Commun. Math. Phys. 389, 121–151 (2022)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-021-04256-y

Communications in
Mathematical Physics

Lyapunov Exponents for Random Perturbations of Coupled Standard Maps

Alex Blumenthal[1], **Jinxin Xue**[2], **Yun Yang**[3]

[1] School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA.
E-mail: ablumenthal6@math.gatech.edu
[2] Department of Mathematics, Yau Mathematical Sciences Center, Tsinghua University, Jingzhai 310, Beijing 100084, China.
E-mail: jxue@tsinghua.edu.cn
[3] Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA.
E-mail: yunyang@vt.edu

Received: 29 April 2020 / Accepted: 18 October 2021
Published online: 16 November 2021 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract: In this paper, we give a quantitative estimate for the first N Lyapunov exponents for random perturbations of a natural class of $2N$-dimensional volume-preserving systems exhibiting strong hyperbolicity on a large but noninvariant subset of the phase space. Concrete models covered by our setting include systems of coupled standard maps, in both 'weak' and 'strong' coupling regimes.

1. Introduction

Many systems, including large classes of those of physical interest, exhibit strong sensitivity with respect to initial conditions. One way to mathematically describe this behavior is through Lyapunov exponents: for a smooth map $F : M \to M$ on a manifold M and a point $x \in M$, the Lyapunov exponents of F along the orbit $\{F^i x\}$ are the possible values of

$$\lambda(x, v) := \lim_{n\to\infty} \frac{1}{n} \log \|D_x F^n(v)\|,$$

when these limits exist, as v ranges over tangent directions in $T_x M$. If $\lambda(x, v) > 0$ for some $(x, v) \in TM$, then 'typical' nearby initial conditions diverge from $\{F^i x\}$ exponentially fast. For more discussion, see, e.g., [1,26,30].

Away from uniformly hyperbolic/Anosov settings, there can be extreme challenges in actually verifying that a given system, even a simple, low-dimensional one, admits positive Lyapunov exponents on an *observable* subset of phase space (i.e., positive-volume). This can be the case even when a positive Lyapunov exponent is "obvious" in numerical experiments.

Exemplifying these challenges is the Chirikov standard map family

$$(I, \theta) \mapsto \Phi_L(I, \theta) = (I + L \sin 2\pi\theta, \theta + I + L \sin 2\pi\theta),$$

where $L \in \mathbb{R}$ is a real parameter and both coordinates I, θ are taken modulo 1. Introduced by Chirikov [9], the standard map is a fundamental toy model describing the dynamics along 'stochastic boundary layers' formed by resonances in perturbed Hamiltonian systems, capturing the intricate interaction in phase space between 'regular' (elliptic) and 'chaotic' (seemingly stochastic) motion. For more discussion, see [8].

When $L \gg 1$, the mapping Φ_L exhibits strong hyperbolicity, except along an $O(L^{-1})$ neighborhood of the vertical lines $\{\theta = \pi/2, 3\pi/2\}$. The volume of these *critical strips*, where hyperbolicity fails, approaches 0 as $L \to \infty$, and so one might expect $\lambda(x, v) > 0$ for most x; this is corroborated by a wealth of numerical evidence. However, to prove this mathematically rigorously is a notorious open problem: to date, no-one has proved that Φ_L admits a positive Lyapunov exponent on a positive-area set for any value of L (equivalently, by Pesin's entropy formula, that Lebesgue measure has positive metric entropy for Φ_L). The primary challenge to overcome is *cone-twisting*, i.e., when previously expanded tangent directions can, upon the trajectory entering the critical set, be 'twisted' into strongly contracting directions. Estimating Lyapunov exponents for models of this kind amounts to an incredibly delicate cancellation problem between phases of growth and decay, all depending on the (time-varying) orientation of tangent directions.

These challenges are real, as evidenced by known results on the relative density of elliptic periodic orbits in phase space (e.g., [10]), which imply that even when L is taken arbitrarily large, there may be positive-area regions of phase space with zero Lyapunov exponents. In the positive direction, Gorodetski has shown that for a residual subset of $[L_0, \infty)$, $L_0 \gg 1$ taken sufficiently large, the set with a positive Lyapunov exponent has Hausdorff dimension 2 [14], although this is quite far from a positive-area set. We also mention the more recent work of Berger and Turaev, whose work on perturbations of elliptic islands for surface diffeomorphisms implies that Φ_L is C^∞ close to a volume-preserving mapping admitting a positive Lyapunov exponent on a positive-area set [3]. Lastly, we note that this brief discussion omits many works and indeed entire subfields related to the standard map, e.g., Schrödinger cocycles. We refer the readers to the introduction of [4] for more discussion.

The aim of this paper is to extend this program to a class of volume-preserving systems of arbitrarily high dimension exhibiting strong hyperbolicity on a large yet noninvariant subset of phase space. We aim to make estimates on all Lyapunov exponents, not just the 'top' exponent. Although our approach in this paper is inspired by that of [4], the higher-dimensional setting introduces several new layers of complexity which must be contended with, e.g., to estimate all Lyapunov exponents we must estimate the stationary statistics of N-dimensional planes in $T\mathbb{T}^{2N}$ (see Sects. 3 and 4).

The setting we introduce below includes systems of coupled Chirikov standard maps in a variety of coupling regimes: for $N > 1$ we consider N standard map oscillators $(I_i, \theta_i) \in \mathbb{T}^1 \times \mathbb{T}^1$, $i = 1, \cdots, N$, with a time evolution $(I_i, \theta_i) \mapsto (\bar{I}_i, \bar{\theta}_i)$ defined by

$$\begin{aligned} \bar{I}_i &= I_i + L \sin 2\pi\theta_i + \sum_{j \neq i} \mu_{ij} \sin 2\pi(\theta_j - \theta_i), \\ \bar{\theta}_i &= \theta_i + \bar{I}_i, \end{aligned} \tag{1}$$

where $\mathbb{T}^1$ is parametrized as $[0, 1)$, with all quantities above regarded "modulo 1". This is a completely integrable, uncoupled system when L, (μ_{ij}) are zero; in this paper, we will instead be interested in the so-called *anti-integrable* regime where $L \gg 1$ and the (μ_{ij}) can be potentially quite large. Note that the above mapping is symplectic, hence volume-preserving, iff $\mu_{ij} = \mu_{ji}$ for all i, j.

Coupled standard maps appear in the physical literature as toy models of Arnold diffusion [9] as well as the statistical properties of chaotic maps. These maps exhibit strong evidence of chaotic behavior in experiments, while mathematically rigorous verification of this chaotic behavior is hopelessly out of reach in the absence of noise. We refer the readers to, e.g., [5,18,21,27] and the references therein for more physics background and research on coupled standard maps.

Related to our work is that of Berger and Carrasco [2], which considered the Lyapunov exponents of a skew product of a hyperbolic CAT map with a Chirikov standard map. This was generalized recently by Carrasco [7] to estimate the Lyapunov exponents of arbitrarily many coupled standard maps. Applying a symbolic coding to the CAT map, one can view the models in [2,7] as random perturbations by discrete noise (by comparison, [4] and this paper both use the absolutely continuous noise). We note, however, that both the models and techniques in [2,7] are very different from the setting in the present manuscript. A key difference is that the perturbations in [2,7] are necessarily of order 1, and so the perturbed and unperturbed mappings have completely different dynamics even after 1 timestep. In contrast, the perturbations in [4] and the present paper, although absolutely continuous, may be extremely small.

The tractability of Lyapunov exponents for randomly perturbed systems is suggestive of the possibility of using computer-assisted techniques: when enough 'nondegenerate' noise is present in the random system, it is possible to estimate asymptotic quantities such as Lyapunov exponents rigorously by approximating the full random dynamics by, e.g., a finite-state Markov chain. To our knowledge, this connection has only been pursued recently: the work [12] uses computer-assisted proof (CAP) to study noise-induced order for a Poincaré section of a model of the famous Belousov-Zhabotinsky chemical reaction; and the work [6] uses CAP to estimate Lyapunov exponents for a stochastically perturbed Hopf system conditioned on remaining a bounded distance from the origin.

Farther from our work, there is a wealth of literature on Lyapunov exponents. We mention, for instance, Furstenberg's famous 1963 paper [11] on positivity of Lyapunov exponents for IID products of determinant 1 matrices, and the vigorous activity that followed extending this work to random products of matrices driven by more general processes (e.g., [15,25]) and to simplicity of the Lyapunov spectrum (e.g., [13]). We emphasize, though, that these works are qualitative and a priori provide no concrete estimates of Lyapunov exponents. We have only emphasized here works which directly address nonuniform hyperbolicity (in the presence of cone twisting) only in high-dimensional systems. For a broader discussion, we refer the reader to the introduction of [4].

1.1. Summary of results. We provide here an incomplete statement of the results in this paper emphasizing applications to coupled standard maps, deferring statements at full generality to the next section.

Maps from which we perturb Let $\mathbb{T} = \mathbb{T}^1$ denote the circle, parametrized as $[0, 1) \cong \mathbb{R}/\mathbb{Z}$. For $N \geq 1$ we consider dynamics on the torus $\mathbb{T}^{2N} \cong \mathbb{R}^{2N}/\mathbb{Z}^{2N}$, which we regard with the flat metric coming from $\mathbb{R}^{2N}$. We consider mappings $F : \mathbb{T}^{2N} \to \mathbb{T}^{2N}$ of the form

$$F(x, y) = (f(x) - y, x) \tag{2}$$

where $x = (x_1, \cdots, x_N) \in \mathbb{T}^N$, $y = (y_1, \cdots, y_N) \in \mathbb{T}^N$ and $f : \mathbb{T}^N \to \mathbb{T}^N$ is a given smooth mapping. Below and throughout, all addition in $\mathbb{T}^N$ is carried out 'modulo 1' in each coordinate. Note that F is always invertible and volume-preserving, irrespective of the mapping f.

This class is a natural setting for high-dimensional volume-preserving systems with strong expansion, and includes the coupled standard map systems from (1): the change of coordinates $x_i = \theta_i$, $y_i = \theta_i - I_i \pmod 1$ conjugates the mapping $(I, \theta) \mapsto (\bar{I}, \bar{\theta})$ to F as in (2) with

$$f(x) = \Big(2x_i + L \sin 2\pi x_i + \sum_{j \neq i} \mu_{ij} \sin 2\pi (x_j - x_i)\Big)_{i=1}^N. \tag{3}$$

Let us comment briefly on the hyperbolicity of the mappings F. Throughout we identify $T\mathbb{T}^{2N} \cong \mathbb{T}^{2N} \times \mathbb{R}^{2N}$ and write $\mathbb{R}^{2N} = \mathbb{R}^x \oplus \mathbb{R}^y$, where $\mathbb{R}^x = \text{Span}\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_N}\}$ and $\mathbb{R}^y = \text{Span}\{\frac{\partial}{\partial y_1}, \cdots, \frac{\partial}{\partial y_N}\}$, each of which is parametrized by $\mathbb{R}^N$. For $\alpha > 0$, we define

$$\mathcal{C}^x_\alpha := \{(u, v) \in \mathbb{R}^{2N} : \|v\| \leq \alpha \|u\|\}$$

of vectors α-close to the 'horizontal' space $\mathbb{R}^x$. We show that for all L sufficiently large and under rather general conditions on the coupling coefficients μ_{ij} that f as above expands N-dimensional volumes to order $\approx L^N$ on a large (but noninvariant) subset of phase space. As a consequence, when L is taken large enough, we have that

$$D_{(x,y)} F \mathcal{C}^x_{1/10} \subset \mathcal{C}^x_{1/20},$$

i.e., the cone $\mathcal{C}^x_{1/10}$ is mapped well inside itself. We note, however, that cone invariance can fail badly in some parts of phase space: vectors in $\mathcal{C}^x_{1/10}$ can be rotated to a vicinity of the 'vertical' space $\mathbb{R}^y$.

Noise model We next randomize the system F_L. Fix a probability space $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ and let $\omega \mapsto R_\omega \in C^2_{\text{vol}}(\mathbb{T}^{2N}, \mathbb{T}^{2N})$ be a measurable assignment to each $\omega \in \Omega_0$ of a C^2, volume-preserving diffeomorphism $R_\omega : \mathbb{T}^{2N} \to \mathbb{T}^{2N}$, to be interpreted as the 'noise' applied to the dynamics at each timestep. Define $\Omega = \Omega_0^{\otimes \mathbb{N}}$, $\mathcal{F} = \mathcal{F}_0^{\otimes \mathbb{N}}$, $\mathbb{P} = \mathbb{P}_0^{\otimes \mathbb{N}}$ and let $\theta : \Omega \to \Omega$ be the leftward shift (which is automatically invariant and ergodic for $\mathbb{P}$). Elements $\underline{\omega} \in \Omega$ are written $\underline{\omega} = (\omega_1, \omega_2, \cdots)$ for $\omega_i \in \Omega_0$, $i \geq 1$. In this paper, we consider random compositions

$$F^n_{\underline{\omega}} = F_{\omega_n} \circ \cdots \circ F_{\omega_1}, \quad n \geq 1, \quad \underline{\omega} = (\omega_i)_{i \in \mathbb{N}} \in \Omega$$

of the (IID) random maps $F_\omega = R_\omega \circ F$, $\omega \in \Omega_0$, where F is as in (2). We also consider the Markov chain $(Z_n) = (X_n, Y_n)$ on $\mathbb{T}^{2N}$ defined by setting $(X_n, Y_n) = F^n_{\underline{\omega}}(X_0, Y_0)$ for fixed initial $Z_0 = (X_0, Y_0) \in \mathbb{T}^{2N}$.

For the noise model R_ω we shall impose three sets of assumptions. The first two are as follows:

(E) The only stationary measure for the Markov chain (X_n, Y_n) is $\text{Leb}_{\mathbb{T}^{2N}}$, the Lebesgue measure on $\mathbb{T}^{2N}$.

(C) The noise model R_ω satisfies $D_z R_\omega(\mathcal{C}^x_{1/20}) \subset \mathcal{C}^x_{1/10}$ and $\|D_z R_\omega\|, \|(D_z R_\omega)^{-1}\| \leq 2$ for all $z = (x, y) \in \mathbb{T}^{2N}$ with probability 1.

Assumption (E) (short for *ergodicity*) is natural in the context of random systems, and is satisfied by a wide variety of noise models R_ω, e.g., if for all fixed $z \in \mathbb{T}^{2N}$, the law of $R_\omega(z)$ has a density q_z which varies L^1-continuously in z– see Lemma 9 for details. Condition (E) implies that the Lyapunov exponents

$$\lambda_i := \lim_n \frac{1}{n} \log \sigma_i(D_z F^n_{\underline{\omega}}) \tag{4}$$

exist with probability 1 and are deterministic and constant over all $z \in \mathbb{T}^{2N}$ (Theorem 11). Here, $\sigma_i(\cdot)$ refers to the i-th singular value ("Appendix A"). See Sect. 3.1 for more background on Markov chains and conditions for existence of the limit (4). Condition (C) is reasonable in the context of this paper, and ensures the noise model R_ω does not introduce any additional cone-twisting to the resulting random system.

The third, assumption (ND) (short for *nondegeneracy*) is somewhat more technical, and has to do with the way that N-dimensional planes are randomized by the noise model R_ω. For additional discussion and motivation, see the Additional Comments at the end of this section. Below, $\mathrm{Gr}_N(\mathbb{R}^{2N})$ denotes the Grassmanian, the manifold of N-dimensional subspaces of $\mathbb{R}^{2N}$ (for more on the Grassmanian, see Sect. 5.1).

