TOTALLY ERGODIC GENERALISED MATRIX EQUILIBRIUM STATES HAVE THE BERNOULLI PROPERTY

IAN D. MORRIS

ABSTRACT. We show that every totally ergodic generalised matrix equilibrium state is ψ-mixing with respect to the natural partition into cylinders and hence is measurably isomorphic to a Bernoulli shift in its natural extension. This implies that the natural extensions of ergodic generalised matrix equilibrium states are measurably isomorphic to Bernoulli processes extended by finite rotations. This resolves a question of Gatzouras and Peres in the special case of self-affine repelling sets with generic translations.

MSC2020 codes: 28A80, 37D35 (primary); 37D20, 37C45 (secondary)

1. BACKGROUND AND MOTIVATION

Given a dynamical system $f\colon M \to M$ defined on a manifold M it is a matter of fundamental interest to be able to describe the behaviour of typical trajectories. Here "typical" is often understood in measure-theoretic terms, leading us to ask what happens to trajectories whose starting point belongs to a set of full, or at least positive, Lebesgue measure. There are however situations in which this is insufficient: for example, a dynamical system may admit a repelling invariant set such that Lebesgue almost every point in an open neighbourhood of the invariant set eventually leaves that open set never to return; but it may still be of interest to understand which behaviours are typical among those points whose trajectories remain on the repelling set at all future times. Since the repelling set itself will usually have zero Lebesgue measure, a natural way to understand this question is to look for invariant measures supported on the repelling set with the largest possible dimension, which in this article will always be taken to mean Hausdorff dimension. This raises obvious fundamental questions: do such measures exist, are they unique, and what are their ergodic properties? Questions and conjectures in this direction have been raised on numerous occasions by various authors (for example [17, 37, 38, 51]). The following conjecture of D. Gatzouras and Y. Peres is representative of this general type of question:

Conjecture 1 ([30]). *Let $f\colon M \to M$ be an expanding map and $K \subseteq M$ a compact invariant set which satisfies specification. Then K supports a unique ergodic f-invariant measure with the same Hausdorff dimension as K. This measure is mixing for f and, perhaps, its natural extension is measurably isomorphic to a Bernoulli shift.*

Here an invariant set is said to satisfy *specification* if it admits a Markov partition which satisfies a certain quantitative topological mixing property; for details we refer the reader to [30]. This property is in particular satisfied if the dynamical system $f\colon K \to K$ is topologically conjugate to the full shift on finitely many symbols, which will be the case for all of the examples considered in this article.

Remarkably, the conjecture of Gatzouras and Peres has been answered negatively in every particular. It has been shown that the measure of maximal dimension can fail to exist, and even that the supremum of the dimensions of invariant measures can fall short of the dimension of K itself (see [18]); it has been shown that the measure of maximal dimension can exist but fail to be unique (see [6, 46]); and it has been shown that the measure of maximal dimension can exist but fail to be totally ergodic (and in particular fail to be mixing), a result which is implied by previous work of the author [42, §2].

The fundamental difficulty of Conjecture 1 is as follows. If $f\colon M \to M$ is an expanding map, its absolutely continuous invariant measures can be characterised as the invariant measures μ which maximise the quantity $h(\mu) - \int \log|\det D_x f|\, d\mu(x)$, where $h(\mu)$ denotes the entropy of μ with respect to f; we call such measures *equilibrium states* of the function $x \mapsto -\log|\det D_x f|$, which we call a *potential*. This definition can be alternatively presented by saying that the absolutely continuous invariant measure μ maximises the entropy minus the total of the $d := \dim M$ different Lyapunov exponents of f with respect to μ. When the dimension of the set K in Conjecture 1 is instead equal to $s \in (0, d)$, the measure of maximal dimension is believed to be typically characterised by the property of maximising the entropy minus the sum of the $\lfloor s \rfloor$ least expanding Lyapunov exponents, minus $(s - \lfloor s \rfloor)$ times the next least expanding Lyapunov exponent. If all of the Lyapunov exponents are equal then this sum of weighted Lyapunov exponents is simply (s/d) times the logarithm of the Jacobian, and the potential can then be realised as a continuous real-valued function. This makes the classical thermodynamic formalism of Bowen, Ruelle and Sinai, which applies to Hölder continuous real-valued potentials, applicable to the problem. For this reason Conjecture 1 has for some time been substantially understood in the special case of repelling sets of *conformal* expanding maps in which all Lyapunov exponents of a given invariant measure are guaranteed to be equal. Outside this special case the problem becomes more difficult in that we are obliged to consider equilibrium states of a potential which is defined in terms of several distinct Lyapunov exponents and cannot be reduced to the classical thermodynamic formalism of continuous potentials such as $x \mapsto -\log|\det D_x f|$. The investigation of the candidate measures of maximal dimension in this case thus requires the development of a subadditive thermodynamic formalism capable of dealing with Lyapunov exponents in place of the ergodic average of a function. This project has seen substantial progress in the last few years (see e.g. [10, 24, 26, 35, 45, 47, 50]) and this article is concerned with the description in detail of the equilibrium states yielded by this thermodynamic formalism in the case of locally constant cocycles over the full shift.

To address the complete generality of Conjecture 1 would require a theory of equilibrium states which allowed the consideration of arbitrary differentiable (or perhaps just Hölder continuous) linear cocycles defined over repelling sets. Such a theory appears to be significantly beyond the range of current techniques, and so far the development of this thermodynamic formalism has focused principally on the simplest nontrivial context, namely the equilibrium states of locally constant linear cocycles over full symbolic shifts. This is precisely the thermodynamic formalism needed to understand the (candidate) invariant measures of maximum dimension for *self-affine sets*, a class of fractal objects of independent interest which (under certain assumptions) correspond to the case of Conjecture 1 in which $M = \mathbb{R}^d$

and in which $D_x f \in \mathrm{GL}_d(\mathbb{R})$ takes only finitely many values when x belongs to the invariant set K. In the present work we completely describe the qualitative mixing properties of equilibrium states of linear cocycles of this type: we will show that every ergodic *generalised matrix equilibrium state* has the property that its natural extension is measurably isomorphic to the product of a Bernoulli process and a rotation of a finite set. In particular the natural extension of every totally ergodic generalised matrix equilibrium state is measurably isomorphic to a Bernoulli process. This completely resolves that part of Conjecture 1 which is concerned with mixing and the Bernoulli property in the special case where K is a self-affine set which is already known to support an invariant measure whose dimension is equal to a theoretical maximum value defined by Falconer in [20]. This property is known to hold for self-affine sets which are "typical" in certain precise senses (see [34, Theorem 4] and [23, Theorem 1.9]).

This motivates us to ask the following question, which is essentially a modification of Conjecture 1 in view of subsequent developments:

Question 1. *Let $f\colon M \to M$ be a C^2 expanding map and $K \subseteq M$ a compact invariant set which satisfies specification and supports a unique ergodic f-invariant measure with the same Hausdorff dimension as K. Is the natural extension of this measure measurably isomorphic to the product of a Bernoulli measure with a rotation on a finite set?*

For self-affine repelling sets which support a measure of dimension equal to the theoretical maximum defined by Falconer, the results in this article suffice to answer Question 1 affirmatively. However, the full range of possible behaviours outside this class of repellers is far from being completely understood even in the self-affine case. Beyond the self-affine class we anticipate that it might not be difficult to extend our methods and results to the case of typical repellers which satisfy a *fibre-bunching condition* on the derivative cocycle $(x, n) \mapsto D_x f^n$ as in [15, 21, 47], particularly if a strong additional assumption is used such as the "pinching and twisting" conditions introduced by Bonatti and Viana in [11]. The removal of the fibre-bunching condition may be a more substantial obstacle to further development of these ideas.

2. Generalised matrix equilibrium states

2.1. **Fundamental definitions and notation.** The class of measures which we investigate in this article, which we call generalised matrix equilibrium states, are defined on abstract symbolic spaces and can be related to self-affine sets via a coding procedure which is described later in this section. In order to describe these objects we require some fundamental definitions. For each $N \geq 2$ let us define $\Sigma_N := \{1, \ldots, N\}^{\mathbb{N}}$ and equip this set with the infinite product topology with respect to which it is compact and metrisable. We define the *shift transformation* $\sigma\colon \Sigma_N \to \Sigma_N$ by $\sigma[(x_k)_{k=1}^{\infty}] := (x_{k+1})_{k=1}^{\infty}$ and we denote the set of all σ-invariant Borel probability measures on Σ_N by $\mathcal{M}_\sigma(\Sigma_N)$. For convenience we will refer to such measures simply as *shift-invariant measures* on Σ_N. We equip $\mathcal{M}_\sigma(\Sigma_N)$ with the weak-* topology, which is compact and metrisable and has the property that $\mu \mapsto \int f \, d\mu$ defines a continuous function $\mathcal{M}_\sigma(\Sigma_N) \to \mathbb{R}$ for every $f \in C(\Sigma_N)$. We likewise define $\hat{\Sigma}_N := \{1, \ldots, N\}^{\mathbb{Z}}$ with the infinite product topology, $\hat{\sigma}\colon \hat{\Sigma}_N \to \hat{\Sigma}_N$ by $\hat{\sigma}[(x_k)_{k\in\mathbb{Z}}] := (x_{k+1})_{\mathbb{Z}}$, and let $\mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)$ denote the set of all $\hat{\sigma}$-invariant

measures on $\hat{\Sigma}_N$ equipped with its weak-* topology with respect to which it is compact and metrisable.

If $\mathtt{i} = (i_k)_{k=1}^n \in \{1, \ldots, N\}^n$ is a finite sequence over the symbols $1, \ldots, N$ then we refer to $\mathtt{i}$ as a *word* over $\{1, \ldots, N\}$; we call n the *length* of the word $\mathtt{i}$ and denote it by $|\mathtt{i}|$. If $\mathtt{i} = (i_k)_{k=1}^n$ and $\mathtt{j} = (j_k)_{k=1}^m$ are words then we let $\mathtt{ij}$ denote the word of length $n+m$ whose first n symbols are $i_1, \ldots, i_n$ and whose next m symbols are $j_1, \ldots, j_m$, and call $\mathtt{ij}$ the *concatenation* of $\mathtt{i}$ with $\mathtt{j}$. If $\mathtt{i}$ is a word then for each $n \geq 1$ we let $\mathtt{i}^n$ denote the concatenation of n successive copies of $\mathtt{i}$ and call this word the nth *power of* $\mathtt{i}$. We denote the set of all words over $\{1, \ldots, N\}$ by Σ_N^* and observe that the map $(\mathtt{i}, \mathtt{j}) \mapsto \mathtt{ij}$ defines a semigroup operation on Σ_N^*. If $x = (x_k)_{k=1}^\infty \in \Sigma_N$ and $n \geq 1$ are given, we let $x|_n$ denote the word $(x_k)_{k=1}^n$; if $\mathtt{i} \in \Sigma_N^*$ is given, we let $[\mathtt{i}]$ denote the set of all $x \in \Sigma_N$ such that $x|_n = \mathtt{i}$. We will also write $x|_n := (x_k)_{k=1}^n$ when $x \in \hat{\Sigma}_N$ and denote the set $\{x \in \hat{\Sigma}_N : x|_n = \mathtt{i}\}$ by $[\mathtt{i}]$ when the difference of context is clear. We refer to sets of the form $[\mathtt{i}]$ as *cylinder sets*. Cylinder sets generate the topology of Σ_N, and shifted cylinder sets $\hat{\sigma}^n[\mathtt{i}]$ suffice to generate the topology of $\hat{\Sigma}_N$. We will usually denote words of length 1 simply by the symbol in $\{1, \ldots, N\}$ which appears in that word, and the cylinders defined by words of length 1 are therefore denoted $[1], \ldots, [N]$.

We define the natural projection $\pi \colon \hat{\Sigma}_N \to \Sigma_N$ by $\pi[(x_k)_{k \in \mathbb{Z}}] := (x_k)_{k=1}^\infty$ which is clearly continuous and surjective. It is clear that $\hat{\mu} \mapsto \pi_* \hat{\mu}$ defines a continuous function $\mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N) \to \mathcal{M}_\sigma(\Sigma_N)$ and since shift-invariant measures on Σ_N and on $\hat{\Sigma}_N$ are in both cases characterised by their values on cylinder sets this map is bijective. Given $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ we will simply write $\hat{\mu}$ for the unique element of $\mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)$ such that $\mu = \pi_* \hat{\mu}$, and we call $\hat{\mu}$ the natural extension of the measure μ. Since properties such as ergodicity, total ergodicity and mixing can be characterised in terms of correlations between cylinder sets it is not difficult to see that each of those properties holds for an invariant measure $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ if and only if the corresponding property holds for $\hat{\mu} \in \mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)$. A measure $\hat{\mu}$ on $\hat{\Sigma}_N$ will be called a Bernoulli measure if it has the form $\hat{\mu} = (\sum_{i=1}^N p_i \delta_i)^{\mathbb{Z}}$ for some probability vector $(p_1, \ldots, p_N)$. We will say $\hat{\mu}$ *has the Bernoulli property* if there exist a Bernoulli measure $\hat{\nu}$ on $\hat{\Sigma}_N$ and a measure-space isomorphism $\phi \colon \hat{\Sigma}_N \to \hat{\Sigma}_N$ such that $\phi \circ \hat{\sigma} = \hat{\sigma} \circ \phi$ and $\phi_* \hat{\mu} = \hat{\nu}$. (This isomorphism must be understood with respect to the completions of the relevant Borel σ-algebras: see §7 for details.) Clearly every Bernoulli measure trivially has the Bernoulli property, but the reverse is in general false.

2.2. **Potentials and equilibrium states.** For the remainder of this article a *potential* will be any function $\Phi \colon \Sigma_N^* \to (0, +\infty)$, where $N \geq 2$ is arbitrary. We call a potential *submultiplicative* if it satisfies the inequality $\Phi(\mathtt{ij}) \leq \Phi(\mathtt{i})\Phi(\mathtt{j})$ for all $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$ and *quasimultiplicative* if there exist $\delta > 0$ and $m \geq 1$ such that $\max_{|\mathtt{k}| \leq m} \Phi(\mathtt{ikj}) \geq \delta \Phi(\mathtt{i})\Phi(\mathtt{j})$ for all $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$. If $\Phi \colon \Sigma_N^* \to (0, +\infty)$ is a submultiplicative potential then we define its *pressure* to be the limit

$$P(\Phi) := \lim_{n \to \infty} \frac{1}{n} \log \sum_{\substack{\mathtt{i} \in \Sigma_N^* \\ |\mathtt{i}| = n}} \Phi(\mathtt{i})$$

which exists by subadditivity. If additionally μ is a shift-invariant measure on Σ_N then we define the *ergodic average* of Φ to be the quantity

$$
\begin{aligned}
\Lambda(\Phi, \mu) &:= \lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N} \log \Phi(x|_n) d\mu(x) \\
&= \lim_{n \to \infty} \frac{1}{n} \int_{\hat{\Sigma}_N} \log \Phi(x|_n) d\hat{\mu}(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{\substack{\mathbf{i} \in \Sigma_N^* \\ |\mathbf{i}|=n}} \mu([\mathbf{i}]) \log \Phi(\mathbf{i});
\end{aligned}
$$

this limit likewise exists by subadditivity. When we wish to emphasise that we are working on the two-sided shift space $\hat{\Sigma}_N$ we may also denote this quantity by $\Lambda(\Phi, \hat{\mu})$. If Φ is a submultiplicative potential defined on Σ_N^* then the pressure of Φ admits the characterisation

$$
P(\Phi) = \sup_{\mu \in \mathcal{M}_\sigma(\Sigma_N)} [h(\mu) + \Lambda(\Phi, \mu)] = \sup_{\hat{\mu} \in \mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)} [h(\hat{\mu}) + \Lambda(\Phi, \hat{\mu})]
$$

a fact which follows from more general results obtained in [16]. We will prefer to say that an *equilibrium state* of Φ is a measure $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ such that $P(\Phi) = h(\mu) + \Lambda(\Phi, \mu)$, and in this case we call the measure $\hat{\mu}$ the natural extension of an equilibrium state. However, this choice of terminology is somewhat arbitrary and is chosen solely in order to have distinct names for μ and for $\hat{\mu}$. Since $\mathcal{M}_\sigma(\Sigma_N)$ is a compact metrisable topological space with respect to its weak-* topology and the quantities $h(\mu)$ and $\Lambda(\phi, \mu)$ depend upper semi-continuously on $\mu \in \mathcal{M}_\sigma(\Sigma_N)$, the existence of at least one equilibrium state for an arbitrary submultiplicative potential follows by elementary topological considerations. Since both $h(\mu)$ and $\Lambda(\Phi, \mu)$ are also affine as functions of the measure μ, the set of equilibrium states of a submultiplicative potential is moreover convex and its extreme points are ergodic measures with respect to σ.

We may now define generalised matrix equilibrium states. If μ is a shift-invariant measure on Σ_N, where $N \geq 2$, and $(A_1, \dots, A_N) \in \mathrm{GL}(V)^N$ is a tuple of linear maps defined on a real finite-dimensional vector space V, we write $A_{\mathbf{i}} := A_{i_1} A_{i_2} \cdots A_{i_n}$ for every $\mathbf{i} = (i_k)_{j=1}^n \in \Sigma_N^*$ and define

$$
\begin{aligned}
\lambda_1(A, \mu) &:= \lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N} \log \|A_{x|_n}\| d\mu(x) \\
&= \lim_{n \to \infty} \frac{1}{n} \int_{\hat{\Sigma}_N} \log \|A_{x|_n}\| d\hat{\mu}(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{|\mathbf{i}|=n} \mu([\mathbf{i}]) \log \|A_{\mathbf{i}}\|
\end{aligned}
$$

which we call the *top Lyapunov exponent* of $(A_1, \dots, A_N)$ with respect to μ. We say that a shift-invariant measure μ on Σ_N, where $N \geq 2$, is a *generalised matrix equilibrium state* if for some integer $k \geq 1$ there exist for each $j = 1, \dots, k$ a finite-dimensional real vector space V_j, a tuple $(A_1^{(j)}, \dots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ of invertible linear maps, and a real number $\beta_j > 0$, such that μ is an equilibrium state of the potential $\Phi \colon \Sigma_{N*} \to (0, +\infty)$ defined by

$$
\tag{1} \Phi(\mathbf{i}) := \prod_{j=1}^k \left\| A_{\mathbf{i}}^{(j)} \right\|^{\beta_j}.
$$

It is easy to check that μ is an equilibrium state of Φ if and only if it maximises the quantity $h(\nu) + \sum_{j=1}^k \beta_j \lambda_1(A^{(j)}, \nu)$ over all $\nu \in \mathcal{M}_\sigma(\Sigma_N)$. The relation between

tuples of linear maps $A_i^{(j)}$ and constants β_j on the one hand, and generalised matrix equilibrium states on the other hand, is not bijective: a potential of the above form may have multiple equilibrium states (see for example [6, 10, 24, 46]) and a single measure may be the equilibrium state of more than one such potential. Indeed, when proving theorems concerning generalised matrix equilibrium states it is often advantageous to look for spaces V_j and tuples $(A_1^{(j)}, \ldots, A_N^{(j)})$ which yield the same equilibrium state but have additional properties to those strictly required by the definition.

In the case $k = 1$ generalised matrix equilibrium states as defined above have been previously referred to simply as *matrix equilibrium states* or *matrix Gibbs states*, which motivates our choice of terminology: see for example [42, 43, 50]. Matrix equilibrium states are sufficient to study the (candidate) measures of maximal dimension for self-affine subsets of the plane, but for self-affine subsets of $\mathbb{R}^d$ with $d > 2$ it seems to be unavoidably necessary to consider the case $k = 2$: specifically, one must consider potentials of the form

$$(2) \qquad \Phi^s(\mathtt{i}) := \left\| A_{\mathtt{i}}^{\wedge \lfloor s \rfloor} \right\|^{1 + \lfloor s \rfloor - s} \left\| A_{\mathtt{i}}^{\wedge \lceil s \rceil} \right\|^{s - \lfloor s \rfloor}$$

for an appropriate tuple $(A_1, \ldots, A_N) \in \mathrm{GL}_d(\mathbb{R})^N$ and real number $s \in (0, d)$, see §2.4 below for details. Matrix equilibrium states (i.e. the case $k = 1$) are substantially easier to handle and results in the general case have typically been preceded in the literature by results in the case $k = 1$; the reason for this difference in difficulty can be attributed to the fact that the case $k = 1$ can be understood using linear-algebraic techniques by embedding the linear maps $A_1, \ldots, A_N \in \mathrm{GL}(V)$ in the subalgebra of $\mathrm{End}(V)$ which they generate, whereas for general k substantial progress has only been made by embedding the tuples $(A_1^{(j)}, \ldots, A_N^{(j)})$ in a linear algebraic group and applying ideas from algebraic geometry (see for example [10, 45]). While the definition of a generalised matrix equilibrium state also makes sense in some cases in which the linear maps $A_i^{(j)}$ are not assumed to be invertible, this is more difficult to handle mathematically when $k > 1$ and in the present work we will always assume the invertibility of the linear maps $A_i^{(j)}$. We will also find the following terminology helpful: if V is a finite-dimensional real vector space then $(A_1, \ldots, A_N) \in \mathrm{GL}(V)^N$ will be called *irreducible* if there does not exist a nonzero proper linear subspace U of V which is preserved by every A_i, and *strongly irreducible* if there does not exist a finite collection $\{U_1, \ldots, U_m\}$ of nonzero proper linear subspaces of V such that every A_i induces a permutation of the set $\{U_1, \ldots, U_m\}$.

2.3. **Main results and previous literature.** The early literature on matrix equilibrium states focused on studying the associated pressure function, proving the existence of equilibrium states, characterising their uniqueness (or otherwise) and describing their supports (see for example [22, 25, 24, 42]) with results on potentials of the form Φ^s as in (2) available only in special cases [19, 28, 33, 35]. The broader concept of a generalised matrix equilibrium state was introduced in [10] where an upper bound was given for the number of ergodic generalised matrix equilibrium states that can correspond to a single potential, and where it was also shown that all generalised matrix equilibrium states as defined above are fully supported on the relevant symbolic space Σ_N; these results in particular apply to potentials of the form (2) and subsumed many prior results on that topic. In parallel with this

work the author investigated several aspects of the ergodic properties of matrix equilibrium states in [42, 44], showing in particular that matrix equilibrium states are mixing if and only if they are totally ergodic and giving necessary and sufficient conditions for mixing. This left open the question of whether mixing matrix equilibrium states admit stronger properties such as being exact endomorphisms or having the Bernoulli property, and also left open the question of whether similar results hold for generalised matrix equilibrium states. The former question was partially addressed by M. Piraino in [50] in the case of matrix equilibrium states, where a sufficient condition was given for a matrix equilibrium state to have the Bernoulli property. This nonetheless left unresolved the question of whether *every* matrix equilibrium state which is mixing also has the Bernoulli property and did not address the corresponding questions for generalised matrix equilibrium states. In the present work we give complete answers to all of these questions in the following result:

Theorem 1. *Let $N \geq 2$ and let μ be a totally ergodic generalised matrix equilibrium state on Σ_N. Then μ is ψ-mixing:*

$$\lim_{n \to \infty} \sup_{\mathtt{i},\mathtt{j} \in \Sigma_N^*} \left| \frac{\mu([\mathtt{i}] \cap \sigma^{-n-|\mathtt{i}|}[\mathtt{j}])}{\mu([\mathtt{i}])\mu([\mathtt{j}])} - 1 \right| = 0,$$

and its natural extension $\hat{\mu}$ has the Bernoulli property.

It is interesting to ask whether the rate of convergence in the above limit is exponential as is the case in the classical thermodynamic formalism of additive potentials (see for example [2, 13, 48]); this may have implications for the statistical properties of typical trajectories with respect to μ as in, for example, [49, §7]. A sufficient condition for this exponential rate of convergence in the case of matrix equilibrium states was given by M. Piraino in [50], but to answer this question in the generality of Theorem 1 seems likely to require additional developments in the transfer operator theory of linear cocycles.

It is not difficult to extend Theorem 1 to the case where μ is ergodic but not totally ergodic, although the fundamental result in this direction is cumbersome to state:

Theorem 2. *Let $k \geq 1$ and $N \geq 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space and let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ and $\beta_j > 0$. For all $\mathtt{i} \in \Sigma_N^*$ define*

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}$$

and let μ be an ergodic equilibrium state of Φ. If μ is not totally ergodic then there exists an integer n satisfying $1 < n \leq \prod_{j=1}^{k} \dim V_j$ with the following properties. Let $\eta \colon \{\mathtt{i} \in \Sigma_N^* \colon |\mathtt{i}| = n\} \to \{1, \ldots, N^n\}$ be the map which takes each word $\mathtt{i} \in \Sigma_N^*$ of length n to the integer representing its position in the lexicographical ordering on $\{\mathtt{i} \in \Sigma_N^* \colon |\mathtt{i}| = n\}$ and define a homeomorphism $\iota \colon \Sigma_N \to \Sigma_{N^n}$ satisfying $\sigma \circ \iota = \iota \circ \sigma^n$ by $\iota[(x_\ell)_{\ell=1}^\infty] := (\eta(x_{(q-1)n+1} \cdots x_{qn}))_{q=1}^\infty$ for every $(x_\ell)_{\ell=1}^\infty \in \Sigma_N$. For each $j = 1, \ldots, k$ define an N^n-tuple $(B_1^{(j)}, \ldots, B_{N^n}^{(j)}) \in \mathrm{GL}(V_j)^{N^n}$ by $B_i^{(j)} := A_{\eta^{-1}(i)}^{(j)}$ for

every $i = 1, \ldots, N^n$ and $j = 1, \ldots, k$, and define a potential $\Psi \colon \Sigma_{N^n}^* \to (0, +\infty)$ by

$$\Psi(\mathsf{j}) = \prod_{j=1}^{k} \left\| B_{\mathsf{j}}^{(j)} \right\|^{\beta_j}$$

for all $\mathsf{j} \in \Sigma_{N^n}^*$. Then we may write $\mu = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_*^i \nu$ where ν is a σ^n-invariant measure on Σ_N and where for every $i = 0, \ldots, n-1$ the measure $(\iota \circ \sigma^i)_* \nu \in \mathcal{M}_\sigma(\Sigma_{N^n})$ is a distinct totally ergodic equilibrium state of Ψ.

The proof of Theorem 2 is technically straightforward but involves a substantial volume of checking and is given in §7 below. Since each measure $(\iota \circ \sigma^i)_* \nu$ is a totally ergodic equilibrium state of Ψ, by Theorem 1 its natural extension is measurably isomorphic to a Bernoulli process. It is not difficult to deduce:

Corollary 2.1. *Let $k \geq 1$ and $N \geq 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space and let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ and $\beta_j > 0$. For all $\mathsf{i} \in \Sigma_N^*$ define*

$$\Phi(\mathsf{i}) = \prod_{j=1}^{k} \left\| A_{\mathsf{i}}^{(j)} \right\|^{\beta_j}$$

and let μ be an ergodic equilibrium state of Φ. Then there exists an integer n satisfying $1 \leq n \leq \prod_{j=1}^{k} \dim V_j$ such that the natural extension of μ is measurably isomorphic to the product of a Bernoulli process with the rotation map $m \mapsto m+1$ mod n on $\mathbb{Z}_n$.

The proof of this corollary is likewise presented in §7. We also note the following:

Corollary 2.2. *Let $k \geq 1$ and $N \geq 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space and let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ and $\beta_j > 0$. For all $\mathsf{i} \in \Sigma_N^*$ define*

$$\Phi(\mathsf{i}) = \prod_{j=1}^{k} \left\| A_{\mathsf{i}}^{(j)} \right\|^{\beta_j}.$$

If every $(A_1^{(j)}, \ldots, A_N^{(j)})$ is strongly irreducible then there is a unique equilibrium state for Φ and that equilibrium state is ψ-mixing and has the Bernoulli property.

Proof. By [10, Corollary 2.2] if every $(A_1^{(j)}, \ldots, A_N^{(j)})$ is strongly irreducible then Φ must have a unique equilibrium state μ. If μ is not totally ergodic, let $n > 1$, Ψ and $(B_1^{(j)}, \ldots, B_{N^n}^{(j)})$ be as given by Theorem 2. It is easy to see that the tuples $(B_1^{(j)}, \ldots, B_{N^n}^{(j)})$ must also be strongly irreducible by construction, so by the same reasoning Ψ has a unique equilibrium state. But Theorem 2 implies that Ψ has at least n distinct ergodic equilibrium states, which is a contradiction. We conclude that μ must be totally ergodic, so Theorem 1 applies and μ is ψ-mixing and has the Bernoulli property. $\square$

In the case of matrix equilibrium states total ergodicity has already been fully characterised in the following sense. If V is a finite-dimensional real vector space and $(A_1, \ldots, A_N) \in \mathrm{GL}(V)^N$ is irreducible then for each $\beta > 0$ there exists a unique matrix equilibrium state for the potential $\Phi(\mathsf{i}) := \|A_{\mathsf{i}}\|^\beta$, see for example [24]. (Moreover, every ergodic matrix equilibrium state is the unique equilibrium state of such a potential.) In this situation it was shown in [44] that if this matrix

equilibrium state is *not* totally ergodic then there necessarily exists a *cyclic splitting* for V: we may write $V = \bigoplus_{j=1}^{m} U_j$ where each U_j is a linear subspace of V and where $A_i U_j = U_{j+1 \mod m}$ for all $i = 1, \ldots, N$ and $j = 1, \ldots, m$. (Examples in which total ergodicity of a matrix equilibrium state fails had already been constructed in [42].) It is natural to ask whether this result extends to generalised matrix equilibrium states: if a generalised matrix equilibrium state as in Theorem 1 is not totally ergodic, is it the case that one of the tuples $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ preserves a cyclic splitting for the associated vector space V_j? For generalised matrix equilibrium states the situation seems to be more complicated than this, and we are able to show that this result does not hold. We give the following example in which total ergodicity fails but the matrix tuples do not admit cyclic splittings:

Proposition 2.3. *Define two irreducible pairs of linear maps* $(A_1, A_2), (B_1, B_2) \in \mathrm{GL}_2(\mathbb{R})^2$ *by*

$$A_1 := \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \qquad A_2 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$B_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B_2 := \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

and let $\beta_1, \beta_2 > 0$ *be arbitrary. Define a potential* $\Phi \colon \Sigma_2^* \to (0, +\infty)$ *by* $\Phi(\mathtt{i}) := \|A_{\mathtt{i}}\|^{\beta_1} \|B_{\mathtt{i}}\|^{\beta_2}$. *Then* Φ *has a unique equilibrium state and this equilibrium state is not totally ergodic.*

The proof of Proposition 2.3 is also given in §7.

