and $m \in \mathbb{N}$.

Proof. We only prove (i). The proof of (ii) is similar.

Assume that $\alpha \in \mathbb{R}$. Let $\epsilon > 0$ and take $0 < \delta < \epsilon/4$. By the Kingman's sub-additive ergodic theorem, for μ-almost every $x \in X$, there exists $n_0(x)$ such that

$$|\log f_n(x) - n\alpha| \le n\delta$$

for all $n \ge n_0(x)$, and

$$|\log f_m(x) - m\alpha| \le (n_0(x) + m)\delta$$

for all $m \in \mathbb{N}$. Hence by (2.1), we have for $n \ge n_0(x)$ and $m \in \mathbb{N}$,

$$\begin{aligned}
\log f_n(T^m x) &\ge \log f_{n+m}(x) - \log f_m(x) \\
(2.4) \qquad &\ge (n + m)(\alpha - \delta) - m(\alpha + \delta) - n_0(x)\delta \\
&\ge n\alpha - 2(n + m)\delta \ge n\alpha - (n + m)\epsilon.
\end{aligned}$$

To see the opposite inequality, take k large enough such that $|\beta - \alpha| < \delta$, where

$$\beta = \frac{1}{k} \int \log f_k \, d\mu.$$

Since $\{f_n(x)\}_{n=1}^\infty$ is sub-multiplicative, by [3, Lemma 2.2], we have for any $n \geq 2k$,

$$(f_n(x))^k \leq C^{2k^2} \prod_{j=0}^{n-k} f_k(T^j x)$$

for all $x \in X$, where $C = \max\{1, \sup_{x \in \Sigma} f_1(x)\}$. It follows that for $n \geq 2k$ and $m \in \mathbb{N}$,

$$\log f_n(T^m x) \leq 2k \log C + \sum_{i=0}^{n-k+m} \frac{1}{k} \log f_k(T^i x) - \sum_{i=0}^{m-1} \frac{1}{k} \log f_k(T^i x).$$

Applying the Birkhoff ergodic theorem to the function $\frac{1}{k} \log f_k$, and combining it with the above inequality, we find for μ-almost every $x \in X$ an integer $\tilde{n}_0(x) \geq 2k\delta^{-1} \log C$ such that

$$\begin{aligned}
\log f_n(T^m x) &\leq n\delta + (n - k + m)(\beta + \delta) - m(\beta - \delta) \\
&\leq n\beta + 2(n + m)\delta \leq n\alpha + 3(n + m)\delta \\
&\leq n\alpha + (n + m)\epsilon
\end{aligned}$$

for all $n \geq \tilde{n}_0(x)$ and $m \in \mathbb{N}$. This together with (2.4) yields (2.2). $\qquad\square$

As a direct corollary of Proposition 2.1, we have the following.

Corollary 2.2. *Under the assumptions of Proposition 2.1, for any $\epsilon, N > 0$ and for μ-almost every $x \in X$, there is $C(x) > 0$ such that*

$$|f_n(T^m x)| \leq C(x) \exp(n \max\{\alpha, -N\}) \exp((n + m)\epsilon)$$

for all $n, m \in \mathbb{N}$.

Proof of Theorem 1.7. We only need to prove part (i), since parts (ii) and (iii) follow immediately from (i), the variational principle (1.3), and Proposition 1.2.

Fix an ergodic measure $\mu \in \mathcal{M}_\sigma(\Sigma)$. The direction $M_*(\mu) \geq \max\{A_*^{(j)}(\mu) : j \in \Lambda\}$ follows from the fact that

$$\|A_{i_1}^{(j,j)} \cdots A_{i_n}^{(j,j)}\| \leq \|T^{-1} M_{i_1} \cdots M_{i_n} T\| \leq \|T^{-1}\| \|T\| \|M_{i_1} \cdots M_{i_n}\|$$

for any $j \in \Lambda$ and $i_1, \ldots, i_n \in \{1, \ldots, \ell\}$. We only need to prove the other direction.

