STOCHASTIC STABILITY OF LYAPUNOV EXPONENTS AND OSELEDETS SPLITTINGS FOR SEMI-INVERTIBLE MATRIX COCYCLES

GARY FROYLAND, CECILIA GONZÁLEZ-TOKMAN AND ANTHONY QUAS

ABSTRACT. We establish (i) stability of Lyapunov exponents and (ii) convergence in probability of Oseledets spaces for semi-invertible matrix cocycles, subjected to small random perturbations. The first part extends results of Ledrappier and Young [18] to the semi-invertible setting. The second part relies on the study of evolution of subspaces in the Grassmannian, where the analysis developed, based on higher-dimensional Möbius transformations, is likely to be of wider interest.

1. INTRODUCTION

The landmark Oseledets Multiplicative Ergodic Theorem (MET) plays a central role in modern dynamical systems, providing a basis for the study of non-uniformly hyperbolic dynamical systems. Oseledets' theorem has been extended in many ways beyond the original context of products of finite-dimensional matrices, for instance to certain classes of operators on Banach spaces and more abstractly to non-expanding maps of non-positively curved spaces.

The original Oseledets theorem [20] was formulated in both an invertible version (both the base dynamics and the matrices are assumed to be invertible) and a non-invertible version (neither the base dynamics nor the matrices are assumed to be invertible). The conclusion in the non-invertible case is much weaker than in the invertible case: in the invertible version, the theorem gives a *splitting* (that is, a direct sum decomposition) of $\mathbb{R}^d$ into equivariant subspaces, each with a characteristic exponent that is used to order the splitting components from largest to smallest expansion rate; whereas in the non-invertible version, the theorem gives an equivariant *filtration* (that is, a decreasing nested sequence of subspaces) of $\mathbb{R}^d$.

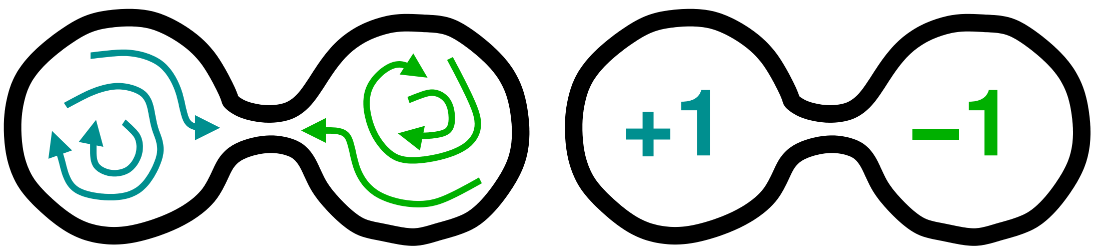
In various combinations, the current authors and collaborators have been working on extensions of the MET to what we have called the *semi-invertible* setting [10, 11, 16]. This refers to the assumption that one has an invertible underlying base dynamical system (also known as driving or forcing), but that the matrices or operators that are composed may fail to be invertible. In this setting, our theorems yield an equivariant splitting as in the invertible case of the MET, rather than the equivariant filtration that the previous theorems would have given.

We are interested in applications where the operators are Perron-Frobenius operators of dynamical systems acting on suitable Banach spaces. Here, the 'suitable' Banach spaces are spaces that are mapped into themselves by the Perron-Frobenius operator, and on which the Perron-Frobenius operator is quasi-compact. These Banach spaces have been widely studied in the case of a single dynamical system.

An Ansatz that first appeared in a paper of Dellnitz, Froyland and Sertl [4] in the context of Perron-Frobenius operators of a single dynamical system is the following:

Ansatz. *The peripheral spectrum (that is, spectrum of the Perron-Frobenius operator outside the essential spectral radius) corresponds to global features of the system (such as bottlenecks or almost-invariant regions) whereas the essential spectrum corresponds to local features of the system, such as rates of expansion.*

In a series of papers, they take this idea further by showing that level sets of eigenfunctions with eigenvalues peripheral to the essential spectral radius can be used to locate almost-invariant sets in the dynamical system [5, 6, 17, 14, 15, 8]. Figure 1 gives a schematic illustration of such a system: The left and right halves are almost-invariant under the dynamics, but the bottleneck joining them allows small but non-negligible interaction between them.



(a) (b)

FIGURE 1. (a) Schematic representation of a dynamical system with almost-invariant regions. (b) Approximate values of eigenfunction corresponding to the bottleneck.

While this Ansatz was initially made in the context of a single dynamical system, it seems to apply equally in the case of random and time-dependent dynamical systems [12, 7, 13], and this is the central motivation for our research in this area. It is well known that Perron-Frobenius operators of non-invertible maps are essentially never invertible, but it is often reasonable to assume that the base dynamics are invertible. Indeed, even if the driving system is non-invertible, one can make use of canonical mathematical techniques to extend it to an invertible one. Hence, we naturally find ourselves in the semi-invertible category. The principal object that we are interested in understanding is the second Oseledets subspace (or more generally the first few Oseledets subspaces).

The significance of our extensions to the MET is that the second subspace that we obtain is low-dimensional (typically one-dimensional) instead of $(d-1)$-dimensional, which is what would come from the standard non-invertible MET. In numerical applications, where d may be 10^5 or greater, one cannot expect to say anything reasonable about level sets of functions belonging to a high-dimensional subspace, whereas using the semi-invertible version of the theorem, we are once again in a position to make sense of the level sets.

In practice, of course, one cannot numerically study the action of Perron-Frobenius operators on infinite-dimensional Banach spaces. Nor can one find a finite-dimensional subspace preserved by the operators. A remarkably fruitful approach is the so-called *Ulam method*. Here, the state space is cut into small reasonably regular pieces and a single dynamical system is treated as a Markov chain, by applying the dynamical system and then randomizing over the cell in which the point lands. This also makes sense for random dynamical systems.

In [9], we showed that applying the Ulam method to certain random dynamical systems, the top Oseledets space of the truncated system converges in probability to the true top Oseledets space of the random dynamical system as the size of the partition is shrunk to 0. The top Oseledets space is known to correspond to the random absolutely continuous invariant measure of the system. It is natural to ask whether the subsequent Oseledets spaces for the truncated systems converge to the corresponding Oseledets spaces for the full system. We are not yet able to answer this, although the current paper represents a substantial step in this direction.

In [9], we viewed the Ulam projections of the Perron-Frobenius operator as perturbations of the original operator, and showed that the top Oseledets space was robust to the kind of projections that were being considered. In this paper, we prove convergence the of subsequent Oseledets spaces under certain perturbations, but do this in the context of matrices instead of infinite-dimensional operators.

A related work in this direction is due to Bogenschütz [3]. The context there is considerably more general, dealing with stability of Oseledets subspaces on Banach spaces rather than just $\mathbb{R}^n$. However, his results only hold under a number of very strong a priori assumptions, including that there is uniform separation between the subspaces - that is, he covers the projectively hyperbolic case, which is known to be robust under all small perturbations, not just the stochastic ones considered in this article. We refer the reader to the article of Bochi and Viana [2] for more information. Our results are established in the finite-dimensional stochastic setting, without any uniformity assumptions on the splitting.

In general, Lyapunov exponents and Oseledets subspaces are known to be highly sensitive to perturbations. A mechanism responsible for this is attributed to Mañé; see also [1]. Ledrappier and Young considered the case of perturbations of random products of uniformly invertible matrices [18][1]. This followed related work of Young in the two-dimensional setting [22]. In view of the sensitivity results, it was necessary to restrict the class of perturbations that they considered, and they dealt with the situation where the distribution of the perturbation of the matrix at time 0 was absolutely continuous (with control on the density) conditioned on all previous perturbations. The simplest instance of this is the case where the matrices to be multiplied are subjected to additive i.i.d. absolutely continuous noise. In this situation, they showed that the perturbed exponents converge almost surely to the true exponents as the noise is shrunk to 0. While they did not directly address the Oseledets subspaces, work of Ochs [19] shows that convergence of the Lyapunov exponents implies convergence in probability of the Oseledets subspaces in the invertible setting.

In this paper, we deal with the case of uniform i.i.d. additive noise in the matrices, but make no assumption on invertibility of the unperturbed matrices. The conclusions that we obtain are the same as may be obtained in the invertible case. Our argument first demonstrates stability of the Lyapunov exponents, and then shows stability of the Oseledets subspaces. The first part is closely based on Ledrappier and Young's approach, although we need to do some non-trivial extra work to deal with the lack of uniform invertibility (in Ledrappier-Young's argument, in one step, there is an upper bound to the amount of damage that can be done to the exponents, whereas in the non-invertible case there is no such bound).

[1] We refer to this situation as *stochastic stability* of the Lyapunov exponents and subspaces, as they change continuously as the matrices are perturbed stochastically, while Ledrappier and Young referred to this just as stability.

The second part of the argument is completely new. The methods of Ochs cannot be made to work here because they are based on finding the space with the smallest exponent and then using exterior powers to move up the ladder. In the case where the smallest exponent is $-\infty$, when one takes exterior powers, all products with this subspace have exponent $-\infty$ so there is no distinguished second-smallest subspace. To get around this problem, we use the Grassmannian in place of the exterior algebra. We study evolution of subspaces in the Grassmannian, and show that this is controlled by fractional linear transformations. An important role is played by a higher-dimensional analogue of the cross ratio.

We are hopeful that the techniques that we introduce to control evolution of these Oseledets subspaces under the matrices may be applied much more widely.

1.1. **Statements of the main results.** If $\sigma\colon (\Omega,\mathbb{P}) \to (\Omega,\mathbb{P})$ is a measure-preserving transformation of a probability space and $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ is a measurable matrix-valued function, we let $A^{(n)}_\omega$ denote the product $A(\sigma^{n-1}\omega)A(\sigma^{n-2}\omega)\cdots A(\omega)$. We call the tuple $(\Omega,\mathbb{P},\sigma,A)$ a *matrix cocycle*. Throughout the article, we shall write $\|\cdot\|$ for the spectral norm of a matrix (that is its operator norm with respect to the Euclidean norm on $\mathbb{R}^d$).

Let $\mathcal{U}$ be the collection of $d\times d$ matrices with entries in $[-1,1]$. We equip $\mathcal{U}$ with the uniform measure, λ, that is volume measure scaled by 2^{-d^2}. Let $\bar\Omega = \Omega\times\mathcal{U}^{\mathbb{Z}}$. We write $\bar\omega = (\omega,\Delta)$ for an element of $\bar\Omega$. and put the measure $\bar{\mathbb{P}} = \mathbb{P}\times\lambda^{\mathbb{Z}}$ on $\bar\Omega$. Given $\epsilon>0$, then for an element $\bar\omega\in\bar\Omega$, the corresponding sequence of matrices is $(A^\epsilon_n(\bar\omega))_{n\in\mathbb{Z}} = (A(\sigma^n\omega)+\epsilon\Delta_n)_{n\in\mathbb{Z}}$. This paper is concerned with a comparison of the properties of the matrix cocycle $(\Omega,\mathbb{P},\sigma,A)$ with those of the matrix cocycle $(\bar\Omega,\bar{\mathbb{P}},\bar\sigma,A^\epsilon)$ as $\epsilon\to 0$.

The main result of this paper is the following.

Theorem 1. *Let σ be an ergodic, invertible measure-preserving transformation of $(\Omega,\mathbb{P})$ and let $A\colon\Omega\to M_{d\times d}(\mathbb{R})$ be a measurable map such that $\int \log^+\|A(\omega)\|\,d\mathbb{P}(\omega)<\infty$.*

Let the Lyapunov exponents of the matrix cocycle be $\lambda_1>\ldots>\lambda_p\ge-\infty$ with multiplicities $d_1,d_2,\ldots,d_p$ and let the corresponding Oseledets decomposition be $\mathbb{R}^d = Y_1(\omega)\oplus\ldots\oplus Y_p(\omega)$.

Let $D_0=0$, $D_i=d_1+\ldots+d_i$ and let the Lyapunov exponents (with multiplicity) be $\infty>\mu_1\ge\mu_2\ge\ldots\ge\mu_d\ge-\infty$, so that $\mu_j=\lambda_i$ if $D_{i-1}<j\le D_i$.

