Transfer Operators and Limit Laws for Typical Cocycles

Kiho Park¹, Mark Piraino²

¹ School of Mathematics, Korea Institute for Advanced Study, Seoul, South Korea.
E-mail: kihopark12@kias.re.kr
² Department of Mathematics, Northwestern University, Evanston, IL 60208, USA.
E-mail: mark.piraino@northwestern.edu

Received: 25 September 2020 / Accepted: 13 December 2021
Published online: 7 January 2022 – © The Author(s), under exclusive licence to Springer-Verlag GmbH
Germany, part of Springer Nature 2022

Abstract: We show that typical cocycles (in the sense of Bonatti and Viana) over irreducible subshifts of finite type obey several limit laws with respect to the unique equilibrium states for Hölder potentials. These include the central limit theorem and the large deviation principle. We also establish the analytic dependence of the top Lyapunov exponent on the underlying equilibrium state. The transfer operator and its spectral properties play key roles in establishing these limit laws.

1. Introduction

There is a vast literature on the asymptotic properties for products of independent random matrices, including the strong laws of large numbers [FK60], the simplicity of the top Lyapunov subspace [FK83,GLP04], various limit laws [LP82], and the regularity of the Lyapunov exponent with respect to the data [LP89, DK16, Section 3, BNV17]. The goal of this article is to extend some of these results to the case in which the matrices are not independent but instead driven by a sufficiently nice dynamical system.

Let $T$ be an irreducible adjacency matrix and $\Sigma_T$ be the two-sided subshift of finite type defined by $T$. For a given Hölder potential $\hat{\psi} : \Sigma_T \to \mathbb{R}$ and its unique equilibrium state $\hat{\mu} = \hat{\mu}_{\hat{\psi}}$, it is well known that $\hat{\mu}$ satisfies various ergodic and statistical properties; see [Bow75,Bal00]. These systems generalize the i.i.d. setting in the sense that they have conditional probabilities which may depend on the entire past. The assumption that $\hat{\psi}$ is Hölder is equivalent to the assumption that the conditional probabilities vary Hölder continuously.

Given a continuous cocycle $\hat{A} : \Sigma_T \to \text{GL}_d(\mathbb{R})$, there are already known results about some of the listed properties above for the long term behavior of $\frac{1}{n} \log \| \hat{A}^n(\hat{x}) \|$ for

M.P. was supported in part by the National Science Foundation Grant “RTG: Analysis on manifolds” at Northwestern University.
\(\hat{\mu}\)-generic points \(\hat{x}\). The strong law of large numbers for example is known to hold in great generality \([\text{Kin73}]\); see also \([\text{Ose68}]\). Sufficient conditions which imply the simplicity of the top Lyapunov subspace have received much attention \([\text{BV04, AV07}]\). There has also been a significant amount of work concerning the continuity of the Lyapunov exponents \([\text{BV05, DK16, BNV17, VY19}]\). However, other than few exceptions such as \([\text{Bou88}]\), less attention has been paid to the problem of establishing the limit laws.

In the scalar case, it is well known that for a Hölder potential \(\hat{\varphi}: \Sigma_T \to \mathbb{R}\) whose \(n\)-th Birkhoff sum is denoted by \(S_n(\hat{\varphi})\), the distribution of 

\[
\frac{S_n(\hat{\varphi}(\hat{x}) - n\hat{\mu}(\hat{\psi})}{\sqrt{n}}
\]

with respect to \(\hat{\mu}\) satisfies various limit laws such as the central limit theorem, the large deviations principle, and a collection of other properties; see \([\text{Rat73, You90}]\) for instance.

The main goal of this article is to extend some of these limit theorems to a class of \(\text{GL}_d(\mathbb{R})\)-cocycles. More specifically, we consider 1-typical cocycles among fiber-bunched cocycles. For \(\hat{A}: \Sigma_T \to \text{GL}_d(\mathbb{R})\), we say \(\hat{A}\) is fiber-bunched if it is nearly conformal, and here “nearly” depends on the metric defined on \(\Sigma_T\) and the Hölder exponent of \(\hat{A}\). The 1-typicality assumption was first introduced by Bonatti and Viana \([\text{BV04}]\). They showed that 1-typical cocycles have simple top Lyapunov exponent with respect to any ergodic measures with continuous local product structure. The set of all such measures includes unique equilibrium states for Hölder potentials such as \(\hat{\mu}\). In some sense, the 1-typicality assumption is a suitable analogue for the proximality and the strong irreducibility that commonly appear in the study of the products of i.i.d. random matrices. Moreover, the set of 1-typical cocycles forms an open and dense subset of fiber-bunched cocycles. See Sect. 2 for its precise definition.

We denote the Lyapunov exponents of \(\hat{A}\) with respect to \(\hat{\mu}\) by \(\lambda_1(\hat{A}, \hat{\mu}) \geq \ldots \geq \lambda_d(\hat{A}, \hat{\mu})\). For the next two theorems, fix any unit vector \(u \in \mathbb{R}^d\) and define the variance as follows:

\[
\text{Var} := \text{Var}(\hat{A}, \hat{\mu}) = \lim_{n \to \infty} \frac{1}{n} \int \left( \log \left\| \hat{A}^n(\hat{x})u \right\| - n\lambda_1(\hat{A}, \hat{\mu}) \right)^2 d\hat{\mu}. \tag{1}
\]

We will see that under the 1-typicality assumption on \(\hat{A}\), the value of \(\text{Var}\) is independent of the choice of \(u\); see Proposition 10.

The first main theorem is the following version of the central limit theorem.

**Theorem 1 (Central Limit theorem).** Let \(\Sigma_T\) be a shift of finite type defined by an irreducible adjacency matrix \(T\), and let \(\hat{A}: \Sigma_T \to \text{GL}_d(\mathbb{R})\) be a 1-typical cocycle. Let \(\hat{\psi}: \Sigma_T \to \mathbb{R}\) be a Hölder potential and \(\hat{\mu} \in \mathcal{M}(\sigma)\) its unique equilibrium state. If \(\text{Var}(\hat{A}, \hat{\mu}) > 0\), then for every unit vector \(u \in \mathbb{R}^d\),

\[
\frac{\log \left\| \hat{A}^n(\hat{x})u \right\| - n\lambda_1(\hat{A}, \hat{\mu})}{\sqrt{n}} \xrightarrow{\text{dist}} N(0, \text{Var}).
\]

If \(\text{Var}(\hat{A}, \hat{\mu}) = 0\), then

\[
\frac{\log \left\| \hat{A}^n(\hat{x})u \right\| - n\lambda_1(\hat{A}, \hat{\mu})}{\sqrt{n}} \xrightarrow{\text{dist}} 0.
\]
The next theorem establishes the large deviation principle in the same setting.

**Theorem 2** (Large Deviation Principle). *In the setting of Theorem 1, suppose that \( \text{Var}(\hat{A}, \hat{\mu}) > 0 \). Then there exist \( \eta > 0 \) and a strictly convex function \( \Lambda^*: [0, \eta] \to \mathbb{R} \) such that \( \Lambda^* \) is positive on \( (0, \eta) \) and that for any \( \varepsilon \in (0, \eta) \) and any unit vector \( u \in \mathbb{R}^d \),

\[
\lim_{n \to \infty} \frac{1}{n} \log \hat{\mu} \left\{ \hat{x} \in \Sigma_T : \left| \log \| \hat{A}^n(\hat{x})u \| - n\lambda_1(\hat{A}, \hat{\mu}) \right| > n\varepsilon \right\} = -\Lambda^*(\varepsilon).
\]

As a corollary, we also obtain the following version of the large deviation principle.

**Corollary 1.** *In the setting of Theorem 1, suppose that \( \text{Var}(\hat{A}, \hat{\mu}) > 0 \). Then for all sufficiently small \( \varepsilon > 0 \), there exists \( C > 0 \) such that for all \( n \geq 0 \),

\[
\hat{\mu} \left\{ \hat{x} \in \Sigma_T : \left| \log \| \hat{A}^n(\hat{x}) \| - n\lambda_1(\hat{A}, \hat{\mu}) \right| > n\varepsilon \right\} \leq Ce^{-C^{-1}n}.
\]

We note that this version of the large deviation principle has previously been obtained by Gouëzel and Stoyanov [GS19] in the same setting. Furthermore, during the writing process of this paper, we were made aware that Duarte, Klein, and Poletti [DKP21] also establish such large deviation principle in a related setting. Both works then use this result to show interesting applications.

We also remark that for the products of i.i.d. random matrices, both the central limit theorem and the large deviation principle have been previously established under various assumptions; see [LP82, BL85, BQ16, Ser19]. Under a suitable moment condition together with an irreducibility assumption (i.e., Zariski density of the semi-group generated by the support of the distribution), the latest versions of the central limit theorem and the large deviation are due to Benoist and Quint [BQ16] and Sert [Ser19], respectively.

Our next result establishes the analytic dependence of the Lyapunov exponent \( \lambda_1(\hat{A}, \hat{\mu}) \) on the measure \( \hat{\mu} \).

**Theorem 3.** (Analyticity of the exponent) *In the setting of Theorem 1, let \( \{\hat{\varphi}_n\}_{n \in \mathbb{N}_0} \) be a sequence of positive real-valued Hölder continuous functions on \( \Sigma_T \), and let

\[
\hat{\psi}_t := \sum_{n=0}^\infty t^n \hat{\varphi}_n.
\]

Suppose that \( \sum_{n=0}^\infty e^n \| \hat{\varphi}_n \|_{C^\alpha(\Sigma_T)} < \infty \), for some \( \varepsilon, \alpha > 0 \). Then \( t \mapsto \lambda_1(\hat{A}, \mu_{\hat{\psi}_t}) \) is real analytic in a neighborhood of \( 0 \).

These theorems are established via introducing and studying the spectral properties of the operator \( \mathcal{L} \), which we describe below. From the given Hölder continuous function \( \hat{\psi}: \Sigma_T \to \mathbb{R} \), we construct a Hölder continuous function \( \psi \) in the cohomology class of \( \hat{\psi} \) that depends only on the future coordinates and realize \( \psi \) as a function on the one-sided subshift \( \Sigma_T^+ \). Associated to \( \psi \), the normalized Ruelle–Perron–Frobenius operator \( L\psi \) on \( C(\Sigma_T^+) \) is defined by

\[
L\psi f(x) := \frac{1}{\lambda} \sum_{y: \sigma^y = x} e^{\psi(y)} f(y)
\]

where \( \lambda = e^{P(\psi)} \) and \( P(\psi) \) is the topological pressure of \( \psi \).
In this setting, the operator $L_\psi$ acting on a suitable space of Hölder functions is quasi-compact and has a unique eigenfunction $h$ and a unique eigenmeasure $\eta$ corresponding to its spectral radius. Such eigendata give rise to the unique equilibrium state $\mu := \mu_\psi$ for $\psi$ defined by $d\mu = h \cdot d\eta$; in fact, the equilibrium state $\hat{\mu}$ for $\hat{\psi}$ can be retrieved as the natural extension of $\mu$. Historically, exploiting such spectral property of $L_\psi$ has been a popular method for establishing statistical properties for $\hat{\mu}$ mentioned above, see [Bal00] and the references contained therein.

By extending $L_\psi$, we define an operator $\mathcal{L} : C(\Sigma_T^+ \times \mathbb{P}^{d-1}) \to C(\Sigma_T^+ \times \mathbb{P}^{d-1})$ defined by

$$\mathcal{L} f(x, \bar{u}) := \frac{1}{\lambda} \sum_{y : \sigma y = x} e^{\psi(y)} f(y, \bar{A}(y)^\ast \bar{u}),$$

where we denote the projection of $u \in \mathbb{R}^d \setminus \{0\}$ onto $\mathbb{P}^{d-1}$ by $\bar{u}$.

In Theorem 4 we establish a Lasota-Yorke inequality for $\mathcal{L}$, and we use it to deduce further spectral properties of $\mathcal{L}$ in Theorem 5. The proofs of the main theorems stated above are based on such spectral properties of $\mathcal{L}$. In proving these spectral properties of $\mathcal{L}$, we exploit the gap in the top Lyapunov exponents $\lambda_1(\hat{A}, \hat{\mu}) > \lambda_2(\hat{A}, \hat{\mu})$ ensured by the 1-typicality assumption on $\hat{A}$. One difficulty in exploiting such a gap comes from the non-uniform nature of Lyapunov exponents. Such difficulty is overcome with the aid of the 1-typicality assumption: we show that non-uniform convergence of some expressions that limit to $\lambda_1(\hat{A}, \hat{\mu})$ becomes uniform when suitably averaged. See Sect. 3 for details.

We note that in the case where $\Sigma_T$ is a full shift and $A$ is locally constant, the value of $\mathcal{L} f(x, \bar{u})$ only depends on the second coordinate $\bar{u}$, and $\mathcal{L}$ descends to an operator on $C(\mathbb{P}^{d-1})$. Similar operators, such as the Markov operators and transfer operators, have appeared in the literature for similar purposes. For instance, such operators have been employed to study the subadditive equilibrium states [Pir18] and the products of random matrices and its associated stationary measures; see [Gui15] and references therein.

The paper is organized as follows. In Sect. 2, we elaborate on the setting for the main theorems and survey relevant preliminary results. Then in Sect. 3, we prove quasi-compactness of the operator $\mathcal{L}$ under the 1-typicality assumption on $\hat{A}$. Using the quasi-compactness of $\mathcal{L}$, we establish spectral properties of $\mathcal{L}$ in Sect. 4. In Sect. 5, we prove the main results of this paper. Section 6 is the Appendix containing the results useful in establishing the main theorems.

2. Preliminaries

2.1. Set up. We elaborate on the setting for the main theorems and set up the relevant notations here.

**Definition 1.** We say that a non-negative matrix $T$ is irreducible if for any $i, j$ there exists an $M$ such that $(T^M)_{ij} > 0$. We say that a matrix is primitive if there exists an $M \in \mathbb{N}$ such that $T^M > 0$.

Throughout the paper, $T$ is an irreducible $q \times q$-square matrix with entries in $\{0, 1\}$. We denote by $\Sigma_T$ and $\Sigma_T^+$ the corresponding two-sided and one-sided subshift of finite type defined by $T$, respectively. We will use $\sigma$ to denote the left shift operator for both $\Sigma_T$ and $\Sigma_T^+$. The points $\hat{x}$ of the two-sided subshift $\Sigma_T$ will be denoted with the hat notation in order to distinguish them from the points of the one-sided subshift $x \in \Sigma_T^+$. 
In particular, letting \( P : \Sigma_T \to \Sigma_T^+ \) be the standard projection, we have \( P\hat{x} = x \) where \( \hat{x} = (x_i)_{i \in \mathbb{Z}} \) and \( x = (x_i)_{i \in \mathbb{N}_0} \).

For any \( x \in \Sigma_T \), we define cylinder around \( \hat{x} \) of length \( n \) by

\[
[\hat{x}]_n := \{(y_i)_{i \in \mathbb{Z}} \in \Sigma_T : x_i = y_i \text{ for all } 0 \leq i \leq n - 1\}.
\]

For any \( x \in \Sigma_T^+ \), we likewise define the cylinder \([x]_n\) as \( P([\hat{x}]_n) \) for any \( \hat{x} \in P^{-1}(x) \).

We equip \( \Sigma_T \) with the following metric \( \rho_0 \): for \( \hat{x} = (x_i)_{i \in \mathbb{Z}} \) and \( \hat{y} = (y_i)_{i \in \mathbb{Z}} \),

\[
\rho_0(\hat{x}, \hat{y}) := 2^{-k}
\]

where \( k \) is the largest integer such that \( x_i = y_i \) for all \( |i| < k \). Similarly, we equip the one-sided subshift \( \Sigma_T^+ \) with the same metric. With such metric \( \rho_0 \), the subshift \( (\Sigma_T, \sigma) \) becomes a hyperbolic homemorphism. Its local stable sets are defined as

\[
W_{\text{loc}}^s(\hat{x}) := \{\hat{y} \in \Sigma_T : x_i = y_i \text{ for all } i \in \mathbb{N}_0\},
\]

and the local unstable sets \( W_{\text{loc}}^u(\hat{x}) \) are likewise defined as the set of \( \hat{y} \) where \( x_i = y_i \) for all \( i \leq 0 \). In the standard manner, we then extend the definition to global stable and unstable sets \( W_{\text{loc}}^{s/u}(\hat{x}) \). By an abuse of notations, for any \( x \in \Sigma_T^+ \) we define the local stable set of \( x \) as

\[
W_{\text{loc}}^s(x) := P^{-1}(x) = W_{\text{loc}}^s(\hat{x}) \text{ for any } \hat{x} \in P^{-1}(x).
\]

For any \( \hat{x}, \hat{y} \in \Sigma_T \) with \( x_0 = y_0 \), define the bracket of \( \hat{x} \) and \( \hat{y} \) by

\[
[\hat{x}, \hat{y}] := W_{\text{loc}}^u(\hat{x}) \cap W_{\text{loc}}^s(\hat{y}).
\]

For any \( \hat{x} = (x_i)_{i \in \mathbb{Z}} \in \Sigma_T \) and \( y = (y_i)_{i \in \mathbb{N}_0} \in \Sigma_T^+ \) with \( x_0 = y_0 \), we define \([\hat{x}, y] := [\hat{x}, \hat{y}]\) by choosing any \( \hat{y} \in W_{\text{loc}}^s(y) \). Note \([\hat{x}, y]\) is well-defined independent of the choice of \( \hat{y} \in W_{\text{loc}}^s(y) \), and we may often denote \([\hat{x}, y]\) by \( \hat{y}_\circ \).

For any continuous cocycle \( A : \Sigma_T \to \text{GL}_d(\mathbb{R}) \) and \( n \in \mathbb{N} \), we set

\[
A^n(\hat{x}) := A(\sigma^{-1}\hat{x}) \ldots A(\hat{x}).
\]

From the definition, it satisfies the cocycle equation: for any \( m, n \in \mathbb{N} \) and \( \hat{x} \in \Sigma_T \),

\[
A^{m+n}(\hat{x}) = A^m(\sigma^n\hat{x})A^n(\hat{x}).
\]

We can associate to \( A \) two related cocycles: the adjoint cocycle and the inverse cocycle. Both are cocycles over \( (\Sigma_T, \sigma^{-1}) \) and they are defined by

\[
A_*^*(\hat{x}) := A(\sigma^{-1}\hat{x})^* \text{ and } A^{-1}(\hat{x}) := A(\sigma^{-1}\hat{x})^{-1}.
\]

In particular, their iterations are given by \( A_*^n(\hat{x}) = [A^n(\sigma^{-n}\hat{x})]^* \) and \( A^{-n}(\hat{x}) = A^n(\sigma^{-n}\hat{x})^{-1} \). We denote the adjoint of the inverse cocycle by \( A_*^{-1}(\hat{x}) \); this defines a cocycle over \( (\Sigma_T, \sigma) \) by \( A_*^{-1}(\hat{x}) = [A(\hat{x})^{-1}]^* \).

Another system related to a given cocycle \( A : \Sigma_T \to \text{GL}_d(\mathbb{R}) \) is the skew product \( \hat{F}_A : \Sigma_T \times \mathbb{R}^{d-1} \to \Sigma_T \times \mathbb{R}^{d-1} \) defined by

\[
\hat{F}_A(\hat{x}, v) = (\sigma\hat{x}, A(\hat{x})v).
\]
It is clear that the action of $A^n(x)$ on the projective space $\mathbb{P}^{d-1}$ is encoded in the second coordinate of the iterations of the skew product $\tilde{F}_A$. Similarly, we denote the skew product on $\Sigma_T^+ \times \mathbb{P}^{d-1}$ by $F'_A$.

A cocycle $A: \Sigma_T \to GL_d(\mathbb{R})$ is $\theta$-Hölder if there exists $C > 0$ such that for all $\hat{x}, \hat{y} \in \Sigma_T$,

$$\|A(\hat{x}) - A(\hat{y})\| \leq C \cdot \rho_0(\hat{x}, \hat{y})^\theta.$$

We denote the set of all $\theta$-Hölder cocycles by $C^\theta(\Sigma_T, GL_d(\mathbb{R}))$. Throughout the paper, we will work with $\theta$-Hölder cocycles $A: \Sigma_T \to GL_d(\mathbb{R})$ for some $\theta \in (0, 1]$ satisfying an extra condition called the fiber-bunching.

**Definition 2.** A $\theta$-Hölder cocycle $A$ is fiber-bunched if

$$\|A(x)\| \cdot \|A(x)^{-1}\| < 2^\theta$$

for every $x \in \Sigma_T$.

Without loss of generality, we may assume that $\theta = 1$. Indeed, for the general case when $\theta \neq 1$ we may rescale the metric on $\Sigma_T$ to

$$\rho(x, y) := \rho_0(x, y)^\theta$$

so that $A$ belongs to $\text{Lip}(\Sigma_T, GL_d(\mathbb{R})) := C^1(\Sigma_T, GL_d(\mathbb{R}))$.

Conformal cocycles are cocycles $A$ that satisfy $\|A(x)\| \cdot \|A(x)^{-1}\| = 1$ for all $x \in \Sigma_T$. Clearly, Hölder continuous conformal cocycles are fiber-bunched. In fact, small perturbations of conformal cocycles are also fiber-bunched. An alternate way to think of fiber-bunched cocycles is that their skew products $\tilde{F}_A: \Sigma_T \times \mathbb{P}^{d-1} \to \Sigma_T \times \mathbb{P}^{d-1}$ defined above are partially hyperbolic systems. The set of fiber-bunched cocycles

$$C_b(\Sigma_T, GL_d(\mathbb{R})) := \{A \in \text{Lip}(\Sigma_T, GL_d(\mathbb{R})): A \text{ is fiber-bunched}\}$$

forms an open subset of $\text{Lip}(\Sigma_T, GL_d(\mathbb{R}))$.