2.4. **Connections with self-affine sets.** We now describe in more detail the connections between Theorem 1 and self-affine sets. If $T_1, \ldots, T_N \colon \mathbb{R}^d \to \mathbb{R}^d$ are invertible affine contractions (with respect to some fixed norm on $\mathbb{R}^d$ which need not be the Euclidean norm) then there exists a unique nonempty compact set $X \subset \mathbb{R}^d$ satisfying $X = \bigcup_{i=1}^{N} T_i X$. Such sets X are referred to as *self-affine sets*. In the situation where the images $T_1 X, \ldots, T_N X$ are pairwise disjoint it is not difficult to define an expanding map $f \colon \mathbb{R}^d \to \mathbb{R}^d$ such that X is a repelling set for f and such that $D_x f = T_i^{-1}$ whenever $x \in T_i X$, so self-affine sets with this disjointness property (which is called the strong separation condition in the fractal geometry literature) are a particular case of the expanding repellers discussed in §1. Besides their connection with questions such as Conjecture 1 self-affine sets are the subject of a deep and substantial literature in their own right, beginning in the 1980s with such works as [8, 20, 40] and flowering into a highly active contemporary research topic (see for example [3, 4, 7, 10, 18, 23, 26, 27, 34]). As well as in its connection to Conjecture 1 the construction of high-dimensional measures on self-affine sets is important to the problem of obtaining sharp lower bounds on the Hausdorff dimension of the set itself. Theorem 1 in particular has implications for the structure of certain high-dimensional measures on self-affine sets, called *Käenmäki measures*, which we now describe.

If V is a d-dimensional vector space equipped with an inner product, we recall that the *singular values* of $A \in \mathrm{GL}(V)$ are defined to be the positive square roots of the eigenvalues of the positive definite linear map $A^\top A$. We write the singular values as $\sigma_1(A), \ldots, \sigma_d(A)$ in decreasing order with repetition in the case of multiple

eigenvalues. For each $s \geq 0$ and $A \in \mathrm{GL}_d(\mathbb{R})$ the *singular value function*, introduced by Falconer in [20], is the function $\varphi^s \colon \mathrm{GL}_d(\mathbb{R}) \to \mathbb{R}$ defined by

$$\varphi^s(A) := \begin{cases} \sigma_1(A) \cdots \sigma_{\lfloor s \rfloor}(A) \sigma_{\lceil s \rceil}(A)^{s - \lfloor s \rfloor} & \text{if } 0 \leq s \leq d, \\ |\det A|^{\frac{s}{d}} & \text{if } s \geq d, \end{cases}$$

where $\varphi^0(A)$ is understood to equal 1. The singular value function satisfies $\varphi^s(AB) \leq \varphi^s(A)\varphi^s(B)$ for all $A, B \in \mathrm{GL}_d(\mathbb{R})$. If $T_1, \ldots, T_N \colon \mathbb{R}^d \to \mathbb{R}^d$ are invertible affine contractions with respect to some fixed norm on $\mathbb{R}^d$, let us write each T_i in the form $T_i(u) := A_i u + v_i$ for all $u \in \mathbb{R}^d$, where $A_i \in \mathrm{GL}(\mathbb{R}^d)$ and $v_i \in \mathbb{R}^d$ for each $i = 1, \ldots, N$.

For each $s \geq 0$ we say that a φ^s-*equilibrium state* for $(T_1, \ldots, T_N)$ is an equilibrium state of the submultiplicative potential $\Phi^s(\mathtt{i}) := \varphi^s(A_\mathtt{i})$. It is not particularly difficult to show that the function $s \mapsto P(\Phi^s)$ is continuous and strictly decreasing with $P(\Phi^0) > 0$ and $\lim_{s \to \infty} P(\Phi^s) = -\infty$, so in particular there exists a unique $s > 0$ such that $P(\Phi^s) = 0$, called the *affinity dimension* of $(T_1, \ldots, T_N)$. By definition a *Käenmäki measure* for $(T_1, \ldots, T_N)$ is a φ^s-equilibrium state for $(T_1, \ldots, T_N)$ where s is the affinity dimension. Crucially every Käenmäki measure is a generalised matrix equilibrium state, since we have

$$\Phi^s(\mathtt{i}) = \begin{cases} \left\| A_\mathtt{i}^{\wedge \lfloor s \rfloor} \right\|^{1 + \lfloor s \rfloor - s} \left\| A_\mathtt{i}^{\wedge \lceil s \rceil} \right\|^{s - \lfloor s \rfloor} & \text{if } 0 \leq s \leq d, \\ |\det A_\mathtt{i}|^{\frac{s}{d}} & \text{if } s \geq d, \end{cases}$$

where $A^{\wedge k}$ denotes the kth exterior power of the linear map A; for details see the following section. (Here $A^{\wedge 0}$ is always understood to equal the identity linear map on $\mathbb{R}$.) It is not difficult to show that there exists a well-defined continuous function $\Pi \colon \Sigma_N \to \mathbb{R}^d$ which satisfies

$$\Pi\left[(x_k)_{k=1}^\infty\right] = \lim_{n \to \infty} T_{x_1} T_{x_2} \cdots T_{x_n} v$$

for all $v \in \mathbb{R}^d$, and indeed the image $\Pi(\Sigma_N)$ is precisely the attractor of $(T_1, \ldots, T_N)$. (It is for this reason that in this article we multiply matrices on the right – we define $A_\mathtt{i} := A_{i_1} \cdots A_{i_n}$ and not $A_\mathtt{i} := A_{i_n} \cdots A_{i_1}$ – and not on the left as is more natural in many other contexts.) It follows from a result of Jordan, Pollicott and Simon ([34], see also [36]) that if a shift-invariant measure μ on Σ_N has the property that $\Pi_* \mu$ has Hausdorff dimension equal to the affinity dimension then it is necessarily a Käenmäki measure for $(T_1, \ldots, T_N)$, and in this sense Käenmäki measures are the natural candidates for the measures of maximal dimension on self-affine sets.

Theorem 1 and Corollary 2.2 together yield the following result for Käenmäki measures:

Corollary 2.4. *Let* $T_1, \ldots, T_N \colon \mathbb{R}^d \to \mathbb{R}^d$ *be invertible affine maps which are all contracting with respect to some fixed norm on* $\mathbb{R}^d$ *and let* $s > 0$ *denote the affinity dimension of* $(T_1, \ldots, T_N)$. *If* μ *is a totally ergodic Käenmäki measure for* $(T_1, \ldots, T_N)$ *then*

$$\lim_{n \to \infty} \sup_{\mathtt{i}, \mathtt{j} \in \Sigma_N^*} \left| \frac{\mu([\mathtt{i}] \cap \sigma^{-n - |\mathtt{i}|}[\mathtt{j}])}{\mu([\mathtt{i}])\mu([\mathtt{j}])} - 1 \right| = 0$$

and the natural extension of μ *is measurably isomorphic to a Bernoulli measure. This holds in particular if the tuples* $(A_1^{\wedge \lfloor s \rfloor}, \ldots, A_N^{\wedge \lfloor s \rfloor})$ *and* $(A_1^{\wedge \lceil s \rceil}, \ldots, A_N^{\wedge \lceil s \rceil})$ *are both strongly irreducible.*

In several works on the dimension theory of Käenmäki measures it has been possible to obtain stronger results if an additional assumption is made, called the *quasi-Bernoulli property*. A measure μ on Σ_N is called quasi-Bernoulli if there exists a constant $C > 0$ such that $C^{-1}\mu([\mathtt{i}])\mu([\mathtt{j}]) \leq \mu([\mathtt{ij}]) \leq C\mu([\mathtt{i}])\mu([\mathtt{j}])$ for all $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$. (In other literatures this property is sometimes called *local product structure*: see for example [11]). It follows from the results of [10] that every ergodic generalised matrix equilibrium state satisfies the upper bound $\mu([\mathtt{ij}]) \leq C\mu([\mathtt{i}])\mu([\mathtt{j}])$, but the lower bound does not hold in general (see for example [5]). If $T_1, \ldots, T_N$ are affine contractions of $\mathbb{R}^d$ with respect to some fixed norm, let us say that the *n-step recoding* of $(T_1, \ldots, T_N)$ is the N^n-tuple $(\hat{T}_1, \ldots, \hat{T}_{N^n}) := (T_1^n, T_1^{n-1}T_2, T_1^{n-1}T_3, \ldots, T_N^{n-1}T_{N-1}, T_N^n)$ which lists all compositions of the form $T_{i_1} \cdots T_{i_n}$ in lexicographical order. It is easy to see that if $X = \bigcup_{i=1}^N T_i X$ then $X = \bigcup_{i=1}^{N^n} \hat{T}_i X$, so the tuples $(T_1, \ldots, T_N)$ and $(\hat{T}_1, \ldots, \hat{T}_{N^n})$ describe the same self-affine set. Moreover one may show that the affinity dimensions of $(T_1, \ldots, T_N)$ and $(\hat{T}_1, \ldots, \hat{T}_{N^n})$ are equal. By recoding $(T_1, \ldots, T_N)$ by the integer n_0 given by Theorem 2 we may recode any $(T_1, \ldots, T_N)$ into a new tuple all of whose ergodic Käenmäki measures are totally ergodic and therefore are ψ-mixing. By recoding a second time we may for any prescribed $\varepsilon > 0$ guarantee that for every ergodic Käenmäki measure μ of the twice-recoded system $(\hat{T}_1, \ldots, \hat{T}_{N^n})$ we have

$$\sup_{\mathtt{i}, \mathtt{j} \in \Sigma_{N^n*}} \left| \frac{\nu([\mathtt{i}] \cap \sigma^{-1-|\mathtt{i}|}[\mathtt{j}])}{\nu([\mathtt{i}])\nu([\mathtt{j}])} - 1 \right| < \varepsilon$$

which is to say

$$(1 - \varepsilon)\mu([\mathtt{i}])\mu([\mathtt{j}]) \leq \sum_{\ell=1}^{N^n} \mu([\mathtt{i}\ell\mathtt{j}]) \leq (1 + \varepsilon)\mu([\mathtt{i}])\mu([\mathtt{j}])$$

for all $\mathtt{i}, \mathtt{j} \in \Sigma_{N^n}^*$. It is interesting to ask whether this property may have dimension-theoretic applications similar to those of the quasi-Bernoulli property.

2.5. **Strategy of proof and structure of the paper.** The fundamental objective in the proof of Theorem 1 is to establish, given a totally ergodic generalised matrix equilibrium state μ on Σ_N, the following property which we refer to as the *precondition for ψ-mixing*: there exist an integer $m \geq 1$ and a real number $\delta > 0$ depending only on μ such that

(3) $$\max_{\substack{\mathtt{k} \in \Sigma_N^* \\ |\mathtt{k}|=m}} \mu([\mathtt{ikj}]) \geq \delta\mu([\mathtt{i}])\mu([\mathtt{j}])$$

for all $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$. By combining this result with a theorem of R.C. Bradley [14] it can easily be deduced that the natural extension $\hat{\mu}$ is ψ-mixing, which implies the same result for μ. A celebrated theorem of N.A. Friedman and D.S. Ornstein [29] on isomorphism with Bernoulli processes then allows us to pass directly from the ψ-mixing property for $\hat{\mu}$ to the Bernoulli property. This basic strategy for proving ψ-mixing and deducing the Bernoulli property follows that used by M. Piraino in [50].

The route to the condition (3) divides naturally into three principal stages. In the first stage, which is relatively elementary, we show that every ergodic generalised matrix equilibrium state μ can be represented by a potential Φ defined in terms of tuples $(A_1^{(j)}, \ldots, A_N^{(j)})$ all of which are irreducible and all of which have

simple top Lyapunov exponent with respect to μ. This to some extent reprises arguments already used in [10] but with the additional detail of the top Lyapunov exponent to be considered. In the second stage we use analytic arguments to further show that μ is the *unique* equilibrium state of a potential of the form $\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^k \|A_{\mathtt{i}}^{(j)}|_{W_j}\|^{\beta_j}$ where $\mathcal{W}$ is a finite invariant set of tuples $(W_j)_{j=1}^k$ of subspaces W_j of the respective vector space V_j, and such that $\mathcal{W}$ has an additional combinatorial property called *primitivity*: this is the stage at which total ergodicity is used. In the third stage, which is more algebraic, these ingredients are combined to obtain the inequality (3). We may then deduce Theorem 1 from (3) in a fairly straightforward manner. This division into parts is reflected in similar divisions in the proofs of other major results on generalised matrix equilibrium states given in [10, 45]: in the first stage of the argument we obtain irreducibility, in the second stage we treat complications arising from the possibility of irreducibility without strong irreducibility, and in the last stage we deal with a reduced case in which the arguments applicable to the strongly irreducible case are available. To illustrate this we remark that in the strongly irreducible case, the arguments in the second stage mostly collapse to trivialities; and in the case where for each j there exists a one-dimensional space with finite orbit under $(A_1^{(j)}, \ldots, A_N^{(j)})$, the arguments in the second stage become of fundamental importance whereas those in the third stage become trivial instead.

The remainder of the paper is therefore structured as follows. In the following section we recall various foundational results in linear algebra, ergodic theory and algebraic geometry which will be used in various parts of the proof of Theorem 1. The three stages in the proof of (3) just described are given successively in sections 4 through 6. In §7 we combine these results to obtain Theorem 1 and also prove the various minor additional results described in this section.

3. PRELIMINARIES

3.1. **Linear algebra.** We first recall some concepts and identities from linear and multilinear algebra which will be used in various sections of this article. Here and throughout the article $\text{End}(V)$ denotes the vector space of linear endomorphisms of the vector space V, and $\rho(A)$ denotes the largest of the absolute values of the eigenvalues of the linear map $A \in \text{End}(V)$. Proofs of the following statements concerning exterior powers and tensor products may be found in, for example, [39, §XVI]; the material on singular values is more commonly found in texts on matrix analysis such as [31].

3.1.1. *Exterior powers.* If V is a finite-dimensional real (or complex) vector space of dimension d then for every $k = 1, \ldots, d$ there exists a vector space $\wedge^k V$ of dimension $\binom{d}{k}$, called the kth exterior power of V, which is spanned by all expressions of the form $v_1 \wedge v_2 \wedge \cdots \wedge v_k$ such that $v_1, \ldots, v_k \in V$. These objects are subject to the identities

$$(\lambda v_1 + u_1) \wedge v_2 \wedge \cdots \wedge v_k = \lambda(v_1 \wedge v_2 \wedge + \cdots + \wedge v_k) + u_1 \wedge v_2 \wedge + \cdots + \wedge v_k,$$

$$v_1 \wedge \cdots \wedge v_i \wedge v_{i+1} \wedge \cdots \wedge v_k = -v_1 \wedge \cdots \wedge v_{i+1} \wedge v_i \wedge \cdots \wedge v_k$$

for all $v_1, \ldots, v_k, u_1 \in V$, all $i \in \{1, \ldots, k-1\}$ and all λ in $\mathbb{R}$ (or $\mathbb{C}$). If $e_1, \ldots, e_d$ is a basis for V then the vectors $e_{i_1} \wedge \cdots \wedge e_{i_k}$ such that $1 \leq i_1 < \cdots < i_k \leq d$ form a basis for $\wedge^k V$. If $A \in \text{End}(V)$ then the kth exterior power of A is the unique

linear map $A^{\wedge k} \in \mathrm{End}(\wedge^k V)$ characterised by the property $A^{\wedge k}(v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k$ for every $v_1, \ldots, v_k \in V$. The identity $(AB)^{\wedge k} = A^{\wedge k} B^{\wedge k}$ for all $A, B \in \mathrm{End}(V)$ is clear. By considering appropriate bases it is not difficult to see that if the eigenvalues of A are $\lambda_1, \ldots, \lambda_d$ then the eigenvalues of $A^{\wedge k}$ are precisely the products $\lambda_{i_1} \cdots \lambda_{i_k}$ such that $1 \leq i_1 < \cdots < i_k \leq d$. If V is additionally equipped with an inner product $\langle \cdot, \cdot \rangle$ then it induces an inner product on $\wedge^k V$ by $\langle u_1 \wedge \cdots \wedge u_k, v_1 \wedge \cdots \wedge v_k \rangle := \det([\langle u_i, v_j \rangle]_{i,j=1}^k)$, and with respect to these inner products it is clear that $(A^\top)^{\wedge k} \equiv (A^{\wedge k})^\top$.

3.1.2. *Tensor products.* If $V_1, \ldots, V_k$ are finite-dimensional real (or complex) vector spaces then their tensor product $\bigotimes_{i=1}^k V_i$ is a vector space of dimension $\prod_{i=1}^k \dim V_i$ spanned by all expressions of the form $v_1 \otimes \cdots \otimes v_k$ such that $v_i \in V_i$ for every $i = 1, \ldots, k$, subject to the identity

$$v_1 \otimes \cdots \otimes (\lambda v_i + u) \otimes \cdots \otimes v_k = \lambda(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_k) + v_1 \otimes \cdots \otimes u \otimes \cdots \otimes v_k$$

for all $v_1 \in V_1$, $v_2 \in V_2, \ldots, v_k \in V_k$, all λ in $\mathbb{R}$ (or $\mathbb{C}$) and $u \in V_i$, for all $i = 1, \ldots, k$. If for each $i = 1, \ldots, k$ we are given a basis $e_{1,i}, \ldots, e_{d_i,i}$ for V_i then the $\prod_{i=1}^k d_i$ vectors of the form $e_{j_1,1} \otimes e_{j_2,2} \otimes \cdots \otimes e_{j_k,k}$ with $1 \leq j_i \leq d_i$ for each $i = 1, \ldots, k$ form a basis for $\bigotimes_{i=1}^k V_i$. If linear maps $A_1 \in \mathrm{End}(V_1), \ldots, A_k \in \mathrm{End}(V_k)$ are given then they induce a linear map $\bigotimes_{i=1}^k A_i$ on $\bigotimes_{i=1}^k V_i$ by $(\bigotimes_{i=1}^k A_i)(u_1 \otimes \cdots \otimes u_k) := A_1 u_1 \otimes \cdots \otimes A_k u_k$. By considering appropriate bases it is not difficult to show that the $\prod_{i=1}^k \dim V_i$ eigenvalues of $\bigotimes_{i=1}^k A_i$ are precisely the products of the form $\prod_{i=1}^k \lambda_i$ where for each i the number λ_i is an eigenvalue of A_i. In particular we have $\rho(\bigotimes_{i=1}^k A_i) = \prod_{i=1}^k \rho(A_i)$ whenever $A_i \in \mathrm{End}(A_i)$ for every $i = 1, \ldots, k$. If for each i we are given an inner product $\langle \cdot, \cdot \rangle_{V_i}$ on V_i then we may define an inner product on $\bigotimes_{i=1}^k V_i$ by defining $\langle u_1 \otimes \cdots \otimes u_k, v_1 \otimes \cdots \otimes v_k \rangle := \prod_{i=1}^k \langle u_i, v_i \rangle_{V_i}$ for all $v_1 \in V_1, \ldots, v_k \in V_k$ and extending linearly. It is not difficult to see that with respect to this inner product we have $(\bigotimes_{i=1}^k A_i)^\top \equiv \bigotimes_{i=1}^k (A_i^\top)$ and the identity $\|\bigotimes_{i=1}^k A_i\| = \prod_{i=1}^k \|A_i\|$ follows.

3.1.3. *Singular values.* If V is a d-dimensional real or complex vector space equipped with an inner product, the singular values of a linear map $A \in \mathrm{End}(V)$ are defined to be the non-negative square roots of the eigenvalues of the positive semidefinite linear map $A^\top A$ listed in decreasing order with repetition in the case of multiple eigenvalues. All vector spaces in this article will be assumed to be equipped with inner products. If $\dim V = d$ we denote the singular values of A by $\sigma_1(A), \ldots, \sigma_d(A)$. The singular values are well known to satisfy the alternative characterisation

$$\sigma_i(A) = \min\{\|A - F\| \colon \mathrm{rank}\, F < i\}$$

for all $A \in \mathrm{End}(V)$ and $i = 1, \ldots, d$. It may be easily demonstrated using these two descriptions that the singular values satisfy the identities $\prod_{i=1}^d \sigma_i(A) = |\det A|$ and $\|A\| = \sigma_1(A)$ and also satisfy the inequality $\sigma_i(X_1 A X_2) \leq \|X_1\| \cdot \sigma_i(A) \cdot \|X_2\|$ for all $A, X_1, X_2 \in \mathrm{End}(V)$ and $i = 1, \ldots, d$. For every $k = 1, \ldots, d$ and $A \in \mathrm{End}(V)$ the singular values of $A^{\wedge k}$ (relative to the inner product on $\wedge^k V$ induced by the inner product on V) are the square roots of the eigenvalues of $(A^{\wedge k})^\top A^{\wedge k} = (A^\top A)^{\wedge k}$ and hence are precisely the products $\sigma_{i_1}(A) \cdots \sigma_{i_k}(A)$ such that $1 \leq i_1 < \cdots < i_k \leq d$. In particular the largest singular value of $A^{\wedge k}$ is $\sigma_1(A) \cdots \sigma_k(A)$, so we have $\|A^{\wedge k}\| = \sigma_1(A) \cdots \sigma_k(A)$ for all $A \in \mathrm{End}(V)$ and

$k = 1, \ldots, d$. The inequality $\prod_{i=1}^{k} \sigma_i(AB) \leq (\prod_{i=1}^{k} \sigma_i(A))(\prod_{i=1}^{k} \sigma_i(B))$ for all $A, B \in \mathrm{End}(V)$ and all $k = 1, \ldots, d$ follows.

In general the singular values of A are defined only relative to a specified inner product on V and may change if a different inner product is used. If $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are distinct inner products on V, $X \in \mathrm{GL}(V)$ is an isometry from $(V, \langle \cdot, \cdot \rangle_1)$ to $(V, \langle \cdot, \cdot \rangle_2)$, and $\sigma_1(A), \ldots, \sigma_d(A)$ denote the singular values of $A \in \mathrm{GL}(V)$ as calculated with respect to $\langle \cdot, \cdot \rangle_1$ then the singular values of A as calculated with respect to $\langle \cdot, \cdot \rangle_2$ are precisely $\sigma_1(XAX^{-1}), \ldots, \sigma_d(XAX^{-1})$. In particular the ratio between the two values of $\sigma_i(A)$ as calculated according to the two distinct inner products is bounded above by $\|X\| \cdot \|X^{-1}\|$ and below by $\|X\|^{-1}\|X^{-1}\|^{-1}$ for all $A \in \mathrm{GL}(V)$. This will be significant in §3.3 below, and in general implies that when we are interested in limits of sequences of the form $\frac{1}{n} \log \sigma_i(A_n)$ for some sequence $(A_n)_{n=1}^{\infty}$ of elements of $\mathrm{End}(V)$, the value of the limit will be independent of the choice of inner product on V with respect to which the sequence of terms is calculated.

3.1.4. *Proximality.* Let V be a finite-dimensional real vector space equipped with an inner product. We will call a linear endomorphism $A \in \mathrm{End}(V)$ *proximal* if it has a unique eigenvalue of maximum modulus and that eigenvalue is simple. (Note that every linear endomorphism of a one-dimensional space is proximal.) We denote the set of all proximal endomorphisms of V by $\mathrm{Prox}(V)$. If $A \in \mathrm{Prox}(V)$ then we write $V^+(A)$ for the leading eigenspace of A and $V^-(A)$ for the unique A-invariant hyperplane which is complementary to $V^+(A)$. Using the perturbation theory of finite-dimensional linear maps it is not difficult to show that $\mathrm{Prox}(V)$ is an open subset of $\mathrm{End}(V)$ and that the functions V^+ and V^- are continuous on it. We also note that $A \in \mathrm{Prox}(V)$ if and only if $\lambda A \in \mathrm{Prox}(V)$ for every nonzero $\lambda \in \mathbb{R}$, if and only if $A^n \in \mathrm{Prox}(V)$ for every $n \geq 1$, and that the identities $V^+(A) = V^+(\lambda A) = V^+(A^n)$ and $V^-(A) = V^-(\lambda A) = V^-(A^n)$ are valid for all $A \in \mathrm{Prox}(V)$, all nonzero $\lambda \in \mathbb{R}$ and all positive integers n. By considering the Jordan form of A it is not difficult to see that $A \in \mathrm{Prox}(V)$ if and only if the limit $P := \lim_{n \to \infty} \|A^n\|^{-1} A^n$ exists, is not nilpotent, and has rank one. Under these conditions the limit P is clearly also proximal and satisfies $V^+(P) = V^+(A)$ and $V^-(P) = V^-(A)$. We observe that if $A \in \mathrm{End}(V)$ has rank one then $A \in \mathrm{Prox}(V)$ if and only if $A^2 \neq 0$, if and only if A is not nilpotent. We lastly remark that if $A \in \mathrm{End}(V)$ and $\rho(A)^2 > \sigma_1(A)\sigma_2(A)$ then A is necessarily proximal, since in this case by Gelfand's formula

$$\rho(A^{\wedge 2}) = \lim_{n \to \infty} \left\| \left(A^{\wedge 2}\right)^n \right\|^{\frac{1}{n}} = \inf_{n \geq 1} \left\| \left(A^{\wedge 2}\right)^n \right\|^{\frac{1}{n}} \leq \|A^{\wedge 2}\| = \sigma_1(A)\sigma_2(A) < \rho(A)^2,$$

and since $\rho(A^{\wedge 2})$ is the product of the absolute values of the two largest eigenvalues of A this implies that A has a unique, simple eigenvalue with absolute value $\rho(A)$ as required for A to be proximal.

3.2. **Linear algebraic groups.** In §6 we will need to consider the Zariski topology on the general linear group $\mathrm{GL}(V)$ of invertible linear transformations of a finite-dimensional real vector space V. We briefly summarise here, without proofs, the definition and important features of this topology which will be needed later. Proofs of the statements described in this section may be found in standard textbooks on linear algebraic groups such as [12, 32]; the introductory treatment of this

subject in [9] may be particularly helpful for readers approaching the subject from a background in ergodic theory.

If V_1 and V_2 are finite-dimensional real vector spaces then a function $p\colon V_1 \to V_2$ is called a *polynomial* if for some (then for every) choice of basis on V_1 and V_2, the coefficients of the vector $p(v) \in V_2$ with respect to the basis on V_2 are consistent polynomial functions of the coefficients of v with respect to the basis on V_1. A subset Z of a finite-dimensional real vector space V is called an *affine variety* if it is the common zero locus of some family of polynomial functions $V \to \mathbb{R}$. In particular V itself is an affine variety. When a (proper) subset of an affine variety is also an affine variety we call it a (proper) subvariety. If Z_1 and Z_2 are affine varieties which are subvarieties of real vector spaces V_1 and V_2 then we define a polynomial $Z_1 \to Z_2$ to be a function from Z_1 to Z_2 which can be realised as the restriction to Z_1 of a polynomial $V_1 \to V_2$.

The *Zariski topology* on an affine variety Z is defined to be the topology generated by declaring the affine subvarieties of Z to be the closed sets for the topology. The Zariski topology is much coarser than the standard (Euclidean) topology which Z inherits as a subset of its ambient vector space V, having far fewer open sets; in particular, it is not a Hausdorff topology. An affine variety is called an *irreducible variety* if it cannot be written as the union of a finite collection of proper subvarieties. In an irreducible variety, every Zariski open set is dense. One may show that every affine variety is equal to the union of finitely many irreducible subvarieties.

Importantly for our arguments, if V is a finite-dimensional real vector space then $\mathrm{GL}(V)$ may be given the structure of an affine variety by identifying it with the set of all linear operators on $V \oplus \mathbb{R}$ which have the form $A \oplus x$ for some $A \in \mathrm{End}(V)$ and $x \in \mathbb{R}$ such that $x \cdot (\det A) = 1$. This condition is clearly polynomial and therefore defines an affine subvariety of $\mathrm{End}(V \oplus \mathbb{R})$. This gives $\mathrm{GL}(V)$ the structure of an affine variety; in this structure a function $p\colon \mathrm{GL}(V) \to \mathbb{R}$ is a polynomial if $p(A)$ is a polynomial function of the matrix entries of A in some basis together with the additional variable $1/\det A$. We note that for every $B \in \mathrm{GL}(V)$ the maps $A \mapsto AB$ and $A \mapsto BA$ are homeomorphisms in the Zariski topology on $\mathrm{GL}(V)$, as is the map $A \mapsto A^{-1}$.

For the purposes of this article a linear algebraic group will be any Zariski-closed subgroup of $\mathrm{GL}(V)$, where V is a finite-dimensional real vector space. Importantly, the Zariski closure of a *subsemigroup* of $\mathrm{GL}(V)$ is always a linear algebraic group. Every linear algebraic group $G \leq \mathrm{GL}(V)$ has only finitely many connected components in the Zariski topology. These components are disjoint and there exists a unique component of G containing the identity, which we denote by G^0 and call the *identity component* of G. We note that since every $A \in G$ induces a Zariski homeomorphism of G by left (or right) multiplication, left or right multiplication by A induces a permutation of the connected components of G. It is not difficult to show that the identity component of G is a normal subgroup of G. If $G_1 \leq \mathrm{GL}(V_1)$ and $G_2 \leq \mathrm{GL}(V_2)$ are linear algebraic groups, a *regular representation* $\phi\colon G_1 \to G_2$ will be any group homomorphism which is also a polynomial. We call ϕ an *irreducible* representation if there is no proper nonzero linear subspace of V_2 which is preserved by every element of $\phi(G_1)$.