- *(I) (Convergence of Lyapunov exponents) Let the Lyapunov exponents of the perturbed matrix cocycle $(\bar\Omega,\bar P,\bar\sigma,A^\epsilon)$ (with multiplicity) be $\mu_1^\epsilon\ge\mu_2^\epsilon\ge\ldots\ge\mu_d^\epsilon$. Then $\mu_i^\epsilon\to\mu_i$ for each i as $\epsilon\to0$.*
- *(II) (Convergence in probability of Oseledets spaces) Let $\mathcal{N}_i$ be a neighbourhood of μ_i in the extended real line $\bar{\mathbb{R}}$. Let ϵ_0 be such that for each $\epsilon\le\epsilon_0$, $\mu_j^\epsilon\in\mathcal{N}_i$ for each $D_{i-1}<j\le D_j$. For $\epsilon<\epsilon_0$, let $Y_i^\epsilon(\omega)$ denote the sum of the Lyapunov subspaces having exponents in $\mathcal{N}_i$. Then $Y_i^\epsilon(\bar\omega)$ converges in probability to $Y_i(\omega)$ as $\epsilon\to0$.*

1.2. **Outline of the paper.** Section 2 introduces terminology, background results and a collection of lemmas that will be used in the proof of the main result. Theorem 1(I) is established in Section 3, and part (II) is proven in Section 4.

2. Preliminaries

For two subspaces, U and V, of $\mathbb{R}^d$ of the same dimension, we define $\angle(U,V) = d_H(U\cap B, V\cap B)$, where d_H denotes Hausdorff distance and B is the unit ball. For two subspaces U

and W of complementary dimensions, we define $\perp(U,W) = (1/\sqrt{2})\inf_{\{u\in U\cap S,\, w\in W\cap S\}} \|u - w\|$, where S denotes the unit sphere. Thus $\perp(U,W)$ is a measure of complementarity of subspaces, taking values between 0 and 1, with 0 indicating that the spaces intersect and 1 indicating that the spaces are orthogonal complements. Note that $\perp(U,V) \geq \perp(U,W) - \angle(W,V)$.

Let $s_j(A)$ denote the jth singular value of the matrix A and let $\Xi^j(A)$ denote $\log s_1(A) + \ldots + \log s_j(A)$. Note that $\Xi^j(A) = \log \|\Lambda^j A\|$, so that $\Xi^j(AB) \leq \Xi^j(A) + \Xi^j(B)$.

The structure of the proof of the first part of the theorem closely follows that of Ledrappier and Young, in which the orbit of ω is divided into blocks of length $\approx |\log\epsilon|$. These are classified as good if a number of conditions hold (separation of Lyapunov spaces, closeness of averages to integrals etc.) and bad otherwise. The crucial modifications that we make are in estimations for the bad blocks. In the case of [18], the matrices (and hence their perturbations) have uniformly bounded inverses, so that for bad blocks one can give uniform lower bounds on the contribution to the singular value. By contrast, here, there is no uniform lower bound. Upper bounds are straightforward, so all of the work is concerned with establishing lower bounds for the exponents. Absent the invertibility, a similar argument to [18] would yield (random) bounds of order $\log\epsilon$, which turn out to be too weak to give the lower bounds that we need.

It is helpful for notation to assume that we are dealing with an unperturbed system with Lyapunov exponents $\lambda_1 > \ldots > \lambda_p$ with multiplicities $d_1, \ldots, d_p$. We define $D_i = d_1 + \ldots + d_i$. Given a matrix A with the property that $s_{D_i}(A) < s_{D_i+1}(A)$, we define $E_i(A)$ to be the space spanned by the (D_i+1)st to dth singular vectors and $F_i(A)$ to be the space spanned by the *images* of the 1st to D_ith singular vectors under A. If one has a matrix cocycle with base space Ω and matrices A_ω, we use the very similar notation $E_i(\omega)$ and $F_i(\omega)$ to refer to the Oseledets subspaces. The convention will be that if the argument is a matrix, then they refer to the span of the bottom singular vectors or the images of the top singular vectors, while if the argument is a point of the base space, they refer to spaces appearing in the Oseledets theorem. These spaces are obtained simply as limits of spans of singular vectors, as explained in Lemma 2 below, justifying the notation. Lemma 3 collects properties of singular-value decompositions retained under perturbations. Lemma 4 shows how the smallness of the perturbations in the matrix cocycle is used in later arguments. Lemma 5 is a finitary version of Lemma 2.

Lemma 2 (Singular Vectors of blocks of the unperturbed system)**.** *Let σ be an ergodic measurable transformation of Ω and let A be a matrix cocycle with exponents $\infty > \lambda_1 > \ldots > \lambda_p \geq -\infty$ with multiplicities $d_1, \ldots, d_p$. Let $D_i = d_1 + \ldots + d_i$. For almost every ω and each $1 \leq i \leq p$, $E_i(A^{(n)}_\omega) \to E_i(\omega)$ and $F_i(A^{(n)}_{\sigma^{-n}\omega}) \to F_i(\omega)$ as $n \to \infty$.*

Proof. The statement that $E_i(A^{(n)}_\omega) \to E_i(\omega)$ follows from Raghunathan's proof of the Oseledets theorem ([21, Claim I]). The singular value decomposition ensures that $F_i(A^{(n)}_{\sigma^{-n}\omega}) = \left(E_i(A^{(n)*}_{\sigma^{-n}\omega})\right)^\perp$, where A^* denotes the adjoint of A.

On the other hand, a similar statement is true for the spaces $E_i^*(\omega)$ and $F_i(\omega)$. More precisely, we claim that if we let $E_i^*(\omega)$ be the Oseledets spaces for the dual cocycle with base σ^{-1} and generator $G(\omega) = A(\sigma^{-1}\omega)^*$, then $F_i(\omega) = (E_i^*(\omega))^\perp$. Applying Oseledets' theorem to the dual cocycle, we obtain $E_i(A^{(n)*}_{\sigma^{-n}\omega}) \to E_i^*(\omega)$.

To prove the claim, suppose for a contradiction that there exist $f(\omega) \in F_i(\omega)$ and $e^*(\omega) \in E_i^*(\omega)$ of unit length such that $\langle f(\omega), e^*(\omega)\rangle \neq 0$. By invertibility of $A^{(n)}_{\sigma^{-n}\omega}$ as a map from $F_i(\sigma^{-n}\omega)$ to $F_i(\omega)$ there exist for all n, unit vectors $f^{(-n)}(\omega) \in F_i(\sigma^{-n}\omega)$ such that $A^{(n)}_{\sigma^{-n}\omega} f^{(-n)}(\omega)$ is a multiple of $f(\omega)$. Then,

$$\langle A^{(n)}_{\sigma^{-n}\omega} f^{(-n)}(\omega), e^*(\omega)\rangle = \langle f^{(-n)}(\omega), A^{(n)*}_{\sigma^{-n}\omega} e^*(\omega)\rangle. \tag{1}$$

The right hand side grows at a rate slower than λ_i. The left hand side grows at a rate at least λ_i, by [10, Eq. (18)]. This yields a contradiction, so $F_i(\omega) \subset \big(E_i^*(\omega)\big)^\perp$. Since the dimensions agree, they coincide.

Now the statement that $F_i(A^{(n)}_{\sigma^{-n}\omega}) \to F_i(\omega)$ follows directly from the fact that $E_i(A^{(n)*}_{\sigma^{-n}\omega}) \to E_i^*(\omega)$, and continuity of $V \mapsto V^\perp$.

∎

Lemma 3. *[Lemmas 3.3, 3.6 & 4.3, [18]] For any $\delta > 0$, there exists a K such that if (i) the D_ith singular value of a matrix A exceeds K; (ii) the (D_i+1)st singular value at most 1; and (iii) $\|B - A\| \le 1$, then the following hold:*

(a) $\angle(F_i(A), F_i(B))$ and $\angle(E_i(A), E_i(B))$ are less than $\delta/3$;

(b) $\frac{1}{3} \le s_j(A)/s_j(B) \le 3$ for each $j \le D_i$ and $s_j(B) \le 2$ for each $j > D_i$;

(c) If V is any subspace of dimension D_i such that $\perp(V, E_i(A)) > \delta/6$, then $\angle(BV, F_i(A)) < \delta/3$;

(d) If V is a subspace of dimension D_i and $\perp(V, E_i(A)) > \delta$, then $|\det(A|V)| \ge (D\delta)^{D_i} \exp \Xi_{D_i}(A)$, where D is an absolute constant.

This is all proved in [18] except for the second part of (b). For this part, let V be the subspace of $\mathbb{R}^d$ spanned by the (D_i+1)st to dth singular vectors. Then for $v \in V$ of norm 1, we have $\|Bv\| \le 2$. Since V is a $(d - D_i)$-dimensional subspace on which B uniformly expands vectors by at most 2, we have $s_j(B) \le 2$ for all $j > D_i$.

Lemma 4. *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$ and let $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ be a measurable map such that $\log^+\|A(\omega)\|$ is integrable. There exists C such that for all $\eta_0 > 0$, there exists ϵ_0 such that for all $\epsilon < \epsilon_0$, there exists $G \subseteq \Omega$ of measure at least $1 - \eta_0$ such that for all $\omega \in G$, and all $(\Delta_n) \in U^{\mathbb{Z}}$*

$$\|(A^\epsilon_{\bar\omega})^{(N)} - A^{(N)}_\omega\| \le 1,$$

where $\bar\omega = (\omega, (\Delta_n))$, $N = \lfloor C|\log\epsilon|\rfloor$ and $(A^\epsilon_{\bar\omega})^{(N)} = A^\epsilon_{N-1}(\bar\omega)\dots A^\epsilon_1(\bar\omega)A^\epsilon_0(\bar\omega)$.

Proof. Let $g(\omega) = \log^+(\|A_\omega\| + 1)$ and let $C > 0$ satisfy $\int g(\omega)\, d\mathbb{P}(\omega) < 1/C$. Notice that provided $\epsilon < 1$ (and using the fact that the perturbations have norm bounded by ϵ), $\log^+\|A^\epsilon_{\bar\omega}\| \le g(\omega)$, and

$$\begin{aligned}\|(A^\epsilon_{\bar\omega})^{(N)} - A^{(N)}_\omega\| &\le \sum_{i=0}^{N-1} \|(A^\epsilon_{\bar\sigma^i\bar\omega})^{(N-i-1)}(A^\epsilon_{\bar\sigma^i\bar\omega} - A_{\sigma^i\omega})A^{(i)}_\omega\| \\ &\le N\epsilon\exp(g(\omega) + \ldots + g(\sigma^{N-1}\omega)).\end{aligned}$$

There exists n_0 such that for $N \ge n_0$, $N\epsilon\exp(g(\omega)+\ldots+g(\sigma^{N-1}\omega)) \le \epsilon\exp(N/C)$ on a set of measure at least $1-\eta_0$. In particular, provided $\lfloor C|\log\epsilon_0|\rfloor > n_0$, taking $N = \lfloor C|\log\epsilon|\rfloor$, the conclusion follows.

∎

Lemma 5. *Let σ be an ergodic measure-preserving transformation of $(\Omega, \mathbb{P})$ and let $A\colon \Omega \to M_{d\times d}(\mathbb{R})$ be a measurable map such that $\int \log^+ \|A(\omega)\| \, \mathbb{P}(\omega) < \infty$. Let the Lyapunov exponents be $\infty > \lambda_1 > \lambda_2 > \ldots > \lambda_p \geq -\infty$ with multiplicities $d_1, \ldots, d_p$. Suppose $1 \leq i < p$ is such that $\lambda_i > 0 > \lambda_{i+1}$, and let $D_i = d_1 + \ldots + d_i$.*

Let $\eta_0 > 0$ and $\delta_1 > 0$ be given. Then there exist $n_0 > 0$, $\kappa > 0$ and $\delta \leq \min(\delta_1, \kappa)$ such that: for all $n \geq n_0$, there exists a set $G \subseteq \Omega$ with $\mathbb{P}(G) > 1 - \eta_0$ such that for $\omega \in G$, we have

(a) $\perp (E_i(\omega), F_i(\omega)) > 10\kappa$;
(b) $\angle(F_i(A_\omega^{(n)}), F_i(\sigma^n\omega)) < \delta$;
(c) $\angle(E_i(A_\omega^{(n)}), E_i(\omega)) < \delta$;
(d) $s_{D_i}(A_\omega^{(n)}) > K(\delta)$ and $s_{D_i+1}(A_\omega^{(n)}) < 1$, where $K(\delta)$ is as given in Lemma 3.

Proof. From the proof of Oseledets' theorem, we know $\perp (E_i(\omega), F_i(\omega))$ is a positive measurable function. Hence there exists $\kappa > 0$ such that (a) occurs on a set of measure at least $1 - \eta_0/4$. Let $\delta = \min(\delta_1, \kappa)$.

From the proof of Oseledets' theorem, there exists an $n_1 > 0$ such that for all $n \geq n_1$, $\angle(F_i(A^{(n)}_{\sigma^{-n}\omega}), F_i(\omega)) < \delta$ and $\angle(E_i(A_\omega^{(n)}), E_i(\omega)) < \delta$ hold on sets of measure at least $1-\eta_0/4$. Hence there is a set of measure at least $1 - \eta_0/4$ where (c) holds. Similarly, using shift-invariance, there is a set of measure at least $1 - \eta_0/4$ where (b) holds.