One of the most important property of fiber-bunched cocycles is the convergence of the canonical stable/unstable holonomy $H^s,u_{\hat{x},\hat{y}}$: for any $\hat{y} \in W^s,u_{\text{loc}}(\hat{x})$,

$$H^s_{\hat{x},\hat{y}} := \lim_{n \to \infty} A^n(\hat{y})^{-1} A^n(\hat{x}) \quad \text{and} \quad H^u_{\hat{x},\hat{y}} := \lim_{n \to -\infty} A^n(\hat{y})^{-1} A^n(\hat{x}).$$

Moreover, the canonical holonomies vary Hölder continuously in the basepoints $\hat{x}, \hat{y} \in \Sigma_T$: there exists $C > 0$ such that for any $\hat{y} \in W^s,u_{\text{loc}}(\hat{x})$,

$$\|H^s_{\hat{x},\hat{y}} - I\| \leq C \cdot \rho(\hat{x}, \hat{y}).$$

See [KS13] for further details.

It can be easily checked that the canonical stable holonomies $H^s_{\hat{x},\hat{y}}$ satisfy the following properties:

1. $H^s_{\hat{x},\hat{x}} = I$ and $H^s_{\hat{y},\hat{z}} \circ H^s_{\hat{z},\hat{x}} = H^s_{\hat{x},\hat{y}}$ for any $\hat{y}, \hat{z} \in W^s_{\text{loc}}(\hat{x})$.
2. $A(\hat{x}) = H^s_{\hat{y},\hat{z},\hat{x}} \circ A(\hat{y}) \circ H^s_{\hat{x},\hat{z}}$.
3. $H^s: (\hat{x}, \hat{y}) \mapsto H^s_{\hat{x},\hat{y}}$ is continuous as $\hat{x}$ and $\hat{y}$ vary continuously while satisfying the relation $\hat{y} \in W^s_{\text{loc}}(\hat{x})$. 

See [KS13] for further details.
Likewise, the canonical unstable holonomies $H^u_{\hat{x},\hat{y}}$ satisfy the analogous properties.

The existence and continuity of the canonical holonomies guarantee that $A$ has the following \textit{bounded distortion property}: there exists $C > 0$ such that for any $n \in \mathbb{N}$ and $\hat{x}, \hat{y} \in \Sigma_T$ with $\hat{y} \in [\hat{x}]_n$, we have

$$C^{-1} \leq \frac{\|A^n(\hat{x})\|}{\|A^n(\hat{y})\|} \leq C.$$  

(5)

We put a metric $d$ on $\mathbb{P}^{d-1}$ defined by

$$d(\overline{u}, \overline{v}) := \frac{\|u \wedge v\|}{\|u\| \cdot \|v\|}$$

(6)

where $u \wedge v$ is the exterior product of $u$ and $v$. It can be easily verified that $d(\overline{u}, \overline{v})$ is equal to the sine of the angle between $\overline{u}$ and $\overline{v}$.

For such metric $d$, there exists $C > 0$ (twice the constant from (4), for instance) such that for any $\overline{u} \in \mathbb{P}^{d-1}$ and $\hat{y} \in W^s_{\text{loc}}(\hat{x})$,

$$d\left(H^s_{\hat{x},\hat{y}} \overline{u}, \overline{u}\right) \leq C \cdot \rho(\hat{x}, \hat{y}).$$

(7)

By an abuse of notation, we also denote by $d$ the metric on the product space $\Sigma^+ T \times \mathbb{P}^{d-1}$ and $\Sigma_+ T \times \mathbb{P}^{d-1}$ defined as the maximum of the distance in each coordinate:

$$d((\hat{x}, \overline{v}), (\hat{y}, \overline{u})) := \max\{\rho(\hat{x}, \hat{y}), d(\overline{u}, \overline{v})\}.$$  

Lastly, we describe how to conjugate and realize a given fiber-bunched cocycle $\hat{A}: \Sigma_T \to \text{GL}_d(\mathbb{R})$ as a $\text{GL}_d(\mathbb{R})$-cocycle $A$ over the one-sided subshift $(\Sigma^+_T, \sigma)$. Assuming that $\Sigma_T$ has $q$ symbols, for each $i \in \{1, 2, \ldots, q\}$, fix $\hat{\eta}^i \in \Sigma_T$ with $(\hat{\eta}^i)_0 = i$. For each $\hat{x} = (x_i)_{i \in \mathbb{Z}} \in \Sigma_T$, we set

$$\hat{x}_\eta := [\hat{\eta}^0, \hat{x}].$$

We then define a new cocycle $A$ on $\Sigma_T$ given by

$$A(\hat{x}) := \hat{A}(\hat{x}_\eta).$$

From its definition, $A$ is constant along the local stable set:

$$A(\hat{x}) = A(\hat{y}) \text{ for } \hat{y} \in W^s_{\text{loc}}(\hat{x}).$$

(8)

Setting $C(\hat{x}) := \hat{H}^s_{\hat{x},\hat{y}}$ where $\hat{H}$ denotes the local holonomies of $\hat{A}$, it serves as a conjugacy between $A$ and $\hat{A}$:

$$A(\hat{x}) \circ C(\hat{x}) = C(\sigma \hat{x}) \circ \hat{A}(\hat{x}).$$

For further details on this conjugacy, see [BGMV03, Corollary 1.15] and [BBB18].

Denoting its local holonomies of $A$ by $H^s_{\hat{x},\hat{y}}$, we have $H^s \equiv I$, and hence, we may realize $A$ as a cocycle over the one-sided subshift $(\Sigma^+_T, \sigma)$.

We will always assume that all fiber-bunched cocycles from here on have already been conjugated as above, and use $A$ to denote the cocycle over $(\Sigma^+_T, \sigma)$ as well as $(\Sigma_T, \sigma)$ when the context is clear.
2.2. Typicality assumption on $\mathcal{A}$. In order to introduce the 1-typicality assumption on $\mathcal{A}$, we need to introduce the notion of holonomy loops. For any periodic point $p \in \Sigma_T$, the homoclinic points of $p$ are points that belong to the intersection between $W^s(p)$ and $W^u(p)$. For uniformly hyperbolic systems such as $(\Sigma_T, \sigma)$, the set of $p$-homoclinic points is dense for any periodic point $p$. To each $p$-homoclinic point $z$, we can associate a homoclinic loop defined by
\[
\psi_z := H^s_{z,p} \circ H^u_{p,z}.
\]

**Definition 3.** We say $\mathcal{A} \in C_b(\Sigma_T, \text{GL}_d(\mathbb{R}))$ is 1-typical if there exist a periodic point $p$ of some period $n$ and a $p$-homoclinic point $z$ such that

1. $P := \mathcal{A}^n(p)$ has simple eigenvalues of distinct norms, and
2. for any $I, J \subseteq \{1, \ldots, d\}$ with $|I| + |J| \leq d$, 
\[
\{\psi_z(v_i) : i \in I\} \cup \{v_j : j \in J\}
\]
are linearly independent, where $\{v_i\}_{i=1}^d$ are the eigenvectors of $P$.

The above two conditions are called pinching and twisting, respectively, which can be seen as generalizations of the proximality and the strong irreducibility of i.i.d. products of random matrices. Given a random walk on $\text{GL}_d(\mathbb{R})$ according to a probability measure $\nu$ on $\text{GL}_d(\mathbb{R})$, let $\Gamma_\nu$ be the smallest closed semigroup containing the support of $\nu$. We say $\Gamma_\nu$ is proximal if it contains an element which has an eigenvalue whose modulus is strictly larger than the modulus of any other eigenvalue, and strongly irreducible if there does not exist a finite set of proper subspaces that is preserved under $\Gamma_\nu$.

Indeed, when the base is a full shift $\Sigma_q$ equipped with a Bernoulli measure and the given cocycle $\mathcal{A} \in C_b(\Sigma_q, \text{GL}_d(\mathbb{R}))$ is locally constant so that it can be viewed as a random walk on $\text{GL}_d(\mathbb{R})$ generated by a distribution $\nu$ finitely supported on the image of $\mathcal{A}$, then the pinching and the twisting conditions necessarily imply proximality and strong irreducibility of $\Gamma_\nu$.

Moreover, as mentioned in the introduction, the pinching and twisting conditions were introduced by Bonatti and Viana [BV04] to generalize the simplicity of Lyapunov exponents for i.i.d. products of random matrices under proximality and strong irreducibility (see for instance, Furstenberg [Fur63] for the positivity of the Lyapunov top exponent). They showed that it is a sufficient (topological) condition on the cocycle which guarantees that the top and bottom Lyapunov exponents with respect to any ergodic measures with continuous local product structure are simple. If the pinching and twisting conditions hold for all exterior product cocycles $\mathcal{A}^\wedge$, $t \in \{1, \ldots, \lfloor d/2 \rfloor\}$, such cocycles $\mathcal{A}$ are called typical. Using standard tricks, the simplicity of the top and bottom Lyapunov exponents for 1-typicality can be modified to show the simplicity of all Lyapunov exponents for such cocycles. Moreover, typical cocycles form an open and dense subset of $C_b(\Sigma_T, \text{GL}_d(\mathbb{R}))$.

Typicality, in slightly altered forms, have recently been employed to study subadditive thermodynamic formalism; see [Par20].

2.3. Invariant measures. Let $\hat{\psi}$ be a Hölder potential on $\Sigma_T$ and $\hat{\mu} \in \mathcal{M}(\sigma)$ be its unique equilibrium state; the uniqueness of $\hat{\mu}$ is well-known when $T$ is primitive [Bow75, 1.4], but the result still holds when $T$ is irreducible [Bal00, Theorem 1.5, Exercise 1.24]. We fix $\hat{\psi}$ and $\hat{\mu}$ once and for all throughout the paper.
Let $\psi$ be a Hölder potential cohomologous to $\hat{\psi}$ defined as in [Bow75, Lemma 1.6], and we realize $\psi$ as a potential on $\Sigma_T^+$. Then the projection of $\hat{\mu}$ onto $\Sigma_T^+$ denoted by $\mu$ is the unique equilibrium state for $\psi$.

Instead of working directly with $\psi$, we will work with $g$-functions, which helps simplify the presentations of the proofs of the main theorems. We say $g : \Sigma_T^+ \to \mathbb{R}$ is a $g$-function if $g$ is a positive function satisfying

$$\sum_{y : \sigma y = x} g(y) = 1$$

for all $x \in \Sigma_T^+$. This is equivalent to the condition that $L \log g 1 = 1$.

Recalling that $h$ is the unique eigenfunction corresponding to the spectral radius of $L_\psi$, a Hölder continuous function $\log g$ defined by

$$g(x) := \frac{e^{\psi(x)}}{\lambda} \cdot \frac{h(x)}{h(\sigma x)}$$

is cohomologous to $\psi$. Hence, we may work with $\log g$ instead of the given $\psi$, and its Ruelle–Perron–Frobenius operator is then defined by

$$L_{\log g} f(x) := \sum_{y : \sigma y = x} g(y) f(y).$$

Then its unique eigenmeasure corresponding to the spectral radius coincides with $\mu$, the unique equilibrium state for $\psi$. Replacing $\psi$ with $\log g$, we can see that $\mathcal{L}$ takes the form

$$\mathcal{L} f(x, \bar{u}) := \sum_{y : \sigma y = x} g(y) f(y, \overline{A(y)^*u}).$$

It is clear that the spectral radius $\rho(\mathcal{L})$ is equal to 1, and the constant function 1 is an eigenfunction of $\mathcal{L}$. Moreover, setting $g^{(n)}(x) := g(\sigma^{n-1}x) \ldots g(x)$, the Gibbs property of $\mu$ gives

$$\mu(\bar{I}) \asymp g^{(n)}(x)$$

for any cylinder $I$ of length $n$ and any $x \in [I]$. In this context the relation $\asymp$ means that there is a constant $C > 0$ such that

$$C^{-1} g^{(n)}(x) \leq \mu(\bar{I}) \leq C g^{(n)}(x).$$

By considering $\Sigma_T^+$ as the parameter space for the local stable sets of $\Sigma_T$, we may consider a disintegration $\{\hat{\mu}_x\}_{x \in \Sigma_T^+}$ of $\hat{\mu}$ where each $\hat{\mu}_x$ is a probability measure on $\mathcal{W}_{\text{loc}}^s(x)$ for $\mu$-a.e. $x$. For any such disintegration, we have

$$\hat{\mu} = \int_{\Sigma_T^+} \hat{\mu}_x^s d\mu(x).$$

It is well-known that equilibrium states of Hölder potentials, such as $\hat{\mu}$, have continuous local product structure; see [Lep00]. From such property, there exists a disintegration $\{\hat{\mu}_x^s\}_{x \in \Sigma_T^+}$ of $\hat{\mu}$ whose unstable holonomies between local stable sets defined by

$$h_{x, y} : (\mathcal{W}_{\text{loc}}^s(x), \hat{\mu}_x^s) \to (\mathcal{W}_{\text{loc}}^s(y), \hat{\mu}_y^s)$$

$\hat{x} \mapsto [\hat{x}, y]$
are absolutely continuous for every \( x, y \in \Sigma_T \) with \( x_0 = y_0 \). Its Jacobian \( J_{x,y} \) depends Hölder continuously in \( x, y \in \Sigma_T^+ \), and we have

\[
\hat{\mu}^s_{y} = J_{x,y}(h_{x,y})_*\hat{\mu}^s_{x}.
\] (10)

We now survey relevant results from Bonatti and Viana [BV04]. Let \( A : \Sigma_T \to \text{GL}_d(\mathbb{R}) \) be a fiber-bunched cocycle, and \( \hat{m} \) be a \( \hat{F}_A \)-invariant measure on \( \Sigma_T \times \mathbb{P}^{d-1} \) that projects to \( \hat{\mu} \) under the canonical projection \( \hat{\pi} : \Sigma_T \times \mathbb{P}^{d-1} \to \Sigma_T \). Setting \( m := (P \times \text{id})_*\hat{m} \), a Martingale argument shows that \( \hat{m} \) can be retrieved from \( m \) in the following sense: letting \( \hat{x}_n := P(\sigma^{-n}\hat{x}) \), for \( \hat{\mu} \)-a.e. \( \hat{x} \),

\[
\lim_{n \to \infty} \mathcal{A}^n(\hat{x}_n)_*m_{\hat{x}_n} = \hat{m}_{\hat{x}}.
\] (11)

**Definition 4.** A probability measure \( \hat{m} \) on \( \Sigma_T \times \mathbb{P}^{d-1} \) is \( H^u \)-invariant if it projects to \( \hat{\mu} \) and there exists a disintegration \( \{m_{\hat{x}}\}_{\hat{x} \in \Sigma_T} \) along the fibers such that

\[
(H^u_{\hat{x},\hat{y}})_*\hat{m}_{\hat{x}} = \hat{m}_{\hat{y}},
\]

for every \( \hat{x} \) and \( \hat{y} \) in the same local unstable set. We say \( \hat{m} \) is \((A, \hat{H}^u)\)-invariant if, in addition, \( \hat{m} \) is \( \hat{F}_A \)-invariant. Such measure is also known as a u-state.

We say that a probability measure \( m \) on \( \Sigma_T^+ \times \mathbb{P}^{d-1} \) is \((A, \hat{H}^u)\)-invariant if there exists an \((A, \hat{H}^u)\)-invariant probability measure \( \hat{m} \) on \( \Sigma_T \times \mathbb{P}^{d-1} \) with \( (P \times \text{id})_*\hat{m} = m \). We similarly define the \((A, \hat{H}^u)\)-invariance. While the existence of \((A, \hat{H}^u)\)-invariant measures is not a priori obvious, it is shown in [BV04] that the set of \((A, \hat{H}^u)\)-invariant measures is necessarily non-empty.

The main result of Bonatti and Viana [BV04] is that if \( A \) is 1-typical, then the top and bottom Lyapunov exponents of \( A \) with respect to \( \hat{\mu} \) are simple. Let \( \xi(\hat{x}) \in \mathbb{P}^{d-1} \) be the projectivization of the top Oseledets subspace at \( \hat{x} \) with respect to \( \hat{A} \) and \( \hat{\mu} \); when there is no confusion, we will denote a unit vector in its direction also by \( \xi(\hat{x}) \). Then consider a probability measure \( \hat{m} \) on \( \Sigma_T \times \mathbb{P}^{d-1} \) which projects to \( \hat{\mu} \) and whose conditional measures are defined by

\[
\hat{m}_{\hat{x}} := \delta_{\xi(\hat{x})}.
\]

Since \( \xi(\hat{x}) \) is both \( \hat{F}_A \) and \( \hat{H}^u \)-invariant, it follows that \( \hat{m} \) is a \((A, \hat{H}^u)\)-invariant measure.

Bonatti and Viana then proceed to show that \( \hat{m} \) is the unique \((A, \hat{H}^u)\)-invariant measure over \( \hat{\mu} \) when \( A \) is 1-typical. Setting \( m := (P \times \text{id})_*\hat{m} \), we summarize the properties of \( \hat{m} \) and \( m \) in the following proposition:

**Proposition 1** [BV04] Suppose \( A \) is 1-typical. Then \( m \) constructed as above satisfies the following properties:

1. \( m \) admits a disintegration \( \{m_x\}_{x \in \Sigma_T^+} \) such that \( x \mapsto m_x \) is continuous for every \( x \in \Sigma_T^+ \).
2. For \( \mu \)-a.e. \( x \in \Sigma_T^+ \), we have

\[
\sum_{\sigma y = x} \frac{1}{J_{\mu} \sigma(y)} \mathcal{A}(y)_*m_y = m_x,
\]

where \( J_{\mu} \sigma : \Sigma_T^+ \to (0, \infty) \) is the Jacobian for \( \mu \).
3. For any $x \in \Sigma_T^+$ and proper subspace $V \subset \mathbb{R}^d$, we have $m_x(V) = 0$.

4. For every $x \in \Sigma_T^+$, we have

$$m_x = \int \hat{m}_x \, d\hat{\mu}_x(\hat{x}) = \int \delta_{\hat{x}} \, d\hat{\mu}_x(\hat{x}).$$

We end this subsection by commenting on the construction of the top Oseledets subspace $\xi(\hat{x})$. The following remark will be useful in the proof of Theorem 4.

**Remark 1.** Under 1-typicality, Bonatti and Viana obtain two measurable splittings of $\mathbb{R}^d$: for $\hat{\mu}$-a.e. $\hat{x}$,

$$\xi(\hat{x}) \oplus V_2(\hat{x}) = \mathbb{R}^d = \omega(\hat{x}) \oplus W_2(\hat{x}).$$

The 1-dimensional top Oseledets subspace equals to the span of $\xi(\hat{x})$ and its complement $V_2(\hat{x})$ consists of all vector $v \in \mathbb{R}^d$ whose asymptotic growth rate $\lim_{n \to \infty} \frac{1}{n} \log \| A^n(\hat{x}) v \|$ is strictly less than $\lambda_1(A, \hat{\mu})$. Similarly, $\omega(\hat{x})$ spans the 1-dimensional bottom Oseledets subspace of $A$ with respect to $\hat{\mu}$ and $W_2(\hat{x})$ is its complementary subspace. Note $\xi(\hat{x})$ is preserved under $H^u$ whereas $\omega(\hat{x})$ is preserved under $H^s$. These subspaces can be constructed as follow:

Letting $\hat{x}_n := P(\sigma^{-n} \hat{x})$, for $\hat{\mu}$-a.e. $\hat{x}$, $\frac{A^n(\hat{x}_n)}{\|A^n(\hat{x}_n)\|}$ converges to a rank 1 quasi-projective map whose image coincides with $\xi(\hat{x}) \in \mathbb{P}^{d-1}$. Denoting the $KAK$-decomposition, also known as the singular value decomposition, of $A^n(\hat{x}_n)$ by $K_n A_n U_n$, let $K_\infty \in O(d)$ be a subsequent limit of $K_n$. Then, we have $\xi(\hat{x}) = K_\infty e_1$ for $\hat{\mu}$-a.e. $\hat{x}$. If we apply this construction of $\xi(\hat{x})$ to $A_n$, then the corresponding direction $\xi_*(\hat{x})$ is used to define the slower subspace of the Lyapunov splitting $V_2(\hat{x}) := \xi_*(\hat{x}) \perp$, necessarily transverse to $\xi(\hat{x})$.

Applying this construction to the inverse cocycle $A^{-1}$ defines $\omega(\hat{x})$. Moreover, applying this construction to the adjoint of the inverse cocycle $A_n^{-1}$, we obtain $\omega_*(\hat{x})$ whose orthogonal complement defines the complementary subspace $W_2(\hat{x})$. We notice that $\omega_*(\hat{x})$ coincides with $K_\infty e_d$, where $K_\infty$ is defined as in the previous paragraph. This is because keeping in mind the $KAK$-decomposition of $A^n(\sigma^{-n} \hat{x})$ from the previous paragraph, the image of the rank 1 quasi-projective limit of $(A_n^{-1})^n(\sigma^{-n} \hat{x}) = ((A^n(\sigma^{-n} \hat{x}))^{-1})^*$ is equal to $K_\infty e_d \in \mathbb{P}^{d-1}$.

3. Quasi-compactness of $\mathcal{L}$

For $f \in C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})$, we set

$$|f|_{\alpha} := \sup_{(x, \overline{u}) \neq (y, \overline{v})} \frac{|f(x, \overline{u}) - f(y, \overline{v})|}{d((x, \overline{u}), (y, \overline{v}))^\alpha}.$$ 

With this notation, we define the $\alpha$-norm $\|f\|_{\alpha} := |f|_{\alpha} + \|f\|_{\infty}$.