(ND) For any fixed $z \in \mathbb{T}^{2N}$ and N-dimensional subspace $E \subset \mathbb{R}^{2N}$ (viewed as $\subset T_z\mathbb{T}^{2N}$), we have that the measure

$$\hat{Q}_{(z,E)}(\cdot) := \mathbb{P}_0((R_\omega z,\, D_z R_\omega(E)) \in \cdot) \tag{5}$$

on $\mathrm{Gr}_N(\mathbb{T}^{2N}) := \mathbb{T}^{2N} \times \mathrm{Gr}_N(\mathbb{R}^{2N})$ is absolutely continuous w.r.t. Lebesgue measure $\mathfrak{m} := \mathrm{Leb}_{\mathbb{T}^{2N}} \times \mathrm{Leb}_{\mathrm{Gr}_N(\mathbb{R}^{2N})}$, and that the resulting density $\hat{q}_{(z,E)} := \frac{d\hat{Q}_{(z,E)}}{d\mathfrak{m}}$ satisfies

$$\|\hat{q}_{(z,E)}\|_{L^\infty} \le M$$

where $M > 0$ is a constant independent of (z, E).

Above, the measure $\mathrm{Leb}_{\mathrm{Gr}_N(\mathbb{R}^{2N})}$ is the Riemannian volume on $\mathrm{Gr}_N(\mathbb{R}^{2N})$; see Sect. 5.1 for details.

The value M itself has the connotation of an 'inverse noise amplitude': if R_ω is typically C^1 close to the identity, then the value M must be large. An explicit noise model satisfying (E), (C), and (ND) is constructed in Sect. 2.

Results for coupled standard maps. The following two results estimate all Lyapunov exponents (λ_i) as in (4) of the random coupled standard maps introduced above.

Theorem 1. *Let $\alpha, \delta \in (0, 1)$. Let $N \ge 2$ and consider the coupled standard map F as in (2) with*

$$f(x) = \left(2x_i + L \sin 2\pi x_i + \sum_{j \ne i} \mu_{ij} \sin 2\pi (x_j - x_i) \right)_{i=1}^{N},$$

where the coefficients μ_{ij} are fixed and L is sufficiently large depending on (μ_{ij}). Assume that the randomizations R_ω satisfy (E), (C) and (ND), and that the noise parameter M satisfies

$$M \le L^{\frac{\delta}{12} L^{1-\delta}} \tag{6}$$

Then, the Lyapunov exponents (λ_i) *of this random composition satisfy*

$$\lambda_1 \geq \ldots \geq \lambda_N > 0 > \lambda_{N+1} \geq \ldots \geq \lambda_{2N}, \quad \textit{and} \quad \min\{|\lambda_i|\} \geq \alpha \log L. \tag{7}$$

Remark 2. Note that the parameter L can be freely taken arbitrarily large. In particular, given an R_ω with a particular value of M, one is free to choose L large enough so that (6) is satisfied.

The setting of Theorem 1 can be thought of as describing a kind of 'weak' to 'moderate' coupling regime: the strength of the hyperbolicity L of each individual oscillator overshadows the coupling amplitude $\max_{ij} |\mu_{ij}|$. The following applies in a regime when the strength of the coupling matches that of the individual oscillators.

Theorem 3. *Let* $\alpha, \delta \in (0, 1)$. *Let* $N = 2$ *and consider the coupled standard map* F *as in* (2) *with*

$$f(x_1, x_2) = \begin{pmatrix} 2x_1 + L \sin 2\pi x_1 + L \sin 2\pi (x_2 - x_1) \\ 2x_2 + L \sin 2\pi x_2 + L \sin 2\pi (x_1 - x_2) \end{pmatrix}.$$

Let L *be sufficiently large. Assume* R_ω *satisfies* (E), (C) *and* (ND) *as well as*

$$M \leq L^{\frac{1}{2}\delta L^{1-\delta}}.$$

Then, the Lyapunov exponents (λ_i) *satisfy the estimate* (7).

Theorems are consequences of our general main result Theorem 5, presented below in full generality in Sect. 6.

Remark 4. We emphasize the order of quantifiers in our results: throughout, we fix the dimension N and choose L sufficiently large depending on N, but not vice versa. It would, though, be of interest to fix L, take the 'hydrodynamic limit' $N \to \infty$ and study the resulting Lyapunov spectrum. Limits of this kind provide toy models for, e.g., gases of particles in the hydrodynamic limit; see, e.g., [28,29]. However, this is beyond the scope of the present paper, since our analysis here does not take into account quantitative dependence in N.

Additional comments We end this section with some discussion on the relationship between this paper and the previous work [4], as well as some discussion on the assumption (ND) and how it could potentially be relaxed.

The paper [4] estimates Lyapunov exponents for random standard maps ($N = 1$) on $\mathbb{T}^2$ subjected to IID 'additive' noise of the form $R_\omega(x, y) = (x + \omega, y)$ with ω uniformly distributed in $[-\epsilon, \epsilon]$ for some $\epsilon > 0$ small. The rough idea there is to relate Lyapunov exponents to stationary statistics of the Markov chain

$$(Z_n, V_n) = \left(F^n_{\underline{\omega}}(Z_0), \frac{D_{Z_0} F^n_{\underline{\omega}}(V_0)}{|D_{Z_0} F^n_{\underline{\omega}}(V_0)|} \right)$$

on the unit tangent bundle $S\mathbb{T}^2 \cong \mathbb{T}^2 \times S^1$. The key idea in that paper is that the presence of noise implies that *stationary measures are absolutely continuous, and that stationary mass cannot concentrate too much parallel to* $\mathbb{R}^y$, where strong contraction can occur. This was used directly to show that the vast majority of stationary mass was nearly parallel to $\mathbb{R}^x$, where the strong expansion ensures a large Lyapunov exponent. To check

absolute continuity of stationary measures for (Z_n, V_n), it was shown that the noise in the base 'propagates' to tangent directions in three iterates of the random system (one for each dimension of the projective bundle $P\mathbb{T}^2 \cong \mathbb{T}^2 \times P^1$), in the sense that for fixed (Z_0, V_0), the law of (Z_3, V_3) is absolutely continuous on $\mathbb{T}^2 \times S^1$.

The goal of this paper is demonstrate that the basic mechanism introduced in [4] generalizes to higher-dimensional systems, e.g., coupled standard maps, and to provide estimates on all Lyapunov exponents, not just λ_1. To this end, for $N > 1$ we study an analogous Markov chain (Z_n, E_n) on the Grassmanian bundle $\mathrm{Gr}_N\, \mathbb{T}^{2N} \cong \mathbb{T}^{2N} \times \mathrm{Gr}_N(\mathbb{R}^{2N})$ of N-dimensional subspaces of $T\mathbb{T}^{2N}$ and relate stationary measures for (Z_n, E_n) to the sum of the first N Lyapunov exponents (with multiplicity). This brings us to the purpose of condition (ND), which plays the same role that the 3-step noise propagation played in [4]: this condition ensures that stationary measures for (Z_n, E_n) are absolutely continuous and cannot concentrate too much in any one place.

On the other hand, condition (ND) is *not* satisfied by 'additive' noise models of the form $R_\omega(x, y) = (x, y)+\omega$ where ω is some random vector: condition (ND) *requires* that the noise model R_ω induces some 'twisting' of tangent directions. That being said, the methods in this paper are entirely compatible with the following hypothetical scheme, more in line with [4]: start with additive noise $R_\omega(x, y) = (x, y)+\omega$ and show that there is some k such that for all fixed $(Z_0, E_0) \in \mathrm{Gr}_N(\mathbb{T}^{2N})$, we have that the law of (Z_k, E_k) is absolutely continuous on $\mathrm{Gr}_N(\mathbb{T}^{2N})$. Although highly plausible, actually carrying this out appears to be quite challenging for general N: the Grassmanian $\mathrm{Gr}_N(\mathbb{R}^{2N})$ has dimension N^2, and so *even if the random vector ω is absolutely continuous along all $2N$ coordinates in $\mathbb{T}^{2N}$*, it would take a minimum of $k \approx N/4$ iterates for the noise to propagate to all directions in $\mathrm{Gr}_N(\mathbb{R}^{2N})$. Already in the simple case $N = 2$ the computations involved are quite involved, requiring separate estimates across multiple charts in $\mathrm{Gr}_2(\mathbb{R}^4)$. So, for the sake of unity of focus and brevity, we have opted to use noise models satisfying condition (ND), as this framework highlights the essential features of the proof. Extensions to additive noise models are left to future work.

2. General Framework and Full Statement of Results; An Explicit Noise Model

We begin by providing the full general framework we work in and the main result from which Theorems 1 and 3 are derived.

Let us start with a full description of the deterministic maps from which we perturb. We consider one-parameter mappings of the form $F = F_L : \mathbb{T}^{2N} \circlearrowleft$, $F_L(x, y) = (f_L(x) - y, x)$, $x, y \in \mathbb{T}^N$, $L > 0$, where the family $f_L : \mathbb{T}^N \to \mathbb{T}^N$ is assumed to satisfy the following:

(F1) There exists $C_0 > 0$ such that $\|D_x f_L\| \le C_0 L$ for all $x \in \mathbb{T}^N$ and $L \ge 1$; and
(F2) For any $\beta \in (0, 1)$, there exist $C_\beta, c_\beta, L_\beta > 0$ so that for all $L \ge L_\beta$, we have that

$$B_\beta = \{x \in \mathbb{T}^N : |\det D_x f_L| \le L^{N-(1-\beta)}\} \subset \mathbb{T}^N$$

has Lebesgue measure $\le C_\beta L^{-c_\beta}$.

Condition (F2) implies that for a large Lebesgue-measure set $x \in \mathbb{T}^N$, we have $|\det D_x f_L|$ is of order L^N, and since $\|D_x f_L\|$ is no larger than order L by (F1), we see that $\|D_x f_L(v)\|$ must be approximately $L\|v\|$ at all $v \in T_x\mathbb{T}^N \cong \mathbb{R}^N$ and for all such x. As a result, the corresponding mappings F_L expand all directions roughly parallel to $\mathbb{R}^x$ by a factor of L. For more details, see Sect. 3.3 below.

We can now state our main abstract result.

Theorem 5 (Main Theorem). *Assume that the one parameter family $f_L : \mathbb{T}^N \to \mathbb{T}^N$ satisfies conditions (F1) and (F2). Fix $\alpha, \beta \in (0,1)$ and $\delta \in (0, c_\beta)$, where c_β is as in (F2). Let $L \geq L_\beta$ be sufficiently large in terms of these parameters. Lastly, let R_ω be any noise model satisfying (E), (C) and (ND), and assume that the noise parameter M as in condition (ND) satisfies*

$$M \leq L^{\frac{1}{2}\beta L^{c_\beta - \delta}}.$$

Then, the Lyapunov exponents (λ_i) of $F^n_{\underline{\omega}} = F_{\omega_n} \circ \cdots \circ F_{\omega_1}$, $F_\omega := R_\omega \circ F$, satisfy

$$\lambda_1 \geq \ldots \geq \lambda_N > 0 > \lambda_{N+1} \geq \ldots \geq \lambda_{2N}, \quad \textit{and} \quad \min\{|\lambda_i|\} \geq \alpha \log L.$$

Once Theorem 5 is proved, to prove Theorems 1 and 3 it will suffice to check that the corresponding one-parameter families f_L satisfy conditions (F1) and (F2); this is carried out in Sect. 4.

We note that conditions (F1) and (F2) are not limited just to coupled standard maps. Indeed, for general $\psi, \varphi : \mathbb{T}^N \to \mathbb{R}^N$, we can consider the one-parameter families

$$f_L := L\psi + \varphi.$$

Condition (F1) is evident for arbitrary ψ, φ, while we are able to show (Lemma 25) that condition (F2) holds under the transversality-type condition

$$\{\det D_x \psi = 0\} \cap \{\nabla \det D_x \psi = 0\} = \emptyset. \tag{8}$$

This is clearly a C^2 open condition; for $N = 2$, we show (Proposition 27) that (8) in fact holds for a C^2-generic set of ψ. On the other hand, (8) *fails* for ψ in the setting of Theorem 1; for this reason we instead check (F1), (F2) directly for this model in Sect. 6.

An explicit noise model. Below we sketch the construction of a noise model $R_\omega : \mathbb{T}^{2N} \to \mathbb{T}^{2N}$ satisfying properties (E), (C) and (ND). Full proofs are deferred to "Appendix C". We write $d = 2N$ for brevity. Below, we write $\mathcal{O}(d)$ for the space of orthogonal $d \times d$ matrices. Let $\mathrm{Skew}(d) = T_{\mathrm{Id}}\mathcal{O}(d)$ denote the Lie algebra of skew-symmetric $d \times d$ matrices, and recall that the matrix exponential e^U of any $U \in \mathrm{Skew}(d)$ is an orthogonal matrix.

Let $\{z_i\}_{i=1}^K$ be a collection of points with the property that

$$\mathbb{T}^d = \cup_i B_{1/20}(z_i), \tag{9}$$

i.e., the balls $\{B_{1/20}(z_i)\}$ of radius $1/20$ cover $\mathbb{T}^d$. For brevity, we write $d = 2N$ below.

Let $\psi : [0,\infty) \to [0,1]$ be a C^∞ bump function such that $\psi|_{[0,1/10]} \equiv 1$ and $\psi|_{[1/5,\infty)} \equiv 0$. For $z \in \mathbb{T}^d$, define $\Delta_i(z) \in \mathbb{R}^d$ to be the unique vector[1] in $[-1/2, 1/2)^d$ such that $\Delta_i(z) = z - z_i$ modulo 1 (recall that $\mathbb{T}^d$ is parametrized by $[0,1)^d$).

For $U \in \mathrm{Skew}(d)$, define

$$\Phi_U^{(i)} : \mathbb{T}^d \to \mathbb{T}^d, \quad \Phi_U^{(i)}(z) = z_i + \exp(\psi(d(z,z_i))U)\Delta_i(z)$$

[1] In other words, for each $1 \leq j \leq d$, we set the j-th component $(\Delta_i(z))_j$ to be $\{(z - z_i)_j\} - 1/2$, where here for $\alpha \in \mathbb{R}$ we write $\{\alpha\} = \alpha - \lfloor \alpha \rfloor \in [0,1)$ for the fractional part of α, and $(z - z_i)_j$ is the j-th coordinate of $z - z_i$.

This yields a defined, continuous mapping of $\mathbb{T}^d$ into itself, which rigidly rotates the shells of constant distance from z_i by some fractional power of $e^U \in \mathcal{O}(d)$. With this picture in mind, it is straightforward to show that $\Phi_U^{(i)}$ preserves volumes (e.g., using polar coordinates we write the volume form as $d\text{vol} = r^{n-1} dr d\Omega$ where Ω is the volume form on the unit sphere. It is clear by definition that $\Phi_U^{(i)}$ preserves $d\text{vol}$ since both dr and $d\Omega$ are preserved).