We finish this section by highlighting for the reader some important instances of Zariski closed sets which will be used in our arguments. If $G \leq \mathrm{GL}(V)$ is a linear algebraic group and $U_1, U_2 \subseteq V$ are linear subspaces then the set $\{A \in G \colon AU_1 =$

U_2} is Zariski closed, because if $u_1, \ldots, u_k$ is a basis for U_1 and $v_1, \ldots, v_\ell$ a basis for $U_2^\perp$ then this set is equal to the intersection of the sets $\{A \in G \colon \langle Au_i, v_j \rangle = 0\}$ over all $i = 1, \ldots, k$ and $j = 1, \ldots, \ell$, which is clearly a subvariety of G. Similarly if $v \in V$ is arbitrary then the set $\{A \in G \colon Av \in U_2\}$ is Zariski closed since A belongs to this set if and only if $\langle Av, v_j \rangle = 0$ for every $j = 1, \ldots, \ell$. We also note that if $B \in \mathrm{End}(V)$ is arbitrary then the set $\{A \in G \colon (AB)^2 = 0\}$ is Zariski closed since each of the finitely many entries of the matrix $(AB)^2$ is a polynomial function of the entries of A.

3.3. Lyapunov exponents.
In all sections of this article except §6 we will have frequent need to refer to Lyapunov exponents. Let V be a finite-dimensional real vector space, let $(A_1, \ldots, A_N) \in \mathrm{GL}(V)^N$ and let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$. If $\langle \cdot, \cdot \rangle$ is any inner product on V then we define the *Lyapunov exponents* of $(A_1, \ldots, A_N)$ to be the quantities

$$\lambda_i(A, \mu) := \lim_{n \to \infty} \frac{1}{n} \int \log \sigma_i(A_{x|_n}) \, d\mu(x)$$

for $i = 1, \ldots, \dim V$. For every $k = 1, \ldots, \dim V$ the limit

$$(4) \qquad \sum_{i=1}^k \lambda_i(A, \mu) = \lim_{n \to \infty} \frac{1}{n} \int \log \prod_{i=1}^k \sigma_i(A_{x|_n}) \, d\mu(x)$$

exists by subadditivity as a consequence of the inequality

$$\prod_{i=1}^k \sigma_i(AB) \leq \left(\prod_{i=1}^k \sigma_i(A) \right) \left(\prod_{i=1}^k \sigma_i(B) \right)$$

noted in §3.1.3, and it follows that the limit in the definition of $\lambda_i(A, \mu)$ is well-defined for every $i = 1, \ldots, k$ since it is a difference of two limits of the form (4). It is clear that $\lambda_1(A, \mu) \geq \lambda_2(A, \mu) \geq \cdots \geq \lambda_{\dim V}(A, \mu)$ as a consequence of the corresponding inequality for singular values. We say that $(A_1, \ldots, A_N) \in \mathrm{GL}(V)^N$ has *simple top Lyapunov exponent with respect to* μ if $\lambda_1(A, \mu) > \lambda_2(A, \mu)$.

The Lyapunov exponents are independent of the choice of inner product on V which is used to define the singular values: if $\sigma_i(A)$ and $\hat{\sigma}_i(A)$ denote the ith singular value of A calculated using two different inner products on V then as remarked in §3.1.3 there is a constant $C > 0$ such that $|\log \sigma_i(A) - \log \hat{\sigma}_i(A)| \leq C$ for all $A \in \mathrm{GL}(V)$, and consequently

$$\lim_{n \to \infty} \frac{1}{n} \int \log \sigma_i(A_{x|_n}) \, d\mu(x) = \lim_{n \to \infty} \frac{1}{n} \int \log \hat{\sigma}_i(A_{x|_n}) \, d\mu(x).$$

In particular we are at liberty to change the inner product on V without affecting the Lyapunov exponents, if there is advantage in doing so. We will take advantage of this principle in §5 below.

We lastly note the following useful result which will be applied in §4 and §5:

Proposition 3.1 ([41]). *Let $N \geq 2$, let V be a finite-dimensional real vector space, let $(B_1, \ldots, B_N) \in \mathrm{GL}(V)^N$ and let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be ergodic. Then*

$$\limsup_{n \to \infty} \frac{1}{n} \log \rho\left(B_{x|_n}\right) = \lambda_1(B, \mu)$$

for μ-a.e. $x \in \Sigma_N$.

Proof. The result follows by applying the subadditive ergodic theorem and [41, Theorem 1.5] to the cocycle $\mathcal{A}\colon \Sigma_N \times \mathbb{N} \to \mathrm{GL}(V)$ defined by $\mathcal{A}(x,n) := B_{x_n}^\top \cdots B_{x_1}^\top = B_{x|_n}^\top$. $\qquad\square$

3.4. **Prior results on generalised matrix equilibrium states.** Throughout this article we will require various facts on the structure of generalised matrix equilibrium states which were established in [10] and which we collect here for the reader's convenience.

3.4.1. *Subspace classes.* Ergodic generalised matrix equilibrium states were characterised in [10] via an algebraic object which we now describe. If V is a finite-dimensional real vector space then for the purposes of this article the Grassmannian of V, denoted $\mathrm{Gr}(V)$, is defined to be the set of all nonzero linear subspaces of V. Note that $\mathrm{Gr}(V)$ thus defined includes the space V itself. If $1 \le \ell \le \dim V$ then we let $\mathrm{Gr}_\ell(V)$ denote the set of all ℓ-dimensional linear subspaces of $\mathrm{Gr}(V)$.

Let $k \ge 1$, let $V_1, \ldots, V_k$ be finite-dimensional real vector spaces and let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ for every $j = 1, \ldots, k$. We define a *subspace class* to be any finite nonempty subset of $\prod_{j=1}^k \mathrm{Gr}(V_j)$. We will say that a subspace class $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}(V_j)$ is *equivariant* if for every $(W_j)_{j=1}^k \in \mathcal{W}$ we have

$$\left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} \subseteq \mathcal{W},$$

transitive if for every $(W_j)_{j=1}^k \in \mathcal{W}$

$$\left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} = \mathcal{W},$$

and *primitive* if for some integer $p \ge 1$ we have for every $(W_j)_{j=1}^k \in \mathcal{W}$

$$\left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \text{ and } |\mathtt{i}| = p \right\} = \mathcal{W}.$$

In other words $\mathcal{W}$ is equivariant if and only if for every $\mathtt{i} \in \Sigma_N^*$ and $(W_j)_{j=1}^k \in \mathcal{W}$ the tuple $(A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k$ also belongs to $\mathcal{W}$; is transitive if and only if for every $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}$ there exists $\mathtt{i} \in \Sigma_N^*$ such that $(A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k = (W_j')_{j=1}^k$; and is primitive if and only if the word $\mathtt{i}$ in this definition of transitivity can be chosen so as to have the same length for all choices of $(W_j)_{j=1}^k$ and $(W_j')_{j=1}^k$. A third description in terms of Perron-Frobenius theory may also be helpful. Suppose that we were to define a non-negative integer matrix M, whose rows and columns are indexed by the elements of $\mathcal{W}$, by placing 1 at the intersection of the row $(W_j)_{j=1}^k$ and the column $(W_j')_{j=1}^k$ if there exists $i \in \{1, \ldots, N\}$ such that $(A_i^{(j)} W_j)_{j=1}^k = (W_j')_{j=1}^k$, and 0 otherwise. Equivariance of $\mathcal{W}$ ensures that this definition makes sense and implies that every row of M has at least one nonzero entry; transitivity asserts precisely that M is an irreducible matrix in the standard sense of Perron-Frobenius theory; and primitivity asserts that M is a primitive matrix in the sense of Perron-Frobenius theory.

3.4.2. *Properties of generalised matrix equilibrium states.* The following result from [10] characterises ergodic generalised matrix equilibrium states in terms of subspace classes in the case where every $(A_1^{(j)}, \ldots, A_N^{(j)})$ is irreducible:

Theorem 3 ([10]). *Let $k \geq 1$ and $N \geq 2$ and for each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space, $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ an irreducible N-tuple of linear maps and $\beta_j > 0$ a real number. For each $j = 1, \ldots, k$ let $\ell_j \in \{1, \ldots, \dim V_j\}$ be the smallest integer such that there exists a nonzero linear subspace $U_j \subseteq V_j$ which has finite orbit under the action of the semigroup $\{A_{\mathtt{i}}^{(j)} \colon \mathtt{i} \in \Sigma_N^*\}$. Then:*

(i) If $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V)$ is a transitive subspace class, define a potential $\Phi_{\mathcal{W}} \colon \Sigma_N^* \to (0, +\infty)$ by

$$\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^k \left\| A_{\mathtt{i}}^{(j)} |_{W_j} \right\|^{\beta_j}.$$

Then $\Phi_{\mathcal{W}}$ is submultiplicative and quasimultiplicative and has a unique equilibrium state $\mu \in \mathcal{M}_\sigma(\Sigma_N)$. There exists $C > 1$ such that

$$C^{-1} \Phi_{\mathcal{W}}(\mathtt{i}) \leq e^{|\mathtt{i}| P(\Phi_{\mathcal{W}})} \mu([\mathtt{i}]) \leq C \Phi_{\mathcal{W}}(\mathtt{i})$$

for all $\mathtt{i} \in \Sigma_N^*$, and in particular μ is fully supported on Σ_N.

(ii) If we define a potential $\Phi \colon \Sigma_N^* \to (0, +\infty)$ by

$$\Phi(\mathtt{i}) := \prod_{j=1}^k \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}$$

for all $\mathtt{i} \in \Sigma_N^*$, then for every ergodic equilibrium state μ of Φ there exists a transitive subspace class $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V)$ such that μ is the unique equilibrium state of the potential $\Phi_{\mathcal{W}}$ defined as in (i). In particular every equilibrium state of Φ is fully supported on Σ_N.

In general a potential of the form $\Phi(\mathtt{i}) := \prod_{j=1}^k \|A_{\mathtt{i}}^{(j)}\|^{\beta_j}$ may admit multiple ergodic equilibrium states corresponding to different choices of transitive subspace class $\mathcal{W}$: the number of ergodic equilibrium states is bounded above by the quantity $(\prod_{1 \leq j \leq k} \dim V_j)/(\max_{1 \leq j \leq k} \dim V_j)$ and in at least some situations this bound can be attained, see [10]. Moreover, in general there may be infinitely many choices of subspace class $\mathcal{W}$ which generate the same equilibrium state. One of the major components of the proof of Theorem 1 will be an extension of Theorem 3(ii) in §5. In this result we will show that if additionally every $(A_1^{(j)}, \ldots, A_N^{(j)})$ has simple top Lyapunov exponent with respect to μ then the choice of transitive subspace class $\mathcal{W}$ is unique, and if μ is totally ergodic then moreover $\mathcal{W}$ must be primitive.

Outside the irreducible case, the following additional result of [10] implies that every ergodic generalised matrix equilibrium state can be expressed as the equilibrium state of a potential satisfying the hypotheses of Theorem 3, and will also be applied in the following section. In order to prove Theorem 1 we will likewise need to extend the below result to include a statement on simple top Lyapunov exponents.

Theorem 4 ([10]). *Let $k \geq 1$ and $N \geq 2$ and for each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space, $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ an N-tuple of linear maps and $\beta_j > 0$ a real number. Then for each $j = 1, \ldots, k$ there exist an integer $r_j \geq 1$, integers $d_{j,1}, \ldots, d_{j,r_j} \geq 1$ satisfying $\sum_{t=1}^{r_j} d_{j,t} = \dim V_j$ and a basis*

for V_j in which we may write

$$A_i^{(j)} = \begin{pmatrix} A_i^{(j,1)} & * & \cdots & * & * \\ 0 & A_i^{(j,2)} & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_i^{(j,r_j-1)} & * \\ 0 & 0 & \cdots & 0 & A_i^{(j,r_j)} \end{pmatrix}$$

for every $i = 1, \ldots, N$, where $(A_1^{(j,t)}, \ldots, A_N^{(j,t)}) \in \mathrm{GL}_{d_{j,t}}(\mathbb{R})^N$ is irreducible for every $t = 1, \ldots, r_j$. If μ is an ergodic equilibrium state of the potential $\Phi \colon \Sigma_N^* \to (0, +\infty)$ defined by

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j},$$

then there exist $t_1, \ldots, t_k$ such that $1 \leq t_j \leq r_j$ for every $j = 1, \ldots, k$ and such that μ is an equilibrium state of the potential $\hat{\Phi} \colon \Sigma_N^* \to (0, +\infty)$ defined by

$$\hat{\Phi}(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j,t_j)} \right\|^{\beta_j}$$

and satisfies $P(\Phi) = P(\hat{\Phi})$. Furthermore, the number of distinct ergodic equilibrium states of Φ is not greater than $\prod_{j=1}^{k} \dim V_j$.

Remark. Theorem 4 follows from the statement of [10, Theorem 5] except for the fact that $P(\Phi) = P(\hat{\Phi})$, which is not made explicit in the statement of that theorem but appears in the theorem's proof. Similarly, in Theorem 3 the fact that every potential of the form $\Phi_{\mathcal{W}}$ where $\mathcal{W} \subseteq \prod_{j=1}^{k} \mathrm{Gr}_{\ell_j}(V_j)$ is a transitive subspace class is quasimultiplicative and satisfies a Gibbs inequality was not made explicit in the statement of [10, Theorem 4] but features prominently in the proof.

4. REDUCTION TO THE CASE OF SIMPLE TOP LYAPUNOV EXPONENTS

As was described in the introduction the first, and by far the simplest, step in the proof of Theorem 1 is to reduce the problem to the case where the generalised matrix equilibrium state μ is defined by tuples $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ which are all irreducible and all have simple top Lyapunov exponent with respect to μ. In this section we prove:

Theorem 5. *Let $N \geq 2$ and let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be an ergodic generalised matrix equilibrium state. Then there exist finite-dimensional real vector spaces V_j, irreducible tuples $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ each having simple top Lyapunov exponent with respect to μ, and real numbers $\beta_j > 0$ for each $j = 1, \ldots, k$ such that μ is an ergodic equilibrium state of the potential*

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}.$$

By definition μ admits at least one representation in the above form but without each $(B_1^{(j)}, \ldots, B_N^{(j)})$ necessarily being irreducible or having simple top Lyapunov exponent. The result is proved by starting with such a representation, passing to

an appropriate exterior power for each j and then finding a block upper triangu-
larisation of each tuple such that for each j one of the tuples of diagonal blocks
yields the desired new tuple. We separate the first part of this argument into a
proposition as follows:

Proposition 4.1. *Let $N \geq 2$ and let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be a generalised matrix equi-
librium state. Then there exist finite-dimensional real vector spaces U_j, tuples of
linear maps $(B_1^{(j)}, \ldots, B_N^{(j)}) \in \mathrm{GL}(U_j)^N$ each having simple top Lyapunov exponent
with respect to μ, and real numbers $\gamma_j > 0$ for each $j = 1, \ldots, k$ such that μ is an
ergodic equilibrium state of the potential*

$$\Psi(\mathtt{i}) := \prod_{j=1}^{k} \left\| B_{\mathtt{i}}^{(j)} \right\|^{\gamma_j}.$$

Proof. Since μ is a generalised matrix equilibrium state, by definition there exist
$k \geq 1$, finite-dimensional real vector spaces V_j, tuples $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$
and real numbers $\beta_j > 0$ for each $j = 1, \ldots, k$ such that μ is an equilibrium state
for the potential $\Phi \colon \Sigma_N \to (0, +\infty)$ defined by

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}.$$

Choose for each $j = 1, \ldots, k$ the largest integer $\ell_j \in \{1, \ldots, \dim V_j\}$ such that
$\lambda_1(A^{(j)}, \mu) = \cdots = \lambda_{\ell_j}(A^{(j)}, \mu)$. Define $U_j := V_j^{\wedge \ell_j}$ for each $j = 1, \ldots, k$ and
$B_i^{(j)} := (A_i^{(j)})^{\wedge \ell_j}$ for each $i = 1, \ldots, N$ and $j = 1, \ldots, k$. For every $A \in \mathrm{GL}(V_j)$
the singular values of $A^{\wedge \ell_j}$ are precisely the numbers $\sigma_{i_1}(A) \cdots \sigma_{i_{\ell_j}}(A)$ such that
$1 \leq i_1 < \cdots < i_{\ell_j} \leq \dim V_j$, listed in decreasing order, so in particular the
largest singular value is $\sigma_1(A) \cdots \sigma_{\ell_j}(A)$ and the second-largest singular value is
$\sigma_1(A) \cdots \sigma_{\ell_j-1}(A)\sigma_{\ell_j+1}(A)$ if $\ell_j < \dim V_j$ and zero otherwise. It follows directly
that

$$\lambda_1(B^{(j)}, \mu) = \sum_{i=1}^{\ell_j} \lambda_i(A^{(j)}, \mu),$$

$$\lambda_2(B^{(j)}, \mu) = \left(\sum_{i=1}^{\ell_j-1} \lambda_i(A^{(j)}, \mu) \right) + \lambda_{\ell_j+1}(A^{(j)}, \mu)$$

for each $j = 1, \ldots, k$, where $\lambda_{1+\dim V_j}(A^{(j)}, \mu)$ is interpreted as being equal to
$-\infty$. By the maximality of each ℓ_j we have $\lambda_{\ell_j+1}(A^{(j)}, \mu) < \lambda_{\ell_j}(A^{(j)}, \mu)$ for each
$j = 1, \ldots, k$ and consequently $\lambda_1(B^{(j)}, \mu) > \lambda_2(B^{(j)}, \mu)$ for every $j = 1, \ldots, k$ so
that every $(B_1^{(j)}, \ldots, B_N^{(j)})$ has simple top Lyapunov exponent with respect to μ.
Define $\gamma_j := \beta_j/\ell_j$ for each $j = 1, \ldots, k$. We claim that μ is an equilibrium state
for the potential $\Psi \colon \Sigma_N^* \to (0, +\infty)$ defined by

$$\Psi(\mathtt{i}) := \prod_{j=1}^{k} \left\| B_{\mathtt{i}}^{(j)} \right\|^{\gamma_j}.$$

We have

$$\Psi(\mathtt{i}) := \prod_{j=1}^{k} \left\| B_{\mathtt{i}}^{(j)} \right\|^{\gamma_j} = \prod_{j=1}^{k} \sigma_1 \left(B_{\mathtt{i}}^{(j)} \right)^{\gamma_j}$$

$$= \prod_{j=1}^{k} \left(\sigma_1 \left(A_{\mathtt{i}}^{(j)} \right) \cdots \sigma_{\ell_j} \left(A_{\mathtt{i}}^{(j)} \right) \right)^{\frac{\beta_j}{\ell_j}} \leq \prod_{j=1}^{k} \sigma_1 \left(A_{\mathtt{i}}^{(j)} \right)^{\beta_j} = \Phi(\mathtt{i})$$

for every $\mathtt{i} \in \Sigma_N^*$ so in particular $P(\Psi) \leq P(\Phi)$. Thus

$$P(\Psi) \geq h(\mu) + \Lambda(\Psi, \mu) = h(\mu) + \sum_{j=1}^{k} \gamma_j \lambda_1 (B^{(j)}, \mu)$$

$$= h(\mu) + \sum_{j=1}^{k} \frac{\beta_j}{\ell_j} \sum_{i=1}^{\ell_j} \lambda_i (A^{(j)}, \mu)$$

$$= h(\mu) + \sum_{j=1}^{k} \beta_j \lambda_1 (A^{(j)}, \mu)$$

$$= h(\mu) + \Lambda(\Phi, \mu) = P(\Phi) \geq P(\Psi)$$

where we have used the definition of ℓ_j in the third equation, and μ is an equilibrium state of Ψ as claimed. $\square$

Proof of Theorem 5. By Proposition 4.1 we may assume without loss of generality that there exist finite-dimensional real vector spaces U_j, tuples of linear maps $(B_1^{(j)}, \ldots, B_N^{(j)}) \in \mathrm{GL}(U_j)^N$ each having simple top Lyapunov exponent with respect to μ, and real numbers $\gamma_j > 0$ for each $j = 1, \ldots, k$ such that μ is an ergodic equilibrium state of the potential

$$\Psi(\mathtt{i}) := \prod_{j=1}^{k} \left\| B_{\mathtt{i}}^{(j)} \right\|^{\gamma_j}.$$

By Theorem 4 there exist integers $r_1, \ldots, r_k \geq 1$ such that in a suitable basis for each U_j we may write

$$B_i^{(j)} = \begin{pmatrix} B_i^{(j,1)} & * & \cdots & * & * \\ 0 & B_i^{(j,2)} & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_i^{(j,r_j-1)} & * \\ 0 & 0 & \cdots & 0 & B_i^{(j,r_j)} \end{pmatrix}$$

for every $i = 1, \ldots, N$, where each $B_i^{(j,t)}$ is a real square matrix whose dimension $d_{t,j} \geq 1$ depends on t and j but not on i, where $(B_1^{(j,t)}, \ldots, B_N^{(j,t)}) \in \mathrm{GL}_{d_{j,t}}(\mathbb{R})^N$ is irreducible for every $t = 1, \ldots, r_j$, and where for some choice of integers $t_1, \ldots, t_k$ satisfying $1 \leq t_j \leq r_j$ for every $j = 1, \ldots, k$ the measure μ is an equilibrium state for the potential

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| B_{\mathtt{i}}^{(j,t_j)} \right\|^{\gamma_j}$$

and addtionally $P(\Phi) = P(\Psi)$. Define $V_j := \mathbb{R}^{d_{j,t_j}}$, $\beta_j := \gamma_j$ and $(A_1^{(j)}, \ldots, A_N^{(j)}) :=$ $(B_1^{(j,t_j)}, \ldots, B_N^{(j,t_j)}) \in \mathrm{GL}(V_j)^N$ for every $j = 1, \ldots, k$. To complete the proof of the theorem we must verify that every $(B_1^{(j,t_j)}, \ldots, B_N^{(j,t_j)})$ has simple top Lyapunov exponent with respect to μ. We will show that $(B_1^{(j,t_j)}, \ldots, B_N^{(j,t_j)})$ inherits this property from $(B_1^{(j)}, \ldots, B_N^{(j)})$.

To this end we first claim that $\lambda_1(B^{(j,t_j)}, \mu) = \lambda_1(B^{(j)}, \mu)$ for every $j = 1, \ldots, k$. On the one hand we have by definition

$$\lambda_1(B^{(j,t_j)}, \mu) = \lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N} \log \left\| B_{x|_n}^{(j,t_j)} \right\| \, d\mu(x)$$

$$\leq \lim_{n \to \infty} \frac{1}{n} \int_{\Sigma_N} \log \left\| B_{x|_n}^{(j)} \right\| \, d\mu(x) = \lambda_1(B^{(j)}, \mu)$$

so that $\lambda_1(B^{(j,t_j)}, \mu) \leq \lambda_1(B^{(j)}, \mu)$ for every $j = 1, \ldots, k$. On the other hand we may apply this estimate to obtain

$$P(\Psi) = P(\Phi) = h(\mu) + \Lambda(\Phi, \mu)$$

$$= h(\mu) + \sum_{j=1}^{k} \gamma_j \lambda_1(B^{(j,t_j)}, \mu)$$

$$\leq h(\mu) + \sum_{j=1}^{k} \gamma_j \lambda_1(B^{(j)}, \mu) = h(\mu) + \Lambda(\Psi, \mu) = P(\Psi)$$

and therefore $\sum_{j=1}^{k} \gamma_j \lambda_1(B^{(j,t_j)}, \mu) = \sum_{j=1}^{k} \gamma_j \lambda_1(B^{(j)}, \mu)$. It follows that

$$\sum_{j=1}^{k} \gamma_j \left(\lambda_1(B^{(j)}, \mu) - \lambda_1(B^{(j,t_j)}, \mu) \right)$$

is a sum of non-negative terms which is equal to zero, so all of the summands must be zero and therefore $\lambda_1(B^{(j,t_j)}, \mu) = \lambda_1(B^{(j)}, \mu)$ for every $j = 1, \ldots, k$ as claimed.

We secondly claim that

$$\lambda_1(B^{(j,t_j)}, \mu) + \lambda_2(B^{(j,t_j)}, \mu) \leq \lambda_1(B^{(j)}, \mu) + \lambda_2(B^{(j)}, \mu)$$

for every $j = 1, \ldots, k$. Indeed, for fixed j we have for μ-a.e. $x \in \Sigma_N$

$$\lambda_1(B^{(j,t_j)}, \mu) + \lambda_2(B^{(j,t_j)}, \mu) = \lambda_1 \left(\left(B^{(j,t_j)} \right)^{\wedge 2}, \mu \right) = \limsup_{n \to \infty} \frac{1}{n} \log \rho \left(\left(B_{x|_n}^{(j,t_j)} \right)^{\wedge 2} \right)$$

and

$$\lambda_1(B^{(j)}, \mu) + \lambda_2(B^{(j)}, \mu) = \lambda_1 \left(\left(B^{(j)} \right)^{\wedge 2}, \mu \right) = \limsup_{n \to \infty} \frac{1}{n} \log \rho \left(\left(B_{x|_n}^{(j)} \right)^{\wedge 2} \right)$$

by Proposition 3.1. For every x and n the quantity $\rho((B_{x|_n}^{(j,t_j)})^{\wedge 2})$ is the product of the absolute values of the two largest eigenvalues of $B_{x|_n}^{(j,t_j)}$ and similarly $\rho((B_{x|_n}^{(j)})^{\wedge 2})$ is the product of the absolute values of the two largest eigenvalues of $B_{x|_n}^{(j)}$. But the set of eigenvalues of $B_{x|_n}^{(j,t_j)}$ is a subset of the set of eigenvalues of $B_{x|_n}^{(j)}$ so we have

$$\limsup_{n \to \infty} \frac{1}{n} \log \rho \left(\left(B_{x|_n}^{(j,t_j)} \right)^{\wedge 2} \right) \leq \limsup_{n \to \infty} \frac{1}{n} \log \rho \left(\left(B_{x|_n}^{(j)} \right)^{\wedge 2} \right)$$

for all $x \in \Sigma_N$. The claim follows.

We may now show that every $(B_1^{(j,t_j)}, \ldots, B_N^{(j,t_j)})$ has simple top Lyapunov exponent with respect to μ. Combining the two claims we deduce that

$$\lambda_2(B^{(j,t_j)}, \mu) \leq \lambda_2(B^{(j)}, \mu)$$

and consequently

$$\lambda_2(B^{(j,t_j)}, \mu) \leq \lambda_2(B^{(j)}, \mu) < \lambda_1(B^{(j)}, \mu) = \lambda_1(B^{(j,t_j)}, \mu)$$

for every $j = 1, \ldots, k$ so that every $(B_1^{(j,t_j)}, \ldots, B_N^{(j,t_j)})$ has simple top Lyapunov exponent with respect to μ as required. The proof of the theorem is complete. $\square$

5. Generalised matrix equilibrium states in the case of simple top Lyapunov exponents

The result of the previous section shows that every ergodic generalised matrix equilibrium state μ can be assumed without loss of generality to be generated by a potential $\Phi(\mathtt{i}) := \prod_{j=1}^{k} \|A_{\mathtt{i}}^{(j)}\|^{\beta_j}$ defined in terms of tuples $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)$ which are all irreducible and have simple top Lyapunov exponent with respect to μ. By Theorem 3 irreducibility and ergodicity together imply that μ is the equilibrium state of a potential $\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^{k} \|A_{\mathtt{i}}^{(j)}|_{W_j}\|^{\beta_j}$ for some transitive subspace class $\mathcal{W} \subseteq \prod_{j=1}^{k} \mathrm{Gr}(V_j)$. In this section we extend Theorem 3 by showing that if μ is totally ergodic then irreducibility and the simplicity of Lyapunov exponents allow us to choose $\mathcal{W}$ so as to additionally be primitive. The same arguments which yield this result incidentally provide a pivotal technical lemma on simultaneously proximal words which recalls some results of Abels, Margulis and Soifer ([1], for related results see also [9, §6]), although in those works an algebraic rather than analytic method is used. In this section we prove:

Theorem 6. *Let $k \geq 1$ and $N \geq 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space, $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ an irreducible N-tuple and $\beta_j > 0$ a real number. For each $j = 1, \ldots, k$ let $\ell_j \in \{1, \ldots, \dim V_j\}$ be the smallest dimension of any nonzero linear subspace of V_j which has finite orbit under the action of $(A_1^{(j)}, \ldots, A_N^{(j)})$. Suppose that $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ is an ergodic equilibrium state of the potential $\Phi \colon \Sigma_N^* \to (0, +\infty)$ defined by*

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}$$

such that for every $j = 1, \ldots, k$ the tuple $(A_1^{(j)}, \ldots, A_N^{(j)})$ has a simple top Lyapunov exponent with respect to μ. Then:

(i) *There exists a* unique *transitive subspace class $\mathcal{W} \subseteq \prod_{j=1}^{k} \mathrm{Gr}_{\ell_j}(V_j)$ such that μ is an equilibrium state of the potential $\Phi_{\mathcal{W}} \colon \Sigma_N^* \to (0, +\infty)$ defined by*

$$\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)}|_{W_j} \right\|^{\beta_j}.$$

(ii) *The transitive subspace class $\mathcal{W}$ defined in (i) has the following additional property: there exist $\mathtt{k} \in \Sigma_N^*$ and $(W_j^0)_{j=1}^k \in \mathcal{W}$ such that for every $j = 1, \ldots, k$ we have $A_{\mathtt{k}}^{(j)} W_j^0 = W_j^0$ and $A_{\mathtt{k}}^{(j)}|_{W_j^0} \in \mathrm{Prox}(W_j^0)$.*

(iii) If μ is totally ergodic then the transitive subspace class $\mathcal{W}$ defined in (i) is primitive.

Theorem 6(i) is not used in this article, but is provided for interest. This result contrasts strongly with the situation in which simple top Lyapunov exponents are not assumed, where it can be the case that uncountably many choices of $\mathcal{W}$ exist which all have the same equilibrium state: a simple example of this, as mentioned in [10], is the case in which $k = 1$, $V = \mathbb{R}^2$ and every A_i is a rational rotation matrix, in which case $\mathcal{W}$ can be taken to be the orbit of any one-dimensional subspace whatsoever. Theorem 6(ii) and (iii) both contribute to the proof of Theorem 1; while the former has more of the character of a lemma than of a main result, we include it in Theorem 6 since it emerges directly from the proof of the other clauses of the theorem.