The main goal of this section is to prove the following theorem establishing a Lasota-Yorke inequality for $\mathcal{L}$.
**Theorem 4.** Suppose $\Sigma_T^+$ is a shift of finite type defined by an irreducible $T$, and $\hat{A}$ is 1-typical. Then for all $\alpha > 0$ sufficiently small, there exist $C > 0$ and $\beta \in (0, 1)$ such that for any $f \in C^\alpha(\Sigma_T^+ \times \mathbb{R}^{d-1})$, we have

$$\| \mathcal{L}^n f \|_\alpha \leq \beta^n \| f \|_\alpha + C \| f \|_\infty$$

(13)

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

As done in Sect. 2, we will work with the cocycle $A : \Sigma_T \to \text{GL}_d(\mathbb{R})$ that is constant along the local stable sets obtained from $\hat{A}$ according to (8). The canonical holonomies of $A$ are denoted by $H_s^{i/u}$ where $H^s \equiv I$.

For $x \in \Sigma_T^+$ and $n \in \mathbb{N}$, we define

$$A[n](x) := A(x)^*A(\sigma x)^* \cdots A(\sigma^{n-1} x)^* = [A^n(x)]^*.$$ 

Then the iterates of $\mathcal{L}$ take the following form

$$\mathcal{L}^n f(x, \overline{u}) := \sum_{y : \sigma^n y = x} g^{(n)}(y) \cdot f(y, A[n](y)u).$$

### 3.1. Controlling the Hölder norm of $\mathcal{L}^n f$

We will make a series of deductions to establish sufficient conditions for (13). First, we set

$$t_{n, \alpha}(x) := \sup_{\overline{u} \neq \overline{v}} \sum_{\sigma^n y = x} g^{(n)}(y) \cdot \left( \frac{d(A[n](y)u, A[n](y)v)}{d(\overline{u}, \overline{v})} \right)^\alpha,$$

and

$$w_{n, \alpha} := \max_{x \in \Sigma_T^+} t_{n, \alpha}(x)$$

whose existence is guaranteed from the compactness of $\Sigma_T^+$.

For any $x_1$ and $x_2$ with the same 0-th symbol, consider any $y_1 \in \sigma^{-n} x_1$. Then there exists $y_2 \in \sigma^{-n} x_2$ such that $y_1$ and $y_2$ belong to the same $n$-cylinder. We set

$$\tau_{n, \alpha} := \sup_{x_1 \neq x_2} \sum_{\sigma^n y_i = x_i, i \in [1, 2]} g^{(n)}(y_1) \cdot \left( \frac{d(A[n](y_1)v, A[n](y_2)v)}{\rho(x_1, x_2)} \right)^\alpha.$$

The supremum is taken over all distinct $x_1, x_2$ with the common 0-th symbol and all $\overline{u} \in \mathbb{R}^{d-1}$, and the summation is over all $y_1 \in \sigma^{-n} x_1$ paired with the corresponding $y_2 \in \sigma^{-n} x_2$ in the same $n$-cylinder.

The following proposition formulates a sufficient condition to establish Theorem 4.

**Proposition 2.** For all $\alpha > 0$ sufficiently small, there exists $C > 0$ such that for any $(x_1, \overline{u}), (x_2, \overline{v}) \in \Sigma_T^+ \times \mathbb{R}^{d-1}$, $f \in C^\alpha(\Sigma_T^+ \times \mathbb{R}^{d-1})$, and $n \in \mathbb{N}$, we have

$$|\mathcal{L}^n f(x_1, \overline{u}) - \mathcal{L}^n f(x_2, \overline{v})| \leq |f|_\alpha \left[ w_{n, \alpha} \cdot d(\overline{u}, \overline{v})^\alpha + (\tau_{n, \alpha} + 2^{-n\alpha}) \cdot \rho(x_1, x_2)^\alpha \right] + C \cdot \| f \|_\infty \cdot \rho(x_1, x_2)^\alpha.$$
Proof. If \( x_1 \) and \( x_2 \) do not have the same 0-th symbol, then \( \rho(x_1, x_2) = 1 \) and
\[
|\mathcal{L}^n f(x_1, \bar{u}) - \mathcal{L}^n f(x_2, \bar{v})| \\
\leq \sum_{\sigma^n y_1 = x_1} g^{(n)}(y_1)\|f\|_\infty + \sum_{\sigma^n y_2 = x_2} g^{(n)}(y_2)\|f\|_\infty = 2\|f\|_\infty
\]
where the last equality follows from the fact that \( \sum_{\sigma^n y = x} g^{(n)}(y) = (L^n_{\log g} 1)(x) = 1 \) for any \( x \in \Sigma_T^+ \) and \( n \in \mathbb{N} \).

So we may assume that the 0-th symbols of \( x_1 \) and \( x_2 \) agree. In order to control the Hölder norm of \( \mathcal{L}^n f \), we consider the difference
\[
\mathcal{L}^n f(x_1, \bar{u}) - \mathcal{L}^n f(x_2, \bar{v}) = \sum_{\sigma^n y_i = x_i, i \in \{1, 2\}} \left[ g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_1)u) - g^{(n)}(y_2)f(y_2, \mathcal{A}[n](y_2)v) \right].
\]
The difference within the rectangular bracket may be estimated by the sum of four differences:
\[
g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_1)u) - g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_1)v) \tag{14}
\]
and
\[
g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_1)v) - g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_2)v) \tag{15}
\]
and
\[
g^{(n)}(y_1)f(y_1, \mathcal{A}[n](y_2)v) - g^{(n)}(y_1)f(y_2, \mathcal{A}[n](y_2)v) \tag{16}
\]
and
\[
g^{(n)}(y_1)f(y_2, \mathcal{A}[n](y_2)v) - g^{(n)}(y_2)f(y_2, \mathcal{A}[n](y_2)v). \tag{17}
\]

Since \( f \) is \( \alpha \)-Hölder, the absolute value of (14) is bounded above by
\[
g^{(n)}(y_1) \cdot |f|_\alpha \cdot d(\mathcal{A}[n](y_1)u, \mathcal{A}[n](y_1)v)^{\alpha}.
\]
The sum of these terms over all \( y_1 \in \sigma^{-n}x_1 \) is bounded above by \( \omega_n, \alpha \cdot |f|_\alpha \cdot d(\bar{u}, \bar{v})^{\alpha} \).

Similarly, the absolute value of (15) is bounded above by
\[
g^{(n)}(y_1) \cdot |f|_\alpha \cdot d(\mathcal{A}[n](y_1)v, \mathcal{A}[n](y_2)v)^{\alpha},
\]
and the sum of these terms is bounded above by \( \tau_n, \alpha \cdot |f|_\alpha \cdot \rho(x_1, x_2)^{\alpha} \).

Likewise, the absolute value of (16) is bounded above by
\[
g^{(n)}(y_1) \cdot |f|_\alpha \cdot \rho(y_1, y_2)^{\alpha} = g^{(n)}(y_1) \cdot |f|_\alpha \cdot (2^{-n} \rho(x_1, x_2))^{\alpha}.
\]

Since \( \log g \) is Hölder continuous, for sufficiently small \( \alpha > 0 \) (i.e., smaller than the Hölder exponent of \( \log g \)) we have \( |\log g^{(n)}(y_1) - \log g^{(n)}(y_2)| \leq K \rho(x_1, x_2)^{\alpha} \) for some \( K \). Then, the absolute value of (17) is bounded above by
\[
\|f\|_\infty \cdot g^{(n)}(y_1) \cdot |e^{K\rho(x_1, x_2)^{\alpha}} - 1| \leq \|f\|_\infty \cdot g^{(n)}(y_1) \cdot K_1 \rho(x_1, x_2)^{\alpha}
\]
for some $K_1 > 0$.

Putting it all together, this gives

$$|\mathcal{L}^n f(x_1, \overline{u}) - \mathcal{L}^n f(x_2, \overline{v})| \leq |f|_\alpha \cdot \left(\tau_{n, \alpha} + 2^{-n\alpha} \sum_{\sigma^n y_1 = x_1} g^{(n)}(y_1)\right) \cdot \rho(x_1, x_2)^\alpha$$

$$+ |f|_\alpha \cdot w_{n, \alpha} \cdot d(\overline{u}, \overline{v})^\alpha$$

$$+ \|f\|_\infty \cdot K_1 \sum_{\sigma^n y_1 = x_1} g^{(n)}(y_1) \cdot \rho(x_1, x_2)^\alpha.$$}

Since $\sum_{\sigma^n y_1 = x_1} g^{(n)}(y_1) = 1$, this proves the proposition. □

**Corollary 2.** Suppose for $\alpha > 0$ sufficiently small that we have

$$\limsup_{n \to \infty} \frac{1}{n} \log w_{n, \alpha} < 0$$

and

$$\limsup_{n \to \infty} \frac{1}{n} \log \tau_{n, \alpha} < 0.$$  \hfill (18)$$\hfill (19)

Then Theorem 4 holds for such $\alpha > 0$. Moreover, there exist $\beta \in (0, 1)$ and $C > 0$ such that for any $(x, \overline{u}), (y, \overline{v}) \in \Sigma_T^+ \times \mathbb{P}^{d-1},$

$$|\mathcal{L}^n f(x, \overline{u}) - \mathcal{L}^n f(y, \overline{v})| \leq \beta^n |f|_\alpha + C \rho(x, y)^\alpha \|f\|_\infty.$$  \hfill (20)

**Proof.** The first statement is an immediate corollary of Proposition 2. The second claim also easily follows since the diameter of $\mathbb{P}^{d-1}$ is finite and $\rho(x, y)$ is bounded above by 1 for any $x, y \in \Sigma_T^+$. □

We have reduced proving Theorem 4 to establishing (18) and (19) for all $\alpha > 0$ sufficiently small. We will first focus on establishing (18), and deal with (19) in Sect. 3.4. The following lemma shows that the lim sup in (18) is in fact a limit. This fact allows us to manipulate the condition (18) more effectively.

**Lemma 1.** The sequence $\{\log w_{n, \alpha}\}_{n \in \mathbb{N}}$ is subadditive, and hence

$$\lim_{n \to \infty} \frac{1}{n} \log w_{n, \alpha} = \inf_{n \to \infty} \frac{1}{n} \log w_{n, \alpha}.$$  \hfill (21)

In particular, (18) would follow if we can show that there exists $n \in \mathbb{N}$ such that

$$\log w_{n, \alpha} = \log \max_{x \in \Sigma_T^+} \tau_{n, \alpha}(x) < 0.$$  \hfill (20)

**Proof.** We demonstrate the idea for $w_n := w_{n, 1}$. The proof readily extends to $w_{n, \alpha}$ for any $\alpha > 0$.

Let $m, n \in \mathbb{N}$, and $x \in \Sigma_T^+$ be the point that achieves the maximum in $w_{m+n}$. Then

$$w_{m+n} = \sup_{\overline{u} \neq \overline{v}} \sum_{\sigma^{m+n} y = x} g^{(m+n)}(y) \cdot \frac{d(A^{m+n}(y)u, A^{m+n}(y)v)}{d(\overline{u}, \overline{v})}.$$
\[
\sup_{\tilde{\mu} \neq \tilde{\nu}} \sum_{\sigma^m y = x} \left[ g^{(n)}(\sigma^m y) \cdot \frac{d(\mathcal{A}([n] \sigma^m y)u, \mathcal{A}([n] \sigma^m y)v)}{d(\tilde{u}, \tilde{v})} \right. \\
\times \left. g^{(m)}(y) \cdot \frac{d(\mathcal{A}([n+m] y)u, \mathcal{A}([n+m] y)v)}{d(\mathcal{A}([n] \sigma^m y)u, \mathcal{A}([n] \sigma^m y)v)} \right] \\
= \sup_{\tilde{\mu} \neq \tilde{\nu}} \sum_{\sigma^m z = x} \left[ g^{(n)}(z) \cdot \frac{d(\mathcal{A}([n] z)u, \mathcal{A}([n] z)v)}{d(\tilde{u}, \tilde{v})} \right. \\
\times \left. \left( \sum_{\sigma^m y = z} g^{(m)}(y) \cdot \frac{d(\mathcal{A}([m] y)u, \mathcal{A}([m] y)v)}{d(\tilde{u}, \tilde{v})} \right) \right]
\]

where \( u_z := \mathcal{A}([n] z)u \) and \( v_z := \mathcal{A}([n] z)v \). It then follows that

\[
w_{m+n} \leq \sup_{\tilde{\mu} \neq \tilde{\nu}} \sum_{\sigma^m z = x} g^{(n)}(z) \cdot \frac{d(\mathcal{A}([n] z)u, \mathcal{A}([n] z)v)}{d(\tilde{u}, \tilde{v})} \cdot t_m(z) \leq w_n \cdot w_m.
\]

This shows that \( \{ \log w_n \}_{n \in \mathbb{N}} \) is subadditive. The second claim follows from Fekete’s subadditive lemma. The last claim is a trivial consequence of the second claim. \( \square \)

3.2. Passing to the integral with respect to \( \hat{\mu}_x^s \). We will express \( t_{n,a} \) as an integral over \( \hat{\mu}_x^s \). Recalling that \( \hat{\mu} = \{ \hat{\mu}_x^s \}_{x \in \Sigma_T^+} \) is the disintegration of \( \hat{\mu} \) along the local stable sets and \( P : \Sigma_T \to \Sigma_T^+ \) is the canonical projection, we begin by establishing the Gibbs-like property of \( \hat{\mu}_x^s \).

**Lemma 2.** For any \( x = (x_i)_{i \in \mathbb{N}_0} \in \Sigma_T^+, n \in \mathbb{N}, \) and \( l = i_0 \ldots i_{n-1} \), consider \( y \in [l] \) with \( \sigma^n y = x \). We have

\[
\hat{\mu}_x^s([\hat{y} \in \mathcal{W}^{st}_{loc}(x) : P(\sigma^{-n} \hat{y}) = y]) \asymp g^{(n)}(y).
\]

**Proof.** From \( \sigma \)-invariance of \( \hat{\mu} \), \( P_* \hat{\mu} = \mu \), and the Gibbs property of \( \mu \), we have

\[
\hat{\mu}([\hat{\omega} \in \Sigma_T : \omega_{i-n} = y_i \text{ for } 0 \leq i \leq n-1 \text{ and } \omega_0 = x_0]) = \mu([x_0]) \asymp g^{(n+1)}(y).
\]

Since the term on the left is equal to

\[
\int_{[x_0]} \hat{\mu}_x^s([\hat{y} \in \mathcal{W}^{st}_{loc}(x) : P(\sigma^{-n} \hat{y}) = y]) d\mu(x)
\]

and the disintegrated measures \( \{ \hat{\mu}_x^s \}_{x \in \Sigma_T^+} \) are absolutely continuous \( (10) \) with respect to each other, this integral is uniformly comparable to

\[
\hat{\mu}^s([\hat{y} \in \mathcal{W}^{st}_{loc}(x) : P(\sigma^{-n} \hat{y}) = y])
\cdot \mu([x_0]) \asymp \hat{\mu}_x^s([\hat{y} \in \mathcal{W}^{st}_{loc}(x) : P(\sigma^{-n} \hat{y}) = y]) \cdot g(x).
\]

Noting that \( \sigma^n y = x \), canceling off \( g(x) \) from both sides proves the lemma. \( \square \)
We will now switch to a notation more suited to the two-sided subshift \( \Sigma_T \) by using the adjoint cocycle \( A_\sigma \) over \( (\Sigma_T, \sigma^{-1}) \).

Consider any \( x, y \in \Sigma_T^+ \) with \( \sigma^n y = x \) and any \( \sigma^{-n} \hat{y}_1, \sigma^{-n} \hat{y}_2 \in \mathcal{W}_n^\Sigma (y) \). Since \( A \) is constant along the local stable sets \( (8) \), we have \( A^n_\sigma (\hat{y}_1) = A^n_\sigma (\hat{y}_2) \). This observation together with the previous lemma give

\[
t_{n, \alpha} (x) \propto \sup_{\bar{u}, \bar{v}} \int \left[ \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right]^\alpha d\hat{\mu}_x (\hat{x}) .
\]

Similar to the arguments from [BL85, Ch. 5, Lemma 2.3], in order to establish (20) for all \( \alpha > 0 \) sufficiently small, it suffices to show that there exists \( n \in \mathbb{N} \) such that

\[
\sup_{\bar{u}, \bar{v}} \int_{x \in \Sigma_T^+} \log \left( \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right) d\hat{\mu}_x (\hat{x}) < 0 . \tag{21}
\]

Indeed assuming (21), we establish (20) as follows: let \( \alpha_0 = \alpha_0(n) > 0 \) be a sufficiently small constant such that \( \alpha \log \left( \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right) \) is sufficiently close to 0 for all \( \alpha \in (0, \alpha_0) \), \( \hat{x} \in \Sigma_T \), and \( (\bar{u}, \bar{v}) \in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1} \setminus \Delta \). Applying the identity

\[
er^r \leq 1 + r + \frac{r^2}{2} e^{|r|} \text{ for } |r| \approx 0
\]

to \( r = \alpha \log \left( \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right) \) and integrating with respect to \( \hat{\mu}_x \), we have

\[
\int \left( \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right)^\alpha d\hat{\mu}_x (\hat{x}) \leq 1 + \alpha \int \log \left( \frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \right) d\hat{\mu}_x (\hat{x}) .
\]

Taking the supremum over all \( x \in \Sigma_T^+ \) and \( (\bar{u}, \bar{v}) \in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1} \setminus \Delta \) followed by the logarithm, (21) then translates to the required inequality \( \log w_{n, \alpha} < 0 \) of (20).

So far, we have reduced proving Theorem 4 to establishing (19) and (21). Note that we did not yet make use of the 1-typicality assumption on \( A \) in deductions thus far.

### 3.3. Making use of the gap in Lyapunov exponents

Denoting the \( i \)-th singular value of \( A \) by \( \sigma_i (A) \), the angular metric \( d \) defined as in (6) satisfy

\[
\frac{d(A^n_\sigma (\hat{x}))u, A^n_\sigma (\hat{x})v}{d(\bar{u}, \bar{v})} \leq \frac{\sigma_1 (A^n_\sigma (\hat{x})) \cdot \sigma_2 (A^n_\sigma (\hat{x}))}{\| A^n_\sigma (\hat{x})u \| \cdot \| A^n_\sigma (\hat{x})v \|} . \tag{22}
\]

Using this inequality together with the 1-typicality assumption on \( A \), we will establish (21) in this subsection.

We begin with a general result which follows from Kingman’s subadditive ergodic theorem. In particular, it does not require any assumption on \( A \).

**Proposition 3.** The Lyapunov exponents \( \{ \lambda_i \}_{1 \leq i \leq d} \) of \( A \) with respect to \( \hat{\mu} \) listed in a decreasing order satisfy the following:
1. For $\hat{\mu}$-a.e. $\hat{x}$,

$$
\lim_{n \to \infty} \frac{1}{n} \log \| A_n^k (\hat{x}) \| = \sum_{i=1}^{k} \lambda_i.
$$

2. For every $x \in \Sigma_T^+$ and $\hat{\mu}_x^s$-a.e. $\hat{x}$,

$$
\lim_{n \to \infty} \frac{1}{n} \log \| A_n^k (\hat{x}) \| = \sum_{i=1}^{k} \lambda_i.
$$

3. The convergence in

$$
\lim_{n \to \infty} \frac{1}{n} \int \log \| A_n^k (\hat{x}) \| \, d\hat{\mu}_x^s (\hat{x}) = \sum_{i=1}^{k} \lambda_i.
$$

is uniform in $x \in \Sigma_T^+$.

Proof. 1. This is a consequence of the work of Furstenberg and Kesten [FK60]. It can also be deduced from the Kingman’s subadditive ergodic theorem applied to the adjoint cocycle $A^*$.

2. The full $\hat{\mu}$-measure subset in which the statement of 1. holds is unstable-saturated; that is, if $\hat{x}$ belongs to such a subset, then its entire local unstable set $W^u_{\text{loc}} (\hat{x})$ also belongs to the subset. Hence, we can promote “$\hat{\mu}$-a.e. $\hat{x}$” from 1. into “every $x \in \Sigma_T^+$ and $\hat{\mu}_x^s$-a.e. $\hat{x}$.”

3. It is clear from 2. that the integral converges for every $x \in \Sigma_T^+$. For the uniform convergence, we will suppose that $k = 1$ for simplicity. Since $\Sigma_T^+$ is compact, it suffices to show that given any $x_n \to x \in \Sigma_T^+$ we have $f_n (x_n) \to \lambda_1$ where

$$
f_n (x) = \frac{1}{n} \int \log \| A_n^k (\hat{x}) \| \, d\hat{\mu}_x^s (\hat{x}).
$$

From the absolute continuity (10) of $\{ \mu_x^s \}_{x \in \Sigma_T^+}$ and the Hölder continuity of the Jacobian $J_{x,y}$, we have a following bound on $f_n (x_n)$:

$$
f_n (x_n) = \frac{1}{n} \int \log \| A_n^k (\hat{x}) \| \, d\hat{\mu}_x^s (\hat{x})
$$

$$
= \frac{1}{n} \int J_{x_n, x} (\hat{x}) \cdot \log \| A_n^k (\hat{x}) \| \, d\hat{\mu}_x^s (\hat{x})
$$

$$
\leq \left( 1 + Cd (x_n, x)^{\gamma} \right) \cdot \frac{1}{n} \int \log \| A_n^k (\hat{x}) \| \, d\hat{\mu}_x^s (\hat{x})
$$

for some uniform constants $C, \gamma > 0$. As $x_n$ limits to $x$, the upper bound limits to $\lambda_1$ as $n \to \infty$. Likewise, $f_n (x_n)$ admits a similar lower bound which also limits to $\lambda_1$. This completes the proof. ⊓⊔

The following proposition relies on the construction and the properties of the measures considered in Sect. 2.3.
Proposition 4. If \( A \) is 1-typical, then for every \( x \in \Sigma_T^+ \), \( \hat{\mu}^x \)-a.e. \( \hat{x} \), and any unit vector \( u \in \mathbb{R}^d \setminus \{0\} \),

\[
\lim_{n \to \infty} \frac{1}{n} \int \log \| A^u_n(\hat{x})u \| \, d\hat{\mu}^x_n(\hat{x}) = \lambda_1. \tag{23}
\]

Moreover, the convergence is uniform in \( x \) and \( u \).

