

# UNIVERSAL GAP GROWTH FOR LYAPUNOV EXPONENTS OF PERTURBED MATRIX PRODUCTS

JASON ATNIP, GARY FROYLAND, CECILIA GONZÁLEZ-TOKMAN,  
AND ANTHONY QUAS

ABSTRACT. We study the quantitative simplicity of the Lyapunov spectrum of  $d$ -dimensional bounded matrix cocycles subjected to additive random perturbations. In dimensions 2 and 3, we establish explicit lower bounds on the gaps between consecutive Lyapunov exponents of the perturbed cocycle, depending only on the scale of the perturbation. In arbitrary dimensions, we show existence of a universal lower bound on these gaps. A novelty of this work is that the bounds provided are uniform over all choices of the original sequence of matrices. Furthermore, we make no stationarity assumptions on this sequence. Hence, our results apply to random and sequential dynamical systems alike.

## 1. INTRODUCTION

We study the Lyapunov spectrum of sequences of matrices with additive noise-like perturbations: we start with a sequence of matrices, which we assume to be of norm at most 1. The entries of each matrix are independently perturbed by adding a random number uniform in the range  $[-\varepsilon, +\varepsilon]$ . The unperturbed matrices are denoted  $(A_i)_{i \in \mathbb{N}}$ , and we write  $A_{i,\varepsilon}$  to denote the perturbed matrix  $A_i + \varepsilon \Xi_i$  where  $\Xi_i$  are independent matrix random variables with independent entries uniformly distributed on  $[-1, 1]$ . We are interested in universal quantitative simplicity of the Lyapunov spectrum. That is, we establish lower bounds on the exponential growth rate of the ratio of the  $j$ th and  $(j + 1)$ st singular values of the product  $A_\varepsilon^{(n)} = A_{n,\varepsilon} \cdots A_{1,\varepsilon}$ . Our bounds depend only on the maximum amplitude  $\varepsilon$  of the noise, and not on the particular sequence of matrices.

Our results are in part motivated by questions concerning the behaviour of Lyapunov exponents of differentiable dynamical systems in the presence of noise. If Lyapunov exponents are estimated using measurements of the derivative cocycle, it is natural to model the measurement error as a random perturbation of the underlying derivative cocycle. We focus on the abstracted question: given an arbitrary bounded sequence of matrices, subjected to noise-like perturbations, what can be said about simplicity of Lyapunov spectrum? Since our results are stated in this general context, it

follows that for *all* unperturbed orbits of a dynamical system, and for almost all choices of the perturbations, we obtain universal gaps between consecutive Lyapunov exponents. By contrast, standard methods based on ergodic theory would likely yield results for almost all orbits. Finally, we remark that our methods are highly robust, and should apply to other kinds of absolutely continuous noise, including situations where the structure of the noise is dependent on the point in the dynamical system.

Since we place no stationarity assumption on the original sequence, the limits defining the Lyapunov exponents may not exist. Specifically, we show the following.

**Theorem A.** *Let  $0 < \varepsilon < 1$ . There exists  $c_2(\varepsilon) > 0$  such that for any sequence  $(A_n)$  of  $2 \times 2$  matrices, each of norm at most 1, the random variable*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_1(A_\varepsilon^{(n)})}{s_2(A_\varepsilon^{(n)})}$$

*is almost surely constant; and the constant is at least  $c_2(\varepsilon)$ . Further  $c_2(\varepsilon) > \exp(-1/\varepsilon^{35})$  for sufficiently small  $\varepsilon$ .*

**Theorem B.** *Let  $0 < \varepsilon < 1$ . There exists  $c_3(\varepsilon) > 0$  such that for any sequence  $(A_n)$  of  $3 \times 3$  matrices, each of norm at most 1, and for  $j = 1$  or  $2$ , the random variable*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_{j+1}(A_\varepsilon^{(n)})}$$

*is almost surely constant; and the constant is at least  $c_3(\varepsilon)$ . Further  $c_3(\varepsilon) > \exp(-1/\varepsilon^{867})$  for sufficiently small  $\varepsilon$ .*

In the case  $d > 3$ , we have the following theorem showing that the Lyapunov spectrum is (quantitatively) non-trivial.

**Theorem C.** *Let  $d > 3$  and  $0 < \varepsilon < 1$ . There exists  $c'_d(\varepsilon) > 0$  such that for any sequence  $(A_n)$  of  $d \times d$  matrices, each of norm at most 1, the random variable*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_1(A_\varepsilon^{(n)})}{s_d(A_\varepsilon^{(n)})}$$

*is almost surely constant; and the constant is at least  $c'_d(\varepsilon)$ . For small  $\varepsilon$ ,  $c'_d(\varepsilon) > \exp(-1/\varepsilon^{16d+3})$ .*

**Theorem D.** *Let  $d > 1$  be arbitrary. If the unperturbed  $d \times d$  matrices  $A_n$  are orthogonal for each  $n \in \mathbb{N}$ , then there exists a constant  $c(\varepsilon) > 0$  depending only on  $\varepsilon$  such that for each  $1 \leq j < d$  we have*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_{j+1}(A_\varepsilon^{(n)})} \geq c(\varepsilon).$$

*For any  $\delta > 0$ , for all small  $\varepsilon$ ,  $c(\varepsilon) > \exp(-1/\varepsilon^{2+\delta})$ .*

We remark that this result should be compared with [7], where the matrices  $(A_n)$  are orthogonal and the perturbations  $\Xi$  are taken to be matrices with independent standard normal entries. In that paper, using symmetry properties of multi-variate normal random variables, an exact expression for the Lyapunov exponents is obtained, as well as the leading terms of a Taylor-like expansion. In any dimension in that setting, the gaps obtained are asymptotically  $\varepsilon^2$ .

The next theorem provides universal positive lower bounds on all Lyapunov exponent gaps for perturbed cocycles. This result gives no quantitative information on the gaps, beyond positivity, because it relies on compactness, and does not exhibit an explicit mechanism for the gaps.

**Theorem E.** *Let  $d \geq 2$  and  $0 < \varepsilon < 1$ . There exists  $c_d(\varepsilon) > 0$  such that for any sequence  $(A_n)$  of  $d \times d$  matrices, each of norm at most 1, and for each  $1 \leq j \leq d - 1$ , the random variable*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_{j+1}(A_\varepsilon^{(n)})}$$

*is almost surely constant; and the constant is at least  $c_d(\varepsilon)$ .*

Finally it is worth pointing out that while we focus on the case of uniform i.i.d. perturbations, our methods would also apply in the case of absolutely continuous i.i.d. perturbations with density bounded away from infinity; and away from zero around 0.

**1.1. Context.** The foundational work of Lyapunov [34] relates quantities of the form

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\varphi(t)v\|$$

to stability of solutions of differential equations in  $\mathbb{R}^d$ , where  $v \in \mathbb{R}^d$  and  $\varphi(t)$  is a  $d \times d$  matrix—the fundamental solution matrix. These quantities, now called *Lyapunov exponents*, have become fundamental for the study of stability and chaos in dynamical systems. Roughly speaking, negative Lyapunov exponents correspond to stable behaviour and positive Lyapunov exponents are a sign of chaos.

For systems observed in discrete time, Lyapunov exponents take the form

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n \dots A_1 v\|,$$

where  $A_1, A_2, \dots$  are  $d \times d$  matrices and  $v \in \mathbb{R}^d$ . Furstenberg and Kesten showed in [23] that if the matrices  $A_j$  are drawn from an ergodic process, then the limit  $\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n \dots A_1\|$  exists and is almost surely constant, although not necessarily finite. In a similar setting, Oseledets showed in [36] that there are between 1 and  $d$  distinct Lyapunov exponents,  $\lambda_1 > \dots > \lambda_k$ , with multiplicities  $m_1, \dots, m_k$  summing to  $d$ , quantifying the exponential rates of expansion or contraction that different vectors can

experience, asymptotically. We refer the reader to the monograph of Viana [39] for classical theory and recent developments on this topic.

Lyapunov exponents are generally difficult to compute and approximate. In fact, determining whether Lyapunov exponents are non-zero or if there are multiplicities remain difficult problems in general. A  $d$ -dimensional system is said to have *simple Lyapunov spectrum* if it has  $d$  distinct Lyapunov exponents. The question of simplicity of Lyapunov spectrum has received considerable attention in the literature, first in the case of i.i.d. matrices and then under increasingly more general stationary driving processes. Works in this direction include Guivarc'h and Raugi [27]; Gol'dsheid and Margulis [24]; Arnold and Cong [1, 2]; Bonatti and Viana [15]; Avila and Viana [3]; Matheus, Möller and Yoccoz [35]; Poletti and Viana [37]; Backes, Poletti, Varandas and Lima [5]. Establishing simplicity of Lyapunov spectra allowed Avila and Viana to resolve the Zorich–Kontsevich conjecture in [4].

The works above build on the theory developed by Furstenberg [21]. This theory has been used to obtain qualitative rather than quantitative results. That is, simplicity is established without providing explicit information on the gaps between Lyapunov exponents. It is also worth mentioning that a non-stationary version of Furstenberg's theorem for random matrix products has been recently proved by Gorodetski and Kleptsyn in [26].

Despite significant progress, Lyapunov exponents for products of matrices are still a source of major challenges, even in dimension two. For instance, the study of Lyapunov exponents for two-dimensional maps, such as the Hénon map and the Chirikov standard map, falls in this category. Benedicks and Carleson have shown positive Lyapunov exponents for Hénon maps in [8]. While the problem remains unsolved for the standard map, Blumenthal, Xue and Young in [9, 10], have established positive lower bounds on Lyapunov exponents for very small stochastic perturbations of the standard map, and other two-dimensional maps.

From a broader perspective, our results may be contrasted with previous works on genericity of trivial Lyapunov spectrum. The papers of Bochi [12] and of Bochi and Viana [14] show that for a generic non-hyperbolic cocycle (over a continuous base), the Lyapunov spectrum is trivial. That is, they show that by making very careful perturbations to the cocycle, one can get the Lyapunov exponents to collapse. In contrast, our results show that for any initial cocycle (including non-stationary) and almost every sequence of random perturbations, the Lyapunov spectrum is simple, and in dimension two or three, they provide explicit bounds on the gaps, only depending on the perturbation size.

The key source of difficulty in the general setting with Lyapunov exponents is that singular values may be far from multiplicative. One has bounds such as  $s_1(A)s_d(B) \leq s_1(AB) \leq s_1(A)s_1(B)$  which are far too weak to control limits. Indeed, this issue underlies the theorems of Bochi and Bochi–Viana. On the other hand, in our randomly perturbed setting, singular values are approximately multiplicative (as embodied here by Lemma

3.4) and this is what makes our theorems work. This approach builds on previous work of Froyland, González-Tokman and Quas [20], and is ultimately inspired by work of Ledrappier and Young [32]. The idea of controlling sub-multiplicative quantities with nearly multiplicative ones has also been exploited in other studies of Lyapunov exponents, in the form of an avalanche principle, introduced by Goldstein and Schlag in [25] and expanded upon in Duarte and Klein’s monograph [18]. Apart from specific cases where Lyapunov exponents have been fully computed (see [7] and references therein), this work seems to be the first one where explicit lower bounds on Lyapunov exponent gaps have been found for random perturbations of cocycles with arbitrary Lyapunov spectrum.

The use of *stochastic perturbations* as a tool for understanding dynamical systems has a long history, described e.g. by Kifer in [31]. More recently, the potential of this approach to provide theoretical and practical insights into the long-term behaviour of complex dynamical systems has been highlighted by Young in [42, 43]. On one hand, random perturbations are a natural way to model phenomena evolving in the presence of noise. On the other hand, by taking zero-noise limits, stochastic perturbations have been effectively used to gain information about dynamical systems, since the work of Khas’minskii [30]. Results about stochastic stability and Lyapunov exponents for randomly perturbed dynamical systems include the works of Young [41], Ledrappier and Young [32]; Imkeller and Lederer [29]; Baxendale and Goukasian [6]; Cowieson and Young [17]; Lian and Stenlund [33]; Froyland, González-Tokman and Quas [20, 19]; Blumenthal and Yun [11]; Chemnitz and Engel [16]; Bednarski and Quas [7]. Progress on the related question of continuity of Lyapunov exponents has been recently reviewed by Viana in [40].

**Acknowledgments.** This research has been funded by the Australian Research Council (ARC DP220102216) and NSERC. The authors thank Alex Blumenthal and Paulo Varandas for bibliographic suggestions.

## 2. PRELIMINARIES AND NOTATION

For a  $d \times d$  matrix  $M$ , we write  $\|M\|$  for its operator norm, that is  $\|M\| = \sup_{\|x\|_2=1} \|Mx\|_2$  and  $|M|_\infty$  for  $\max |m_{ij}|$ . Let  $(A_n)_{n \in \mathbb{N}}$  be a fixed sequence of  $d \times d$  matrices of (operator) norm at most 1 and let  $(\Xi_n)_{n \in \mathbb{N}}$  be a family of independent identically distributed  $d \times d$  matrix random variables with mutually independent entries, each distributed as  $U([-1, 1])$ . Let  $A_{n,\varepsilon} = A_n + \varepsilon \Xi_n$  for  $\varepsilon > 0$ . Define  $A^{(n)} := A_n A_{n-1} \cdots A_1$  and (for  $\varepsilon > 0$ )  $A_\varepsilon^{(n)} := A_{n,\varepsilon} A_{n-1,\varepsilon} \cdots A_{1,\varepsilon}$ .

For a  $d \times d$  matrix  $A$ , let  $s_1(A) \geq \cdots \geq s_d(A)$  be its singular values and for  $1 \leq j \leq k \leq d$  set  $S_j^k(A) = s_j(A) \cdots s_k(A)$  be the product of the  $j$ th singular value through the  $k$ th singular value. For non-zero vectors  $u$  and  $v$  in  $\mathbb{R}^d$ , we define  $\angle(u, v) = \|u/\|u\| - v/\|v\|\|$ . For subsets  $C_1$  and  $C_2$  of  $\mathbb{R}^d$ , we define  $d(C_1, C_2) = \inf_{x \in C_1, y \in C_2} \|x - y\|$ . If the sets are disjoint, one is

compact and the other is closed, this quantity is positive and the infimum is attained. We denote the unit sphere in  $\mathbb{R}^d$  by  $S$ .

**Lemma 2.1.** *Let  $\mathbb{R}^d$  be expressed as a direct sum (not necessarily orthogonal)  $E \oplus F$  and let  $\Pi_{E\|F}$  and  $\Pi_{F\|E}$  be projections of  $\mathbb{R}^d$  onto  $E$  and  $F$  respectively so that for any  $x \in \mathbb{R}^d$ ,  $\Pi_{E\|F}(x) + \Pi_{F\|E}(x)$  is the unique decomposition of  $x$  in  $E \oplus F$ . Let  $\delta = d(E \cap S, F \cap S)$  where  $S$  is the unit sphere. Then  $\|\Pi_{E\|F}\|$  and  $\|\Pi_{F\|E}\|$  are at most  $2/\delta$ .*

*Proof.* Let  $x \in S$  and write  $x$  as  $e - f$ . We then have  $\|e/\|e\| - f/\|e\|\| = 1/\|e\|$  and  $\|f/\|e\| - f/\|f\|\| = \|\|f\|/\|e\| - 1\| = \|\|f\| - \|e\|\|/\|e\| \leq \|f - e\|/\|e\| = 1/\|e\|$ . Hence by the triangle inequality we see  $\|e/\|e\| - f/\|f\|\| \leq 2/\|e\|$ . Since  $e = \Pi_{E\|F}(x)$ , this gives  $\|\Pi_{E\|F}(x)\| \leq 2/\|(e/\|e\|) - (f/\|f\|)\| \leq 2/\delta$ .  $\square$

We observe that this proof does not make use of the fact that  $\mathbb{R}^d$  is Euclidean, so it applies in arbitrary normed spaces.

**Lemma 2.2.** *Let  $u \in S$  be a unit vector in  $\mathbb{R}^d$  and let  $V$  be a subspace of  $\mathbb{R}^d$ . Then  $d(u, V \cap S) \leq 2d(u, V)$ .*

*Proof.* If  $d(u, V) = 1$ , the conclusion follows from the triangle inequality. Otherwise, let  $v \in V \setminus \{0\}$  be such that  $d(u, V) = \|u - v\|$ . Then  $d(u, V \cap S) \leq \|u - v/\|v\|\| \leq \|u - v\| + \|v - v/\|v\|\| = \|u - v\| + \|\|v\| - 1\| = \|u - v\| + \|\|v\| - \|u\|\| \leq 2\|u - v\| = 2d(u, V)$ .  $\square$

This proof also works in normed spaces. In a Euclidean space, the constant can be improved from 2 to  $\sqrt{2}$ .