Given $h \in \mathbb{R}^d$, define $T_h : \mathbb{T}^d \to \mathbb{T}^d$ to be the translation $T_h z = z + h$. Given $U^{(1)}, \cdots, U^{(K)} \in \text{Skew}(d)$ and $v \in \mathbb{R}^d$, we define

$$R_{(v;\{U^{(i)}\})} := T_v \circ \Phi_{U^{(K)}}^{(K)} \circ \cdots \circ \Phi_{U^{(1)}}^{(1)}.$$

With $\Omega_0 = \mathbb{R}^d \times \text{Skew}(d)^K$, we see that a Borel probability $\mathbb{P}_0$ on $\Omega_0 \ni \omega \mapsto R_\omega$ yields a 'random' volume-preserving diffeomorphism of $\mathbb{T}^d$. This diffeomorphism is a composition of a sequence of 'twists' of $\mathbb{T}^d$ by rigid rotations, post-composed by a rigid translation in $\mathbb{T}^d$.

We now provide sufficient conditions on the law $\mathbb{P}_0$ for the noise model R_ω to satisfy conditions (E), (C) and (ND). Below, we regard Ω_0 as a copy of $\mathbb{R}^{d+Kd(d-1)/2}$ equipped with Lebesgue measure Λ and the standard Euclidean norm.

Proposition 6. *There exists $c = c_{K,d} > 0$ sufficiently small so that the following holds. Let $\mathbb{P}_0$ be any Borel probability measure on Ω_0 such that*

(i) $\text{Supp}(\mathbb{P}_0)$ *is contained in the ball of radius c centered at the origin;*
(ii) $\mathbb{P}_0 \ll \Lambda$ *with $\|d\mathbb{P}_0/d\Lambda\|_{L^\infty} < \infty$; and*
(iii) $\exists \zeta > 0$ *such that $d\mathbb{P}_0/d\Lambda > 0$ on the ball of radius ζ centered at the origin.*

Then, R_ω equipped with $\mathbb{P}_0$ satisfies conditions (E), (C) and (ND).

To summarize these conditions: item (i) ensures that the R_ω are not too far from the identity in the C^1 norm, hence (C) holds. Item (ii) is used to affirm condition (ND), that N-dimensional planes are randomized by R_ω. Finally, item (iii) is used to check ergodicity of $\text{Leb}_{\mathbb{T}^d}$ as in condition (E). See "Appendix C" for a full proof.

Organization of the paper In Sect. 3, we give some preliminary results on the Markov chain on $\mathbb{T}^{2N}$ and on $\text{Gr}_N(\mathbb{T}^{2N})$, while in Sects. 4 and 5 we prove Theorem 5. Theorems 1 and 3 are proved in Sect. 6. Sufficient conditions for (F2) and genericity results, as well as applications to coupled standard maps, are worked out in Sect. 6. Included in "Appendix A" is a version of the standard singular value decomposition used in this paper. In "Appendix B", we give the proofs of Lemma 19 and 21 on Grassmannian. In "Appendix C", we prove that the noise model R_ω constructed above satisfies conditions (E), (C) and (ND).

3. Preliminaries

Here we provide some preliminaries used in the rest of the paper. Section 3.1 recalls elements of random dynamics and formulations in terms of Markov chains, while Sect. 3.2 relates stationary measures of these Markov chains to Lyapunov exponents, crucial to the approach we take in this paper. In Sect. 3.3 we provide some preliminary estimates characterizing hyperbolic behavior of our random maps F_ω.

3.1. Background on random dynamics and Markov chains. Let us recall the definition and basic properties of a Markov chain. Let $(S, \mathcal{S})$ be a measurable space, and let (X_n) be a sequence of S-valued random variables, i.e., there is some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a sequence of measurable mappings $X_n : \Omega \to S$. We say that (X_n) is a (time-homogeneous) *Markov chain* if for all $n \geq 1$ and sets $A_0, \cdots, A_n \in \mathcal{S})$, we have that

$$\mathbb{P}(X_n \in A_n | X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) = \mathbb{P}(X_n \in A_n | X_{n-1} \in A_{n-1}).$$

Under general conditions on $(S, \mathcal{S})$ (e.g., S is a complete metric space and $\mathcal{S} = \mathrm{Bor}(S)$), the law of X_n conditioned on X_{n-1} is proscribed by a *transition kernel* P, i.e., an assignment to each $s \in S$ of a probability measure $P(s, \cdot)$ on $(S, \mathcal{S})$ such that for all $n \geq 1$ and $A \in \mathcal{S}$,

$$\mathbb{P}(X_n \in A | X_{n-1}) = P(X_{n-1}, A) \quad \text{with probability 1,}$$

where above $\mathbb{P}(\cdot | X_{n-1})$ refers to the probability of an event conditioned on the σ-algebra generated by X_{n-1}.

In this paper we consider Markov chains arising from *random dynamical systems*. Assume (as will always be the case for us) that S is a compact metric space and $\mathcal{S} = \mathrm{Bor}(S)$. Let $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ be a probability space, and let $G : \Omega_0 \to C(S, S)$ be a measurable mapping into the space $C(S, S)$ of continuous self-maps of S (equipped with the uniform norm and corresponding Borel σ-algebra). As in the previous section, form $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_0, \mathcal{F}_0, \mathbb{P}_0)^{\otimes \mathbb{N}}$ and write elements $\underline{\omega} \in \Omega$ as $\underline{\omega} = (\omega_1, \omega_2, \cdots)$, $\omega_i \in \Omega_0$, noting that $\{G_{\omega_i}\}_i$ is an IID sequence of $C(S, S)$-valued random variables. For fixed initial $X_0 \in S$, this gives rise to the Markov chain $X_n = G_{\omega_n} \circ \cdots \circ G_{\omega_1}(X_0)$ with transition kernel

$$P(s, A) = \mathbb{P}(s \in G_{\omega}^{-1}(A)) .$$

Recall that a probability measure μ on $(S, \mathcal{S})$ is called *stationary* if

$$\mu(A) = \int P(s, A) d\mu(s)$$

for all $A \in \mathcal{S}$. Note that if μ is stationary, then for all $\varphi : S \to \mathbb{R}$ bounded and measurable and for all $n \geq 1$, we have that

$$\int \varphi(s) d\mu(s) = \int (\mathbb{E}_s \varphi(X_n)) \, d\mu(s),$$

where $\mathbb{E}_s$ denotes the expectation w.r.t. $\mathbb{P}$ conditioned on $X_0 = s$. When $X_n = G_{\omega_n} \circ \cdots \circ G_{\omega_1}(X_0)$ as above, we can view $X_n = X_n(s; \underline{\omega})$ as a function of the random sample $\underline{\omega} \in \Omega$ and the initial condition $X_0 = s$; with these conventions, for φ bounded-measurable and $n \geq 1$ we have by Fubini's Theorem that

$$\int \varphi(s) d\mu(s) = \int_{\Omega} \left(\int_S \varphi(X_n) d\mu(s) \right) d\mathbb{P}(\underline{\omega}). \tag{10}$$

Lastly, we recall that the random dynamics induced by G_{ω} can be thought of as a "deterministic" dynamical system via the skew-product construction, which we recall

below. Let $\theta : \Omega \to \Omega$ denoting the leftward shift, and define the mapping $\tau : \Omega \times S \circlearrowleft$ by

$$\tau(\omega, s) = (\theta\omega, G_\omega s).$$

We record below a well-known 'glossary' between invariant measures of τ and stationary measures of (X_n). Recall that a stationary measure μ for (X_n) is called ergodic any P-invariant set $A \subset S$ satisfies $\mu(A) = 0$ or 1. Here, a measurable set A is called P-invariant if $P(s, A) = 1$ for all $s \in A$.

Proposition 7. *Assume S is a compact metric space. Let μ be any Borel probability measure on S.*

(a) *The measure $\mathbb{P} \times \mu$ is τ-invariant iff μ is a stationary measure for (X_n).*
(b) *(Ohno's Theorem) The measure $\mathbb{P} \times \mu$ is τ-ergodic iff μ is an ergodic stationary measure.*

Lastly, we recall the following consequence of the Krein-Milman Theorem applied to the (convex, weak* compact) set of stationary measures for (X_n).

Corollary 8. *Assume $(S, \mathcal{S})$ is a compact metric space. Then, the set of ergodic stationary measures for (X_n) is nonempty.*

In particular, if (X_n) admits a unique stationary measure, then it must be ergodic.

3.2. Stationary measures and Lyapunov exponents. We now specialize to the random dynamics of interest in this paper: let $(\Omega, \mathcal{F}, \mathbb{P})$ be as in Sect. 1 corresponding to a noise model R_ω, and let $F : \mathbb{T}^{2N} \to \mathbb{T}^{2N}$, $F(x, y) = (f(x) - y, x)$. Throughout, we write $F_\omega = R_\omega \circ F$ and $F^n_{\underline{\omega}} = F_{\omega_n} \circ \cdots \circ F_{\omega_1}$. Lastly, for fixed initial $Z_0 = (X_0, Y_0) \in \mathbb{T}^{2N}$, we write

$$Z_n = (X_n, Y_n) = F^n_{\underline{\omega}}(X_0, Y_0)$$

for the Markov chain generated by the system $F^n_{\underline{\omega}}$ and

$$P(z, A) := \mathbb{P}(z \in F_\omega^{-1}(A))$$

for the corresponding transition kernel. Below and throughout, $\tau : \Omega \times \mathbb{T}^{2N} \circlearrowleft$ denotes the corresponding skew product $\tau(\underline{\omega}, z) := (\theta\underline{\omega}, F_{\omega_1}(z))$.

3.2.1. A sufficient condition for (E): $\mathrm{Leb}_{\mathbb{T}^{2N}}$ *is ergodic* Since R_ω, F are always volume-preserving, we see that the volume $\mathrm{Leb}_{\mathbb{T}^{2N}}$ is always a stationary measure for (Z_n). The following is a useful criterion for ensuring that the ergodicity condition (E) holds.

Lemma 9. *Assume the following:*

(a) *For all $z \in \mathbb{T}^{2N}$, the law Q_z of $R_\omega(z)$ is absolutely continuous w.r.t.* $\mathrm{Leb}_{\mathbb{T}^{2N}}$*; and*
(b) *We have that $z \mapsto q_z := \frac{dQ_z}{d\,\mathrm{Leb}_{\mathbb{T}^{2N}}}$ is L^1 continuous, i.e., for any $z \in \mathbb{T}^{2N}, \eta > 0$ there exists $\delta > 0$ such that $d(z, z') < \delta$ implies $\|q_z - q_{z'}\|_{L^1(\mathrm{Leb}_{\mathbb{T}^{2N}})} < \eta$.*

Then, condition (E) holds, i.e., Lebesgue measure $\mathrm{Leb}_{\mathbb{T}^{2N}}$ *is the unique, hence ergodic, stationary measure for the Markov chain (Z_n).*

Proof. The result is similar to Lemma 5 of [4], but the proof we present here is somewhat more general. To start, by Corollary 8, we may fix some ergodic stationary measure m for (Z_n). Define

$$G = \{z \in \mathbb{T}^{2N} : \lim_n \frac{1}{n} \sum_{k=0}^{n-1} \varphi(F^n_{\underline{\omega}}(z)) = \int \varphi dm \text{ for all } \varphi : \mathbb{T}^{2N} \to \mathbb{R} \text{ continuous and for } \mathbb{P} - \text{ a.e. } \underline{\omega}\},$$

noting that G is Borel since the space $C(\mathbb{T}^{2N})$ of continuous functions on $\mathbb{T}^{2N}$ with the uniform norm is separable. By the Birkhoff ergodic theorem, applied to the skew product $\tau : \Omega \times \mathbb{T}^{2N} \circlearrowleft$ with ergodic measure $\mathbb{P} \times m$, we see that $m(G) = 1$. Note that $P(z, G) = 1$ must hold for all $z \in G$.

The following is useful for checking membership in G.

Claim 10. $P(z, G) > 0$ implies that $z \in G$

Proof of Claim. For fixed $z \in \mathbb{T}^{2N}$ and $\varphi : \mathbb{T}^{2N} \to \mathbb{R}$ continuous, we have that the event $E_{z,\varphi} := \{\underline{\omega} \in \Omega : \lim_n \frac{1}{n} \sum_0^{n-1} \varphi \circ F^i_{\underline{\omega}}(z) = \int \varphi dm\}$ is a tail event in Ω, hence $\mathbb{P}(E_{z,\varphi}) = 0$ or 1 by the Kolmogorov 0-1 Law. Since $C(\mathbb{T}^{2N})$ is separable, we see that $E_z = \{\underline{\omega} \in \Omega : \lim_n \frac{1}{n} \sum_0^{n-1} \varphi \circ F^i_{\underline{\omega}}(z) = \int \varphi dm \text{ for all } \varphi \in C(\mathbb{T}^{2N})\}$ is also a tail event. The proof is complete on noting that $\mathbb{P}(E_z) > 0$ if $P(z, G) > 0$. □

We now set about checking that G is open and closed, hence $G = \mathbb{T}^{2N}$. This precludes the existence of ergodic stationary measures for (Z_n) distinct from m, hence $m = \text{Leb}_{T^{2N}}$ follows.

For openness, observe that (b) implies $z \mapsto P(z, G)$ is uniformly continuous, hence $\exists \delta > 0$ such that for all $z, z' \in \mathbb{T}^{2N}$, $d(z, z') < \delta$ implies $|P(z, G) - P(z', G)| < 1/2$. With $z \in G$ fixed, we conclude that $P(z', G) > 1/2$ for all $z' \in B_\delta(z)$. By the Claim, we conclude $B_\delta(z) \subset G$, hence G is open. For closedness, let $z \in \mathbb{T}^{2N}$ be the limit of a sequence $\{z_n\} \subset G$. With $\delta > 0$ as in the previous paragraph and N large enough so that $d(z_n, z) < \delta$ for all $n < N$, we conclude that $P(z, G) > 1/2$, hence $z \in G$ by the Claim. We conclude G is closed, hence $G = \mathbb{T}^{2N}$. □

3.2.2. Lyapunov exponents and stationary measures We now turn our attention to the relation between Lyapunov exponents and stationary measures. The following is an abridged version of the multiplicative ergodic theorem (MET) suitable for our purposes; for more on the MET and its consequences in smooth ergodic theory, see, e.g. [20].

Below, for a matrix A we write $\sigma_1(A), \sigma_2(A), \cdots$ for the singular values of A. See "Appendix A" for more discussion.

Theorem 11 (Multiplicative Ergodic Theorem). *Assume condition* (E) *holds. Then, there is a* τ*-invariant set* $\Gamma \subset \Omega \times \mathbb{T}^{2N}$ *of full* $\mathbb{P} \times \text{Leb}_{\mathbb{T}^{2N}}$*-measure such that the following holds.*

(a) *For each* $i = 1, \cdots, 2N$*, there is a (deterministic) constant* $\lambda_i \in \mathbb{R}$ *such that the limit*

$$\lambda_i = \lim_{n \to \infty} \frac{1}{n} \log \sigma_i(D_z F^n_{\underline{\omega}}) \tag{11}$$

holds for all $(\underline{\omega}, z) \in \Gamma$.