By Furstenberg-Kesten's theorem [12] on random matrices, or Kingman's sub-additive ergodic theorem (see e.g. [19]), we have for μ-almost every $x = (x_i)_{i=1}^\infty \in \Sigma$,

$$(2.5) \qquad\qquad \lim_{n \to \infty} \frac{1}{n} \log \|M_{x_1 \cdots x_n}\| = M_*(\mu).$$

For any $i \in \{1, \ldots, t\}$, define a sequence $\{f_n^{(j)}\}_{n=1}^{\infty}$ of non-negative functions on Σ by setting

$$f_n^{(j)}(x) = \|A_{x_1}^{(j,j)} \cdots A_{x_n}^{(j,j)}\|$$

for all $x = (x_i)_{i=1}^{\infty} \in \Sigma$. Let $\epsilon, N > 0$. Apply Corollary 2.2 for $\{f_n^{(j)}\}_{n=1}^{\infty}$ to obtain that, for μ-almost every $x = (x_i)_{i=1}^{\infty} \in \Sigma$, there exists $C(x) \geq 1$ such that

$$(2.6) \quad \begin{aligned} \|A_{x_{m+1}x_{m+2}\cdots x_{m+n}}^{(j,j)}\| &\leq C(x) \exp(n \max\{A_*^{(j)}(\mu), -N\}) \exp((n+m)\epsilon) \\ &\leq C(x) \exp(n \max\{W, -N\}) \exp((n+m)\epsilon), \end{aligned}$$

for all $j \in \{1, \ldots, t\}$ and $n, m \in \mathbb{N}$, where

$$W = \max\{A_*^{(j)}(\mu) : j \in \Lambda\}.$$

For the rest of the proof, we take a point $x = (x_i)_{i=1}^{\infty} \in \Sigma$ such that both (2.5) and (2.6) hold for x.

Fix $n \in \mathbb{N}$. According to (1.8), $T^{-1}M_{x_1\cdots x_n}T$ is a partitioned matrix of the form $(B^{(j,k)})_{1 \leq j,k \leq t}$, where each $B^{(j,k)}$ is a $d_j \times d_k$ matrix given by

$$(2.7) \quad B^{(j,k)} = \sum_{j \leq y_1 \leq y_2 \leq \cdots \leq y_{n-1} \leq k} A_{x_1}^{(j,y_1)} A_{x_2}^{(y_1,y_2)} \cdots A_{x_n}^{(y_{n-1},k)}.$$

It is easy to check that the number of words $y_1 y_2 \cdots y_{n-1} \in \{1, \ldots, t\}^{n-1}$, satisfying the restriction $j \leq y_1 \leq y_2 \leq \cdots \leq y_{n-1} \leq k$, is bounded above by $h(n) = (2n)^t$. Furthermore, each such a word $jy_1y_2\cdots y_{n-1}k$ can be written as $a_1^{n_1} a_2^{n_2} \cdots a_s^{n_s}$, where $s \in \{1, \ldots, t\}$, $j = a_1 < \cdots < a_s = k$, and $n_1, \ldots, n_s \in \mathbb{N}$ with $n_1 + \cdots + n_s = n+1$. Hence

$$(2.8) \quad A_{x_1}^{(j,y_1)} A_{x_2}^{(y_1,y_2)} \cdots A_{x_n}^{(y_{n-1},k)} = W_1 A_{x_{n_1}}^{(a_1,a_2)} W_2 A_{x_{n_1+n_2}}^{(a_2,a_3)} \cdots W_{s-1} A_{x_{n_1+n_2+\cdots+n_{s-1}}}^{(a_{s-1},a_s)} W_s,$$

where

$$W_i = \begin{cases} \mathbf{I}_{d_{a_i} \times d_{a_i}} & \text{if } n_i = 1, \\ A_{x_{n_0+\cdots+n_{i-1}+1}}^{(a_i,a_i)} \cdots A_{x_{n_0+\cdots+n_i-1}}^{(a_i,a_i)} & \text{if } n_i > 1 \end{cases}$$

for all $i \in \{1, \ldots, s\}$. Here $\mathbf{I}_{d \times d}$ is the $d \times d$ identity matrix and $n_0 = 0$. Observe that (2.6) gives