The proof of Theorem 6 is heavily inclined towards ergodic theory, unlike the results of the following section which are essentially algebraic. The essential idea of the proof is as follows. The separation of Lyapunov exponents allows us fairly easily to construct words $\mathbf{j} \in \Sigma_N^*$ such that $A_{\mathbf{j}}^{(j)} \in \text{Prox}(V_j)$ for all $j = 1, \ldots, k$, and the existence of these words implies that each V_j can be written as a splitting $V_j = \bigoplus_{i=1}^{r_j} U_j^i$ into ℓ_j-dimensional spaces in such a way that for each j the spaces $U_j^1, \ldots, U_j^{r_j}$ are permuted by the linear maps $A_i^{(j)}$. By replacing the inner product on each V_i if necessary this splitting can without loss of generality be taken to be orthogonal and it follows that the norm of the restriction of each product $A_{\mathbf{i}}^{(j)}$ to each of the various spaces U_j^i is a singular value of $A_{\mathbf{i}}^{(j)}$. Together with the separation of Lyapunov exponents this implies that for each $j = 1, \ldots, k$ the growth of $\|A_{x|_n}^{(j)}\|$ for almost every x is concentrated on a single one of the subspaces U_j^i, which in general will depend on both x and n. Using this observation we construct measurable functions $\mathsf{U}_j \colon \hat{\Sigma}_N \to \{U_j^1, \ldots, U_j^{r_j}\}$ which are equivariant with respect to the action of $A_{x|_n}^{(j)}$ and capture the maximal growth of each $\|A_{x|_n}^{(j)}\|$ in the sense that

$$\lim_{n \to \infty} \frac{1}{n} \log \left\| A_{x|_n}^{(j)} |_{\mathsf{U}_j(\hat{\sigma}^n x)} \right\| = \lambda_1(A^{(j)}, \mu),$$

$$\lim_{n \to \infty} \frac{1}{n} \log \left\| A_{x|_n}^{(j)} |_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu)$$

for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$. (These functions may be thought of as resembling Oseledets spaces, although they do not always correspond precisely to Oseledets spaces and it is not clear whether they are necessarily given by direct sums of Oseledets spaces either.) The ergodicity of $\hat{\mu}$ implies that the tuple $(\mathsf{U}_j(x))_{j=1}^k$ belongs almost everywhere to a single transitive subspace class $\mathcal{W}$; if $\hat{\mu}$ is also totally ergodic, this can be applied to show that for every integer $n \geq 1$, every two values taken by the tuple $(\mathsf{U}_j(x))_{j=1}^k$ on sets of positive measure can be linked by a word $\mathbf{i}$ with length divisible by n. This allows us to show that

$$\left\{ (A_{\mathbf{i}}^{(j)} W_j)_{j=1}^k \colon \mathbf{i} \in \Sigma_N^* \text{ and } n \text{ divides } |\mathbf{i}| \right\} = \mathcal{W}$$

for every $(W_j)_{j=1}^k \in \mathcal{W}$ and $n \geq 1$, and this is sufficient to deduce that $\mathcal{W}$ is primitive in the totally ergodic case.

Before beginning the proof we require three lemmas, one of which is primarily ergodic-theoretic in character, one primarily algebraic and one somewhat more

combinatorial; these respectively treat the existence of proximal elements, the algebraic consequences of their existence for splittings of each V_j, and the criterion for primitivity just mentioned.

Lemma 5.1. *Let $k \geq 1$ and $N \geq 2$ and let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be ergodic. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real inner product space and $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ a tuple which has simple top Lyapunov exponent with respect to μ. Then for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have*

$$
\text{(5)} \qquad \limsup_{n \to \infty} \min_{1 \leq j \leq k} \left(\frac{1}{n} \log \rho\left(A_{x|_n}^{(j)}\right) - \frac{1}{2n} \log \sigma_1\left(A_{x|_n}^{(j)}\right) \sigma_2\left(A_{x|_n}^{(j)}\right) \right) > 0.
$$

Furthermore there exists $\mathtt{k} \in \Sigma_N^*$ such that for every $j = 1, \ldots, k$ we have $A_{\mathtt{k}}^{(j)} \in \mathrm{Prox}(V_j)$.

Proof. The inequality (5) implies that we may find $\mathtt{k} := x|_n \in \Sigma_N^*$ such that $\rho(A_{\mathtt{k}}^{(j)})^2 > \sigma_1(A_{\mathtt{k}}^{(j)})\sigma_2(A_{\mathtt{k}}^{(j)})$ for all $j = 1, \ldots, k$ and as noted in §3.1.4 this yields $A_{\mathtt{k}}^{(j)} \in \mathrm{Prox}(V_j)$ for every $j = 1, \ldots, k$. To prove the lemma it is thus sufficient to show that (5) holds for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ and this clearly follows if the same result is shown for μ-a.e. $x \in \Sigma_N$. We therefore proceed to prove the latter statement.

We first claim that for μ-a.e. $x \in \Sigma_N$

$$
\text{(6)} \qquad \limsup_{n \to \infty} \sum_{j=1}^{k} \left(\frac{1}{n} \log \rho\left(A_{x|_n}^{(j)}\right) - \frac{1}{n} \log \left\| A_{x|_n}^{(j)} \right\| \right) = 0.
$$

To prove this let $V := \bigotimes_{j=1}^{k} V_j$ be the tensor product of the vector spaces V_j and define a tuple $(B_1, \ldots, B_N) \in \mathrm{GL}(V)^N$ by $B_i := \bigotimes_{j=1}^{k} A_i^{(j)}$ for each $i = 1, \ldots, N$. We equip V with the inner product induced by the inner products on the spaces V_j. As noted in §3.1 we have $\|B_{\mathtt{i}}\| = \prod_{j=1}^{k} \|A_{\mathtt{i}}^{(j)}\|$ for all $\mathtt{i} \in \Sigma_N^*$ with respect to the corresponding inner product norms and also $\rho(B_{\mathtt{i}}) = \prod_{j=1}^{k} \rho(A_{\mathtt{i}}^{(j)})$ for all $\mathtt{i} \in \Sigma_N^*$. It follows from the first statement in particular that $\lambda_1(B, \mu) = \sum_{j=1}^{k} \lambda_1(A^{(j)}, \mu)$. Since by the subadditive ergodic theorem

$$
\lim_{n \to \infty} \sum_{j=1}^{k} \frac{1}{n} \log \left\| A_{x|_n}^{(j)} \right\| = \lim_{n \to \infty} \frac{1}{n} \log \|B_{x|_n}\| = \lambda_1(B, \mu)
$$

for μ-a.e. $x \in \Sigma_N$, and by Proposition 3.1

$$
\limsup_{n \to \infty} \sum_{j=1}^{k} \frac{1}{n} \log \rho\left(A_{x|_n}^{(j)}\right) = \limsup_{n \to \infty} \frac{1}{n} \log \rho\left(B_{x|_n}\right) = \lambda_1(B, \mu)
$$

for μ-a.e. $x \in \Sigma_N$, the claimed result (6) follows. Now, each of the summands $\frac{1}{n} \log \rho\left(A_{x|_n}^{(j)}\right) - \frac{1}{n} \log \left\| A_{x|_n}^{(j)} \right\|$ in (6) is non-positive since clearly $\rho(A_{\mathtt{i}}^{(j)}) \leq \|A_{\mathtt{i}}^{(j)}\|$ for every $\mathtt{i} \in \Sigma_N^*$, and it follows from this observation that for μ-a.e. $x \in \Sigma_N$

$$
\text{(7)} \qquad \limsup_{n \to \infty} \min_{1 \leq j \leq k} \left(\frac{1}{n} \log \rho\left(A_{x|_n}^{(j)}\right) - \frac{1}{n} \log \left\| A_{x|_n}^{(j)} \right\| \right) = 0
$$

since this sequence is bounded below by the sequence in (6) and is bounded above by 0. But for μ-a.e. $x \in \Sigma_N$ we have

$$
(8) \qquad \lim_{n \to \infty} \min_{1 \leq j \leq k} \left(\frac{1}{n} \log \left\| A^{(j)}_{x|_n} \right\| - \frac{1}{2n} \log \sigma_1(A^{(j)}_{x|_n}) \sigma_2(A^{(j)}_{x|_n}) \right)
$$

$$
= \min_{1 \leq j \leq k} \left(\frac{1}{2} \lambda_1(A^{(j)}, \mu) - \frac{1}{2} \lambda_2(A^{(j)}, \mu) \right) > 0
$$

because by the subadditive ergodic theorem

$$
\lim_{n \to \infty} \frac{1}{n} \log \| A^{(j)}_{x|_n} \| = \lambda_1(A^{(j)}, \mu)
$$

and

$$
\lim_{n \to \infty} \frac{1}{n} \log \sigma_1 \left(A^{(j)}_{x|_n} \right) \sigma_2 \left(A^{(j)}_{x|_n} \right) = \lambda_1 \left(A^{(j)}, \mu \right) + \lambda_2 \left(A^{(j)}, \mu \right)
$$

for every $j = 1, \ldots, k$, and each of the differences $\lambda_1(A^{(j)}, \mu) - \lambda_2(A^{(j)}, \mu)$ is positive by the hypothesis of simplicity of the top Lyapunov exponent. Adding the sequences in (7) and (8) gives a sequence which is a lower bound for the left-hand side of (5) and has positive limit superior almost everywhere. This yields (5) for μ-a.e. $x \in \Sigma_N$ as required and the result follows. $\qquad \square$

The following algebraic consequence of proximality is perhaps already known but we have been unable to find a reference:

Lemma 5.2. *Let V be a finite-dimensional real vector space, let $(B_1, \ldots, B_N) \in \mathrm{GL}(V)^N$ be irreducible and let $\ell \in \{1, \ldots, \dim V\}$ be the smallest integer such that there exists an ℓ-dimensional linear subspace $U \subseteq V$ with finite orbit under the action of $(B_1, \ldots, B_N)$. Suppose that there exists $\mathtt{k} \in \Sigma_N^*$ such that $B_{\mathtt{k}} \in \mathrm{Prox}(V)$. Then $\dim V = \ell r$ for some integer $r \geq 1$, there are exactly r subspaces $U_1, \ldots, U_r \in \mathrm{Gr}_\ell(V)$ which have finite orbit under the action of $(B_1, \ldots, B_N)$, and these subspaces form a direct sum $\bigoplus_{t=1}^r U_t$ which is equal to V.*

Proof. Let $U \subseteq V$ be an ℓ-dimensional subspace with finite orbit under $(B_1, \ldots, B_N)$ and let $\mathcal{U} := \{B_{\mathtt{i}} U : \mathtt{i} \in \Sigma_N^*\}$ denote its orbit. We claim that $(B_1, \ldots, B_N)$ acts transitively on $\mathcal{U}$ in the following sense: for every $U_1, U_2 \in \mathcal{U}$ there exists $\mathtt{i} \in \Sigma_N^*$ such that $B_{\mathtt{i}} U_1 = U_2$.

To see this we argue as follows. For every $\mathtt{i} \in \Sigma_N^*$ the map $\mathrm{Gr}_\ell(V) \to \mathrm{Gr}_\ell(V)$ defined by $W \mapsto B_{\mathtt{i}} W$ is clearly bijective, so in particular each $B_{\mathtt{i}}$ induces an injective map from $\{U\} \cup \mathcal{U}$ to $\mathcal{U}$. Since $\mathcal{U}$ is finite this is only possible if $\{U\} \cup \mathcal{U} = \mathcal{U}$ and thus necessarily $U \in \mathcal{U}$. If $\mathtt{i} \in \Sigma_N^*$ is arbitrary then the map $B_{\mathtt{i}} \colon \mathcal{U} \to \mathcal{U}$ is a bijection of a finite set, so in particular some power of that map is the identity permutation. In particular if $\mathtt{i} \in \Sigma_N^*$ is arbitrary then $B_{\mathtt{i}}^p U = U$ for some integer $p \geq 1$ and hence also $B_{\mathtt{i}}^{np} U = U$ for every $n \geq 1$. It follows that if $U_1 \in \mathcal{U}$ is arbitrary then since by definition $U_1 = B_{\mathtt{j}} U$ for some $\mathtt{j} \in \Sigma_N^*$, the set $\{B_{\mathtt{i}} U_1 : \mathtt{i} \in \Sigma_N^*\}$ contains $B_{\mathtt{j}}^n U$ for every $n \geq 2$ and therefore contains U and hence contains $\{B_{\mathtt{i}} U : \mathtt{i} \in \Sigma_N^*\}$. But the reverse inclusion is obvious, so $\{B_{\mathtt{i}} U_1 : \mathtt{i} \in \Sigma_N^*\}$ is equal to $\mathcal{U}$ for every $U_1 \in \mathcal{U}$ and this proves the result claimed.

We next claim that there exist $r \geq 1$ and subspaces $U_1, \ldots, U_r \in \mathcal{U}$ such that $V = \bigoplus_{t=1}^r U_t$. To see this we observe that there exists at least one list of elements $U_1, \ldots, U_m \in \mathcal{U}$ which forms a direct sum, namely the list of length 1 consisting of the single space U. There therefore exists a *maximal* list of elements $U_1, \ldots, U_r \in \mathcal{U}$ which forms a direct sum. If $U' \in \mathcal{U}$ is arbitrary then the subspace $(U_1 \oplus \cdots \oplus U_r) \cap U'$

cannot have dimension zero since then $U_1, \ldots, U_r, U'$ would form a direct sum, contradicting maximality. It also cannot have dimension nonzero but strictly less than ℓ, since it clearly has finite orbit under the action of $(B_1, \ldots, B_N)$ and such a space must have dimension at least ℓ. It cannot have dimension strictly greater than ℓ since it is a subset of U', so by elimination its dimension is precisely ℓ and therefore it is equal to the subspace U'. Since $U' \in \mathcal{U}$ was arbitrary this shows that every element of $\mathcal{U}$ is a subspace of $U_1 \oplus \cdots \oplus U_r$. Hence $U_1 \oplus \cdots \oplus U_r = \mathrm{span} \bigcup_{W \in \mathcal{U}} W$. The latter space is clearly preserved by every B_i and has nonzero dimension, hence equals V by irreducibility. The claim is proved. We deduce in particular that $r\ell = \dim V$. We have proved every part of the lemma except for the claim that $U_1, \ldots, U_r$ are the *only* subspaces of V with dimension ℓ and with finite orbit under the action of $(B_1, \ldots, B_N)$.

In pursuit of this result we now claim that for every $t = 1, \ldots, r$ there exists a linear map $A_t \in \mathrm{Prox}(V)$ which belongs to the group generated by the maps $B_1, \ldots, B_N$, fixes every $U_1, \ldots, U_r$ and whose leading eigenvalue has an eigenvector in U_t. By hypothesis there exists $\mathtt{k} \in \Sigma_N^*$ such that $B_{\mathtt{k}}$ is proximal. Since $B_{\mathtt{k}}$ induces a permutation of $\mathcal{U}$ it has a power which induces the identity permutation of $\mathcal{U}$ and hence fixes all of the subspaces $U_1, \ldots, U_t$. Without loss of generality we replace $B_{\mathtt{k}}$ with this power, so we have $B_{\mathtt{k}} U_t = U_t$ for every $t = 1, \ldots, r$ and also $B_{\mathtt{k}} \in \mathrm{Prox}(V)$. The determinant of $(B_{\mathtt{k}} - \lambda I)$ on V is equal to the product of the determinants of the linear maps $(B_{\mathtt{k}} - \lambda I)|_{U_t}$, so if λ is the leading eigenvalue of $B_{\mathtt{k}}$ then one of the restrictions $(B_{\mathtt{k}} - \lambda I)|_{U_t}$ must have zero determinant, which is to say there exists an eigenvalue for this eigenvector in the space U_t. Consequently there exists an integer $t_0 \in \{1, \ldots, r\}$ such that we may take $A_{t_0} := B_{\mathtt{k}}$. By the transitivity property demonstrated earlier, for any other $t \in \{1, \ldots, r\}$ we may choose $\mathtt{i} \in \Sigma_N^*$ such that $B_{\mathtt{i}} U_t = U_{t_0}$ and it is clear that the map $A_t := B_{\mathtt{i}}^{-1} B_{\mathtt{k}} B_{\mathtt{i}}$ has the desired properties.

It remains to show that $U_1, \ldots, U_r$ are the *only* subspaces of V with finite orbit under the action of $(B_1, \ldots, B_N)$ and with dimension ℓ. To this end suppose that $\hat{U}$ is an arbitrary ℓ-dimensional subspace of V with finite orbit under the action of $(B_1, \ldots, B_N)$. As with the space U we see easily that every element of the semigroup generated by $B_1, \ldots, B_N$ induces a permutation on the orbit of $\hat{U}$, which is finite. It follows directly that every element of the group generated by $B_1, \ldots, B_N$ also induces a permutation on the orbit of $\hat{U}$. We now claim that $\hat{U}$ must be one of the spaces $U_1, \ldots, U_r$. To see this it is clearly sufficient to show that $\hat{U}$ has nontrivial intersection with one of the spaces U_t, since then $\hat{U} \cap U_t$ is a nonzero subspace of V with dimension at most ℓ and with finite orbit under the action of $(B_1, \ldots, B_N)$; hence it is necessarily ℓ-dimensional and this yields $U_t = \hat{U}$ as required. Let us therefore show that $\hat{U}$ intersects one of the spaces U_t. Since $V = \bigoplus_{t=1}^r U_t$ there exists a subcollection of $U_1, \ldots, U_r$ whose span intersects $\hat{U}$ nontrivially, namely the whole collection. There therefore exists a *minimal* subcollection whose span intersects $\hat{U}$ nontrivially; by relabelling if necessary, call this subcollection $U_1, \ldots, U_m$, say. Since $\hat{U} \cap (U_1 \oplus \cdots \oplus U_m)$ is a nonzero subspace of V with finite orbit under $(B_1, \ldots, B_N)$ it has dimension at least ℓ and therefore equals $\hat{U}$, and we deduce that $\hat{U} \subseteq U_1 \oplus \cdots \oplus U_m$. By minimality the span of the smaller subcollection $U_1, \ldots, U_{m-1}$ must intersect $\hat{U}$ trivially. Therefore $U_1, \ldots, U_{m-1}, \hat{U}$ forms a direct sum which is contained in $U_1 \oplus \cdots \oplus U_m$ and so by dimension considerations must be equal to $U_1 \oplus \cdots \oplus U_m$. We deduce that

$V = U_1 \oplus \cdots \oplus U_{m-1} \oplus \hat{U} \oplus U_{m+1} \oplus \cdots \oplus U_r = U_1 \oplus \cdots \oplus U_r$. Now, the linear map A_m constructed earlier induces a permutation on the orbit of $\hat{U}$ under $(B_1, \ldots, B_N)$ and therefore has a power which induces the identity permutation on that set. Replacing A_m with a power of itself if necessary we may therefore choose $A_m \in \mathrm{Prox}(V)$ so as to fix every space $U_1, \ldots, U_r$ and also so as to fix $\hat{U}$. Now since the leading eigenspace of A_m is one-dimensional and intersects U_m, it does not intersect any of $U_1, \ldots, U_{m-1}, U_{m+1}, \ldots, U_r$. Since A_m preserves every space in the splitting $V = U_1 \oplus \cdots \oplus U_{m-1} \oplus \hat{U} \oplus U_{m+1} \oplus \cdots \oplus U_r$, by the same argument used before its leading eigenspace must intersect one of those spaces; but this space must be $\hat{U}$ since none of the other spaces in the splitting intersects the leading eigenspace. Therefore $\hat{U} \cap U_m \neq \{0\}$ and we deduce that $\hat{U} = U_m$. The proof is complete. $\qquad\square$

We lastly require the following simple combinatorial facts which will be used to construct transitive and primitive subspace classes:

Lemma 5.3. *Let $k \geq 1$ and $N \geq 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space and $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ an irreducible N-tuple. Then:*

(i) *If $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}(V)$ is an equivariant subspace class then it is equal to the disjoint union of finitely many transitive subspace classes.*

(ii) *Suppose that $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}(V)$ is an equivariant subspace class with the following property: for every $(W_j)_{j=1}^k \in \mathcal{W}$ and $n \geq 1$ we have*

$$\left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \text{ and } n \text{ divides } |\mathtt{i}| \right\} = \mathcal{W}.$$

Then $\mathcal{W}$ is primitive.

Proof. It is easy to see that if two transitive subspace classes intersect then they must contain each other and hence be equal, so to establish (i) it is sufficient to show that every $(W_j)_{j=1}^k \in \mathcal{W}$ belongs to a transitive subspace class. Given an arbitrary $(W_j^0)_{j=1}^k \in \mathcal{W}$ define

$$\hat{\mathcal{W}} := \left\{ (A_{\mathtt{i}}^{(j)} W_j^0)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} \subseteq \mathcal{W}.$$

We will show that $\hat{\mathcal{W}}$ is a transitive subspace class and contains $(W_j^0)_{j=1}^k$, which by the arbitrariness of $(W_j^0)_{j=1}^k \in \mathcal{W}$ clearly suffices to prove (i).

To show that $\hat{\mathcal{W}}$ is transitive we must show that for every $(W_j)_{j=1}^k \in \hat{\mathcal{W}}$ we have

$$(9) \qquad\qquad \left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} = \hat{\mathcal{W}}.$$

The left-hand side above is clearly a subset of the right-hand side, so we need only prove the reverse inclusion. Fix an arbitrary $(W_j)_{j=1}^k = (A_{\mathtt{j}}^{(j)} W_j^0)_{j=1}^k \in \hat{\mathcal{W}}$. The map $\mathcal{W} \to \mathcal{W}$ defined by $(W_j')_{j=1}^k \mapsto (A_{\mathtt{j}}^{(j)} W_j')_{j=1}^k$ is clearly injective and acts on a finite set, hence induces a permutation of $\mathcal{W}$, hence there exists an integer $p \geq 1$ such that the pth power of this map is the identity permutation of $\mathcal{W}$. We therefore

have $(W_j^0)_{j=1}^k = (A_{\mathtt{j}^p}^{(j)} W_j^0)_{j=1}^k = (A_{\mathtt{j}^{(p-1)}}^{(j)} W_j)_{j=1}^k \in \hat{\mathcal{W}}$. Hence

$$\hat{\mathcal{W}} = \left\{ (A_{\mathtt{i}}^{(j)} W_j^0)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\}$$
$$= \left\{ (A_{\mathtt{i}\mathtt{j}^{p-1}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} \subseteq \left\{ (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \right\} \subseteq \hat{\mathcal{W}}.$$

We have established (9) which proves the transitivity of $\hat{\mathcal{W}}$ and the result (i) follows.

Let us now prove (ii). If $\mathtt{i} \in \Sigma_N^*$ is arbitrary then it induces a permutation of $\mathcal{W}$ via the map $(W_j)_{j=1}^k \mapsto (A_{\mathtt{i}}^{(j)} W_j)_{j=1}^k$ and in particular some power of $\mathtt{i}$ induces the identity permutation. We may therefore choose $\mathtt{j} \in \Sigma_N^*$ such that $(A_{\mathtt{j}}^{(j)} W_j)_{j=1}^k = (W_j)_{j=1}^k$ for all $(W_j)_{j=1}^k \in \mathcal{W}$. Let $\ell := |\mathtt{j}|$. By hypothesis for every $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}$ there exists $\mathtt{i} \in \Sigma_N^*$ such that $(W_j)_{j=1}^k = (A_{\mathtt{i}}^{(j)} W_j')_{j=1}^k$ and ℓ divides $|\mathtt{i}|$. Since $\mathcal{W}$ is finite it follows directly that there exists an integer t such that for every $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}$ we may choose $\mathtt{i} \in \Sigma_N^*$ satisfying $(W_j)_{j=1}^k = (A_{\mathtt{i}}^{(j)} W_j')_{j=1}^k$, such that ℓ divides $|\mathtt{i}|$, and such that $|\mathtt{i}| = s\ell$ for some integer s such that $1 \leq s < t$; we then have $(W_j)_{j=1}^k = (A_{\mathtt{j}^{t-s}\mathtt{i}}^{(j)} W_j')_{j=1}^k$ and $|\mathtt{j}^{t-s}\mathtt{i}| = t\ell$. Since the length of the word $\mathtt{j}^{t-s}\mathtt{i}$ is independent of the choice of $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}$ we have proved that $\mathcal{W}$ is primitive. $\square$

Proof of Theorem 6. By Lemma 5.1 there exists $\mathtt{k} \in \Sigma_N^*$ such that for every $j = 1, \ldots, k$ we have $A_{\mathtt{k}}^{(j)} \in \mathrm{Prox}(V_j)$, so the hypotheses of Lemma 5.2 are met by $(A_1^{(j)}, \ldots, A_N^{(j)})$ and V_j for every $j = 1, \ldots, k$. Hence for each $j = 1, \ldots, k$ there exist an integer $r_j \geq 1$ and a splitting $V_j = \bigoplus_{i=1}^{r_j} U_j^i$ such that $A_{\mathtt{i}}^{(j)} U_j^i \in \{U_j^1, \ldots, U_j^{r_j}\}$ for every $\mathtt{i} \in \Sigma_N^*$ and $i \in \{1, \ldots, r_j\}$. Without loss of generality we adopt an inner product structure on each V_j such that the splitting $V_j = \bigoplus_{i=1}^{r_j} U_j^i$ is orthogonal.

Fix $\mathtt{i} \in \Sigma_N^*$ and $j = 1, \ldots, k$. The linear map $A_{\mathtt{i}}^{(j)}$ induces a permutation of the set $\{U_j^1, \ldots, U_j^{r_j}\}$. We assert that $(A_{\mathtt{i}}^{(j)})^\top$ also permutes this set and moreover induces the inverse permutation on it. Indeed, if $A_{\mathtt{i}}^{(j)} U_j^{i_1} = U_j^{i_2}$ then we have

$$(10) \qquad A_{\mathtt{i}}^{(j)} \left((U_j^{i_1})^\perp \right) = A_{\mathtt{i}}^{(j)} \left(\bigoplus_{\substack{1 \leq i \leq r_j \\ i \neq i_1}} U_j^i \right) = \bigoplus_{\substack{1 \leq i \leq r_j \\ i \neq i_2}} U_j^i = (U_j^{i_2})^\perp.$$

A standard and easy calculation shows that if $B \in \mathrm{GL}(V)$ and $U, U' \in \mathrm{Gr}(V)$ then $BU = U'$ if and only if $B^\top (U')^\perp = U^\perp$, so (10) yields $(A_{\mathtt{i}}^{(j)})^\top U_j^{i_2} = U_j^{i_1}$. This demonstrates that $(A_{\mathtt{i}}^{(j)})^\top$ induces the inverse permutation on $\{U_j^1, \ldots, U_j^{r_j}\}$ to that induced by $A_{\mathtt{i}}^{(j)}$. It follows that $(A_{\mathtt{i}}^{(j)})^\top A_{\mathtt{i}}^{(j)}$ preserves U_j^i for every $i = 1, \ldots, r_j$. Since $V_j = \bigoplus_{i=1}^{r_j} U_j^i$ it follows that every eigenvalue of $(A_{\mathtt{i}}^{(j)})^\top A_{\mathtt{i}}^{(j)}$ is realised as an eigenvalue of its restriction to some U_j^i. In particular the numbers $\|A_{\mathtt{i}}^{(j)}|_{U_j^i}\| = \rho((A_{\mathtt{i}}^{(j)})^\top A_{\mathtt{i}}^{(j)}|_{U_j^i})^{1/2}$ for $i = 1, \ldots, r_j$ form a subset of the singular values of $A_{\mathtt{i}}^{(j)}$. We conclude that for every $\mathtt{i} \in \Sigma_N^*$ and $j = 1, \ldots, k$ we have

$$(11) \qquad \left\| A_{\mathtt{i}}^{(j)} \right\| = \max_{1 \leq i \leq r_j} \left\| A_{\mathtt{i}}^{(j)}|_{U_j^i} \right\|$$

and moreover for each $i = 1, \ldots, r_j$

$$\tag{12} \max\left\{ \left\| A_{\mathtt{i}}^{(j)}|_{U_j^i} \right\|, \left\| A_{\mathtt{i}}^{(j)}|_{(U_j^i)^\perp} \right\| \right\} = \left\| A_{\mathtt{i}}^{(j)} \right\|,$$

$$\tag{13} \min\left\{ \left\| A_{\mathtt{i}}^{(j)}|_{U_j^i} \right\|, \left\| A_{\mathtt{i}}^{(j)}|_{(U_j^i)^\perp} \right\| \right\} \leq \sigma_2\left(A_{\mathtt{i}}^{(j)} \right).$$

For the remainder of the proof let $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be an ergodic equilibrium state of Φ and let $\hat{\mu} \in \mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)$ be its natural extension. A large initial part of the proof will be taken up by the following construction: we claim that for every $j = 1, \ldots, k$ there exists a Borel measurable function $\mathsf{U}_j \colon \hat{\Sigma}_N \to \{U_j^1, \ldots, U_j^{r_j}\}$ such that for $\hat{\mu}$-a.e $x \in \hat{\Sigma}_N$ and every $m \geq 1$,

$$\tag{14} \lim_{n\to\infty} \frac{1}{n} \log \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{\mathsf{U}_j(x)} \right\| = \lambda_1(A^{(j)}, \mu),$$

$$\tag{15} \limsup_{n\to\infty} \frac{1}{n} \log \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{\mathsf{U}_j(x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu)$$

and

$$\tag{16} A_{(\hat{\sigma}^{-m}x)|_m}^{(j)} \mathsf{U}_j(x) = \mathsf{U}_j(\hat{\sigma}^{-m}x).$$

To begin this construction fix $j \in \{1, \ldots, k\}$, define Z to be the set

$$\bigcap_{\ell=1}^{\dim V_j} \left\{ x \in \hat{\Sigma}_N \colon \lim_{n\to\infty} \frac{1}{n} \log \sigma_\ell\left(A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} \right) = \lambda_\ell(A^{(j)}, \mu) \right\}$$

and for each $i = 1, \ldots, r_j$ define

$$X_i := \left\{ x \in \hat{\Sigma}_N \colon \lim_{n\to\infty} \frac{1}{n} \log \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{U_j^i} \right\| = \lambda_1(A^{(j)}, \mu) \right\}.$$

Clearly all of these sets are Borel sets. For each $\ell = 1, \ldots, \dim V_j$ the functions $f_n^\ell \colon \hat{\Sigma}_N \to \mathbb{R}$ defined by

$$f_n^\ell(x) := \log \prod_{i=1}^{\ell} \sigma_i\left(A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} \right)$$

satisfy $f_{n+m}^\ell(x) \leq f_n^\ell(\hat{\sigma}^{-m}x) + f_m^\ell(x)$ for all $x \in \hat{\Sigma}_N$ and $n, m \geq 1$, and since clearly

$$Z = \bigcap_{\ell=1}^{\dim V_j} \left\{ x \in \hat{\Sigma}_N \colon \lim_{n\to\infty} \frac{1}{n} \log \prod_{i=1}^{\ell} \sigma_\ell\left(A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} \right) = \sum_{i=1}^{\ell} \lambda_i(A^{(j)}, \mu) \right\}$$

it follows by the subadditive ergodic theorem applied to the functions f_n^ℓ and the transformation $\hat{\sigma}^{-1}$ that $\mu(Z) = 1$.