### 3. GENERAL STRATEGY

In this section we present the general strategy for proving Theorems [A](#) to [D](#). We begin with the following definition: Given a  $d \times d$  matrix  $A$  and  $1 \leq j < k \leq d$ , we say that  $A$  has a  $(j, k)$ -gap if  $\frac{s_j(A)}{s_k(A)} \geq \mathcal{G}$ , where  $\mathcal{G}$  is a constant to be determined depending on  $\varepsilon$ .

The main idea of the proof is to break the sequence of matrices  $(A_n)$  into blocks of length  $N$ ,  $B_i := A_{(i+1)N} \cdots A_{iN+1}$ ,  $i \geq 0$ , where  $N$  is a constant to be determined based on  $\varepsilon$ . We then show for each such block  $B_i$ , how to construct a *target block*  $B_i^\circ$  where  $B_i^\circ$  is of the form  $A_{(i+1)N}^\circ \cdots A_{iN+1}^\circ$  where  $\|A_l^\circ - A_l\|_\infty$  is small for each  $l$  and the target blocks  $B_i^\circ$  have a  $(j, k)$ -gap.

First, in [§3.1](#) we present a procedure for perturbing a matrix to make it sufficiently non-singular. The full details of the target block construction will be given in [§4](#). In [§3.2](#), for a given target block  $B_i^\circ$  with a  $(j, k)$ -gap, we show that any block  $B_i^\times$  sufficiently close to  $B_i^\circ$  must also have a  $(j, k)$ -gap. Then in [§3.3](#), we show that, when two blocks of any length where the random perturbations have been fixed are separated by a single matrix where the random perturbation is not yet determined, the logarithmic  $(j, k)$  gap of the

combined block is the sum of logarithmic gaps of the outside two blocks plus a random variable with a tight distribution. The gap size is chosen to ensure that the gains from hitting targets dominate the losses from this joining procedure. In §3.4, we use the Kolmogorov 0–1 law to show that the growth rate of the  $(j, k)$  gap is constant almost surely. In §3.5, we present a book-keeping procedure, Theorem 3.8, relying on the gluing lemma (Lemma 3.4), which provides a quantitative lower bound on the exponential growth rate of  $(j, k)$  gaps, for small perturbations of target blocks.

**3.1. Non-Singular Initialization.** Our target blocks in Lemma 3.2 will be products of non-singular matrices. The following lemma provides a simple way to construct non-singular matrices near possibly singular matrices.

**Lemma 3.1.** *Let  $d \geq 2$ . Let  $A$  be a  $d \times d$  matrix with  $\|A\| \leq 1$  and  $0 < \varepsilon < 1$ . Then, there exists  $A'$  such that  $\|A' - A\| < \frac{\varepsilon}{2}$ ,  $\|A'\| \leq \max(\frac{1}{2}, 1 - \frac{\varepsilon}{2})$  and  $s_d(A') \geq \frac{\varepsilon}{2}$ .*

*Proof.* Let  $A = VDU^T$  be a singular value decomposition of  $A$ . Then,  $A' = V((1 - \varepsilon)D + \varepsilon(\frac{1}{2}))U^T$  satisfies the required property. Indeed, if  $s_1, \dots, s_d$  are the singular values of  $A$ , then the singular values of  $A'$  are  $(1 - \varepsilon)s_1 + \varepsilon\frac{1}{2}, \dots, (1 - \varepsilon)s_d + \varepsilon\frac{1}{2}$ . Also,  $\|A' - A\| \leq \frac{\varepsilon}{2}$ .  $\square$

In fact by [38, Lemma 1], the singular value decomposition is measurable, so that  $U$ ,  $D$  and  $V$  may be chosen to depend measurably on  $A$ , so that  $A'$  may be taken to be a measurable function of  $A$ .

**3.2. Triangle Inequality for Target Block.** The following continuity lemma shows that under arbitrary suitably small perturbations, the singular value ratios for subproducts of matrices of length less than  $N$  cannot decrease by more than half.

**Lemma 3.2.** *Let  $d > 1$ ,  $1 \leq m \leq m' \leq N - 1$  and  $0 < \varepsilon < 1$ . Suppose  $A_m^\circ, \dots, A_{m'}^\circ$ , called targets, are  $d \times d$  matrices of norm at most  $1 - \frac{\varepsilon}{4}$  such that  $s_d(A_i^\circ) \geq \frac{\varepsilon}{4}$  for every  $m \leq i \leq m'$ . For each  $m \leq i \leq m'$ , let  $A_i^\times$  be such that  $|A_i^\times - A_i^\circ|_\infty < (\frac{\varepsilon}{4})^N / (3dN)$ . Then, for every  $1 \leq j < k \leq d$  we have*

$$(3.1) \quad \frac{s_j(A_{m'}^\times \dots A_m^\times)}{s_k(A_{m'}^\times \dots A_m^\times)} \geq \frac{1}{2} \frac{s_j(A_{m'}^\circ \dots A_m^\circ)}{s_k(A_{m'}^\circ \dots A_m^\circ)}.$$

*Proof.* Note that  $\|A_i^\times - A_i^\circ\| \leq d|A_i^\times - A_i^\circ|_\infty < (\frac{\varepsilon}{4})^N / (3N)$ . Since  $\|A_i^\times\| \leq \|A_i^\circ\| + \|A_i^\times - A_i^\circ\|$ , we have  $\|A_i^\times\|$  and  $\|A_i^\circ\|$  are at most 1 for each  $m \leq i \leq m'$ . Thus, the triangle inequality implies

$$\begin{aligned} \|A_{m'}^\circ \dots A_m^\circ - A_{m'}^\times \dots A_m^\times\| &\leq \sum_{i=m}^{m'} \|A_{m'}^\circ \dots A_i^\circ A_{i-1}^\times \dots A_m^\times - A_{m'}^\circ \dots A_{i+1}^\circ A_i^\times \dots A_m^\times\| \\ &= \sum_{i=m}^{m'} \|A_{m'}^\circ \dots A_{i+1}^\circ (A_i^\circ - A_i^\times) A_{i-1}^\times \dots A_m^\times\| \end{aligned}$$

$$\leq \sum_{i=m}^{m'} \|A_i^\odot - A_i^\times\| \leq (\frac{\varepsilon}{4})^N / 3.$$

Using the well-known fact<sup>1</sup> that  $|s_j(A) - s_j(A')| \leq \|A - A'\|$  for all  $j$ , we see

$$\begin{aligned} \frac{s_j(A_{m'}^\times \dots A_m^\times)}{s_k(A_{m'}^\times \dots A_m^\times)} &\geq \frac{s_j(A_{m'}^\odot \dots A_m^\odot) - (\frac{\varepsilon}{4})^N / 3}{s_k(A_{m'}^\odot \dots A_m^\odot) + (\frac{\varepsilon}{4})^N / 3} \\ &= \frac{s_j(A_{m'}^\odot \dots A_m^\odot)}{s_k(A_{m'}^\odot \dots A_m^\odot)} \left( \frac{1 - (\frac{\varepsilon}{4})^N / (3s_j(A_{m'}^\odot \dots A_m^\odot))}{1 + (\frac{\varepsilon}{4})^N / (3s_k(A_{m'}^\odot \dots A_m^\odot))} \right). \end{aligned}$$

Since  $s_j(A_{m'}^\odot \dots A_m^\odot)$  and  $s_k(A_{m'}^\odot \dots A_m^\odot)$  are at least  $s_d(A_{m'}^\odot) \dots s_d(A_m^\odot) \geq (\frac{\varepsilon}{4})^N$ , we obtain

$$\frac{s_j(A_{m'}^\times \dots A_m^\times)}{s_k(A_{m'}^\times \dots A_m^\times)} \geq \frac{s_j(A_{m'}^\odot \dots A_m^\odot)}{s_k(A_{m'}^\odot \dots A_m^\odot)} \left( \frac{1 - \frac{1}{3}}{1 + \frac{1}{3}} \right),$$

so that the desired inequality holds.  $\square$

We denote by  $\mathbb{P}$  the i.i.d. probability measure on sequences of matrices with independent uniformly distributed entries in  $[-1, 1]$ . The probability that the perturbed matrices are close enough along a sampled block to satisfy the hypothesis of Lemma 3.2 has a straightforward lower bound.

**Corollary 3.3.** *Let  $d > 1$ . For all  $0 < \varepsilon < 1$  there exists  $p \in (0, 1)$  (for example  $p = (\frac{\varepsilon}{4})^{N^2 d^2} / (3dN)^{Nd^2}$  works) such that for any target,  $A_m^\odot, \dots, A_{m'}^\odot$  as in Lemma 3.2 with  $1 \leq m \leq m' \leq N - 1$ ,*

$$\mathbb{P}(|A_{i,\varepsilon} - A_i^\odot|_\infty < (\frac{\varepsilon}{4})^N / (3dN) \text{ for all } i = m, \dots, m') \geq p.$$

**3.3. Gluing Lemma.** For  $1 \leq j < k \leq d$  and a  $d \times d$  matrix  $A$ , let  $q_{j,k}(A) = \frac{s_j(A)}{s_k(A)}$ . We define a function  $F$  whose arguments are  $d \times d$  matrices by  $F(L, A, R) = \log q_{j,k}(LAR) - \log q_{j,k}(L) - \log q_{j,k}(R)$ .

**Lemma 3.4.** *Let  $d > 1$  and let  $1 \leq j < k \leq d$ . Then there exists  $\zeta > 0$  such that for all invertible  $d \times d$  matrices  $L$  and  $R$  and for all  $A$  such that  $\|A\| \leq 1$ , all  $X \geq 0$  and all  $0 < \varepsilon < 1$ , we have*

$$\mathbb{P}(|F(L, A + \varepsilon \Xi, R)| > X) \leq \min\{1, \zeta e^{-\frac{X}{4d}} \varepsilon^{-1}\},$$

where  $\Xi$  is a matrix-valued random variable with independent  $U[-1, 1]$  entries.

*Proof.* Recall that  $S_1^\ell(A) = s_1(A) \dots s_\ell(A)$  for each  $1 \leq \ell \leq d$ , and note that

$$(3.2) \quad \log q_{j,k}(A) = \log S_1^j(A) - \log S_1^{j-1}(A) + \log S_1^{k-1}(A) - \log S_1^k(A).$$

<sup>1</sup>This follows from the so-called max-min inequality:  $s_j(A) = \max_{\dim V=j} \min_{v \in V \cap S} \|(A - A' + A')v\| \leq \|A - A'\| + \max_{\dim V=j} \min_{v \in V \cap S} \|A'v\| = \|A - A'\| + s_j(A')$ .



Using the fact that  $\|A\| \leq 1$ , and the fact that the entries of  $\Xi$  are all in  $[-1, 1]$ , we first note that

$$(3.3) \quad \log S_1^\ell(L(A + \varepsilon\Xi)R) \leq \log S_1^\ell(L) + \log S_1^\ell(R) + d \log(1 + \varepsilon d).$$

Following the proof of Lemma 3.5 in [20] we write  $L = O_1 D_1 O_2$  where  $O_1$  and  $O_2$  are orthogonal and  $D_1$  is diagonal with non-negative entries arranged in decreasing order. Similarly we write  $R = O_3 D_2 O_4$ , where  $O_3$  and  $O_4$  are orthogonal and  $D_2$  is diagonal. Let  $A' = O_2 A O_3$  and  $\Xi' = (1/d) O_2 \Xi O_3$  and  $C_\ell$  be the diagonal matrix with 1's in the first  $\ell$  diagonal positions and 0's elsewhere. (The choice of normalization of  $\Xi'$  is to ensure that its entries remain in  $[-1, 1]$ ). In this case we have

$$S_1^\ell(L(A + \varepsilon\Xi)R) = S_1^\ell(D_1(A' + \varepsilon d \Xi')D_2), \quad S_1^\ell(L) = S_1^\ell(D_1), \quad \text{and} \quad S_1^\ell(R) = S_1^\ell(D_2).$$

We briefly summarize the next steps of the proof of Lemma 3.5 in [20]. Using the facts that  $S_1^\ell(OM) = S_1^\ell(M) = S_1^\ell(MO)$  for an orthogonal matrix  $O$  and  $S_1^\ell(CM), S_1^\ell(MC) \leq S_1^\ell(M)$ , we have  $S_1^\ell(L(A + \varepsilon\Xi)R) = S_1^\ell(O_1 D_1 O_2 (A + \varepsilon\Xi) O_3 D_2 O_4) = S_1^\ell(D_1 O_2 (A + \varepsilon\Xi) O_3 D_2) \geq S_1^\ell(C_\ell D_1 O_2 (A + \varepsilon\Xi) O_3 D_2 C_\ell)$ . Now since  $C_\ell D_1 = C_\ell D_1 C_\ell$  and since  $S_1^\ell$  is multiplicative for matrices whose non-zero entries are in the top left  $\ell \times \ell$  submatrix, we obtain

$$(3.4) \quad \log S_1^\ell(L(A + \varepsilon\Xi)R) \geq \log S_1^\ell(L) + \log S_1^\ell(R) + \log S_1^\ell(C_\ell(A' + \varepsilon d \Xi')C_\ell).$$

Combining (3.3) and (3.4), we deduce for each  $\ell \leq d$  we have

$$(3.5) \quad \left| \log S_1^\ell(L(A + \varepsilon\Xi)R) - \log S_1^\ell(L) - \log S_1^\ell(R) \right| \leq d \log(1 + \varepsilon d) + \left| \log S_1^\ell(C_\ell(A' + \varepsilon d \Xi')C_\ell) \right|.$$

For each  $1 \leq \ell \leq d$  let  $A_\ell''$  and  $\Xi_\ell''$  be the top-left  $\ell \times \ell$  submatrices of  $A'$  and  $\Xi'$  respectively and note that  $S_1^\ell(C_\ell(A' + \varepsilon d \Xi')C_\ell) = |\det(A_\ell'' + \varepsilon d \Xi_\ell'')|$ .

Hence (3.2) gives

$$(3.6) \quad |F(L, A + \varepsilon\Xi, R)| \leq 4d \log(1 + \varepsilon d) + \sum_{\ell \in \{j-1, j, k-1, k\}} \left| \log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')| \right|.$$

In order to have  $|F(L, A + \varepsilon\Xi, R)| > X$ , it is therefore necessary that one of the  $|\log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')||$  terms is greater than  $\frac{X}{4} - d \log(1 + \varepsilon d)$ .

Since  $A_\ell'' + \varepsilon d \Xi_\ell'' = C_\ell O_2 (A + \varepsilon\Xi) O_3 C_\ell$ ,  $\|A_\ell'' + \varepsilon d \Xi_\ell''\| \leq \|A + \varepsilon\Xi\| \leq 1 + \varepsilon d$  for each  $\ell \leq d$ , so that  $\log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')| \leq d \log(1 + \varepsilon d)$  for all  $\Xi$  and all  $\ell \leq d$ . We have shown in the case  $X > 8d \log(1 + \varepsilon d)$ , that it is impossible for  $\log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')|$  to exceed  $\frac{X}{4} - d \log(1 + \varepsilon d)$ . Hence in order to have  $|F(L, A + \varepsilon\Xi, R)| > X$ , it is necessary that one of the  $\log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')|$  terms is *less than*  $-(\frac{X}{4} - d \log(1 + \varepsilon d))$ .

That is for  $X > 8d \log(1 + \varepsilon d)$ ,

$$(3.7) \quad \mathbb{P}(|F(L, A + \varepsilon\Xi, R)| > X) \leq 4 \max_{\ell \in \{j-1, j, k-1, k\}} \mathbb{P}\left(\log |\det(A_\ell'' + \varepsilon d \Xi_\ell'')| < -\left(\frac{X}{4} - d \log(1 + \varepsilon d)\right)\right).$$

The distribution of  $\Xi'_\ell$  is absolutely continuous with respect to the uniform measure on  $\ell \times \ell$  matrices with entries in  $[-1, 1]$ , and has a bounded density, where the bound only depends on  $d$ . Indeed, the linear map  $T : \mathbb{R}^{d^2} \rightarrow \mathbb{R}^{d^2}$  given by  $T(\Xi) = O_2 \Xi O_3$  is an isometry in the Frobenius norm, so the density of  $\Xi'$  (thought of as a vector in  $[-1, 1]^{d^2} \subset \mathbb{R}^{d^2}$ ) is uniform on its image. Thus, by taking marginals, the density of the top-left  $\ell \times \ell$  submatrix of  $\Xi'$  is scaled by at most a factor of  $2^{d^2 - \ell^2}$  in each coordinate.