(b) *Let* $\chi_1 > \cdots > \chi_r$ *denote the* distinct *values of the Lyapunov exponents* $\{\lambda_i\}$ *(note* $r = 1$ *is possible). Then, for all* $(\underline{\omega}, z) \in \Gamma$*, there is a flag of subspaces*

$$\begin{aligned}\mathbb{R}^{2N} =: F_1(\underline{\omega}, z) \supsetneq F_2(\underline{\omega}, z) \supsetneq \cdots \supsetneq F_r(\underline{\omega}, z)\\ \supsetneq \mathbb{F}_{r+1}(\underline{\omega}, z) := \{0\}\end{aligned}$$

such that for all $i \in \{1, \cdots, r\}$*, we have that (i)* $D_z F_{\underline{\omega}}(F_i(\underline{\omega}, z)) = F_i(\tau(\underline{\omega}, z))$ *and (ii) for all* $v \in F_i(\underline{\omega}, z) \backslash F_{i+1}(\underline{\omega}, z)$*, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \|D_z F^n_{\underline{\omega}}(v)\| = \chi_i \,.$$

Lastly, (iii) there are (deterministic) natural numbers $m_1, \cdots, m_r$, $\sum_i m_i = 2N$ *such that for all,* $1 < i \leq r$ *and* $(\underline{\omega}, z) \in \Gamma$*, we have that the codimension of* $F_i(\underline{\omega}, z)$ *in* $F_{i-1}(\underline{\omega}, z)$ *is equal to* m_{i-1}*.*

Proof. Condition (E) implies $m := Leb_{\mathbb{T}^{2N}}$ is ergodic as a stationary measure, hence $\mathbb{P} \times m$ is τ-ergodic (Proposition 7). From here, the usual multiplicative ergodic theorem for ergodic measure-preserving transformations applies to the linear cocycle $A^n_{\underline{\omega},z} := D_z F^n_{\underline{\omega}}$ over $\tau : \Omega \times \mathbb{T}^{2N} \circlearrowleft$, c.f. [20]. □

3.2.3. Lyapunov exponents from statistics of N-dimensional planes We now turn our attention to the method used in this paper to estimate Lyapunov exponents from below: by considering the stationary statistics of N-dimensional subspaces of tangent space evolving under the derivative $D_z F_\omega$. More precisely, let $\mathrm{Gr}_m(\mathbb{R}^k)$ denote the Grassmanian of m-dimensional subspaces of $\mathbb{R}^k$. For fixed initial $(Z_0, E_0) \in \mathrm{Gr}_N(\mathbb{T}^{2N}) := \mathbb{T}^{2N} \times \mathrm{Gr}_N(\mathbb{R}^{2N})$, the Markov chain (Z_n, E_n) on $\mathrm{Gr}_N(\mathbb{T}^{2N})$ is defined by setting

$$E_n := D_z F^n_{\underline{\omega}}(E_0).$$

This gives rise to the corresponding skew-product construction $\hat{\tau} : \Omega \times \mathrm{Gr}_N(\mathbb{T}^{2N}) \circlearrowleft$ given by $\hat{\tau}(\underline{\omega}, z, E) = (\theta\underline{\omega}, F_{\omega_1}(z), D_z F_{\omega_1}(E))$. Recall that a probability measure $\hat{\nu}$ on $\mathrm{Gr}_N(\mathbb{T}^{2N})$ is stationary for (Z_n, E_n) if

$$\hat{\nu}(A) = \int \hat{P}((z, E), A) d\hat{\nu}(z, E)$$

for all measurable $A \subset \mathrm{Gr}_N(\mathbb{T}^{2N})$, where the transition kernel $\hat{P}$ is defined by

$$\hat{P}((z, E), \cdot) := \mathbb{P}((F_{\omega_1}(z), D_z F_{\omega_1}(E)) \in \cdot). \tag{12}$$

The following relates stationary measures $\hat{\nu}$ for the (Z_n, E_n) chain to Lyapunov exponents.

Lemma 12. *Assume condition (E). Let* $\hat{\nu}$ *be any ergodic stationary measure for* (Z_n, E_n)*. Then,*

$$\sum_{i=1}^{N} \lambda_i \geq \int \log \det(D_z F_{\omega_1}|_E)\, d(\mathbb{P} \times \hat{\nu})(\underline{\omega}, z, E). \tag{13}$$

Above, for a $2N \times 2N$ matrix A and $E \subset \mathbb{R}^{2N}$, $\dim E = N$, we write $A|_E : E \to A(E)$ for the linear mapping of E to $A(E)$ obtained by restricting A to E. From this standpoint, $\det(A|_E)$ is defined as usual, e.g., as the volume ratio

$$\det(A|_E) := \frac{\mathrm{Leb}_{A(E)}\, A(B_E)}{\mathrm{Leb}_E(B_E)},$$

where $B_E \subset E$ is the unit ball, and Leb_E denotes Lebesgue measure on E.

Proof. To start, it is straightforward to check that the projection of $\hat{\nu}$ onto $\mathbb{T}^{2N}$ is a stationary measure for (Z_n), hence by (E) this projected measure coincides with $m = \mathrm{Leb}_{\mathbb{T}^{2N}}$. Next, by Proposition 7, we have that $\mathbb{P} \times \hat{\nu}$ is $\hat{\tau}$-ergodic. By the Birkhoff ergodic theorem applied to $\hat{\tau}$, we have

$$\begin{aligned}&\int \log\det(D_z F_\omega|_E)\, d(\mathbb{P}\times\hat{\nu})(\underline{\omega}, z, E)\\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \varphi\circ\hat{\tau}^i(\underline{\omega}, z, E) = \lim_{n\to\infty}\frac{1}{n}\log\det(D_z F^n_{\underline{\omega}}|_E)\end{aligned}$$

for $\mathbb{P} \times \hat{\nu}$-almost every $(\underline{\omega}, z, E)$, where $\varphi(\underline{\omega}, z, E) := \log\det(D_z F_{\omega_1}|_E)$. Note that $\det(D_z F^n_{\underline{\omega}}|_E) \le \prod_{i=1}^N \sigma_i(D_z F^n_{\underline{\omega}})$ (Lemma 30 in Appendix A). By Lemma 11, we conclude

$$\int \log\det(D_z F_\omega|_E)\, d\hat{\nu}(z, E) \le \limsup_{n\to\infty}\frac{1}{n}\sum_{i=1}^N \log\sigma_i(D_z F^n_{\underline{\omega}}) = \sum_{i=1}^N \lambda_i.$$

□

Remark 13. It is straightforward to check that equality holds in (13) if $\hat{\nu} \ll \mathfrak{m}$, the Riemannian volume on $\mathrm{Gr}_N(\mathbb{T}^{2N})$ (see Sect. 5.1 for details). However, this fact is not used, and so we omit a proof.

A key component of our analysis is the use of the nondegeneracy condition (ND) to provide a priori control on the density of stationary measures $\hat{\nu}$ for the Grassmanian Markov chain (Z_n, E_n).

Lemma 14. *Assume (E) and (ND) hold. Let $\hat{\nu}$ be any stationary measure for (Z_n, E_n). Then, $\hat{\nu} \ll \mathfrak{m}$, where $\mathfrak{m}$ is the Riemannian volume on $\mathrm{Gr}_N(\mathbb{T}^{2N})$, and satisfies*

$$\left\| \frac{d\hat{\nu}}{d\mathfrak{m}} \right\|_{L^\infty} \le M$$

where all notation is as in condition (ND).

Proof. For $(z, E) \in \mathrm{Gr}_N(\mathbb{T}^{2N})$ and $K \subset \mathrm{Gr}_N(\mathbb{T}^{2N})$, the transition kernel $\hat{P}$ for (Z_n, E_n) satisfies

$$\hat{P}((z, E), K) = \hat{Q}((Fz, D_z F(E)), K)$$

where $\hat{Q}$ is the kernel for R_ω as in (5). In particular, by (ND), we have that $\hat{P}((z, E), \cdot) \ll \mathfrak{m}$, where $\mathfrak{m}$ is normalized Lebesgue measure on $\mathrm{Gr}_N(\mathbb{T}^{2N})$, while $d\hat{P}((z, E), \cdot)/d\mathfrak{m} = d\hat{Q}((Fz, D_zF(E)), \cdot)/d\mathfrak{m}$ satisfies

$$\left\| \frac{d\hat{P}((z, E), \cdot)}{d\mathfrak{m}} \right\|_{L^\infty} \leq M$$

uniformly in $(z, E) \in \mathrm{Gr}_N(\mathbb{T}^{2N})$. Moreover, stationarity of $\hat{\nu}$ (see (12)) implies that for $K \subset \mathrm{Gr}_N(\mathbb{T}^{2N})$ measurable, we have

$$\hat{\nu}(K) = \int_{\mathrm{Gr}_N(\mathbb{T}^{2N})} \hat{P}((z, E), K) d\hat{\nu}(z, E) \leq M\mathfrak{m}(K).$$

Therefore, $\hat{\nu} \ll \mathfrak{m}$ and $d\hat{\nu}/d\mathfrak{m}$ is essentially bounded from above by M. □

3.3. Hyperbolicity estimates assuming (F1), (F2). Let us record some estimates describing the quality of the predominant hyperbolicity of the family $F = F_L$. Fix $\beta \in (0, 1)$. Recall the notation

$$\begin{aligned} \mathcal{C}^x_\alpha &:= \{(u, v) \in \mathbb{R}^{2N} : \|v\| \leq \alpha\|u\|\}, \\ B_\beta &= \{x \in \mathbb{T}^N : |\det D_x f_L| \leq L^{N-(1-\beta)}\} \subset \mathbb{T}^N. \end{aligned}$$

We define $G_\beta = B^c_\beta$.

Lemma 15. *Fix $\beta \in (0, 1)$ and let L be sufficiently large. Let $z = (x, y) \in \mathbb{T}^{2N}$ be such that $x \in G_\beta$.*

(a) *Let $w = (u, v) \in T_z\mathbb{T}^{2N} \cong \mathbb{R}^{2N}$ be such that $w \in \mathcal{C}^x_{1/10}$. Then, $D_zF(w) \in \mathcal{C}^x_{1/10}$, and*

$$\|D_zF(w)\| \geq L^{\frac{2}{3}\beta}\|w\|.$$

(b) *Let $E \subset \mathbb{R}^{2N}$ be an N-dimensional subspace such that $E \subset \mathcal{C}^x_{1/10}$. Then,*
   (i) *$E' := D_zF(E)$ is an N-dimensional subspace satisfying $E' \subset \mathcal{C}^x_{1/20}$, and*
   (ii) $\det(D_zF|_E) \geq \frac{1}{2^N} L^{N-(1-\beta)}$.

Proof. For an $N \times N$ matrix A, write $m(A) = \|A^{-1}\|^{-1} = \min\{\|Av\|/\|v\| : v \in \mathbb{R}^N \setminus \{0\}\}$ for the minimum norm of A (setting $m(A) = 0$ if A is not invertible). To start, the estimate

$$m(D_x f) \geq C_0^{-(N-1)} L^\beta \tag{14}$$

follows from (F1) and the standard fact that $m(D_x f) \geq \det(D_x f)/\|D_x f\|^{N-1}$.

For the estimate in (a), assume $w = (u, v) \in \mathcal{C}^x_\alpha$ for some $\alpha > 0$. Then $\|v\| \leq \alpha\|u\|$ and $\|u\| \leq \|w\| \leq \sqrt{1 + \alpha^2}\|u\|$. So,

$$\begin{aligned} \|D_zF(w)\| &\geq \|D_x f(u)\| - \|v\| \geq (m(D_x f) - \alpha)\|u\| \\ &\geq \frac{C_0^{-(N-1)} L^\beta - \alpha}{\sqrt{1 + \alpha^2}} \|w\| \\ &\geq L^{2\beta/3}\|w\|. \end{aligned}$$

The last inequality above holds when L is large enough and $\alpha = 1/10$. Similar estimates imply $D_z F(w) \in \mathcal{C}^x_{1/10}$ when L is large enough.

For (b)(i): by hypothesis, we can express $E = \operatorname{graph} G = \{(u, G(u)) : u \in \mathbb{R}^x\}$, where $G : \mathbb{R}^x \to \mathbb{R}^y$ is a linear map with $\|G\| \le 1/10$. To express $E' = D_z F(E)$ in the form $E' = \operatorname{graph} G'$, we would need to have that for all $u' \in \mathbb{R}^x$ there exists $u \in \mathbb{R}^x$ so that

$$\begin{pmatrix} u' \\ G'(u') \end{pmatrix} = \begin{pmatrix} D_x f & -I_N \\ I_N & 0_N \end{pmatrix} \begin{pmatrix} u \\ G(u) \end{pmatrix} = \begin{pmatrix} D_x f(u) - G(u) \\ u \end{pmatrix}$$

Formally, then, we ought to have $G' = (D_x f - G)^{-1}$. That this exists follows from (14); moreover,

$$\|G'\| \le \frac{1}{m(D_x f) - 1/10} \le 2L^{-\beta} \ll 1/20$$

when L is sufficiently large, hence $E' = D_z F(E) \subset \mathcal{C}^x_{1/20}$ as desired.

For (b)(ii), define $\Pi^x : \mathbb{R}^{2N} \cong \mathbb{R}^x \times \mathbb{R}^y \to \mathbb{R}^x$ to be the orthogonal projection onto $\mathbb{R}^x$. Then,

$$\det(D_z F|_E) = \det(D_x f - G) \cdot \frac{\det(\Pi^x|_E)}{\det(\Pi^x|_{E'})} = \det(D_x f - G) \cdot \frac{\det(I + G')}{\det(I + G)} \tag{15}$$

on noting that $(\Pi^x|_E)^{-1} = (I + G) : \mathbb{R}^x \to \mathbb{R}^{2N}$, and similarly for $\Pi^x|_{E'}$. For these terms we have $(1 - 1/10)^N \le \det(I + G), \det(I + G') \le (1 + 1/10)^N$, while for the remaining $D_x f - G$ term we have

$$\det(D_x f - G) \ge \det D_x f - \frac{\|G\|}{m(D_x f) - \|G\|} \ge \frac{1}{2} L^{N-(1-\beta)}$$

using the elementary estimate $|\det(A + B) - \det(A)| \le \|B\|/(m(A) - \|B\|)$. □

Condition (C) says that the randomizations $R_\omega, \omega \in \Omega_0$ do not 'disrupt' the hyperbolicity of the system too much. The following is an immediate consequence of (C) and Lemma 15.

Lemma 16. *The following holds for $\mathbb{P}_0$-a.e. $\omega \in \Omega_0$. Fix $\beta \in (0, 1)$ and let L be sufficiently large. Let $z = (x, y) \in \mathbb{T}^{2N}$, $x \in G_\beta$. Then,*

(i) *Let $w = (u, v) \in T_z \mathbb{T}^{2N} \cong \mathbb{R}^{2N}$ be such that $w \in \mathcal{C}^x_{1/10}$. Then,*

$$\|D_z F_\omega(w)\| \ge L^{\frac{1}{2}\beta} \|w\|$$

(ii) *Let $E \subset \mathcal{C}^x_{1/10}$ be an N-dimensional subspace. Then, $E' = D_z F_\omega(E)$ is an N-dimensional subspace with $E' \subset \mathcal{C}^x_{1/10}$.*

4. Proof of Theorem 5

In brief, our method will be to obtain a lower bound of the form

$$\sum_{i=1}^{N} \lambda_i \geq (1-\varepsilon) N \log L$$

on the sum of the first N Lyapunov exponents for $\varepsilon > 0$ small and L sufficiently large. This directly implies $\lambda_i \geq (1-(2N-1)\varepsilon)\log L$ for each $1 \leq i \leq N$, in view of the fact that $\lambda_i \leq \lambda_1 \leq (1+\varepsilon)\log L$ for all $i, \varepsilon > 0$ and $L \gg 1$ (see condition (F1)). Since $\sum_{i=1}^{2N} \lambda_i = 0$, similar considerations apply to the exponents $\lambda_{N+1}, \cdots, \lambda_{2N}$. These proofs are straightforward and omitted for brevity.