We claim that the sets $Z \cap X_i$ are pairwise disjoint for distinct values of i. Indeed, if $x \in Z \cap X_{i_1} \cap X_{i_2}$ with $i_1 \neq i_2$ then since $U_j^{i_2} \subseteq (U_j^{i_1})^\perp$,

$$\lambda_1(A^{(j)}, \mu) = \lim_{n\to\infty} \frac{1}{n} \log \min\left\{ \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{U_j^{i_1}} \right\|, \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{U_j^{i_2}} \right\| \right\}$$

$$\leq \limsup_{n\to\infty} \frac{1}{n} \log \min\left\{ \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{U_j^{i_1}} \right\|, \left\| A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} |_{(U_j^{i_1})^\perp} \right\| \right\}$$

$$\leq \lim_{n\to\infty} \frac{1}{n} \log \sigma_2\left(A_{(\hat{\sigma}^{-n}x)|_n}^{(j)} \right) = \lambda_2(A^{(j)}, \mu) < \lambda_1(A^{(j)}, \mu)$$

where in the first line we have used the definition of X_{i_1} and X_{i_2}, and in the third line we have used (13), the definition of Z and the hypothesis that the top Lyapunov exponent of each $(A_1^{(j)}, \ldots, A_N^{(j)})$ is simple. But the above chain of inequalities ends in a contradiction, and we conclude that $Z \cap X_{i_1} \cap X_{i_2}$ must be empty. The sets $Z \cap X_i$ are thus pairwise disjoint for distinct values of i as claimed.

We now claim that $Z \cap (X_1 \cup \cdots \cup X_{r_j})$ has full measure. Define

$$C := \max_{1 \le i \le N} \max \left\{ \left\| \left(A_i^{(j)} \right)^{-1} \right\|, \left\| A_i^{(j)} \right\| \right\} \ge 1.$$

If $x \in Z$ then there exists n_0 such that for all $n \ge n_0$

$$\left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right\| = \sigma_1 \left(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right) > C^2 \sigma_2 \left(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right)$$

by the definition of Z and the hypothesis that the top Lyapunov exponent of each $(A_1^{(j)}, \ldots, A_n^{(j)})$ is simple. If $x \in Z$ is fixed then for all $n \ge n_0$ we in particular have $\sigma_1(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)}) > \sigma_2(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)})$ and in view of (12) and (13) there exists a unique $i(x, n)$ such that $\|A_{(\hat{\sigma}^{-n} x)|_n}^{(j)}\| = \|A_{(\hat{\sigma}^{-n} x)|_n}^{(j)}|_{U_j^{i(x,n)}}\|$. We assert that $i(x, n)$ does not depend on the value of $n \ge n_0$. Given arbitrary $n \ge n_0$ we have

$$\left\| A_{(\hat{\sigma}^{-n-1} x)|_{n+1}}^{(j)} |_{U_j^{i(x,n)}} \right\| = \left\| A_{x_{-n}}^{(j)} A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{U_j^{i(x,n)}} \right\|$$

$$\ge C^{-1} \left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{U_j^{i(x,n)}} \right\|$$

$$= C^{-1} \sigma_1 \left(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right)$$

$$> C \sigma_2 \left(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right)$$

$$\ge C \left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{(U_j^{i(x,n)})^\perp} \right\|$$

$$= C \max_{i \ne i(x,n)} \left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{U_j^i} \right\|$$

$$\ge \max_{i \ne i(x,n)} \left\| A_{x_{-n}}^{(j)} A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{U_j^i} \right\| = \max_{i \ne i(x,n)} \left\| A_{(\hat{\sigma}^{-n-1} x)|_{n+1}}^{(j)} |_{U_j^i} \right\|$$

and in view of (11) it follows that $i(x, n)$ satisfies the definition of $i(x, n+1)$. We conclude that $i(x, n+1) = i(x, n)$ for all $n \ge n_0$ and therefore $i(x, n)$ in fact takes a constant value $i(x)$ for all $n \ge n_0$. We then have

$$(17) \qquad \lambda_1(A^{(j)}, \mu) = \lim_{n \to \infty} \frac{1}{n} \log \sigma_1 \left(A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right)$$

$$= \lim_{n \to \infty} \frac{1}{n} \log \left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} \right\| = \lim_{n \to \infty} \frac{1}{n} \log \left\| A_{(\hat{\sigma}^{-n} x)|_n}^{(j)} |_{U_j^{i(x)}} \right\|$$

so that $x \in X_{i(x)}$. Combining this with the preceding result we have shown that every element of Z belongs to a unique X_i. Since Z has full measure and the sets X_i are Borel measurable we conclude that the sets $X_1, \ldots, X_{r_j}$ partition $\hat{\Sigma}_N$ up to $\hat{\mu}$-measure zero. If we define a function $\mathsf{U}_j \colon \hat{\Sigma}_N \to \{U_j^1, \ldots, U_j^{r_j}\}$ by $\mathsf{U}_j(x) := U_j^i$ if $x \in Z \cap X_i$ and $\mathsf{U}_j(x) := U_j^1$ otherwise, it is obvious that U_j is a measurable function. The equation (17) precisely asserts the truth of (14) and combining this with (12), (13) and the definition of Z yields (15). In particular (14) and (15) hold for all $x \in Z \cap (X_1 \cup \cdots \cup X_{r_j})$.

It remains to prove (16). Suppose that x belongs to the set $\bigcap_{\ell \in \mathbb{Z}} \hat{\sigma}^\ell (Z \cap (X_1 \cup \cdots \cup X_{r_j}))$, which clearly has full measure. It is easy to see that for every fixed $m \geq 1$ we have

$$\lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_n} \big|_{A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} \mathsf{U}_j(x)} \right\|$$

$$= \lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_n} A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} \big|_{\mathsf{U}_j(x)} \right\|$$

$$= \lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_{n+m}} \big|_{\mathsf{U}_j(x)} \right\| = \lambda_1(A^{(j)}, \mu)$$

and

$$\limsup_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_n} \big|_{\left(A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} \mathsf{U}_j(x) \right)^\perp} \right\|$$

$$= \limsup_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_n} \big|_{A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} (\mathsf{U}_j(x)^\perp)} \right\|$$

$$= \limsup_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_n} A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} \big|_{\mathsf{U}_j(x)^\perp} \right\|$$

$$= \limsup_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{(\hat{\sigma}^{-n-m}x)|_{n+m}} \big|_{\mathsf{U}_j(x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu).$$

Since $\hat{\sigma}^{-m}x \in Z \cap (X_1 \cup \cdots \cup X_{r_j})$ both (14) and (15) apply with $\hat{\sigma}^{-m}x$ in place of x and in view of the above we can only have $A^{(j)}_{(\hat{\sigma}^{-m}x)|_m} \mathsf{U}_j(x) = \mathsf{U}_j(\hat{\sigma}^{-m}x)$, which is (16). This completes the construction of the measurable functions $\mathsf{U}_j \colon \hat{\Sigma}_N \to \{U_j^1, \ldots, U_j^{r_j}\}$.

We next claim that for fixed $j \in \{1, \ldots, k\}$, for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have

$$(18) \qquad\qquad \lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{x|_n} \big|_{\mathsf{U}_j(\hat{\sigma}^n x)} \right\| = \lambda_1(A^{(j)}, \mu),$$

$$(19) \qquad\qquad \lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{x|_n} \big|_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu).$$

Define two sequences of bounded measurable functions $f_n, g_n \colon \hat{\Sigma}_N \to \mathbb{R}$ by $f_n(x) := \log \| A^{(j)}_{x|_n} |_{\mathsf{U}_j(\hat{\sigma}^n x)} \|$ and $g_n(x) := \log \| A^{(j)}_{x|_n} |_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \|$. For all $n, m \geq 1$ we have $f_{n+m}(x) \leq f_n(\hat{\sigma}^m x) + f_m(x)$ and $g_{n+m}(x) \leq g_n(\hat{\sigma}^m x) + g_m(x)$ for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$, where we have used (16). It follows by the subadditive ergodic theorem that the limits in (18) and (19) exist $\hat{\mu}$-a.e. and are constant $\hat{\mu}$-a.e, so we need only show that these almost sure values are equal to $\lambda_1(A^{(j)}, \mu)$ and bounded by $\lambda_2(A^{(j)}, \mu)$ respectively. But from (14) and (15) we have

$$\lim_{n \to \infty} \hat{\mu} \left(\left\{ x \in \hat{\Sigma}_N \colon \left| \frac{1}{n} \log \left\| A^{(j)}_{x|_n} \big|_{\mathsf{U}_j(\hat{\sigma}^n x)} \right\| - \lambda_1(A^{(j)}, \mu) \right| < \varepsilon \right\} \right) = 1$$

and

$$\lim_{n \to \infty} \hat{\mu} \left(\left\{ x \in \hat{\Sigma}_N \colon \frac{1}{n} \log \left\| A^{(j)}_{x|_n} \big|_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu) + \varepsilon \right\} \right) = 1$$

for all $\varepsilon > 0$ since convergence almost everywhere implies convergence in measure and since $\hat{\mu}$ is invariant with respect to $\hat{\sigma}$. This is only possible if the almost sure limit in (18) is equal to $\lambda_1(A^{(j)}, \mu)$ and the almost sure limit in (19) is bounded by $\lambda_2(A^{(j)}, \mu)$. We have proved the claim.

Let $\mathcal{U}$ denote the set of all k-tuples of the form $(U_j^{i_j})_{j=1}^k$ where $1 \le i_j \le r_j$ for every $j = 1, \ldots, k$. It is clear that $\mathcal{U} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ is an equivariant subspace class, so by Lemma 5.3(i) it is equal to the disjoint union of finitely many transitive subspace classes $\mathcal{W}_1, \ldots, \mathcal{W}_r$, say. Consider the sets $Y_t := \{x \in \hat{\Sigma}_N \colon (\mathsf{U}_j(x))_{j=1}^k \in \mathcal{W}_t\}$ for $t = 1, \ldots, r$, which clearly form a measurable partition of $\hat{\Sigma}_N$. For $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have $(\mathsf{U}_j(x))_{j=1}^k \in \mathcal{W}_t$ if and only if $(\mathsf{U}_j(\hat{\sigma}x))_{j=1}^k \in \mathcal{W}_t$ as a consequence of (16), so each Y_t is invariant under $\hat{\sigma}$ up to $\hat{\mu}$-measure zero. By the ergodicity of $\hat{\mu}$ it follows that there exists $t_0 \in \{1, \ldots, r\}$ such that $\hat{\mu}(Y_{t_0}) = 1$ and $\hat{\mu}(Y_t) = 0$ for all other t.

We may now prove (i). If $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ is any transitive subspace class then we define a potential $\Phi_{\mathcal{W}} \colon \Sigma_N^* \to (0, +\infty)$ by

$$\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^k \|A_{\mathtt{i}}^{(j)}|_{W_j}\|^{\beta_j}.$$

Since μ is an equilibrium state of Φ we have

$$P(\Phi) = h(\mu) + \Lambda(\Phi, \mu) = h(\mu) + \sum_{j=1}^k \beta_j \lambda_1(A^{(j)}, \mu).$$

Since clearly $\Phi_{\mathcal{W}} \le \Phi$ for every transitive subspace class $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ we have $P(\Phi_{\mathcal{W}}) \le P(\Phi)$ for all such classes, so if $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ is a transitive subspace class then μ is an equilibrium state of $\Phi_{\mathcal{W}}$ if and only if $\Lambda(\Phi_{\mathcal{W}}, \mu) = \Lambda(\Phi, \mu)$. To prove (i) we will show that this is the case precisely when $\mathcal{W} = \mathcal{W}_{t_0}$.

Let us first show that μ is an equilibrium state for $\Phi_{\mathcal{W}_{t_0}}$. We clearly have

$$\sum_{j=1}^k \beta_j \lambda_1(A^{(j)}, \mu) = \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^k \beta_j \log \left\| A_{x|_n}^{(j)}|_{\mathsf{U}_j(\hat{\sigma}^n x)} \right\|$$

$$= \lim_{n \to \infty} \frac{1}{n} \log \prod_{j=1}^k \left\| A_{x|_n}^{(j)}|_{\mathsf{U}_j(\hat{\sigma}^n x)} \right\|^{\beta_j}$$

$$\le \lim_{n \to \infty} \frac{1}{n} \log \max_{(W_j)_{j=1}^k \in \mathcal{W}_{t_0}} \prod_{j=1}^k \left\| A_{x|_n}^{(j)}|_{W_j} \right\|^{\beta_j}$$

$$= \lim_{n \to \infty} \frac{1}{n} \log \Phi_{\mathcal{W}_{t_0}}(x|_n)$$

$$= \Lambda(\Phi_{\mathcal{W}_{t_0}}, \mu) \le \Lambda(\Phi, \mu) = \sum_{j=1}^k \beta_j \lambda_1(A^{(j)}, \mu)$$

for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$, where we have used (18) together with the subadditive ergodic theorem applied to $\Phi_{\mathcal{W}_{t_0}}$ and Φ. It follows that $\Lambda(\Phi_{\mathcal{W}_{t_0}}, \mu) = \Lambda(\Phi, \mu)$ as required for μ to be an equilibrium state of $\Phi_{\mathcal{W}_{t_0}}$. This proves the existence part of (i).

Now suppose instead that $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ is a transitive subspace class such that μ is an equilibrium state of the potential $\Phi_{\mathcal{W}}$. If $(W_j)_{j=1}^k \in \mathcal{W}$ then by definition each W_j has finite orbit under the action of $(A_1^{(j)}, \ldots, A_N^{(j)})$ and therefore must be one of the spaces $U_j^1, \ldots, U_j^{r_j}$, since Lemma 5.2 implies that those are the only ℓ_j-dimensional subspaces of V_j with finite orbit. Thus necessarily $\mathcal{W} \subseteq \mathcal{U}$. It

follows that $\mathcal{W}$ intersects at least one $\mathcal{W}_t$ and since both are transitive we must have $\mathcal{W} = \mathcal{W}_t$. To complete the uniqueness part of the proof of (i) we will show that if $t \neq t_0$ then a contradiction occurs.

Fix an arbitrary $t \neq t_0$. For almost every $x \in \hat{\Sigma}_N$ we have $(\mathsf{U}_j(\hat{\sigma}^n x))_{j=1}^k \in \mathcal{W}_{t_0}$ for all $n \geq 1$. For such x and n, if $(W_j)_{j=1}^k \in \mathcal{W}_t$ is arbitrary then $(W_j)_{j=1}^k \notin \mathcal{W}_{t_0}$ and in particular $(W_j)_{j=1}^k \neq (\mathsf{U}_j(\hat{\sigma}^n x))_{j=1}^k$, so there exists j_0 depending on x, n and $(W_j)_{j=1}^k$ such that $W_{j_0} \neq \mathsf{U}_{j_0}(\hat{\sigma}^n x)$ and therefore $W_{j_0} \subseteq \mathsf{U}_{j_0}(\hat{\sigma}^n x)^\perp$. Consequently

$$\prod_{j=1}^k \left\| A_{x|_n}^{(j)}|_{W_j} \right\|^{\beta_j} \leq \left(\prod_{\substack{1 \leq j \leq k \\ j \neq j_0}} \left\| A_{x|_n}^{(j)} \right\|^{\beta_j} \right) \left\| A_{x|_n}^{(j_0)}|_{\mathsf{U}_{j_0}(\hat{\sigma}^n x)^\perp} \right\|^{\beta_{j_0}}$$

for this particular x, n and $(W_j)_{j=1}^k$, and therefore

$$\max_{(W_j)_{j=1}^k \in \mathcal{W}_t} \prod_{j=1}^k \left\| A_{x|_n}^{(j)}|_{W_j} \right\|^{\beta_j} \leq \max_{1 \leq \ell \leq k} \left[\left(\prod_{\substack{1 \leq j \leq k \\ j \neq \ell}} \left\| A_{x|_n}^{(j)} \right\|^{\beta_j} \right) \left\| A_{x|_n}^{(\ell)}|_{\mathsf{U}_\ell(\hat{\sigma}^n x)^\perp} \right\|^{\beta_\ell} \right]$$

for this particular x and n. Thus for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have

$$\Phi_{\mathcal{W}_t}(x|_n) \leq \max_{1 \leq \ell \leq k} \left[\left(\prod_{\substack{1 \leq j \leq k \\ j \neq \ell}} \left\| A_{x|_n}^{(j)} \right\|^{\beta_j} \right) \left\| A_{x|_n}^{(\ell)}|_{\mathsf{U}_\ell(\hat{\sigma}^n x)^\perp} \right\|^{\beta_\ell} \right]$$

for all $n \geq 1$. But for all $j = 1, \ldots, k$ by the subadditive ergodic theorem we have for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$

$$\lim_{n \to \infty} \frac{1}{n} \log \left(\left\| A_{x|_n}^{(j)} \right\|^{\beta_j} \right) = \beta_j \lambda_1(A^{(j)}, \mu)$$

and by (19)

$$\lim_{n \to \infty} \frac{1}{n} \log \left(\left\| A_{x|_n}^{(j)}|_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\|^{\beta_j} \right) \leq \beta_j \lambda_2(A^{(j)}, \mu)$$

from which it follows that for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$

$$\Lambda(\Phi_{\mathcal{W}_t}, \mu) = \lim_{n \to \infty} \frac{1}{n} \log \Phi_{\mathcal{W}_t}(x|_n)$$

$$\leq \max_{1 \leq \ell \leq k} \left[\left(\sum_{\substack{1 \leq j \leq k \\ j \neq \ell}} \beta_j \lambda_1(A^{(j)}, \mu) \right) + \beta_\ell \lambda_2(A^{(\ell)}, \mu) \right]$$

$$< \sum_{j=1}^k \beta_j \lambda_1(A^{(j)}, \mu) = \Lambda(\Phi, \mu).$$

Thus if $t \neq t_0$ we have $\Lambda(\Phi_{\mathcal{W}_t}, \mu) < \Lambda(\Phi, \mu)$ and we have completed the proof of (i).

The proof of (ii) is now straightforward. By Lemma 5.1, for $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have

$$\limsup_{n \to \infty} \min_{1 \leq j \leq k} \left(\frac{1}{n} \log \rho \left(A^{(j)}_{x|_n} \right) - \frac{1}{2n} \log \sigma_1 \left(A^{(j)}_{x|_n} \right) \sigma_2 \left(A^{(j)}_{x|_n} \right) \right) > 0,$$

and for all $j = 1, \ldots, k$ and $\hat{\mu}$-a.e. $x \in \hat{\Sigma}_N$ we have by (19)

$$\lim_{n \to \infty} \frac{1}{n} \log \left\| A^{(j)}_{x|_n}|_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\| \leq \lambda_2(A^{(j)}, \mu)$$

and by the subadditive ergodic theorem

$$\lim_{n \to \infty} \frac{1}{2n} \log \sigma_1 \left(A^{(j)}_{x|_n} \right) \sigma_2 \left(A^{(j)}_{x|_n} \right) = \frac{1}{2}\lambda_1(A^{(j)}, \mu) + \frac{1}{2}\lambda_2(A^{(j)}, \mu) > \lambda_2(A^{(j)}, \mu).$$

It follows that there exist $x \in \hat{\Sigma}_N$ and $n \geq 1$ such that $(\mathsf{U}_j(\hat{\sigma}^n x))_{j=1}^k \in \mathcal{W}_{t_0}$ and such that for all $j = 1, \ldots, k$

(20) $$\rho \left(A^{(j)}_{x|_n} \right) > \sigma_1 \left(A^{(j)}_{x|_n} \right)^{\frac{1}{2}} \sigma_2 \left(A^{(j)}_{x|_n} \right)^{\frac{1}{2}} > \left\| A^{(j)}_{x|_n}|_{\mathsf{U}_j(\hat{\sigma}^n x)^\perp} \right\|.$$

Let $\mathbf{j} := x|_n$ and $(W_j^0)_{j=1}^k := (\mathsf{U}_j(\hat{\sigma}^n x))_{j=1}^k \in \mathcal{W}_{t_0}$. Since $\rho(A^{(j)}_{\mathbf{j}})^2 > \sigma_1(A^{(j)}_{\mathbf{j}})\sigma_2(A^{(j)}_{\mathbf{j}})$ for every $j = 1, \ldots, k$ we have $A^{(j)}_{\mathbf{j}} \in \operatorname{Prox}(V_j)$ for every $j = 1, \ldots, k$. The map $(W_j)_{j=1}^k \mapsto (A^{(j)}_{\mathbf{j}} W_j)_{j=1}^k$ induces a permutation of $\mathcal{W}_{t_0}$ and therefore some power of this map induces the identity permutation, so let $\mathbf{k} := \mathbf{j}^p$ be a power of p such that $A^{(j)}_{\mathbf{k}} W_j^0 = W_j^0$ for every $j = 1, \ldots, k$. For every $j = 1, \ldots, k$ we clearly have $A^{(j)}_{\mathbf{k}} \in \operatorname{Prox}(V_j)$, and since $A^{(j)}_{\mathbf{k}}$ preserves the splitting $V_j = U_j^1 \oplus \cdots \oplus U_j^{r_j}$ its leading eigenvector belongs to one of these spaces. In particular this eigenvector belongs either to W_j^0 or to $(W_j^0)^\perp$. If it belongs to the latter then the leading eigenvector of $A^{(j)}_{\mathbf{j}}$, being the same vector, belongs to $(W_j^0)^\perp$ and therefore we have $\rho(A^{(j)}_{\mathbf{j}}) \leq \|A^{(j)}_{\mathbf{j}}|_{(W_j^0)^\perp}\|$, but this contradicts (20). Hence the leading eigenvector of $A^{(j)}_{\mathbf{k}}$ belongs to W_j^0 for every $j = 1, \ldots, k$. For each $j = 1, \ldots, k$ the linear map $A^{(j)}_{\mathbf{k}}$ fixes W_j^0 and its eigenvalues when restricted to W_j^0 are a subset of its eigenvalues on V_j which includes the largest eigenvalue. Thus $A^{(j)}_{\mathbf{k}}|_{W_j^0} \in \operatorname{Prox}(W_j^0)$ for every $j = 1, \ldots, k$ and we have proved (ii).

It remains to prove (iii). We first claim that for every $(W_j)_{j=1}^k \in \mathcal{W}_{t_0}$ we have

$$\hat{\mu} \left(\{x \in \hat{\Sigma}_N \colon (\mathsf{U}_j(x))_{j=1}^k = (W_j)_{j=1}^k \} \right) > 0.$$

To see this we argue as follows. Each $\mathbf{i} \in \Sigma_N^*$ induces a permutation of $\mathcal{W}_{t_0}$ via the map $(W_j)_{j=1}^k \mapsto (A^{(j)}_{\mathbf{i}} W_j)_{j=1}^k$, and some power of the same word consequently induces the inverse permutation. The set of all permutations of $\mathcal{W}_{t_0}$ which can be realised in this way thus forms a finite group. We may therefore find a finite list of words $\mathbf{i}_1, \ldots, \mathbf{i}_p \in \Sigma_N^*$ such that every permutation of $\mathcal{W}_{t_0}$ of the form $(W_j)_{j=1}^k \mapsto (A^{(j)}_{\mathbf{i}} W_j)_{j=1}^k$ is realised by $(W_j)_{j=1}^k \mapsto (A^{(j)}_{\mathbf{i}_1 \mathbf{i}_2 \cdots \mathbf{i}_q} W_j)_{j=1}^k$ for some $q \in \{1, \ldots, p\}$. (For example, we could take p to be an even number such that the words $\mathbf{i}_1, \mathbf{i}_3, \mathbf{i}_5, \ldots, \mathbf{i}_{p-1}$ induce all of the various permutations and such that $\mathbf{i}_{q+1}$ induces the inverse of the permutation induced by $\mathbf{i}_q$ for each odd number q.) The list $\mathbf{i}_1, \ldots, \mathbf{i}_p \in \Sigma_N^*$ has the property that if $(W_j)_{j=1}^k \in \mathcal{W}_{t_0}$ is arbitrary, then

(21) $$\left\{ ((A^{(j)}_{\mathbf{i}_1 \cdots \mathbf{i}_q})^{-1} W_j)_{j=1}^k \colon 1 \leq q \leq p \right\} = \mathcal{W}_{t_0}.$$

Now by Theorem 3 the measure μ is fully supported on Σ_N, so $\hat{\mu}([\mathtt{i}]) > 0$ for every $\mathtt{i} \in \Sigma_N^*$ and in particular $\hat{\mu}([\mathtt{i}_1 \cdots \mathtt{i}_p]) > 0$. Since the set of all $x \in \hat{\Sigma}_N$ such that $(\mathsf{U}_j(x))_{j=1}^k \in \mathcal{W}_{t_0}$ has full measure, we may choose $(W_j)_{j=1}^k \in \mathcal{W}_{t_0}$ such that

$$\hat{\mu}\left(\left\{x \in \hat{\Sigma}_N \colon (\mathsf{U}_j(x))_{j=1}^k = (W_j)_{j=1}^k \text{ and } x \in [\mathtt{i}_1 \cdots \mathtt{i}_p]\right\}\right) > 0.$$

For every $q = 1, \ldots, p$ we have

$$\hat{\mu}\left(\left\{x \in \hat{\Sigma}_N \colon (\mathsf{U}_j(\hat{\sigma}^{-|\mathtt{i}_1 \cdots \mathtt{i}_q|} x))_{j=1}^k = (W_j)_{j=1}^k \text{ and } \hat{\sigma}^{-|\mathtt{i}_1 \cdots \mathtt{i}_q|} x \in [\mathtt{i}_1 \cdots \mathtt{i}_p]\right\}\right) > 0$$

by the $\hat{\sigma}$-invariance of $\hat{\mu}$. Now using (16) we have

$$\left(\mathsf{U}_j(\hat{\sigma}^{-|\mathtt{i}_1 \cdots \mathtt{i}_q|} x)\right)_{j=1}^k = \left(A^{(j)}_{(\hat{\sigma}^{-|\mathtt{i}_1 \cdots \mathtt{i}_q|} x)|_{|\mathtt{i}_1 \cdots \mathtt{i}_q|}} \mathsf{U}_j(x)\right)_{j=1}^k = \left(A^{(j)}_{\mathtt{i}_1 \cdots \mathtt{i}_q} \mathsf{U}(x)\right)_{j=1}^k$$

for $\hat{\mu}$-a.e. $x \in \hat{\sigma}^{|\mathtt{i}_1 \cdots \mathtt{i}_q|}[\mathtt{i}_1 \cdots \mathtt{i}_p]$, so

$$\hat{\mu}\left(\left\{x \in \hat{\Sigma}_N \colon (A^{(j)}_{\mathtt{i}_1 \cdots \mathtt{i}_q} \mathsf{U}(x))_{j=1}^k = (W_j)_{j=1}^k\right\}\right) > 0$$

and therefore

$$\hat{\mu}\left(\left\{x \in \hat{\Sigma}_N \colon (\mathsf{U}(x))_{j=1}^k = ((A^{(j)}_{\mathtt{i}_1 \cdots \mathtt{i}_q})^{-1} W_j)_{j=1}^k\right\}\right) > 0$$

for every $q = 1, \ldots, p$. In view of (21) this proves the claim.

We may now prove (iii). We will show that the hypothesis of Lemma 5.3(ii) is satisfied by $\mathcal{W}_{t_0}$. Fix arbitrary $n \geq 1$ and suppose that μ, and hence $\hat{\mu}$, is totally ergodic. Let $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}_{t_0}$ be arbitrary. Since $\hat{\mu}$ is ergodic with respect to $\hat{\sigma}^n$, for every two sets $Z_1, Z_2 \subseteq \hat{\Sigma}_N$ such that $\hat{\mu}(Z_1), \hat{\mu}(Z_2) > 0$ we may find $\ell \geq 1$ such that $\hat{\mu}(Z_1 \cap \hat{\sigma}^{n\ell} Z_2) > 0$. In particular using the claim just proved it follows that there exists $\ell \geq 1$ such that

$$\hat{\mu}\left(\{x \in \hat{\Sigma}_N \colon (\mathsf{U}_j(x))_{j=1}^k = (W_j)_{j=1}^k \text{ and } (\mathsf{U}_j(\hat{\sigma}^{-n\ell} x))_{j=1}^k = (W_j')_{j=1}^k\}\right) > 0.$$

Hence for a positive-measure set of $x \in \hat{\Sigma}_N$ we have

$$(W_j')_{j=1}^k = (\mathsf{U}_j(\hat{\sigma}^{-n\ell} x))_{j=1}^k = (A^{(j)}_{(\hat{\sigma}^{-n\ell} x)|_{n\ell}} \mathsf{U}_j(x))_{j=1}^k = (A^{(j)}_{(\hat{\sigma}^{-n\ell} x)|_{n\ell}} W_j)_{j=1}^k$$

using (16). In particular $(W_j')_{j=1}^k = (A^{(j)}_{\mathtt{i}} W_j)_{j=1}^k$ for some $\mathtt{i} \in \Sigma_N^*$ such that n divides $|\mathtt{i}|$. Since $(W_j')_{j=1}^k \in \mathcal{W}_{t_0}$ was arbitrary, we have shown that for every $(W_j)_{j=1}^k \in \mathcal{W}_{t_0}$ we have

$$\left\{(A^{(j)}_{\mathtt{i}} W_j)_{j=1}^k \colon \mathtt{i} \in \Sigma_N^* \text{ and } n \text{ divides } |\mathtt{i}|\right\} = \mathcal{W}_{t_0}.$$

Since n and $(W_j)_{j=1}^k \in \mathcal{W}_{t_0}$ were also arbitrary this shows that the hypothesis of Lemma 5.3(ii) is satisfied by $\mathcal{W}_{t_0}$. Hence $\mathcal{W}_{t_0}$ is primitive as required to prove (iii) and complete the proof of the theorem. $\square$

6. Towards the precondition for ψ-mixing in the primitive proximal case

In this section we prove the following final major component of Theorem 1:

Theorem 7. *Let $N \geq 2$ and $k \geq 1$ and for each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space, let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ be irreducible and let $\beta_j > 0$. For each $j = 1, \ldots, k$ let $\ell_j \geq 1$ be the least possible dimension of a nonzero subspace of V_j which has finite orbit under the action of $(A_1^{(j)}, \ldots, A_N^{(j)})$. Let $\mathcal{W} \subseteq \prod_{j=1}^k \mathrm{Gr}_{\ell_j}(V_j)$ be a primitive subspace class, and for every $\mathtt{i} \in \Sigma_N^*$ define*

$$\Phi_{\mathcal{W}}(\mathtt{i}) := \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^k \left\| A_{\mathtt{i}}^{(j)}|_{W_j} \right\|^{\beta_j} .$$

Suppose that there exist $\mathtt{k} \in \Sigma_N^*$ and $(W_j^0)_{j=1}^k \in \mathcal{W}$ such that $A_{\mathtt{k}}^{(j)}|_{W_j^0} \in \mathrm{Prox}(W_j^0)$ for all $j = 1, \ldots, k$. (In particular we suppose that $A_{\mathtt{k}}^{(j)} W_j^0 = W_j^0$ for all $j = 1, \ldots, k$.) Then there exist $m \geq 1$ and $\delta > 0$ such that

(22)
$$\max_{|\mathtt{k}|=m} \Phi_{\mathcal{W}}(\mathtt{ikj}) \geq \delta \Phi_{\mathcal{W}}(\mathtt{i}) \Phi_{\mathcal{W}}(\mathtt{j})$$

for all $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$.