Now, from the proof of Lemma 3.5 of [20], there exists  $C > 0$  such that for all  $\ell \leq d$ , for every  $\ell \times \ell$  matrix  $B$  of  $\|\cdot\|$ -norm at most 1, and for all  $K > 0$ ,

$$\mathbb{P}(\{Z_\ell : \log |\det(B + \varepsilon Z_\ell)| < -K\}) \leq \min\{1, C e^{-K/d} \varepsilon^{-1}\},$$

where the random variable  $Z_\ell$  is an  $\ell \times \ell$  matrix random variable with independent entries uniformly distributed in  $[-1, 1]$ .

Substituting the terms of (3.7) into this, replacing  $\varepsilon$  with  $\varepsilon d$  and absorbing the factors 4,  $d$  and the upper bound for the density into  $C$ , we obtain

$$\mathbb{P}(|F(L, A + \varepsilon \Xi, R)| > X) \leq C e^{-(X/4 - d \log(1 + \varepsilon d))/d} \varepsilon^{-1}.$$

which is of the required form.  $\square$

**Corollary 3.5.** *In the setting of Lemma 3.4, the random variables  $|F(L, A + \varepsilon \Xi, R)|$  are uniformly stochastically dominated by the random variable  $4d \log(\frac{\zeta}{\varepsilon}) + \text{Exp}((4d)^{-1})$  which has exponentially decaying tails.*

*In particular, there is a constant  $K$  (depending only on  $d$ ) such that for any  $0 < \varepsilon < 1$ , any invertible  $L$  and  $R$ , and any  $A$  of norm at most 1,  $\mathbb{E}(F(L, A + \varepsilon \Xi, R)) \geq 4d \log \varepsilon - K$ .*

*Proof.* Observe that if  $Y \sim 4d \log(\frac{\zeta}{\varepsilon}) + \text{Exp}((4d)^{-1})$ , then for any  $X > 0$ ,  $\mathbb{P}(Y > X) = \min(1, \exp[-(4d)^{-1}(X - 4d \log(\frac{\zeta}{\varepsilon}))]) = \min(1, \zeta e^{-\frac{X}{4d}} \varepsilon^{-1})$ .

Hence for any  $L, R$  and  $A$  as in the statement of Lemma 3.4, and  $\Xi$  a matrix random variable with independent  $U[-1, 1]$  entries,

$$\mathbb{E}F(L, A + \varepsilon \Xi, R) \geq -|\mathbb{E}F(L, A + \varepsilon \Xi, R)| \geq -4d \log \frac{\zeta}{\varepsilon} - 4d = 4d \log \varepsilon - K,$$

where  $K = -4d(\log \zeta + 1)$ .  $\square$

In the sequel we will make use of the following theorem.

**Theorem 3.6** (Theorem 2.19 of [28]). *Let  $Z_n$  be a sequence of random variables and  $\mathcal{F}_n$  an increasing sequence of  $\sigma$ -fields such that  $Z_n$  is measurable with respect to  $\mathcal{F}_n$  for each  $n \geq 1$ . Suppose that there exists a random variable  $Z$  with  $\mathbb{E}(|Z| \log^+ |Z|) < \infty$  and a constant  $c > 0$  such that for each  $X > 0$  and each  $n \geq 1$*

$$(3.8) \quad \mathbb{P}(|Z_n| > X) \leq c \mathbb{P}(|Z| > X).$$

*Then the Strong Law of Large Numbers holds for  $Z_n$ , i.e.*

$$\frac{1}{n} \sum_{k=1}^n (Z_k - \mathbb{E}(Z_k | \mathcal{F}_{k-1})) \xrightarrow{a.s.} 0$$

as  $n \rightarrow \infty$ .

**3.4. Constancy of limits.** We now show that the quantities appearing in Theorems **A-E** are almost surely constant.

**Lemma 3.7.** *Let  $d \geq 2$ , let  $(A_n)$  be an arbitrary sequence of  $d \times d$  matrices of norm at most 1 and let  $\varepsilon > 0$ . Then for any  $1 \leq j < k \leq d$ , the quantity  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)}))$  is almost surely constant.*

*Proof.* The proof is an application of the Kolmogorov 0–1 law. Write  $A_{i,i',\varepsilon}$  for the product  $A_{i',\varepsilon} \cdots A_{i+1,\varepsilon}$ . We rely on the following well-known facts about singular values: firstly  $s_1(A)s_d(B) \leq s_1(AB) \leq s_1(A)s_1(B)$  for any matrices  $A$  and  $B$ ; and secondly  $s_1(A^{\wedge l}) = S_1^l(A)$  and  $s_D(A^{\wedge l}) = S_{d-l+1}^l(A)$  where  $A^{\wedge l}$  is the  $l$ -fold exterior power of  $A$  acting on the  $l$ -fold exterior power of  $\mathbb{R}^d$  and  $D = \binom{d}{l}$  is the dimension of the  $l$ -fold exterior power of  $\mathbb{R}^d$ .

Let  $m$  be an arbitrary fixed natural number. We have  $A_{0,m,\varepsilon}$  is almost surely non-singular. Let  $c$  and  $C$  be random variables depending only on  $\Xi_1, \dots, \Xi_m$  defined by  $c = \min_{i,j} S_i^j(A_{0,m,\varepsilon})$  and  $C = \max_{i,j} S_i^j(A_{0,m,\varepsilon})$  so that  $0 < c < C < \infty$ . For  $n > m$ , and any  $1 \leq l \leq d$ , applying the above inequalities to  $(A_{0,n,\varepsilon})^{\wedge l} = (A_{m,n,\varepsilon})^{\wedge l} (A_{0,m,\varepsilon})^{\wedge l}$ , we obtain

$$S_1^l(A_{m,n,\varepsilon}) S_{d-l+1}^l(A_{0,m,\varepsilon}) \leq S_1^l(A_{0,n,\varepsilon}) \leq S_1^l(A_{m,n,\varepsilon}) S_1^l(A_{0,m,\varepsilon}),$$

so that for each  $l$ , we have

$$c S_1^l(A_{m,n,\varepsilon}) \leq S_1^l(A_\varepsilon^{(n)}) \leq C S_1^l(A_{m,n,\varepsilon})$$

for all  $n$ . Dividing the  $l$ th system of inequalities by the  $(l+1)$ st, we obtain for each  $l$ ,

$$\frac{c}{C} s_l(A_{m,n,\varepsilon}) \leq s_l(A_\varepsilon^{(n)}) \leq \frac{C}{c} s_l(A_{m,n,\varepsilon}),$$

and hence, dividing the  $j$ th system of these inequalities by the  $k$ th, we obtain

$$\frac{c^2}{C^2} \frac{s_j(A_{m,n,\varepsilon})}{s_k(A_{m,n,\varepsilon})} \leq \frac{s_j(A_\varepsilon^{(n)})}{s_k(A_\varepsilon^{(n)})} \leq \frac{C^2}{c^2} \frac{s_j(A_{m,n,\varepsilon})}{s_k(A_{m,n,\varepsilon})}.$$

Taking logs, and dividing by  $n$ , we see

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_k(A_\varepsilon^{(n)})} = \liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_{m,n,\varepsilon})}{s_k(A_{m,n,\varepsilon})}.$$

Letting  $\mathcal{F}_m$  be the smallest  $\sigma$ -algebra with respect to which the sequence of perturbations  $\Xi_{m+1}, \Xi_{m+2}, \dots$  is measurable, we deduce that  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)}))$  is  $\mathcal{F}_m$ -measurable. Since  $m$  was arbitrary, we deduce that  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)}))$  is tail-measurable. Since the perturbations are independent, it follows from the Kolmogorov 0–1 law that  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)}))$  is almost surely constant as required.  $\square$

**3.5. Bookkeeping.** The main result in this section is a meta-theorem, showing that if every block of matrices of a fixed length has a nearby “target” block with a large gap between its  $j$ th and  $k$ th singular values, then there is a uniform lower bound on the gap between the  $j$ th and  $k$ th Lyapunov exponents of the cocycle  $A_\varepsilon^{(n)}$  depending only on  $\varepsilon$ .

Somewhat more precisely, the hypothesis is that for every  $\varepsilon > 0$ , there exists an  $N$  such that for every block  $M_1, \dots, M_{N-1}$  of  $d \times d$  matrices, there exists a *target*  $M_m^\circ, \dots, M_{m'}^\circ$  with  $1 \leq m \leq m' \leq N-1$  consisting of nearby matrices, for which there is a *gap* between the  $j$ th and  $k$ th singular values (that is  $s_j(M_{m'}^\circ \cdots M_m^\circ)/s_k(M_{m'}^\circ \cdots M_m^\circ)$  exceeding a threshold determined based on the Gluing Lemma above). The conclusion is that there is an explicit lower bound  $c(\varepsilon) > 0$ , such that for all sequences  $(A_n)_{n \in \mathbb{N}}$  of  $d \times d$  matrices of norm 1,  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)})) \geq c(\varepsilon)$ .

If  $0 < \varepsilon < 1$  is fixed, we say that  $\tilde{M}$  is a *near-identity perturbation* of  $M$  if  $\tilde{M} = RM$  or  $MR$  for some matrix  $R$  satisfying  $\|R - I\| < \frac{\varepsilon}{4}$  and  $\|R^{-1} - I\| < \frac{\varepsilon}{4}$ .

**Theorem 3.8.** *Let  $d > 1$  be fixed and let  $1 \leq j < k \leq d$ . Suppose that for all  $0 < \varepsilon < 1$  and all  $\mathcal{G} > 2$ , there exists an  $N \in \mathbb{N}$  such that for every sequence  $M_1, \dots, M_{N-1}$  of invertible  $d \times d$  matrices there exists a target: numbers  $1 \leq m \leq m' \leq N-1$  and near-identity perturbations  $M_m^\circ, \dots, M_{m'}^\circ$  of  $M_m, \dots, M_{m'}$  such that*

$$\frac{s_j(M_{m'}^\circ \cdots M_m^\circ)}{s_k(M_{m'}^\circ \cdots M_m^\circ)} > \mathcal{G}.$$

*Then there exists  $c(\varepsilon) > 0$  such that for any sequence  $(A_n)_{n \in \mathbb{N}}$  of  $d \times d$  matrices of norm at most 1, the almost sure constant  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log (s_j(A_\varepsilon^{(n)})/s_k(A_\varepsilon^{(n)}))$  satisfies*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_k(A_\varepsilon^{(n)})} > c(\varepsilon).$$

*Furthermore, the asymptotic behaviour of the lower bound  $c(\varepsilon)$  can be explicitly computed based on the expectations of the random variables appearing in Section 3.3.*

*Proof.* Let  $d, j$  and  $k$  be as in the statement of the theorem and let  $0 < \varepsilon < 1$ . We recall the definition

$$(3.9) \quad F(L, A, R) = G(LAR) - G(L) - G(R),$$

where  $G(M) = \log q_{j,k}(M)$ .

Let  $\lambda = \inf_{L,A,R} \mathbb{E}_\Xi F(L, A + \varepsilon \Xi, R)$  where the random variable  $\Xi$  runs over matrices with independent  $U[-1, 1]$ -valued entries, and the infimum is taken over all non-singular matrices  $L$  and  $R$  and all matrices  $A$  of norm at most 1. By Corollary 3.5,  $\lambda \geq 4d \log \varepsilon - K$ , where the constant  $K$  depends only on  $d$ . We then apply the hypothesis of the theorem with  $\mathcal{G}$  taken to be

$2e^{2+2|\lambda|}$  to obtain an  $N$  such that for every  $(N-1)$ -block,  $M_1, \dots, M_{N-1}$  there is a target  $M_m^\odot, \dots, M_{m'}^\odot$ .

Starting with a sequence of matrices  $A_1, A_2, \dots$ , each of norm at most 1, we initially make perturbations of size at most  $\frac{\varepsilon}{2}$  as in Lemma 3.1 to obtain matrices  $A'_1, A'_2, \dots$  each of norm at most  $1 - \frac{\varepsilon}{2}$  and with smallest singular value at least  $\frac{\varepsilon}{2}$ .

For each  $(N-1)$ -block  $A'_{lN+1}, \dots, A'_{lN+N-1}$  where  $l$  ranges over  $\mathbb{N}_0$ , we next apply the hypothesis of the theorem to produce a target block  $A_{lN+m_l}^\odot, \dots, A_{lN+m'_l}^\odot$  with  $1 \leq m_l < m'_l \leq N-1$  where  $A_i^\odot$  is a near-identity perturbation of  $A'_i$  for each  $i \in [lN+m_l, lN+m'_l]$  such that the product has at least the required gap of  $2e^{2+2|\lambda|}$  between the  $j$ th and  $k$ th singular values. Note that if  $A_i^\odot = R_i A'_i$ , then  $\|A_i^\odot - A'_i\| = \|(R_i - I)A'_i\| < \frac{\varepsilon}{4}$  (and similarly if  $A_i^\odot = A'_i R_i$ ), which ensures that  $\|A_i^\odot - A_i\| < \frac{3\varepsilon}{4}$ ,  $\|A_i^\odot\| < 1 - \frac{\varepsilon}{4}$  and  $s_d(A_i^\odot) > \frac{\varepsilon}{4}$ .

As established in Lemma 3.2 if the random sub-block  $A_{lN+m_l, \varepsilon}, \dots, A_{lN+m'_l, \varepsilon}$  falls in the *target area*, that is, if  $|A_{i, \varepsilon} - A_i^\odot|_\infty \leq (\frac{\varepsilon}{4})^N / (3dN)$  for each  $lN+m_l \leq i \leq lN+m'_l$ , then the ratio of the  $j$ th and  $k$ th singular values of  $A_{lN+m'_l, \varepsilon} \cdots A_{lN+m_l, \varepsilon}$  exceeds  $e^{2+2|\lambda|}$ . Since  $|A_i^\odot - A_i|_\infty \leq \|A_i^\odot - A_i\| < \frac{3\varepsilon}{4}$ , the  $(\frac{\varepsilon}{4})^N / (3dN)$ -ball in the  $|\cdot|_\infty$ -norm around  $A_i^\odot$  lies within the  $\varepsilon$ -ball around  $A_i$ . That is, the target area is contained in the support of  $A_{i, \varepsilon}$ . Hence each target is independently hit with probability at least  $p = ((\frac{\varepsilon}{4})^N / (3dN))^{d^2 N} = (\frac{\varepsilon}{4})^{d^2 N^2} / (3dN)^{d^2 N}$ . The quantity  $c(\varepsilon)$  will end up being  $p/N$ .

Note that the  $m_l$  and  $m'_l$  are determined solely by the unperturbed sequence of matrices (that is, they are deterministic). We let  $\mathcal{B}_0$  denote the  $\sigma$ -algebra generated by  $\{\Xi_i : i \in \bigcup_l [lN+m_l, lN+m'_l]\}$ . That is  $\mathcal{B}_0$  is the smallest  $\sigma$ -algebra with respect to which the matrix random variables  $\Xi_i$  in the target sub-blocks are measurable. We then define a  $\mathcal{B}_0$ -measurable random subset of  $\mathbb{N}$  by

$$\text{TargHit} = \bigcup_{l \in \mathbb{N}_0} \{[lN+m_l, lN+m'_l] : |A_{i, \varepsilon} - A_i^\odot|_\infty < (\frac{\varepsilon}{4})^N / (3dN) \text{ for all } i \in [lN+m_l, lN+m'_l]\}.$$

That is, **TargHit** consists of the coordinates of those matrices in blocks where the target is successfully hit. Note that since  $1 \leq m_l < m'_l < N$  for each  $l$ , there is a gap of length at least 1 between successive targets that are hit. We now define a second subset called **Trans** (for transition coordinates) consisting of those  $i \in \mathbb{N}$  immediately preceding or following a block in **TargHit**. This is again  $\mathcal{B}_0$ -measurable. We enumerate the elements of **Trans** as  $T_1 < T_2 < T_3 < \dots$ , so that the coordinates  $(T_n)$  are  $\mathcal{B}_0$ -measurable random variables. Lemma 3.7 implies that the limit in the statement is almost surely constant.

Further, we assume without loss of generality (using Lemma 3.7 again) that  $T_1 > 1$  and define  $T_0 = 0$ .

At this point we introduce a second  $\sigma$ -algebra:  $\mathcal{F}_0$  which is generated by  $\{\Xi_i : i \notin \text{Trans}\}$ . This is a refinement of  $\mathcal{B}_0$  since all of the transition coordinates lie outside the target sub-blocks (each transition coordinate immediately precedes or follows a target sub-block). We claim that conditioned on  $\mathcal{F}_0$ , the random variables  $\Xi_{T_i}$  are independent and identically distributed with entries uniformly distributed over  $[-1, 1]$ . The reason for this is that the  $T_i$  are  $\mathcal{B}_0$ -measurable (hence  $\mathcal{F}_0$ -measurable) and the  $\Xi_{T_i}$  are independent of all of the  $\Xi_m$ 's that generate  $\mathcal{F}_0$ . (Informally, we could say that we never ‘‘looked at’’ the  $\Xi_{T_i}$ 's when identifying the transition coordinates, so that the  $\Xi$  values at these times are independent of everything that we have looked at).