In order to prove this proposition, we need the following preliminary lemma where we use the same notations from Remark 1. In particular, \( \xi(\hat{x}) \) is the one-dimensional top Lyapunov subspace of \( A \) with respect to \( \hat{\mu} \), and we denote by \( H^{s,*}_\xi \) the local stable holonomies of the adjoint cocycle \( A \) over \((\Sigma_T, \sigma^{-1})\).

Lemma 3. Let \( A \) be 1-typical. Then for any \( x \in \Sigma_T^+ \), \( \hat{\mu}^x \)-a.e. \( \hat{x} \), and any sequence of unit vectors \( \{u_n\}_{n \in \mathbb{N}} \) converging to some \( u \in \mathbb{R}^d \setminus \{0\} \), we have

\[
\lim_{n \to \infty} \frac{\| A^u_n(\hat{x})u_n \|}{\| A^u_n(\hat{x}) \|} = |\langle u, \xi(\hat{x}) \rangle|. 
\]

Moreover, for any \( y \in \Sigma_T^+ \) with the same 0-symbol as \( \hat{x} \), we have

\[
\lim_{n \to \infty} \frac{\| A^u_n(\hat{x})H^{s,*}_\xi \hat{y} u \|}{\| A^u_n(\hat{x}) \|} = |\langle v, \xi(\hat{y}) \rangle| 
\]

where \( \hat{y} := [\hat{x}, y] \).

Proof. Let \( B := \{ \hat{x} \in \Sigma_T : \xi(\hat{x}) \text{ is well-defined} \} \). Then \( \hat{\mu}(B) = 1 \) and \( B \) is unstable-saturated as \( \xi(\hat{x}) \) is \( H^u \)-invariant.

For \( \hat{x} \in B \), we have seen in Remark 1 that \( \xi(\hat{x}) \) is equal to \( K_\infty e_1 \), where \( K_n \) is from the \( KAK \)-decomposition \( A^n(\sigma^{-n}\hat{x}) = K_n A_n U_n \) and \( K_\infty \) is the subsequent limit of \( K_n \). Then \( A^u_n(\hat{x}) \) is equal to \( (K_n A_n U_n)^* = U_n^* A_n^* K_n^* \). Since \( U_n \in O(d) \), we have

\[
\frac{\| A^u_n(\hat{x})u_n \|^2}{\| A^u_n(\hat{x}) \|^2} = \frac{\| (U^*_n A_n K_n^*)u_n \|^2}{\| A^u_n(\hat{x}) \|^2} = \frac{\| A_n^* K_n^* u_n \|^2}{\| A^u_n(\hat{x}) \|^2} \xrightarrow{n \to \infty} \langle K_n^* u, e_1 \rangle^2 = \langle u, \xi(\hat{x}) \rangle^2. 
\]

Since \( B \) is unstable-saturated, this holds true for every \( x \) and \( \hat{\mu}^x \)-a.e. \( \hat{x} \).

For the second statement, using the identity \( H^{s,*}_{\hat{x},\hat{y}} = (H^{u}_{\hat{x},\hat{y}})^* \) we have

\[
\lim_{n \to \infty} \frac{\| A^u_n(\hat{x})H^{s,*}_{\hat{x},\hat{y}} \hat{y} u \|}{\| A^u_n(\hat{x}) \|} = |\langle H^{s,*}_{\hat{x},\hat{y}} \hat{y} v, \xi(\hat{x}) \rangle| = |\langle v, H^u_{\hat{x},\hat{y}} \hat{y} \xi(\hat{x}) \rangle|. 
\]

The claim then follows by noticing \( \xi(\hat{y}) = H^u_{\hat{x},\hat{y}} \xi(\hat{x}) \) since the top Lyapunov subspace \( \xi(\hat{x}) \) is preserved under \( H^u \). \( \square \)

Proof of Proposition 4. For any \( x \in \Sigma_T^+ \) and unit \( u \in \mathbb{R}^d \setminus \{0\} \), we define

\[
B_{x,u} := \{ \hat{x} \in \mathcal{W}^s_{\text{loc}}(\hat{x}) : u \perp \xi(\hat{x}) \} = \{ \hat{x} \in \mathcal{W}^s_{\text{loc}}(\hat{x}) : \xi(\hat{x}) \in u^\perp \}. 
\]

Recalling the \( F_A \)-invariant probability measure \( m \) from Proposition 1, we have

\[
\hat{\mu}^x(B_{x,u}) = \hat{\mu}^x(\{ \hat{x} \in \mathcal{W}^s_{\text{loc}}(x) : \xi(\hat{x}) \in u^\perp \}) = \int_{\xi(\hat{x})} \delta_{\xi(\hat{x})}(u^\perp) d\hat{\mu}^x(\hat{x}) = m_x(u^\perp) = 0. 
\]
The last equality is due to the fact that \( m_x \) is proper for every \( x \in \Sigma_T^\pm \). Then Proposition 3 (2) and Lemma 3 give

\[
\lim_{n \to \infty} \frac{1}{n} \log \| A^n(x) u \| = \lambda_1
\]

for every \( x \in \Sigma_T^\pm \) and \( \hat{\mu}_x \)-a.e. \( \hat{x} \), which then yields (23).

Since \( \Sigma_T^\pm \times S^{d-1} \) is compact, the uniform convergence of (23) in \( x \in \Sigma_T^\pm \) and \( u \in S^{d-1} \) would follow if for any \( x_n \to x \in \Sigma_T^\pm \) and \( u_n \to u \in S^{d-1} \), we have

\[
\lim_{n \to \infty} \frac{1}{n} \int \log \| A^n(x) u_n \| d\hat{\mu}_{x_n}(\hat{x}) = \lambda_1.
\]  

(24)

In order to establish (24), we consider a full \( \hat{\mu}_x \)-measure subset \( U_x \) of \( \mathcal{W}_{\text{loc}}^s(x) \) such that for any \( \hat{x} \in U_x \), we have

1. \( \xi(\hat{x}) \) exists, and
2. \( \hat{x} \) does not belong to \( B_{x,u} \), and
3. \( \lim_{n \to \infty} \frac{1}{n} \log \| A^n(\hat{x}) \| = \lambda_1. \)

For any \( \hat{x} \in U_x \) and \( n \in \mathbb{N} \), let \( \hat{x}_n := [\hat{x}, x_n] \) be the unique point in the intersection between \( \mathcal{W}_{\text{loc}}^s(x_n) \) and \( \mathcal{W}_{\text{loc}}^u(\hat{x}) \). Then we have

\[
A^n(\hat{x}_n) u_n = H^{s,*}_{\sigma^{-n} \hat{x}, \sigma^{-n} \hat{x}_n} A^n_*(\hat{x}) H^{s,*}_{\hat{x}_n, \hat{x}} u_n.
\]

Denoting \( H^{s,*}_{\hat{x}_n, \hat{x}} u_n \) by \( \tilde{u}_n \), we have \( \tilde{u}_n \to u \) since \( \hat{x}_n \to \hat{x} \) and \( u_n \to u \). Moreover, \( H^{s,*}_{\sigma^{-n} \hat{x}, \sigma^{-n} \hat{x}_n} \to I \) as \( n \to \infty \). The first statement of Lemma 3 then gives

\[
\frac{\| A^n(\hat{x}_n) u_n \|}{\| A^n_*(\hat{x}) \|} = \frac{\| H^{s,*}_{\sigma^{-n} \hat{x}, \sigma^{-n} \hat{x}_n} A^n_*(\hat{x}) \tilde{u}_n \|}{\| A^n_*(\hat{x}) \|} \to_{n \to \infty} |(u, \xi(\hat{x}))|.
\]

Since the disintegration \( \{ \hat{\mu}_x \}_{x \in \Sigma_T^\pm} \) is continuous in \( x \in \Sigma_T^\pm \) from the absolute continuity (10) of \( \{ \hat{\mu}_x \}_{x \in \Sigma_T^\pm} \) and \( \frac{1}{n} \log \| A^n_*(\hat{x}) \| \) limits to \( \lambda_1 \) from the choice of \( \hat{x} \in U_x \), together we obtain (24) as required. \( \square \)

Then (21) easily follows: from (22) and Proposition 3 and 4, there exists \( n \in \mathbb{N} \) such that

\[
\frac{1}{n} \int \log \left( \frac{d(\mathcal{L}(\hat{x}, A^n_*(\hat{x}) u), A^n_*(\hat{x}) v)}{d(\bar{u}, \bar{v})} \right) d\hat{\mu}_x(\hat{x})
\]

\[
= \frac{1}{n} \int \left( \log \| A^n_*(\hat{x}) \|^2 \| - \log \| A^n_*(\hat{x}) u \| - \log \| A^n_*(\hat{x}) v \| \right) d\hat{\mu}_x(\hat{x})
\]

\[
< \frac{1}{2} (\lambda_2 - \lambda_1)
\]

\[
< 0
\]

for every \( x \in \Sigma_T^\pm \) and \( (\bar{u}, \bar{v}) \in \mathbb{P}^{d-1} \times \mathbb{P}^{d-1} \setminus \Delta \). In the last inequality, we have used the result of Bonatti and Viana [BV04] that the top Lyapunov exponent \( \lambda_1 \) is simple when \( \mathcal{A} \) is 1-typical. This establishes (21) as required.
3.4. Controlling $\tau_{n, \alpha}$. We will complete the proof of Theorem 4 by showing (19) which states that

$$\limsup_{n \to \infty} \frac{1}{n} \log \tau_{n, \alpha} < 0.$$ 

Similar to how we passed from $t_{n, \alpha}$ to the integral, we have

$$\tau_{n, \alpha} \propto \sup_{x, y \in \Sigma_T^+ \cap \mathbb{T}^{d-1}} \int \left( \frac{d(A^\alpha_*(\hat{x}))}{\rho(x, y)} \right) d\mu^x_\tau(\hat{x})$$

(25)

where the supremum is taken over all $x, y \in \Sigma_T^+$ with the common 0-symbol and all $\bar{v} \in \mathbb{T}^{d-1}$. This is because for any $x, y \in \Sigma_T^+$ with the common 0-symbol and $z_1 \in \sigma^{-n}x$ and $z_2 \in \sigma^{-n}y$ in the same $n$-cylinder, we have

$$d(A^{[\alpha]}(z_1), A^{[\alpha]}(z_2)) = d(A^{[\alpha]}_*(\hat{x}), A^{[\alpha]}_*(\hat{y}))$$

for any $\hat{x} \in \sigma^n(\mathcal{A}^{[\alpha]}_{\text{loc}}(z_1))$. Lemma 2 then gives (25).

Writing $A^{[\alpha]}_*(\hat{y}) = H^1_{\hat{y}} A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}}$ where $H^1_{\hat{y}} = H^{s, \ast}_{\sigma^{-n} \hat{y}, \sigma^{-n} \hat{y}}$ and $H^2_{\hat{y}} = H^{s, \ast}_x$, the integral appearing in (25) is bounded above by the sum

$$\int \left( \frac{d(H^1_{\hat{y}} A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v, A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v)}{\rho(x, y)} \right)^\alpha d\mu^x(\hat{x})$$

$$+ \int \left( \frac{d(A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v, A^{[\alpha]}_*(\hat{x}) v)}{\rho(x, y)} \right)^\alpha d\mu^x(\hat{x}),$$

and it suffices to show that both terms are decreasing exponentially fast to 0.

For the first term, we apply (7) with $u = A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v$:

$$d(H^1_{\hat{y}} u, \bar{v}) \leq C \cdot \rho(\sigma^{-n} \hat{x}, \sigma^{-n} \hat{y}) = C \cdot 2^{-n} \cdot \rho(x, y).$$

Hence, the first term is bounded above by $C^\alpha \cdot 2^{-n \alpha}$.

For the second term, apply (22):

$$d(A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v, A^{[\alpha]}_*(\hat{x}) v) \leq d(H^2_{\hat{y}} v, \bar{v}) \cdot \frac{\sigma_1(A^{[\alpha]}_*(\hat{x}))}{\|A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v\|} \cdot \frac{\sigma_2(A^{[\alpha]}_*(\hat{x}))}{\|A^{[\alpha]}_*(\hat{x}) v\|}$$

$$\leq C \cdot \rho(x, y) \cdot \frac{\sigma_1(A^{[\alpha]}_*(\hat{x}))}{\|A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v\|} \cdot \frac{\sigma_2(A^{[\alpha]}_*(\hat{x}))}{\|A^{[\alpha]}_*(\hat{x}) v\|}.$$ 

All terms in the product except $\|A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} (v)\|$ can be dealt with by proceeding similar to the above subsection. So it suffices to show the following analogue of (23): for any $x, y \in \Sigma_T^+$ with $x_0 = y_0$ and $v \in \mathbb{S}^{d-1}$, we have

$$\lim_{n \to \infty} \frac{1}{n} \int \log \|A^{[\alpha]}_*(\hat{x}) H^2_{\hat{y}} v\| d\mu^x(\hat{x}) = \lambda_1$$

(26)

and the convergence is uniform in $x, y \in \Sigma_T^+$ and $v \in \mathbb{S}^{d-1}$. 
Recalling that \( \hat{x}_y = [\hat{x}, y] \), we have
\[
\hat{\mu}_\lambda^x (\{ \hat{x} \in \mathcal{W}^\nu_{\text{loc}} (x) : v \perp \xi (\hat{x}_y) \}) = \int \delta_{\xi([\hat{x}, y])} (v^{-1}) d\hat{\mu}_\lambda^x (\hat{x}) \\
= \int J_{y,x} (\hat{y}) \delta_{\xi([h_{y,x}(\hat{y}), y])} (v^{-1}) d\hat{\mu}_\lambda^x (\hat{y})
\]
where the second equality is due to the absolute continuity (10) of \( \{ \hat{\mu}_\lambda^x \}_{x \in \Sigma_T^+} \). Since \([h_{y,x}(\hat{y}), y] = [\hat{y}, x], y = \hat{y} \), this gives
\[
\hat{\mu}_\lambda^x (\{ \hat{x} \in \mathcal{W}^\nu_{\text{loc}} (x) : v \perp \xi (\hat{x}_y) \}) = \int J_{y,x} (\hat{y}) \delta_{\xi(\hat{y})} (v^{-1}) d\hat{\mu}_\lambda^x (\hat{y}) = 0
\]
using the properness of \( m_v \) as in the proof for Proposition 4. This together with Proposition 3 and the second statement of Lemma 3 establishes (26). The uniform convergence in (26) can also be shown as in the proof of Proposition 4 using the absolute continuity of \( \{ \hat{\mu}_\lambda^x \}_{x \in \Sigma_T^+} \).

3.5. Quasi-compactness and the eigenmeasure of \( \mathcal{L} \). Using the Lasota-Yorke inequality established in Theorem 4, we have the following result:

**Proposition 5.** \( \mathcal{L} \) is quasi-compact.

In fact, this proposition can be deduced by applying the following result of Hennion to \( X = C(\Sigma_T^+ \times \mathbb{R}^{d-1}) \), \( T = \mathcal{L} \), \( \| \cdot \|_1 = \| \cdot \|_\sigma \), \( \| \cdot \|_2 = \| \cdot \|_\infty \). For details, see [Hen93] and references therein.

**Proposition 6 [Hen93].** Let \( (X, \| \cdot \|_1) \) be a Banach space and \( T : (X, \| \cdot \|_1) \to (X, \| \cdot \|_1) \) a bounded linear operator with spectral radius \( \rho(T) \). Suppose there exists a norm \( \| \cdot \|_2 \) such that \( T : (X, \| \cdot \|_1) \to (X, \| \cdot \|_2) \) is compact and there exist \( \{ R_n \}_{n \in \mathbb{N}}, \{ r_n \}_{n \in \mathbb{N}} \subset \mathbb{R} \) such that \( r := \liminf_{n \to \infty} (r_n)^{1/n} < \rho(T) \) and for every \( f \in X \)
\[
\| T^nf \|_1 \leq r_n \| f \|_1 + R_n \| f \|_2.
\]
Then \( T \) is quasi-compact and the essential spectrum of \( T \) is less than or equal to \( r \).

We now establish properties of the eigenmeasure for \( \mathcal{L} \). For any probability measure \( \xi \) on \( \Sigma_T^+ \times \mathbb{R}^{d-1} \) projecting to \( \mu \) and any continuous function \( f \in C(\Sigma_T^+ \times \mathbb{R}^{d-1}) \), consider the function \( \tilde{f} \) on \( \Sigma_T^+ \) defined by \( \tilde{f} (x) := \int f(x, \mathcal{A}(x)^* u) d\xi_{\sigma_x}(u) \). Recalling that \( \mu \) is the eigenfunction of \( L_{\log g} \), we have
\[
\int \tilde{f} \, d\mu = \int \tilde{f} (L_{\log g}^* \mu) = \int L_{\log g} \tilde{f} \, d\mu \\
= \int \sum_{\sigma y = x} g(y) \tilde{f} (y) \, d\mu (x) \\
= \int \sum_{\sigma y = x} g(y) \int f (y, \mathcal{A}(y)^* u) \, d\xi_x (u) \, d\mu (x) \\
= \int \mathcal{L} f \, d\xi.
\]
Comparing the conditional measures of both sides (with the definition of \( \tilde{f} \) in mind), this is equivalent to

\[
(\mathcal{L}^* \xi)_x = (\mathcal{A}(x)^*)_x \xi_x. \tag{27}
\]

This may also be observed from the identity

\[
(\mathcal{L} f_1) \cdot f_2 = \mathcal{L}(f_1 \cdot f_2 \circ F_{A^{-1}})
\]

where \( A^{-1} \) is the adjoint of the inverse cocycle for \( \mathcal{A} \).

**Proposition 7.** There exists a unique probability measure \( \nu \) on \( \Sigma^+_f \times \mathbb{P}^{d-1} \) such that \( \mathcal{L}^* \nu = \nu \). Moreover, \( \nu \) is \( F_{A^{-1}} \)-invariant and \( \nu_x \) is equal to \( \delta_{\xi^*_s(x)} \) for \( \mu \)-a.e. \( x \).

**Proof.** The existence of such \( \nu \) is standard: it follows from the fact that \( \mathcal{L}^* \) maps the space of probability measures to itself and that the set of probability measures is weak* compact. From (27) we have \( \nu_x = (A^{-1}_*(x))_x \nu_x \), which shows that \( \nu \) is \( F_{A^{-1}} \)-invariant.

For the remaining statements claimed in the proposition, it suffices to show that \( \nu \) is \((A^{-1}_*, H^s)\)-invariant. This is because the \((A^{-1}_*, H^s)\)-invariant measure is unique when \( \mathcal{A} \) (hence \( A^{-1}_* \)) is 1-typical, and it is characterized by its conditional measures supported on the slowest Lyapunov subspace \( \xi^*_s(x) \) with respect to \( A^{-1}_* \) and \( \mu \).

Being an eigenfunction of \( \mathcal{L} \), \( \nu \) has to project to \( \mu \) under \( \pi : \Sigma^+_f \times \mathbb{P}^{d-1} \to \Sigma^+_f \), and (11) applied to \( A^{-1}_* \) gives

\[
\hat{\nu}_x = \lim_{n \to \infty} (A^{-1}_*)^n (P(\sigma^{-n} \hat{x})) \nu_{P(\sigma^{-n} \hat{x})} = \lim_{n \to \infty} \nu_x = \nu_x.
\]

This implies that \( \hat{\nu}_x = \hat{\nu}_\hat{y} \) is for any \( \hat{y} \in \mathcal{W}^s_{loc}(\hat{x}) \). Since the local stable holonomies for \( A^{-1}_* \) are identically equal to \( I \) (due to the same reasoning why \( H^s = I \) for \( \mathcal{A} \)), it follows that \( \hat{\nu} \) is \((A^{-1}_*, H^s)\)-invariant. \( \square \)

3.6. \( \mathcal{L} \) defined with the inverse. In this subsection, we remark that all results obtained thus far equally hold if we had defined the operator with \( \mathcal{A}(y)^{-1} \) in place of \( \mathcal{A}(y)^* \):

\[
\mathcal{L}_{\text{inv}} f(x, \overline{u}) := \sum_{y : \sigma y = x} g(y) f(y, \mathcal{A}(y)^{-1} u).
\]

Theorem 4 holds for \( \mathcal{L}_{\text{inv}} \) with the statements of Proposition 4 and Lemma 3 replaced by

\[
\lim_{n \to \infty} \frac{1}{n} \int \log \| A^{-n}(\hat{x}) u \| d\hat{\mu}_s^*(\hat{x}) = -\lambda_d
\]

and for \( u_n \to u \),

\[
\lim_{n \to \infty} \frac{\| A^{-n}(\hat{x}) u_n \|}{\| A^{-n}(\hat{x}) \|} = |\langle u, \omega_s(\hat{x}) \rangle|.
\]

Moreover, the proof would make use of the gap in the bottom exponents \( \lambda_{d-1} > \lambda_d \) instead of the gap in the top exponents \( \lambda_1 > \lambda_2 \).