The inequality (22) refines the statement

(23)
$$\max_{|\mathtt{k}|\leq m} \Phi_{\mathcal{W}}(\mathtt{ikj}) \geq \delta \Phi_{\mathcal{W}}(\mathtt{i}) \Phi_{\mathcal{W}}(\mathtt{j})$$

which was the core result required in the proof of Theorem 3 above in the earlier article [10]. To some extent, therefore, Theorem 7 above adapts the ideas used in [10] in such a manner as to obtain a stronger conclusion under stronger hypotheses. In particular this argument makes heavy use of the properties of the Zariski topology described for the reader in §3.2. We will briefly compare these two results and their proofs.

The earlier result [10, Theorem 6] began by defining $V := \bigoplus_{j=1}^k V_j$ and $g_i := \bigoplus_{j=1}^k A_i^{(j)} \in \mathrm{GL}(V)$ for each $i = 1, \ldots, N$, defining $g_{\mathtt{i}} := g_{i_1} \cdots g_{i_n}$ for every $\mathtt{i} = (i_\ell)_{\ell=1}^n \in \Sigma_N^*$, defining $\Gamma \subset \mathrm{GL}(V)$ to be the semigroup generated by the linear maps $g_1, \ldots, g_N$, taking $G \leq \mathrm{GL}(V)$ to be the Zariski closure of Γ, defining homomorphisms $\phi_j \colon \Gamma \to \mathrm{GL}(V_j)$ by $\phi_j(g_{i_1} \cdots g_{i_n}) := A_{i_1}^{(j)} \cdots A_{i_n}^{(j)}$ for each $j = 1, \ldots, k$ and observing that these extend to regular representations $\phi_j \colon G \to \mathrm{GL}(V_j)$ for each j. The key objective was then to show that there exist an integer $m \geq 1$ and real number $\kappa > 0$ such that for every $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$ there exists $\mathtt{k} \in \Sigma_N^*$ such that $|\mathtt{k}| \leq m$ and $\|\phi_j(g_{\mathtt{i}} g_{\mathtt{k}} g_{\mathtt{j}})\| \geq \kappa \|\phi_j(g_{\mathtt{i}})\| \cdot \|\phi_j(g_{\mathtt{j}})\|$ simultaneously for $j = 1, \ldots, k$; the main result (23) then followed directly. In order to find $\mathtt{k}$ satisfying these simultaneous conditions it is useful to be able to pass to the identity component G^0 of G, which has the advantageous property of being an irreducible variety: in an irreducible variety every nonempty open set is dense, and so to find a word solving k algebraic conditions simultaneously it is sufficient to show that each condition separately is satisfied on a nonempty open subset of G^0, since the intersection of these dense open sets must be nonempty. The analytic result $\|\phi_j(g_{\mathtt{i}} g_{\mathtt{k}} g_{\mathtt{j}})\| \geq \kappa \|\phi_j(g_{\mathtt{i}})\| \cdot \|\phi_j(g_{\mathtt{j}})\|$ is proved by applying the previous reasoning to an algebraic property of $g_{\mathtt{k}}$ and combining this with a compactness argument. In order to reduce this task to that of studying elements of G^0 only, the proof of [10, Theorem 6] exploited the transitivity of $\mathcal{W}$: by appending a small word to $\mathtt{i}$ and prepending a small word to $\mathtt{j}$ if necessary, we could assume that $\|\phi_j(g_{\mathtt{i}})\|$ and $\|\phi_j(g_{\mathtt{j}})\|$ could be understood using the behaviour of $\phi_j(g_{\mathtt{i}})$ and $\phi_j(g_{\mathtt{j}})$ restricted to $W_j \subseteq V_j$ for a consistent

choice of $(W_j)_{j=1}^k \in \mathcal{W}$ not depending on $\mathtt{i}$ or $\mathtt{j}$. This allowed us the freedom to specialise to considering only those $g_\mathtt{k}$ such that $\phi_j(g_\mathtt{k})W_j = W_j$, in particular allowing us to work only in G^0 and not in the whole of G, making arguments based on the irreducibility of the variety G^0 available. Crucially, since only transitivity of $\mathcal{W}$ was assumed, the lengths of the appended and prepended words could be bounded *a priori* via the finiteness and transitivity of $\mathcal{W}$ but their precise length could not be specified in advance.

In order to prove Theorem 7 we improve this argument in two respects so as to make the length of the word $\mathtt{k}$ interposed between $\mathtt{i}$ and $\mathtt{j}$ consistent across all $\mathtt{i}$ and $\mathtt{j}$. Firstly, by assuming primitivity of $\mathcal{W}$ instead of transitivity we can control the length of the small word appended to $\mathtt{i}$ and the small word prepended to $\mathtt{j}$ so as to make their lengths equal to some *a priori* constant p which is independent of the choice of $\mathtt{i}$ and $\mathtt{j}$. This reduces the problem to that of controlling the length of the word $\mathtt{k}$ which is chosen so as to satisfy $\|\phi_j(g_\mathtt{i}g_\mathtt{k}g_\mathtt{j})\| \geq \kappa\|\phi_j(g_\mathtt{i})\| \cdot \|\phi_j(g_\mathtt{j})\|$ simultaneously for $j = 1, \ldots, k$. To solve the latter problem we use the following intuition: if we knew that $\mathtt{k}$ could be chosen with length less than or equal to some number m_0 but also with the property that each $\phi_j(g_\mathtt{k})$ was of rank one and non-nilpotent, then we would know that every power of $\phi_j(g_\mathtt{k})$ is just a scalar multiple of $\phi_j(g_\mathtt{k})$ and satisfies a relation similar to $\|\phi_j(g_\mathtt{i}g_\mathtt{k}g_\mathtt{j})\| \geq \kappa\|\phi_j(g_\mathtt{i})\| \cdot \|\phi_j(g_\mathtt{j})\|$ but with an additional scalar constant. Thus if we took m to be the least common multiple of the lengths of the finitely many different words $\mathtt{k}$ needed to connect the full range of possible pairs of words $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$, and replaced each $\mathtt{k}$ with a power of itself having length m, we could obtain the desired inequality (22) with the word in the middle having length $m + 2p$. But it is obvious that $\phi_j(g_\mathtt{k})$ actually has full rank, so this idea must be modified. The key modification is to choose each $\mathtt{k}$ so that each $\phi_j(g_\mathtt{k})$ is *proximal*, and hence is close in norm to a scalar multiple of its powers. This explains the additional hypothesis of Theorem 7 regarding the existence of a simultaneously proximal word.

Proof of Theorem 7. Define $V := \bigoplus_{j=1}^k V_j$ and for each $i = 1, \ldots, N$ define $g_i :=$ $\bigoplus_{j=1}^k A_i^{(j)} \in \mathrm{GL}(V)$. Define $g_\mathtt{i} := g_{i_1} \cdots g_{i_n}$ for every $\mathtt{i} = (i_\ell)_{\ell=1}^n \in \Sigma_N^*$ in the obvious fashion. Let $\Gamma := \{g_\mathtt{i} \colon \mathtt{i} \in \Sigma_N^*\}$, which is a semigroup. Let $G \leq \mathrm{GL}(V)$ denote the Zariski closure of Γ, which is a group. For each $j = 1, \ldots, k$ we may define a regular representation $\phi_j \colon \Gamma \to \mathrm{GL}(V_j)$ by $\phi_j(g_\mathtt{i}) := A_\mathtt{i}^{(j)}$ and this extends to a regular representation $\phi_j \colon G \to \mathrm{GL}(V_j)$ which is irreducible since its image contains the linear maps $A_1^{(j)}, \ldots, A_N^{(j)}$.

Let $(W_j^0)_{j=1}^k \in \mathcal{W}$ be as in the statement of the theorem. For each $j = 1, \ldots, k$ the subspace W_j^0 has finite orbit under the action of $\phi_j(G)$. Denote this orbit by $\{U_j^1, \ldots, U_j^{r_j}\}$. If $i_0 \in \{1, \ldots, r_j\}$ is fixed then each of the sets $\{g \in G \colon \phi_j(g)U_j^{i_0} = U_j^i\}$ is closed in the Zariski topology, and since these sets are disjoint and their union over $i = 1, \ldots, r_j$ is equal to G, these sets are also open. It follows that they are unions of connected components of G and in particular every connected component of G is contained in a unique set of the form $\{g \in G \colon \phi_j(g)U_j^{i_0} = U_j^i\}$ for some $i \in \{1, \ldots, r_j\}$. Since the identity is in G^0 we must have $G^0 \subseteq \{g \in G \colon \phi_j(g)U_j^{i_0} = U_j^{i_0}\}$, and it follows that every U_j^i is stabilised by $\phi_j(G^0)$. In particular we have $\phi_j(g)W_j^0 = W_j^0$ for every $j = 1, \ldots, k$ for every $g \in G^0$. For each $j = 1, \ldots, k$ let us define a regular representation $\hat{\phi}_j \colon G^0 \to \mathrm{GL}(W_j^0)$ by $\hat{\phi}_j(g) := \phi_j(g)|_{W_j^0}$. By

hypothesis there exists $\mathbf{k} \in \Sigma_N^*$ such that $\hat{\phi}_j(g_{\mathbf{k}}) \in \mathrm{Prox}(W_j^0)$ for every $j = 1, \ldots, k$. The map $g \mapsto g_{\mathbf{k}} g$ is a Zariski homeomorphism of G and induces a permutation of the connected components of G; in particular there exists an iterate of this map which induces the identity permutation. By replacing $\mathbf{k}$ with a suitable power of $\mathbf{k}$ if necessary we may therefore assume without loss of generality that $g_{\mathbf{k}} \in G^0$. Define $q_{\mathbf{k},j} := \lim_{n \to \infty} \|\hat{\phi}_j(g_{\mathbf{k}}^n)\|^{-1} \hat{\phi}_j(g_{\mathbf{k}}^n) \in \mathrm{End}(W_j^0)$ for each $j = 1, \ldots, k$; by proximality each $q_{\mathbf{k},j}$ is well-defined, of rank one, and not nilpotent.

We make the following first claim: if nonzero elements $b_{1,j}, b_{2,j}$ of $\mathrm{End}(W_j^0)$ are given for each $j = 1, \ldots, k$ then there exists $\mathbf{i} \in \Sigma_N^*$ such that $g_{\mathbf{i}} \in G^0$, such that $\hat{\phi}_j(g_{\mathbf{i}}) \in \mathrm{Prox}(W_j^0)$ for each $j = 1, \ldots, k$ and such that for each $j = 1, \ldots, k$ the linear map $p_{\mathbf{i},j} := \lim_{n \to \infty} \|\hat{\phi}_j(g_{\mathbf{i}}^n)\|^{-1} \hat{\phi}_j(g_{\mathbf{i}}^n) \in \mathrm{End}(W_j^0)$ satisfies $b_{1,j} p_{\mathbf{i},j} b_{2,j} \neq 0$. Clearly it will be sufficient to choose a nonzero vector $v_j \in b_{2,j} W_j^0$ for each $j = 1, \ldots, k$ and find $\mathbf{i} \in \Sigma_N^*$ such that $g_{\mathbf{i}} \in G^0$, such that $\hat{\phi}_j(g_{\mathbf{i}})$ is proximal for all $j = 1, \ldots, k$ and such that $p_{\mathbf{i},j} v_j \notin \ker b_{1,j}$ for all $j = 1, \ldots, k$. We therefore fix nonzero linear maps $b_{1,j}$ and $b_{2,j}$ and nonzero vectors v_j for each $j = 1, \ldots, k$ and prove the claim in this form.

We assert that there exists $\mathbf{j}_2 \in \Sigma_N^*$ such that $g_{\mathbf{j}_2} \in G^0$ and such that for all $j = 1, \ldots, k$ we have $q_{\mathbf{k},j} \hat{\phi}_j(g_{\mathbf{j}_2}) v_j \neq 0$. To see this it suffices to show that

$$\bigcap_{j=1}^{k} \left\{ g \in G^0 \colon q_{\mathbf{k},j} \hat{\phi}_j(g) v_j \neq 0 \right\}$$

is nonempty and Zariski open, since by the Zariski density of Γ in G it must then contain some $g_{\mathbf{i}} \in \Gamma$. This set is the intersection of the sets

(24) $$\left\{ g \in G^0 \colon q_{\mathbf{k},j} \hat{\phi}_j(g) v_j \neq 0 \right\}$$

over $j = 1, \ldots, k$ and each of these sets is clearly Zariski open. Since G^0 is an irreducible variety, all of its nonempty open subsets are also dense, so the intersection of the sets (24) will be nonempty and open as long as each individual set is nonempty. The assertion will therefore be proved if each of the sets in (24) is shown to be nonempty. But if this set is empty for some j then the vector space $U := \mathrm{span}\{\hat{\phi}_j(g) v_j \colon g \in G^0\}$ is a subspace of W_j^0 (and hence of V_j) which is invariant under $\phi_j(G^0)$ and has smaller dimension than W_j^0, since it is contained in the proper subspace $\ker q_{\mathbf{k},j}$ of W_j^0. If $h_1, h_2 \in G$ belong to the same connected component G_i of G then $h_1^{-1} G_i$ is a connected component of G which contains the identity, so $h_1^{-1} h_2 \in h_1^{-1} G_i = G^0$, hence $\phi_j(h_1)^{-1} \phi_j(h_2) U = \phi_j(h_1^{-1} h_2) U = U$ by the $\phi_j(G^0)$-invariance of U. Thus $g \mapsto \phi_j(g) U$ is constant on each connected component of G and therefore $\{\phi_j(g) U \colon g \in G\}$ is finite. But then U has finite orbit under $\phi_j(G)$ and in particular has finite orbit under $\phi_j(\Gamma) = \{A_{\mathbf{i}}^{(j)} \colon \mathbf{i} \in \Sigma_N^*\}$ whilst having dimension smaller than $\ell_j = \dim W_j^0$; this contradicts the definition of ℓ_j. We conclude that such a subspace U cannot exist, so the set (24) must be nonempty for every $j = 1, \ldots, k$ and we deduce the existence of the claimed element $\mathbf{j}_2 \in \Sigma_N^*$.

We next assert that there exists $\mathbf{j}_1 \in \Sigma_N^*$ such that $g_{\mathbf{j}_1} \in G^0$ and such that for all $j = 1, \ldots, k$ the endomorphism $\hat{\phi}_j(g_{\mathbf{j}_1}) q_{\mathbf{k},j} \hat{\phi}_j(g_{\mathbf{j}_2})$ is proximal and satisfies $\hat{\phi}_j(g_{\mathbf{j}_1}) q_{\mathbf{k},j} \hat{\phi}_j(g_{\mathbf{j}_2}) v_j \notin \ker b_{1,j}$. Clearly $\hat{\phi}_j(g) q_{\mathbf{k},j} \hat{\phi}_j(g_{\mathbf{j}_2}) \in \mathrm{End}(W_j^0)$ has rank one for every $g \in G^0$ since $q_{\mathbf{k},j}$ is of rank one and $\hat{\phi}_j(g)$ and $\hat{\phi}_j(g_{\mathbf{j}_2})$ are invertible, so for

$\hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})$ to be proximal it is necessary and sufficient that it be non-nilpotent, which by rank considerations is equivalent to the condition $(\hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}))^2 \neq 0$. Thus to obtain the existence of j_1 it suffices to show that

(25)
$$\bigcap_{j=1}^{k} \left\{ g \in G^0 \colon \hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})v_j \notin \ker b_{1,j} \right\} \cap \left\{ g \in G^0 \colon \hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}))^2 \neq 0 \right\}$$

is nonempty and Zariski open. We must therefore likewise show that each set

$$\left\{ g \in G^0 \colon \hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})v_j \notin \ker b_{1,j} \right\}$$

is nonempty and Zariski open, and that each set

$$\left\{ g \in G^0 \colon (\hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}))^2 \neq 0 \right\}$$

is also nonempty and Zariski open. A vector belongs to $\ker b_{1,j}$ if and only if it orthogonal to every element of a basis for the orthogonal complement of $\ker b_{1,j}$, and the latter is obviously a Zariski closed condition, so the first of the two sets is Zariski open. The Zariski openness of the second set is obvious.

If the first set is empty for some j, then by the same arguments as were used previously the vector space $\mathrm{span}\{\hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})v_j \colon g \in G^0\}$ would be a $\phi_j(G^0)$-invariant proper subspace of W_j^0, the existence of which would contradict the definition of ℓ_j. We conclude that the first set is nonempty for each $j = 1, \ldots, k$. If $(\hat{\phi}_j(g)q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}))^2 = 0$ for some $g \in G^0$ then since $q_{\mathsf{k},j}$ has rank one and $\hat{\phi}_j(g)$ and $\hat{\phi}_j(g_{\mathsf{j}_2})$ are invertible, it must be the case that the one-dimensional image subspace $q_{\mathsf{k},j}W_j^0$ is mapped into the kernel of $q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})$ by $\hat{\phi}_j(g)$. If this holds for every $g \in G^0$ then we deduce that $\mathrm{span}\bigcup_{g \in G^0} \hat{\phi}_j(g)q_{\mathsf{k},j}W_j^0 \subseteq \ker q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}) \neq W_j^0$ is a $\phi_j(G^0)$-invariant proper subspace of W_j^0, which is again impossible. We deduce the nonemptiness and Zariski openness of the set (25) and the existence of j_1 follows.

We have shown that there exist $\mathsf{j}_1, \mathsf{j}_2 \in \Sigma_N^*$ such that $g_{\mathsf{j}_1}, g_{\mathsf{j}_2} \in G^0$ and such that for every $j = 1, \ldots, k$, $\hat{\phi}_j(g_{\mathsf{j}_1})q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}) \in \mathrm{End}(W_j^0)$ is proximal and of rank one and satisfies $\hat{\phi}_j(g_{\mathsf{j}_1})q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})v_j \notin \ker b_{1,j}$. In particular we necessarily have $V^+(\hat{\phi}_j(g_{\mathsf{j}_1})q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2})) \cap \ker b_{1,j} = \{0\}$ and $v_j \notin V^-(\hat{\phi}_j(g_{\mathsf{j}_1})q_{\mathsf{k},j}\hat{\phi}_j(g_{\mathsf{j}_2}))$. By the openness of the set of proximal endomorphisms in the analytic topology on $\mathrm{End}(W_j^0)$ together with the continuity of V^+ and V^- on that set, it follows that for all sufficiently large n the element $\|\hat{\phi}_j(g_{\mathsf{k}}^n)\|^{-1}\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2})$ is proximal and satisfies $V^+(\|\hat{\phi}_j(g_{\mathsf{k}}^n)\|^{-1}\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2})) \cap \ker b_{1,j} = \{0\}$ and $v_j \notin V^-(\|\hat{\phi}_j(g_{\mathsf{k}}^n)\|^{-1}\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2}))$ for all $j = 1, \ldots, k$. Fix n large enough that these properties hold and define $\mathsf{i} := \mathsf{j}_1\mathsf{k}^n\mathsf{j}_2$. We then have $\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2}) \in \mathrm{Prox}(W_j^0)$ for all $j = 1, \ldots, k$, and for every $j = 1, \ldots, k$ we also have $V^+(\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2})) \cap \ker b_{1,j} = \{0\}$ and $v_j \notin V^-(\hat{\phi}_j(g_{\mathsf{j}_1}g_{\mathsf{k}}^n g_{\mathsf{j}_2}))$. The limit $p_{\mathsf{i},j} := \lim_{n \to \infty} \|\hat{\phi}_j(g_{\mathsf{i}}^n)\|^{-1}\hat{\phi}_j(g_{\mathsf{i}}^n) \in \mathrm{Prox}(W_j^0)$ has image $V^+(\hat{\phi}_j(g_{\mathsf{i}}))$ and kernel $V^-(\hat{\phi}_j(g_{\mathsf{i}}))$ for each $j = 1, \ldots, k$, so in particular $p_{\mathsf{i},j}v_j \notin \ker b_{1,j}$ and it follows that $b_{1,j}p_{\mathsf{i},j}b_{2,j} \neq 0$. The proof of the first claim is complete.

We secondly claim that there exist $\kappa > 0$ and $m \geq 1$ with the following property: if for each $j = 1, \ldots, k$ we are given $b_{1,j}, b_{2,j} \in \mathrm{End}(W_j^0)$, then

$$\max_{\substack{|\mathbf{k}|=m \\ g_{\mathbf{k}} \in G^0}} \left\| b_{1,j} \hat{\phi}_j(g_{\mathbf{k}}) b_{2,j} \right\| \geq \kappa \|b_{1,j}\| \cdot \|b_{2,j}\|$$

for every $j = 1, \ldots, k$. By homogeneity it is clearly sufficient to consider only the case where $\|b_{1,j}\| = \|b_{2,j}\| = 1$ for every $j = 1, \ldots, k$, and we will prove the claim in this form.

For each $j = 1, \ldots, k$ let $S_{\mathrm{End}(W_j^0)}$ denote the unit sphere of $\mathrm{End}(W_j^0)$. If $((b_{1,j}, b_{2,j}))_{j=1}^k \in \prod_{j=1}^k S_{\mathrm{End}(W_j^0)} \times S_{\mathrm{End}(W_j^0)}$ is given, then by the preceding step there exists $\mathtt{i} \in \Sigma_N^*$ such that $g_{\mathtt{i}} \in G^0$ and such that for each $j = 1, \ldots, k$ the element $p_{\mathtt{i},j} := \lim_{n \to \infty} \|\hat{\phi}_j(g_{\mathtt{i}}^n)\|^{-1} \hat{\phi}_j(g_{\mathtt{i}}^n) \in \mathrm{End}(W_j^0)$ is well-defined and satisfies $b_{1,j} p_{\mathtt{i},j} b_{2,j} \neq 0$. If $((b_{1,j}', b_{2,j}'))_{j=1}^k \in \prod_{j=1}^k S_{\mathrm{End}(W_j^0)} \times S_{\mathrm{End}(W_j^0)}$ is chosen in a sufficiently small open neighbourhood of $((b_{1,j}, b_{2,j}))_{j=1}^k$ then we clearly also have $b_{1,j}' p_{\mathtt{i},j} b_{2,j}' \neq 0$ for every $j = 1, \ldots, k$ for the same word $\mathtt{i}$. By the compactness of $\prod_{j=1}^k S_{\mathrm{End}(W_j^0)} \times S_{\mathrm{End}(W_j^0)}$ it follows that there exist finitely many words $\mathtt{i}_1, \ldots, \mathtt{i}_r \in \Sigma_N^*$ such that $g_{\mathtt{i}_t} \in G^0$ for every $t = 1, \ldots, r$, such that $p_{\mathtt{i}_t,j} := \lim_{n \to \infty} \|\hat{\phi}_j(g_{\mathtt{i}_t}^n)\|^{-n} \hat{\phi}_j(g_{\mathtt{i}_t}^n) \in \mathrm{End}(W_j^0)$ is well-defined for each $j = 1, \ldots, k$ and $t = 1, \ldots, r$ and such that for every $((b_{1,j}, b_{2,j}))_{j=1}^k \in S_{\mathrm{End}(W_j^0)}$ there exists $t \in \{1, \ldots, r\}$ such that $b_{1,j} p_{\mathtt{i}_t,j} b_{2,j} \neq 0$ for all $j = 1, \ldots, k$. By compactness and continuity the function $\prod_{j=1}^k S_{\mathrm{End}(W_j^0)} \times S_{\mathrm{End}(W_j^0)} \to \mathbb{R}$ defined by

$$((b_{1,j}, b_{2,j}))_{j=1}^k \mapsto \min_{1 \leq j \leq k} \max_{1 \leq t \leq r} \|b_{1,j} p_{\mathtt{i}_t,j} b_{2,j}\|$$

therefore has a nonzero minimum value $\tau > 0$, say.

Let $m \geq 1$ be a natural number divisible by each of $|\mathtt{i}_1|, \ldots, |\mathtt{i}_t|$ and choose natural numbers $n_1, \ldots, n_r$ such that $m = n_1|\mathtt{i}_1| = n_2|\mathtt{i}_2| = \cdots = n_r|\mathtt{i}_r|$. By choosing a large integer $\ell \geq 1$ and replacing m with ℓm and each n_t with ℓn_t if required, we may without loss of generality suppose that $n_1, \ldots, n_r$ are large enough that

$$\max_{1 \leq j \leq k} \max_{1 \leq t \leq r} \left\| p_{\mathtt{i}_t,j} - \|\hat{\phi}_j(g_{\mathtt{i}_t}^{n_t})\|^{-1} \hat{\phi}_j(g_{\mathtt{i}_t}^{n_t}) \right\| < \frac{\tau}{2}.$$

Define $\mathbf{k}_t := \mathtt{i}_t^{n_t} \in \Sigma_N^*$ for each t and observe that each $\mathbf{k}_t$ has the same length m and satisfies $g_{\mathbf{k}_t} = g_{\mathtt{i}_t}^{n_t} \in G^0$. We easily see that

$$\min_{1 \leq j \leq k} \max_{1 \leq t \leq r} \|\hat{\phi}_j(g_{\mathbf{k}_t})\|^{-1} \|b_{1,j} \hat{\phi}_j(g_{\mathbf{k}_t}) b_{2,j}\| > \frac{\tau}{2}$$

for all $((b_{1,j}, b_{2,j}))_{j=1}^k \in \prod_{j=1}^k S_{\mathrm{End}(W_j^0)} \times S_{\mathrm{End}(W_j^0)}$, $j = 1, \ldots, k$ and $t = 1, \ldots, r$, so if we define

$$\kappa := \frac{\tau}{2} \min_{1 \leq j \leq k} \min_{1 \leq t \leq r} \|\hat{\phi}_j(g_{\mathbf{k}_t})\|$$

then we have proved our second claim.