We are now ready to define a sequence of  $\sigma$ -algebras that will be used when we apply Theorem 3.6: Let  $\mathcal{F}_n = \mathcal{F}_0 \vee \sigma(\Xi_{T_1}, \dots, \Xi_{T_n})$ . That is,  $\mathcal{F}_n$  is the smallest  $\sigma$ -algebra such that all non-transition  $\Xi$  variables as well as the first  $n$  transition  $\Xi$  variables are measurable. By the observation above,  $\Xi_{T_n}$  is independent of  $\mathcal{F}_{n-1}$ .

We give names to blocks appearing in  $A_\varepsilon^{(n)}$  by defining  $B_{i,\varepsilon} = A_{T_i-1,\varepsilon} \cdots A_{T_i-1+1,\varepsilon}$ . (In the unusual case where  $T_i = T_{i-1} + 1$ ,  $B_{i,\varepsilon}$  is just the identity). With this notation,

$$A_\varepsilon^{(n)} = R_{n,\varepsilon} A_{T_{L_n},\varepsilon} B_{L_n,\varepsilon} A_{T_{L_n-1},\varepsilon} \cdots A_{T_2,\varepsilon} B_{2,\varepsilon} A_{T_1,\varepsilon} B_{1,\varepsilon},$$

where  $L_n = \#(\text{Trans} \cap \{1, \dots, n\})$  is an  $\mathcal{F}_0$ -measurable random variable and  $R_{n,\varepsilon}$  is the tail  $A_{n,\varepsilon} \cdots A_{T_{L_n}+1,\varepsilon}$ . Using (3.9), we see  $G(BAC) = G(B) + G(C) + F(B, A, C)$ . Applying it twice, first separating out  $B$  and  $A_1$ , we obtain  $G(BA_1CA_2D) = G(B) + G(C) + G(D) + F(C, A_2, D) + F(B, A_1, CA_2D)$ . Using induction, we have

$$(3.10) \quad G(A_\varepsilon^{(n)}) = G(B_{1,\varepsilon}) + \dots + G(B_{L_n,\varepsilon}) + G(R_{n,\varepsilon}) + \sum_{i=1}^{L_n-1} F(B_{i+1,\varepsilon}, A_{T_i,\varepsilon}, \bar{B}_{i,\varepsilon}) + F(R_{n,\varepsilon}, A_{T_{L_n},\varepsilon}, \bar{B}_{L_n,\varepsilon}),$$

where  $\bar{B}_{i,\varepsilon} = B_{i,\varepsilon} A_{T_{i-1},\varepsilon} \cdots A_{T_1,\varepsilon} B_{1,\varepsilon}$ . Since each transition coordinate is immediately followed or preceded by a target that was hit, at least half of the  $B_{i,\varepsilon}$  consist of blocks where the target area was hit. By Lemma 3.2,  $G(B_{i,\varepsilon}) \geq 2 + 2|\lambda|$  for those blocks. For the other blocks,  $G(B_{i,\varepsilon}) \geq 0$  by definition. Hence we see

$$(3.11) \quad \liminf_{n \rightarrow \infty} \frac{1}{L_n} \sum_{i=1}^{L_n} G(B_{i,\varepsilon}) \geq 1 + |\lambda| \text{ a.s.}$$

We let  $Z_i = F(B_{i+1,\varepsilon}, A_{T_i,\varepsilon}, \bar{B}_{i,\varepsilon})$  and note that it is  $\mathcal{F}_i$ -measurable. However since  $\Xi_{T_i}$  is independent of  $\mathcal{F}_{i-1}$ ,

$$\mathbb{E}(F(B_{i+1,\varepsilon}, A_{T_i} + \varepsilon \Xi_{T_i}, \bar{B}_{i,\varepsilon}) | \mathcal{F}_{i-1}) = \mathbb{E}_\Xi F(B_{i+1,\varepsilon}, A_{T_i} + \varepsilon \Xi, \bar{B}_{i,\varepsilon}).$$

That is,  $\mathbb{E}(Z_i | \mathcal{F}_{i-1})$  is exactly the expectation of a random variable as in Lemma 3.4, so that by definition,  $\mathbb{E}(Z_i | \mathcal{F}_{i-1}) \geq \lambda$ .

Hence by the version of the Strong Law of Large Numbers, Theorem 3.6,

$$(3.12) \quad \liminf_{n \rightarrow \infty} \frac{1}{L_n} \sum_{i=1}^{L_n-1} F(B_{i+1,\varepsilon}, A_{T_i,\varepsilon}, \bar{B}_{i,\varepsilon}) \geq \lambda \text{ a.s.}$$

Since each successive target is independently hit with probability at least  $p$ , we may apply Theorem 3.6 one more time (with  $\sigma$ -algebras  $\mathcal{F}'_n = \sigma(\{\Xi_i : i \in \bigcup_{l=0}^{n-1} \{lN+1, \dots, lN+N-1\}\})$ ) to obtain

$$(3.13) \quad \liminf_{n \rightarrow \infty} \frac{L_n}{n} \geq \frac{p}{N} \text{ a.s.}$$

Adding (3.11) and (3.12); and then multiplying by (3.13), we obtain

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \left( \sum_{i=1}^{L_n} G(B_{i,\varepsilon}) + \sum_{i=1}^{L_n-1} F(B_{i+1,\varepsilon}, A_{T_i,\varepsilon}, \bar{B}_{i,\varepsilon}) \right) \geq \frac{p}{N} \text{ a.s.}$$

Comparing this with (3.10), it remains to control  $G(R_{n,\varepsilon})$  and  $F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon})$ .

By definition,  $G(R_{n,\varepsilon})$  is non-negative. Finally we will show that  $\frac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon}) \rightarrow 0$  a.s. Using (3.10) and the inequalities above, this is sufficient to lead to the conclusion

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_k(A_\varepsilon^{(n)})} \geq \frac{p}{N} \text{ a.s.}$$

To show that  $\frac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon}) \rightarrow 0$  a.s., let  $\eta > 0$  be arbitrary. Then

$$\begin{aligned} \mathbb{P}(|\tfrac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon})| > \eta) &= \mathbb{E}(\mathbb{P}(|\tfrac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon})| > \eta | \mathcal{F}_{L_n-1})) \\ &= \mathbb{E}(\mathbb{P}(|F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon})| > n\eta | \mathcal{F}_{L_n-1})) \\ &= \mathbb{E}(\mathbb{P}_\Xi(|F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}} + \varepsilon\Xi, \bar{B}_{L_n,\varepsilon})| > n\eta | \mathcal{F}_{L_n-1})). \end{aligned}$$

By the bounds in Lemma 3.4, the above quantities are summable, so that by the first Borel-Cantelli lemma, almost surely  $|\frac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon})| \leq \eta$  for all but finitely many  $n$ . Since  $\eta$  is an arbitrary positive number, we see  $\frac{1}{n}F(R_{n,\varepsilon}, A_{T_{L_n,\varepsilon}}, \bar{B}_{L_n,\varepsilon}) \rightarrow 0$  a.s. as required.  $\square$

We also state an alternative formulation of Theorem 3.8 that will be useful in Section 5.

**Corollary 3.9.** *Let  $d > 1$  be fixed and let  $1 \leq j < k \leq d$ . Let  $0 < \varepsilon < 1$  and suppose that for all  $\mathcal{G} > 2$ , there exists an  $N \in \mathbb{N}$  such that for every sequence  $M_1, \dots, M_{N-1}$  of  $d \times d$  matrices of norm at most  $1 - \frac{\varepsilon}{2}$  and  $d$ th singular value at least  $\frac{\varepsilon}{2}$ , there exists a target: numbers  $1 \leq m \leq m' \leq N-1$  and perturbations  $M_m^\circ, \dots, M_{m'}^\circ$  of  $M_m, \dots, M_{m'}$  such that  $|M_i^\circ - M_i|_\infty \leq \frac{\varepsilon}{4d}$  for each  $m \leq i \leq m'$  and*

$$\frac{s_j(M_{m'}^\circ \cdots M_m^\circ)}{s_k(M_{m'}^\circ \cdots M_m^\circ)} > \mathcal{G}.$$

*Then  $c(\varepsilon) > 0$ , where  $c(\varepsilon)$  is as in the statement of Theorem 3.8.*



The proof is almost exactly the same as the proof of Theorem 3.8. Unlike in that theorem, there is no need to apply Lemma 3.1. The new hypothesis ensures that  $\|M_i^\circ - M_i\| \leq \frac{\varepsilon}{4}$  for each  $m \leq i \leq m'$  which is exactly the conclusion that was obtained from the “near identity perturbation” hypothesis in the proof of Theorem 3.8. From that point in the proof onwards, that is the only fact about the  $M_i^\circ$  that is used, so that the remainder of the proof of Theorem 3.8 applies directly.

#### 4. CONCRETE TARGETS

This section identifies suitable targets for our general strategy put forth in §3. First, in §4.1, we show that if we can build a  $(j, j+1)$ -gap then we can build a  $(d-j, d-j+1)$ -gap. In §4.2, we give a general argument to identify targets in the case where singular vectors remain *spread* as they evolve within a block. Finally, §4.3 and §4.4 complete the target identification step for the case of  $2 \times 2$  – or more generally, extremal singular values– and  $3 \times 3$  matrices, respectively.

Let  $B = M_{N-1} \cdots M_1$  be a product of  $N - 1$  invertible matrices. Let the singular vectors of  $B$  be  $v_1, \dots, v_d$ , listed as usual in order of decreasing singular values. Let  $E_j(B)$  denote the span of  $(v_i)_{i \leq j}$  and  $F_j(B)$  denote the span of  $(v_i)_{i > j}$ . For  $1 \leq j < d$  and for each  $1 \leq n < N$ , we let  $E_{j,n}$  denote the linear span of  $(M_n \dots M_1 v_i)_{i \leq j}$  and let  $F_{j,n}$  denote the span of  $(M_n \dots M_1 v_i)_{i > j}$ . We write  $v_i^n$  for the unit normalization of  $M_n \dots M_1 v_i$ . Let  $\delta_{n,j}$  denote the minimal distance between norm-1 vectors in  $E_{j,n}$  and  $F_{j,n}$  and let  $\delta_n = \min_{1 \leq j < d} \delta_{n,j}$ . We will say the block  $M_{N-1}, \dots, M_1$  is  $\eta$ -*spread* if  $\delta_n > \eta$  for each  $1 \leq n < N$  and *nearly aligned* otherwise.

**4.1. Complementarity.** Recall that for given a matrix  $M$ , a fixed  $\varepsilon > 0$ , and a matrix  $R$  satisfying  $\|R^{\pm 1} - I\| < \frac{\varepsilon}{4}$ , we call  $RM$  or  $MR$  a *near-identity perturbation*.

Let  $d \in \mathbb{N}$  and  $\varepsilon > 0$  be fixed. We assume there is a quantity  $\mathcal{G}$ , depending on  $d$  and  $\varepsilon$ , called the *required gap size*. We say that we can *build a gap between the  $j$ th and  $(j+1)$ st singular values* if there is an  $N$  such that for every block  $M_1, \dots, M_{N-1}$  of invertible  $d \times d$  matrices, there exists a sub-block  $M_m, \dots, M_{m'}$  with  $1 \leq m \leq m' \leq N - 1$  and near-identity targets  $M_m^\circ, \dots, M_{m'}^\circ$  of the sub-block such that  $s_j(M_m^\circ \cdots M_{m'}^\circ) / s_{j+1}(M_m^\circ \cdots M_{m'}^\circ) > \mathcal{G}$ .

**Lemma 4.1.** *Let  $d \in \mathbb{N}$  and  $\varepsilon > 0$ . If we can build a gap between the  $j$ th and  $(j+1)$ st singular values, then we can build a gap between the  $(d-j)$ th and  $(d-j+1)$ st singular values.*

*Proof.* Let  $M_1, \dots, M_{N-1}$  be a block of invertible matrices and set  $\tilde{M}_n = M_{N-n}^{-1}$  so that we obtain a new block  $\tilde{M}_1, \dots, \tilde{M}_{N-1}$  of invertible matrices. By assumption, there exist a sub-block and near-identity perturbations



$\tilde{M}_m^\circ, \dots, \tilde{M}_{m'}^\circ$  such that

$$\frac{s_j(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)}{s_{j+1}(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)} > \mathcal{G}.$$

Now let  $M_l^\circ = (\tilde{M}_{N-l}^\circ)^{-1}$  for  $l = N - m', \dots, N - m$ . These are near-identity perturbations of the matrices  $M_{N-m'}, \dots, M_{N-m}$ . Since  $M_{N-m}^\circ \cdots M_{N-m'}^\circ = (\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)^{-1}$ , we see  $s_{d-j}(M_{N-m}^\circ \cdots M_{N-m'}^\circ) = s_{j+1}(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)^{-1}$  and  $s_{d-j+1}(M_{N-m}^\circ \cdots M_{N-m'}^\circ) = s_j(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)^{-1}$ . This ensures that

$$\frac{s_{d-j}(M_{N-m}^\circ \cdots M_{N-m'}^\circ)}{s_{d-j+1}(M_{N-m}^\circ \cdots M_{N-m'}^\circ)} = \frac{s_j(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)}{s_{j+1}(\tilde{M}_{m'}^\circ \cdots \tilde{M}_m^\circ)} > \mathcal{G}$$

as required.  $\square$

#### 4.2. Target when singular vectors remain spread.

**Lemma 4.2** (Target for spread block). *Let  $1 \leq j < d$ ,  $\mathcal{G} > 1$ ,  $0 < \eta < 1$  and let  $0 < \varepsilon < 1$ . Consider a product of  $N - 1 \geq \frac{16 \log \mathcal{G}}{\varepsilon \eta}$  invertible matrices,  $M_{N-1} \cdots M_1$  (a “block”). If the block is  $\eta$ -spread, then there exists a product  $M_{N-1}^\circ, \dots, M_1^\circ$  (a “target block”) with the properties:*

- (1)  $M_n^\circ = R_n M_n$  with  $\|R_n^{\pm 1} - I\| \leq \frac{\varepsilon}{4}$  for each  $1 \leq n < N$ ;
- (2)

$$\frac{s_j(M_{N-1}^\circ \cdots M_1^\circ)}{s_{j+1}(M_{N-1}^\circ \cdots M_1^\circ)} \geq \mathcal{G}.$$

*Proof.* We first describe the simple idea: after applying each  $M_n$ , we expand the images of the leading  $j$  singular vectors by a factor of  $1 + \frac{\varepsilon \eta}{8}$ . The spread condition implies that the fast and slow subspaces are far enough apart that the perturbation to each matrix is of size at most  $\frac{\varepsilon}{4}$ . In  $N - 1$  steps, the cumulative effect of this is to increase the ratio between the  $j$ th and  $(j + 1)$ st singular values by a factor of  $\mathcal{G}$ , by the choice of  $N$ .

Let the spaces  $E_n$  and  $F_n$  be as above so that  $\mathbb{R}^d = E_n \oplus F_n$ . By Lemma 2.1, we have  $\|\Pi_{E_n|F_n}\| \leq 2/\delta_n < 2/\eta$ .