Since $\frac{1}{n} \log s_{D_i}(A_\omega^{(n)}) \to \lambda_i$ and $\frac{1}{n} \log s_{D_i+1}(A_\omega^{(n)}) \to \lambda_{i+1}$, (d) holds on a set of measure at least $1 - \eta_0/4$ for all $n \geq n_2$ for some $n_2 > 0$. Now let $n \geq n_0 = \max(n_1, n_2)$. Intersecting the above sets gives a set G satisfying the conclusions of the lemma. $\square$

3. Convergence of Lyapunov exponents

Proof of Theorem 1(I). Most of the work in this part is concerned with showing the inequality

$$\liminf_{\epsilon\to 0}(\mu_1^\epsilon + \ldots + \mu_{D_i}^\epsilon) \geq \mu_1 + \ldots + \mu_{D_i}, \text{ for any } 1 \leq i \leq p. \tag{2}$$

We also prove

$$\limsup_{\epsilon\to 0}(\mu_1^\epsilon + \ldots + \mu_j^\epsilon) \leq \mu_1 + \ldots + \mu_j \text{ for any } 1 \leq j \leq d \tag{3}$$

which is fairly straightforward using sub-additivity. These facts, combined with the fact that the μ_j^ϵ and μ_j are decreasing in j are sufficient to establish the claim that $\mu_j^\epsilon \to \mu_j$ for each j.

To see this, suppose that (2) and (3) hold. Let $h_j = \mu_1 + \ldots + \mu_j$ and let $H_j(\epsilon) = \mu_1^\epsilon + \ldots + \mu_j^\epsilon$. By (2) and (3), we have $\lim_{\epsilon\to 0} H_{D_i}(\epsilon) = h_{D_i}$. If $\lambda_{i+1} = -\infty$, we see $\lim_{\epsilon\to 0}\mu_j^\epsilon = -\infty$ for all $j > D_i$ from (3). Hence we may assume that $\lambda_{i+1} > -\infty$. Since the exponents are arranged in decreasing order, $(H_j(\epsilon))_{j=1}^d$ is a 'concave' sequence for each ϵ (that is $H_{j+1}(\epsilon) - H_j(\epsilon) \leq H_j(\epsilon) - H_{j-1}(\epsilon)$ for each j in range), as is $(h_j)_{j=1}^d$. However, h_j is an arithmetic progression for j in the range D_i to D_{i+1}. Since a concave function is bounded below by its secant, we deduce $\liminf_{\epsilon\to 0} H_j(\epsilon) \geq h_j$ for $D_i \leq j \leq D_{i+1}$. Hence we see $H_j(\epsilon) \to h_j$ as $\epsilon \to 0$ for each j, from which the statement follows.

To show (3), let $\chi > 0$ and let $1 \leq j \leq d$. By sub-additivity, $\int (1/N)\Xi_j((A_{\bar\omega}^\epsilon)^{(N)}) \, d\bar{\mathbb{P}}(\bar\omega)$ is an upper bound for $H_j(\epsilon)$. By the sub-additive ergodic theorem, there exists an $N > 0$ such that $\int (1/N)\Xi_j(A_\omega^{(N)})\,d\mathbb{P} < h_j + \chi/2$. Now for arbitrary $\epsilon < \frac{1}{d}$ and an arbitrary sequence $(\Delta_k) \in U^{\mathbb{Z}}$, $\Xi_j((A_{\bar\omega}^\epsilon)^{(N)}) \le j\sum_{k=0}^{N-1}\log(\|A_{\sigma}^k\omega\| + d\epsilon)$ giving domination by an integrable function. Now as ϵ is shrunk to 0, $\Xi_j((A_{\bar\omega}^\epsilon)^{(N)}) \to \Xi_j(A_\omega^{(N)})$ for all $\bar\omega$. Hence, the dominated convergence theorem gives that $H_j(\epsilon) < h_j + \chi$ for all sufficiently small ϵ as required. Notice that this part of the argument is completely general, whereas the lower bound depends on the particular properties of the matrix perturbations.

We now focus on proving (2). In the case $i = p$, this is automatic since the sum of the characteristic exponents is the integral of the log of the determinant of the matrices generating the cocycle, so we assume $i < p$. Let $j = D_i$. We therefore have $\lambda_i > -\infty$. By multiplying the entire family of matrices by a positive constant, we may assume that $\lambda_i > 0$ and $\lambda_{i+1} < 0$.

Let $\chi > 0$ be arbitrary. Let D be the absolute constant occurring in the statement of Lemma 3, C be as in the statement of Lemma 4 and K be the constant occurring in the statement of Lemma 10. Define a constant $\eta > 0$ by

$$\eta = \min\left(\frac{\chi}{4 \times 1.28 d^2 j}, \frac{\chi C}{8K}\right). \tag{4}$$

Let n_0, κ and δ be the quantities given by Lemma 5 using $\delta_1 = \frac{1}{2}$ and $\eta_0 = \eta/2$. Since $\int \log\|A_\omega\|\,d\mathbb{P}(\omega) < \infty$, it follows that $\int \Xi_j^+(A_\omega)\,d\mathbb{P}(\omega) < \infty$.

Let $N(\epsilon) = \lfloor C|\log\epsilon|\rfloor$, where C is as above. The fact that N scales like $|\log\epsilon|$ will be of crucial importance later. Let ϵ_0 be the quantity appearing in Lemma 4 with η_0 taken to be $\eta/2$.

Let ϵ be sufficiently small that

$$\begin{aligned} N(\epsilon) &> \frac{4j\log(3/(\delta D))}{\chi}, \\ \frac{|\log\epsilon|}{N(\epsilon)} &< 2/C, \\ N(\epsilon) &> \frac{8}{\chi}\int \Xi_j^+(A_\omega)\,d\mathbb{P}(\omega), \\ \epsilon &< \epsilon_0. \end{aligned} \tag{5}$$

Let G be the intersection of the good set given by Lemma 4 with the good set given by Lemma 5 with n taken to be $N = N(\epsilon)$, so that $\mathbb{P}(G) > 1 - \eta$. If $\omega \in G$, we say the matrix product $A_{\sigma^{N-1}\omega}\cdots A_\omega$ is a *good block*.

Now we divide everything into blocks of length N and estimate the sum of the logarithms of the first j singular values of the ϵ-perturbed cocycle.

We will bound from above the difference between the sum of the logs of the first j singular values in the unperturbed system and this sum in the perturbed version. We informally speak of the *costs* due to various contributions. That is, estimates of various contributions to an upper bound for the difference (*unperturbed*)−(*perturbed*). These costs are estimated in the following parts.

i. To deal with the concatenation of good blocks, we give an upper bound for the difference (*sum of individual block exponents*)−(*exponent of concatenated block*). This is estimated using Lemma 3. Over the whole block there is a cost of at most $\log(3/D\delta)$, so a cost per index of $O(1/|\log\epsilon|)$.

ii. Reduction of singular values within bad blocks. There is an expected cost of at worst $1.28d^2j$ per index in a bad block from (8).

iii. Reduction of singular values at the first and last matrix of a string of bad blocks. Here, there is an upper bound in expected cost of approximately $|\log\epsilon|$ per bad block. Here is where it is crucial that the blocks are of length $O(|\log\epsilon|)$. The upper bound for the cost averages out at $O(1)$ per index in each bad block. The argument is saved by the fact that most blocks are good blocks.

The sum of the costs is $O(\eta)+O(1/|\log\epsilon|)$ per index (η being the frequency of bad blocks), which will allow us to derive (2). Let us proceed with the details.

Step 1. *A lower bound for Ξ_j for concatenations of good blocks.*

Suppose $k < l$ and $\sigma^{kN}\omega, \sigma^{(k+1)N}\omega, \ldots, \sigma^{(l-1)N}\omega \in G$. Let $B_n = A^{(N)}_{\sigma^{nN}\omega}$ and $\tilde{B}_n = (A^{\epsilon}_{\sigma^{nN}\omega})^{(N)}$. We then claim that

$$\begin{aligned}\Xi_j(\tilde{B}_{l-1}\cdots\tilde{B}_k) &\geq \sum_{i=k}^{l-1}\Xi_j(B_i) + (l-k)j\log\delta - (l-k)j\log(3/D) \\ &\geq \Xi_j(B_{l-1}\cdots B_k) + (l-k)j\log(D\delta/3),\end{aligned} \tag{6}$$

where D is the absolute constant appearing in Lemma 3.

This is proved inductively using Lemma 3. Recall that $\|B_n - \tilde{B}_n\| \leq 1$. We let $\tilde{V}_k = V_k = E_j(B_k)^{\perp}$ and define $V_{n+1} = B_nV_n$ and $\tilde{V}_{n+1} = \tilde{B}_n\tilde{V}_n$.

We claim that the following hold:

i. $\angle(V_n, \tilde{V}_n) < \delta$ for each n;

ii. $\perp(V_n, E_i(B_n)) > \delta$ and $\perp(\tilde{V}_n, E_i(\tilde{B}_n)) > \delta$ for each n.

Item (i) and the first part of (ii) hold immediately for the case $n = k$. The second part of (ii) holds because $\tilde{V}_k = V_k = E_i(B_k)^{\perp}$ and $\angle(E_i(B_k), E_i(\tilde{B}_k)) < \delta$ by Lemma 3.

Given that (i) and (ii) hold for $n = m$ and that B_m is a good block, Lemma 3 implies that $\angle(V_{m+1}, F_i(B_m)) < \delta/3$, $\angle(\tilde{V}_{m+1}, F_i(\tilde{B}_m)) < \delta/3$ and $\angle(F_i(B_m), F_i(\tilde{B}_m)) < \delta/3$, so that $\angle(\tilde{V}_{m+1}, V_{m+1}) < \delta$, yielding (i) for $n = m + 1$.

Finally, by the induction hypothesis and Lemma 5, we have $\angle(F_i(\sigma^{(m+1)N}\omega), F_i(B_m)) < \delta$ and $\perp(F_i(\sigma^{(m+1)N}\omega), E_j(\sigma^{(m+1)N}\omega)) > 10\delta$. Thus, we obtain (ii) for $n = m + 1$.

Hence using Lemma 3(d), we see that $\det(\tilde{B}_n|_{\tilde{V}_n}) \geq (D\delta)^j \det(\tilde{B}_n|_{E_i(\tilde{B}_n)^{\perp}}) = (D\delta)^j e^{\Xi_j(\tilde{B}_n)} \geq (D\delta/3)^j e^{\Xi_j(B_n)}$.

Since $\Xi_j(\tilde{B}_{l-1}\cdots\tilde{B}_k) \geq \prod_{n=k}^{l-1}\det(\tilde{B}_n|\tilde{V}_n)$, multiplying the inequalities and taking logarithms gives the result.

Step 2. *A lower bound for Ξ_j for concatenations of arbitrary blocks.*

Let $A_1, \ldots, A_n$ be an arbitrary sequence of matrices. We write

$$g^{\epsilon}(A_1, \ldots, A_n, \Delta_1, \ldots, \Delta_n) = \Xi_j(A^{\epsilon}_n \ldots A^{\epsilon}_1) - \Xi_j(A_n \ldots A_1) \tag{7}$$

and prove that $(g^{\epsilon})^-$ is integrable in $\Delta_1, \ldots, \Delta_n$ and that

$$\int g^{\epsilon}(A_1, \ldots, A_n, \Delta_1, \ldots, \Delta_n)\, d\lambda^n(\Delta_1, \ldots, \Delta_n) \geq -1.28d^2nj \tag{8}$$

for all $A_1, \ldots, A_n$.

Lemma 6. *There exists $B \approx -1.28$ such that for all $z \in \mathbb{C}$, and all $l \geq 0$*

$$\int_0^1 t^l \log|1 - tz|\, dt \geq B.$$

Proof. Let us show there exists a lower bound; its precise value is irrelevant for our purposes. Since for every $z \in \mathbb{C}, t \in [0,1]$, we have that $|1 - tz| \geq |1 - t\operatorname{Re}(z)| \geq \big|1 - t|\operatorname{Re}(z)|\big|$, it suffices to show the lemma holds for $z \in \mathbb{R}^+ \cup \{0\}$. For $z = 0$ the integral is 0. Let us assume $z \in \mathbb{R}^+$, and let $\log^- x := \min(0, \log x)$. Then,

$$\begin{aligned}\int_0^1 t^l \log|1 - tz|\, dt &\geq \int_0^1 t^l \log^-|1 - tz|\, dt \geq \int_0^1 \log^-|1 - tz|\, dt \\ &= \frac{1}{z}\int_0^z \log^-|1 - y|\, dy \geq \inf_{z \in (0,2]} \frac{1}{z}\int_0^z \log^-|1 - y|\, dy.\end{aligned}$$

The function $g(z) := \frac{1}{z}\int_0^z \log^-|1 - y|\, dy$ for $z \neq 0$ and $g(0) := 0$ is continuous on $\mathbb{R}^+$, and hence bounded on $[0, 2]$. The statement follows. □

Lemma 7. *Let B be as in Lemma 6 and p be a polynomial. Then, for all $l \geq 0$,*

$$\int_0^1 t^l (\log|p(t)| - \log|p(0)|)\, dt \geq B \cdot \deg(p).$$

Proof. If $p(0) = 0$, then the result is clear. Otherwise, we consider the polynomial $f(t) = p(t)/p(0)$ and demonstrate that $\int_0^1 t^l \log|f(t)|\, dt \geq B \cdot \deg(f)$.