We have

\[
(\mathcal{L}_{\text{inv}} f_1) \cdot f_2 = \mathcal{L}_{\text{inv}}(f_1 \cdot f_2 \circ F_{\mathcal{A}}),
\]
and its eigenmeasure $\nu_{\text{inv}}$ of $\mathcal{L}_{\text{inv}}$ obtained as in Proposition 7 is unique and $F_A$-invariant. Moreover, $\nu_{\text{inv}}$ is $(A, H^s)$-invariant, so its conditional measure $\nu_{\text{inv},x}$ sits on the slowest Lyapunov subspace $\omega(x)$ with respect to $A$ and $\mu$. This observation is compatible with the fact that $\mathcal{L}_{\text{inv}}$ acts on the fiber $\mathbb{P}^{d-1}$ by the inverse cocycle $A^{-1}$.

4. The Peripheral Spectrum of $\mathcal{L}$

In this section, we will prove the following theorem assuming the results and terminologies from the Appendix (Sect. 6).

Theorem 5. Suppose $\Sigma^+_T$ is a shift of finite type defined by an irreducible $T$, and $\hat{A}$ is $1$-typical. Then for all $\alpha > 0$ sufficiently small, $\mathcal{L}$ is quasi-compact and the following are true: there exists $h \in \mathbb{N}$ such that

1. There exists a unique probability measure $\nu$ with $\mathcal{L}^* \nu = \nu$.
2. The only eigenvalues of $\mathcal{L}$ of modulus 1 are $\left\{ e^{\frac{2\pi i k}{h}} \right\}_{k=0}^{h-1}$ and each is simple.
3. $\mathcal{L}$ can be written as

$$\mathcal{L} = \sum_{k=0}^{h-1} \lambda_k P_k + S$$

where $\lambda_k = \exp\left( \frac{2k\pi i}{h} \right)$.

Moreover, $P_k^2 = P_k$, $P_i P_j = 0$ for $i \neq j$, $P_0 f = \langle f, \nu \rangle$, and $S$ is a bounded linear operator with $\rho(S) < 1$ and $P_i S = SP_i = 0$.

As $T$ is irreducible, so is $T^*$. Hence, up to reordering the symbols, we may assume that $T^*$ can written as a block matrix

$$T^* = \begin{bmatrix} 0 & T_{12}^* & 0 & \cdots & 0 \\ \vdots & 0 & T_{23}^* & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & T_{h-1h}^* & 0 \\ T_{h1}^* & 0 & \cdots & 0 & \end{bmatrix}$$

where the diagonal blocks are square. This implies a decomposition of $\mathbb{R}^n$ as

$$\mathbb{R}^n = \bigoplus_i W_i \text{ where } W_i = e_k \oplus e_{k+1} \oplus \cdots \oplus e_{k'}$$

and $e_i$ is the $i$-th standard basis vector of $\mathbb{R}^n$. This decomposition has the property that $T_{23}^* W_3 \subseteq W_2$. This then naturally partitions the symbols in classes $\{(p)\}_{p=1}^h$ based on the implied decomposition of $\mathbb{R}^n$. That is two elements of the symbols $i$ and $j$ are in the same partition element if the corresponding standard basis vectors $e_i$, $e_j$ are in the same subspace in this decomposition of $\mathbb{R}^n$. For convenience we will write

$$\{(p)\} = \bigcup_{i \in (p)} [i].$$
For any $f \in C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})$, we define the multiplication operator $M_f$ by

$$M_f h = fh.$$  

(28)

It is clear that $M_f$ is a bounded linear operator on the space $C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})$ for any $0 < \alpha \leq 1$. In the following proposition, along with $M_f$ we will also make use of the indicator function denoted by $\chi$.

**Proposition 8.**  
1. The space $C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})$ is isomorphic to the space $\bigoplus_i C^\alpha([[i]] \times \mathbb{P}^{d-1})$.  
2. When thought of as acting on $\bigoplus_i C^\alpha([[i]] \times \mathbb{P}^{d-1})$, $L$ can be written as $L = [L_{ij}]$ where

$$L_{ij} f(x, \bar{u}) = \sum_{k \in (j) : kx \in \Sigma_T^+} g(kx) f(\bar{kx}, \bar{\mathcal{A}(kx)^*u}) \chi_{[[i]] \times \mathbb{P}^{d-1}}(x, \bar{u}).$$

Notice that $L_{ij} = 0$ unless there exist $a \in (i)$ and $b \in (j)$ such that $T_{ab}^* = T_{ba} = 1$; this is equivalent to $j \equiv i + 1 \mod h$.

3. For each $1 \leq p \leq h$, we have

$$\mathcal{L} \chi_{[[p]] \times \mathbb{P}^{d-1}} = \begin{cases} \chi_{[[h]] \times \mathbb{P}^{d-1}} & p = 1 \\ \chi_{[[p-1]] \times \mathbb{P}^{d-1}} & p \geq 2. \end{cases}$$

**Proof.**  
1. Define the map

$$\text{id} : C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1}) \to \bigoplus_{i=1}^h C^\alpha([[i]] \times \mathbb{P}^{d-1})$$

by

$$\text{id}(f) = (f \cdot \chi_{[[1]] \times \mathbb{P}^{d-1}}, f \cdot \chi_{[[2]] \times \mathbb{P}^{d-1}}, \ldots, f \cdot \chi_{[[h]] \times \mathbb{P}^{d-1}}).$$

One can check that $\text{id}$ is linear and that $\|\text{id}(f)\| \leq \|f\|_\alpha$. Moreover the function

$$(f_1, f_2, \ldots, f_h) \mapsto \sum_{i=1}^h f_i$$

defines an inverse for $\text{id}$.  
2. We expand $L_{ij}$ out as follows, where the summations in the second and the third lines are over all $k$ with $kx \in \Sigma_T^+$:

$$L_{ij} f(x, \bar{u}) = M_{\chi_{[[i]] \times \mathbb{P}^{d-1}}} \mathcal{L} M_{\chi_{[[j]] \times \mathbb{P}^{d-1}}} f(x, \bar{u})$$

$$= \chi_{[[i]] \times \mathbb{P}^{d-1}}(x, \bar{u}) \sum_k g(kx) M_{\chi_{[[j]] \times \mathbb{P}^{d-1}}} f(kx, \bar{\mathcal{A}(kx)^*u})$$

$$= \chi_{[[i]] \times \mathbb{P}^{d-1}}(x, \bar{u}) \sum_k g(kx) \chi_{[[j]] \times \mathbb{P}^{d-1}}(kx, \bar{\mathcal{A}(kx)^*u}) f(kx, \bar{\mathcal{A}(kx)^*u})$$

$$= \sum_{k : k \in (j), kx \in \Sigma_T^+} g(kx) f(kx, \bar{\mathcal{A}(kx)^*u}) \chi_{[[i]] \times \mathbb{P}^{d-1}}(x, \bar{u}).$$
3. As above, the summation in the first line is over all \( k \) with \( kx \in \Sigma_T^+ 
olinebreak: \)
\[
L \chi_{\{(p)\} \times \mathbb{P}^{d-1}}(x, \overline{u}) = \sum_k g(kx) \chi_{\{(p)\} \times \mathbb{P}^{d-1}}(kx, A(kx)^* \overline{u})
= \sum_{k, k \in (p), kx \in \Sigma_T^+} g(kx)
= \begin{cases} 
\chi_{\{(h)\} \times \mathbb{P}^{d-1}}(x, \overline{u}) & p = 1 \\
\chi_{\{(p-1)\} \times \mathbb{P}^{d-1}}(x, \overline{u}) & p \geq 2 .
\end{cases}
\]

This completes the proof. \( \square \)

**Lemma 4.** Suppose that \( T \) is primitive with \( T^m > 0 \) for some \( m \in \mathbb{N} \). Given \( f \geq 0 \), suppose further that there exists a point \((z, \overline{w})\) such that \( L^n f(z, \overline{w}) = 0 \) for some \( n \in \mathbb{N} \). Then for any \( 0 \leq k \leq n \) and \( x \in \Sigma_T^+ \), there exists \((y, \overline{v})\) such that \( L^k f(y, \overline{v}) = 0 \) and \( \rho(x, y) \leq 2^{-(n-k-m)} \).

**Proof.** Notice that
\[
0 = L^n f(z, \overline{w}) = \sum_{\sigma^{n-k} y = z} g(n-k)(y)L^k f(y, A^{[n-k]}(y)w) .
\]
Thus \( L^k f(y, A^{[n-k]}(y)w) = 0 \) for all \( y \) such that \( \sigma^{n-k} y = z \). Among all such \( y \), there necessarily exists a \( y \) such that \( \rho(x, y) \leq 2^{-(n-k-m)} \). \( \square \)

**Lemma 5.** For each \( 1 \leq p \leq h \), the restriction \( L_{\langle p \rangle} := L^{|p|}_{C^a((p) \times \mathbb{P}^{d-1})} \) has a spectral gap.

**Proof.** Since \( L_{\langle p \rangle} \) is quasi-compact from Proposition 6, in view of Theorem 10 it suffices to show that \( L_{\langle p \rangle} \) is semi-positive with respect to the cone of non-negative functions.

Notice that if \( L^N_{\langle p \rangle} f > 0 \) for some \( N \in \mathbb{N} \), then \( L^n_{\langle p \rangle} f > 0 \) for all \( n \geq N \). Therefore, in view of Definition 7 and 8, it suffices to show the following statement: if \( f \geq 0 \) and there exists \((z_n, \overline{w}_n)_{n \in \mathbb{N}} \subset \{(p)\} \times \mathbb{P}^{d-1} \) with \( L^n_{\langle p \rangle} f(z_n, \overline{w}_n) = 0 \) for all \( n \in \mathbb{N} \), then \( L^n_{\langle p \rangle} f \) converges to 0.

Let \( x \) be any given point in \( \{(p)\} \). By Theorem 11 there exists \( M \in \mathbb{N} \) such that for any \( i, j \in (p) \), there exists a word \( l \) of length \( M \) such that \( ilj \) is a valid word of \( \Sigma_T \).

Thus for any \( n \in \mathbb{N} \), we can apply Lemma 4 to \( L^{2n+M}_{\langle p \rangle} f(z_{2n+M}, \overline{w}_{2n+M}) = 0 \) with \( M, 2n + M \), and \( n \) playing the roles of \( m, n, \) and \( k \) respectively. This gives \((y, \overline{v})\) such that \( L^n_{\langle p \rangle} (y, \overline{v}) = 0 \) and \( \rho(x, y) \leq 2^{-nh} \).

Noticing that \( L^{|p|}_{\langle p \rangle} f \) agrees with \( L^{nh} f \) for \( f \in C((\{p\}) \times \mathbb{P}^{d-1}) \), Corollary 2 then gives for any \( \overline{u} \in \mathbb{P}^{d-1} \)
\[
\left| L^n_{\langle p \rangle} f(x, \overline{u}) \right| = \left| L^{nh} f(x, \overline{u}) - L^{nh} f(y, \overline{v}) \right|
\leq \| f \|_\alpha \beta^{nh} + C \| f \|_\infty \rho(x, y)^\alpha
\leq \| f \|_\alpha \beta^{nh} + C \| f \|_\infty 2^{-an h} .
\]

In particular, this shows that \( L^n_{\langle p \rangle} f(x, \overline{u}) \) converges to 0 for any \((x, \overline{u})\).

It is clear that point-mass measures \( \{ \delta(x, \overline{u}) \} \) form a sufficient collection for the cone of non-negative functions, and hence we conclude that \( L_{\langle p \rangle} \) is semi-positive. \( \square \)
Lemma 6. The algebraic multiplicity of the eigenvalue 1 for \( L \) is one.

**Proof.** Given \( f = (f_1, \ldots, f_h) \in \ker(I - L) \), we have \( L_p f_p = f_p \) for each \( 1 \leq p \leq h \). From Lemma 5, this implies that \( f_p \) is a constant multiple of the function \( \chi_{\{(p)\times \mathbb{P}^{d-1}\}} \). Hence \( f \in \text{span}\ \{\chi_{\{(p)\times \mathbb{P}^{d-1}\}} : 1 \leq p \leq h\} \).

Now consider the action of \( L \) on the subspace \( \text{span}\ \{\chi_{\{(p)\times \mathbb{P}^{d-1}\}} : 1 \leq p \leq h\} \). Notice from Proposition 8 and Lemma 5 that

\[
L \chi_{\{(p)\times \mathbb{P}^{d-1}\}} = \chi_{\{(p-1)\times \mathbb{P}^{d-1}\}} \text{ for } p \geq 2 \text{ and } L \chi_{\{(1)\times \mathbb{P}^{d-1}\}} = \chi_{\{(h)\times \mathbb{P}^{d-1}\}}.
\]

When written as a matrix with basis \( \{\chi_{\{(p)\times \mathbb{P}^{d-1}\}} : 1 \leq p \leq h\} \), the operator \( L \) is exactly

\[
\begin{bmatrix}
0 & 1 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
1 & 0 & \cdots & 0
\end{bmatrix}
\]

This matrix is irreducible and its only eigenvector corresponding to the eigenvalue 1 is the constant function 1. Therefore \( f = \alpha 1 \) for some \( \alpha \in \mathbb{C} \).

Next we show that the order of the eigenvalue 1 is one. We need to show that \( \ker(I - L)^2 = \ker(I - L) \).

The inclusion \( \ker(I - L) \subseteq \ker(I - L)^2 \) is clear, so we must show the other inclusion. If \( f \in \ker(I - L)^2 \), then \( f - L f \in \ker(I - L) \) and thus \( f - L f = \alpha 1 \) for some \( \alpha \in \mathbb{C} \). Then

\[
0 = \langle f - L f, v \rangle = \alpha \langle 1, v \rangle = \alpha,
\]

where \( v \) is the eigenmeasure of \( L \) from Proposition 7. Therefore \( f - L f = 0 \) and \( f \in \ker(I - L) \). \( \square \)

**Proof of Theorem 5.** The quasi-compactness of \( L \) is already shown in Proposition 6.

1. This follows from Proposition 7.
2. It is clear that 1 is an eigenvalue of \( L \). From Proposition 17 the spectrum of \( L \) is invariant under multiplication by \( e^{\frac{2\pi i}{h}} \). Thus for each \( 0 \leq k \leq h - 1 \), \( e^{\frac{2\pi i k}{h}} \) is an eigenvalue of \( L \). As the only eigenvalue of modulus 1 for \( L^h \) is 1, this implies that the only eigenvalues of modulus 1 are \( \left\{ e^{\frac{2\pi i k}{h}} \right\}_{k=0}^{h-1} \). By Lemma 6 the eigenvalue 1 is simple, and thus by Proposition 17 the eigenvalue \( e^{\frac{2\pi i k}{h}} \) is simple for each \( k \).
3. Using spectral projections, \( L \) can be written as

\[
L = \sum_{k=0}^{h-1} e^{\frac{2\pi i k}{h}} P_k + S
\]

where \( P_k \) is the spectral projection onto the eigenspace associated to \( e^{\frac{2\pi i k}{h}} \) and \( \rho(S) < 1 \).
Notice that as the eigenvalues \( \left\{ e^{\frac{2\pi i}{n}k} \right\}_{k=0}^{h-1} \) are simple, \( P_k \) is a projection onto the one-dimensional subspace \( \ker(e^{\frac{2\pi i}{n}k}I - \mathcal{L}) \) and it can be written as
\[
P_k = \lim_{\xi \to e^{\frac{2\pi i}{n}k}} (\xi - e^{\frac{2\pi i}{n}k}) R(\xi, \mathcal{L})
\]
where \( R(\xi, \mathcal{L}) = (\xi - \mathcal{L})^{-1} \) is the resolvent of \( \mathcal{L} \). As \( \mathcal{L} \) commutes with \( R(\xi, \mathcal{L}) \) we have that \( P_k \mathcal{L} = \mathcal{L} P_k = e^{\frac{2\pi i}{n}k} P_k \). Notice
\[
e^{\frac{2\pi i}{n}k} P_k P_k' = \mathcal{L} P_k P_k' = P_k \mathcal{L} P_k' = e^{\frac{2\pi i}{n}k'} P_k P_k'
\]
for \( k \neq k' \), so it must be that \( P_k P_k' = 0 \). Finally notice that
\[
SP_j = \left( \mathcal{L} - \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n}k} P_k \right) P_j = \mathcal{L} P_j - e^{\frac{2\pi i}{n}j} P_j = 0
\]
and similarly \( P_j S = 0 \). \( \square \)

**Corollary 3.** Suppose \( T \) is primitive, \( \Sigma_T^+ \) is the shift of finite type defined by \( T \), and \( \hat{A} \) is 1-typical. Then for all \( \alpha > 0 \) sufficiently small the operator \( \mathcal{L} : C^\alpha(\Sigma_T^+ \times \mathbb{R}^{d-1}) \to C^\alpha(\Sigma_T^+ \times \mathbb{R}^{d-1}) \) has a spectral gap.

**Proof.** If \( T \) is primitive, then \( h = 1 \) and the result follows from Theorem 5. \( \square \)

### 5. Applications of Theorem 5

In this section, we will prove Theorem 1 and 2. However, instead of directly proving them for \( A \), we will prove them for the adjoint cocycle \( A_* \). Recall that \( A \) is 1-typical if and only if \( A_* \) is.

First, we make a few simplifications as done in Sect. 2; that is, we assume that \( g : \Sigma_T^+ \to \mathbb{R} \) is a Hölder continuous \( g \)-function, and let \( \mu \) be its equilibrium state obtained as the eigenmeasure of \( L_{\log g} \). We denote the \( \sigma \)-invariant measure on \( \Sigma_T \) that projects to \( \mu \) by \( \hat{\mu} \). We will use \( \lambda_1 \) to denote the common top Lyapunov exponent:
\[
\lambda_1 := \lambda_1(A, \mu) = \lambda_1(A, \hat{\mu}) = \lambda_1(A_*, \hat{\mu}).
\]

Recalling the notation \( A^{[n]}(x) = [A^n(x)]^* \) from Sect. 3, we have
\[
\int \log \| A^{[n]}(x) u \| \, d\mu = \int \log \| A(\hat{x})^* \ldots A(\sigma^{n-1}\hat{x})^* u \| \, d\hat{\mu} = \int \log \| A_n^*(\hat{x}) u \| \, d\hat{\mu},
\]
where the first equality is because \( A \) is constant along the local stable sets and the second equality is due to \( \sigma \)-invariance of \( \hat{\mu} \). Hence, for the adjoint cocycle \( A_* \), \( \text{Var}(\hat{A}, \hat{\mu}) \) defined in (1) may alternatively be described by
\[
\text{Var} = \lim_{n \to \infty} \frac{1}{n} \int \left( \log \| A^{[n]}(x) u \| - n\lambda_1 \right)^2 \, d\mu.
\]

In the same vein, because the distribution of \( \left( \log \| A_n^*(\hat{x}) u \| - n\lambda_1 \right) \) with respect to \( \hat{\mu} \) agrees with that of \( \left( \log \| A^{[n]}(x) u \| - n\lambda_1 \right) \) with respect to \( \mu \), it suffices to establish the claimed results for the later distribution; see Theorem 6 and 7.
5.1. Central limit theorem. In this section we will prove the central limit theorem (Theorem 1). The proof uses the spectral properties of the operator $L$ as well as the standard functional analytic proof for central limit theorems.

**Theorem 6.** Suppose $\hat{A} : \Sigma_T \to GL_d(\mathbb{R})$ is $1$-typical. If $\text{Var} > 0$, then

$$\frac{\log \| A^{[n]}(x)u \| - n\lambda_1}{\sqrt{n}} \xrightarrow{\text{dist}} \mathcal{N}(0, \text{Var}).$$

If $\text{Var} = 0$, then

$$\frac{\log \| A^{[n]}(x)u \| - n\lambda_1}{\sqrt{n}} \xrightarrow{n \to \infty} 0.$$  

We begin by defining a function $\psi_n : \Sigma_T^+ \times \mathbb{P}^{d-1} \to \mathbb{R}$ by

$$\psi_n(x, \overline{u}) := \log \left\| A^{[n]}(x) \frac{u}{\|u\|} \right\|.$$  

Moreover, for $n \geq 1$ and $z \in \mathbb{C}$, we set

$$S_{n,z}(x, \overline{u}) := \left\| A^{[n]}(x) \frac{u}{\|u\|} \right\|^z.$$  

For $z \in \mathbb{C}$ we define the operator

$$\mathcal{L}_z f(x, \overline{u}) := \sum_{\sigma_{y=x}} g(y) \left\| A(y)^* \frac{u}{\|u\|} \right\|^z f(y, \overline{A(y)^*u}).$$  

Notice that when $z = 0$, $\mathcal{L}_0$ agrees with the operator $L$ considered in the previous sections.