We then set  $R_n = I + \frac{\varepsilon \eta}{8} \Pi_{E_n|F_n}$ . The calculation above shows that  $\|R_n - I\| = \|\frac{\varepsilon \eta}{8} \Pi_{E_n|F_n}\| \leq \frac{\varepsilon \eta}{8} (2/\eta) \leq \frac{\varepsilon}{4}$ . Similarly  $R_n^{-1} = I - (\frac{\varepsilon \eta}{8} / (1 + \frac{\varepsilon \eta}{8})) \Pi_{E_n|F_n}$  so that  $\|R_n^{-1} - I\| \leq \frac{\varepsilon}{4}$  also. Since by definition,  $M_n(E_{n-1}) = E_n$  and  $M_n(F_{n-1}) = F_n$ , we have

$$\begin{aligned} M_n \Pi_{E_{n-1}|F_{n-1}} &= \Pi_{E_n|F_n} M_n \quad \text{and} \\ M_n \Pi_{F_{n-1}|E_{n-1}} &= \Pi_{F_n|E_n} M_n, \end{aligned}$$

so that

$$\begin{aligned} M_n \left( I + \frac{\varepsilon \eta}{8} \Pi_{E_{n-1}|F_{n-1}} \right) &= M_n \left( \Pi_{F_{n-1}|E_{n-1}} + \left( 1 + \frac{\varepsilon \eta}{8} \right) \Pi_{E_{n-1}|F_{n-1}} \right) \\ &= \left( \Pi_{F_n|E_n} + \left( 1 + \frac{\varepsilon \eta}{8} \right) \Pi_{E_n|F_n} \right) M_n. \end{aligned}$$

Using this inductively, along with the fact that  $\Pi_{E_n \| F_n}$  and  $\Pi_{F_n \| E_n}$  are complementary projections we see that

$$M_{N-1}^{\odot} \cdots M_1^{\odot} = (\Pi_{F_{N-1} \| E_{N-1}} + (1 + \frac{\varepsilon\eta}{8})^{N-1} \Pi_{E_{N-1} \| F_{N-1}}) M_{N-1} \cdots M_1.$$

Hence

$$s_i(M_{N-1}^{\odot} \cdots M_1^{\odot}) = \begin{cases} (1 + \frac{\varepsilon\eta}{8})^{N-1} s_i(M_{N-1} \cdots M_1) & \text{for } i \leq j; \\ s_i(M_{N-1} \cdots M_1) & \text{for } i > j. \end{cases}$$

In particular,

$$\frac{s_j(M_{N-1}^{\odot} \cdots M_1^{\odot})}{s_{j+1}(M_{N-1}^{\odot} \cdots M_1^{\odot})} = \left(1 + \frac{\varepsilon\eta}{8}\right)^{N-1} \frac{s_j(M_{N-1} \cdots M_1)}{s_{j+1}(M_{N-1} \cdots M_1)} \geq \left(1 + \frac{\varepsilon\eta}{8}\right)^{N-1}.$$

Since  $0 < \varepsilon < 1$  and  $0 < \eta < 1$ , we have  $1 + \frac{\varepsilon\eta}{8} > \exp(\frac{\varepsilon\eta}{16})$ . Hence the choice of  $N$  ensures that  $\frac{s_j(M_{N-1}^{\odot} \cdots M_1^{\odot})}{s_{j+1}(M_{N-1}^{\odot} \cdots M_1^{\odot})} \geq \mathcal{G}$ , as required.  $\square$

**4.3. Targets for  $d = 2$  and extremal singular value gaps.** This section describes how to identify targets for blocks of  $2 \times 2$  matrices. We start by presenting an auxiliary result about extremal singular value gaps, valid for arbitrary dimension. The main result of this section is Lemma 4.4 which, specialised to  $d = 2$ , yields targets for  $2 \times 2$  matrices.

**Lemma 4.3.** *Let  $B$  be a non-singular  $d \times d$  matrix and suppose there exist orthogonal vectors  $u$  and  $v$  such that  $\angle(Bu, Bv) \leq \eta$ . Then  $s_1(B)/s_d(B) \geq 1/\eta$ .*

*The same conclusion holds if there exist  $u$  and  $v$  with  $\angle(u, v) \leq \eta$  but  $Bu$  and  $Bv$  are orthogonal.*

*Proof.* We assume without loss of generality that  $\|Bu\| \geq \|Bv\|$ . We then have  $\|B^{\wedge 2}(u \wedge v)\| = \|Bu \wedge Bv\| \leq \|Bu\| \|Bv\| \angle(Bu, Bv) \leq \|Bu\|^2 \eta$ . Letting  $L$  be the restriction of  $B$  to  $\text{lin}(u, v)$  and  $L^{\wedge 2}$  be its exterior square, it follows that  $\|L^{\wedge 2}\| \leq \eta s_1(L)^2$ . Since  $\|L^{\wedge 2}\| = s_1(L)s_2(L)$ , it follows that  $s_2(L) \leq \eta s_1(L)$ . Finally  $s_1(B) \geq s_1(L)$  and  $s_d(B) \leq s_2(L)$  giving the required conclusion. The second statement follows from taking inverses of  $B$ .  $\square$

**Lemma 4.4** (Target for  $d \times d$  matrices). *Let  $0 < \varepsilon < 1$  and  $\mathcal{G} > 1$  be given. Consider a block of  $N - 1 \geq \frac{16}{\varepsilon} \mathcal{G} \log \mathcal{G}$   $d \times d$  invertible matrices,  $M_1, \dots, M_{N-1}$ . Then there exist  $1 \leq m \leq m' < N$  and a sequence  $M_m^{\odot}, \dots, M_{m'}^{\odot}$  of near-identity perturbations of  $M_m, \dots, M_{m'}$  with the property*

$$\frac{s_1(M_{m'}^{\odot} \cdots M_m^{\odot})}{s_d(M_{m'}^{\odot} \cdots M_m^{\odot})} \geq \mathcal{G}.$$

*Proof.* We consider two cases:

- (1) The block  $M_{N-1}, \dots, M_1$  is  $1/\mathcal{G}$ -spread. That is,  $\delta_n > 1/\mathcal{G}$  for each  $1 \leq n < N$ .

- (2) The block  $M_{N-1}, \dots, M_1$  is nearly aligned. That is, there exists  $1 \leq n < N$  such that  $\delta_n \leq 1/\mathcal{G}$ .

In the first case, applying Lemma 4.2, with  $j = 1$  and  $\eta = 1/\mathcal{G}$ , identifies a target for  $M_{N-1}, \dots, M_1$ ; namely  $B^\circ := M_{N-1}^\circ \dots M_1^\circ$ , as constructed in Lemma 4.2, satisfies  $\frac{s_1(B^\circ)}{s_d(B^\circ)} \geq \frac{s_1(B^\circ)}{s_2(B^\circ)} \geq \mathcal{G}$ .

In the second case, the result follows from Lemma 4.3, with  $B^\circ := M_n \dots M_1$ .  $\square$

**Remark 4.5.** Note that Lemma 4.4, specialised to  $d = 2$ , yields targets for  $2 \times 2$  matrices.

**4.4. Targets for  $d = 3$ .** In this section, we describe how to identify targets for blocks of  $3 \times 3$  matrices. We start with two auxiliary lemmas. The main result of this section is Lemma 4.8. The reader should recall the definitions of  $v_j, v_j^n, E_j^n, F_j^n$  and  $\delta_{n,j}$  from the beginning of Section 4.

**Lemma 4.6.** *Let  $M_1, \dots, M_{N-1}$  be invertible  $3 \times 3$  matrices and let  $B = M_{N-1} \dots M_1$ . Suppose  $s_1(B)/s_2(B) < \mathcal{G}$ . Suppose further that for some  $1 \leq n < N$ ,  $\angle(v_1^n, v_2^n) < \mathcal{G}^{-5}$ . Then, setting  $B_1 = M_n \dots M_1$  and  $B_2 = M_{N-1} \dots M_{n+1}$ , either  $s_1(B_1)/s_2(B_1) \geq \mathcal{G}$  or  $s_1(B_2)/s_2(B_2) \geq \mathcal{G}$ .*

*Proof.* We use the (known) inequality  $s_1(B_2 B_1) \geq s_{1+k}(B_2) s_{d-k}(B_1)$ , valid in all dimensions. We briefly indicate the proof: by the max-min characterization of singular values, namely  $s_k(B) = \max_{\{W: \dim W = k\}} \min_{\{w \in W, \|w\|=1\}} \|Bw\|$ , there is a  $(d-k)$ -dimensional subspace  $U$  of  $\mathbb{R}^d$  such that  $\|B_1 u\| \geq s_{d-k}(B_1)$  for all  $u \in U \cap S$  and a  $(1+k)$ -dimensional subspace  $V$  of  $\mathbb{R}^d$  such that  $\|B_2 v\| \geq s_{1+k}(B_2)$  for all  $v \in V \cap S$ . Since  $\dim(B_1 U) + \dim V = \dim U + \dim V > d$ , the spaces  $B_1 U$  and  $V$  must intersect. That is, there exists  $u \in U \cap S$  such that  $B_1 u \in V$ . Hence  $\|Bu\| = \|B_2(B_1 u)\| \geq s_{1+k}(B_2) \|B_1 u\| \geq s_{1+k}(B_2) s_{d-k}(B_1)$  as required.

Specializing to the case  $d = 3$  and  $k = 1$  gives the inequality

$$(4.1) \quad s_1(B) \geq s_2(B_2) s_2(B_1),$$

valid for  $3 \times 3$  matrices. Combining this with the assumption that  $B$  does not have a (1,2) gap, we see  $s_2(B) \geq \mathcal{G}^{-1} s_2(B_2) s_2(B_1)$  so that  $S_1^2(B) \geq \mathcal{G}^{-1} s_2(B_2)^2 s_2(B_1)^2$ .

On the other hand, we have

$$\begin{aligned} S_1^2(B) &= s_1(B) s_2(B) \\ &= \|Bv_1 \wedge Bv_2\| \\ &\leq s_1(B_2)^2 \|B_1 v_1 \wedge B_1 v_2\| \\ &\leq s_1(B_2)^2 \|B_1 v_1\| \|B_1 v_2\| \mathcal{G}^{-5} \\ &\leq s_1(B_2)^2 s_1(B_1)^2 \mathcal{G}^{-5}. \end{aligned}$$

Combining this with the preceding inequality gives

$$\frac{s_1(B_2)^2 s_1(B_1)^2}{s_2(B_2)^2 s_2(B_1)^2} \geq \mathcal{G}^4,$$

so that, by taking square roots, at least one of  $s_1(B_2)/s_2(B_2) \geq \mathcal{G}$  and  $s_1(B_1)/s_2(B_1) \geq \mathcal{G}$  must hold as required.  $\square$

**Lemma 4.7.** *Let  $0 < \varepsilon < 1$ ,  $\mathcal{G} > 2$  and  $N > 2$ . Let  $M_1, \dots, M_{N-1}$  be a sequence of invertible  $3 \times 3$  matrices. If there exists  $1 \leq n \leq N-1$  such that  $\delta_{n,2} < \min\{\mathcal{G}^{-9}, \frac{\varepsilon}{4}\}$ , then there exist  $1 \leq m \leq m' < N$ , a sequence of matrices  $R_m, \dots, R_{m'}$  and target matrices  $M_i^\circ = R_i M_i$  such that*

- (1)  $\|R_i^{\pm 1} - I\| \leq \frac{\varepsilon}{4}$  for each  $m \leq i \leq m'$ ;
- (2)

$$\frac{s_1(M_{m'}^\circ \cdots M_m^\circ)}{s_2(M_{m'}^\circ \cdots M_m^\circ)} \geq \mathcal{G}.$$

*Proof.* Let  $B = M_{N-1} \cdots M_1$ ,  $B_1 = M_n \cdots M_1$ , and  $B_2 = M_{N-1} \cdots M_{n+1}$ . If  $s_1(B)/s_2(B) \geq \mathcal{G}$ ,  $s_1(B_1)/s_2(B_1) \geq \mathcal{G}$  or  $s_1(B_2)/s_2(B_2) \geq \mathcal{G}$ , we can use either  $B$ ,  $B_1$  or  $B_2$  as the target block (with  $m = 1$ ,  $m' = N-1$ ;  $m = 1$ ,  $m' = n$ ; or  $m = n+1$ ,  $m' = N-1$ , respectively, and  $R_i = I$  for all  $i$ ) to satisfy the conclusion of the lemma.

Hence we assume  $B, B_1$  and  $B_2$  have no (1,2)-gap. That is, we assume  $s_1(B)/s_2(B)$ ,  $s_1(B_1)/s_2(B_1)$  and  $s_1(B_2)/s_2(B_2)$  are all less than  $\mathcal{G}$ . Using (4.1) and the absence of a (1,2)-gap for  $B_1$  and  $B_2$ , we see

$$(4.2) \quad s_1(B) \geq s_2(B_1) s_2(B_2) \geq \mathcal{G}^{-2} s_1(B_1) s_1(B_2).$$

Since  $B$  does not have a (1,2)-gap, we see from (4.1)

$$(4.3) \quad s_2(B) \geq \mathcal{G}^{-1} s_1(B) \geq \mathcal{G}^{-1} s_2(B_1) s_2(B_2).$$

Finally, using the fact that  $S_1^3(B) = |\det B| = |\det B_1| |\det B_2| = S_1^3(B_1) S_1^3(B_2)$  together with (4.2) and (4.3) we see

$$(4.4) \quad s_3(B) \leq \mathcal{G}^3 s_3(B_1) s_3(B_2).$$

By hypothesis, there exist  $e \in E_{n,2} \cap S$  and  $f \in F_{n,2} \cap S$  satisfying  $\|e - f\| < \min\{\mathcal{G}^{-9}, \frac{\varepsilon}{4}\}$ . We let  $m = 1$ ,  $m' = N-1$  and let  $R_i = I$  for each  $i \neq n$  so that  $M_i^\circ = M_i$ , for each  $i \neq n$ . We define  $M_n^\circ = R_n M_n$ , where  $R_n$  is a near identity orthogonal transformation such that  $R_n e = f$ . Since  $\|f - e\| < \min\{\mathcal{G}^{-9}, \frac{\varepsilon}{4}\}$ ,  $R_n$  may be chosen so that  $\|R_n^{\pm 1} - I\| \leq \frac{\varepsilon}{4}$ . We next show that  $B^\circ = B_2 R_n B_1$  satisfies condition (2). Let  $e_0$  and  $f_0$  be such that  $B_1 e_0 = e$  and  $B_1 f_0 = f$ .

Since  $f_0 \in F_2(B)$ ,  $\|B f_0\| = s_3(B) \|f_0\|$ . By properties of singular values,  $\|B_1 f_0\| \geq s_3(B_1) \|f_0\|$ . Hence  $\|B_2 R_n B_1 \frac{e_0}{\|e_0\|}\| = \|B_1 \frac{e_0}{\|e_0\|}\| \frac{\|B f_0\|}{\|B_1 f_0\|} \leq s_1(B_1) \frac{s_3(B)}{s_3(B_1)} \leq s_1(B_1) s_3(B_2) \mathcal{G}^3$ , where the last inequality follows from (4.4).

Lemma 4.3 ensures that  $s_1(B_2)/s_3(B_2) \geq \mathcal{G}^9$ . Thus, using (4.2) in the last step, we get

$$(4.5) \quad \|B_2 R_n B_1 \frac{e_0}{\|e_0\|}\| \leq \mathcal{G}^{-6} s_1(B_1) s_1(B_2) \leq \mathcal{G}^{-4} s_1(B).$$

Next,  $\|B_2 R_n B_1 \frac{f_0}{\|f_0\|}\| \leq s_1(B_2) \|B_1 \frac{f_0}{\|f_0\|}\| \leq s_1(B_2) \frac{\|B f_0\|/\|f_0\|}{s_3(B_2)} = s_1(B_2) \frac{s_3(B)}{s_3(B_2)} \leq s_1(B_2) s_3(B_1) \mathcal{G}^3$ , where the last inequality follows from (4.4). Lemma 4.3 ensures that  $s_1(B_1)/s_3(B_1) \geq \mathcal{G}^9$ . Thus, using (4.2) again, we have

$$(4.6) \quad \left\| B_2 R_n B_1 \frac{f_0}{\|f_0\|} \right\| \leq s_1(B_2) s_1(B_1) \mathcal{G}^{-6} \leq s_1(B) \mathcal{G}^{-4}.$$

Note that, by construction,  $e_0 \perp f_0$ . Then, (4.5) and (4.6) imply that for every  $0 \neq v \in \text{lin}(e_0, f_0)$ ,  $\|B_2 R_n B_1 v\|/\|v\| \leq \sqrt{2} \mathcal{G}^{-4} s_1(B) \leq \mathcal{G}^{-3} s_1(B)$ . Therefore by the min-max characterization of singular values,

$$(4.7) \quad s_2(B_2 R_n B_1) \leq \mathcal{G}^{-3} s_1(B).$$

In view of (4.1) and (4.2),

$$(4.8) \quad s_1(B_2 R_n B_1) \geq s_2(B_2 R_n) s_2(B_1) = s_2(B_2) s_2(B_1) \geq \mathcal{G}^{-2} s_1(B_2) s_1(B_1) \geq \mathcal{G}^{-2} s_1(B).$$

Since  $B^\circ = B_2 R_n B_1$ , (4.7) and (4.8) imply  $s_1(B^\circ)/s_2(B^\circ) \geq \mathcal{G}$ , as required.  $\square$

**Lemma 4.8** (Target for  $3 \times 3$  matrices). *Let  $\mathcal{G} > 2$ ,  $0 < \varepsilon < 1$ ,  $\eta = \min\{\mathcal{G}^{-9}, \frac{\varepsilon}{4}\}$  and consider a product of  $3 \times 3$  invertible matrices,  $M_{N-1}, \dots, M_1$ , with  $N-1 \geq \frac{16}{\varepsilon \eta^2} \log \mathcal{G}$ . Then there exists  $1 \leq m \leq m' < N$  and a sequence  $M_m^\circ, \dots, M_{m'}^\circ$  with the properties:*

- (1)  $M_n^\circ$  is a near-identity perturbation of  $M_n$  for each  $m \leq n \leq m'$ ;
- (2) The target  $B^\circ := M_{m'}^\circ \cdots M_m^\circ$  satisfies

$$\frac{s_1(B^\circ)}{s_2(B^\circ)} \geq \mathcal{G}.$$

*Proof.* For each  $1 \leq j < 3$  and  $1 \leq n < N$  we recall that  $\delta_{n,j}$  denotes the minimal distance between points in  $E_{j,n} \cap S$  and  $F_{j,n} \cap S$ . We consider three partially overlapping cases that cover all situations:

- (1) The block  $M_{N-1}, \dots, M_1$  is  $\eta^2$ -spread. That is,  $\delta_n > \eta^2$  for each  $1 \leq n < N$ .
- (2) There exists  $1 \leq n < N$  such that  $\delta_{n,1} \leq \eta^2$ , and  $\delta_{n,2} \geq \eta$  for every  $1 \leq n < N$ .
- (3) There exists  $1 \leq n < N$  such that  $\delta_{n,2} < \eta$ .