To see this, notice that $f(t)$ may be expressed as $\prod_{i=1}^{\deg(f)}(1 - tz_i)$, where $\{z_i^{-1}\}_{i=1}^{\deg(f)}$ is the set of roots of f, and hence p, with multiplicity. Applying Lemma 6 then gives the result. □

Lemma 8. *Let B be the constant from the statement of Lemma 6. Let $P(t)$ be a degree j matrix-valued polynomial. That is, $P(t)$ may be expressed as $\sum_{k=0}^{j} A_k t^k$ for some collection of $d \times d$ matrices A_k. Then, for all $l \geq 0$,*

$$\int_0^1 t^l (\log\|P(t)\| - \log\|P(0)\|)\, dt \geq B \cdot \deg(P).$$

Proof. If $P(0)$ is the zero matrix, the result is trivial. Otherwise, there exist unit vectors $\mathbf{e}$ and $\mathbf{f}$ such that $P(0)\mathbf{e} = \|P(0)\|\mathbf{f}$.

If we set $p(t) = \langle P(t)\mathbf{e}, \mathbf{f}\rangle$, then we have $p(0) = \|P(0)\|$ and $\|P(t)\| \geq |p(t)|$, so the result follows from Lemma 7. □

Lemma 9. *Let $B \approx -1.28$ be the constant from the statement of Lemma 6. Let L, M, A and R be arbitrary $d \times d$ matrices. Then*

$$\int_0^1 t^l \left(\log\|\Lambda^j(L(A + tM)R)\| - \log\|\Lambda^j(LAR)\|\right)\, dt \geq jB.$$

Proof. Notice that $\Lambda^j(L(A+tM)R)$ is a polynomial family of operators on $\Lambda^j\mathbb{R}^d$. Taking the standard orthogonal basis of $\Lambda^j\mathbb{R}^d$, let $P(t)$ be the matrix of $\Lambda^j(L(A + tM)R)$. The result then follows by applying Lemma 8. □

We obtain (8) by a telescoping argument:

$$\begin{aligned}&\Xi_j(A_n \dots A_1) - \Xi_j(A_n^\epsilon \dots A_1^\epsilon)\\ &= \sum_{k=1}^{n} \left(\Xi_j(A_n^\epsilon \dots A_{k+1}^\epsilon A_k \dots A_1) - \Xi_j(A_n^\epsilon \dots A_k^\epsilon A_{k-1} \dots A_1)\right)\end{aligned}$$

Recall that $\Xi_j(A) = \log \|\Lambda^j A\|$. We estimate the integral of the kth term in the sum. Let $L = A_n^\epsilon \dots A_{k+1}^\epsilon$ and $R = A_{k-1} \dots A_1$. Regarding $\Delta_{k+1}, \dots, \Delta_n$ as fixed, we need to estimate:

$$\int_{\mathbf{B}} \left(\log \|\Lambda^j(LA_kR)\| - \log \|\Lambda^j(L(A_k + \epsilon\Delta)R)\|\right) \, d\lambda(\Delta),$$

where $\mathbf{B} = \{M \colon \|M\| \le 1\}$. We then disintegrate the measure λ radially, so that $d\lambda = d^2 t^{d^2-1}\, dt \cdot d(\partial\lambda)(H)$ where $\Delta = tH$, H takes values in $\partial\mathbf{B}$ and $\partial\lambda$ is the boundary measure. For a fixed H, the quantity to estimate is

$$\int_0^1 \left(\log \|\Lambda^j(LA_kR)\| - \log \|\Lambda^j(L(A_k + \epsilon tH)R)\|\right) t^{d^2-1}\, dt.$$

Since this quantity is uniformly bounded above, by Lemma 9, we obtain (8).

Step 3. *Gluing blocks.*

Lemma 10. *Let L, R and A be given matrices. Then $\Xi_j(L(A+\epsilon\Delta)R) - (\Xi_j(L) + \Xi_j(R))$ has integrable negative part as a function of Δ and has integral bounded below by $K \log \epsilon$, where K is independent of L, A and R.*

Proof. Write $L = O_1 D_1 O_2$ where D_1 is diagonal with entries arranged in decreasing order and O_1 and O_2 are orthogonal. Similarly write $R = O_3 D_2 O_4$. Let $A' = O_2 A O_3$ and $\Delta' = O_2 \Delta O_3$. Then we have

$$\begin{aligned}\Xi_j(L(A+\epsilon\Delta)R) &= \Xi_j(D_1(A' + \epsilon\Delta')D_2);\\ \Xi_j(L) &= \Xi_j(D_1); \text{ and}\\ \Xi_j(R) &= \Xi_j(D_2).\end{aligned}$$

Using the inequality $\Xi_j(AB) \le \Xi_j(A) + \Xi_j(B)$ and setting C to be the diagonal matrix with 1's in the first j elements of the diagonal and 0's elsewhere, we have

$$\begin{aligned}\Xi_j(D_1(A' + \epsilon\Delta')D_2) &\ge \Xi_j(CD_1(A' + \epsilon\Delta')D_2C)\\ &= \Xi_j(D_1C(A' + \epsilon\Delta')CD_2)\\ &= \Xi_j(D_1C) + \Xi_j(C(A' + \epsilon\Delta')C) + \Xi_j(CD_2)\\ &= \Xi_j(L) + \Xi_j(R) + \Xi_j(C(A' + \epsilon\Delta')C).\end{aligned}$$

The equality between the second and third lines arises because the matrices D_1C, $C(A' + \epsilon\Delta')C$ and CD_2 and their product have non-zero entries only in the top left $j \times j$ submatrix. For such matrices, $\Xi_j(\cdot)$ is numerically equal to the logarithm of the absolute value of the determinant of the submatrix. Since the determinant is multiplicative, the equality follows.

Since Lebesgue measure on $\mathbf{B} = \{M \colon \|M\| \le 1\}$ is preserved by the operations of pre- and post-multiplying by an orthogonal matrix, it suffices to show that there exists $K > 0$

such that for any matrix A,

$$\int_{\mathbf{B}} \Xi_j(C(A'+\epsilon\Delta)C)\,d\lambda(\Delta) \geq K\log\epsilon \quad \text{for } \epsilon<\tfrac{1}{2}. \tag{9}$$

Let A'' be the top left $j\times j$ submatrix of A' and notice that the measure on the top left $j\times j$ submatrix of Δ is absolutely continuous with respect to the measure on $j\times j$ matrices with uniform entries in $[-1,1]$ with bounded density. As noted above, Ξ_j agrees with the logarithm of the absolute value of the determinant for a $j\times j$ matrix.

Hence to establish (9), it suffices to give a logarithmic lower bound:

$$\int_{\mathbf{U}} \log\det(A''+\epsilon U)\,d\lambda'(U) \geq K\log\epsilon \quad \text{for } \epsilon<\tfrac{1}{2}, \tag{10}$$

where $\mathbf{U}$ is the collection of $j\times j$ matrices with entries in $[-1,1]$ and λ' is the uniform measure on $\mathbf{U}$. One checks, thinking of the columns of U being generated one at a time, that the probability that the ith column lies within a δ-neighbourhood of the span of the previous columns is at most $O(\delta/\epsilon)$, so the probability that the determinant of $A''+\epsilon U$ is less than δ^j is $O(j\delta/\epsilon)$. Hence we obtain

$$\mathbb{P}(-\log\det(A''+\epsilon U)>k) \leq \min(1, Cje^{-k/j}/\epsilon).$$

Using the estimate for non-negative random variables $\mathbb{E}X \leq \sum_{k=0}^{\infty}\mathbb{P}(X\geq k)$, we obtain the bound $\mathbb{E}(-\log\det(A''+\epsilon U)) \lesssim j|\log\epsilon|$.

From this, we obtain the $O(|\log\epsilon|)$ bound as required. □

Step 4. *Putting it all together.*

We apply this by grouping each consecutive string of good blocks into a single matrix (and using (6)) and also grouping strings of consecutive bad blocks minus the first and last matrices into a single matrix (and using (8)). The first and last matrices of a string of bad blocks are then handled with Lemma 10.

More specifically, we condition on $\omega\in\Omega$ and calculate $\int \Xi_j((A^\epsilon_\omega)^{(MN)})\,d\lambda^{MN}$. Let $S=\{0\leq l<M\colon A^{(N)}_{\sigma^{lN}\omega} \text{ is bad}\}$. Let $r(\omega)=|S|$ and $0<b_1<b_2<\ldots<b_r<M$ be the increasing enumeration of S. Also let $b_0=-1$ and $b_{r+1}=M$. Then we factorize $(A^\epsilon_\omega)^{(MN)}$ and $A^{(MN)}_\omega$ as

$$(A^\epsilon_\omega)^{(MN)} = \tilde G_r\tilde B_r\cdots\tilde B_2\tilde G_1\tilde B_1\tilde G_0\,;\text{ and}$$
$$A^{(MN)}_\omega = G_rB_r\cdots B_2G_1B_1G_0,$$

where $G_l=A^{((b_{l+1}-b_l-1)N)}_{\sigma^{(b_l+1)N}\omega}$, $\tilde G_l=(A^\epsilon_{\sigma^{(b_l+1)N}\omega})^{((b_{l+1}-b_l-1)N)}$, $B_l=A^{(N)}_{\sigma^{b_lN}\omega}$ and $\tilde B_l=(A^\epsilon_{\sigma^{b_lN}\omega})^{(N)}$ (so the G_l are products of consecutive good blocks and B_l are (single) bad blocks). We further factorize B_l and $\tilde B_l$ as $\tilde B_l=A^\epsilon_{\sigma^{(b_l+1)N-1}\omega}\tilde C_lA^\epsilon_{\sigma^{b_lN}\omega}$ and $B_l=A_{\sigma^{(b_l+1)N-1}\omega}C_lA_{\sigma^{b_lN}\omega}$, where $\tilde C_l=(A^\epsilon_{\sigma^{b_lN+1}\omega})^{(N-2)}$ and $C_l=A^{(N-2)}_{\sigma^{b_lN+1}\omega}$.

Now using Lemma 10 (and the constant K from its statement), we have

$$\int \Xi_j((A^\epsilon_\omega)^{(MN)})\,d\lambda^{MN} \geq \int\left(\sum_{l=0}^{r(\omega)}\Xi_j(\tilde G_l)+\sum_{l=1}^{r(\omega)}\Xi_j(\tilde C_l)\right)d\lambda^{MN} + r(\omega)K|\log\epsilon|.$$

From (6), we have $\sum_{l=0}^r \Xi_j(\tilde{G}_l) \geq \sum_{l=0}^r \Xi_j(G_l) + Mj\log(D\delta/3)$ for all values of the perturbation matrices that occur inside those blocks. From (8), we have for each $1 \leq l \leq r(\omega)$,

$$\int \Xi_j(\tilde{C}_l)\, d\lambda^{N-2}(\Delta_{b_lN+1}, \ldots, \Delta_{(b_l+1)N-2}) \geq \Xi_j(C_l) - 1.28d^2(N-2)j.$$

Letting $E(\omega) = Mj\log(D\delta/3) - 1.28d^2(N-2)jr(\omega) + r(\omega)K\log\epsilon$ and combining the inequalities together with subadditivity of Ξ_j yields

$$\begin{aligned} \frac{1}{MN}\int \Xi_j((A^\epsilon_\omega)^{(MN)})\, d\Lambda^{MN} &\geq \frac{1}{MN}\sum_{l=0}^{r(\omega)} \Xi_j(G_l) + \frac{1}{MN}\sum_{l=1}^{r(\omega)} \Xi_j(C_l) + \frac{E(\omega)}{MN} \\ &\geq \frac{\Xi_j(A^{(MN)}_\omega)}{MN} - \frac{1}{MN}\sum_{l=1}^{r(\omega)} \left(\Xi^+_j(A_{\sigma^{b_lN}\omega}) + \Xi^+_j(A_{\sigma^{(b_l+1)N-1}\omega})\right) + \frac{E(\omega)}{MN}. \end{aligned} \tag{11}$$

By (4) and (5), we see $(1/MN)\int E(\omega)\, d\mathbb{P}(\omega) > -3\chi/4$. Finally, we have

$$\frac{1}{MN}\int \left(\sum_{l=1}^{r(\omega)} (\Xi^+_j(A_{\sigma^{b_lN}\omega}) + \Xi^+_j(A_{\sigma^{(b_l+1)N-1}\omega}))\right) \leq \frac{2}{N}\int \Xi^+_j(A_\omega)\, d\mathbb{P}(\omega) < \frac{\chi}{4}.$$

Combining these inequalities, we obtain

$$\frac{1}{MN}\int \Xi_j((A^\epsilon_\omega)^{(MN)})\, d\bar{\mathbb{P}}(\bar\omega) \geq \frac{1}{MN}\int \Xi_j(A^{(MN)}_\omega)\, d\mathbb{P}(\omega) - \chi.$$

Taking the limit as $M \to \infty$, we deduce $\liminf_{\epsilon\to 0}(\mu^\epsilon_1 + \ldots + \mu^\epsilon_j) \geq (\mu_1 + \ldots + \mu_j) - \chi$. Since $\chi > 0$ was arbitrary, we deduce (2).