From this point forward, we will focus our attention on $\sum_{i=1}^{N} \lambda_i$, which we shall estimate using an ergodic stationary measure $\hat{\nu}$ for the Markov chain (Z_n, E_n) on $\mathrm{Gr}_N(\mathbb{T}^{2N})$, following Lemma 12. Applying (13), condition (C), and the chain rule $D_z F_\omega = D_{F_{\omega z}} R_\omega \circ D_z F$, we have

$$\sum_{i=1}^{N} \lambda_i \geq -N \log 2 + \int \log |\det(D_z F|_E)| \, d\hat{\nu}(z, E)$$

By stationarity, for any bounded measurable $\phi : \mathrm{Gr}_N(\mathbb{T}^{2N}) \to \mathbb{R}$ we have (by a slight abuse of notation)

$$\int \phi(z, E) d\hat{\nu}(z, E) = \int \left(\mathbb{E}_{(z,E)} \phi(Z_n, E_n)\right) d\hat{\nu}(z, E)$$

for all $n \geq 1$ (recall that $\mathbb{E}_{(z,E)}$ denotes the expectation conditioned on the initial state $(Z_0, E_0) = (z, E)$). Treating (Z_n, E_n) as a function of the initial condition and the random sample $\underline{\omega}$, we also have

$$\int \phi(z, E) d\hat{\nu}(z, E) = \mathbb{E}\left[\int \phi(Z_n, E_n) d\hat{\nu}(z, E)\right]$$

(see (10)). Applying to $\phi(z, E) = \log |\det(D_z F|_E)|$, we conclude

$$\sum_{i=1}^{N} \lambda_i \geq -N \log 2 + \mathbb{E} \underbrace{\int \log |\det(D_{Z_n} F|_{E_n})| d\nu(z, E)}_{(**)}.$$

To prove Theorem 5 it therefore suffices to bound $(**)$ as follows.

Proposition 17. (Main estimate) *Fix $\alpha, \beta \in (0, 1)$ and $\delta \in (0, c_\beta)$. Let L be sufficiently large in terms of these parameters. Then, there exists $n \gg 1$, depending on L, such that for a.e. $\underline{\omega} \in \Omega$, we have*

$$(**) = \int \log |\det(D_{Z_n} F|_{E_n})| d\nu(z, E) \geq \alpha N \log L.$$

The proof of Theorem 5 is complete upon adjusting the parameter α and taking L large enough to absorb the remaining additive term $-N \log 2$.

Proof of Proposition 17: exploiting predominant hyperbolicity. Below $n \geq 1$ is fixed, to be determined later, and $\underline{\omega} \in \Omega$ is an arbitrary random sample. Recall that $D_z F_\omega$ is strongly expanding in the G_β along which $D_z F_\omega$ is strongly expanding in the horizontal cone $\mathcal{C}_\alpha^x = \{(u, v) : \|v\| \leq \alpha\|u\|\}$ for $z \in G_\beta$ (Lemma 15). For $n \geq 1$, define

$$G_\beta^n = \{z \in \mathbb{T}^{2N} : Z_i \in G_\beta \text{ for all } 0 \leq i \leq n-1\}$$

to be the set of trajectories experiencing this hyperbolicity for n timesteps, where as usual we condition on $Z_0 = z$.

Fix $z \in G_\beta^n$. Hyperbolic expansion along the x-direction $\mathbb{R}^x$ implies that the 'bulk' of Grassmanian dynamics is attracted to a close vicinity of $\mathbb{R}^x$. This is, after all, the conceptual picture underlying the $N = 1$ case studied in the previous paper [4]. The following is the analogue of Lemma 10 in [4].

Proposition 18. *Let $\underline{\omega} \in \Omega$ be arbitrary, and let $\beta \in (0, 1)$, $n \geq 1$. Fix $z \in G_\beta^n$. Set $E_n = D_z F_{\underline{\omega}}^n(E)$. Then,*

$$\mathrm{Leb}_{\mathrm{Gr}_N(\mathbb{R}^{2N})}\{E \in \mathrm{Gr}_N(\mathbb{R}^{2N}) \text{ such that } E_n \notin \mathcal{C}_2^x\} \leq L^{-\beta n}.$$

The proof of Proposition 18 is deferred for now. Let us show how it can be used to prove Proposition 17. For $z \in G_\beta^n$ define $\mathcal{G}_z^n = \{E \in \mathrm{Gr}_N(\mathbb{R}^{2N}) : E_n \in \mathcal{C}_2^x\}$. Letting $\beta^* \in (0, 1)$ be a parameter to be chosen later, define

$$\mathcal{G}^n = \{(z, E) \in \mathrm{Gr}_N(\mathbb{T}^{2N}) : z \in G_\beta^n \cap (F_{\underline{\omega}}^n)^{-1} G_{\beta^*},\ E \in \mathcal{G}_z^n\}$$

and $\mathcal{B}^n = \mathrm{Gr}_N(\mathbb{T}^{2N}) \backslash \mathcal{G}^n$. The integral of $(**)$ along $(z, E) \in \mathcal{G}^n \subset \mathrm{Gr}_N(\mathbb{T}^{2N})$ will result in a tight lower bound for $\det(D_{Z_N} F|_{E_n})$, while $\mathcal{B}^n$ is an error set along which we use the poor estimate

$$\log|\det(D_{Z_n} F|_{E_n})| \geq -N \log(2C_0 L)\,, \tag{16}$$

which follows from (F1) and the form of the mapping $F = F_L$.

Splitting $(**)$ along the partition $\mathcal{G}^n, \mathcal{B}^n$, we have

$$\int_{\mathcal{G}^n} \log|\det(D_{Z_n} F|_{E_n})| d\nu(z, E) \geq (1 - \nu(\mathcal{B}^n)) \underbrace{\inf_{z \in G_{\beta^*}, E \in \mathcal{C}_2^x} \log|\det(D_z F|_E)|}_{\dagger}\,.$$

Choosing β^* sufficiently close to 1, we can arrange for $\dagger \geq \frac{1+\alpha}{2} N \log L$ (Lemma 15(b)(ii)) on taking L sufficiently large. Plugging in (16), we obtain

$$(**) \geq \frac{1+\alpha}{2} N \log L - 2\nu(\mathcal{B}^n) \cdot N \log(2C_0 L).$$

It remains to bound $\nu(\mathcal{B}^n)$ from above. We decompose $\mathcal{B}^n = \mathcal{B}^{n,1} \cup \mathcal{B}^{n,2}$, where

$$\mathcal{B}^{n,1} = \left((G_\beta^n)^c \cup (F_{\underline{\omega}}^n)^{-1} G_{\beta^*}^c\right) \times \mathrm{Gr}_N(\mathbb{R}^{2N})$$
$$\mathcal{B}^{n,2} = \{(z, E) : z \in G_\beta^n, E_n \notin \mathcal{C}_2^x\} = \{(z, E) : z \in G_\beta^n, E \notin \mathcal{G}_z^n\}.$$

For $\mathcal{B}^{n,1}$ we have the simple estimate

$$\nu(\mathcal{B}^{n,1}) = \mathrm{Leb}_{\mathbb{T}^{2N}}((G_\beta^n)^c \cup (F_{\underline{\omega}}^n)^{-1} G_{\beta^*}^c) \leq n C_\beta L^{-c_\beta} + C_{\beta^*} L^{-c_{\beta^*}}.$$

For $\mathcal{B}^{n,2}$, we estimate

$$\nu(\mathcal{B}^{n,2}) \le M \operatorname{Leb}_{\mathrm{Gr}_N(\mathbb{T}^{2N})}(\mathcal{B}^{n,2}) \le ML^{-\beta n}$$

using our bound on $\frac{d\nu}{d\mathrm{m}}$ from Lemma 14 and the estimate in Proposition 18.

In total, we have shown that

$$(**) \ge \frac{\alpha+1}{2} N \log L - N \log L \left(nC_\beta L^{-c_\beta} + ML^{-\beta n} + C_{\beta^*} L^{-c_{\beta^*}} \right)$$

Fix $n = \lceil L^{c_\beta - \delta} \rceil$ for some small $\delta \ll c_\beta$. Then, $nC_\beta L^{-c_\beta} = O(L^{-\delta})$, while $ML^{-\beta n} \lesssim L^{-\delta}$ as long as

$$M \le L^{\frac{1}{2}\beta L^{c_\beta - \delta}}.$$

Thus, under this condition relating M and L, we have

$$(**) \ge \alpha N \log L + \left(\frac{1-\alpha}{2} - CL^{-\min\{\delta, c_{\beta^*}\}} \right) N \log L \ge \alpha N \log L$$

assuming L is sufficiently large in terms of $\alpha, \beta, \beta^*, \delta$. This completes the proof of Proposition 17.

5. Proof of Main Proposition (Proposition 18)

In Sect. 5.1 we recall some necessary preliminaries on the Grassmanian as a Riemannian manifold. The proof of Proposition 18 is carried out in Sect. 5.2.

5.1. Grassmanian preliminaries. Fix $m \ge 1$ and $1 \le k < m$. Here we describe the smooth and Riemannian structures of the manifold $\mathrm{Gr}_k(\mathbb{R}^m)$ of k-dimenisonal subspaces of $\mathbb{R}^m$, and give a few preliminary lemmas. The following is all well-known; see, e.g., [23,24].

To fix ideas and avoid dealing with unnecessary cases, we will exclusively deal with the case when $k \le \frac{m}{2}$, hence $k \le m-k$. Otherwise, we can reduce to this case by noting that orthogonal projection provides a natural identification $\mathrm{Gr}_k(\mathbb{R}^m) \cong \mathrm{Gr}_{m-k}(\mathbb{R}^m)$. Throughout, $\mathbb{R}^m$ carries the standard Euclidean inner product $\langle \cdot, \cdot \rangle$.

Manifold structure of $\mathrm{Gr}_k(\mathbb{R}^m)$ Given $E \in \mathrm{Gr}_k(\mathbb{R}^m)$, we define the coordinate patch $\mathcal{U}_E = \{\mathrm{graph}_E H : H \in L(E, E^\perp)\}$, where we write $L(E, E^\perp)$ for the space of linear maps from E to $E^\perp$, and the chart map

$$\mathrm{graph}_E : \mathcal{U}_E \to \mathrm{Gr}_k(\mathbb{R}^m)$$

is defined by $\mathrm{graph}_E H = \{v + H(v) : v \in E\}$.

We highlight the following facts:

(A) We have that $\mathcal{U}_E$ is the set of all $E' \in \mathrm{Gr}_k(\mathbb{R}^m)$ intersecting $E^\perp$ transvsersally. In particular, $\mathcal{U}_E$ is open and dense for any $E \in \mathrm{Gr}_k(\mathbb{R}^m)$.

(B) We have the following basis-independent identification:

$$T_E \operatorname{Gr}_k(\mathbb{R}^m) = L(E, E^\perp).$$

If bases for $E, E^\perp$ are fixed, then we have the parametrization $\mathcal{U}_E \cong M_{m-k,k}(\mathbb{R})$, the space of $(m-k) \times k$ real matrices.

The following is a qualitative geometric description of the complement of any chart $\mathcal{U}_E$.

Lemma 19. *Assume $k \leq m/2$. Then, the set $(\mathcal{U}_E)^c$ is a finite union of closed submanifolds of $\operatorname{Gr}_k(\mathbb{R}^m)$ of codimension ≥ 1.*

For the sake of completeness, we provide the proof of Lemma 19 in Appendix B.

Riemannian metric on $\operatorname{Gr}_m(\mathbb{R}^k)$ With respect to the identification in (B) above, the Riemannian metric g on $T_E \operatorname{Gr}_k(\mathbb{R}^m)$ can be expressed as

$$g_E(H_1, H_2) = \operatorname{Tr}_E(H_2^\top H_1),$$

where Tr_E denotes the trace induced by the inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^m$ restricted to E. As usual, the Riemannian metric induces a volume measure $\operatorname{Leb}_{\operatorname{Gr}_k(\mathbb{R}^m)}$ and a geodesic distance d_{geo} between subspaces in $\operatorname{Gr}_k(\mathbb{R}^m)$. Recall that the orthogonal group $\mathcal{O}(m)$ acts on $M = \operatorname{Gr}_k(\mathbb{R}^m)$ via the action $E \mapsto U(E)$ for $U \in \mathcal{O}(m)$. It is standard that the orthogonal group acts isometrically on (M, g). In fact, (M, g) is the unique (up to scalar) Riemannian metric on M with respect to which $\mathcal{O}(m)$ acts isometrically.

The following alternative metric d_H on Grassmannians is very useful in practice.

Definition 20. Let $E, E' \in \operatorname{Gr}_k(\mathbb{R}^m)$. We define the *Hausdorff distance* $d_H(E, E')$ between them by

$$d_H(E, E') = \max \left\{ \max_{\substack{v' \in E' \\ \|v'\|=1}} d(v', E), \max_{\substack{v \in E \\ \|v\|=1}} d(v, E') \right\},$$

where above $d(v, E)$ denotes the minimal Euclidean distance between $v \in \mathbb{R}^m$ and $E \subset \mathbb{R}^m$.

The distance function d_H is uniformly equivalent to the geodesic distance d_{geo}:

Lemma 21. *For any $E, E' \in \operatorname{Gr}_k(\mathbb{R}^m)$, we have*

$$\frac{2}{\pi} d_{geo}(E, E') \leq d_H(E, E') \leq d_{geo}(E, E')$$

This appears to be well-known, but we are unable to find a proof of Lemma 21 in the literature. For the sake of completeness a sketch is provided below in "Appendix B".

5.2. The proof of Proposition 18. Throughout, the parameter $\beta \in (0, 1)$ is fixed, as are $n \geq 1$, $\underline{\omega} \in \Omega$ and $z \in G^n_\beta$. We now proceed to study the singular-value decomposition for the iterated Jacobian $D_z F^n_{\underline{\omega}}$.

Lemma 22. *Let $\sigma_i = \sigma_i(D_z F^n_\omega), \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{2N}$ denote the singular values of $D_z F^n_\omega$.*

(i) *We have*

$$\sigma_N \geq L^{n\beta/2} \geq L^{-n\beta/2} \geq \sigma_{N+1}$$

(ii) *Let $h_1, \cdots, h_{2N}, h'_1, \cdots, h'_{2N}$ denote orthonormal bases of $\mathbb{R}^{2N}$ for which*

$$D_z F^n_{\underline{\omega}} h_i = \sigma_i h'_i$$

Then, $h_i, h'_i \in \mathcal{C}^x_{1/10}$ for $1 \leq i \leq N$ and $h_i, h'_i \in \mathcal{C}^y_{1/10}$ for all $N + 1 \leq i \leq 2N$.

For $\alpha > 0$, we have written $\mathcal{C}^y_\alpha = \{(u, v) \in \mathbb{R}^{2N} : \|u\| \leq \alpha\|v\|\}$ for the cone of vectors roughly parallel to $\mathbb{R}^y$.