In the first and third cases, Lemma 4.2 and Lemma 4.7, respectively, identify targets for  $M_{N-1}, \dots, M_1$ , namely  $B^\circ := M_{N-1}^\circ \cdots M_1^\circ$  satisfies  $\frac{s_1(B^\circ)}{s_2(B^\circ)} \geq \mathcal{G}$ .

To finish, we show how to identify a target block in the second case. Let  $1 \leq n \leq N-1$  be such that  $\delta_{n,1} \leq \eta^2$ , and let  $\alpha v_2^n + \beta v_3^n \in \text{lin}(v_2^n, v_3^n) \cap S$  be such that  $\|v_1^n + \alpha v_2^n + \beta v_3^n\| \leq \eta^2$ . We wish to show  $|\beta| \leq 2\eta$ . If  $\beta = 0$ , this is immediate. If  $\beta \neq 0$ , then we have  $d(v_3^n, \text{lin}(v_1^n, v_2^n)) \leq \eta^2/|\beta|$ , so

using Lemma 2.2,  $\delta_{n,2} \leq 2\eta^2/|\beta|$ . By assumption,  $\delta_{n,2} \geq \eta$ , so it follows that  $|\beta| \leq 2\eta$ .

Hence,  $\|v_1^n + \alpha v_2^n\| \leq 2\eta + \eta^2$ . Therefore, using Lemma 2.2,  $\angle(v_1^n, v_2^n) \leq 2(2\eta + \eta^2) < \mathcal{G}^{-5}$ . Thus Lemma 4.6 shows that either  $B^\odot = M_{N-1} \cdots M_1$  or  $B^\odot = B_1 = M_n \cdots M_1$  or  $B^\odot = B_2 = M_{N-1} \cdots M_{n+1}$  yield a target for  $M_{N-1}, \dots, M_1$ . □

**Corollary 4.9.** *Let  $\mathcal{G} > 2$ ,  $0 < \varepsilon < 1$  and  $\eta = \min\{\mathcal{G}^{-9}, \frac{\varepsilon}{4}\}$ . Consider a product of  $3 \times 3$  invertible matrices,  $M_{N-1}, \dots, M_1$ , with  $N - 1 \geq \frac{16}{\varepsilon\eta^2} \log \mathcal{G}$ . Then, there exist  $1 \leq m \leq m' < N$  and a sequence  $M_m^\odot, \dots, M_{m'}^\odot$  with the properties:*

- (1)  $M_n^\odot$  is a near-identity perturbation of  $M_n$  for each  $m \leq n \leq m'$ ;
- (2) The target  $B^\odot := M_{m'}^\odot \cdots M_m^\odot$  satisfies

$$\frac{s_2(B^\odot)}{s_3(B^\odot)} \geq \mathcal{G}.$$

*Proof.* The result follows directly from Lemma 4.8 and Lemma 4.1. □

## 5. ABSTRACT TARGETS

In this section, we show that one may find targets in all dimensions  $d$ , and for each pair  $j, j + 1$  of exponents with  $1 \leq j \leq d - 1$ . The proof shows that for all  $0 < \varepsilon < 1$  and all  $\mathcal{G} > 0$ , there exists an  $N$  such that the target property is satisfied. However we do not give any upper bounds for  $N$ . The proof is based on a contradiction argument. One supposes for a contradiction that for some  $0 < \varepsilon < 1$  and some  $\mathcal{G} > 0$  there is no such  $N$ . Then there would be arbitrarily long blocks of matrices for which no perturbation has  $s_j(\cdot)/s_{j+1}(\cdot)$  exceeding  $\mathcal{G}$ . One may then take a limit of these blocks to obtain a closed shift-invariant collection of sequences of matrices with the property that for any perturbation of any sub-block, the ratio  $s_j(\cdot)/s_{j+1}(\cdot)$  remains below  $\mathcal{G}$ . In particular, when perturbing such a matrix cocycle, one gets that  $\lambda_j^{\varepsilon/(4d)} = \lambda_{j+1}^{\varepsilon/(4d)}$ . This construction is reminiscent of the Furstenberg Correspondence Principle [22], used in Additive Combinatorics, to relate questions about configurations in positive density subsets of the integers with questions in dynamical systems. It also has some resemblance to the argument in the paper [13] of Bochi and Gourmelon where a dynamical system was constructed to relate two non-dynamical statements. The conclusion of Lemma 5.1, however, contradicts Theorem 5.3 in which we show that additive noiselike perturbations of cocycles always have simple Lyapunov spectrum. Unlike in the previous section, we give no information on how the targets are constructed. Rather, we infer their existence from the contradiction and compactness argument described above.

In this section, we consider an invertible ergodic measure-preserving transformation,  $\sigma$ , of a probability space  $(\Omega, \rho)$ . If  $A: \Omega \rightarrow \text{Mat}_d(\mathbb{R})$  is a map, we

write  $A^{(n)}(\omega) = A(\sigma^{n-1}\omega) \cdots A(\omega)$ . We let  $S_\infty = \{M \in \text{Mat}_d(\mathbb{R}) : |M|_\infty \leq 1\}$  and let  $S_\infty$  be equipped with the uniform measure, so that each entry is uniformly distributed in the range  $[-1, 1]$  and distinct entries are independent. We also use the notation  $S$  to denote  $\{M \in \text{Mat}_d(\mathbb{R}) : \|M\| \leq 1\}$ . As before, we denote by  $\mathbb{P}$  the measure on  $S_\infty^\mathbb{Z}$  where the matrices in the sequence are mutually independent and each is distributed uniformly as described above. Finally let  $\bar{\Omega} = \Omega \times S_\infty^\mathbb{Z}$  and we equip  $\bar{\Omega}$  with the measure  $\rho \times \mathbb{P}$ , invariant under  $\bar{\sigma} := \sigma \times \text{shift}$ . Given  $(\omega, \zeta) \in \Omega \times S_\infty^\mathbb{Z}$ , we write  $A_\varepsilon(\omega, \zeta) = A(\omega) + \varepsilon\zeta_0$  (that is, we perturb  $A(\omega)$  by  $\varepsilon$  times the zeroth matrix in the sequence  $\zeta$ ). We then write  $A_\varepsilon^{(n)}(\omega, \zeta) = A_\varepsilon(\bar{\sigma}^{n-1}(\omega, \zeta)) \cdots A_\varepsilon(\omega, \zeta)$ . Since  $\rho \times \mathbb{P}$  is ergodic,  $\lim_{n \rightarrow \infty} \frac{1}{n} \log s_j(A_\varepsilon^{(n)}(\omega))$  exists  $\rho \times \mathbb{P}$ -a.e. and is almost surely equal to a constant that we call  $\mu_j^\varepsilon$ . As usual, we let  $\lambda_1^\varepsilon > \dots > \lambda_k^\varepsilon$  be the distinct almost-sure constant values taken by  $\mu_j^\varepsilon$  and we let  $m_j^\varepsilon$  be the multiplicity of  $\lambda_j^\varepsilon$ .

**Lemma 5.1** (Compactness). *Suppose that for some  $0 < \epsilon < 1$ , for all  $\delta > 0$ , there exists a sequence of  $d \times d$  matrices  $(A_i)$  of norm at most 1 and  $1 \leq j < d$ , such that*

$$\liminf \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)})}{s_{j+1}(A_\varepsilon^{(n)})} < \delta \text{ a.s.}$$

*Then there exists an ergodic invariant measure on  $S^\mathbb{Z}$  such that  $\lambda_j^{\varepsilon/(4d)} = \lambda_{j+1}^{\varepsilon/(4d)}$ .*

*Proof.* Notice that the hypothesis of the lemma implies that the conclusion of Corollary 3.9 fails, so that the hypothesis of Corollary 3.9 must fail also. Hence we deduce that there exists  $\mathcal{G} > 0$  such that for all  $N$ , there exists a block  $M_1, \dots, M_N$  of matrices with the properties:

- $\|M_i\| \leq 1 - \frac{\varepsilon}{2}$  for each  $i$ ;
- $s_d(M_i) \geq \frac{\varepsilon}{2}$  for each  $i$ ;
- for all  $1 \leq m \leq m' \leq N$  and all  $M'_m, \dots, M'_{m'}$  such that  $|M'_i - M_i|_\infty \leq \frac{\varepsilon}{4d}$  for each  $m \leq i \leq m'$ , one has

$$\frac{s_j(M'_{m'} \cdots M'_m)}{s_{j+1}(M'_{m'} \cdots M'_m)} \leq \mathcal{G}.$$

For each  $N$ , let  $M_{N,1}, \dots, M_{N,2N+1}$  be such a block with the above property, of length  $2N + 1$ , and build an element  $A_N$  of  $S^\mathbb{Z}$  by  $A_N = (A_{N,i})_{i \in \mathbb{Z}}$  by

$$A_{N,i} = \begin{cases} M_{N,i+N} & \text{if } -N \leq i \leq N; \\ I & \text{otherwise.} \end{cases}$$

Equip  $S$  with the norm topology induced by  $\|\cdot\|$  and  $S^\mathbb{Z}$  with the product topology. Then  $S$  is compact and so is  $S^\mathbb{Z}$ . Let  $A = (A_i)_{i \in \mathbb{Z}}$  be a sub-sequential limit of  $A_N = (A_{N,i})_{i \in \mathbb{Z}}$ . That is, there is a sequence  $(N_q)$  converging to  $\infty$  such that  $A_i = \lim_{q \rightarrow \infty} A_{N_q,i}$  for each  $i$ . Then let

$-\infty < m \leq m' < \infty$  and let  $B_m, \dots, B_{m'}$  be an arbitrary perturbation such that  $|B_i - A_i|_\infty \leq \frac{\varepsilon}{4d}$  for each  $i$ . Let  $B_{q,i} = A_{N_q,i} + (B_i - A_i)$  so that  $B_{q,i} \rightarrow B_i$  for each  $m \leq i \leq m'$  and  $|B_{q,i} - A_{N_q,i}|_\infty \leq \frac{\varepsilon}{4d}$  for each  $m \leq i \leq m'$ . By the choice of the  $A_N$ 's and the assumption, we have that for all sufficiently large  $q$ ,

$$\frac{s_j(B_{q,m'} \cdots B_{q,m})}{s_{j+1}(B_{q,m'} \cdots B_{q,m})} \leq \mathcal{G}.$$

By continuity of singular values,

$$\frac{s_j(B_{m'} \cdots B_m)}{s_{j+1}(B_{m'} \cdots B_m)} \leq \mathcal{G}.$$

Let  $S_0 = \{M \in \text{Mat}_d(\mathbb{R}) : \|M\| \leq 1 - \frac{\varepsilon}{2} \text{ and } s_d(M) \geq \frac{\varepsilon}{2}\}$  and set

$$\Omega = \{A \in S_0^{\mathbb{Z}} : q_j(B_{m'} \cdots B_m) \leq \mathcal{G} \text{ for all } m < m' \text{ whenever } |B_i - A_i|_\infty \leq \frac{\varepsilon}{4d}, \forall i\},$$

where  $q_j(M) = s_j(M)/s_{j+1}(M)$  as before. The arguments above show that  $\Omega$  is a non-empty closed, and hence compact, shift-invariant subset of  $S_0^{\mathbb{Z}}$ . By the Krylov-Bogoliubov theorem, there exists a shift-invariant measure supported on  $\Omega$ . Taking an ergodic component, we may assume that  $\mathbb{P}$  is an ergodic measure supported on  $\Omega$ .

Now for every  $\omega \in \Omega$  and every sequence  $(\Xi_i)$  in  $S_\infty^{\mathbb{Z}}$ , by definition

$$\frac{s_j\left(\left(A(\sigma^{n-1}\omega) + \frac{\varepsilon}{4d}\Xi_{n-1}\right) \cdots \left(A(\omega) + \frac{\varepsilon}{4d}\Xi_0\right)\right)}{s_{j+1}\left(\left(A(\sigma^{n-1}\omega) + \frac{\varepsilon}{4d}\Xi_{n-1}\right) \cdots \left(A(\omega) + \frac{\varepsilon}{4d}\Xi_0\right)\right)} \leq \mathcal{G}.$$

It follows that  $\lambda_j^{\varepsilon/(4d)} = \lambda_{j+1}^{\varepsilon/(4d)}$ .  $\square$

The following result is essentially well known (see for example [12]). We include the proof for completeness.

**Lemma 5.2** ((Semi-)continuity). *Let  $\sigma$  be an invertible ergodic measure-preserving transformation of a probability space  $(\Omega, \rho)$  and let  $A: \Omega \rightarrow \text{Mat}_m(\mathbb{R})$  be a cocycle of matrices such that  $\|A(\omega)\|^{-1} \leq c$  a.e. and  $\int \log \|A(\omega)\| d\rho < \infty$ . Suppose  $A$  has trivial Lyapunov spectrum. That is, there exists  $\lambda$  such that  $\lim_{n \rightarrow \infty} \frac{1}{n} \log s_j(A^{(n)}(\omega)) = \lambda$  a.e. for  $j = 1, \dots, m$ . Then for all  $\varepsilon > 0$ , there exists a  $\delta > 0$  with the following property:*

*Let  $\bar{\sigma}: (\bar{\Omega}, \bar{\rho}) \rightarrow (\bar{\Omega}, \bar{\rho})$  be an ergodic invertible extension of  $\sigma: (\Omega, \rho) \rightarrow (\Omega, \rho)$  with factor map  $\pi$ . If  $B: \bar{\Omega} \rightarrow \text{Mat}_m(\mathbb{R})$  satisfies  $\|B(\bar{\omega}) - A(\pi(\bar{\omega}))\| \leq \delta$  for  $\bar{\rho}$ -a.e.  $\bar{\omega}$ , then all Lyapunov exponents of the cocycle  $B$  lie in the interval  $(\lambda - \varepsilon, \lambda + \varepsilon)$ .*

The proof here is essentially a semi-continuity result (that holds for arbitrary cocycles), and the additional assumption that the unperturbed Lyapunov spectrum is trivial yields the continuity result.