□

4. Convergence of Oseledets spaces

Let $\mathcal{N}_i$ for $1 \leq i \leq p$ be disjoint neighbourhoods of λ_i as in the statement of the theorem. In Part (I), we established the existence of an $\epsilon_0 > 0$ such that for $\epsilon < \epsilon_0$, in the perturbed matrix cocycle, $\mu^\epsilon_j \in \mathcal{N}_i$ for all j satisfying $D_{i-1} < j \leq D_i$. Recall that $Y^\epsilon_i(\bar\omega)$ was defined to be the sum of the Oseledets spaces corresponding to exponents in $\mathcal{N}_i$, with $Y_i(\omega)$ being the corresponding spaces for the unperturbed matrix cocycle. Let $F_i(\omega) = \bigoplus_{k\leq i} Y_i(\omega)$ be the fast subspace for the unperturbed matrix cocycle and $E_i(\omega) = \bigoplus_{k>i} Y_i(\omega)$ be the slow subspace. We similarly introduce notation $F^\epsilon_i(\bar\omega)$ and $E^\epsilon_i(\bar\omega)$ in the perturbed matrix cocycle, so that $F^\epsilon_i(\bar\omega)$ is a D_i-dimensional space corresponding to the top D_i exponents (counted with multiplicity) and $E^\epsilon_i(\bar\omega)$ is a $(d-D_i)$-dimensional space corresponding to the smallest $d - D_i$ exponents. Notice that $\bar\omega = (\omega, \Delta)$, and so $F_i(\omega)$ and $F^\epsilon_i(\bar\omega)$ may be regarded as living on the same probability space $(\bar\Omega, \bar{\mathbb{P}})$.

The proof of Theorem 1(II) will follow relatively straightforwardly from the following lemma whose proof will occupy this section.

Lemma 11. *Let $0 < \chi < 1$. Let $F_i(\omega)$ and $F^\epsilon_i(\bar\omega)$ be as above. Then, for every ϵ sufficiently small,*

$$\bar{\mathbb{P}}\big(\bar\omega : \angle(F^\epsilon_i(\bar\omega), F_i(\omega)) > \chi\big) < \chi.$$

That is, $F^\epsilon_i(\bar\omega)$ converges in probability to $F_i(\omega)$ as $\epsilon \to 0$.

Notice that the conclusion of Lemma 11 trivially holds for $i = p$ because $F_i^\epsilon(\bar\omega) = F_i(\omega) = \mathbb{R}^d$ in this case. Thus, we will only be concerned with the case $i < p$. In this case none of the Lyapunov exponents of vectors in $F_i(\omega)$ are $-\infty$.

Recalling that $E_i(\omega) = \left(F_i^*(\omega)\right)^\perp$, where $F_i^*(\omega)$ denotes the Oseledets space of the cocycle dual to A, Lemma 11 immediately implies the following.

Corollary 12. *Let $E_i(\omega)$ and $E_i^\epsilon(\bar\omega)$ denote the slow Oseledets subspaces of the unperturbed and perturbed cocycles, respectively, as described above. Then $E_i^\epsilon(\bar\omega)$ converges in probability to $E_i(\omega)$ as $\epsilon \to 0$.*

Proof of Theorem 1(II) from Lemma 11 . Notice that $Y_i(\omega) = F_i(\omega) \cap E_{i-1}(\omega)$ and $Y_i^\epsilon(\bar\omega) = F_i^\epsilon(\bar\omega) \cap E_{i-1}^\epsilon(\bar\omega)$, so we want to show that $E_{i-1}^\epsilon(\bar\omega) \cap F_i^\epsilon(\bar\omega)$ converges in probability to $E_{i-1}(\omega) \cap F_i(\omega)$ as $\epsilon \to 0$. We also have

$$E_{i-1}^\epsilon(\bar\omega) \cap F_i^\epsilon(\bar\omega) = \mathrm{Pr}_{F_i^\epsilon(\bar\omega)\|E_i^\epsilon(\bar\omega)}(E_{i-1}^\epsilon(\bar\omega)); \text{ and}$$
$$E_{i-1}(\omega) \cap F_i(\omega) = \mathrm{Pr}_{F_i(\omega)\|E_i(\omega)}(E_{i-1}(\omega)).$$

Now, lemma 6 of [11], together with Lemma 11 and the separation of $F_i(\omega)$ and $E_i(\omega)$ guaranteed by Lemma 5 implies the result. □

4.1. **Strategy and notation.** Throughout, we shall let $j = D_i$, so that we are studying evolution of j-dimensional subspaces. In order to show Lemma 11, we will assume that all of the perturbations (Δ_n) are fixed except for the -1 time coordinate. That is, we compute the probability that the perturbed and unperturbed fast spaces are close conditioned on $(\Delta_n)_{n\neq -1}$ and ω.

It is well known that $F_i^\epsilon(\bar\omega)$ depends only on the matrices $A^\epsilon_{\bar\sigma^n\bar\omega}$ for $n < 0$. Once ω and $(\Delta_n)_{n\neq-1}$ are fixed, we think of $F_i^\epsilon(\bar\omega)$ as a random variable (depending on Δ_{-1}), then applying the sequence of matrices (all already fixed), $(A^\epsilon_{\bar\omega})^{(nN)}$, we show that the resulting j-dimensional subspace is highly likely to be closely aligned to $F_i(\sigma^{nN}\omega)$.

To control the evolution, we successively apply $(A^\epsilon_{\bar\omega})^{(N)}$, $(A^\epsilon_{\bar\sigma^N\bar\omega})^{(N)}, \ldots, (A^\epsilon_{\bar\sigma^{(n-1)N}\bar\omega})^{(N)}$, where $\bar\omega = (\omega, \Delta)$, s denotes the left shift and $\bar\sigma(\bar\omega) = (\sigma\omega, s\Delta)$. We will assume that the underlying blocks of unperturbed A's are good blocks. The number, n, of steps will be fixed. In fact, n will depend only on the difference $\lambda_j - \lambda_{j+1}$ and the quantity, C, appearing in Lemma 4. Hence provided that the probability of bad blocks is very small, it will be likely that one has n consecutive good blocks.

We shall use the following parameterization of the Grassmannian of j-dimensional subspaces of $\mathbb{R}^d$. Let $f_1, \ldots, f_j$ be a basis for F, a j-dimensional subspace, and $e_1, \ldots, e_{d-j}$ a basis for E, a complementary subspace. Now for any j-dimensional vector space V with the property that $V \cap E = \{0\}$, each f_k can be uniquely expressed in the form $v_k - \sum_i b_{ik}e_i$ where $v_k \in V$. The parameterization of V with respect to the $(f_1, \ldots, f_j; e_1, \ldots, e_{d-j})$ chart (we shall mainly speak more informally of the (F, E) chart) is the $(d-j) \times j$ matrix $B = (b_{lk})_{1\le l\le d-j, 1\le k\le j}$. Conversely, given the matrix B, one can easily recover a basis for V: $v_k = f_k + \sum_l b_{lk}e_l$ and hence the subspace V.

Lemma 13. *Let F and E be orthogonal complements in $\mathbb{R}^d$ and let $(f_k)_{1\le k\le j}$ and $(e_k)_{1\le k\le d-j}$ be orthonormal bases. Let V be a j-dimensional subspace of $\mathbb{R}^d$ such that $V \cap E = \{0\}$. Let* *the parameterization of V with respect to the (F, E) chart be B. Then*

$$\perp(V, E) = \sqrt{1 - \frac{\|B\|}{\sqrt{1 + \|B\|^2}}}.$$

In particular for any $M > 1$, $\|B\| \leq M$ implies $\perp(V, E) \geq 1/(2M)$.

Proof. Let $v_k = f_k + \sum_l b_{lk} e_l$ so that (v_k) forms a basis for V. Now let $v = \sum_k c_k v_k = \sum c_k f_k + \sum_k (Bc)_k e_k$ belong to $V \cap S$, so that $\|c\|^2 + \|Bc\|^2 = 1$. The closest point in $E \cap S$ to v is $(1/\|Bc\|) \sum_k (Bc)_k e_k$, which is at a square distance $\|c\|^2 + (\|Bc\| - 1)^2 = 2(1 - \|Bc\|)$ from v. This distance is minimized when c is the multiple of the dominant singular vector of B for which $\|c\|^2 + \|Bc\|^2 = 1$. That is, $\|c\| = 1/\sqrt{\|B\|^2 + 1}$ and $\|Bc\| = \|B\|/\sqrt{\|B\|^2 + 1}$. Substituting this, we obtain the claimed formula for $\perp(V, E)$. □

Given a matrix M, for which $s_j(M) > s_{j+1}(M)$, let $e_1, \ldots, e_d$ be the singular vectors. We will refer to the $(e_1, \ldots, e_j; e_{j+1}, \ldots, e_d)$ chart as the $(E(M)^\perp, E(M))$ chart. Similarly, the $(F(M), F(M)^\perp)$ chart refers to the $(f_1, \ldots, f_j; f_{j+1}, \ldots, f_d)$ chart where $f_1, \ldots, f_d$ are an orthonormal basis given by the normalized Me_l for those l where this is non-zero; and chosen arbitrarily to ensure orthonormality otherwise. The vectors then satisfy $\langle Me_l, f_m \rangle = \delta_{lm} s_l(M)$. There may be some non-uniqueness of charts in the case that there are repeated singular values. However, this does not affect the conclusion and occurs with probability zero in any case.

Given $\bar{\omega}$, we write C_l to mean $(A^\epsilon_{\bar{\sigma}^{lN}\bar{\omega}})^{(N)}$. We will do the iteration using the following steps:

- (S0) Express $V = F^\epsilon_i(\bar{\omega})$ as a matrix B using the $(E(C_0)^\perp, E(C_0))$ chart. Set $l = 0$.
- (S1) Compute $C_l(V)$ in the $(F(C_l), F(C_l)^\perp)$ chart; this is straightforward as C_l is diagonal with respect to the pair of bases on the domain and range spaces.
- (S2) Change bases to the $(E(C_{l+1})^\perp, E(C_{l+1}))$ chart. Increase l and repeat steps S1 and S2 a total of n times.

The group $GL_d(\mathbb{R})$ acts on the $(d-j) \times j$ matrices in the following way: Let $Q = \begin{pmatrix} W & X \\ Y & Z \end{pmatrix}$, where W, X, Y and Z are respectively of dimensions $j \times j$, $j \times (d-j)$, $(d-j) \times j$ and $(d-j) \times (d-j)$. Then Q defines a self-map of the collection of $(d-j) \times j$ matrices by $f_Q(B) = (Y + ZB)(W + XB)^{-1}$ (providing the inverse is defined). One can check $f_Q \circ f_{Q'} = f_{QQ'}$. We then show that the update rules (S1) and (S2) correspond to Möbius transformations, so that their n-step composition is another Möbius transformation. Notice also that $f_Q(B) = YW^{-1} + (Z - YW^{-1}X)B(W + XB)^{-1}$. The strategy is to show that for ω belonging to a good set, and any sequence of perturbations, the non-constant term takes small values for most arguments B. In our application, these B's are the image under the chart of the $F^\epsilon_i(\bar{\omega})$ as in step (S0). We use properties of multivariate normal distributions to show that for good ω, most B's are in the part of the space where the Möbius transformation is almost constant.

The matrix $Z - YW^{-1}X$ plays a key role in this, and is analogous to a cross ratio.

Proof of Lemma 11. As pointed out above, the result is trivial if $i = p$, so that we assume $i < p$ and therefore $\lambda_i > -\infty$. Let $\tau < \frac{1}{4}(\lambda_i - \lambda_{i+1})$.

By multiplying the family of matrices by a positive constant, we may assume without loss of generality that $\lambda_i > 4\tau$ and $\lambda_{i+1} < 0$. Let C be the constant guaranteed for the family of matrices by Lemma 4. Fix $n > 1/(C\tau)$ and let ϵ_0 be the constant arising when Lemma 4 is applied with $\eta_0 = \chi/(8n+8)$.