Proof. It follows from Lemma 16(ii) that for any $E \in \mathrm{Gr}_N(\mathbb{R}^{2N})$, $E \subset \mathcal{C}^x_{1/10}$, we have

$$D_{Z_0} F^n_{\underline{\omega}}(E) \subset \mathcal{C}^x_{1/10}.$$

A mild variation of the arguments for Lemma 16 similarly implies that $(D_{Z_0} F^n_{\underline{\omega}})^\top(E) \subset \mathcal{C}^x_{1/10}$. The same then holds for $(D_{Z_0} F^n_{\underline{\omega}})^\top D_{Z_0} F^n_{\underline{\omega}}$. By Lemma 31 in Appendix A, it follows that there are N orthonormal eigenvectors $h_{i_1}, \cdots, h_{i_N}$ in $\mathcal{C}^x_{1/10}$ for $(D_{Z_0} F^n_{\underline{\omega}})^\top D_{Z_0} F^n_{\underline{\omega}}$ spanning an N-dimensional space $E \subset \mathcal{C}^x_{1/10}$, such that $(D_{Z_0} F^n_{\underline{\omega}})^\top D_{Z_0} F^n_{\underline{\omega}}(h_{i_j}) = \sigma^2_{i_j}$.

We want to check that $\{i_1, \cdots, i_N\} = \{1, \cdots, N\}$. Note that Lemma 16(i) implies that the singular values $\sigma_{i_1}, \cdots, \sigma_{i_N}$ satisfy $\sigma_{i_j} \geq L^{n\beta/2}$. We combine this with the fact (straightforward to check) that $(D_{Z_0} F^n_{\underline{\omega}})^\top D_{Z_0} F^n_{\underline{\omega}}$ is symplectic with symplectic form $J = \begin{pmatrix} 0 & -I_N \\ I_N & 0 \end{pmatrix}$ (recall that a matrix A is symplectic for J if $A^\top J A = J$). It is well-known that the eigenvalues of a self-adjoint symplectic matrix come in pairs of the form λ, λ^{-1}; since we already have N real eigenvalues $\sigma^2_{i_1}, \cdots, \sigma^2_{i_N}$ which are > 1, this argument implies the remaining N real eigenvalues for $(D_{Z_0} F^n_{\underline{\omega}})^\top D_{Z_0} F^n_{\underline{\omega}}$ are in the interval $(0, 1)$. Consequently,

$$\sigma_1 = \sigma^{-1}_{2N}, \quad \sigma_2 = \sigma^{-1}_{2N-1}, \quad \cdots, \quad \sigma_N = \sigma^{-1}_{N+1}.$$

This implies $\{i_1, \cdots, i_N\} = \{1, \cdots, N\}$, and moreover, that the estimate $\sigma_N \geq L^{n\beta/2} \geq L^{-n\beta/2} \geq \sigma_{N+1}$ holds as in item (i).

A repetition of the above argument implies that there are N orthonormal eigenvectors $h'_1, \cdots, h'_N$ for $D_{Z_0} F^n_{\underline{\omega}}(D_{Z_0} F^n_{\underline{\omega}})^\top$ contained in $\mathcal{C}^x_{1/10}$, for which $D_{Z_0} F^n_{\underline{\omega}}(D_{Z_0} F^n_{\underline{\omega}})^\top(h'_i) = \sigma^2_i$. By Theorem 29, it follows that (up to changing signs and possible rearrangement of indices of the h_i, h'_i) that

$$D_{Z_0} F^n_{\underline{\omega}}(h_i) = \sigma_i h'_i$$

holds. It is now straightforward to complete $\{h_1, \cdots, h_N\}$ and $\{h'_1, \cdots, h'_N\}$ to orthonormal bases of $\mathbb{R}^{2N}$ with the desired properties. □

Define

$$\mathcal{H} = \mathcal{H}(D_z F^n_{\underline{\omega}}) = \mathrm{Span}\{h_1, \cdots, h_N\}$$
$$\mathcal{H}' = \mathcal{H}'(D_z F^n_{\underline{\omega}}) = \mathrm{Span}\{h'_1, \cdots, h'_N\},$$

noting that $\mathcal{H}^\perp = \mathrm{Span}\{h_{N+1}, \cdots, h_{2N}\}$, $(\mathcal{H}')^\perp = \mathrm{Span}\{h'_{N+1}, \cdots, h'_{2N}\}$. By Lemma 22, we have that $\mathcal{H}, \mathcal{H}' \subset \mathcal{C}^x_{1/10}$. Below, given $\eta > 0$ and $S \subset \mathrm{Gr}_N(\mathbb{R}^{2N})$, we write $\mathcal{N}_\eta(S)$ for the (open) η-neighborhood of S with respect to the geodesic distance d_{geo}.

Lemma 23. *There exists a universal constant $c > 0$ depending only on N such that*

$$\{E \in \mathrm{Gr}_N(\mathbb{R}^{2N}) : D_z F^n_{\underline{\omega}}(E) \text{ is not contained in } \mathcal{C}^x_2\} \subset \mathcal{N}_\eta((\mathcal{U}_\mathcal{H})^c), \tag{17}$$

where $\eta = cL^{-\beta n}$.

Proposition 18 follows, since $(\mathcal{U}_\mathcal{H})^c$ is the finite union of a collection of closed submanifolds of $\mathrm{Gr}_N(\mathbb{R}^{2N})$ (Lemma 19). Here, we use the standard fact that if $M' \subset M$ is a closed submanifold of a compact Riemannian manifold M with strictly positive codimension, then the Lebesgue measure of any neighborhood $\mathcal{N}_\eta(M')$ is $\leq C\eta$, where $C > 0$ depends only on M.

Proof of Lemma 23. Let $E \in \mathrm{Gr}_N(\mathbb{R}^{2N})$ be such that $E' := D_x F^n_{\underline{\omega}}(E)$ is not contained in $\mathcal{C}^x_2$. Let $v' \in E' \backslash \mathcal{C}^x_2$; since $\mathcal{C}^x_2$ is a cone and E' is a subspace, we can assume without loss of generality that v' is a unit vector.

Now, let

$$v' = v'_\parallel + v'_\perp,$$

where $v'_\parallel \in \mathcal{H}'$, $v'_\perp \in (\mathcal{H}')^\perp$. Since $v' \notin \mathcal{C}^x_2$ and $\mathcal{H}' \subset \mathcal{C}^x_{1/10}$, it follows that $\|v'_\perp\| \approx 1$. Now, let $v = (D_z F^n_{\underline{\omega}})^{-1}(v')/\|(D_z F^n_{\underline{\omega}})^{-1}(v')\|$, so that

$$v = v_\parallel + v_\perp = \frac{(D_z F^n_{\underline{\omega}})^{-1}(v'_\parallel)}{\|(D_z F^n_{\underline{\omega}})^{-1}(v')\|} + \frac{(D_z F^n_{\underline{\omega}})^{-1}(v'_\perp)}{\|(D_z F^n_{\underline{\omega}})^{-1}(v')\|}$$

since $D_z F^n_{\underline{\omega}}(\mathcal{H}) = \mathcal{H}'$, $D_z F^n_{\underline{\omega}}(\mathcal{H}^\perp) = (\mathcal{H}')^\perp$ (see "Appendix A"). To estimate these components, we have

$$\|(D_z F^n_{\underline{\omega}})^{-1}(v'_\parallel)\| \leq (\sigma_N(D_z F^n_{\underline{\omega}}))^{-1}\|v'_\parallel\| \leq L^{-\beta n/2},$$
$$\|(D_z F^n_{\underline{\omega}})^{-1}(v'_\perp)\| \geq (\sigma_{N+1}(D_z F^n_{\underline{\omega}}))^{-1}\|v'_\perp\| \gtrsim L^{\beta n/2}$$

using that $\|v'_\perp\| \approx 1$. Continuing,

$$\|(D_z F^n_{\underline{\omega}})^{-1}(v')\| \geq \|(D_z F^n_{\underline{\omega}})^{-1}(v'_\perp)\| - \|(D_z F^n_{\underline{\omega}})^{-1}(v'_\parallel)\| \gtrsim L^{\beta n/2} - L^{-\beta n/2} \gtrsim L^{\beta n/2}.$$

Thus, we conclude that $\|v_\parallel\| \lesssim L^{-\beta n}$, while $1 - \|v_\perp\| = O(L^{-\beta n})$. This immediately implies that

$$d(v, \mathcal{H}^\perp) \leq \|v - \hat{v}_\perp\| = O(L^{-\beta n}), \tag{18}$$

where $\hat{v}_\perp = \frac{v_\perp}{\|v_\perp\|} \in \mathcal{H}^\perp$ is a unit vector.

Fix a basis $w_1, \cdots, w_{N-1}$ for the orthogonal complement of v in E and define

$$\hat{E} := \mathrm{Span}\{\hat{v}, w_1, \cdots, w_{N-1}\}.$$

It follows from (18) that $d_H(\hat{E}, E) \lesssim L^{-\beta n}$. Since $\hat{E} \in (\mathcal{U}_\mathcal{H})^c$ and d_H, d_{geo} are uniformly equivalent by Lemma 21, we conclude that $d_{geo}(E, (\mathcal{U}_\mathcal{H})^c) \lesssim L^{-\beta n}$. □

6. Proof of Theorems 1 and 3

6.1. Proof of Theorem 1. It suffices to check that condition (F2) holds for the family

$$f_L(x_1, \cdots, x_N) = (2x_i + L\sin(2\pi x_i) + \sum_{j\neq i} \mu_{ij} \sin 2\pi(x_j - x_i))_{i=1}^N,$$

where (μ_{ij}) is a fixed family of coefficients.

Write $f = (f^1, \cdots, f^N)$ in component form. By the Leibniz formula for the determinant,

$$\det L^{-1} D_x f = L^{-N} \sum_{\tau\in S_N} \prod_{i=1}^N \operatorname{sgn}(\tau) \frac{\partial f^i}{\partial x_{\tau(i)}}$$

where the outer summation is over the permutations $\tau \in S_N$ of the set $\{1, \cdots, N\}$ and the sgn is the sign function of permutations in the permutation group, which returns +1 and -1 for even and odd permutations, respectively. The dominant term is the product $\prod_{i=1}^N 2\pi L \cos 2\pi x_i$; precisely,

$$\det L^{-1} D_x f - (2\pi)^N \prod_{i=1}^N \cos 2\pi x_i = O(L^{-1})$$

when L is taken sufficiently large relative to $\max_{ij} |\mu_{ij}|$. Therefore

$$B_\beta \subset \left\{ \left| \prod_{i=1}^N \cos 2\pi x_i \right| \leq L^{2\beta-1} \right\}$$

for L sufficiently large. To estimate the volume of the RHS, we bound $|\cos 2\pi x| \geq 4\min\{|x - 1/4|, |x - 3/4|\}$, so that $|\prod_{i=1}^N \cos 2\pi x_i| \geq \min\{\prod_{i=1}^N |x_i - r_i|\}$, where the minimum is taken over all possible choices of $(r_1, \cdots, r_N) \in \{1/4, 3/4\}^N$. Thus,

$$B_\beta \subset \bigcup_{(r_i)_{i=1}^N \in \{1/4,3/4\}^N} \left\{ \prod_{i=1}^N |x_i - r_i| \leq L^{2\beta-1} \right\}$$

where the union is again over all possible configurations of $(r_1, \cdots, r_N)$. We conclude that

$$\operatorname{Leb} B_\beta \lesssim \operatorname{Leb}\{(x_1, \cdots, x_N) \in [0, 1]^N : \prod_i x_i \leq L^{2\beta-1}\}.$$

For the RHS, we have the asymptotic $\lesssim L^{3\beta-1}$ (see Lemma 24 below). Therefore, condition (F2) is satisfied with $c_\beta = 1 - 3\beta$.

Lemma 24. *Define* $S_N(\delta) = \{(x_1, \cdots, x_N) \in [0, 1]^N : \prod_i x_i \leq \delta\}$. *Then, For any* $\delta > 0$, *we have*

$$\operatorname{Leb} S_N(\delta) = \delta \sum_{i=0}^{N-1} \frac{(-\log \delta)^i}{i!}$$

Proof. Define

$$J = J(x_1, \cdots, x_N) := \min\{1 \le j \le N : 0 \le x_j \le \frac{\delta}{x_1 \cdots x_{j-1}}\},$$

and note that J is defined on $S_N(\delta)$. In particular,

$$\{J = j\} \cap S_N(\delta) = \left\{ \frac{\delta}{\prod_{i \le \ell - 1} x_i} \le x_\ell \le 1 \text{ for all } 1 \le \ell \le j-1, 0 \le x_j \le \frac{\delta}{x_1 \cdots x_{j-1}} \right\}$$

with no constraint on $x_{j+1}, \cdots, x_N$. Thus $\mathrm{Leb}(S_N(\delta) \cap \{J = j\}) = \mathrm{Leb}(S_j(\delta) \cap \{J = j\})$ for all j. It therefore suffices to compute $\mathrm{Leb}(S_N(\delta) \cap \{J = N\})$. This is given by

$$\int_{x_1=\delta}^1 \int_{x_2=\frac{\delta}{x_1}}^1 \cdots \int_{x_{N-1}=\frac{\delta}{x_1 \cdots x_{N-2}}}^1 \frac{\delta}{x_1 \cdots x_{N-1}} dx_{N-1} \cdots dx_1$$

For $c \in (0, 1)$, $k \ge 0$, we have

$$\int_{y=c}^1 \frac{c}{y} \left(\log \frac{c}{y}\right)^k dy = -\int_{\log \frac{c}{y}=0}^{\log c} \left(\log \frac{c}{y}\right)^k d \log \frac{c}{y} = -\frac{c}{k+1} (\log c)^{k+1}.$$

Thus, after k iterated integrals, we have

$$\begin{aligned} &\int_{x_{N-k}=\frac{\delta}{x_1 \cdots x_{N-k-1}}}^1 \cdots \int_{x_{N-1}=\frac{\delta}{x_1 \cdots x_{N-2}}}^1 \frac{\delta}{x_1 \cdots x_{N-1}} dx_{N-1} \cdots dx_{N-k} \\ &= \frac{(-1)^k}{k!} \frac{\delta}{x_1 \cdots x_{N-k-1}} \left(\log \frac{\delta}{x_1 \cdots x_{N-k-1}}\right)^k \end{aligned}$$

and after $k = N - 1$ such integrals, we deduce $\mathrm{Leb}(S_N(\delta) \cap \{J = N\}) = \frac{\delta}{(N-1)!}(-\log \delta)^{N-1}$. This completes the proof. □

6.2. Transversality criterion and genericity for (F2) .

6.2.1. Transversality criterion for (F2) Below, we derive the general transversality criterion (8) for condition (F2) for families of the form $f_L = L\psi + \varphi$.

Lemma 25 (Transversality Criterion). *Let $\psi, \varphi : \mathbb{T}^N \to \mathbb{R}^N$ be C^2 mappings and assume that*

$$\{\det D_x \psi = 0\} \cap \{D_x \det D\psi = 0\} = \emptyset. \tag{19}$$

Then, $f_L = L\psi + \varphi$ satisfies condition (F2) with $c_\beta = 1 - \beta$ for all L sufficiently large. Precisely, for any $\beta > 0$, there exists $C_\beta = C_\beta(\psi, \varphi)$ so that for any L sufficiently large (in terms of ψ, φ), we have that

$$\mathrm{Leb}\{\det(L^{-1} D_x f_L) \le L^{-(1-\beta)}\} \le C_\beta L^{-(1-\beta)}.$$

Remark 26. Note that $\psi(x) = (\sin 2\pi x_i)_i$ does *not* satisfy the transversality condition (19); this is why we had to check (F2) by hand in the proof of Theorem 1. However, (19) does hold for a large class of models: as we check below in Proposition 27, it is satisfied by a C^2 generic set of ψ.