*Proof.* By the Multiplicative ergodic theorem and the Kingman sub-additive ergodic theorem,

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|A^{(n)}(\omega)\| d\rho = \inf_n \frac{1}{n} \int \log \|A^{(n)}(\omega)\| d\rho \quad \text{and}$$



$$-\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|(A^{(n)}(\omega))^{-1}\| d\rho = \inf_n \frac{1}{n} \int \log \|(A^{(n)}(\omega))^{-1}\| d\rho.$$

Let  $n$  be such that

$$\begin{aligned} \frac{1}{n} \int \log \|A^{(n)}(\omega)\| d\rho &< \lambda + \varepsilon \text{ and} \\ \frac{1}{n} \int \log \|(A^{(n)}(\omega))^{-1}\| d\rho &< -\lambda + \varepsilon. \end{aligned}$$

Define

$$\begin{aligned} f_\delta(\omega) &= \max_{\|\Xi_0\|, \dots, \|\Xi_{n-1}\| \leq \delta} \log \left\| (A(\sigma_{n-1}(\omega)) + \Xi_{n-1}) \cdots (A(\omega) + \Xi_0) \right\| \text{ and} \\ g_\delta(\omega) &= \max_{\|\Xi_0\|, \dots, \|\Xi_{n-1}\| \leq \delta} \log \left\| ((A(\sigma_{n-1}(\omega)) + \Xi_{n-1}) \cdots (A(\omega) + \Xi_0))^{-1} \right\|. \end{aligned}$$

Let  $\delta < \min(1, \frac{1}{2c})$ . For  $0 < \varepsilon < \delta$  and  $\Xi$  such that  $\|\Xi\| \leq 1$  and  $A$  such that  $s_d(A) \geq \frac{1}{c}$ , we have  $(\frac{1}{c} - \frac{1}{2c})\|x\| \leq \|(A + \varepsilon\Xi)x\| \leq ((\max(\|A\|, 1) + 1)\|x\|$ . Since  $\log(\max(t, 1) + 1) \leq \log^+ t + 1$ , the right inequality gives that  $\log \|(A + \varepsilon\Xi)\| \leq \log^+ \|A\| + 1$ . The left inequality gives  $\log \|(A + \varepsilon\Xi)^{-1}\| \leq \log(2c)$ . This yields  $f_\delta(\omega) \leq \sum_{i=0}^{n-1} (1 + \log^+ \|A(\sigma^i \omega)\|)$  and  $g_\delta(\omega) \leq n \log 2c$ . Since  $\|A\| \|A\|^{-1} \geq 1$ , we also obtain  $f_\delta(\omega) + g_\delta(\omega) \geq 0$ , giving the bounds

$$\begin{aligned} n \log \frac{1}{2c} \leq f_\delta(\omega) &\leq \sum_{i=0}^{n-1} (1 + \log^+ \|A(\sigma^i \omega)\|) \text{ and} \\ - \sum_{i=0}^{n-1} (1 + \log^+ \|A(\sigma^i \omega)\|) &\leq g_\delta(\omega) \leq n \log(2c). \end{aligned}$$

Also  $f_\delta(\omega) \rightarrow \log \|A^{(n)}(\omega)\|$  and  $g_\delta(\omega) \rightarrow \log \|(A^{(n)}(\omega))^{-1}\|$  for each  $\omega \in \Omega$  as  $\delta \rightarrow 0$ . Hence by dominated convergence, for all sufficiently small  $\delta$ ,  $\frac{1}{n} \int f_\delta d\rho < \lambda + \varepsilon$  and  $\frac{1}{n} \int g_\delta d\rho < -\lambda + \varepsilon$ .

Now let  $(\bar{\Omega}, \bar{\rho})$  be an extension of  $(\Omega, \rho)$  and let  $B$  be a cocycle on  $\bar{\Omega}$  such that  $\|B(\bar{\omega}) - A(\pi(\bar{\omega}))\| \leq \delta$  a.e. Then by definition  $\log \|B^{(n)}(\bar{\omega})\| \leq f_\delta(\pi(\bar{\omega}))$  and  $\log \|(B^{(n)}(\bar{\omega}))^{-1}\| \leq g_\delta(\pi(\bar{\omega}))$ . It follows that  $\lambda_1(B) \leq \lambda + \varepsilon$  and  $\lambda_m(B) \geq -\frac{1}{n} \int g_\delta(\pi(\bar{\omega})) d\rho \geq \lambda - \varepsilon$  as required.  $\square$

**Theorem 5.3** (Simplicity). *Let  $\sigma$  be an invertible ergodic measure-preserving transformation of a probability space  $(\Omega, \rho)$  and let  $A: \Omega \rightarrow \text{Mat}_d(\mathbb{R})$  such that  $\|A(\omega)\| \leq 1$  for every  $\omega \in \Omega$ . Then for all  $0 < \varepsilon < 1$ , for all  $1 \leq j < d$ , and  $\rho \times \mathbb{P}$ -almost surely,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{s_j(A_\varepsilon^{(n)}(\omega))}{s_{j+1}(A_\varepsilon^{(n)}(\omega))} > 0.$$

*That is  $\lambda_j^\varepsilon > \lambda_{j+1}^\varepsilon$  for each  $j$ : the noiselike perturbation of an arbitrary bounded matrix cocycle has simple Lyapunov spectrum.*

*Proof.* The proof is divided in three steps.

*Step 0: Initial perturbation*

Let  $\sigma$  be an ergodic invertible measure-preserving transformation of  $(\Omega, \rho)$  and  $A: \Omega \rightarrow \text{Mat}_d(\mathbb{R})$ , such that  $\|A(\omega)\| \leq 1$  for every  $\omega \in \Omega$ . Let  $0 < \varepsilon < 1$  and let  $1 \leq j \leq d-1$  be fixed. We first apply Lemma 3.1 to obtain a cocycle  $A_0$  where  $\|A_0(\omega) - A(\omega)\| \leq \frac{\varepsilon}{2}$ ,  $\|A_0(\omega)\| \leq 1 - \frac{\varepsilon}{2}$  and  $s_d(A_0(\omega)) \geq \frac{\varepsilon}{2}$  for each  $\omega \in \Omega$ .

*Step 1: A nearby extension with simple Lyapunov spectrum*

In this step, we find an extension  $\Omega'$  of  $\Omega$ , together with a factor map  $\pi: \Omega' \rightarrow \Omega$  and a cocycle  $A'$  on  $\Omega'$  such that  $\|A'(\omega') - A(\pi(\omega'))\| < \frac{3\varepsilon}{4}$  such that the cocycle  $A'$  has simple Lyapunov spectrum.

We inductively build a sequence of extensions of  $\Omega$ , each one breaking a tie between a pair of Lyapunov exponents at the previous level, taking care not to create any new ties. We let  $\Omega_0 = \Omega$ . Then we build an extension  $\Omega_1$  of  $\Omega_0$  with a new cocycle  $A_1$ ; an extension  $\Omega_2$  of  $\Omega_1$  with a cocycle  $A_2$  etc.

More precisely, we claim the following: There exist  $0 \leq K \leq d-1$ , spaces  $\Omega_0, \dots, \Omega_K$ ; transformations  $\sigma_0, \dots, \sigma_K$ ; probability measures  $\rho_0, \dots, \rho_K$ ; factor maps  $\pi_1, \dots, \pi_K$  and cocycles  $A_0, \dots, A_K$  defined on  $\Omega_0, \dots, \Omega_K$  such that:

- The map  $\sigma_k$  is an invertible ergodic measure-preserving transformation of  $(\Omega_k, \rho_k)$  for each  $k = 0, \dots, K$ ;
- For each  $1 \leq k \leq K$  the map  $\pi_k$  is a factor map from  $(\Omega_k, \rho_k)$  to  $(\Omega_{k-1}, \rho_{k-1})$ :  $\rho_k(\pi_k^{-1}B) = \rho_{k-1}(B)$  for all measurable subsets  $B$  of  $\Omega_{k-1}$  and  $\sigma_{k-1} \circ \pi_k = \pi_k \circ \sigma_k$ ;
- Let the Lyapunov exponents of the cocycle  $A_k$  over  $\Omega_k$  *with repetition* be  $\mu_{k,1} \geq \dots \geq \mu_{k,d}$ . For any  $1 \leq k \leq K$ , if  $\mu_{k-1,j} > \mu_{k-1,j+1}$ , then  $\mu_{k,j} > \mu_{k,j+1}$ . Further there exists at least one  $j$  such that  $\mu_{k-1,j} = \mu_{k-1,j+1}$  while  $\mu_{k,j} > \mu_{k,j+1}$ ;
- Each cocycle is a small perturbation of the previous one:  $\|A_k(\omega) - A_{k-1}(\pi_k(\omega))\| \leq \frac{\varepsilon}{4d}$  for each  $1 \leq k \leq K$ ;
- The  $K$ th cocycle has simple Lyapunov spectrum:  $\mu_{K,1} > \dots > \mu_{K,d}$ .

For the base case,  $\Omega_0 = \Omega$ ,  $\rho_0 = \rho$  and we already constructed  $A_0$ . Now suppose  $\Omega_0, \dots, \Omega_k$ ;  $\rho_0, \dots, \rho_k$ ;  $\sigma_0, \dots, \sigma_k$  and  $A_0, \dots, A_k$  have already been constructed to satisfy the above conditions.

Let the Lyapunov exponents of the  $A_k$  cocycle over  $\sigma_k$  *without repetition* be  $\lambda_{k,1} > \dots > \lambda_{k,l_k}$  with multiplicities  $m_{k,1}, \dots, m_{k,l_k}$ . If the cocycle  $A_k$  over  $\sigma_k$  has simple Lyapunov spectrum, we let  $K = k$  and the induction is complete. Otherwise, fix a  $j$  such that  $m_{k,j} > 1$ . Applying the Multiplicative Ergodic Theorem, let  $V(\omega)$  be the Oseledets space corresponding to the  $j$ th Lyapunov exponent of the  $A_k$  cocycle, and  $U(\omega)$  be the direct sum of all of the other Oseledets spaces. For  $\mathbb{P}_k$ -a.e.  $\omega$ ,  $V(\omega)$  and  $U(\omega)$  have trivial intersection and satisfy  $A_k(\omega)V(\omega) = V(\sigma_k\omega)$  and  $A_k(\omega)U(\omega) = U(\sigma_k\omega)$ . The function  $g(\omega) := d(U(\omega) \cap S, V(\omega) \cap S)$  is measurable and almost everywhere positive. Hence there exists an  $\eta > 0$  such that  $\mathbb{P}_k(g(\omega) >$

$\eta) \geq \frac{2}{3}$ . We let  $G = \{\omega \in \Omega_k : g(\omega) > \eta\}$  and say this is the *good set*. Given  $\omega \in \Omega_k$ , the  $n$ 's such that  $\sigma_k^n(\omega) \in G$  are called the *good times*.

We modify the cocycle only on the good set. Also our modifications change only what is happening on  $V(\omega)$ , leaving all of the cocycle alone on  $U(\omega)$ . That is, we are considering perturbations  $A'(\omega')$  of  $A(\omega)$  such that  $A'(\omega')|_{U(\omega)} = A(\omega)|_{U(\omega)}$ . In fact, we will also require  $A'(\omega')(V(\omega)) = V(\sigma\omega)$  so that the perturbed cocycle preserves the equivariant direct sum  $\mathbb{R}^d = U(\omega) \oplus V(\omega)$ . It follows from the Multiplicative Ergodic Theorem that the multiset that is the Lyapunov spectrum (with repetition) of the cocycle  $A'$  is the union of the Lyapunov spectra of the restriction of  $A'$  to  $U(\omega)$  and the restriction of  $A'$  to  $V(\omega)$ . Also, the Lyapunov spectrum of the restriction of  $A'$  to  $U(\omega)$  agrees with the Lyapunov spectrum of the restriction of  $A$  to  $U(\omega)$ .

Notice that  $A'(\omega') - A(\omega) = (A'(\omega') - A(\omega)) \circ \Pi_{V(\omega)|_{U(\omega)}}$ , so that  $\|A'(\omega') - A(\omega)\| \leq \|(A'(\omega') - A(\omega))|_{V(\omega)}\| \|\Pi_{V(\omega)|_{U(\omega)}}\|$ . By Lemma 2.1, at good times  $\|\Pi_{V(\omega)|_{U(\omega)}}\| < \frac{2}{\eta}$ . To ensure that  $\|A'(\omega') - A(\omega)\| \leq \frac{\varepsilon}{4d}$ , it suffices to ensure that  $\|(A'(\omega') - A(\omega))|_{V(\omega)}\| \leq \frac{\eta\varepsilon}{8d}$ .

We also want to ensure that the Lyapunov exponents of the restriction of the perturbed cocycle to  $V(\omega)$  remain strictly bigger than  $\lambda_{j+1}$  and strictly smaller than  $\lambda_{j-1}$  (where we take  $\lambda_0$  to be  $\infty$  and  $\lambda_{l_k+1}$  to be  $-\infty$ ). To this end, we use Lemma 5.2 to obtain a  $\delta$  such that if we build an extension  $\Omega_{k+1}$  and a cocycle such that  $|(B_{\bar{\omega}} - A_{\pi(\bar{\omega})})|_{V(\pi(\bar{\omega}))}|_{\infty} \leq \delta$  for all  $\bar{\omega} \in \Omega_{k+1}$ , then the Lyapunov exponents of the restriction of  $B$  to  $V(\pi(\bar{\omega}))$  are in the range  $(\lambda_{j+1}, \lambda_{j-1})$ . Hence this paragraph and the previous paragraph taken together allow us to choose a  $\kappa$  so that if the restriction of the extension cocycle to  $V(\omega)$  differs at good times by at most  $\kappa$  and the extension cocycle restricted to  $U(\omega)$  is unchanged, then the exponents of the restriction of the cocycle to  $V(\omega)$  are in the range  $(\lambda_{j+1}, \lambda_{j-1})$  and  $\|A'(\omega') - A(\omega)\| \leq \frac{\varepsilon}{4d}$ .

We use a very similar strategy to that in Section 4.3 where we found a target for  $s_1/s_d$ . Let  $\mathcal{G}_k$  be as computed in the start of Theorem 3.8 with  $\varepsilon$  replaced by  $\kappa$  and  $d$  replaced by  $m_{k,j}$ . Here we aim to create a gap between the fastest and slowest vectors in a block within  $V(\omega)$ . In a block, either there are two orthogonal vectors whose angle of separation at some stage within the block is less than  $1/\mathcal{G}_k$ ; otherwise one may incrementally boost one direction.

Now let  $N > \frac{32}{\kappa} \mathcal{G}_k \log \mathcal{G}_k$ . We apply Rokhlin's lemma to build a Rokhlin tower with base  $H$  of positive  $\mathbb{P}_k$ -measure contained in the good set  $G$  such that the return time to the base always exceeds  $N$ . Let  $r_H$  denote the return time to  $H$  under  $\sigma_k$ . An element  $\omega \in H$  of the base is said to be *good* if  $\sum_{i=0}^{r_H(\omega)-1} \mathbf{1}_G(\sigma_k^i \omega) > \frac{r_H(\omega)}{2}$  and let  $H' \subset H$  denote the set of good elements of the base. Since  $\rho_k(G) \geq \frac{2}{3}$ ,  $H'$  has positive  $\rho_k$ -measure. For  $\omega \in H'$ , let  $B^i(\omega) = A_k(\sigma_k^i \omega) \cdots A_k(\sigma_k \omega)$ . If there exists a good time  $1 < i \leq r_H(\omega)$  such that there exist two orthogonal vectors  $v$  and  $v'$  in  $V(\sigma_k^i(\omega))$  for which  $\angle(B^{i-1}(\omega)v, B^{i-1}(\omega)v') < \frac{1}{\mathcal{G}_k}$ , then

$s_1(B^{i-1}(\omega)|_{V(\sigma_k(\omega))})/s_{m_{k,j}}(B^{i-1}(\omega)|_{V(\sigma_k(\omega))}) > \mathcal{G}_k$  by Lemma 4.3. In that case, the block  $B^{i-1}(\omega)$  is taken to be the target block.

Otherwise, let  $v_1, \dots, v_{m_{k,j}}$  be a (measurably-chosen) family of singular vectors for the restriction of  $B^{r_H(\omega)-1}(\omega)$  to  $V(\sigma_k(\omega))$  with singular values  $s_1 \geq \dots \geq s_{m_{k,j}}$ . As in Lemma 4.2, we modify the cocycle (but only at the good times), each time expanding the image of  $v_1$  by a factor of  $1 + \frac{\kappa}{8\mathcal{G}_k}$ , leaving the other  $v$  images (as well as  $U(\sigma_k^i(\omega))$ ) unchanged. We write  $A'(\sigma_k(\omega)), \dots, A'(\sigma_k^{r_H(\omega)-1}(\omega))$  for the matrices in the perturbed block and  $B^{r_H(\omega)-1}(\omega)$  for the product  $A'(\sigma_k^{r_H(\omega)-1}(\omega)) \cdots A'(\sigma_k(\omega))$ . By the end of the block, as in the lemma, since the number of good times exceeds  $N/2$ , we have  $s_1(B^{r_H(\omega)-1}(\omega))/s_{m_{k,j}}(B^{r_H(\omega)-1}(\omega)) > \mathcal{G}_k$ . We call this a Type II target.

The coordinates immediately preceding and following a target block are called transition coordinates as in the proof of Theorem 3.8. We then define an extension system  $\Omega_{k+1} = \Omega_k \times S_\infty^{\mathbb{Z}}$  equipped with the measure  $\rho_k \times \mathbb{P}$  where  $\mathbb{P}$  is the i.i.d. measure on matrices with independent uniform  $[-1, 1]$  elements. We define the cocycle  $A_{k+1}$  by

$$A_{k+1}(\omega, \Xi) = \begin{cases} A'(\omega) & \text{if } \omega \text{ lies inside a Type II target block;} \\ A(\omega) + \kappa \Xi_0 & \text{if } \omega \text{ is a transition coordinate;} \\ A(\omega) & \text{otherwise.} \end{cases}$$

A calculation exactly analogous to the calculation in Theorem 3.8 shows that the restriction of the  $A_{k+1}$  cocycle to the equivariant  $V(\omega)$  block has non-trivial Lyapunov spectrum. By the choice of  $\kappa$ , all Lyapunov exponents lie in  $(\lambda_{j-1}, \lambda_{j+1})$ , completing this step of the induction.