We may now prove the theorem. Let $\mathtt{i}, \mathtt{j} \in \Sigma_N^*$ be arbitrary. There exist $(W_j)_{j=1}^k, (W_j')_{j=1}^k \in \mathcal{W}$ such that

$$\Phi_{\mathcal{W}}(\mathtt{i}) = \prod_{j=1}^k \left\| \phi_j(g_{\mathtt{i}})|_{W_j} \right\|^{\beta_j}, \qquad \Phi_{\mathcal{W}}(\mathtt{j}) = \prod_{j=1}^k \left\| \phi_j(g_{\mathtt{j}})|_{W_j'} \right\|^{\beta_j}.$$

By the hypothesis that $\mathcal{W}$ is primitive there exists $p \geq 1$ not depending on $\mathtt{i}, \mathtt{j},$ $(W_j)_{j=1}^k$ or $(W_j')_{j=1}^k$ such that we may choose $\mathtt{i}_1, \mathtt{i}_2, \mathtt{j}_1, \mathtt{j}_2 \in \Sigma_N^*$ with $|\mathtt{i}_1| = |\mathtt{i}_2| = |\mathtt{j}_1| = |\mathtt{j}_2| = p$ satisfying

$$(\phi_j(g_{\mathtt{j}_2})W_j^0)_{j=1}^k = (W_j')_{j=1}^k,$$
$$(\phi_j(g_{\mathtt{j}_1})\phi_j(g_{\mathtt{j}})W_j')_{j=1}^k = (W_j^0)_{j=1}^k,$$
$$(\phi_j(g_{\mathtt{i}_2})W_j^0)_{j=1}^k = (W_j)_{j=1}^k,$$
$$(\phi_j(g_{\mathtt{i}_1})\phi_j(g_{\mathtt{i}})W_j))_{j=1}^k = (W_j^0)_{j=1}^k.$$

In particular $\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2})|_{W_j^0}$ and $\phi_j(g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}$ are endomorphisms of W_j^0 for each $j = 1, \ldots, k$, and

$$\prod_{j=1}^k \|\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2})|_{W_j^0}\|^{\beta_j} \geq \left(\min_{|\mathtt{l}|=p} \prod_{j=1}^k \|\phi_j(g_{\mathtt{l}})^{-1}\|^{-\beta_j}\right)^2 \prod_{j=1}^k \|\phi_j(g_{\mathtt{i}})|_{W_j}\|^{\beta_j} = \varepsilon\Phi_{\mathcal{W}}(\mathtt{i})$$

and

$$\prod_{j=1}^k \|\phi_j(g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}\|^{\beta_j} \geq \left(\min_{|\mathtt{l}|=p} \prod_{j=1}^k \|\phi_j(g_{\mathtt{l}})^{-1}\|^{-\beta_j}\right)^2 \prod_{j=1}^k \left\|\phi_j(g_{\mathtt{j}})|_{W_j'}\right\|^{\beta_j} = \varepsilon\Phi_{\mathcal{W}}(\mathtt{j}),$$

say, where

$$\varepsilon := \left(\min_{|\mathtt{l}|=p} \prod_{j=1}^k \|\phi_j(g_{\mathtt{l}})^{-1}\|^{-\beta_j}\right)^2.$$

Now by the previous step there exists $\mathtt{k} \in \Sigma_N^*$ with $|\mathtt{k}| = m$ such that $g_{\mathtt{k}} \in G^0$ (and hence $\phi_j(g_{\mathtt{k}})W_j^0 = W_j^0$ for every $j = 1, \ldots, k$) and such that

$$\prod_{j=1}^k \left\|\left(\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2})|_{W_j^0}\right)\left(\phi_j(g_{\mathtt{k}})|_{W_j^0}\right)\left(\phi_j(g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}\right)\right\|^{\beta_j}$$

$$\geq \kappa^{\sum_{j=1}^k \beta_j} \left(\prod_{j=1}^k \left\|\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2})|_{W_j^0}\right\|^{\beta_j}\right)\left(\prod_{j=1}^k \left\|\phi_j(g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}\right\|^{\beta_j}\right)$$

$$\geq \kappa^{\sum_{j=1}^k \beta_j} \varepsilon^2 \Phi_{\mathcal{W}}(\mathtt{i})\Phi_{\mathcal{W}}(\mathtt{j}).$$

Defining $K := \max_{|\mathtt{l}|=p} \Phi_{\mathcal{W}}(\mathtt{l}) < \infty$ we have

$$\prod_{j=1}^k \left\|\left(\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2})|_{W_j^0}\right)\left(\phi_j(g_{\mathtt{k}})|_{W_j^0}\right)\left(\phi_j(g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}\right)\right\|^{\beta_j}$$

$$= \prod_{j=1}^k \left\|\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2}g_{\mathtt{k}}g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j^0}\right\|^{\beta_j}$$

$$\leq \max_{(W_j)_{j=1}^k \in \mathcal{W}} \prod_{j=1}^k \left\|\phi_j(g_{\mathtt{i}_1}g_{\mathtt{i}}g_{\mathtt{i}_2}g_{\mathtt{k}}g_{\mathtt{j}_1}g_{\mathtt{j}}g_{\mathtt{j}_2})|_{W_j}\right\|^{\beta_j}$$

$$= \Phi_{\mathcal{W}}(\mathtt{i}_1\mathtt{i}\mathtt{i}_2\mathtt{k}\mathtt{j}_1\mathtt{j}\mathtt{j}_2) \leq \Phi_{\mathcal{W}}(\mathtt{i}_1)\Phi_{\mathcal{W}}(\mathtt{j}_2)\Phi_{\mathcal{W}}(\mathtt{i}\mathtt{i}_2\mathtt{k}\mathtt{j}_1\mathtt{j}) \leq K^2\Phi_{\mathcal{W}}(\mathtt{i}\mathtt{i}_2\mathtt{k}\mathtt{j}_1\mathtt{j}),$$

and since $|\mathtt{i}_2\mathtt{k}\mathtt{j}_1| = m + 2p$, we have obtained

$$\max_{|\mathtt{l}|=m+2p} \Phi_{\mathcal{W}}(\mathtt{i}\mathtt{l}\mathtt{j}) \geq \Phi_{\mathcal{W}}(\mathtt{i}\mathtt{i}_2\mathtt{k}\mathtt{j}_1\mathtt{j}) \geq K^{-2}\varepsilon^2\kappa^{\sum_{j=1}^{k}\beta_j}\Phi_{\mathcal{W}}(\mathtt{i})\Phi_{\mathcal{W}}(\mathtt{j})$$

where m, p, K, κ and ε do not depend on $\mathtt{i},\mathtt{j} \in \Sigma_N^*$. The theorem is proved. $\square$

7. Proofs of main results

7.1. **Proof of Theorem 1.** Let the integers k and N, vector spaces V_j, tuples $(A_1^{(j)},\ldots,A_N^{(j)}) \in \mathrm{GL}(V_j)^N$, real numbers $\beta_j > 0$, potential $\Phi\colon \Sigma_N^* \to (0,+\infty)$ and totally ergodic equilibrium state $\mu \in \mathcal{M}_\sigma(\Sigma_N)$ be as in the statement of Theorem 1. By Theorem 5 we may without loss of generality assume that every $(A_1^{(j)},\ldots,A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ is irreducible and has simple top Lyapunov exponent with respect to μ. By Theorem 6(i) we find that μ is the equilibrium state of a unique potential $\Phi_{\mathcal{W}}$ such that $\mathcal{W} \subseteq \prod_{j=1}^{k} \mathrm{Gr}_{\ell_j}(V_j)$ is a transitive subspace class, where for each $j = 1,\ldots,k$ the integer ℓ_j is the dimension of the smallest nonzero subspace of V_j with finite orbit under the action of $(A_1^{(j)},\ldots,A_N^{(j)})$. It follows by Theorem 3(i) that there exists $C > 0$ such that

$$C^{-1}\Phi_{\mathcal{W}}(\mathtt{i}) \leq e^{|\mathtt{i}|P(\Phi_{\mathcal{W}})}\mu([\mathtt{i}]) \leq C\Phi_{\mathcal{W}}(\mathtt{i})$$

for every $\mathtt{i} \in \Sigma_N^*$. By Theorem 6(ii)–(iii) the potential $\Phi_{\mathcal{W}}$ satisfies the hypotheses of Theorem 7, so there exist an integer m and constant $\delta > 0$ such that

$$\max_{|\mathtt{k}|=m} \Phi_{\mathcal{W}}(\mathtt{i}\mathtt{k}\mathtt{j}) \geq \delta\Phi_{\mathcal{W}}(\mathtt{i})\Phi_{\mathcal{W}}(\mathtt{j})$$

for all $\mathtt{i},\mathtt{j} \in \Sigma_N^*$. Hence for every $\mathtt{i},\mathtt{j} \in \Sigma_N^*$ we have

$$\begin{aligned}
\delta\mu([\mathtt{i}])\mu([\mathtt{j}]) &\leq C^2\delta e^{-(|\mathtt{i}|+|\mathtt{j}|)P(\Phi_{\mathcal{W}})}\Phi_{\mathcal{W}}(\mathtt{i})\Phi_{\mathcal{W}}(\mathtt{j}) \\
&\leq C^2 e^{-(|\mathtt{i}|+|\mathtt{j}|)P(\Phi_{\mathcal{W}})}\max_{|\mathtt{k}|=m}\Phi_{\mathcal{W}}(\mathtt{i}\mathtt{k}\mathtt{j}) \\
&\leq C^3 e^{|\mathtt{k}|P(\Phi_{\mathcal{W}})}\max_{|\mathtt{k}|=m}\mu([\mathtt{i}\mathtt{k}\mathtt{j}]) \\
&\leq C^3 e^{mP(\Phi_{\mathcal{W}})}\sum_{|\mathtt{k}|=m}\mu([\mathtt{i}\mathtt{k}\mathtt{j}]) \\
&= C^3 e^{mP(\Phi_{\mathcal{W}})}\mu([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}])
\end{aligned}$$

so that

(26) $$\mu([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}]) \geq \kappa\mu([\mathtt{i}])\mu([\mathtt{j}])$$

where $\kappa := C^{-3}\delta e^{-mP(\Phi_{\mathcal{W}})}$, and also

$$
\begin{aligned}
\mu([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}]) &= \sum_{|\mathbf{k}|=m} \mu([\mathtt{ikj}]) \\
&\leq C \sum_{|\mathbf{k}|=m} e^{-(|\mathtt{i}|+|\mathbf{k}|+|\mathtt{j}|)P(\Phi_{\mathcal{W}})} \Phi_{\mathcal{W}}(\mathtt{ikj}) \\
&\leq Ce^{-(|\mathtt{i}|+|\mathtt{j}|)P(\Phi_{\mathcal{W}})} \Phi_{\mathcal{W}}(\mathtt{i})\Phi_{\mathcal{W}}(\mathtt{j}) \left(\sum_{|\mathbf{k}|=m} e^{-|\mathbf{k}|P(\Phi_{\mathcal{W}})} \Phi_{\mathcal{W}}(\mathbf{k}) \right) \\
&\leq C^4 \mu([\mathtt{i}])\mu([\mathtt{j}]) \left(\sum_{|\mathbf{k}|=m} \mu([\mathbf{k}]) \right) \\
&= C^4 \mu([\mathtt{i}])\mu([\mathtt{j}])
\end{aligned}
$$

so that

$$(27) \qquad\qquad \mu([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}]) \leq K\mu([\mathtt{i}])\mu([\mathtt{j}])$$

where $K := C^4$. To prove the theorem we will combine inequalities (26) and (27) with theorems of R.C. Bradley, N.A. Friedman and D.S. Ornstein in a manner similar to earlier works such as [50, 52].

For every integer $n \geq 1$ let $\mathcal{B}_{1,n}$ be the finite σ-algebra on $\hat{\Sigma}_N$ generated by the set

$$\left\{ [\mathtt{i}] \subset \hat{\Sigma}_N \colon |\mathtt{i}| = n \right\}.$$

For every pair of integers $n, m \in \mathbb{Z}$ such that $n \leq m$ define $\mathcal{B}_{n,m} := \hat{\sigma}^{n-1}\mathcal{B}_{1,m+1-n}$. Thus $\mathcal{B}_{n,m}$ is precisely the σ-algebra generated by cylinders of the form

$$\{(x_\ell)_{\ell \in \mathbb{Z}} \colon x_i = y_i \text{ for all } i = n, \dots, m\}$$

where the finite sequence $(y_i)_{i=n}^m$ varies over $\{1, \dots, N\}^{m-n+1}$. For every $n \in \mathbb{Z}$ define also

$$\mathcal{B}_{-\infty,n} := \bigvee_{m=-\infty}^{n} \mathcal{B}_{m,n}, \qquad \mathcal{B}_{n,+\infty} := \bigvee_{m=n}^{\infty} \mathcal{B}_{n,m}.$$

The following is a special case of a theorem of R.C. Bradley ([14, Theorem 4.1(2)]):

Theorem 8. *Let $\hat{\mu}$ be a $\hat{\sigma}$-invariant measure on $\hat{\Sigma}_N$ such that for some integer $m \geq 1$ the conditions*

$$\inf_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_m \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} > 0, \qquad \sup_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_m \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} < \infty$$

are both satisfied. Then

$$\lim_{n \to \infty} \sup_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_n \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \left| \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} - 1 \right| = 0.$$

Now, the natural extension $\hat{\mu} \in \mathcal{M}_{\hat{\sigma}}(\hat{\Sigma}_N)$ of the equilibrium state μ satisfies

$$\inf_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_{m+1} \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} = \inf_{n_1, n_2 \geq 1} \inf_{\substack{A \in \mathcal{B}_{1-n_1, 0} \\ B \in \mathcal{B}_{m+1, m+n_2}}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)}$$

$$= \inf_{n_1, n_2 \geq 1} \inf_{\substack{A \in \mathcal{B}_{1, n_1} \\ B \in \mathcal{B}_{m+n_1+1, m+n_1+n_2}}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)}$$

$$= \inf_{n_1, n_2 \geq 1} \inf_{\substack{|\mathtt{i}|=n_1 \\ |\mathtt{j}|=n_2}} \frac{\hat{\mu}([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}])}{\hat{\mu}([\mathtt{i}])\hat{\mu}([\mathtt{j}])} \geq \kappa > 0$$

by (26), and likewise

$$\sup_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_{m+1} \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} = \sup_{n_1, n_2 \geq 1} \sup_{\substack{|\mathtt{i}|=n_1 \\ |\mathtt{j}|=n_2}} \frac{\hat{\mu}([\mathtt{i}] \cap \sigma^{-m-|\mathtt{i}|}[\mathtt{j}])}{\hat{\mu}([\mathtt{i}])\hat{\mu}([\mathtt{j}])} \leq K < \infty$$

by (27). Theorem 8 therefore applies and yields

$$\lim_{n \to \infty} \sup_{\substack{A \in \mathcal{B}^0_{-\infty}, B \in \mathcal{B}^\infty_n \\ \hat{\mu}(A), \hat{\mu}(B) \neq 0}} \left| \frac{\hat{\mu}(A \cap B)}{\hat{\mu}(A)\hat{\mu}(B)} - 1 \right| = 0$$

which clearly implies the result

$$\lim_{n \to \infty} \sup_{\mathtt{i}, \mathtt{j} \in \Sigma^*_N} \left| \frac{\mu([\mathtt{i}] \cap \sigma^{-n-|\mathtt{i}|}[\mathtt{j}])}{\mu([\mathtt{i}])\mu([\mathtt{j}])} - 1 \right| = 0$$

which is the first assertion of Theorem 1.

To deduce the Bernoulli property of $\hat{\mu}$ we will apply a theorem of N.A. Friedman and D.S. Ornstein. We recall that a measure space $(X, \mathcal{F}, m)$ is called a Lebesgue space if there exists a measure space isomorphism between $(X, \mathcal{F}, m)$ and Lebesgue measure on a bounded interval equipped with the σ-algebra of Lebesgue measurable sets. If $(X, \mathcal{F}, m)$ is a Lebesgue space, $Z \subseteq X$ has nonzero measure, $\mathcal{F}_Z := \{A \cap Z \colon A \in \mathcal{F}\}$ and m_Z is the measure on $(X, \mathcal{F}_Z)$ defined by $m_Z(A) := m(Z \cap A)$ then $(Z, \mathcal{F}_Z, m_Z)$ is also a Lebesgue space.

If T is an invertible measure-preserving transformation of a Lebesgue probability space $(X, \mathcal{F}, m)$ then a partition $\mathcal{P}$ of X is defined to be any finite set $\mathcal{P} = \{P_1, \ldots, P_n\} \subset \mathcal{F}$ such that, up to measure zero, X is the disjoint union of the sets $P_1, \ldots, P_n$. Given a partition $\mathcal{P}$, for each $k \in \mathbb{Z}$ we may define a new partition $T^k \mathcal{P} := \{T^k P_1, \ldots, T^k P_n\}$ of X in the obvious fashion. If $\mathcal{P}_1, \ldots, \mathcal{P}_k$ are partitions then we let $\bigvee_{j=1}^k \mathcal{P}_j$ denote the partition $\{A_1 \cap A_2 \cap \cdots \cap A_k \colon A_j \in \mathcal{P}_j\}$. We will say that a partition $\mathcal{P}$ of X is ε-independent of a partition $\mathcal{Q}$ of X if there exists a subset $\mathcal{Q}_\varepsilon$ of $\mathcal{Q}$ such that

$$m\left(\bigcup_{Q \in \mathcal{Q}_\varepsilon} Q \right) > 1 - \varepsilon$$

and

$$\max_{P \in \mathcal{P}} \max_{Q \in \mathcal{Q}_\varepsilon} \left| \frac{m(P \cap Q)}{m(Q)} - m(P) \right| < \varepsilon.$$

A partition $\mathcal{P}$ is called a *weak Bernoulli partition* if for every $\varepsilon > 0$ there exists an integer $k_\varepsilon \geq 0$ such that for all $n \geq 1$ the partition $\bigvee_{j=k_\varepsilon+1}^{k_\varepsilon+n} T^j \mathcal{P}$ is ε-independent of

$\bigvee_{j=1-n}^{0} T^j \mathcal{P}$, or equivalently if $\bigvee_{j=1-n}^{0} T^j \mathcal{P}$ is ε-independent of $\bigvee_{j=-2n-k_\varepsilon+1}^{-n-k_\varepsilon} T^j \mathcal{P}$. The following theorem paraphrases a celebrated result of N.A. Friedman and D.S. Ornstein [29]:

Theorem 9. *Let* $\mathbb{P} = (\sum_{i=1}^{n} p_i \delta_i)^{\mathbb{Z}}$ *be a Bernoulli measure on* $\hat{\Sigma}_N$ *for some* $N \geq 2$ *and some nondegenerate probability vector* $(p_1, \ldots, p_N)$ *and let* $\overline{\mathcal{B}}_{\mathbb{P}}$ *denote the completion of the Borel* σ*-algebra on* $\hat{\Sigma}_N$ *with respect to* $\mathbb{P}$. *Let* T *be an invertible measure-preserving transformation of a Lebesgue probability space* $(X, \mathcal{F}, m)$ *which admits a weak Bernoulli partition. Then there exists a measure space isomorphism* $\phi \colon X \to \hat{\Sigma}_N$ *such that* $\phi \circ T = \hat{\sigma} \circ \phi$.

If μ is a totally ergodic generalised matrix equilibrium state on Σ_N for some $N \geq 2$, let $\hat{\mu}$ denote its natural extension to $\hat{\Sigma}_N$ and let $\overline{\mathcal{B}}_{\hat{\mu}}$ denote the completion of the Borel σ-algebra on $\hat{\Sigma}_N$ with respect to $\hat{\mu}$. Since μ is fully supported on Σ_N it is not a Dirac measure, and since it is additionally totally ergodic, it has no atoms. This implies the corresponding properties for $\hat{\mu}$ on $\hat{\Sigma}_N$ and it follows that $(\hat{\Sigma}_N, \overline{\mathcal{B}}_{\hat{\mu}}, \hat{\mu})$, being the completion of an atomless Borel probability measure on a complete metric space, is a Lebesgue probability space. Let $\mathcal{P}$ denote the partition of $\hat{\Sigma}_N$ into the N sets $[i] := \{(x_\ell)_{\ell \in \mathbb{Z}} \colon x_1 = i\}$ for $i = 1, \ldots, N$. Since

$$\lim_{n \to \infty} \sup_{\mathtt{i},\mathtt{j} \in \Sigma_N^*} \left| \frac{\hat{\mu}([\mathtt{i}] \cap \hat{\sigma}^{-n-|\mathtt{i}|}[\mathtt{j}])}{\hat{\mu}([\mathtt{i}])\hat{\mu}([\mathtt{j}])} - 1 \right| = 0,$$

for every $\varepsilon > 0$ we may choose $k_\varepsilon \geq 1$ such that

$$\sup_{\mathtt{i},\mathtt{j} \in \Sigma_N^*} \left| \frac{\hat{\mu}([\mathtt{i}] \cap \hat{\sigma}^{-k_\varepsilon - |\mathtt{i}|}[\mathtt{j}])}{\hat{\mu}([\mathtt{i}])\hat{\mu}([\mathtt{j}])} - 1 \right| < \varepsilon.$$

For each $n \geq 1$ the partitions $\bigvee_{j=1-n}^{0} \hat{\sigma}^j \mathcal{P}$ and $\bigvee_{j=-2n-k_\varepsilon+1}^{-n-k_\varepsilon} \hat{\sigma}^j \mathcal{P}$ are simply the partitions into sets of the form $[\mathtt{i}]$ and into sets of the form $\sigma^{-k_\varepsilon - |\mathtt{i}|}[\mathtt{i}]$ respectively, where $|\mathtt{i}| = n$. The ε-independence of the first partition from the second is immediate and we conclude that $\mathcal{P}$ is a weak Bernoulli partition for the transformation $\hat{\sigma}$ of $(\hat{\Sigma}_N, \overline{\mathcal{B}}_{\hat{\mu}}, \hat{\mu})$. Applying the theorem of Friedman and Ornstein proves the second assertion of Theorem 1.

7.2. **Proof of Theorem 2.**
Before starting the proof we require the following simple lemma:

Lemma 7.1. *Let* $T \colon X \to X$ *be an ergodic measure-preserving transformation of a probability space* $(X, \mathcal{F}, \mu)$ *which is not totally ergodic, and let* $n > 1$ *be the smallest integer such that* T^n *is not ergodic. Then there exists a measurable set* $Z \subset X$ *such that* $Z, T^{-1}Z, \ldots, T^{-(n-1)}Z$ *partitions* X *up to measure zero and satisfies* $T^{-n}Z = Z$ *up to measure zero.*

Proof. Since T^n is not ergodic there by definition exists a measurable set $Y \subset X$ such that $0 < \mu(Y) < 1$ and $T^{-n}Y = Y$ up to μ-measure zero. Consider a set $R \subseteq \{0, \ldots, n-1\}$ with the properties $0 \in R$ and $\mu(\bigcap_{i \in R} T^{-i}Y) > 0$ and which has maximum cardinality of all such sets. (Clearly at least one set with those two properties exists, namely $\{0\}$, so R is well-defined.) Define $Z := \bigcap_{i \in R} T^{-i}Y$ and note that clearly $T^{-n}Z = Z$ up to μ-measure zero. For each $i \in \{1, \ldots, n-1\}$ we must have either $\mu(Z \cap T^{-i}Z) = 0$ or $R = R + i \mod n$, since if neither of these holds then $R \cup (R + i \mod n)$ would have larger cardinality than R while having the same

characteristic properties, contradicting maximality. But if $R = R + i \mod n$ then $T^{-i}Z = Z$ up to measure zero which implies that T^i is not ergodic, contradicting the definition of n. It follows that $\mu(Z \cap T^{-i}Z) = 0$ for all $i = 1, \ldots, n-1$ and hence by the T-invariance of μ we deduce that $T^{-i}Z$ and $T^{-j}Z$ are pairwise disjoint up to measure zero whenever $0 \le i < j < n$. We also have $\bigcup_{i=0}^{n-1} T^{-i}Z = X$ up to measure zero since this set is T-invariant and has positive measure and T is ergodic with respect to μ, and this completes the proof. $\qquad\square$

The core of the proof of Theorem 2 is contained in the following result, which will be used twice in the proof.

Proposition 7.2. *Let $k \ge 1$ and $N \ge 2$. For each $j = 1, \ldots, k$ let V_j be a finite-dimensional real vector space and let $(A_1^{(j)}, \ldots, A_N^{(j)}) \in \mathrm{GL}(V_j)^N$ and $\beta_j > 0$. For all $\mathtt{i} \in \Sigma_N^*$ define*

$$\Phi(\mathtt{i}) := \prod_{j=1}^{k} \left\| A_{\mathtt{i}}^{(j)} \right\|^{\beta_j}$$

and let μ be an ergodic equilibrium state of Φ.

Suppose that there exist an integer $n > 1$ and Borel set $Z \subset \Sigma_N$ such that $T^{-n}Z = Z$ and $Z, T^{-1}Z, \ldots, T^{-(n-1)}Z$ is a partition of Σ_N, both up to μ-measure zero. Define a measure ν on Σ_N by $\nu(A) := \mu(A \cap Z)/\mu(Z)$ for all Borel sets $A \subseteq \Sigma_N$. Let $\eta \colon \{\mathtt{i} \in \Sigma_N^* \colon |\mathtt{i}| = n\} \to \{1, \ldots, N^n\}$ be the map which takes each word $\mathtt{i} \in \Sigma_N^*$ of length n to the integer representing its position in the lexicographical ordering on $\{\mathtt{i} \in \Sigma_N^* \colon |\mathtt{i}| = n\}$, and define a homeomorphism $\iota \colon \Sigma_N \to \Sigma_{N^n}$ by $\iota[(x_\ell)_{\ell=1}^\infty] := (\eta(x_{(q-1)n+1} \cdots x_{qn}))_{q=1}^\infty$. For each $j = 1, \ldots, k$ define an N^n-tuple $(B_1^{(j)}, \ldots, B_{N^n}^{(j)}) \in \mathrm{GL}(V_j)^{N^n}$ by $B_i^{(j)} := A_{\eta^{-1}(i)}^{(j)}$ for every $i = 1, \ldots, N^n$ and $j = 1, \ldots, k$, and define a potential $\Psi \colon \Sigma_{N^n}^* \to (0, +\infty)$ by

$$\Psi(\mathtt{j}) = \prod_{j=1}^{k} \left\| B_{\mathtt{j}}^{(j)} \right\|^{\beta_j}$$

for all $\mathtt{j} \in \Sigma_{N^n}^*$.

Then $\mu = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_*^i \nu$, each measure $\sigma_*^i \nu$ is σ^n-invariant, the measures $(\iota \circ \sigma^i)_* \nu \in \mathcal{M}_\sigma(\Sigma_{N^n})$ are pairwise mutually singular equilibrium states of Ψ, and $n \le \prod_{j=1}^{k} \dim V_j$.

Proof. Clearly the properties of Z imply $\mu(Z) = 1/n$. For every Borel set $A \subseteq \Sigma_N$ we have

$$\frac{1}{n} \sum_{i=0}^{n-1} (\sigma_*^i \nu)(A) = \frac{1}{n} \sum_{i=0}^{n-1} \frac{\mu(\sigma^{-i}A \cap Z)}{\mu(Z)} = \sum_{i=0}^{n-1} \mu(\sigma^{-i}A \cap Z)$$

$$= \sum_{i=0}^{n-1} \mu(A \cap \sigma^{-(n-i)}Z) = \mu(A)$$

so that $\mu = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_*^i \nu$, and similarly

$$(\sigma_*^i \nu)(\sigma^{-n}A) = \frac{\mu(\sigma^{-i-n}A \cap Z)}{\mu(Z)} = \frac{\mu(\sigma^{-i-n}A \cap \sigma^{-n}Z)}{\mu(Z)} = \frac{\mu(\sigma^{-i}A \cap Z)}{\mu(Z)} = (\sigma_*^i \nu)(A)$$

so that each $\sigma_*^i \nu$ is σ^n-invariant. The equation $\iota \circ \sigma^n = \sigma \circ \iota$ is obvious from the definition of ι. It follows directly that each $(\iota \circ \sigma^i)_* \nu$ is a σ-invariant measure

on Σ_{N^n}. It is easy to check from the construction of ν that the sets which have nonzero measure with respect to $\sigma_*^i \nu$ are precisely those which intersect $\sigma^{-(n-i)} Z$ in a set of nonzero μ-measure, and since the sets $Z, \sigma^{-1} Z, \ldots, \sigma^{-(n-1)} Z$ are pairwise disjoint up to μ-measure zero this implies that the measures $\sigma_*^i \nu$ are pairwise mutually singular for distinct $i \in \{0, \ldots, n-1\}$. Since ι is a homeomorphism this implies that the measures $(\iota \circ \sigma^i)_* \nu$ are pairwise mutually singular for distinct $i \in \{0, \ldots, n-1\}$. By Theorem 4 there can be at most $\prod_{j=1}^k \dim V_j$ distinct ergodic equilibrium states for Ψ, which are necessarily pairwise mutually singular since they are distinct ergodic measures. By standard ergodic decomposition arguments every equilibrium state of Ψ arises as a convex combination of these ergodic equilibrium states. It follows from this that the cardinality of a set of pairwise mutually singular equilibrium states of Ψ cannot be larger than the cardinality of the set of ergodic equilibrium states of Ψ, which is bounded by $\prod_{j=1}^k \dim V_j$. Thus if we can show that every $(\iota \circ \sigma^i)_* \nu$ is an equilibrium state of Ψ then the bound $n \leq \prod_{j=1}^k \dim V_j$ follows and we will have proved the proposition. But it follows easily from the definition of equilibrium state and the fact that $h(\cdot)$ and $\Lambda(\Psi, \cdot)$ are affine functions that if a finite convex combination of invariant measures is an equilibrium state of Ψ, then so must be the measures which are the summands in the convex combination. So to complete the proof we need only show that the invariant measure $\iota_* \mu = \frac{1}{n} \sum_{i=0}^{n-1} (\iota \circ \sigma^i)_* \nu \in \mathcal{M}_\sigma(\Sigma_{N^n})$ is an equilibrium state of Ψ. But this is a straightforward calculation: the equation $\iota \circ \sigma^n = \sigma \circ \iota$ and the fact that ι is a homeomorphism together imply that $h(\iota_* \mu) = n h(\mu)$ by basic ergodic theory, and we have

$$P(\Psi) = \lim_{m \to \infty} \frac{1}{m} \log \sum_{\mathbf{j} \in \Sigma_{N^n}^* \,:\, |\mathbf{j}| = m} \Psi(\mathbf{i}) = \lim_{m \to \infty} \frac{1}{m} \log \sum_{\mathbf{j} \in \Sigma_N^* \,:\, |\mathbf{j}| = nm} \Phi(\mathbf{i}) = n P(\Phi)$$

and $\Lambda(\Psi, \iota_* \mu) = n \Lambda(\Phi, \mu)$ by an almost identical calculation. The result follows. $\square$

We may now prove Theorem 2. Let k, N, V_j, $(A_1^{(j)}, \ldots, A_N^{(j)})$, β_j and μ be as in the statement of the theorem. Since μ is not totally ergodic, it follows from Lemma 7.1 that there exists an integer $n > 1$ with the property that there exists a measurable set $Z \subset X$ such that $Z, T^{-1} Z, \ldots, T^{-(n-1)} Z$ partitions X up to μ-measure zero and such that $T^{-n} Z = Z$ up to measure zero. Proposition 7.2 implies that every integer n with this property is less than or equal to $\prod_{j=1}^k \dim V_j$, so we may choose a *largest* such integer. Let n be the largest integer with the aforementioned property, which clearly satisfies $1 < n \leq \prod_{j=1}^k \dim V_j$. Let $Z \subset \Sigma_N$ be a Borel set with the property that $Z, T^{-1} Z, \ldots, T^{-(n-1)} Z$ partitions Σ_N up to μ-measure zero and define a Borel probability measure ν on Σ_N by $\nu(A) := \mu(A \cap Z) / \mu(Z)$ for all Borel sets $A \subseteq \Sigma_N$. By Proposition 7.2 there exist tuples $(B_1^{(j)}, \ldots, B_{N^n}^{(j)}) \in \mathrm{GL}(V_j)^{N^n}$ and a potential $\Psi \colon \Sigma_{N^n}^* \to (0, +\infty)$ as in the statement of Theorem 2 such that each of the measures $(\iota \circ \sigma^i)_* \nu \in \mathcal{M}_\sigma(\Sigma_{N^n})$ is a distinct equilibrium state of Ψ, and we have $\mu = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_*^i \nu$.