Let χ be as in the statement of the lemma, and apply Lemma 5 with $\eta_0 = \chi/(8n+8)$ and $\delta_1 = \chi/2$. Let κ, δ and n_0 be as in the conclusion of the lemma and let G_1 be the 'good set' of ω's of measure at least $1-\eta_0$ on which the transversality, nearness and separation of singular value conditions of Lemma 5 are satisfied.

We now fix the range of ϵ in which we will obtain the required closeness of the top spaces. We shall set $N(\epsilon) = \lfloor C|\log \epsilon|\rfloor$, and will require that ϵ be small enough (and hence that $N(\epsilon)$ should be large enough) to simultaneously satisfy a number of conditions:

(C1) $\epsilon < \min(\epsilon_0, \frac{1}{d})$;
(C2) $N(\epsilon) > n_0$;
(C3) $e^{N(\epsilon)\tau} > \max(2+2/\delta, [(1+\delta)/(5\delta^2)]^{1/4})$;
(C4) $\exp(N(\epsilon)(n\tau - 1/C)) > 2\delta(e\pi/2)^{d^2/2}j^{3/2}/\chi$;
(C5) $\mathbb{P}(\|A_\omega\| > e^{\tau n N(\epsilon)} - 1) < \chi/8$;
(C6) $\mathbb{P}(s_j(A_\omega^{(N)}) > 6e^{4\tau N}$ and $s_{j+1}(A_\omega^{(N)}) < 1) > 1 - \chi/(8n+8)$;

Let G_2 be the set of measure at least $1-\eta_0$ guaranteed by applying Lemma 4 with the value ϵ as obtained above. Let $G_3 = \{\omega\colon s_j(A_\omega^{(N)}) > 6e^{4\tau N}$ and $s_{j+1}(A_\omega^{(N)}) < 1\}$ and $G_4 = \{\omega\colon \|A_{\sigma^{-1}\omega}\| \le e^{\tau n N} - 1\}$. Let $\tilde{G} = G_4 \cap \bigcap_{j=0}^{n} \sigma^{-jN}(G_1 \cap G_2 \cap G_3)$. Then $\mathbb{P}(\tilde{G}) \ge 1 - \chi/2$. We make the following claim.

Claim 14. *Let $\tilde{G}$ be as above. Then,*

$$\bar{\mathbb{P}}(\angle\big(F_i^\epsilon(\bar\omega), F_i(\omega)\big) > \chi|\omega) \le \chi/2 \text{ for all } \omega \in \sigma^{-(n+1)N}\tilde{G}. \tag{12}$$

With this result at hand, the proof of Lemma 11 goes as follows.

Proof of Lemma 11 using Claim 14.

$$\begin{aligned}
&\bar{\mathbb{P}}\big(\bar\omega : \angle(F_i^\epsilon(\bar\omega), F_i(\omega)) > \chi\big) \\
&\quad \le \bar{\mathbb{P}}(\angle(F_i^\epsilon(\bar\omega), F_i(\omega)) > \chi \mid \omega \in \sigma^{-(n+1)N}\tilde{G}) \cdot \mathbb{P}(\sigma^{-(n+1)N}\tilde{G}) + (1 - \mathbb{P}(\sigma^{-(n+1)N}\tilde{G})) \\
&\quad \le \chi/2 \cdot (1-\chi/2) + \chi/2 < \chi.
\end{aligned}$$

□

4.2. **Proof of Claim 14.** We let ω be a fixed element of $\tilde{G}$ throughout this proof, and demonstrate that $\lambda^{\mathbb{Z}}(\{\Delta\colon \angle(F_i^\epsilon(\bar\sigma^{(n+1)N}(\omega,\Delta)), F_i(\sigma^{(n+1)N}\omega)) > \chi\}) < \chi/2$. In fact, we do more. We let $\omega \in \tilde{G}$ be fixed and let $(\Delta_k)_{k\neq -1}$ be an arbitrary sequence. We then show that $\lambda\{\Delta_{-1}\colon \angle(F_i^\epsilon(\bar\sigma^{(n+1)N}(\omega,\Delta)), F_i(\sigma^{(n+1)N}\omega)) > \chi\} < \chi/2$. Write $\bar\omega$ for $(\omega, (\Delta_k))$.

Let $V_0 = F_i^\epsilon(\bar\omega)$, and let $V_{l+1} = C_l(V_l)$. Write B_l for the matrix of V_l with respect to the $(E(C_l)^\perp, E(C_l))$ basis, as explained in Step S0 above; we will see that B_0 is a random variable with ϵ-variability.

Let R_l be the matrix describing multiplication by C_l with respect to the $(E(C_l)^\perp, E(C_l))$ and $(F(C_l), F(C_l)^\perp)$ bases. This corresponds to Step S1. Let P_l (corresponding to Step S2) be the basis change matrix from the $(F(C_l), F(C_l)^\perp)$ to the $(E(C_{l+1})^\perp, E(C_{l+1}))$ basis.

Then, R_l is diagonal, say

$$R_l := \begin{pmatrix} D_{2,l} & 0 \\ 0 & D_{1,l} \end{pmatrix},$$

where $D_{1,l}$ is the diagonal matrix with entries $s_{j+1}, \ldots, s_d$ and $D_{2,i}$ is the diagonal matrix with entries $s_1, \ldots, s_j$, where $s_1 \geq \ldots \geq s_d$ are the singular values of C_l. Since $\sigma^{lN}\omega \in G_2$, we have $\|A^{(N)}_{\sigma^{lN}\omega} - (A^{\epsilon}_{\bar{\sigma}^{lN}\bar{\omega}})^{(N)}\| \leq 1$. Since $\sigma^{lN}\omega \in G_3$, we have $s_j(A^{(N)}_{\sigma^{lN}\omega}) > 6e^{4\tau N}$ and $s_{j+1}(A^{(N)}_{\sigma^{lN}\omega}) < 1$, so that by Lemma 3,

$$\|D_{2,l}x\| > 2e^{4\tau N}\|x\| \text{ for all } x \in \mathbb{R}^j \text{ and } \|D_{1,l}\| < 2. \tag{13}$$

Notice that P_l is an orthogonal matrix, as it is the change of basis matrix between two orthonormal bases. Let

$$P_l := \begin{pmatrix} \zeta_l & \gamma_l \\ \beta_l & \alpha_l \end{pmatrix}, \text{ and} \tag{14}$$

$$Q_l := P_l R_l =: \begin{pmatrix} q_l & t_l \\ p_l & r_l \end{pmatrix}, \tag{15}$$

so that

$$q_l = \zeta_l D_{2,l},\ t_l = \gamma_l D_{1,l},\ p_l = \beta_l D_{2,l} \text{ and } r_l = \alpha_l D_{1,l}. \tag{16}$$

To estimate $\|\zeta^{-1}\|$ we use a similar argument to that in Lemma 13. Let $F(C_l)$ be spanned by the singular vector images $f_1, \ldots, f_j$; $E(C_{l+1})^{\perp}$ be spanned by the singular vectors $g_1, \ldots, g_j$ and $E(C_{l+1})$ be spanned by $h_1, \ldots, h_{d-j}$. In particular, if $a_1^2 + \cdots + a_j^2 = 1$ and $v = a_1 f_1 + \cdots + a_j f_j$, then with respect to the $((g_k), (h_k))$ basis, v has coordinates $(\zeta_l a, \beta_l a)$. The nearest point in the unit sphere of $E(C_{l+1})$ has coordinates $\beta_l a / \|\beta_l a\|$ with respect to the (h_k) vectors. By the calculation in Lemma 13, the distance squared between the two points is $2 - 2\|\beta_l a\|$.

From the definition of the good set G, we get $\perp (F(C_l), E(C_{l+1})) > 6\delta$. Indeed we have $\perp (E_i(\sigma^{(l+1)n}\omega), F_i(\sigma^{(l+1)n}\omega)) > 10\delta$ from Lemma 5(a); $\angle(F_i(A^{(n)}_{\sigma^{ln}\omega}), F_i(\sigma^{(l+1)n}\omega)) < \delta$ and $\angle(E_j(A^{(n)}_{\sigma^{(l+1)n}\omega}), E_i(\sigma^{(l+1)n}\omega)) < \delta$ by Lemma 5(b) and (c). Finally the δ-closeness of $E(A^{(n)}_{\sigma^{(l+1)n}\omega})$ and $E(C_{l+1})$; and $F(C_l)$ and $F_j(A^{(n)}_{\sigma^{ln}\omega})$ comes from Lemma 3(a).

Combining the two previous paragraphs, we see $2 - 2\|\beta_l a\|$ exceeds $72\delta^2$, so that $1 - \|\beta_l a\| \geq 36\delta^2$ and $\|\zeta_l a\|^2 = (1 - \|\beta_l a\|)(1 + \|\beta_l a\|) \geq 36\delta^2$. In particular, we deduce

$$\|\zeta_l x\| \geq 6\delta\|x\| \text{ for all } x \in \mathbb{R}^j. \tag{17}$$

Let us also note that the p_l, q_l, r_l and t_l depend only on the choice of matrices from time 0 onwards and hence have been fixed by the conditioning, whereas B_0 is a random quantity whose conditional distribution we will study in §4.2.3.

Notice that the matrix Q_l is characterized by the property that if the coordinates of $x \in \mathbb{R}^d$ with respect to the $(E(C_l)^{\perp}, E(C_l))$ basis are given by z, then the coordinates of $(A^{\epsilon}_{\bar{\sigma}^{lN}\bar{\omega}})^{(N)}x$ are given by $Q_l z$ with respect to the $(E(C_{l+1})^{\perp}, E(C_{l+1}))$ basis.

Let

$$\begin{pmatrix} F_l \\ H_l \end{pmatrix} = Q_{l-1} \ldots Q_1 Q_0 \begin{pmatrix} I \\ B_0 \end{pmatrix}. \tag{18}$$

Recall that B_l is the matrix of V_l with respect to the $(E(C_l)^\perp, E(C_l))$ chart. Then, provided $V_l \cap E(C_l) = \{0\}$, we have

$$B_l = H_l F_l^{-1}. \tag{19}$$

To see this, consider a point x of V expressed in terms of the $(E(C_0)^\perp, E(C_0))$ basis as z. Then with respect to the $(E(C_l)^\perp, E(C_l))$ basis, $(A^\epsilon_{\bar\omega})^{(lN)}x$ has coordinates $Q_{l-1}\dots Q_0 z$. $(A^\epsilon_{\bar\omega})^{(lN)}V$ has a basis expressed in coordinates of the $(E(C_l)^\perp, E(C_l))$ basis given by the columns of $\begin{pmatrix} F_l \\ H_l \end{pmatrix}$. Post-multiplying by F_l^{-1} gives an alternative basis for $(A^\epsilon_{\bar\omega})^{(lN)}$ expressed in terms of the $(E(C_l)^\perp, E(C_l))$ basis, as the columns of $\begin{pmatrix} I \\ H_l F_l^{-1} \end{pmatrix}$ as required.

Our strategy for the remainder of the proof is to show that for most choices of the perturbation Δ_{-1}, the matrix B_n is of norm at most $2/\delta$. Since B_n is the matrix of $F^\epsilon_i(\bar\sigma^{nN}(\bar\omega))$ in the $(E(C_{n+1})^\perp, E(C_{n+1}))$ chart, Lemma 13 will ensure that $\perp\big(F^\epsilon_i(\bar\sigma^{nN}(\bar\omega)), E(C_{n+1})\big) > \delta/4$. Lemma 3 (c) will then give that $\angle\big(F^\epsilon_i(\bar\sigma^{(n+1)N}(\bar\omega)), F(C_{n+1})\big) < \delta/3$. Since by the goodness properties, $\angle\big(F(C_{n+1}), F_i(\sigma^{(n+1)N}(\omega))\big) < \delta$, we will obtain the required closeness.

Let

$$Q_{n-1}\dots Q_0 = \begin{pmatrix} W^{(n)} & X^{(n)} \\ Y^{(n)} & Z^{(n)} \end{pmatrix},$$

where the W's are $j\times j$, X's are $j\times(d-j)$, Y's are $(d-j)\times j$ and Z's are $(d-j)\times(d-j)$.

4.2.1. *Singular values and invertibility of $W^{(n)}$.* Let us now show that $\|W^{(n)}x\| \ge e^{3nN\tau}\|x\|$ for all x. Of course, this implies that $W^{(n)}$ is invertible. We start by proving $\|W^{(k)}x\| \ge \delta\|Y^{(k)}x\|$ for all x and $k \ge 0$.