$$\|W_i\| \leq C(x) \exp\left((n_i - 1) \max\{W, -N\}\right) \exp\left((n_1 + \cdots + n_i - 1)\epsilon\right)$$

for all $i \in \{1, \ldots, s\}$. Hence, by (2.8), we have

$$
\begin{aligned}
\|A_{x_1}^{(j,y_1)} A_{x_2}^{(y_1,y_2)} \cdots A_{x_n}^{(y_{n-1},k)}\| &\leq L^{s-1} \prod_{i=1}^{s} \|W_i\| \\
&\leq L^{s-1} C(x)^s \exp((n+1-s)\max\{W, -N\}) \exp(ns\epsilon) \\
&\leq DL^t C(x)^t \exp(n\max\{W, -N\}) \exp(nt\epsilon),
\end{aligned}
$$

(2.9)

where

$$
\begin{aligned}
L &= 1 + \max\{\|A_i^{(j_1,j_2)}\| : \; j_1, j_2 \in \{1, \ldots, t\} \text{ and } i \in \{1, \ldots, \ell\}\}, \\
D &= \max\{1, \exp((t+1)\max\{W, -N\})\}.
\end{aligned}
$$

Therefore, by (2.7)–(2.9), we have the estimate

$$
\begin{aligned}
\|T^{-1} M_{x_1 \cdots x_n} T\| &\leq t^2 \max\{\|B^{(j,k)}\| : j, k \in \{1, \ldots, t\}\} \\
&\leq t^2 h(n) DL^t C(x)^t \exp(n(\max\{W, -N\})) \exp(nt\epsilon)
\end{aligned}
$$

for all $n \in \mathbb{N}$. Combining this estimate and (2.5) yields

$$
M_*(\mu) = \lim_{n \to \infty} \frac{1}{n} \log \|T^{-1} M_{x_1 \cdots x_n} T\| \leq \max\{W, -N\} + t\epsilon.
$$

Letting $N \to \infty$ and $\epsilon \to 0$, we get

$$
M_*(\mu) \leq W = \max\{A_*^{(j)}(\mu) : \; j \in \Lambda\},
$$

which finishes the proof of part (i) of Theorem 1.7. $\square$

3. Extensions and remarks

For an invertible matrix $M \in \mathbb{R}^{d \times d}$, following [4], we define the singular value function of M as

$$
\phi^q(M) = \alpha_1(M) \cdots \alpha_k(M) \alpha_{k+1}(M)^{q-k},
$$

where $0 \leq q < d$, k is the integral part of t, and $\alpha_i(M)$ is the i-th largest singular value of M. For $q > d$, we put $\phi^q(M) = |\det(M)|^{q/d}$. It is known (see [4, Lemma 2.1]) that ϕ^q is sub-multiplicative in the sense that

$$
\phi^q(M_1 M_2) \leq \phi^q(M_1)\phi^q(M_2)
$$

for any two invertible matrices $M_1, M_2 \in \mathbb{R}^{d \times d}$. For a given family of invertible matrices $\{M_i\}_{i=1}^{\ell} \subset \mathbb{R}^{d \times d}$, similar to (1.1), we define

(3.10) $$P^{\phi}(q) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{J \in \Sigma_n} \phi^q(M_J).$$

For $\mu \in \mathcal{M}_\sigma(\Sigma)$, we define

$$(3.11) \qquad \phi_*^q(\mu) = \lim_{n \to \infty} \frac{1}{n} \sum_{J \in \Sigma_n} \mu(J) \log \phi^q(M_J).$$

Then by [15, Theorem 2.6], or more generally by [3, Theorem 1.1], we have the following variational principle

$$P^\phi(q) = \max\{\phi_*^q(\mu) + h(\mu) : \ \mu \in \mathcal{M}_\sigma(\Sigma)\}.$$

Similarly we can study the structure of the equilibrium states of $P^\phi(q)$. It is easy to see that Theorem 1.7 remains true for $P^\phi(q)$ when $0 \leq q \leq 1$ or $q \geq d-1$. Observe also that it is true when q is an integer: if $M^{\wedge q}$ is the q-th exterior product of $M \in \mathbb{R}^{d \times d}$ (i.e. the $\binom{d}{q} \times \binom{d}{q}$ matrix whose entries are the $q \times q$ minors of M), then

$$\alpha_1(M^{\wedge q}) = \alpha_1(M) \cdots \alpha_q(M) = \phi^q(M).$$

This gives a partial answer to [16, Question 6.3].