In addition recall the definition of the transfer operator $L_{\log g} : C(\Sigma_T^+) \to C(\Sigma_T^+)$ defined by

$$L_{\log g} f(x) = \sum_{\sigma_{y=x}} g(y) f(y).$$  

Observe that $L_{\log g}$ could also be considered as acting on $C(\Sigma_T^+ \times \mathbb{P}^{d-1})$ by

$$L_{\log g} f(x, \overline{u}) = \sum_{\sigma_{y=x}} g(y) f(y, \overline{u}).$$

**Lemma 7.** 1. For any $n \geq 1$,

$$\mathcal{L}_z^n f(x, \overline{u}) = \sum_{\sigma^n_{y=x}} g^{(n)}(y) \left\| A^{[n]}(y) \frac{u}{\|u\|} \right\|^z f(y, \overline{A^{[n]}(y)u}).$$

2. For any $n \geq 1$ and $u \in \mathbb{R}^d$ we have that

$$\mathcal{L}_z^n \mathbf{1}(x, \overline{u}) = L_{\log g}^n S_{n,z}(x, \overline{u})$$
3. For any unit vector $u \in \mathbb{R}^d$,

$$
\int \left\| A^{[n]}(x)u \right\|^z d\mu(x) = \int \mathcal{L}_z^n 1(x, \vec{u}) d\mu(x) 
$$

(31)

Proof. 1. Notice that

$$
\mathcal{L}_z^2 f(x, \vec{u}) = \sum_{\sigma y_1 = x} g(y_1) \left\| A(y_1)^* \frac{u}{\|u\|} \right\|^z \mathcal{L}_z f(y, \overline{A(y_1)u})
$$

$$
= \sum_{\sigma y_1 = x} \left[ g(y_1) \left\| A(y_1)^* \frac{u}{\|u\|} \right\|^z \times \left( \sum_{\sigma y_2 = y_1} g(y_2) \left\| A(y_2)^* \frac{A(y_1)^*u}{\|A(y_1)^*u\|} \right\|^z f(y_2, \overline{A(y_2)A(y_1)u}) \right) \right]
$$

$$
= \sum_{\sigma z y = x} g(y) g(\sigma y) \left\| A(y)^* A(\sigma y)^* \frac{u}{\|u\|} \right\|^z f(y, \overline{A(y)^* A(\sigma y)^*u})
$$

$$
= \sum_{\sigma z y = x} g^{(2)}(y) \left\| A^{[2]}(y)^* \frac{u}{\|u\|} \right\|^z f(y, \overline{A^{[2]}(y)u}).
$$

The general case for $n \geq 1$ is similar.

2. Notice that

$$
L_{\log g}^n S_{n,z}(x, \vec{u}) = \sum_{\sigma z y = x} g^{(n)}(y) S_{n,z}(y, \vec{u})
$$

$$
= \sum_{\sigma z y = x} g^{(n)}(y) \left\| A^{[n]}(y)^* \frac{u}{\|u\|} \right\|^z
$$

$$
= \mathcal{L}_z^n 1(x, \vec{u}).
$$

3. Since $(L_{\log g})^* \mu = \mu$, from the definition of $S_{n,z}(x, \vec{u})$ we have

$$
\int \left\| A^{[n]}(x)u \right\|^z d\mu(x) = \int S_{n,z}(x, \vec{u}) d(L_{\log g}^n)^* \mu(x) = \int L_{\log g}^n S_{n,z}(x, \vec{u}) d\mu(x).
$$

Now the claim follows from 2. $\square$

For the following lemma, define

$$
C_{A^{[n]}} f(x, \vec{u}) := f(x, \overline{A^{[n]}(x)u})
$$

for each $n \in \mathbb{N}$. Recall also the definition for $M_f$ and $\psi_n$ from (28) and (29).

Lemma 8. 1. For any $n \in \mathbb{N}$, $C_{A^{[n]}}$ is a bounded linear operator on $C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})$.

2. For any $z \in \mathbb{C}$,

$$
\mathcal{L}_z^n = \sum_{k=0}^{\infty} \frac{z^k}{k!} L_{\log g}^n M_{\psi_n}^k C_{A^{[n]}}
$$

where we interpret $M_{\psi_n}^0 = I$. In particular, $z \mapsto \mathcal{L}_z$ is analytic.
2. Notice that \( C \) is linear. Notice that for \( f \in C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1}) \) and \( (x, \overline{u}), (y, \overline{w}) \in \Sigma_T^+ \times \mathbb{P}^{d-1} \) we have

\[
|C_{\mathcal{A}[1]} f(x, \overline{u}) - C_{\mathcal{A}[1]} f(y, \overline{w})| = \left| f(x, \mathcal{A}(x) u) - f(y, \mathcal{A}(y) w) \right| \\
\leq |f|_\alpha \cdot d((x, \mathcal{A}(x) u), (y, \mathcal{A}(y) w)) \\
= |f|_\alpha \cdot \max \left\{ \rho(x, y), d(\mathcal{A}(x) u, \mathcal{A}(y) w) \right\}^\alpha.
\]

We have that

\[
d(\mathcal{A}(x) u, \mathcal{A}(y) w) = \frac{\|\mathcal{A}(x) u \wedge \mathcal{A}(y) w\|}{\|\mathcal{A}(x) u\| \cdot \|\mathcal{A}(y) w\|} \\
\leq \frac{\|\mathcal{A}(x) u \wedge \mathcal{A}(y) w\|}{\|\mathcal{A}(x) u\| \cdot \|\mathcal{A}(y) w\|} \\
\leq \left( \sup_{x, y} \frac{\|\mathcal{A}(x) u \wedge \mathcal{A}(y) w\|}{\sigma_d(\mathcal{A}(x) u) \sigma_d(\mathcal{A}(y) w)} \right) \cdot d(\overline{u}, \overline{w}).
\]

Therefore

\[
|C_{\mathcal{A}[1]} f|_\alpha \leq |f|_\alpha \left( \sup_{x, y} \frac{\|\mathcal{A}(x) u \wedge \mathcal{A}(y) w\|}{\sigma_d(\mathcal{A}(x) u) \sigma_d(\mathcal{A}(y) w)} \right)^\alpha.
\]

As \( \|C_{\mathcal{A}[1]} f\|_\infty \leq \|f\|_\infty \), we have that \( C_{\mathcal{A}[1]} \) is bounded.

2. Notice that

\[
\left\| L_x^n M_{\psi_n}^k C_{\mathcal{A}[n]} \right\|_{op} \leq \left\| M_{\psi_n} \right\|_{op}^k \left\| L_x^n \right\|_{op} \left\| C_{\mathcal{A}[n]} \right\|_{op}.
\]

Thus for any \( z \in \mathbb{C} \)

\[
\sum_{k=0}^{\infty} \left| \frac{z^k}{k!} L_x^n M_{\psi_n}^k C_{\mathcal{A}[n]} \right| \leq \left\| L_x^n \right\|_{op} \left\| C_{\mathcal{A}[n]} \right\|_{op} \sum_{k=0}^{\infty} \left( \frac{|z| \left\| M_{\psi_n} \right\|_{op}}{k!} \right)^k \\
= \left\| L_x^n \right\|_{op} \left\| C_{\mathcal{A}[n]} \right\|_{op} e^{|z| \left\| M_{\psi_n} \right\|_{op}}.
\]

Hence

\[
\sum_{k=0}^{\infty} \left| \frac{z^k}{k!} L_x^n M_{\psi_n}^k C_{\mathcal{A}[n]} \right| \]

converges for all \( z \in \mathbb{C} \) and defines an analytic function from \( \mathbb{C} \) to \( B(C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1})) \).

Then for \( f \in C^\alpha(\Sigma_T^+ \times \mathbb{P}^{d-1}) \), we have

\[
\left( \sum_{k=0}^{\infty} \frac{z^k}{k!} L_x^n M_{\psi_n}^k C_{\mathcal{A}[n]} \right) f(x, \overline{u}) \\
= \sum_{k=0}^{\infty} \frac{z^k}{k!} \sum_{\sigma^n y = x} g^{(n)}(y) \left( \log \left\| C_{\mathcal{A}[n]}(y) \frac{u}{\|u\|} \right\| \right)^k f(y, \mathcal{A}[n](y) u)
\]
Thus by Proposition 17 the spectrum of \( \mathcal{L}_z \) is invariant under multiplication by \( e^{\frac{2\pi i}{n} k} \) for \( 0 \leq k \leq h - 1 \). Thus \( \mathcal{L}_z \) has \( h \) eigenvalues of modulus \( \rho(\mathcal{L}_z) \); in particular it must be that \( |\lambda_{k,z}| = \rho(\mathcal{L}_z) \) and \( \lambda_{k,z} := e^{\frac{2\pi i}{n} k} \lambda_{0,z} \). The remainder of the proof is similar to that of Theorem 5 (3). Finally to see that for any \( p \geq 1 \) there exist constants \( C \) and \( 0 < \beta < 1 \) such that

\[
\left\| \frac{d^p}{dz^p} S^n_z \right\| \leq C \beta^n,
\]

we refer the reader to [BL85, Chapter V Lemma 3.2(iv)]; while the assumptions in the lemma are different from ours, the proof still applies in our case. \( \square \)
Remark 2. Notice that $\rho_z = \lambda_{0,z}$. Thus for $\varepsilon \in \mathbb{R}$ as $L_\varepsilon$ preserves the cone of non-negative functions we have by Proposition 13 that $\rho_\varepsilon = \rho(L_\varepsilon)$. However, in general $\rho_z \neq \rho(L_\varepsilon)$ as the perturbation may cause $\lambda_{0,z}$ to move off the real line.

Proposition 10. Denoting by $\rho'_z$ and $\rho''_z$ the derivative and the second derivative of $z \mapsto \rho_z$, for any unit $u \in \mathbb{R}^d$ we have

1. $\rho'_0 = \lim_{n \to \infty} \frac{1}{n} \int \log \| A^n(x) u \| \, d\mu = \lambda_1$.
2. $\rho''_0 = \lim_{n \to \infty} \frac{1}{n} \int (\log \| A^n(x) u \| - n \lambda_1)^2 \, d\mu = \text{Var}$.

Proof. 1. For the first equality, recall from (31) that

$$\int \| A^n(x) u \|^z \, d\mu(x) = \int \mathcal{L}_z^n 1(x, \bar{u}) \, d\mu(x)$$

$$= \rho_z^n \int \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n} nk} P_{k,z} 1(x, \bar{u}) + S^n_z 1(x, \bar{u}) \, d\mu(x)$$

for any $z \in \mathbb{C}$. Taking the derivative of both sides give

$$\int \| A^n(x) u \|^z \log \| A^n(x) u \| \, d\mu = n \rho_z^{n-1} \rho'_z \int \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n} nk} P_{k,z} 1(x, \bar{u}) + S^n_z 1(x, \bar{u}) \, d\mu(x) (32)$$

$$+ \rho_z^n \int \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n} nk} P'_{k,z} 1(x, \bar{u}) + (S^n_z)' 1(x, \bar{u}) \, d\mu(x).$$

Noticing that $\rho_0 = 1$ and $S_0 1 = (S_0 P_0) 1 = 0$, evaluating (32) at $z = 0$ gives

$$\rho'_0 = \frac{1}{n} \int \log \| A^n(x) u \| \, d\mu - \frac{1}{n} \int \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n} nk} P'_{k,0} 1(x, \bar{u})$$

$$+ \frac{1}{n} \int (S^n_0)' 1(x, \bar{u}) \, d\mu(x).$$

Now taking the limit as $n \to \infty$ gives the result.

2. By dividing $A(x)$ by $e^{\lambda_1}$ if necessary, we may assume that $\lambda_1 = \rho'_0 = 0$. Differentiating (32) again gives

$$\rho''_0 = \frac{1}{n} \int (\log \| A^n(x) u \|)^2 \, d\mu - \frac{1}{n} \int \sum_{k=0}^{h-1} e^{\frac{2\pi i}{n} nk} P''_{k,0} 1(x, \bar{u})$$

$$+ \frac{1}{n} \int (S^n_0)'' 1(x, \bar{u}) \, d\mu(x).$$

By the same reasoning as in the proof of (1), the second term tends to 0 as $n \to \infty$. \Box

We are now ready to prove Theorem 6.
Proof of Theorem 6. As in the proof of Proposition 10 (2), we may assume that \( \lambda_1 = 0 \). By Proposition 9, we can write

\[ \mathcal{L}_z^n = \rho_z^n \left( \sum_{k=0}^{h-1} e^{2\pi i nk} P_{k,z} + S_z^n \right) \]

for \( |z| \) sufficiently small. By (31) we have that for \( n \) sufficiently large

\[
\int \left\| A^{[n]}(x)u \right\| \frac{it}{\sqrt{n}} d\mu(x) = \int \mathcal{L}_z^n \frac{it}{\sqrt{n}} 1(x, \overline{u}) d\mu(x) = \int \rho^n_{it/\sqrt{n}} \left( \sum_{k=0}^{h-1} e^{2\pi i nk} P_{k, it/\sqrt{n}} + S_{it/\sqrt{n}} \right) 1(x, \overline{u}) d\mu(x)
\]

\[
= \rho^n_{it/\sqrt{n}} \left( \sum_{k=0}^{h-1} \int e^{2\pi i nk} P_{k, it/\sqrt{n}} 1(x, \overline{u}) d\mu(x) \right) + \int S_{it/\sqrt{n}} 1(x, \overline{u}) d\mu(x). \]

Notice that \( z \mapsto P_{k,z} \) is continuous. Thus

\[
\lim_{n \to \infty} P_{0, it/\sqrt{n}} 1 = P_{0,0}1 = 1,
\]

and for \( k \neq 1 \)

\[
\lim_{n \to \infty} P_{k, it/\sqrt{n}} 1 = P_{k,0}1 = 0
\]

where the limit is taken in \( C^\alpha (\Sigma_T^+ \times \mathbb{P}^{d-1}) \). Furthermore

\[
\left\| S_{it/\sqrt{n}} 1(x, \overline{u}) \right\|_{C^\alpha (\Sigma_T^+ \times \mathbb{P}^{d-1})} = O(\beta^n)
\]

for some \( 0 < \beta < 1 \). Note that this \( \beta \) can be chosen uniformly for all \( S_z \) with \( |z| \) sufficiently small. Thus for \( n \) sufficiently large

\[
\int \left\| A^{[n]}(x)u \right\| \frac{it}{\sqrt{n}} d\mu(x) = \rho^n_{it/\sqrt{n}} (1 + o(1)). \tag{33}
\]

Using the Taylor expansion of \( \rho_z \) at \( z = 0 \) we have

\[
\rho^n_{it/\sqrt{n}} = \left( 1 - \frac{\rho''_0}{2} \left( \frac{t}{\sqrt{n}} \right)^2 + O \left( \frac{t^3}{n^{3/2}} \right) \right)^n;
\]

here we have used the assumption that \( \rho'_0 = \lambda_1 = 0 \). Since \( \rho''_0 = \text{Var} \), we have

\[
\int \left\| A^{[n]}(x)u \right\| \frac{it}{\sqrt{n}} d\mu(x) = \left( 1 - \frac{\rho''_0}{2} \left( \frac{t}{\sqrt{n}} \right)^2 + O \left( \frac{t^3}{n^{3/2}} \right) \right)^n (1 + o(1))
\]
If $\text{Var} > 0$, then $\exp \left[ -\frac{1}{2} \text{Var} \cdot t^2 \right]$ is the characteristic function for $\mathcal{N}(0, \text{Var})$. Hence by Levy continuity theorem,

$$\frac{\log \| A^{[n]}(x) u \|}{\sqrt{n}} \xrightarrow{n \to \infty} \mathcal{N}(0, \text{Var}).$$

If $\text{Var} = 0$, then $\exp \left[ -\frac{1}{2} \text{Var} \cdot t^2 \right] = 1$ is the characteristic function of $0$. So Levy continuity theorem implies that

$$\frac{\log \| A^{[n]}(x) u \|}{\sqrt{n}} \xrightarrow{n \to \infty} 0.$$

This completes the proof. □

5.2. Furstenberg’s formula. For i.i.d. products of matrices, Furstenberg [Fur63] gave a formula for the top Lyapunov exponent assuming strong irreducibility:

$$\lambda_1 = \int \int \log \frac{\| Au \|}{\| u \|} \ d\mu(A) d\nu(u)$$

where $\mu$ is the distribution of $A$ and $\nu$ is any stationary measure on the projective space. If, in addition, $A$ is proximal, then the stationary measure $\nu$ is unique.

As a byproduct of the previous subsection, we prove the following analogue of Furstenberg’s formula for the top Lyapunov exponent $\lambda_1$.

**Proposition 11** (Furstenberg’s Formula).

$$\lambda_1 = \rho'_0 = \int \sum_{\sigma y = x} g(y) \log \left\| A(y)^* \frac{u}{\| u \|} \right\| d\nu(x, u)$$

where $\nu$ is the eigenmeasure of $L$ from Proposition 7.

**Proof.** Write $L_z = \rho_z \left( \sum_{k=0}^{h-1} e^{\frac{2\pi i}{h} k} P_{k,z} + S_z \right)$ for $|z|$ small as in Proposition 9. Then differentiating the identity $L_z P_{0,z} = \rho_z P_{0,z}$ gives

$$L'_z P_{0,z} + L_z P'_{0,z} = \rho'_z P_{0,z} + \rho_z P_{0,z}.$$

Since $P_{0,z} \mathcal{L}_z = \rho_z P_{0,z}$ and $P^2_{0,z} = P_{0,z}$, applying $P_{0,z}$ to both sides gives

$$P_{0,z} \mathcal{L}'_z P_{0,z} + \rho_z P_{0,z} P'_{0,z} = \rho'_z P_{0,z} + \rho_z P_{0,z} P'_{0,z}.$$

Canceling $\rho_z P_{0,z} P'_{0,z}$ from both sides gives $P_{0,z} \mathcal{L}'_z P_{0,z} = \rho'_z P_{0,z}$. Evaluating this identity at $z = 0$ and applying the operators to the constant function $1$ gives

$$\int \mathcal{L}'_0 1 d\nu = P_{0,0} \mathcal{L}'_0 P_{0,0} 1 = \rho'_0 P_{0,0} 1 = \rho'_0.$$
From Lemma 8 (2), we have \((L_0')_0 = L_0 \log g M_{\psi_1} C A^{[1]}\). Thus

\[
\lambda_1 = \rho_0' = \int (L_0')_0 \, d\nu = \int L_0 \log g M_{\psi_1} C A^{[1]} \, d\nu = \int \log g \psi_1 \, d\nu = \int \sum_{\sigma y = x} g(y) \log \| A(y)^* \frac{u}{\|u\|} \| \, d\nu(x, u),
\]

completing the proof. \(\square\)

5.3. Large deviation principle. In this subsection, we prove Theorem 2 and Corollary 1. For small enough \(\varepsilon \in \mathbb{R}\), define

\[
\Lambda(\varepsilon) := \log \rho_\varepsilon - \varepsilon \lambda_1
\]

where \(\rho_\varepsilon\) is the spectral radius of the operator

\[
L_\varepsilon f(x, u) := \sum_{\sigma y = x} g(y) \left\| A(y)^* \frac{u}{\|u\|} \right\|^\varepsilon f(y, A(y)^* u).
\]

**Theorem 7** (Large Deviation Principle). Suppose \(\hat{A} : \Sigma_T \to GL_d(\mathbb{R})\) is 1-typical and \(\text{Var}\) is positive. Then there exists \(\eta > 0\) such that for any \(\varepsilon \in \left(0, \frac{\Lambda(\eta)}{\eta}\right)\) and any unit vector \(u \in \mathbb{R}^d\),

\[
\lim_{n \to \infty} \frac{1}{n} \log \mu\left\{ x : \left| \log \left\| A^{[n]}(x)u \right\| - n\lambda_1 \right| > n\varepsilon \right\} = -\Lambda^*(\varepsilon) < 0.
\]

Here \(\Lambda^*\) is the Legendre transform of \(\Lambda\) on the interval \([0, \eta]\).

**Lemma 9** (Local Gärtner–Ellis Theorem). Suppose that \(X_n\) is a sequence of random variables such that there exists \(\eta > 0\) for which

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}(e^{tX_n}) = c(t)
\]

exists in an open set of \([-\eta, \eta]\). If \(c(t)\) is continuously differentiable and strictly convex on \([-\eta, \eta]\) and \(c'(0) = 0\), then for any \(0 < \varepsilon < \frac{c'(\eta)}{\eta}\)

\[
\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(X_n > n\varepsilon) = -c^*(\varepsilon) < 0.
\]

where \(c^*\) is the Legendre transform of \(c\) on the interval \([0, \eta]\).

The proof of Lemma 9 can be found in [BL85, Chapter V Lemma 6.2].
Proof of Theorem 7. Without the loss of generality, we may assume that $\lambda_1 = 0$. Then $A(\varepsilon) = \log \rho_\varepsilon$. We will verify the assumptions of Lemma 9 for $X_n = \log \|A^{[n]}(x)u\|$.

First, $A'(0) = \rho'_0/\rho_0 = 0$; here we have used the assumption that $\rho'_0 = \lambda_1 = 0$. By reasoning as in (33), for sufficiently small $\varepsilon$ we have

$$
\lim_{n \to \infty} \frac{1}{n} \log \int \|A^{[n]}(x)u\| \, d\mu = \lim_{n \to \infty} \frac{1}{n} \log \left( \rho^n_\varepsilon \cdot (1 + o(1)) \right) = \Lambda(\varepsilon).
$$

Since $\text{Var} = \rho''_0$ from Proposition 10 and $\text{Var}$ is positive from the assumption, it follows that

$$
\frac{d^2 A}{d\varepsilon^2} \bigg|_{\varepsilon=0} = \frac{\rho''_\varepsilon \rho_\varepsilon - (\rho'_\varepsilon)^2}{\rho^2_\varepsilon} \bigg|_{\varepsilon=0} = \rho''_0 > 0.
$$

Thus there is some neighborhood of 0 on which $\Lambda(\varepsilon)$ is strictly convex.

The analogous argument also applies to $X_n = -\log \|A^{[n]}(x)u\|$, and the result then follows by applying Lemma 9. □

Proof of Corollary 1. As in Theorem 6 and 7, we will prove the result for $A^{[n]}(x)$ with respect to $\mu$. Again without loss of generality, we may assume that $\lambda_1 = 0$.