Combining (13), (16) and (17) with the fact that $\|\gamma_k\|$, $\|\beta_k\|$ and $\|\alpha_k\|$ are all at most 1 since P_l is orthogonal, we have

$$\begin{aligned} \|q_k x\| &\ge 6\delta\|D_{2,k}x\| \text{ for all } x\in\mathbb{R}^j;\\ \|t_k\| &\le 2\\ \|p_k x\| &\le \|D_{2,k}x\| \text{ for all } x\in\mathbb{R}^j;\ \text{and}\\ \|r_k\| &\le 2. \end{aligned} \tag{20}$$

Notice that

$$\begin{pmatrix} W^{(k+1)} \\ Y^{(k+1)} \end{pmatrix} = \begin{pmatrix} q_k & t_k \\ p_k & r_k \end{pmatrix}\begin{pmatrix} W^{(k)} \\ Y^{(k)} \end{pmatrix}.$$

We have

$$\|W^{(k+1)}x\| \ge \|q_k W^{(k)}x\| - \|t_k Y^{(k)}x\| \ge 6\delta\|D_{2,k}W^{(k)}x\| - 2\|Y^{(k)}x\|, \tag{21}$$

where we used (17) to obtain the third inequality. Similarly

$$\|Y^{(k+1)}x\| \le \|p_k W^{(k)}x\| + \|r_k Y^{(k)}x\| \le \|D_{2,k}W^{(k)}x\| + 2\|Y^{(k)}x\|.$$

Suppose $\|W^{(k)}x\| \ge c_k\|Y^{(k)}x\|$ for all x. Since $\|D_{2,k}y\| \ge 2e^{4N\tau}\|y\|$ for all y, we see that $\|W^{(k+1)}x\| \ge c_{k+1}\|Y^{(k+1)}x\|$, where

$$c_{k+1} = \frac{6\delta c_k e^{4\tau N} - 1}{c_k e^{4\tau N} + 1}. \tag{22}$$

Using (C3), one checks

$$\text{if } c > \delta, \text{ then } \frac{6\delta c e^{4N\tau} - 1}{c e^{4N\tau} + 1} > \delta. \tag{23}$$

It is easy to see that $c_0 = 1$ ($W^{(0)} = I$ and $Y^{(0)} = 0$), so that $c_k > \delta$ for all k. We deduce that

$$\|W^{(k)}x\| \geq \delta \|Y^{(k)}x\|, \tag{24}$$

for all $k > 0$ and all $x \in \mathbb{R}^j$ as required. Hence we see that

$$\begin{aligned}
\|W^{(k+1)}x\| &\geq 12\delta e^{4\tau N} \|W^{(k)}x\| - 2\|Y^{(k)}x\| \\
&\geq (12\delta e^{4\tau N} - (2/\delta))\|W^{(k)}x\| \\
&\geq e^{3\tau N}\|W^{(k)}x\|,
\end{aligned}$$

where we used (21), (24) and the first part of (C3) for the respective inequalities.

In particular we deduce

$$\|W^{(n)}x\| \geq e^{3\tau n N}\|x\| \text{ for all } x \in \mathbb{R}^j. \tag{25}$$

4.2.2. *Recursion for E_n.* Let E_n be the cross ratio $Z^{(n)} - Y^{(n)}{W^{(n)}}^{-1}X^{(n)}$. Notice that

$$\begin{pmatrix} W^{(n)} & X^{(n)} \\ Y^{(n)} & Z^{(n)} \end{pmatrix} \begin{pmatrix} -{W^{(n)}}^{-1}X^{(n)} \\ I \end{pmatrix} = \begin{pmatrix} 0 \\ E_n \end{pmatrix}. \tag{26}$$

In fact, E_n may be defined this way: E_n is the unique lower submatrix M such that there exists A satisfying

$$\begin{pmatrix} W^{(n)} & X^{(n)} \\ Y^{(n)} & Z^{(n)} \end{pmatrix} \begin{pmatrix} A \\ I \end{pmatrix} = \begin{pmatrix} 0 \\ M \end{pmatrix}.$$

Now we have

$$\begin{aligned}
\begin{pmatrix} W^{(n+1)} & X^{(n+1)} \\ Y^{(n+1)} & Z^{(n+1)} \end{pmatrix} \begin{pmatrix} -{W^{(n)}}^{-1}X^{(n)} \\ I \end{pmatrix} &= \begin{pmatrix} q_n & t_n \\ p_n & r_n \end{pmatrix} \begin{pmatrix} W^{(n)} & X^{(n)} \\ Y^{(n)} & Z^{(n)} \end{pmatrix} \begin{pmatrix} -{W^{(n)}}^{-1}X^{(n)} \\ I \end{pmatrix} \\
&= \begin{pmatrix} q_n & t_n \\ p_n & r_n \end{pmatrix} \begin{pmatrix} 0 \\ E_n \end{pmatrix} \\
&= \begin{pmatrix} t_n E_n \\ r_n E_n \end{pmatrix}.
\end{aligned}$$

To finalise the recursion setup, we now seek matrices B and C such that

$$\begin{pmatrix} W^{(n+1)} & X^{(n+1)} \\ Y^{(n+1)} & Z^{(n+1)} \end{pmatrix} \begin{pmatrix} B \\ 0 \end{pmatrix} = \begin{pmatrix} -t_n E_n \\ C \end{pmatrix}. \tag{27}$$

Combining the above, we see

$$\begin{pmatrix} W^{(n+1)} & X^{(n+1)} \\ Y^{(n+1)} & Z^{(n+1)} \end{pmatrix} \begin{pmatrix} B - {W^{(n)}}^{-1}X^{(n)} \\ I \end{pmatrix} = \begin{pmatrix} 0 \\ r_n E_n + C \end{pmatrix},$$

so that $E_{n+1} = r_n E_n + C$.

From (27), we see that $B = -W^{(n+1)^{-1}} t_n E_n$, so that $C = Y^{(n+1)} B = -Y^{(n+1)} W^{(n+1)^{-1}} t_n E_n$. In particular, we obtain

$$E_{n+1} = \left(r_n - Y^{(n+1)} W^{(n+1)^{-1}} t_n\right) E_n.$$

Substituting $x = W^{(k)^{-1}} z$ in (24), we obtain $\|z\| \geq \delta \|Y^{(k)} W^{(k)^{-1}} z\|$, so that $\{\|Y^{(k)} W^{(k)^{-1}}\|\}_{k\in\mathbb{N}}$ is uniformly bounded by $1/\delta$. Furthermore, from the definition of Q_k, (15), and the choice of N, $\|r_k\|, \|t_k\| \leq 2$, so that $\|E_{k+1}\| \leq 2(1+\frac{1}{\delta})\|E_k\|$. Hence by (C3), we obtain

$$\|E_n\| \leq e^{\tau N n}. \tag{28}$$

Finally, we have

$$\begin{aligned} B_n &= (Y^{(n)} + Z^{(n)} B_0)(W^{(n)} + X^{(n)} B_0)^{-1} \\ &= (Y^{(n)} + (Y^{(n)} W^{(n)^{-1}} X^{(n)} + E_n) B_0)(W^{(n)} + X^{(n)} B_0)^{-1} \\ &= Y^{(n)}(I + W^{(n)^{-1}} X^{(n)} B_0)(W^{(n)} + X^{(n)} B_0)^{-1} + E_n B_0 (W^{(n)} + X^{(n)} B_0)^{-1} \\ &= Y^{(n)} W^{(n)^{-1}} + E_n B_0 (W^{(n)} + X^{(n)} B_0)^{-1}, \end{aligned} \tag{29}$$

where we used (18) and (19) in the first equality, and the definition of E_n in the second equality.

4.2.3. *Expression for B_0.* Recall that we conditioned on ω and $(\Delta_l)_{l\neq -1}$. This determines the top subspace at time -1, as well as the $E_j((A^{\epsilon}_{\bar\sigma^{kN}\bar\omega})^{(N)})$ and $E_j((A^{\epsilon}_{\bar\sigma^{kN}\bar\omega})^{(N)})^{\perp}$ spaces for each $k \geq 0$.

Let V be a $d \times j$ matrix whose columns consist of an orthonormal basis in $\mathbb{R}^d$ for the fast space at time -1. The fast space at time 0 has a basis given by the columns of $(A_{\sigma^{-1}\omega} + \epsilon\Delta_{-1})V$ (recall that Δ was assumed to be independent of the other perturbations of $(A(\sigma^n\omega))_{n\in\mathbb{Z}\setminus\{-1\}}$ that have already been fixed). For this section, we write A in place of $A_{\sigma^{-1}\omega}$ and Δ in place of Δ_{-1}. The coordinates of $(A+\epsilon\Delta)V$ in terms of the $(E_j((A^{\epsilon}_{\bar\sigma^{kN}\bar\omega})^{(N)})^{\perp}, E_j((A^{\epsilon}_{\bar\sigma^{kN}\bar\omega})^{(N)}))$ basis are given by

$$\begin{pmatrix} F^T \\ E^T \end{pmatrix} (A+\epsilon\Delta)V =: \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

where F is the matrix whose column vectors are the (orthonormal) basis for $E^{\perp}$ and E is the matrix whose column vectors are the orthonormal basis for E. Specifically, the jth column of this matrix gives the $(E^{\perp}, E)$ coordinates of the image of the jth basis vector of V under $A+\epsilon\Delta$. The matrix B_0 is then given by $Z_2 Z_1^{-1}$, that is $(E^T(A+\epsilon\Delta)V)(F^T(A+\epsilon\Delta)V)^{-1}$.

4.2.4. *Bounds on B_n.* Substituting the expression for B_0 into (29), we get

$$\begin{aligned} B_n = Y^{(n)} W^{(n)^{-1}} &+ E_n E^T (A+\epsilon\Delta)V(F^T(A+\epsilon\Delta)V)^{-1} \cdot \\ &\qquad \left[W^{(n)} + X^{(n)} E^T (A+\epsilon\Delta)V(F^T(A+\epsilon\Delta)V)^{-1}\right]^{-1} \\ = Y^{(n)} W^{(n)^{-1}} &+ E_n E^T (A+\epsilon\Delta)V \left[W^{(n)} F^T (A+\epsilon\Delta)V + X^{(n)} E^T (A+\epsilon\Delta)V\right]^{-1} \\ = Y^{(n)} W^{(n)^{-1}} &+ E_n E^T (A+\epsilon\Delta)V(UAV + \epsilon U\Delta V)^{-1}, \end{aligned}$$

where $U = W^{(n)} F^T + X^{(n)} E^T$.

We make the following definitions:

$$M = E_n E^T (A + \epsilon\Delta)V; \tag{30}$$
$$D = UAV; \qquad \text{and}$$
$$\tilde{\Delta} = D + \epsilon U \Delta V.$$

Now we have

$$B_n = Y^{(n)} {W^{(n)}}^{-1} + M\tilde{\Delta}^{-1}. \tag{31}$$

What remains is to give an upper bound on $\|M\|$ and to show that $\|\tilde{\Delta}^{-1}\|$ is small for a large set of Δ's.

4.2.5. *Bounds on $\|\tilde{\Delta}^{-1}\|$ using multivariate normal random variables.*

Reduction to normal random variables. We want to majorize $\mathbb{P}(\|(D + \epsilon U \Delta V)^{-1}\| \geq T)$. Let Z be a $d \times d$ matrix of independent standard normal random variables. We first show that a bound for $\mathbb{P}(\|(D + \epsilon UZV)^{-1}\| \geq T)$ yields a majorization of $\mathbb{P}(\|(D + \epsilon U\Delta V)^{-1}\| \geq T)$.

For $T > 0$, let R_T be the subset of $d \times d$ matrices C with entries in $[-1, 1]$ such that $\|(D + \epsilon UCV)^{-1}\| \geq T$.