*Step 2: Completion of the proof*

At the end of the induction, we have an ergodic invertible extension  $\bar{\Omega} := \Omega_K$  of  $\Omega$  and a measurable cocycle  $\bar{A} := A_K$  with simple Lyapunov spectrum  $\lambda_1 > \dots > \lambda_d$ . Let  $\bar{\pi} = \pi_1 \circ \dots \circ \pi_K$  denote the factor map from  $\bar{\Omega}$  to  $\Omega$ . We have  $\|A_0(\bar{\pi}(\bar{\omega})) - A(\bar{\pi}(\bar{\omega}))\| < \frac{\varepsilon}{2}$  for each  $\bar{\omega} \in \bar{\Omega}$ . Similarly, by definition of the perturbations, we have  $\|A_k(\pi_{k+1} \circ \dots \circ \pi_K(\bar{\omega})) - A_{k-1}(\pi_k \circ \dots \circ \pi_K(\bar{\omega}))\| \leq \frac{\varepsilon}{4d}$  for each  $k$  and each  $\bar{\omega} \in \bar{\Omega}$ . Summing these, we obtain  $\|\bar{A}(\bar{\omega}) - A(\bar{\pi}(\bar{\omega}))\| \leq \frac{3\varepsilon}{4}$  for each  $\bar{\omega} \in \bar{\Omega}$ .

Let  $\mathcal{G}$  be the required gap identified in the beginning of the proof of Theorem 3.8 for perturbations of size  $\varepsilon$  and dimension  $d$ . From Raghunathan's proof of Oseledets' theorem [38], we know that  $\lim_{n \rightarrow \infty} \frac{1}{n} \log s_j(\bar{A}^{(n)}(\bar{\omega})) \rightarrow \lambda_j$  for each  $1 \leq j \leq d$  and for a.e.  $\bar{\omega} \in \bar{\Omega}$ . Since the Lyapunov spectrum is simple, it follows that  $s_j(\bar{A}^{(n)}(\bar{\omega}))/s_{j+1}(\bar{A}^{(n)}(\bar{\omega})) \rightarrow \infty$  for a.e.  $\bar{\omega} \in \bar{\Omega}$ . Hence, there exists an  $N$  such that the probability that  $s_j(\bar{A}^{(N-1)}(\bar{\omega}))/s_{j+1}(\bar{A}^{(N-1)}(\bar{\omega}))$  exceeds  $\mathcal{G}$  is at least  $\frac{1}{2}$ . Let  $\bar{G} \subset \bar{\Omega}$  be this set.

Since  $\bar{\rho}$  is an ergodic invariant measure,  $\bar{\rho}$ -a.e.  $\bar{\omega}$  hits  $\bar{G}$  with frequency equal to  $\bar{\rho}(\bar{G}) \geq \frac{1}{2}$ . Since  $\bar{\rho}$  may not be ergodic with respect to  $\sigma^N$ , we are unable to conclude that for a.e.  $\bar{\omega} \in \bar{\Omega}$ ,  $\lim_{M \rightarrow \infty} \#\{n \leq M: \bar{\sigma}^{nN} \bar{\omega} \in$

$\bar{G}\}/M \geq \rho(\bar{G})$ , however this inequality holds on a set of measure at least  $\frac{1}{N}$ . Call this set  $\bar{H}$ .

Let  $\bar{\omega} \in \bar{H}$  and consider a realization  $(X_n)_{n \in \mathbb{Z}}$  of the cocycle  $A_\varepsilon$  where  $X_n = A_{\sigma^n(\bar{\omega})} + \varepsilon \Xi_n$ . We say that the realization hits the target area on the  $\ell$ th block if the following two conditions are satisfied:

- (1)  $s_j(\bar{A}^{(N-1)}(\bar{\sigma}^{\ell N} \bar{\omega})) / s_{j+1}(\bar{A}^{(N-1)}(\bar{\sigma}^{\ell N} \bar{\omega})) > \mathcal{G}$ ; and
- (2)  $|X_n - \bar{A}_{\bar{\sigma}^n(\bar{\omega})}|_\infty \leq (\frac{\varepsilon}{4})^N / (3dN)$  for  $n = \ell N, \dots, \ell N + (N-1)$ .

By definition of  $\bar{H}$ , the first condition is satisfied with frequency at least  $\frac{1}{2}$ . Given that the first condition is satisfied, the block  $(A_\varepsilon(\sigma^n \omega) + \varepsilon \Xi_n)_{n=\ell N}^{\ell N+(N-2)}$  hits the target area with probability  $[(\frac{\varepsilon}{4})^N / (3dN)]^{d^2(N-1)}$  (independently of whether other targets are hit). Following the proof of Theorem 3.8, we see that the  $j$ th exponent of the  $(A_{n,\varepsilon}(\omega))$  cocycle is strictly larger than the  $(j+1)$ st.  $\square$

## 6. PROOFS OF MAIN THEOREMS

*Proof of Theorem C.* Lemma 4.4 shows that the hypothesis of Theorem 3.8 is satisfied. From the proof of Theorem 3.8,  $\mathcal{G} = C/\varepsilon^{8d}$ . From Lemma 4.4, the block length is given by  $N = (16/\varepsilon)\mathcal{G} \log \mathcal{G}$ , so that  $N = \varepsilon^{-(8d+1+o(1))}$ . The lower bound is then given by the quantity  $\frac{p}{N}$ , where  $p$  appears in Corollary 3.3. That is,  $c'_d(\varepsilon) = (\frac{\varepsilon}{4})^{d^2 N^2} / [N(3dN)^{d^2 N}]$ , so that  $c'_d(\varepsilon) = \exp(-4\varepsilon^{-(16d+2+o(1))} |\log \frac{\varepsilon}{4}| - 4\varepsilon^{-(8d+1+o(1))} \log(3dN) - \log N)$ . Hence  $c'_d(\varepsilon) = \exp(-\varepsilon^{-(16d+2+o(1))})$  as claimed.  $\square$

Note that in the case  $d = 2$ ,  $s_1(A)/s_d(A) = s_1(A)/s_2(A)$ . Hence  $c_2(\varepsilon) = c'_2(\varepsilon)$ , so that Theorem A is a special case of Theorem C.

*Proof of Theorem B.* Lemma 4.8 in the case  $j = 1$  or Corollary 4.9 in the case  $j = 2$  shows that the hypothesis of Theorem 3.8 is satisfied. From the proof of Theorem 3.8, we take  $\mathcal{G} = C/\varepsilon^{8d} = C/\varepsilon^{24}$ . From Lemma 4.8 or Corollary 4.9, for small  $\varepsilon$ ,  $\eta = \mathcal{G}^{-9} = C\varepsilon^{216}$ , so that  $N = \varepsilon^{-433+o(1)}$ , giving  $c_3(\varepsilon) = \exp(-1/\varepsilon^{866+o(1)})$  as claimed.  $\square$

*Proof of Theorem D.* Let  $1 \leq j < d$ . Let  $\mathcal{G} = C/\varepsilon^{8d}$  as required in Theorem 3.8. Since orthogonal matrices preserve orthogonal frames, any block is  $\sqrt{2}$ -spread. Letting  $N$  be as in the statement of Lemma 4.2, we see  $N = \varepsilon^{-(1+o(1))}$ . Substituting in the expression for  $\frac{p}{N}$  in Theorem 3.8, we see that  $c(\varepsilon) = \exp(-1/\varepsilon^{2+o(1)})$  as required.  $\square$

*Proof of Theorem E.* Let  $\varepsilon > 0$ . Let  $\mathcal{G} = C/\varepsilon^{8d}$  be as constructed at the start of the proof of Theorem 3.8. Suppose for a contradiction that there does not exist a  $c_d(\varepsilon) > 0$  such that for all sequences of matrices  $(A_n)$  of norm at most 1 and all  $1 \leq j \leq d-1$ , one has  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log[s_j(A_\varepsilon^{(n)})/s_{j+1}(A_\varepsilon^{(n)})] \geq c_d(\varepsilon)$ . That is, there exist  $1 \leq j \leq d-1$  and sequences of matrices  $(A_n)$  such that  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log[s_j(A_\varepsilon^{(n)})/s_{j+1}(A_\varepsilon^{(n)})]$  is arbitrarily close to 0. Then the

hypothesis of Lemma 5.1 is satisfied. So one concludes that there exists an ergodic invariant measure on  $S^{\mathbb{Z}}$  for which  $\lambda_j^{\varepsilon/(4d)} = \lambda_{j+1}^{\varepsilon/(4d)}$ . However, this contradicts the conclusion of Theorem 5.3, so that there must exist a positive universal gap  $c_d(\varepsilon) > 0$ .  $\square$

## REFERENCES

- [1] L. Arnold and N. D. Cong. On the simplicity of the Lyapunov spectrum of products of random matrices. *Ergodic Theory Dynam. Systems*, 17:1005–1025, 1997. 4
- [2] L. Arnold and N. D. Cong. Linear cocycles with simple Lyapunov spectrum are dense in  $L^\infty$ . *Ergodic Theory Dynam. Systems*, 19:1389–1404, 1999. 4
- [3] A. Avila and M. Viana. Simplicity of Lyapunov spectra: a sufficient criterion. *Port. Math. (N.S.)*, 64:311–376, 2007. 4
- [4] A. Avila and M. Viana. Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture. *Acta Mathematica*, 198:1–56, 2007. 4
- [5] L. Backes, M. Poletti, P. Varandas, and Y. Lima. Simplicity of Lyapunov spectrum for linear cocycles over non-uniformly hyperbolic systems. *Ergodic Theory Dynam. Systems*, 40:2947–2969, 2020. With an appendix by Lima. 4
- [6] P. H. Baxendale and L. Goukasian. Lyapunov exponents for small random perturbations of Hamiltonian systems. *Ann. Probab.*, 30:101–134, 2002. 5
- [7] S. Bednarski and A. Quas. Lyapunov exponents of orthogonal-plus-normal cocycles, 2023. ArXiv 2309.08193. 3, 5
- [8] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. of Math. (2)*, 133:73–169, 1991. 4
- [9] A. Blumenthal, J. Xue, and L.-S. Young. Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. *Ann. of Math. (2)*, 185:285–310, 2017. 4
- [10] A. Blumenthal, J. Xue, and L.-S. Young. Lyapunov exponents and correlation decay for random perturbations of some prototypical 2D maps. *Comm. Math. Phys.*, 359:347–373, 2018. 4
- [11] A. Blumenthal and Y. Yang. Positive Lyapunov exponent for random perturbations of predominantly expanding multimodal circle maps. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 39:419–455, 2022. 5
- [12] J. Bochi. Genericity of zero Lyapunov exponents. *Ergodic Theory Dynam. Systems*, 22:1667–1696, 2002. 4, 24
- [13] J. Bochi and N. Gourmelon. Some characterizations of domination. *Math. Z.*, 209:221–231, 2009. 22
- [14] J. Bochi and M. Viana. The Lyapunov exponents of generic volume-preserving and symplectic maps. *Ann. of Math. (2)*, 161:1423–1485, 2005. 4
- [15] C. Bonatti and M. Viana. Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergodic Theory Dynam. Systems*, 24:1295–1330, 2004. 4
- [16] D. Chemnitz and M. Engel. Positive Lyapunov exponent in the Hopf normal form with additive noise. *Comm. Math. Phys.*, 402:1807–1843, 2023. 5
- [17] W. Cowieson and L.-S. Young. SRB measures as zero-noise limits. *Ergodic Theory Dynam. Systems*, 25:1115–1138, 2005. 5
- [18] P. Duarte and S. Klein. *Lyapunov exponents of linear cocycles*. Atlantis Press, Paris, 2016. Continuity via large deviations. 5
- [19] G. Froyland, C. González-Tokman, and A. Quas. Hilbert space Lyapunov exponent stability. *Trans. Amer. Math. Soc.*, 372:2357–2388, 2019. 5
- [20] G. Froyland, C. González-Tokman, and A. Quas. Stochastic Stability of Lyapunov Exponents and Oseledets Splittings for Semi-invertible Matrix Cocycles. *Communications on Pure and Applied Mathematics*, 68:2052–2081, 2015. 5, 9, 10

- [21] H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.*, 108:377–428, 1963. [4](#)
- [22] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. Princeton University Press, Princeton, NJ, 1981. [22](#)
- [23] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960. [3](#)
- [24] I. Y. Goldsheĭd and G. A. Margulis. Lyapunov exponents of a product of random matrices. *Uspekhi Mat. Nauk*, 44:13–60, 1989. [4](#)
- [25] M. Goldstein and W. Schlag. Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. *Ann. of Math. (2)*, 154:155–203, 2001. [5](#)
- [26] A. Gorodetski and V. Kleptsyn. Non-stationary version of ergodic theorem for random dynamical systems, 2023. ArXiv 2305.05028. [4](#)
- [27] Y. Guivarc’h and A. Raugi. Products of random matrices: convergence theorems. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 31–54. Amer. Math. Soc., Providence, RI, 1986. [4](#)
- [28] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. [10](#)
- [29] P. Imkeller and C. Lederer. An explicit description of the Lyapunov exponents of the noisy damped harmonic oscillator. *Dynam. Stability Systems*, 14:385–405, 1999. [5](#)
- [30] R. Z. Khasminskiĭ. The averaging principle for parabolic and elliptic differential equations and Markov processes with small diffusion. *Teor. Veroyatnost. i Primenen.*, 8:3–25, 1963. [5](#)
- [31] Y. Kifer. *Random perturbations of dynamical systems*, volume 16 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1988. [5](#)
- [32] F. Ledrappier and L.-S. Young. Stability of Lyapunov exponents. *Ergodic Theory Dynam. Systems*, 11:469–484, 1991. [5](#)
- [33] Z. Lian and M. Stenlund. Positive Lyapunov exponent by a random perturbation. *Dyn. Syst.*, 27:239–252, 2012. [5](#)
- [34] A. M. Lyapunov. The general problem of the stability of motion. *Internat. J. Control*, 55:521–790, 1992. Translated by A. T. Fuller from Édouard Davaux’s French translation (1907) of the 1892 Russian original, With an editorial (historical introduction) by Fuller, a biography of Lyapunov by V. I. Smirnov, and the bibliography of Lyapunov’s works collected by J. F. Barrett, Lyapunov centenary issue. [3](#)
- [35] C. Matheus, M. Möller, and J.-C. Yoccoz. A criterion for the simplicity of the Lyapunov spectrum of square-tiled surfaces. *Invent. Math.*, 202:333–425, 2015. [4](#)
- [36] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Lyapunov exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968. [3](#)
- [37] M. Poletti and M. Viana. Simple Lyapunov spectrum for certain linear cocycles over partially hyperbolic maps. *Nonlinearity*, 32:238–284, 2019. [4](#)
- [38] M. S. Raghunathan. A proof of Oseledec’s multiplicative ergodic theorem. *Israel J. Math.*, 32:356–362, 1979. [7](#), [28](#)
- [39] M. Viana. *Lectures on Lyapunov exponents*, volume 145 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2014. [4](#)
- [40] M. Viana. (Dis)continuity of Lyapunov exponents. *Ergodic Theory Dynam. Systems*, 40:577–611, 2020. [5](#)
- [41] L.-S. Young. Random perturbations of matrix cocycles. *Ergodic Theory Dynam. Systems*, 6:627–637, 1986. [5](#)
- [42] L.-S. Young. Chaotic phenomena in three settings: large, noisy and out of equilibrium. *Nonlinearity*, 21:T245–T252, 2008. [5](#)
- [43] L.-S. Young. Mathematical theory of Lyapunov exponents. *J. Phys. A*, 46:254001, 17, 2013. [5](#)

SCHOOL OF MATHEMATICS AND PHYSICS, THE UNIVERSITY OF QUEENSLAND, BRISBANE, QLD 4072, AUSTRALIA

*Email address:* [j.atnip@uq.edu.au](mailto:j.atnip@uq.edu.au)

SCHOOL OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NEW SOUTH WALES, SYDNEY, NSW 2052, AUSTRALIA

*Email address:* [g.froyland@unsw.edu.au](mailto:g.froyland@unsw.edu.au)

SCHOOL OF MATHEMATICS AND PHYSICS, THE UNIVERSITY OF QUEENSLAND, BRISBANE, QLD 4072, AUSTRALIA

*Email address:* [cecilia.gt@uq.edu.au](mailto:cecilia.gt@uq.edu.au)

DEPARTMENT OF MATHEMATICS & STATISTICS, UNIVERSITY OF VICTORIA, VICTORIA, BC, CANADA V8W 2Y2

*Email address:* [aquas@uvic.ca](mailto:aquas@uvic.ca)