We begin by choosing a basis $\mathcal{U} = \{u_1, \ldots, u_d\}$ of $\mathbb{R}^d$ consisting of unit vectors and defining a new norm

$$
\|A\|_\mathcal{U} := \max_{1 \leq i \leq d} \|Au_i\|.
$$

By relaxing the constant $C > 0$ in the statement of the corollary if necessary, we may prove the result with respect to the new norm $\| \cdot \|_\mathcal{U}$.

Then for any $\varepsilon > 0$,

$$
\left\{ x : \|A^{[n]}(x)\|_\mathcal{U} > n\varepsilon \right\} = \bigcup_{1 \leq i \leq d} \left\{ x : \log \|A^{[n]}(x)u_i\|_\mathcal{U} > n\varepsilon \right\},
$$

and hence,

$$
\mu \left\{ x : \log \|A^{[n]}(x)\|_\mathcal{U} > n\varepsilon \right\} \leq \sum_{1 \leq i \leq d} \mu \left\{ x : \log \|A^{[n]}(x)u_i\|_\mathcal{U} > n\varepsilon \right\}.
$$

Since each term in the summation on the right hand side decreases exponentially from Theorem 7, so does the sum. Hence, the term on the left also has to decrease exponentially fast, as required. □

As mentioned in the introduction, the large deviation principle stated in Corollary 1 is already known. For instance, Gouëzel and Stoyanov [GS19] obtain this result which implies exponential returns to Pesin sets. Such property is then used to deduce exponential mixing of Gibbs measures in certain settings. We remark, however, that the large deviation principle stated in Theorem 2 is new, and we provide a simpler proof via using the spectral properties of the operator $L$.

We also note that Duarte, Klein, and Poletti [DKP21] prove the uniform version of Corollary 1. They then use it to show Hölder continuity of the Lyapunov exponents as a function of the cocycle.
5.4. Analyticity of the Lyapunov exponent $\lambda_1$. By analyticity of a function mapping some domain of $\mathbb{R}$ into a Banach space $\mathcal{B}$ we mean that the function admits an analytic extension to a domain in $\mathbb{C}$ which contains the original domain of $\mathbb{R}$.

**Theorem 8.** Suppose $\hat{A} : \Sigma_T \to GL_d(\mathbb{R})$ is $1$-typical and $t \mapsto g_t$ be a function defined on some interval $(-\varepsilon, \varepsilon)$ such that $g_t$ is a Hölder continuous $g$-function for each $t$. If the function $t \mapsto g_t$ is analytic, then $t \mapsto \lambda_1(\hat{A}, \mu_{g_t})$ is real analytic in a neighborhood of $0$.

**Proof.** Let $z \mapsto g_z$ be the analytic extension of $t \mapsto g_t$. Define

$$\mathcal{L}_{g_z} f(x, \underline{u}) := \sum_{\sigma y = x} g_z(y) f(y, \hat{A}(y)^* u).$$

By Theorem 5, for each $t \in (-\varepsilon, \varepsilon)$ the number $1$ is a simple eigenvalue for the operator $\mathcal{L}_{g_t}$. We denote by $v_t$ the eigenmeasure of $\mathcal{L}_{g_t}$ corresponding to $\rho(\mathcal{L}_{g_t})$. By analytic perturbation theory (see [Kat95] or [Klo19]), there exists an open neighborhood $U \subseteq \mathbb{C}$ of $(-\varepsilon, \varepsilon)$ and an analytic function $z \mapsto v_z$ defined on $U$ which extends $v_t$; that is, $v_z$ is an eigenmeasure of $\mathcal{L}_{g_z}$ for $z \in U$. Moreover the function $z \mapsto L_{\log g_z}$ is analytic and thus $z \mapsto L_{\log g_z}^* v_z$ is also analytic. Hence the function

$$\lambda(z) := \{L_{\log g_z}^* \psi_1, v_z\}$$

is an analytic function on $U$, and by Proposition 11 we have $\lambda(t) = \lambda_1(\hat{A}, \mu_{g_t})$ for all $t \in (-\varepsilon, \varepsilon)$. Therefore $t \mapsto \lambda_1(\hat{A}, \mu_{g_t})$ is real analytic in a neighborhood of $0$. \( \square \)

This result is new even in the case when $\mathcal{A}$ and $g$ depend only on two coordinates. Given an irreducible stochastic matrix $S$, we denote by $\mu_S$ the unique Markov measure determined by $S$. In this special case, the result says the following:

**Corollary 4.** Suppose that $T$ is irreducible and $\{A_{ij}\}_{i, j : T_{ij} = 1} \subseteq GL_d(\mathbb{R})$. If the function $\mathcal{A} : \Sigma_T^+ \to GL_d(\mathbb{R})$ defined by $\mathcal{A}(x) = A_{x_0 x_1}$ is $1$-typical, then there is a relatively open neighborhood $U$ in the hyperplane

$$\left\{(z_{ij}) \in \mathbb{C}^{d^2} : \sum_{i=1}^d z_{ij} = 1 \text{ and } z_{ij} \neq 0 \text{ if and only if } T_{ij} = 1\right\}$$

such that

$$\left\{S \in M_d(\mathbb{R}) : S \text{ is stochastic and } S_{ij} > 0 \text{ if and only if } T_{ij} = 1\right\} \subseteq U$$

and $S \mapsto \lambda_1(\mathcal{A}, \mu_S)$ extends to an analytic function on $U$.

One may compare this result to [Per92, Theorem 2] in which the matrices are assumed to be positive. Finally we use Theorem 8 to prove Theorem 3.

**Proof of Theorem 3.** By assumption, $t \mapsto \hat{\psi}_t$ is analytic. Essentially we must check that the reduction to a one-sided potential and then to a $g$-function preserves this analyticity.

As in [PP90, Proposition 1.2], there is a bounded linear map $W : C^{\alpha/2}(\Sigma_T^+) \to C^{\alpha/2}(\Sigma_T^+)$ such that $\hat{\psi}$ is cohomologous to $W \hat{\psi}$. Writing $\psi_t = W \hat{\psi}_t$, we have that $t \mapsto \psi_t$ is analytic and $\lambda_1(\hat{A}, \mu_{\hat{\psi}_t}) = \lambda_1(\mathcal{A}, \mu_{\psi_t})$. Let $h_t$ be the eigenfunction for
the transfer operator $L_{\psi_t}$ corresponding to $\rho(L_{\psi_t})$. It is known (for instance it can be deduced from [PP90, Proposition 4.6 on pg 56] or [Bal00, Theorem 1.5]) that $h_t$ and $\lambda_t := \rho(L_{\psi_t})$ are analytic functions. Thus if we define

$$g_t := \frac{e^{\psi_t(x)}}{\lambda_t} \cdot \frac{h_t(x)}{h_t(\sigma x)}$$

as in (9), then $t \mapsto g_t$ is analytic and $g_t$ is a $g$-function for each $t \in \mathbb{R}$. The result now follows from Theorem 8. □

Acknowledgement. The authors would like to thank Amie Wilkinson, Aaron Brown, Jairo Bochi, and Silvius Klein for helpful discussions comments. The authors would also like to thank anonymous referees for their comments which helped improved the paper.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

6. Appendix

6.1. The Ruelle–Perron–Frobenius theorem for semi-positive operators. Here we present the abstract Ruelle–Perron–Frobenius theorem that we make use of in this paper. We begin by recalling some definitions and propositions; for more background and proofs of these facts, we refer the reader to [DS88].

Let $X$ be a complex Banach space and $\mathcal{L}(E)$ the space of bounded linear operators on $X$. If $L \in \mathcal{L}(E)$, we define

$$\text{spec}(L) = \{ \lambda \in \mathbb{C} : \lambda - L \text{ is not invertible} \}$$

and

$$\text{res}(L) = \mathbb{C} \setminus \text{spec}(L).$$

**Theorem 9.** The function $\lambda \mapsto (\lambda - L)^{-1}$ is analytic on $\text{res}(L)$.

For $\lambda \in \text{res}(L)$, we define

$$R(\lambda, L) = (\lambda - L)^{-1}.$$

**Definition 5.** The geometric multiplicity of an eigenvalue $\lambda$ is

$$\text{geom}(\lambda) := \dim \text{ker}(\lambda - L)$$

and the algebraic multiplicity

$$\text{alg}(\lambda) := \dim \bigcup_{k=1}^{\infty} \ker(\lambda - L)^k.$$

The order of $\lambda$, provided it exists, is

$$\min \left\{ k : \ker(\lambda - L)^{k+1} = \ker(\lambda - L)^k \right\}.$$
Proposition 12. If \( \lambda_0 \) is an isolated point of \( \text{spec}(L) \) and a pole of \( R(\lambda) \), then \( \lambda_0 \) is an eigenvalue for \( L \). The order of \( \lambda_0 \) as an eigenvalue is the order of \( \lambda_0 \) as a pole of \( R(\lambda) \).

The main ideas of results are essentially contained in Sasser [Sas64]. However, we cannot use the results from those contained in [Sas64] directly. So we will use the ideas to prove a Ruelle–Perron–Frobenius theorem which applies in our case.

Definition 6. Let \( X \) be a real topological vector space a set \( C \subseteq X \) is called a cone if
1. \( C \) is convex, and
2. If \( x \in C \), then \( \lambda x \in C \) for all \( \lambda \geq 0 \).

A cone is called proper if \( -C \cap C = \{0\} \). A cone is called closed if it is a closed set in the topology of \( X \).

Throughout this section we will assume that \( X \) is a real Banach space ordered by a closed proper cone \( C \).

Definition 7. A set \( S \subseteq X^* \) is called sufficient for \( C \) if
\[
C = \{ x \in X : \langle x, s \rangle \geq 0 \text{ for all } s \in S \}.
\]

Definition 8. Let \( L : X \rightarrow X \) be linear and bounded, \( C \subseteq X \) be a closed proper cone with non-empty interior, and \( S \) be a sufficient set for \( C \). We say that \( L \) is semi-positive with respect to \( C \) if
1. \( LC \subseteq C \).
2. For all \( x \in C \) one of the following is true:
   (a) There exists \( N(x) \) such that \( L^n x \in \text{int}(C) \) for all \( n \geq N(x) \).
   (b) For any \( s \in S \), the sequence \( \{ \langle L^n x, s \rangle \}^\infty_{n=1} \) converges to 0.

The main goal of this subsection is to prove the following theorem.

Theorem 10. Let \( X \) be a real Banach space ordered by a proper closed cone \( C \), with non-empty interior and \( L : X \rightarrow X \) be a quasi-compact bounded linear operator which is semi-positive with respect to \( C \). Then
1. There exist \( u \in \text{int}(C) \) and \( u^* \in C^* \) such that \( Lu = \rho(L)u, L^* u^* = \rho(L)u^* \), and \( \langle u, u^* \rangle = 1 \).
2. If \( \lambda \in \text{spec}(L) \) with \( |\lambda| = \rho(L) \), then \( \lambda = \rho(L) \).
3. \( L \) can be written as
\[
L = \rho(L)(P + S)
\]
where \( Px = \langle x, u^* \rangle u \), \( S \) is a bounded linear operator with \( \rho(S) < 1 \) and \( PS = SP = 0 \).
4. There exist constants \( C > 0 \) and \( 0 < \gamma < 1 \) such that for any \( x \in X \)
\[
\| \rho(L)^{-n} L^n x - \langle x, u^* \rangle u \| \leq C \| x \| \gamma^n
\]
for all \( n \geq 0 \).

Lemma 10. Suppose that \( C \) is a closed proper cone and \( S \) is a sufficient set for \( C \). If \( \langle x, s \rangle = 0 \) for all \( s \in S \), then \( x = 0 \).

Proof. Notice that if \( \langle x, s \rangle = 0 \) for all \( s \), then \( x \in C \). On the other hand \( \langle -x, s \rangle = 0 \) for all \( s \in S \), so \(-x \in C \). As \( C \) is proper this implies that \( x = 0 \).
One may in general be concerned that a cone may not have a sufficient set. Note that every cone has a natural candidate for a sufficient set. Define

\[ C^* = \{ x^* \in X^* : \langle x, x^* \rangle \geq 0 \text{ for all } x \in C \} . \]

**Lemma 11.** Suppose that \( C \) is a closed proper cone with non-empty interior. Then

1. \( x \in C \) if and only if \( \langle x, x^* \rangle \geq 0 \) for all \( x^* \in C \); that is \( C^* \) is a sufficient set for \( C \).
2. We have

\[ \text{int}(C) = \{ x \in X : \langle x, x^* \rangle > 0 \text{ for all } x^* \in C^* \} . \]

**Proof.** See [Nau04, Proposition 4.11] for (1).

For (2), we will show inclusions of each set into another. First, suppose that \( x \in \text{int}(C) \).

Take \( \delta > 0 \) such that \( B(x, \delta) \subseteq C \). If \( \| y \| = \delta/2 \), then \( \| (x + y) - x \| = \| y \| < \delta \) and \( \| (x - y) - x \| = \| y \| < \delta \), and this implies that \( x - y, x + y \in C \). Thus for any \( x^* \in C^* \),

\[ 0 \leq \langle x + y, x^* \rangle = \langle x, x^* \rangle + \langle y, x^* \rangle \quad \text{and} \quad 0 \leq \langle x - y, x^* \rangle = \langle x, x^* \rangle - \langle y, x^* \rangle \]

which implies

\[ -\langle x, x^* \rangle \leq \langle y, x^* \rangle \leq \langle x, x^* \rangle . \]

Then we have

\[ \| x^* \| = 2/\delta \sup_{\| y \|=\delta/2} \| y, x^* \| \leq 2/\delta \langle x, x^* \rangle . \quad (34) \]

In particular if \( x^* \neq 0 \), then \( \langle x, x^* \rangle > 0 \).

For the other inclusion, suppose that

\[ x \in \{ x \in X : \langle x, x^* \rangle > 0 \text{ for all } x^* \in C^* \} . \]

Let \( u \in \text{int}(C) \) and take \( \delta > 0 \) such that \( B(u, \delta) \subseteq \text{int}(C) \). Setting

\[ S = \{ x^* \in C^* : \langle u, x^* \rangle = 1 \} , \]

notice that \( S \) is closed in the weak*-topology and by (34) we have that \( S \subseteq B(0, 2/\delta) \). Thus \( S \) is weak*-compact by Banach-Alaoglu, and we can take \( \eta > 0 \) such that \( \langle x, x^* \rangle > \eta \) for all \( x^* \in S \). Then for any \( y \in X \) with \( \| x - y \| < \eta \delta/4 \) and any \( x^* \in S \), we have

\[ \| (x, x^*) - \langle y, x^* \rangle \| \leq \| x - y \| \| x^* \| \leq 2 \| x - y \| /\delta \leq \eta/2 . \]

Hence \( \langle y, x^* \rangle > 0 \). As \( C \) is defined as

\[ C = \{ x \in X : \langle x, x^* \rangle \geq 0 \text{ for all } x^* \in S \} , \]

we have that \( y \in C \). This shows that \( x \in \text{int}(C) \). \( \Box \)
Denote by $\tilde{X}$ the complexification of $X$. That is,

$$\tilde{X} := \{ x + iy : x, y \in X \}.$$ 

With addition and scalar multiplication defined in the natural way, this becomes a complex vector space. The function

$$\|x + iy\|_{\tilde{X}} := \sup \{ \|x \cos \theta + y \sin \theta\| : \theta \in [0, 2\pi] \}$$

defines a norm on $\tilde{X}$ making it into a Banach space.

Any bounded linear operator $T : X \to X$ extends to a bounded linear operator $\tilde{T} : \tilde{X} \to \tilde{X}$ and $\|T\|_X = \|\tilde{T}\|_{\tilde{X}}$. For a bounded linear operator $T$, we define $\text{spec}(T) := \text{spec}(\tilde{T})$.

Finally the dual of the complexification of $X$ is isomorphic to the complexification of $X^*$ and elements $x^* \in X^*$ act on $\tilde{X}$ in the natural way by $(x + iy, x^*) := (x, x^*) + i (y, x^*)$.

Let $\tilde{L} : \tilde{X} \to \tilde{X}$ be the extension of $L$ to $\tilde{X}$. As $L$ is quasi-compact, any $\lambda_0 \in \text{spec}(L) := \text{spec}(\tilde{L})$ with $|\lambda_0| = 1$ is an isolated point of the spectrum and a pole of the resolvent. Thus $R(z) = (zI - \tilde{L})^{-1}$ has an expansion

$$R(z) = \sum_{k=1}^{n} \frac{(\lambda_0 I - \tilde{L})^{k-1} P_{\lambda_0}}{(z - \lambda_0)^k} + \sum_{k=0}^{\infty} (z - \lambda_0)^k A_k$$

which is valid for $0 < |z - \lambda_0| < \delta$ for any $\delta > 0$ with $B(\lambda_0, \delta) \cap \text{spec}(L) = \{ \lambda_0 \}$. Note that $\sum_{k=0}^{\infty} (z - \lambda_0)^k A_k$ converges in the operator norm topology, $n$ is the order of $\lambda$ (by Proposition 12, $n$ is the smallest integer such that $\ker(\lambda_0 I - \tilde{L})^n = \ker(\lambda_0 I - \tilde{L})^{n+1}$), and $P_{\lambda_0}$ is the spectral projection onto the subspace $\ker(\lambda_0 I - \tilde{L})^n$.

Next two results due to Sasser [Sas64] describes the properties of quasi-compact operator preserving a closed cone.

**Proposition 13.** Suppose that $L : X \to X$ is quasi-compact, $C$ is a proper closed cone, $LC \subseteq C$ and $\rho(L) = 1$. Then $1 \in \text{spec}(L)$. Moreover if $\lambda_0 \in \text{spec}(L)$ and $|\lambda_0| = 1$, then the order of $\lambda_0$ is at most the order of 1.

**Proof.** This is [Sas64, Theorem 2]. \qed

**Lemma 12.** Suppose that $L : X \to X$ is quasi-compact, $C$ is a proper closed cone, $LC \subseteq C$, and $\rho(L) = 1$. Then there exist $u \in C$ ($u \neq 0$) and $u^* \in C^*$ ($u^* \neq 0$) such that $Lu = u$ and $L^* u^* = u^*$.

**Proof.** This is [Sas64, Theorem 3]. \qed

**Lemma 13.** If $z \in \text{int}(C)$ and $x \in X$ with $x \neq 0$, then there exists $t \in \mathbb{R} \setminus \{0\}$ such that $z + tx \in \partial C$.

**Proof.** By replacing $x$ with $-x$ if necessary, we may assume that $x \not\in C$. Set

$$t_0 = \sup \{ t : z + tx \in C \}$$

and notice that as $z \in \text{int}(C)$ we have that $t_0 > 0$. As $x \not\in C$, we may choose $x^* \in C^*$ such that $(x, x^*) < 0$ by Lemma 11 (1). Notice that if $t > - (z, x^*)/(x, x^*)$, then $z + tx \not\in C$. Thus $t_0 < \infty$. As $C$ is closed we have that $z + t_0 x \in \bar{C}$, but it is clear that $z + t_0 x \not\in \text{int}(C)$. Hence, $z + t_0 x \in \partial C$. \qed
The following proposition shows that with the added assumption of semi-positivity with respect to a proper closed cone, extra properties (such as some of the listed properties in Theorem 10) of the operator can be established.

**Proposition 14.** Suppose that \( L : X \to X \) is quasi-compact, \( C \) is a proper closed cone, \( L \) is semi-positive with respect to \( C \), and \( \rho(L) = 1 \). Then the following are true:

1. If \( w \in C \setminus \{0\} \) with \( Lw = w \), then \( w \in \text{int}(C) \).
2. The order of the eigenvalue 1 is one.
3. \( \tilde{L} \) can be written as
   \[
   \tilde{L} = \sum_{j=1}^{m} \lambda_j P_j + S
   \]
   where \( \{ \lambda_j \} \) is the collection of eigenvalues of \( L \) with modulus 1, \( P_j \) is a projection for \( j \), \( P_j S = SP_j = 0 \) for \( i \neq j \), and \( \rho(S) < 1 \).
4. The dimension of the eigenspace corresponding to 1 is one.

**Proof.** We can modify the proof of [Sas64, Theorem 4] to fit our assumptions.

1. It is clear that there is some \( s \in S \) such that \( \{ \langle L^n w, s \rangle \}_{n=1}^{\infty} \) does not converge to 0.
   So by the assumption that \( L \) is semi-positive, it must be that \( L^n w \in \text{int}(C) \) for some \( n \), but of course \( L^n w = w \). Hence, \( w \in \text{int}(C) \).
2. Let \( n \) be the order of 1 and set
   \[
   \Gamma = (I - \tilde{L})^{n-1} P_1.
   \]
From the Laurent expansion of \( R(z) \) about 1 we can see that
   \[
   \Gamma = \lim_{z \to 1} (z - 1)^n R(z)
   = \lim_{z \to 1^+} (z - 1)^n R(z)
   = \lim_{z \to 1^+} (z - 1)^n \sum_{k=0}^{\infty} z^{-k-1} \tilde{L}^k.
   \]
From this it follows that \( \Gamma X \subseteq X \) and \( \Gamma C \subseteq C \). Moreover
   \[
   (I - \tilde{L})\Gamma = (I - \tilde{L})^n P_1 = 0.
   \]
This combined with the previous equation implies that \( \Gamma \tilde{L} = \tilde{L} \Gamma = \Gamma \).
Take \( x \in C \) such that \( \Gamma x \neq 0 \). Notice that \( L\Gamma x = \Gamma x \) and \( \Gamma x \in C \), and thus \( \Gamma x \in \text{int}(C) \) by part (1). With \( u^* \in C^* \) from Lemma 12, we have
   \[
   0 < \langle \Gamma x, u^* \rangle = \lim_{z \to 1^+} (z - 1)^n \sum_{k=0}^{\infty} z^{-k-1} \langle \tilde{L}^k x, u^* \rangle
   = \langle x, u^* \rangle \lim_{z \to 1^+} (z - 1)^n \sum_{k=0}^{\infty} z^{-k-1}
   = \langle x, u^* \rangle \lim_{z \to 1^+} (z - 1)^n \frac{z^{-1}}{1 - z^{-1}}
   = \langle x, u^* \rangle \lim_{z \to 1^+} (z - 1)^{n-1}.
   \]
Therefore \( n = 1 \).
3. We enumerate the points in \( \text{spec}(L) \) with modulus 1 as \( \{ \lambda_i \}_{i=1}^{m} \) where \( \lambda_1 = 1 \). Letting \( P_i = P_{\lambda_i} \) be the spectral projection onto \( \ker(\lambda_i I - \widetilde{L}) \), notice from Proposition 13 that the order of \( \lambda_i \) is one. Setting \( \mathcal{P} = \sum_i P_i \) and writing

\[
\widetilde{L} = \sum_i P_i \sum_i (I - \mathcal{P}),
\]

it is easily seen that \( \widetilde{L}P_i = \lambda_i P_i \).