Notice that $\mathbb{P}(Z \in R_T) = \int_{R_T} f_Z(X)\, dX$, where $f_Z(X) = (2\pi)^{-d^2/2} \exp(-\sum_{1 \leq i,j \leq d} X_{ij}^2/2)$ is the density function of the $d \times d$ matrices with $N(0, 1)$ entries. In particular $f_Z(X) \geq (2\pi e)^{-d^2/2}$ for all matrices X with entries in $[-1, 1]$, so that $\mathbb{P}(Z \in R_T) \geq (2\pi e)^{-d^2/2}\mathrm{Vol}(R_T)$ (where $\mathrm{Vol}(R_T)$ is the volume of R_T as a subset of $\mathbb{R}^{d^2}$). Similarly, since R_T is a subset of $\mathcal{U}$, $\mathbb{P}(\Delta \in R_T) = 2^{-d^2}\mathrm{Vol}(R_T)$. We see that $P(\Delta \in R_T) \leq (e\pi/2)^{d^2/2}\mathbb{P}(Z \in R_T)$, or

$$\mathbb{P}(\|(D + \epsilon U\Delta V)^{-1}\| \geq T) \leq (e\pi/2)^{d^2/2}\mathbb{P}(\|(D + \epsilon UZV)^{-1}\| \geq T). \tag{32}$$

Bound in the normal case. Now let $\tilde{Z} = D + \epsilon UZV$ and let $\hat{V}_k = \mathrm{span}\{c_1, \ldots, c_{k-1}, c_{k+1}, \ldots, c_j\}$, where the c_k are the columns of $\tilde{Z}$. Let $d_k = d(c_k, \hat{V}_k)$. Let $x = (x_1, \ldots, x_j)^T$. Now we have $\|\tilde{Z}x\| = \|x_k c_k + (x_1 c_1 + \ldots + x_{k-1}c_{k-1} + x_{k+1}c_{k+1} + \ldots + x_j c_j)\| \geq d(x_k c_k, \hat{V}_k) = |x_k| d_k$. We see that $\|\tilde{Z}x\| \geq \max_k(|x_k| d_k) \geq (\min_k d_k) \max_k |x_k| \geq \|x\| \min_k d_k/\sqrt{j}$. In particular, we deduce that $\|\tilde{Z}^{-1}\| \leq \sqrt{j}/\min_k d_k$. We now have

$$\mathbb{P}(\|\tilde{Z}^{-1}\| \geq T) \leq \mathbb{P}(\min_k d_k \leq \sqrt{j}/T) \leq \sum_k \mathbb{P}(d_k \leq \sqrt{j}/T). \tag{33}$$

Notice that the entries of the matrix $D + \epsilon UZV$ have a multivariate normal distribution. Any two such distributions with the same means and covariances are identically distributed. Recall that V is a $d \times j$ matrix whose columns are pairwise orthogonal. A consequence of this is that ZV has the same distribution as a $d \times j$ matrix of independent standard normal random variables. To see this, we see immediately that the expectation of each entry is 0. We then need to check the covariances, recalling that the columns of V are orthonormal, we get:

$$
\begin{aligned}
\operatorname{Cov}((ZV)_{ab},(ZV)_{cd}) &= \sum_{l,m} \operatorname{Cov}(Z_{al}V_{lb}, Z_{cm}V_{md}) \\
&= \sum_{l,m} V_{lb}V_{md} \operatorname{Cov}(Z_{al}, Z_{cm}) \\
&= \sum_{l,m} V_{lb}V_{md}\delta_{ac}\delta_{lm} \\
&= \delta_{ac}\sum_{m} V^T_{bm}V_{md} = \delta_{ac}(V^TV)_{bd} = \delta_{ac}\delta_{bd},
\end{aligned}
$$

as required.

We next observe (by an identical calculation) that F^TZV and E^TZV are distributed as independent $j \times j$ and $(d-j)\times j$ matrices with independent standard normal entries. Let $Z_1 = F^TZV$ and $Z_2 = E^TZV$. Recall that $\tilde{Z} = D + \epsilon UZV$ and $U = W^{(n)}F^T + X^{(n)}E^T$. By (33), we are interested in the columns of $\tilde{Z} = D + \epsilon W^{(n)}Z_1 + \epsilon X^{(n)}Z_2$.

For a fixed k, we compute the probability that the distance of the kth column of $\tilde{Z}$ is distant at least $\sqrt{j}/T$ from the span of the other columns. We give a uniform estimate on this probability conditioned on the columns of Z_1 other than the kth and the value of Z_2. Having fixed all of this data, let $\mathbf{n}$ be a unit normal vector to the $(j-1)$-dimensional space spanned by the other columns (a constant given the data). We then want to estimate $\mathbb{P}\left(|\mathbf{n}\cdot(D^{(k)} + \epsilon X^{(n)}Z_2^{(k)} + \epsilon W^{(n)}Z_1^{(k)})| < \sqrt{j}/T\right)$, where the superscript (k) indicates we are considering the kth column.

Let $A = \mathbf{n}\cdot(D^{(k)} + \epsilon X^{(n)}Z_2^{(k)})$ and $\mathbf{v} = \epsilon\mathbf{n}^TW^{(n)}$ (both are constant given the data on which we conditioned). We are therefore interested in $\mathbb{P}(|A + \mathbf{v}\cdot Z_1^{(k)}| < \sqrt{j}/T)$. This is bounded above by $\mathbb{P}(|\mathbf{v}\cdot Z_1^{(k)}| < \sqrt{j}/T)$. More multivariate normal machinery tells us that the distribution of $\mathbf{v}\cdot Z_1^{(k)}$ has the same distribution as $\|\mathbf{v}\|$ times a standard normal random variable, so we want to estimate $\mathbb{P}(|Z_0| < \sqrt{j}/(T\|\mathbf{v}\|))$, where Z_0 is a standard normal random variable. Simple estimates show this is less than $\sqrt{j}/(T\|\mathbf{v}\|)$ which, using (25) and the fact that $\mathbf{v} = \epsilon\mathbf{n}^TW^{(n)}$, is bounded above by $\sqrt{j}/(\epsilon e^{3\tau nN}T)$. Hence, $\mathbb{P}(\|\tilde{Z}^{-1}\| > T) \le j^{3/2}/(\epsilon T e^{3\tau nN})$. Hence we obtain

$$
\mathbb{P}\left(\|\tilde{\Delta}^{-1}\| > T\right) \le \frac{(e\pi/2)^{d^2/2} j^{3/2}}{\epsilon e^{3\tau nN}T} \text{ for any } T > 0. \tag{34}
$$

4.2.6. *Final estimates.* From (28) and (30), we have the upper bounds: $\|M\| \le e^{\tau nN}(1 + \|A\|) \le e^{2\tau nN}$ using the facts that $1 + \|A\| \le e^{\tau nN}$ since $\omega \in G_4$; and that $\|E^T\| = \|V\| = 1$. Combining this with (34) we obtain that

$$
\begin{aligned}
\bar{\mathbb{P}}(\|M\tilde{\Delta}^{-1}\| > 1/\delta) &\le \bar{\mathbb{P}}(\|\tilde{\Delta}^{-1}\| > 1/(\delta e^{2\tau nN})) \\
&\le \frac{\delta j^{3/2}(e\pi/2)^{d^2/2}}{\epsilon e^{\tau nN}}
\end{aligned} \tag{35}
$$

Recalling the expression for B_n given in (31) and that $N = C|\log\epsilon|$ and $\|Y^{(n)}{W^{(n)}}^{-1}\| \leq 1/\delta$, we get

$$\begin{aligned}\bar{\mathbb{P}}(\|B_n\| > 2/\delta) &\leq \bar{\mathbb{P}}(\|M\tilde{\Delta}^{-1}\| > 1/\delta)\\ &\leq \frac{j^{3/2}\delta(e\pi/2)^{d^2/2}}{e^{(\tau n-1/C)N}} < \chi/2,\end{aligned} \tag{36}$$

where we used (C4) for the final inequality.

For the last part of the proof, suppose that $\|B_n\| \leq 2/\delta$. Then Lemma 13 shows that $\perp(F_j^\epsilon(\bar{\sigma}^{nN}\bar{\omega}), E_j((A^\epsilon_{\bar{\sigma}^{nN}\bar{\omega}})^{(N)})) \geq \delta/6$. We extract two conclusions from the fact that $\sigma^{nN}\omega \in G_2$. Recall that $\delta \leq \delta_1 = \chi/2$. Lemma 3(c) yields that $\angle(F_j^\epsilon(\bar{\sigma}^{(n+1)N}\bar{\omega}), F_j((A^\epsilon_{\bar{\sigma}^{nN}\bar{\omega}})^{(N)})) \leq \chi/4$. Next, the hypotheses of Lemma 3 are satisfied with $A = A^{(N)}_{\sigma^{nN}\omega}$ and $B = (A^\epsilon_{\sigma^{nN}\omega})^{(N)}$. Conclusion (a) tells us that $\angle(F_j((A^\epsilon_{\bar{\sigma}^{nN}\bar{\omega}})^{(N)}), F_j(A^{(N)}_{\sigma^{nN}\omega})) \leq \chi/4$. Finally, since $\sigma^{nN}\omega \in G_1$, we have $\angle(F_j(\sigma^{(n+1)N}\omega), F_j(A^{(N)}_{\sigma^{nN}\omega})) < \chi/2$. Combining these we get

$$\angle\big(F_j^\epsilon(\bar{\sigma}^{(n+1)N}\bar{\omega}), F_j(\sigma^{(n+1)N}\omega)\big) < \chi.$$

By (36), $\bar{\mathbb{P}}(\|B_n\| > 2/\delta \mid \omega \in \tilde{G}) < \chi/2$. Thus,

$$\bar{\mathbb{P}}(\angle\big(F_j^\epsilon(\sigma^{(n+1)N}\bar{\omega}), F_j(\sigma^{(n+1)N}\omega)\big) > \chi \mid \omega) < \chi/2 \text{ for all } \omega \in \tilde{G}.$$

We have therefore established (12), and Claim 14 is proved.

□

Acknowledgments

The research of GF and CGT is supported by an ARC Future Fellowship and an ARC Discovery Project (DP110100068). AQ acknowledges NSERC, ARC DP110100068 for travel support and UNSW for hospitality during a research visit in 2012.

References

[1] L. Arnold and N. D. Cong. On the simplicity of the Lyapunov spectrum of products of random matrices. *Ergodic Theory Dynam. Systems*, 17(5):1005–1025, 1997.

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

[3] T. Bogenschütz. Stochastic stability of invariant subspaces. *Ergodic Theory Dynam. Systems*, 20(3):663–680, 2000.

[4] M. Dellnitz, G. Froyland, and S. Sertl. On the isolated spectrum of the Perron-Frobenius operator. *Nonlinearity*, 13(4):1171–1188, 2000.

[5] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM J. Numer. Anal.*, 36(2):491–515, 1999.

[6] G. Froyland. Unwrapping eigenfunctions to discover the geometry of almost-invariant sets in hyperbolic maps. *Physica D*, 237(6):840–853, 2008.

[7] G. Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Physica D: Nonlinear Phenomena*, 250(0):1 – 19, 2013.

[8] G. Froyland, C. González-Tokman, and A. Quas. Detecting isolated spectrum of transfer and Koopman operators with Fourier analysis tools. Submitted.

[9] G. Froyland, C. González-Tokman, and A. Quas. Stability and approximation of random invariant densities for Lasota-Yorke map cocycles. *Nonlinearity*, 27:647–660, 2014.

[10] G. Froyland, S. Lloyd, and A. Quas. Coherent structures and isolated spectrum for Perron-Frobenius cocycles. *Ergodic Theory Dynam. Systems*, 30(3):729–756, 2010.

[11] G. Froyland, S. Lloyd, and A. Quas. A semi-invertible Oseledets theorem with applications to transfer operator cocycles. *Discrete Contin. Dyn. Syst.*, 33(9):3835–3860, 2013.

[12] G. Froyland, S. Lloyd, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. *Physica D: Nonlinear Phenomena*, 239(16):1527 – 1541, 2010.

[13] G. Froyland and K. Padberg-Gehle. Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion. In W. Bahsoun, C. Bose, and G. Froyland, editors, *Ergodic Theory, Open Dynamics, and Coherent Structures*, volume 70 of *Proceedings in Mathematics and Statistics*, chapter 9, pages 171–216. Springer, 2014.

[14] G. Froyland and O. Stancevic. Escape rates and Perron-Frobenius operators: Open and closed dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B*, 14(2):457–472, 2010.

[15] C. González-Tokman, B. Hunt, and P. Wright. Approximating invariant densities of metastable systems. *Ergodic Theory and Dynamical Systems*, 31:1345–1361, 2011.

[16] C. González-Tokman and A. Quas. A semi-invertible operator Oseledets theorem. *Ergodic Theory and Dynamical Systems*. To appear.

[17] G. Keller and C. Liverani. Rare events, escape rates and quasistationarity: some exact formulae. *J. Stat. Phys.*, 135(3):519–534, 2009.

[18] F. Ledrappier and L.-S. Young. Stability of Lyapunov exponents. *Ergodic Theory Dynam. Systems*, 11(3):469–484, 1991.

[19] G. Ochs. Stability of Oseledets spaces is equivalent to stability of Lyapunov exponents. *Dynam. Stability Systems*, 14(2):183–201, 1999.

[20] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968.

[21] M. Raghunathan. A proof of Oseledec's multiplicative ergodic theorem. *Israel J. Math.*, 32(4):356–362, 1979.

[22] L.-S. Young. Random perturbations of matrix cocycles. *Ergodic Theory Dynam. Systems*, 6(4):627–637, 1986.

(Froyland and González-Tokman) School of Mathematics and Statistics, University of New South Wales, Sydney, NSW, 2052, Australia

(Quas) Department of Mathematics and Statistics, University of Victoria, Victoria, BC, CANADA, V8W 3R4