Proof. We begin with the following straightforward consequence of the constant rank theorem applied to $x \mapsto \det D_x\psi$: there exist $\hat{C} > 0$ with the property that for any $0 \le \epsilon \le \hat{\epsilon}(\psi)$, we have

$$\text{Leb}\{\det D_x\psi \le \epsilon\} \le \hat{C}\epsilon.$$

We will also need the following estimate: if A, B are $N \times N$ matrices, then there exists $C_{A,B} > 0$, depending only on $\max |A_{ij}|$, $\max |B_{ij}|$, and N, such that

$$\det(A + \eta B) \ge \det(A) - C_{A,B}\eta.$$

This can be obtained, e.g., from the formula

$$\det(A + \eta B) - \det(A) = \int_0^\eta \text{Tr}\left(\text{Adj}(A + sB)B\right)ds,$$

where $\text{Adj}(\cdot)$ denotes the adjugate of a square matrix and $\eta \le 1$. With this notation, define $\tilde{C} = \sup_x C_{D_x\psi, D_x\varphi}$.

To complete the proof, let $x \in \mathbb{T}^N$ be such that $\det D_x\psi \ge 2L^{-(1-\beta)}$. Then,

$$\det D_x(\psi + L^{-1}\phi) \ge \det D_x\psi - \tilde{C}L^{-1} \ge L^{-(1-\beta)}$$

if $L^\beta \gg \tilde{C}$. Thus, $\{\det D_x\psi \ge 2L^{-(1-\beta)}\} \subset \{\det(L^{-1}D_x f_L) \ge L^{-(1-\beta)}\}$. Taking complements, we conclude that

$$\text{Leb}\{\det(L^{-1}D_x f_L) \le L^{-(1-\beta)}\} \le \text{Leb}\{\det D_x\psi \le 2L^{-(1-\beta)}\} \le 2\hat{C}L^{-(1-\beta)}$$

on taking L large enough so that $2L^{-(1-\beta)} \ll \hat{\epsilon}(\psi)$. The proof is complete on setting $C_\beta = 2\hat{C}$. □

6.2.2. Genericity of (F2) when $N = 2$ In this subsection, we consider genericity of the transversality condition (19) used to prove property (F2). We expect that (19) is generic in general. For simplicity, however, we prove this only in the special case $N = 2$.

Proposition 27. *There is a residual set $\mathcal{R}$ in $C^r(\mathbb{T}^2, \mathbb{R}^2)$, $r \ge 2$ such that for all $\psi \in \mathcal{R}$, equation (19) holds, i.e., we have that 0 is a regular value of $x \mapsto \det D_x\psi$.*

Proof. We write $\psi = (\psi_1, \psi_2)$ where $\psi_1, \psi_2 : \mathbb{T}^2 \to \mathbb{R}$. It is well-known that Morse functions are generic; without loss of generality, we may assume that ψ_1 and ψ_2 are Morse functions, hence have finitely many critical points. We also assume that the set of critical points of ψ_1 is disjoint from that of ψ_2, which can be achieved by an arbitrary small translation $\psi_1(\cdot) \mapsto \psi_1(\cdot + a)$, $a \in \mathbb{R}^2$ small. Thus we conclude for all $x \in \mathbb{T}^2$ either $D_x\psi_1$ or $D_x\psi_2$ is nonzero.

Noting that $\det D_x\psi = \|D_x\psi_1 \wedge D_x\psi_2\|$, we introduce the function

$$\Psi(\psi_1, \psi_2, x) = D_x\psi_1 \wedge D_x\psi_2.$$

We seek to apply the following consequence of the Sard-Smale theorem to the functional Ψ.

Theorem 28 (Theorem 5.4 of [17]). *Let Y, Z be separable Banach manifolds, $E \to Y \times Z$ a Banach space fiber bundle and $\Psi : Y \times Z \to E$ a smooth section. Suppose we have for all $(y, z) \in \Psi^{-1}(0)$*

1. *the differential $\nabla\Psi(y, z) : T_yY \times T_zZ \to E_{(y,z)}$ is surjective;*
2. *the partial derivative $\partial_z\Psi(y, z) : T_zZ \to E_{(y,z)}$ is Fredholm of index ℓ;*

Then for generic $y \in Y$, the set $\{z \in Z \mid \Psi(y, z) = 0\}$ is an ℓ-dimensional submanifold of Z.

We apply Theorem 28 with $Y = \left(C^k(\mathbb{T}^2, \mathbb{R}^2)\right)^2$, $Z = \mathbb{T}^2$, and $E = Y \times \Lambda^2(\mathbb{T}^2)$, where $\Lambda^2(\mathbb{T}^2)$ is the vector bundle of differential 2-forms on $\mathbb{T}^2$. Concretely, we identify $E \cong (C^k(\mathbb{T}^2, \mathbb{R}^2))^2 \times \mathbb{T}^2 \times \mathbb{R}$ using $T\mathbb{T}^2 \cong \mathbb{T}^2 \times \mathbb{R}^2$ and $\bigwedge^2(\mathbb{R}^2) \cong \mathbb{R}$. In particular, item 2 is always satisfied since $\partial_x\Psi(\psi_1, \psi_2, x) : T_x\mathbb{T}^2 \to E_{(\psi_1,\psi_2,x)}$ is a linear mapping between two finite-dimensional spaces.

It remains to check item 1. The derivative of Ψ acting on $(h_1, h_2, v) \in (C^k(\mathbb{T}^2, \mathbb{R}))^2 \times \mathbb{R}^2$ is given by

$$D_{(\psi_1,\psi_2,x)}\Psi(h_1, h_2, v) = D_xh_1 \wedge D_x\psi_2 + D_x\psi_1 \wedge D_xh_2 + (D_x^2\psi_1(v)) \wedge D_x\psi_2 + D_x\psi_1 \wedge (D_x^2\psi_2(v)).$$

It suffices to check surjectivity of $D_{(\psi_1,\psi_2,x)}\Psi$ at all (ψ_1, ψ_2, x) such that ψ_1, ψ_2 are Morse. With x fixed, by symmetry we can assume without loss of generality that $D_x\psi_1 \neq 0$. Set $h_1 = 0$, $v = 0$, and construct h_2 so that D_xh_2 is not parallel with $D_x\psi_1$. Then, $D_{(\psi_1,\psi_2,x)}\Psi(h_1, 0, 0) = D_x\psi_1 \wedge D_xh_2 \neq 0$, hence $D\Psi$ is surjective at (ψ_1, ψ_2, x). This completes the proof. □

6.3. Proof of Theorem 3.

Proof. For ease of notation and to avoid factors of 2π, we work below with the parameterization $\mathbb{T}^2 \cong [0, 2\pi)^2$. By Lemma 25, it suffices to check the transversality condition 19 for the function

$$\psi(x_1, x_2) = \begin{pmatrix} \sin x_1 + \sin(x_2 - x_1) \\ \sin x_2 + \sin(x_1 - x_2) \end{pmatrix}.$$

That is, we seek to show that the system $\det D_x\psi = 0$, $D_x \det D\psi = 0$ does not have any solutions. This system of equations is given by

$$\cos(x_1)\cos(x_2) + (\cos(x_1) + \cos(x_2))\cos(x_1 - x_2) = 0 \tag{20}$$

$$-\sin(x_1)\cos(x_2) - \sin x_1 \cos(x_1 - x_2) - (\cos x_1 + \cos x_2)\sin(x_1 - x_2) = 0 \tag{21}$$

$$-\sin(x_2)\cos(x_1) - \sin x_2 \cos(x_1 - x_2) + (\cos x_1 + \cos x_2)\sin(x_1 - x_2) = 0 \tag{22}$$

Adding (21) to (22) gives

$$\sin(x_1)\cos(x_2) + \sin x_1 \cos(x_1 - x_2) + \sin(x_2)\cos(x_1) + \sin x_2 \cos(x_1 - x_2) = 0. \tag{23}$$

Solving (23) and (20) for $\cos(x_1 - x_2)$ separately yields

$$\frac{1}{\sec(x_1) + \sec(x_2)} = \frac{\sin(x_1 + x_2)}{\sin(x_1) + \sin(x_2)};$$

division by $\cos x_1, \cos x_2$ is justified since, as one can easily check, no solution (x_1, x_2) can satisfy either of $\cos x_1 = 0, \cos x_2 = 0$. With some standard algebraic manipulations, this equation can be cast as

$$(\sin x_1 + \sin x_2)(1 - \sin x_1 \sin x_2) = 0.$$

If $\sin x_1 \sin x_2 = 1$, then $\sin x_1 = \sin x_2 = \pm 1$, which is inconsistent with (21). If $\sin x_1 + \sin x_2 = 0$, then $x_1 + x_2 = 2k\pi$ or $x_1 - x_2 = (2k+1)\pi$ for some $k \in \mathbb{Z}$. If $x_1 + x_2 = 2k\pi$, (20) will give us $\cos(2x_1) = -\frac{\cos x_1}{2}$. Plugging this into (22), we get $\sin(2x_1) = \frac{\sin x_1}{4}$. Since $\frac{\cos^2 x_1}{4} + \frac{\sin^2 x_1}{16} = 1$ is a contradiction, we deduce that no solution to (20),(21),(22) can satisfy $x_1 + x_2 = 2k\pi$. Similarly, one can rule out solutions satisfying $x_1 - x_2 = (2k+1)\pi$. This completes the proof. □

Acknowledgements. Alex Blumenthal was supported by the National Science Foundation under Award No. DMS-1604805. Jinxin Xue is supported by NSFC (Significant Project No.11790273) in China and Beijing Natural Science Foundation (Z180003). Yun Yang is supported by the National Science Foundation under Award No. DMS-2000167.

Appendix A: Version of the Singular Value Decomposition

Lyapunov exponents are asymptotic exponential growth rates of singular values. For this reason, we recall here some basic facts about the Singular Value Decomposition and related results used in this paper. Below, $d \geq 1$ and A is a $d \times d$ matrix. The *singular values* $\sigma_1(A) \geq \cdots \geq \sigma_d(A)$ are defined to be the eigenvalues of $\sqrt{A^\top A}$, listed in decreasing order and counted with multiplicity.

Theorem 29 (Singular Value Decomposition, Theorem 3.1.1 of [16]). *There exist orthonormal bases $\{e_1, \cdots, e_d\}$ and $\{e'_1, \cdots, e'_d\}$ of $\mathbb{R}^d$ with the property that*

$$Ae_i = \sigma_i(A) e'_i.$$

These bases are unique up to changes of sign and rearrangements of indices in case of repeated singular values (i.e., $\sigma_i(A) = \sigma_j(A)$ for some $i \neq j$).

Recall that $\{e_i\}$ is an (orthonormal) eigenbasis for $A^\top A$, while $\{e'_i\}$ is an appropriate ordering of an (orthonormal) eigenbasis for $AA^\top$.
The following characterization of singular values is also used.

Lemma 30 (Min-max Principles for Singular Values, Theorem 3.1.2 of [16]). *For all $1 \leq i \leq d$, we have that*

$$\prod_{j=1}^{k} \sigma_j(A) = \max_{\substack{E \subset \mathbb{R}^d \\ \dim E = i}} |\det(A|_E)|,$$

where $A|_E$ is regarded as a linear mapping $E \to A(E)$.

The characterization in Lemma 30 is central to the approach taken in this paper: it directly implies that to control $\prod_1^k \sigma_j(A)$, it suffices to control the volume growth of A along k-dimensional subspaces. Since Lyapunov exponents are asymptotic exponential growth rates of singular values, this motivates why we can control sums of the top-k Lyapunov exponents by studying the 'typical' rate at which k-dimensional volumes grow.
Lastly, we state the following corollary of Theorem 29, which we use in Lemma 22 to estimate singular directions. Below, E_0 is a k-dimensional subspace, $\alpha > 0$ is fixed, $\Pi_0 : \mathbb{R}^d \to E_0$ is orthogonal projection to E_0, and

$$\mathcal{C}_0 := \{v \in \mathbb{R}^d : \|(I - \Pi_0)v\| \le \alpha \|\Pi_0 v\|\}$$

is a cone of vectors roughly parallel to E_0.

Lemma 31. *Assume A is invertible, and has the property that for any ℓ-dimensional subspace $E \subset \mathcal{C}_0$, $\ell \le k$, we have that $A(E) \subset \mathcal{C}_0$ and $A^\top(E) \subset \mathcal{C}_0$. Then, $\exists 1 \le i_1 < \cdots < i_k \le d$ such that $e_{i_j}, e'_{i_j} \in \mathcal{C}_0$ for $1 \le j \le k$.*

Proof. Let $\mathcal{E}_0 := \{E \in \mathrm{Gr}_k(\mathbb{R}^m) : E \subset \mathcal{C}_0\}$ and observe that $\mathcal{E}_0$ is invariant under $B := A^\top A$, which we view as a (continuous) mapping $\mathrm{Gr}_k(\mathbb{R}^m) \to \mathrm{Gr}_k(\mathbb{R}^m)$. In the chart $\mathcal{U}_{E_0}$ (see Sect. 5.1), the set $\mathcal{E}_0$ is convex. Since $\mathcal{E}_0$ is also compact, we see by the Brouwer fixed point theorem that B must have a fixed point in $\mathcal{E}_0$, i.e., a k-dimensional subspace E for which $B(E) = E$. Since B is self-adjoint, it follows that B has k linearly independent, orthogonal eigenvectors in E, from which we conclude $\exists i_j$, $1 \le j \le k$, for which $e_{i_j} \in \mathcal{C}_0$. Applying the hypothesis to the span of e_{i_j} ($\ell = 1$), we conclude that $Ae_{i_j} = \sigma_{i_j}(A)e'_{i_j} \in \mathcal{C}_0$, from which it follows immediately that $e'_{i_j} \in \mathcal{C}_0$ as well (note $\sigma_i(A) \neq 0$ for all $1 \le i \le d$ if A is invertible). □

Appendix B: Proofs of Grassmanian Geometric Lemmas 19 and 21

Proof of Lemma 19. Given $E \in \mathrm{Gr}_k(\mathbb{R}^m)$, write $V = E^\perp$ and define $\mathcal{V} = (\mathcal{U}_E)^c$, which by point (A) at the beginning of Sect. 3 is the set of k-dimensional subspaces intersecting V nontransversally. We will describe $\mathcal{V}$ as the image of a fiber bundle $\mathcal{E}$, to be defined below, via a smooth mapping $\Phi : \mathcal{E} \to \mathrm{Gr}_k(\mathbb{R}^m)$. As we will show, $\dim \mathcal{E} < k(m-k) = \dim \mathrm{Gr}_k(\mathbb{R}^m)$, hence $\mathcal{V}$ can be covered by embedded submanifolds of dimension $< k(m-k)$.
To define Φ and $\mathcal{E}$, we first introduce some notation. Given $v \in \mathbb{R}^m$ let $I_v = \{S \in \mathrm{Gr}_k(\mathbb{R}^m) : v \in S\}$. Then, each $S \in I_v$ is uniquely specified by a corresponding $k-1$-dimensional subspace $S_v := S \cap \langle v\rangle^\perp = (I - \Pi_v)(S)$, where $\Pi_v : \mathbb{R}^m \to \langle v\rangle$ is the orthogonal projection. So, we can (canonically) identify $I_v \cong \mathrm{Gr}_{k-1}(\langle v\rangle^\perp)$. The latter is essentially $\mathrm{Gr}_{k-1}(\mathbb{R}^{m-1})$ and has dimension $(k-1)(m-1-(k-1)) = (k-1)(m-k)$. Let $\pi : \mathcal{E} \to \mathrm{Gr}_1(V)$ denote the fiber bundle over $\mathrm{Gr}_1(V)$ with fibers $\mathrm{Gr}_{k-1}(\langle v\rangle^\perp)$. Write elements of $\mathcal{E}$ as $(v, \hat{S})$, where $v \in V$, $\hat{S} \in \mathrm{Gr}_{k-1}(\langle v\rangle^\perp)$. We define $\Phi : \mathcal{E} \to \mathrm{Gr}_m(\mathbb{R}^k)$ to be the sum of subspaces

$$\Phi(v, \hat{S}) = \langle v\rangle + \hat{S}$$

in $\mathbb{R}^m$. Evidently, the image of Φ coincides with $\mathcal{V}$. Since $\dim \mathcal{E} = (k-1) + (k-1)(m-k) = (k-1)(m-(k-1))$, it follows that $\mathcal{V}$ can be covered by finitely many closed submanifolds of dimension $\le (k-1)(m-(k-1)) < k(m-k)$. □

In what follows, given a subspace $E \subset \mathbb{R}^m$, we write $\Pi_E : \mathbb{R}^m \to E$ for its corresponding orthogonal projection.