To complete the proof of Theorem 2 we must show that for every $i \in \{0, \ldots, n-1\}$ the measure $(\iota \circ \sigma^i)_* \nu$ is totally ergodic. Fix such an i. We will first show that $(\iota \circ \sigma^i)_* \nu$ is ergodic. If this is not the case then there exists a Borel set $A \subset \Sigma_{N^n}$ such that $\sigma^{-1} A = A$ up to $(\iota \circ \sigma^i)_* \nu$-measure zero and such that $0 < ((\iota \circ \sigma^i)_* \nu)(A) < 1$. Define $B := \iota^{-1} A \subset \Sigma_N$ so that $B = \iota^{-1} A = \iota^{-1} \sigma^{-1} A = \sigma^{-n} \iota^{-1} A = \sigma^{-n} B$

up to $\sigma_*^i \nu$-measure zero, and $0 < (\sigma_*^i \nu)(B) < 1$. We have $0 < \mu(\sigma^{-i}B \cap Z) < \mu(Z) = \frac{1}{n}$ and $\sigma^{-n}(\sigma^{-i}B \cap Z) = \sigma^{-i}B \cap Z$ up to μ-measure zero. But then $\bigcup_{j=0}^{n-1} \sigma^{-j}(\sigma^{-i}B \cap Z)$ is σ-invariant up to μ-measure zero but has measure strictly between 0 and 1, contradicting the ergodicity of μ.

We may now show that $(\iota \circ \sigma^i)_* \nu$ is totally ergodic. If it is not then by Lemma 7.1 there exist an integer $m > 1$ and a Borel subset A of Σ_{N^n} such that $\sigma^{-m}A = A$ up to $(\iota \circ \sigma^i)_* \nu$-measure zero and such that $A, \sigma^{-1}A, \ldots, \sigma^{-(m-1)}A$ forms a partition of Σ_{N^n} up to $(\iota \circ \sigma^i)_* \nu$-measure zero. Define $B := \iota^{-1}A \subset \Sigma_N$; then $B, \sigma^{-n}B, \sigma^{-2n}B \ldots, \sigma^{-(m-1)n}B$ forms a partition of Σ_N up to $\sigma_*^i \nu$-measure zero, and $B = \sigma^{-mn}B$ up to $\sigma_*^i \nu$-measure zero. This implies that the sets $\sigma^{-i}B, \sigma^{-n-i}B, \sigma^{-2n-i}B \ldots, \sigma^{-(m-1)n-i}B$ form a partition of Z up to μ-measure zero and are all σ^{-mn}-invariant up to μ-measure zero. But then

$$\bigcup_{\ell=0}^{mn-1} \sigma^{-\ell}B = \bigcup_{j=0}^{n-1} \sigma^{-j} \left(\bigcup_{r=0}^{m-1} \sigma^{-i-rn}B \right) = \bigcup_{j=0}^{n-1} \sigma^{-j}Z = \Sigma_N$$

up to μ-measure zero, and all of these unions are disjoint up to μ-measure zero, which contradicts the maximality of n. This completes the proof that each $(\iota \circ \sigma^i)_* \nu$ is totally ergodic and completes the proof of the theorem.

7.3. **Proof of Corollary 2.1.** Let μ be an ergodic generalised matrix equilibrium state on Σ_N. If μ is totally ergodic then the conclusion follows by Theorem 1, so suppose that μ is not totally ergodic. By Theorem 2 there exist $n > 1$ and a σ^n-invariant measure ν on Σ_N, which is totally ergodic with respect to σ^n and is measurably isomorphic via a homeomorphism $\iota \colon \Sigma_N \to \Sigma_{N^n}$ satisfying $\iota \circ \sigma^n = \sigma \circ \iota$ to a generalised matrix equilibrium state on Σ_{N^n}, such that $\mu = \frac{1}{n} \sum_{i=0}^{n-1} \sigma_*^i \nu$. In particular we may write $\hat{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} \hat{\sigma}_*^i \hat{\nu}$ where each $\hat{\sigma}_*^i \hat{\nu}$ is a distinct ergodic measure with respect to the transformation $\hat{\sigma}^n$ and where $\hat{\nu}$ has the Bernoulli property with respect to the transformation $\hat{\sigma}^n$ as a consequence of Theorem 1. Since the measures $\hat{\sigma}_*^i \hat{\nu}$ are distinct ergodic measures they are pairwise mutually singular, so there exists $Z \subset \hat{\Sigma}_N$ such that $\hat{\nu}(Z) = 1$ and $\hat{\nu}(\hat{\sigma}^i Z) = 0$ for all $i \in \{1, \ldots, n-1\}$, and this set satisfies $\hat{\sigma}^n Z = Z$ up to $\hat{\mu}$-measure zero by the $\hat{\sigma}^n$-invariance of the measure $\hat{\nu}$. It follows that $\hat{\Sigma}_N = Z \cup \hat{\sigma}^i Z \cup \cdots \cup \hat{\sigma}^{n-1}Z$ up to $\hat{\mu}$-measure zero and that these sets are pairwise disjoint up to $\hat{\mu}$-measure zero. By virtue of the equation $\hat{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} \hat{\sigma}_*^i \hat{\nu}$, the measure $\hat{\nu}$ must be precisely the measure $\hat{\mu}_Z$ on Z defined by $\hat{\mu}_Z(A) := \hat{\mu}(A \cap Z)/\hat{\mu}(Z)$ for all Borel sets $A \subseteq \hat{\Sigma}_N$.

Let $\mathbb{P}$ be a Bernoulli measure on $\hat{\Sigma}_N$ which has the same entropy as $\hat{\mu}$. Let $\overline{\hat{\mathcal{B}}_{\hat{\nu}}}$ and $\overline{\hat{\mathcal{B}}_{\mathbb{P}}}$ denote the completion of the Borel σ-algebra on $\hat{\Sigma}_N$ with respect to the measures $\hat{\nu}$ and $\mathbb{P}$ respectively. Since $\hat{\nu}$ and $\mathbb{P}$ both have the Bernoulli property with respect to $\hat{\sigma}^n$, and both have the same entropy as $\hat{\mu}$ with respect to $\hat{\sigma}^n$, they are measurably isomorphic, so there exists a measure space isomorphism ϕ from $(\hat{\Sigma}_N, \overline{\hat{\mathcal{B}}_{\hat{\nu}}}, \hat{\nu})$ to $(\hat{\Sigma}_N, \overline{\hat{\mathcal{B}}_{\mathbb{P}}}, \mathbb{P})$ such that $\phi \circ \hat{\sigma}^n = \hat{\sigma}^n \circ \phi$ and $\phi_* \hat{\nu} = \mathbb{P}$.

Let $\mathbb{Z}_n$ denote the set $\{0, \ldots, n-1\}$ equipped with addition modulo n and define a transformation $T \colon \hat{\Sigma}_N \times \mathbb{Z}_n \to \hat{\Sigma}_N \times \mathbb{Z}_n$ by $T(x, i) := (\hat{\sigma}x, i+1 \mod n)$. To prove the corollary we must construct a measure space isomorphism ψ from $\hat{\Sigma}_N$ to $\hat{\Sigma}_N \times \mathbb{Z}_n$ which satisfies $\psi \circ \hat{\sigma} = T \circ \psi$ and $\psi_* \hat{\mu} = \mathbb{P} \times (\frac{1}{n} \sum_{i=0}^{n-1} \delta_i)$. To this end define $\psi \colon \hat{\Sigma}_N \to \hat{\Sigma}_N \times \mathbb{Z}_n$ by $\psi(x) = (\hat{\sigma}^i \phi(\hat{\sigma}^{-i}x), i)$ whenever $x \in \hat{\sigma}^i Z$ for some $i \in \mathbb{Z}_n$, and define $\psi(x)$ to be an arbitrary constant value in $\hat{\Sigma}_N \times \mathbb{Z}_n$ for all

$x \in \hat{\Sigma}_N \setminus \bigcup_{i=0}^{n-1} \hat{\sigma}^i Z$. When $x \in \hat{\sigma}^i Z$ for $i \in \{0, \ldots, n-2\}$ we have

$$T(\psi(x)) = (\hat{\sigma}^{i+1}\phi(\hat{\sigma}^{-i}x), i+1) = \psi(\hat{\sigma}x)$$

and when $x \in \hat{\sigma}^{n-1}$ we have

$$T(\psi(x)) = (\hat{\sigma}^n \phi(\hat{\sigma}^{-(n-1)}x), 0) = (\phi(\hat{\sigma}x), 0) = \psi(\hat{\sigma}x)$$

so that $T \circ \psi = \psi \circ \hat{\sigma}$ almost everywhere with respect to $\hat{\mu}$. By construction we have $\phi_* \hat{\mu}_Z = \mathbb{P}$ and consequently $\psi_* \hat{\nu} = \mathbb{P} \times \delta_0$. It follows directly that

$$\psi_* \mu = \psi_* \left(\frac{1}{n} \sum_{i=0}^{n-1} \hat{\sigma}^i_* \mu_Z \right) = \frac{1}{n} \sum_{i=0}^{n-1} T_*^i (\mathbb{P} \times \delta_0) = \mathbb{P} \times \left(\frac{1}{n} \sum_{i=0}^{n-1} \delta_i \right)$$

as required. This completes the construction of the isomorphism ψ and proves the corollary.

7.4. **Proof of Proposition 2.3.** Fix $\alpha, \beta > 0$ throughout the proof. It is obvious that (A_1, A_2) and (B_1, B_2) are irreducible since every nonzero proper subspace of $\mathbb{R}^2$ is one-dimensional but neither A_2 nor B_1 preserves any one-dimensional subspace of $\mathbb{R}^2$. On the other hand it is obvious that the horizontal and vertical axes in $\mathbb{R}^2$ both have finite orbit under the action of (A_1, A_2) and similarly for (B_1, B_2). If $\mu \in \mathcal{M}_\sigma(\Sigma_2)$ is an ergodic equilibrium state of the potential $\Phi \colon \Sigma_2^* \to (0, +\infty)$ defined by

$$\Phi(\mathtt{i}) := \|A_\mathtt{i}\|^\alpha \|B_\mathtt{i}\|^\beta$$

then it follows by Theorem 3 and the preceding observations that there exists a transitive subspace class $\mathcal{W} \subset \mathrm{Gr}_1(\mathbb{R}^2) \times \mathrm{Gr}_1(\mathbb{R}^2)$ such that μ is the unique equilibrium state of the potential

$$\Phi_\mathcal{W}(\mathtt{i}) := \max_{(W_1, W_2) \in \mathcal{W}} \|A_\mathtt{i}|_{W_1}\|^\alpha \|B_\mathtt{i}|_{W_2}\|^\beta .$$

We claim that there exists a *unique* transitive subspace class preserved by these pairs of matrices, which is the set

$$\mathcal{W}_0 := \{(\overline{e_1}, \overline{e_1}), (\overline{e_1}, \overline{e_2}), (\overline{e_2}, \overline{e_1}), (\overline{e_2}, \overline{e_2})\}.$$

(Here e_1, e_2 denotes the standard basis for $\mathbb{R}^2$ and $\overline{u}$ the one-dimensional subspace spanned by the nonzero vector u.) Indeed, if $\overline{u} \subset \mathbb{R}^2$ is a one-dimensional space with finite orbit under (A_1, A_2) then the set $\{\overline{A_1^n u} \colon n \geq 1\}$ must be finite, but this is the case only when $\overline{u} \in \{\overline{e_1}, \overline{e_2}\}$; similarly $\{\overline{B_2^n u} \colon n \geq 1\}$ is finite only when $\overline{u} \in \{\overline{e_1}, \overline{e_2}\}$; we conclude that every equivariant subspace class must be a subset of $\mathcal{W}_0$. On the other hand it is easy to see that $\mathcal{W}_0$ is transitive: we may apply the symbol 1 to pass from $(\overline{e_1}, \overline{e_1})$ to $(\overline{e_1}, \overline{e_2})$ and vice versa, or from $(\overline{e_2}, \overline{e_1})$ to $(\overline{e_2}, \overline{e_2})$ and vice versa, and we may apply the symbol 2 to pass from $(\overline{e_1}, \overline{e_1})$ to $(\overline{e_2}, \overline{e_1})$ and vice versa, or from $(\overline{e_1}, \overline{e_2})$ to $(\overline{e_2}, \overline{e_2})$ and vice versa. Thus every element of $\mathcal{W}_0$ can be reached from any other element via a word of length 1 or 2. We conclude that it contains a unique transitive subspace class, which is $\mathcal{W}_0$ itself. It follows that the potential Φ above has a unique equilibrium state, namely the unique equilibrium state of the potential $\Phi_{\mathcal{W}_0}$ as defined above. Let us denote this unique equilibrium state by μ. We wish to show that μ is not totally ergodic.

Define tuples

$$(\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{A}_4) := (A_1 A_1, A_1 A_2, A_2 A_1, A_2 A_2),$$
$$(\hat{B}_1, \hat{B}_2, \hat{B}_3, \hat{B}_4) := (B_1 B_1, B_1 B_2, B_2 B_1, B_2 B_2)$$

so that

$$\hat{A}_1 = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad \hat{A}_2 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \quad \hat{A}_3 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \quad \hat{A}_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\hat{B}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \hat{B}_2 = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \quad \hat{B}_3 = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \quad \hat{B}_4 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

and define a potential $\hat{\Phi} \colon \Sigma_4^* \to (0, +\infty)$ by

$$\hat{\Phi}(\mathtt{i}) := \left\| \hat{A}_{\mathtt{i}} \right\|^{\alpha} \left\| \hat{B}_{\mathtt{i}} \right\|^{\beta}.$$

By arguments similar to that used in Theorem 2 we may define a recoding homeo-morphism $\iota \colon \Sigma_2 \to \Sigma_4$ such that $\iota \circ \sigma^2 = \sigma \circ \iota$ and such that $\hat{\Phi}(\iota(x)|_n) = \Phi(x|_{2n})$ for every $x \in \Sigma_2$. Easy calculations similar to those occurring in the proof of Proposition 7.2 show directly that $P(\hat{\Phi}) = 2P(\Phi)$ and that $\iota_* \mu$ is an equilibrium state of $\hat{\Phi}$. To prove that μ is not totally ergodic we will show that $\iota_* \mu$ is not ergodic with respect to $\sigma \colon \Sigma_4 \to \Sigma_4$, which combined with the identity $\iota \circ \sigma^2 = \sigma \circ \iota$ implies immediately that μ is not ergodic with respect to $\sigma^2 \colon \Sigma_2 \to \Sigma_2$. Since $\iota_* \mu$ is an equilibrium state of $\hat{\Phi}$ it will suffice for us to identify the ergodic equilibrium states of $\hat{\Phi}$ and show that $\iota_* \mu$ cannot be equal to any of them.

Define

$$\mathcal{W}_1 := \{ (\overline{e_1}, \overline{e_1}), (\overline{e_2}, \overline{e_2}) \},$$
$$\mathcal{W}_2 := \{ (\overline{e_1}, \overline{e_2}), (\overline{e_2}, \overline{e_1}) \}.$$

We observe that both $\mathcal{W}_1$ and $\mathcal{W}_2$ are transitive subspace classes for $(\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{A}_4)$ and $(\hat{B}_1, \hat{B}_2, \hat{B}_3, \hat{B}_4)$: the symbols 1 and 4 fix every pair $(\overline{e_i}, \overline{e_j})$ and the symbols 2 and 3 swap $(\overline{e_1}, \overline{e_1})$ with $(\overline{e_2}, \overline{e_2})$ and swap $(\overline{e_1}, \overline{e_2})$ with $(\overline{e_2}, \overline{e_1})$. Similarly to our analysis of Φ, since the co-ordinate axes are the only one-dimensional sub-spaces which have finite orbit under $\hat{A}_1$ and $\hat{B}_4$, every transitive subspace class for $(\hat{A}_1, \hat{A}_2, \hat{A}_3, \hat{A}_4)$ and $(\hat{B}_1, \hat{B}_2, \hat{B}_3, \hat{B}_4)$ must be a subset of $\mathcal{W}_0$. It follows that there exist exactly *two* transitive subspace classes for these tuples, $\mathcal{W}_1$ and $\mathcal{W}_2$. Define potentials $\hat{\Phi}_1, \hat{\Phi}_2 \colon \Sigma_4^* \to (0, +\infty)$ by

$$\hat{\Phi}_i(\mathtt{i}) := \max_{(W_1, W_2) \in \mathcal{W}_i} \left\| \hat{A}_{\mathtt{i}}|_{W_1} \right\|^{\alpha} \left\| \hat{B}_{\mathtt{i}}|_{W_2} \right\|^{\beta}.$$

for $i = 1, 2$. By Theorem 3 each of $\hat{\Phi}_1$ and $\hat{\Phi}_2$ has a unique equilibrium state which we denote respectively μ_1 and μ_2, and furthermore every ergodic equilibrium state of $\hat{\Phi}$ must be equal to one of these two measures. In particular at least one of the two measures is an equilibrium state for $\hat{\Phi}$. Straightforward checking of definitions demonstrates that for $i = 1, 2$ the measure μ_i is an equilibrium state of $\hat{\Phi}$ if and only if $P(\hat{\Phi}_i) = P(\hat{\Phi})$.

Suppose for a contradiction that μ is totally ergodic. In particular μ is ergodic with respect to $\sigma^2 \colon \Sigma_2 \to \Sigma_2$ and therefore $\iota_* \mu$ is ergodic with respect to $\sigma \colon \Sigma_4 \to \Sigma_4$. Hence $\iota_* \mu$ is an ergodic equilibrium state of $P(\hat{\Phi})$ and there exists $i \in \{1, 2\}$ such that $\iota_* \mu = \mu_i$ and $P(\hat{\Phi}_i) = P(\hat{\Phi})$. It follows from Theorem 3 that there exists $C_1 > 0$ such that

$$C_1^{-1} e^{-nP(\Phi)} \Phi(x|_n) \le \mu([x|_n]) \le C_1 e^{-nP(\Phi)} \Phi(x|_n)$$

for every $n \ge 1$ and $x \in \Sigma_2$, and also that there exists $C_2 > 0$ such that

$$C_2^{-1} e^{-nP(\hat{\Phi}_i)} \hat{\Phi}_i(z|_n) \le \mu_i([z|_n]) \le C_2 e^{-nP(\hat{\Phi}_i)} \hat{\Phi}_i(z|_n)$$

for every $z \in \Sigma_4$. Recoding via ι, the former inequalities imply

$$C_1^{-1} e^{-nP(\hat{\Phi})} \hat{\Phi}(z|_n) \leq (\iota_* \mu) \left([z|_n]\right) \leq C_1 e^{-nP(\hat{\Phi})} \hat{\Phi}(z|_n)$$

for every $z \in \Sigma_4$ and $n \geq 1$, and since $\iota_* \mu = \mu_i$ and $P(\hat{\Phi}_i) = P(\hat{\Phi})$ by hypothesis we conclude that necessarily

(28) $$C_1^{-1} C_2^{-1} \leq \frac{\hat{\Phi}(\mathtt{i})}{\hat{\Phi}_i(\mathtt{i})} \leq C_1 C_2$$

for every $\mathtt{i} \in \Sigma_4^*$. But if $\mathtt{i} = 1^n 4^n$ then

$$\hat{A}_{\mathtt{i}} = \hat{A}_1^n \hat{A}_4^n = A_1^{2n} A_2^{2n} = \begin{pmatrix} 4^n & 0 \\ 0 & 1 \end{pmatrix},$$

$$\hat{B}_{\mathtt{i}} = \hat{B}_1^n \hat{B}_4^n = B_1^{2n} B_2^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 4^n \end{pmatrix}$$

and $\hat{\Phi}_1(\mathtt{i}) = \max\{4^{n\alpha}, 4^{n\beta}\}$ whereas $\hat{\Phi}(\mathtt{i}) = 4^{n(\alpha+\beta)}$, so (28) cannot hold for $i = 1$; and if $\mathtt{j} = 31^{n-1} 24^{n-1}$ for some $n \geq 1$ then

$$\hat{A}_{\mathtt{j}} = \hat{A}_3 \hat{A}_1^{n-1} \hat{A}_2 \hat{A}_4^{n-1} = (A_2 A_1)(A_1 A_1)^{n-1}(A_1 A_2)(A_2 A_2)^{n-1}$$
$$= A_2 A_1^{2n} A_2^{2n-1} = \begin{pmatrix} 1 & 0 \\ 0 & 4^n \end{pmatrix}$$

and

$$\hat{B}_{\mathtt{j}} = \hat{B}_3 \hat{B}_1^{n-1} \hat{B}_2 \hat{B}_4^{n-1} = (B_2 B_1)(B_1 B_1)^{n-1}(B_1 B_2)(B_2 B_2)^{n-1}$$
$$= B_2 B_1^{2n} B_2^{2n-1} = \begin{pmatrix} 1 & 0 \\ 0 & 4^n \end{pmatrix}$$

so that $\hat{\Phi}_2(\mathtt{j}) = \max\{4^{n\alpha}, 4^{n\beta}\}$, but clearly we have $\hat{\Phi}(\mathtt{j}) = 4^{n(\alpha+\beta)}$. It follows that (28) also cannot hold for $i = 2$. We conclude that neither μ_1 nor μ_2 can be equal to $\iota_* \mu$, and since this exhausts the ergodic equilibrium states of $\hat{\Phi}$ the equilibrium state $\iota_* \mu$ cannot be ergodic, so μ cannot be ergodic with respect to σ^2 and hence is not totally ergodic as required. This completes the proof.

Remark. If $\hat{\Phi}$ had a unique ergodic equilibrium state then it would have to be equal to either μ_1 or μ_2 and also to $\iota_* \mu$, which has been shown to be impossible, so $\hat{\Phi}$ cannot have a unique ergodic equilibrium state. By elimination the only possibility is that *both* of μ_1 and μ_2 are equilibrium states and that $\iota_* \mu$ is equal to a strict linear combination of these two mutually singular measures. On the other hand Theorem 2 implies that μ must be equal to a *balanced* linear combination of two mutually singular σ^2-invariant measures, and we conclude that necessarily $\iota_* \mu = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$.

8. Acknowledgments

References

[1] ABELS, H., MARGULIS, G. A., AND SOĬFER, G. A. Semigroups containing proximal linear maps. *Israel J. Math. 91*, 1-3 (1995), 1–30.

[2] BALADI, V. *Positive transfer operators and decay of correlations*, vol. 16 of *Advanced Series in Nonlinear Dynamics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[3] BARAŃSKI, K. Hausdorff dimension of the limit sets of some planar geometric constructions. *Adv. Math. 210*, 1 (2007), 215–245.

[4] BÁRÁNY, B., HOCHMAN, M., AND RAPAPORT, A. Hausdorff dimension of planar self-affine sets and measures. *Invent. Math. 216*, 3 (2019), 601–659.

[5] BÁRÁNY, B., KÄENMÄKI, A., AND MORRIS, I. D. Domination, almost additivity and thermodynamical formalism for planar matrix cocycles. *Israel J. Math. 239*, 1 (2020), 173–214.

[6] BARRAL, J., AND FENG, D.-J. Non-uniqueness of ergodic measures with full Hausdorff dimensions on a Gatzouras-Lalley carpet. *Nonlinearity 24*, 9 (2011), 2563–2567.

[7] BARRAL, J., AND FENG, D.-J. Multifractal formalism for almost all self-affine measures. *Comm. Math. Phys. 318*, 2 (2013), 473–504.

[8] BEDFORD, T. *Crinkly curves, Markov partitions and dimension*. 1984. Thesis (Ph.D.)–The University of Warwick.

[9] BENOIST, Y., AND QUINT, J.-F. *Random walks on reductive groups*, vol. 62 of *Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics*. Springer, Cham, 2016.

[10] BOCHI, J., AND MORRIS, I. D. Equilibrium states of generalised singular value potentials and applications to affine iterated function systems. *Geom. Funct. Anal. 28*, 4 (2018), 995–1028.

[11] BONATTI, C., AND VIANA, M. Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergodic Theory Dynam. Systems 24*, 5 (2004), 1295–1330.

[12] BOREL, A. *Linear algebraic groups*, second ed., vol. 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[13] BOWEN, R. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, 1975.

[14] BRADLEY, R. C. Basic properties of strong mixing conditions: a survey and some open questions. *Probab. Surv. 2* (2005), 107–144.

[15] BUTLER, C., AND PARK, K. Thermodynamic formalism of $GL_2(\mathbb{R})$-cocycles with canonical holonomies. *Discrete Contin. Dyn. Syst. 41*, 5 (2021), 2141–2166.

[16] CAO, Y.-L., FENG, D.-J., AND HUANG, W. The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst. 20*, 3 (2008), 639–657.

[17] CHEN, J., AND PESIN, Y. Dimension of non-conformal repellers: a survey. *Nonlinearity 23*, 4 (2010), R93–R114.

[18] DAS, T., AND SIMMONS, D. The Hausdorff and dynamical dimensions of self-affine sponges: a dimension gap result. *Invent. Math. 210*, 1 (2017), 85–134.

[19] FALCONER, K., AND SLOAN, A. Continuity of subadditive pressure for self-affine sets. *Real Anal. Exchange 34*, 2 (2009), 413–427.

[20] FALCONER, K. J. The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc. 103*, 2 (1988), 339–350.

[21] FALCONER, K. J. Bounded distortion and dimension for nonconformal repellers. *Math. Proc. Cambridge Philos. Soc. 115*, 2 (1994), 315–334.

[22] FENG, D.-J. Lyapunov exponents for products of matrices and multifractal analysis. I. Positive matrices. *Israel J. Math. 138* (2003), 353–376.

[23] FENG, D.-J. Dimension of invariant measures for affine iterated function systems. Preprint: arXiv:1901.01691, 2019.

[24] FENG, D.-J., AND KÄENMÄKI, A. Equilibrium states of the pressure function for products of matrices. *Discrete Contin. Dyn. Syst. 30*, 3 (2011), 699–708.

[25] FENG, D.-J., AND LAU, K.-S. The pressure function for products of non-negative matrices. *Math. Res. Lett. 9*, 2-3 (2002), 363–378.

[26] FENG, D.-J., AND SHMERKIN, P. Non-conformal repellers and the continuity of pressure for matrix cocycles. *Geom. Funct. Anal. 24*, 4 (2014), 1101–1128.

[27] FRASER, J. M. On the packing dimension of box-like self-affine sets in the plane. *Nonlinearity 25*, 7 (2012), 2075–2092.

[28] Fraser, J. M. Remarks on the analyticity of subadditive pressure for products of triangular matrices. *Monatsh. Math. 177*, 1 (2015), 53–65.

[29] Friedman, N. A., and Ornstein, D. S. On isomorphism of weak Bernoulli transformations. *Advances in Math. 5* (1970), 365–394.

[30] Gatzouras, D., and Peres, Y. Invariant measures of full dimension for some expanding maps. *Ergodic Theory Dynam. Systems 17*, 1 (1997), 147–167.

[31] Horn, R. A., and Johnson, C. R. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[32] Humphreys, J. E. *Linear algebraic groups*. Springer-Verlag, New York-Heidelberg, 1975. Graduate Texts in Mathematics, No. 21.

[33] Järvenpää, E., Järvenpää, M., Li, B., and Stenflo, O. Random affine code tree fractals and Falconer-Sloan condition. *Ergodic Theory Dynam. Systems 36*, 5 (2016), 1516–1533.

[34] Jordan, T., Pollicott, M., and Simon, K. Hausdorff dimension for randomly perturbed self affine attractors. *Comm. Math. Phys. 270*, 2 (2007), 519–544.

[35] Käenmäki, A., and Morris, I. D. Structure of equilibrium states on self-affine sets and strict monotonicity of affinity dimension. *Proc. Lond. Math. Soc. (3) 116*, 4 (2018), 929–956.

[36] Käenmäki, A., and Reeve, H. W. J. Multifractal analysis of Birkhoff averages for typical infinitely generated self-affine sets. *J. Fractal Geom. 1*, 1 (2014), 83–152.

[37] Kenyon, R., and Peres, Y. Measures of full dimension on affine-invariant sets. *Ergodic Theory Dynam. Systems 16*, 2 (1996), 307–323.

[38] Lalley, S. P., and Gatzouras, D. Hausdorff and box dimensions of certain self-affine fractals. *Indiana Univ. Math. J. 41*, 2 (1992), 533–568.

[39] Mac Lane, S., and Birkhoff, G. *Algebra*, third ed. Chelsea Publishing Co., New York, 1988.

[40] McMullen, C. The Hausdorff dimension of general Sierpiński carpets. *Nagoya Math. J. 96* (1984), 1–9.

[41] Morris, I. D. The generalised Berger-Wang formula and the spectral radius of linear cocycles. *J. Funct. Anal. 262*, 3 (2012), 811–824.

[42] Morris, I. D. Ergodic properties of matrix equilibrium states. *Ergodic Theory Dynam. Systems 38*, 6 (2018), 2295–2320.

[43] Morris, I. D. Some observations on Käenmäki measures. *Ann. Acad. Sci. Fenn. Math. 43*, 2 (2018), 945–960.

[44] Morris, I. D. A necessary and sufficient condition for a matrix equilibrium state to be mixing. *Ergodic Theory Dynam. Systems 39*, 8 (2019), 2223–2234.

[45] Morris, I. D., and Sert, Ç. A converse statement to Hutchinson's theorem and a dimension gap for self-affine measures. *Journal of the European Mathematical Society*. To appear. Preprint: arXiv:1909.08532v2.

[46] Morris, I. D., and Sert, Ç. A strongly irreducible affine iterated function system with two invariant measures of maximal dimension. *Ergodic Theory Dynam. Systems*. To appear. Preprint: arXiv:1905.08299v2.

[47] Park, K. Quasi-multiplicativity of typical cocycles. *Comm. Math. Phys. 376*, 3 (2020), 1957–2004.

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

[49] Philipp, W., and Stout, W. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc. 2*, 161 (1975), iv+140.

[50] Piraino, M. The weak Bernoulli property for matrix Gibbs states. *Ergodic Theory Dynam. Systems 40*, 8 (2020), 2219–2238.

[51] Schmeling, J., and Weiss, H. An overview of the dimension theory of dynamical systems. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, vol. 69 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., Providence, RI, 2001, pp. 429–488.

[52] Walters, P. Regularity conditions and Bernoulli properties of equilibrium states and *g*-measures. *J. London Math. Soc. (2) 71*, 2 (2005), 379–396.