We then take \( S = \widetilde{L}(I - \mathcal{P}) \). As \( I - \mathcal{P} \) is the spectral projection onto \( \text{spec}(L) \) we have that \( \text{spec}(\widetilde{L}|_{\text{Ran}(I - \mathcal{P})}) = \text{spec}(L) \setminus \{ \lambda_i \}_{i=1}^{k} \). In particular, \( \rho(S) < 1 \).

4. Let \( u \in \mathcal{C} \) be as in Lemma 12. Supposing that \( Lx = x \) for some \( x \in X \), we take \( t \) as in Lemma 13 such that \( u + tx \in \partial \mathcal{C} \). Since \( L(u + tx) = u + tx \), it follows from part (1) that \( u + tx = 0 \); that is \( x \) is a scalar multiple of \( u \).

Now if \( x + iy \in \widetilde{X} \) is such that \( L(x + iy) = x + iy \), then \( Lx = x \) and \( Ly = y \). Hence, \( x + iy \) is a scalar multiple of \( u \), as required. \( \square \)

**Lemma 14.** Suppose that \( L : X \to X \) is quasi-compact, \( \mathcal{C} \) is a proper closed cone, \( L \) is semi-positive with respect to \( \mathcal{C} \), and \( \rho(L) = 1 \). Then for any \( z \in \text{int}(\mathcal{C}) \) we have that \( \{L^n z\} \subseteq \text{int}(\mathcal{C}) \).

**Proof.** Let \( u \in \mathcal{C} \) be as in Lemma 12. Notice that for any \( x^* \in \mathcal{C}^* \), we have that

\[
\langle u, x^* \rangle = \langle L^n u, x^* \rangle \leq \|u\| \|(L^*)^n x^*\|.
\]

Suppose \( z \in \text{int}(\mathcal{C}) \) and take \( \delta > 0 \) such that \( B(z, \delta) \subseteq \mathcal{C} \). Then

\[
\langle L^n z, x^* \rangle = \left( z, \frac{(L^*)^n x^*}{\|L^* x^*\|} \right) \left( (L^*)^n x^* \right) \geq \frac{\delta}{2} \|L^* x^*\| \geq \frac{\delta}{2} \frac{\langle u, x^* \rangle}{\|u\|}
\]

where the first inequality is due to (34). In particular, if \( y \in \{L^n z\} \), then \( \langle y, x^* \rangle > 0 \). Thus by Lemma 11, \( y \in \text{int}(\mathcal{C}) \). \( \square \)

**Proposition 15.** Suppose that \( L : X \to X \) is quasi-compact, \( \mathcal{C} \) is a proper closed cone, \( L \) is semi-positive with respect to \( \mathcal{C} \), and \( \rho(L) = 1 \). Then 1 is the only element of \( \text{spec}(L) \) of modulus 1.

**Proof.** Let \( u \) be as in Lemma 12. Suppose that \( e^{i\theta} \) belongs to \( \text{spec}(L) \) and let \( x + iy \neq 0 \) be its corresponding eigenvector; that is, \( L(x + iy) = e^{i\theta}(x + iy) \). Take a sequence \( n_k \) such that \( e^{in_k \theta} \xrightarrow{k \to \infty} 1 \), then

\[
L^{n_k}(x + iy) \xrightarrow{k \to \infty} x + iy.
\]

Thus

\[
L^{n_k}x \xrightarrow{k \to \infty} x \text{ and } L^{n_k}y \xrightarrow{k \to \infty} y.
\]

If \( x \neq 0 \), then by Lemma 13 we can take \( t \neq 0 \) such that \( u + tx \in \partial \mathcal{C} \). Notice that

\[
L^{n_k}(u + tx) = u + tL^{n_k}x \xrightarrow{k \to \infty} u + tx.
\]

By the assumption that \( L \) is semi-positive with respect to \( \mathcal{C} \), either \( u + tx = 0 \) or there exists \( N \) such that \( L^n(u + tx) \in \text{int}(\mathcal{C}) \) for all \( n \geq N \).
It turns out that we must have $u + tx = 0$. This is because if $u + tx \neq 0$, then we would have $u + tx \in \{ L_n(L^n(u + tx)) \}_{n=1}^{\infty} \subseteq \text{int}(C)$ by Lemma 14. Then $u + tx \in \partial C$, and at the same time, $u + tx \in \text{int}(C)$, which is absurd. The same argument shows that if $y \neq 0$, then $y$ is a scalar multiple of $u$. Therefore, $x + iy$ is a scalar multiple of $u$ and $e^{id} = 1$. □

Proof of Theorem 10. 1. This follows by applying Lemma 12 as well as Proposition 14 (1) and (2) to $\rho(L)^{-1}L$.

2. This follows by an application of Proposition 15 to $\rho(L)^{-1}L$.

3. This follows from Proposition 14 (4) and part (2) of this theorem.

4. Notice that by part (3) we have that $\rho(L)^{-n}L^n = P + S^n$. Thus

$$\| \rho(L)^{-n}L^n x - (x, u^*)u \| = \| Px + S^n x - (x, u^*)u \| = \| S^n x \| .$$

Take $\rho(S) < \gamma < 1$, then

$$\| S^n x \| \leq \frac{\| S^n \|}{\gamma^n} \| x \| \gamma^n .$$

As $\| S^n \|$ has an exponential growth rate of $\log \rho(S)$ by Gelfand’s formula, we have that $\frac{\| S^n \|}{\gamma^n}$ tends to 0, and in particular, is bounded by some constant $C > 0$. □

6.2. Direct sums of Banach spaces and spectral properties of non-negative matrices.

Proposition 16. 1. Let $\{X_i\}_{i=1}^{n}$ be a finite collection of Banach spaces. Define the vector space

$$\bigoplus_{i=1}^{n} X_i = \{(x_1, x_2, \ldots, x_n) : x_i \in X_i\}$$

with addition and scalar multiplication being coordinate wise. Then the function

$$\| (x_i)_{i=1}^{n} \| = \max_{i} \| x_i \|_{X_i}$$

defines a norm on $\bigoplus_{i=1}^{n} X_i$ which makes it a Banach space.

2. Suppose that $\{X_i\}_{i=1}^{n}$ is a collection of Banach spaces and that $\{A_{ij}\}_{i=1,j=1}^{n,n}$ is a collection of bounded linear operators $A_{ij} : X_j \to X_i$. The function

$$A(x_i)_{i=1}^{n} := \left( \sum_{j=1}^{n} A_{ij} x_j \right)_{i=1}^{n}$$

defines a bounded linear operator $A : \bigoplus_{i=1}^{n} X_i \to \bigoplus_{i=1}^{n} X_i$. We will denote this operator by $[A_{ij}]$.

3. Suppose that $A : \bigoplus_{i=1}^{n} X_i \to \bigoplus_{i=1}^{n} X_i$ is a bounded linear operator. Then there exists a collection of bounded linear operators $A_{ij} : X_j \to X_i$ such that $A = [A_{ij}]$.

4. Let $A = [A_{ij}]$ and $B = [B_{ij}]$, then $AB = [\sum_k A_{ik} B_{kj}]$.

Proof. 1. This can be found in [Con90, Chapter 3 proposition 4.4].
2. That the function $\mathcal{A}$ is linear can be verified by computation. To see that $\mathcal{A}$ is bounded, we note that
\[
\left\| \sum_{j=1}^{n} A_{ij} x_j \right\|_{X_i} \leq \sum_{j=1}^{n} \| A_{ij} \|_{\text{op}} \| x_j \|_{X_j} \leq \| (x_i) \| \sum_{j=1}^{n} \| A_{ij} \|_{\text{op}}.
\]
Therefore $\| \mathcal{A} \|_{\text{op}} \leq \max_{i} \sum_{j=1}^{n} \| A_{ij} \|_{\text{op}}.$

3. For each $1 \leq i \leq n$, define $P_i : \bigoplus_k X_k \to X_i$ by
\[
P_i(x_1, x_2, \ldots, x_n) = x_i
\]
and $\overline{P}_i : X_i \to \bigoplus_k X_k$ by
\[
\overline{P}_i x = (0, \ldots, 0, x_{i\text{th place}}, 0, \ldots, 0).
\]
Setting $A_{ij} = P_i \mathcal{A} \overline{P}_j$, it can be verified that $\mathcal{A} = [A_{ij}].$

4. Notice that
\[
\mathcal{A} \mathcal{B}(x_i)_{i=1}^{n} = \mathcal{A} \left( \sum_{j=1}^{n} B_{ij} x_j \right)_{i=1}^{n} = \left( \sum_{k=1}^{n} A_{ik} \sum_{j=1}^{n} B_{kj} x_j \right)_{i=1}^{n} = \left( \sum_{j=1}^{n} \sum_{k=1}^{n} A_{ik} B_{kj} x_j \right)_{i=1}^{n}.
\]
Therefore $\mathcal{A} \mathcal{B} = [\sum_k A_{ik} B_{kj}]$. $\square$

When working with an operator $\mathcal{A}$ acting on a direct sum, we will sometimes write
\[
\mathcal{A} = \begin{bmatrix}
A_{11} & \cdots & A_{1n} \\
\vdots & \ddots & \vdots \\
A_{n1} & \cdots & A_{nn}
\end{bmatrix}
\]
to mean that $\mathcal{A} = [A_{ij}]$. By Proposition 16, the operation of multiplication of operators is compatible with the typical formulas for matrix multiplication, justifying such notation.

**Lemma 15.** Let $\{X_i\}_{i=1}^{n}$ be a collection of Banach space and $\mathcal{A} : \bigoplus_i X_i \to \bigoplus_i X_i$. Suppose that $\mathcal{A}$ is block upper triangular
\[
\mathcal{A} = \begin{bmatrix}
A_{11} & A_{12} & \cdots & A_{1n} \\
0 & A_{22} & \cdots & \vdots \\
0 & 0 & \ddots & \vdots \\
0 & \cdots & 0 & A_{nn}
\end{bmatrix}
\]
1. A is invertible if and only if \( A_{ii} \) is invertible for all \( 1 \leq i \leq n \).

2. We have

\[
\text{spec}(A) = \bigcup_{i=1}^{n} \text{spec}(A_{ii}).
\]

**Proof.** 1. \((\Rightarrow)\) This direction is clear; use the invertibility of \( A \) and apply it to vectors of the form \((0, 0, \ldots, x_i, \ldots, 0)\) to define the inverse for \( A_{ii} \).

\((\Leftarrow)\) The proof is by induction on \( n \) (the number of Banach spaces). It is clear that the result holds for \( n = 1 \). Suppose that the result holds for \( n - 1 \), we can think of \( \bigoplus_{i=1}^{n} X_i \) as \( \bigoplus_{i=1}^{n-1} X_i \oplus X_n \) and write \( A \) as a matrix with respect to this direct sum

\[
A = \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & A_{nn}
\end{bmatrix}.
\]

By assumption \( A_{nn} \) is invertible, and by the induction hypothesis \( \tilde{A}_{11} \) is invertible. Now one can check using Proposition 16 (4) that

\[
\begin{bmatrix}
\tilde{A}_{11}^{-1} - \tilde{A}_{11}^{-1} \tilde{A}_{12} A_{nn}^{-1} \\
0
\end{bmatrix}
\]

defines an inverse for \( A \).

2. This follows from part (1). \(\Box\)

**Theorem 11** (Perron–Frobenius). Suppose that \( T \) is an irreducible non-negative matrix. There exist a number \( h \geq 1 \), called the period of \( A \), and a permutation matrix \( P \) such that \( P A P^* \) can be written as a block matrix of the following form

\[
PTP^* = \begin{bmatrix}
0 & T_{12} & 0 & \cdots & 0 \\
\vdots & 0 & T_{23} & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & T_{h-1h} & 0 \\
T_{h1} & 0 & \cdots & 0
\end{bmatrix}
\]

where the diagonal blocks are square. Moreover,

\[
PT^h P^* = \begin{bmatrix}
T_{12} T_{23} T_{34} \cdots T_{h1} & 0 & 0 & \cdots & 0 \\
0 & T_{23} T_{34} \cdots T_{h1} T_{12} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & T_{h1} T_{12} T_{23} \cdots T_{h-1h}
\end{bmatrix}
\]

and each diagonal block is primitive.

**Proof.** For the proof we refer the reader to [HJ13, Section 8.4] and the references therein. \(\Box\)
Proposition 17. Suppose that

\[
A = \begin{bmatrix}
0 & A_{12} & 0 & \cdots & 0 \\
\vdots & 0 & A_{23} & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & A_{h-1, h} \\
A_{h1} & 0 & \cdots & 0
\end{bmatrix}
\]

That is \( A = [A_{ij}] \) where \( A_{ij} = 0 \) unless \( j \equiv i+1 \mod h \). Then for each \( 0 \leq k \leq h-1 \), there exists an isometric isomorphism \( D : \bigoplus_{i=1}^{h} X_i \to \bigoplus_{i=1}^{h} X_i \) such that \( D^{-1} AD = e^{\frac{2\pi i}{h} k} A \).

Proof. Define

\[
D = \begin{bmatrix}
e^{-\frac{2\pi i}{h}} & 0 & \cdots & 0 \\
e^{-\frac{2\pi i}{h}} & e^{\frac{2\pi i}{h}} & 0 & \cdots \\\ne^{-\frac{2\pi i}{h}} & 0 & e^{\frac{2\pi i}{h}} & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
e^{-\frac{2\pi i}{h}} & 0 & \cdots & e^{\frac{2\pi i}{h} (h-1)} \\
e^{-\frac{2\pi i}{h}} & 0 & \cdots & 0 \\
A_{h1} & e^{-\frac{2\pi i}{h}} & \cdots & 0
\end{bmatrix}
\]

Then

\[
D^{-1} AD = \begin{bmatrix}
0 & e^{-\frac{2\pi i}{h}} A_{12} e^{\frac{2\pi i}{h}} & 0 & \cdots & 0 \\
\vdots & 0 & e^{-\frac{2\pi i}{h}} A_{23} e^{\frac{2\pi i}{h}} & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & e^{-\frac{2\pi i}{h} (h-1)} A_{h-1, h} \\
A_{h1} e^{-\frac{2\pi i}{h}} & 0 & \cdots & 0
\end{bmatrix}
\]

which is equal to \( e^{\frac{2\pi i}{h} k} A \). \( \square \)

References

[AV07] Avila, A., Viana, M.: Simplicity of lyapunov spectra: proof of the Zorich–Kontsevich conjecture. Acta Mathematica 198(1), 1–56 (2007)

[Bal00] Baladi, V.: Positive Transfer Operators and Decay of Correlations. Advanced Series in Nonlinear Dynamics, vol. 16. World Scientific Publishing Co., Inc., River Edge (2000)

[BBB18] Backes, L., Brown, A., Butler, C.: Continuity of Lyapunov exponents for cocycles with invariant holonomies. J. Mod. Dyn. 12(1), 223–260 (2018)

[BGMV03] Bonatti, C., Gómez-Mont, X., Viana, M.: Généricité d’exposants de lyapunov non-nuls pour des produits déterministes de matrices. Annales de l’IHP Analyse non linéaire 20, 579–624 (2003)
[BL85] Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrödinger Operators, Progress in Probability and Statistics, vol. 8. Birkhäuser Boston Inc, Boston (1985)

[BNV17] Bocker-Neto, C., Viana, M.: Continuity of Lyapunov exponents for random two-dimensional matrices. Ergod. Theory Dyn. Syst. 37(5), 1413–1442 (2017)

[Bou88] Bougerol, P.: Théorèmes limite pour les systèmes linéaires à coefficients markoviens. Probab. Theory Relat. Fields 78(2), 193–221 (1988)

[Bow75] Bowen, R.: Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Lecture Notes in Mathematics, vol. 470. Springer-Verlag (1975)

[BQ16] Benoist, Y., Quint, J.-F.: Random Walks on Reductive Groups, Random Walks on Reductive Groups, pp. 153–167. Springer (2016)

[BV04] Bonatti, C., Viana, M.: Lyapunov exponents with multiplicity 1 for deterministic products of matrices. Ergod. Theory Dyn. Syst. 24(5), 1295–1330 (2004)

[BV05] Bochi, J., Viana, M.: The Lyapunov exponents of generic volume-preserving and symplectic maps. Ann. Math. 161, 1423–1485 (2005)

[Con90] Conway, J.B.: A Course in Functional Analysis, Graduate Texts in Mathematics, vol. 96, 2nd edn. Springer-Verlag, New York (1990)

[DK16] Duarte, P., Klein, S.: Lyapunov exponents of linear cocycles, continuity via large deviations. Atl. Stud. Dyn. Syst. 3 (2016)

[DKP21] Duarte, P., Klein, S., Poletti, M.: Holder continuity of the Lyapunov exponents of linear cocycles over hyperbolic maps. arXiv preprint arXiv:2110.10265 (2021)

[DS88] Dunford, N., Schwartz, J.T.: Linear Operators, Part I. Wiley Classics Library, Wiley, New York (1988)

[Fur63] Furstenberg, H.: Noncommuting random products. Trans. Am. Math. Soc. 108(3), 377–428 (1963)

[GLP04] Guivarc’h, Y., Le Page, É.: 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, pp. 181–259. Walter de Gruyter, Berlin (2004)

[GS19] Gouëzel, S., Stoyanov, L.: Quantitative Pesin theory for Anosov diffeomorphisms and flows. Ergod. Theory Dyn. Syst. 39(1), 159–200 (2019)

[Gui15] Guivarc’h, Y.: Spectral gap properties and limit theorems for some random walks and dynamical systems, hyperbolic dynamics, fluctuations and large deviations. Proc. Sympos. Pure Math. 89, 279–310 (2015)

[Henn93] Hennion, H.: Sur un théorème spectral et son application aux noyaux lipchitziens. Proc. Am. Math. Soc. 118(2), 627–634 (1993)

[HJ13] Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd edn. Cambridge University Press, Cambridge (2013)

[Kat95] Kato, T.: Perturbation Theory for Linear Operators, Classics in Mathematics. Springer-Verlag, Berlin (1995)

[Kin73] Kingman, J.F.C.: Subadditive ergodic theory. Ann. Probab. 1, 883–909 (1973)

[Klo19] Kloeckner, B.R.: Effective perturbation theory for simple isolated eigenvalues of linear operators. J. Oper. Theory 81(1), 175–194 (2019)

[KS13] Kalinin, B., Sadovskaya, V.: Cocycles with one exponent over partially hyperbolic systems. Geometriea Dedicata 167(1), 167–188 (2013)

[Lep00] Leplaideur, R.: Local product structure for equilibrium states. Trans. Am. Math. Soc. 352(4), 1889–1912 (2000)

[LP82] Le Page, É.: Théorèmes limites pour les produits de matrices aléatoires, Probability measures on groups (Oberwolfach, 1981) Lecture Notes in Mathematics, vol. 928, pp. 258–303. Springer, Berlin, New York (1982)

[LP89] Le Page, É.: Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. Annales de l’IHP Probabilités et statistiques 25, 109–142 (1989)

[Nau04] Naud, F.: Birkhoff cones, symbolic dynamics and spectrum of transfer operators. Discrete Contin. Dyn. Syst. 11(2–3), 581–598 (2004)

[Ose68] Oseledets, V.I.: A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. Trudy Moskovskogo Matematicheskogo Obshchestva 19, 179–210 (1968)

[Par20] Park, K.: Quasi-multiplicativity of typical cocycles. Commun. Math. Phys. 376(3), 1957–2004 (2020)

[Per92] Peres, Y.: Domains of analytic continuation for the top Lyapunov exponent. Ann. Inst. H. Poincaré Probab. Stat. 28(1), 131–148 (1992)
Piraino, M.: The weak Bernoulli property for matrix Gibbs states. Ergod. Theory Dyn. Syst. 40, 1–20 (2018)

Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. Astérisque 187(188), 1-268 (1990)

Ratner, M.: The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. Israel J. Math. 16(2), 181–197 (1973)

Sasser, D.W.: Quasi-positive operators. Pac. J. Math. 14, 1029–1037 (1964)

Sert, C.: Large deviation principle for random matrix products. Ann. Probab. 47(3), 1335–1377 (2019)

Viana, M., Yang, J.: Continuity of Lyapunov exponents in the c0 topology. Israel J. Math. 229(1), 461–485 (2019)

Young, L.-S.: Large deviations in dynamical systems. Trans. Am. Math. Soc. 318(2), 525–543 (1990)

Communicated by C. Liverani