Proof of Lemma 21. Let $E, E' \in \mathrm{Gr}_k(\mathbb{R}^m)$. Then, $d_{geo}(E, E') = \sqrt{\psi_1^2 + \cdots + \psi_k^2}$ where each $\psi_i = \psi_i(E, E') \in [0, \pi/2]$ is the i-th *Jordan angle* between E, E', defined by, e.g.,

$$\cos \psi_i = \min_{\substack{P \subset E \\ \dim P = i}} \max_{\substack{v \in P \\ \|v\|=1}} \max_{\substack{w \in E' \\ \|w\|=1}} \langle v, w \rangle \tag{24}$$

(see Proposition 3(b) of [22]). We have $\psi_1 \le \psi_2 \le \cdots \le \psi_k$, hence $\psi_k \le d_{geo}(E, E') \le k\psi_k$.

To connect this with the Hausdorff metric, by [19] Theorem I-6.34 and some elementary arguments, we have

$$\begin{aligned} d_H(E, E') &= \|(I - \Pi_{E'})\Pi_E\| = \sup_{v \in E} \|(I - \Pi_{E'})v\| = \sup_{v \in E} d(v, E') \\ &= \sin \angle(v, \Pi_{E'} v). \end{aligned}$$

On the other hand, by (24), we have

$$\max_{\substack{w \in E' \\ \|w\|=1}} \langle v, w \rangle = \left\langle v, \frac{\Pi_{E'} v}{\|\Pi_{E'} v\|} \right\rangle = \cos \angle(v, \Pi_{E'} v),$$

hence $\psi_k = \max_{v \in E, \|v\|=1} \angle(v, \Pi_{E'} v)$. We conclude, then, that $d_H(E, E') = \sin \psi_k$. In particular, $\frac{2}{\pi}\psi_k \le d_H(E, E') \le \psi_k$. This completes the proof. □

Appendix C: Proof of Proposition 6: Explicit Noise Model R_ω Satisfying (E), (C) and (ND)

Below, we write $d = 2N$ for short. Throughout, $\{z_i\}_{i=1}^K$ is a set such that the balls $\{B_{1/20}(z_i)\}$ of radius $1/20$ centered at the z_i form an open cover of $\mathbb{T}^d$.

Proof. For $(z, E) \in \mathrm{Gr}_{d/2}(\mathbb{T}^d)$ fixed, we write

$$\Psi_{(z,E)} : \Omega_0 \to \mathrm{Gr}_{d/2}(\mathbb{T}^d), \quad \Psi_{(z,E)}(\omega) := (R_\omega z, D_z R_\omega(E)).$$

Here and throughout, elements $\omega \in \Omega_0$ are written $\omega = (v, (U^{(i)}))$. Observe that $\Psi = \Psi_{(z,E)}$ is a continuous mapping sending the origin to (z, E). From this, we see that (C) is a straightforward consequence of hypothesis (i). Condition (E) follows from Lemma 9, the hypotheses of which are guaranteed by (iii) and the fact $(z, E) \mapsto \Psi_{(z,E)}$ is continuous as a mapping from $\mathrm{Gr}_{d/2}(\mathbb{T}^d)$ into the space of continuous mappings $\Omega_0 \to \mathrm{Gr}_{d/2}(\mathbb{T}^d)$ in the compact-open topology.

It remains to check condition (ND). For this, by a compactness argument and the Constant Rank Theorem, it suffices to show that for $(z, E) \in \mathrm{Gr}_{d/2}(\mathbb{T}^d)$ fixed, the mapping $\Psi = \Psi_{(z,E)}$ is a submersion. To simplify the argument, we begin with the following observation regarding the 'upper triangular' structure of $D\Psi$: writing $v = (v_1, \cdots, v_d) \in \mathbb{R}^d$ and $z = (z_1, \cdots, z_d)$,[2] we have that

$$D_{(t,(U^{(i)}))}\Psi \left(\frac{\partial}{\partial v_i} \right) = \frac{\partial}{\partial z_i}.$$

[2] That is, we are abandoning for the moment the distinction between x and y coordinates in $\mathbb{T}^d = \mathbb{T}^{2N}$.

That is, varying v does not change at all the $\mathrm{Gr}_{d/2}(\mathbb{R}^d)$ coordinate in the image. Therefore, it suffices to show that when v is held fixed, we have that

$$(U^{(i)}) \mapsto D_z R_{(v,(U^{(i)}))}(E)$$

is a submersion $\mathrm{Skew}(d)^K \to \mathrm{Gr}_{d/2}(\mathbb{R}^d)$.
In fact, we will show that for each z, it suffices to consider tangent directions corresponding to a single $U^{(i)}$. To see this, we make the following claim.

Claim 32. There exists $c = c_{K,d} > 0$ with the following property. For any $z \in \mathbb{T}^d$ there exists $1 \le j \le K$ such that for any $(U^{(i)}) \in \mathrm{Skew}(d)^K$, $\|(U^{(i)})\| \le c$, we have that

$$d(\Phi^{(j-1)}_{U^{(j-1)}} \circ \cdots \circ \Phi^{(1)}_{U^{(1)}} z, z_j) \le \frac{1}{10}.$$

Indeed, the claim holds with j any index for which $d(z, z_j) \le 1/20$ (see (9)), assuming $c = c_{K,d}$ is taken small enough.
With (z, E) and the above value of j fixed, we now set about checking that

$$U^{(j)} \mapsto D_z R_{(t,U^{(i)})}(E)$$

is a submersion. Since T_v, $\Phi^{(i)}_{U^{(i)}}$ are all diffeomorphisms of $\mathbb{T}^d$, it suffices to check that $U^{(j)} \mapsto D_{z'}\Phi^{(j)}_{U^{(j)}}(E')$ is a submersion $\mathrm{Skew}(d) \to \mathrm{Gr}_{d/2}(\mathbb{T}^d)$, where $z' = \Phi^{(j-1)}_{U^{(j-1)}} \circ \Phi^{(1)}_{U^{(1)}}(z)$ and $E' = D_z\Phi^{(j-1)}_{U^{(j-1)}} \circ \Phi^{(1)}_{U^{(1)}}(E)$. By our choice of j, Claim 32 ensures $d(z', z_j) \le 1/10$, hence $D_{z'}\Phi^{(j)}_{U^{(j)}} = \exp(U^{(j)})$. In view of the composition

$$U^{(j)} \mapsto \exp(U^{(j)}) \mapsto \exp(U^{(j)})(E')$$

and the fact that $U \mapsto \exp(U)$ is a local diffeomorphism $\mathrm{Skew}(d) \to \mathcal{O}(d)$, it suffices to check that $O \mapsto O(E)$ is a submersion $\mathcal{O}(d) \mapsto \mathrm{Gr}_{d/2}(\mathbb{R}^d)$. Since surjectivity of a derivative is an open property, it suffices to check that the differential of $O \mapsto O(E)$ is a submersion at the identity $\mathrm{Id} \in \mathcal{O}(d)$. □

Claim 33. Fix $1 \le k \le d$ and $E_0 \in \mathrm{Gr}_k(\mathbb{R}^d)$. Define $\Xi : \mathcal{O}(d) \to \mathrm{Gr}_k(\mathbb{R}^d)$, $\Xi(O) := O(E_0)$. Then, $D_{\mathrm{Id}}\Xi : \mathrm{Skew}(d) \to T_{E_0}\mathrm{Gr}_k(\mathbb{R}^d)$ is surjective.

Proof of Claim. We evaluate the differential explicitly in coordinates. Recall the chart $\mathcal{U}_{E_0} \cong L(E_0, E_0^\perp)$ for $\mathrm{Gr}_k(\mathbb{R}^d)$ at E_0. As one can check, in this chart, $\Xi(O) = O(E_0)$ is represented as

$$\Xi(O) = \mathrm{graph}_{E_0} g(O), \quad g(O) := (\mathrm{Id} - \Pi_{E_0})O(\Pi_{E_0} O|_{E_0})^{-1}.$$

Therefore, in these coordinates we have (writing $\Pi_{E_0} = \Pi$, $\Pi^\perp = \mathrm{Id} - \Pi_{E_0}$)

$$D_O g(U) = \Pi^\perp U(\Pi O|_{E_0})^{-1} + \Pi^\perp O(\Pi O|_{E_0})^{-1} U(\Pi O|_{E_0})^{-1}$$

for $U \in T_O\mathcal{O}(d)$. Evaluating at $O = \mathrm{Id}$, we see that

$$D_{\mathrm{Id}} g(U) = \Pi^\perp U|_{E_0}$$

This is clearly surjective as a linear mapping $\mathrm{Skew}(d) \mapsto L(E_0, E_0^\perp)$; given an arbitrary $B \in L(E_0, E_0^\perp)$, we have $D_{\mathrm{Id}} g(U) = B$ for any U of the form

$$U = \begin{pmatrix} \Pi U|_{E_0} & \Pi U|_{E_0^\perp} \\ \Pi^\perp U|_{E_0} & \Pi^\perp U|_{E_0^\perp} \end{pmatrix} = \begin{pmatrix} * & -B^\top \\ B & * \end{pmatrix}.$$

□

References

1. Barreira, L., Pesin, Y.B.: Lyapunov Exponents and Smooth Ergodic Theory, vol. 23. American Mathematical Society, Providence (2002)
2. Berger, P., Carrasco, P.: Non-uniformly hyperbolic diffeomorphisms derived from the standard map. Commun. Math. Phys. **329**, 239–262 (2014)
3. Berger, P., Turaev, D.: On Herman's positive entropy conjecture. Adv. Math. **349**, 1234–1288 (2019)
4. Blumenthal, A., Xue, J.X., Young, L.S.: Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. Ann. Math. **185**, 285–310 (2018)
5. Boffetta, G., del Castillo-Negrete, D., López, C., Pucacco, G., Vulpiani, A.: Diffusive transport and self-consistent dynamics in coupled maps. Phys. Rev. E **67**(2), 026–224 (2003)
6. Breden, M., Engel, M.: Computer-assisted proof of shear-induced chaos in stochastically perturbed hopf systems. (2021). arXiv preprint arXiv:2101.01491
7. Carrasco, P.D.: Random products of standard maps. (2019). https://arxiv.org/abs/1705.09705
8. Chirikov, B., Shepelyansky, D.: Chirikov standard map. Scholarpedia **3**(3), 3550 (2008)
9. Chirikov, B.V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**(5), 263–379 (1979)
10. Duarte, P.: Plenty of elliptic islands for the standard family of area preserving maps. In: Annales de l'IHP Analyse nonlinéaire, vol. 11, pp. 359–409 (1994)
11. Furstenberg, H.: Noncommuting random products. Transactions of the American Mathematical Society, pp. 377–428 (1963)
12. Galatolo, S., Monge, M., Nisoli, I.: Existence of noise induced order, a computer aided proof. Nonlinearity **33**(9), 4237 (2020)
13. Gol'dsheid, I.Y., Margulis, G.A.: Lyapunov indices of a product of random matrices. Russ. Math. Surv. **44**(5), 11–71 (1989)
14. Gorodetski, A.: On stochastic sea of the standard map. Commun. Math. Phys. **309**(1), 155–192 (2012)
15. Guivarch, Y., Raugi, A.: Products of random matrices: convergence theorems. Random Matrices and Their Applications (Brunswick, Maine, 1984) 31–54. Contemp. Math. **50**, 18
16. Horn, R.A., Johnson, C.R.: Topics in matrix analysis. Bull. Am. Math. Soc. **27**, 191–198 (1992)
17. Hutchings, M.: Lecture notes on Morse homology (with an eye towards floer theory and pseudoholomorphic curves) (2002)
18. Kantz, H., Grassberger, P.: Internal Arnold diffusion and chaos thresholds in coupled symplectic maps. J. Phys. A: Math. Gen. **21**(3), 127–133 (1988)
19. Kato, T.: Perturbation Theory for Linear Operators, vol. 132. Springer, Berlin (2013)
20. Kifer, Y.: Random perturbations of dynamical systems. Nonlinear Problems in Future Particle Accelerators. World Scientific, pp. 189 (1988)
21. Manos, T., Skokos, C., Bountis, T.: Global Dynamics of Coupled Standard Maps, Chaos in Astronomy, pp. 367–371. Springer, Berlin (2008)
22. Neretin, Y.A.: On Jordan angles and the triangle inequality in Grassmann manifolds. Geom. Dedic. **86**(1–3), 81–91 (2001)
23. Nicolaescu, L.I.: Lectures on the Geometry of Manifolds. World Scientific, Singapore (2007)
24. Piccione, P., Tausk, D.V. et al.: On the geometry of Grassmannians and the symplectic group: the Maslov index and its applications. UFF (2000)
25. Virtser, A.D.: On products of random matrices and operators. Theory Probab. Appl. **24**(2), 367–377 (1980)
26. Wilkinson, A.: What are Lyapunov exponents, and why are they interesting? Bull. Am. Math. Soc. **54**(1), 79–105 (2017)
27. Wood, B.P., Lichtenberg, A.J., Lieberman, M.A.: Arnold diffusion in weakly coupled standard maps. Phys. Rev. A **42**(10), 58–85 (1990)
28. Yang, H., Radons, G.: Dynamical behavior of hydrodynamic Lyapunov modes in coupled map lattices. Phys. Rev. E **73**(1), 016208 (2006)
29. Yang, H., Radons, G.: Lyapunov modes in extended systems. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **367**(1901), 3197–3212 (2009)
30. Young, L.-S.: Mathematical theory of Lyapunov exponents. J. Phys. A: Math. Theor. **46**(25), 254001 (2013)