Question 3.1. *When using* (3.10) *and* (3.11) *instead of* (1.1) *and* (1.2), *does Theorem 1.7 hold for* $q \in [1, d-1] \setminus \mathbb{N}$?

We remark that some assumption was given in [5] so that an analogue of (1.5) (where $\|\cdot\|$ is replaced by $\phi^q(\cdot)$) holds; and for such case, an analogue of Proposition 1.2 holds for P^ϕ (cf. [9, Theorem 5.5]).

References

[1] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, *Lecture notes in Math.*, No. **470**, Springer-Verlag, 1975.

[2] P. Bougerol and J. Lacroix, *Products of random matrices with applications to Schrödinger operators*, Birkhäuser, 1985.

[3] Y. L. Cao, D. J. Feng and W. Huang, The thermodynamical formalism for submultiplicative potentials. *Discrete Contin. Dyn. Syst.* **20** (2008), 639–657.

[4] K. J. Falconer, The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 339–350.

[5] K. Falconer and A. Sloan, Continuity of subadditive pressure for self-affine sets, *Real Analysis Exchange*, to appear.

[6] D. J. Feng, Lyapunov exponents for products of matrices and multifractal analysis, Part I: Positive matrices, *Israel J. Math.* **138** (2003), 353–376.

[7] D. J. Feng, The variational principle for products of non-negative matrices. *Nonlinearity* **17** (2004), 447–457.

[8] D. J. Feng, Lyapunov exponents for products of matrices and multifractal analysis, Part II: General matrices, *Israel J. Math.* **170** (2009), 355-394.

[9] D. J. Feng, Weighted equilibrium states for factor maps between subshifts. Preprint. arXiv:0909.4250v1 [math.DS]

[10] D. J. Feng and W. Huang, Lyapunov spectrum of asymptotically sub-additive potentials. Preprint. arXiv:0905.2680v1 [math.DS]

[11] D. J. Feng and K. S. Lau, The pressure function for products of non-negative matrices. *Math. Res. Lett.* **9** (2002), 363–378.

[12] H. Furstenberg and H. Kesten, Products of random matrices. *Ann. Math. Statist.* **31** (1960), 457-468.

[13] Y. Guivarc'h and E. Le Page, Simplicité de spectres de Lyapounov et propriété d'isolation spectrale pour une famille d'opérateurs de transfert sur l'espace projectif. *Random walks and geometry*, 181–259, Walter de Gruyter GmbH & Co. KG, Berlin, 2004.

[14] Y. Heurteaux, Estimations de la dimension inférieure et de la dimension supérieure des mesures. *Ann. Inst. Henri Poincaré* **34** (1998), 309–338.

[15] A. Käenmäki. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math.* **29** (2004), 419–458.

[16] A. Käenmäki and M. Vilppolainen. Dimension and measures on sub-self-affine sets. *Monatsh. Math.*, doi:10.1007/s00605-009-0144-9.

[17] E. Le Page, Théorèmes limites pour les produits de matrices aléatoires. *Lecture Notes in Math.*, **928**, Springer, Berlin-New York, 1982.

[18] D. Ruelle, *Thermodynamic formalism. The mathematical structures of classical equilibrium statistical mechanics.* Encyclopedia of Mathematics and its Applications, 5. Addison-Wesley Publishing Co., Reading, Mass., 1978.

[19] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, 1982.

DEPARTMENT OF MATHEMATICS, THE CHINESE UNIVERSITY OF HONG KONG, SHATIN, HONG KONG, P. R. CHINA

E-mail address: djfeng@math.cuhk.edu.hk

DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. BOX 35 (MAD), FI-40014 UNIVERSITY OF JYVÄSKYLÄ, FINLAND

E-mail address: antti.kaenmaki@jyu.fi