The Annals of Probability
2023, Vol. 51, No. 4, 1559–1601
https://doi.org/10.1214/23-AOP1627
© Institute of Mathematical Statistics, 2023

EXPONENTIAL MIXING FOR RANDOM DYNAMICAL SYSTEMS AND AN EXAMPLE OF PIERREHUMBERT

BY ALEX BLUMENTHAL[1,a], MICHELE COTI ZELATI[2,b] AND RISHABH S. GVALANI[3,c]

[1]*School of Mathematics, Georgia Institute of Technology,* [a]*ablumenthal6@gatech.edu*
[2]*Department of Mathematics, Imperial College London,* [b]*m.coti-zelati@imperial.ac.uk*
[3]*Max-Planck-Institut für Mathematik in den Naturwissenschaften,* [c]*gvalani@mis.mpg.de*

We consider the question of exponential mixing for random dynamical systems on arbitrary compact manifolds without boundary. We put forward a robust, dynamics-based framework that allows us to construct space-time smooth, uniformly bounded in time, universal exponential mixers. The framework is then applied to the problem of proving exponential mixing in a classical example proposed by Pierrehumbert in 1994, consisting of alternating periodic shear flows with randomized phases. This settles a longstanding open problem on proving the existence of a space-time smooth (universal) exponentially mixing incompressible velocity field on a two-dimensional periodic domain while also providing a toolbox for constructing such smooth universal mixers in all dimensions.

1. Introduction. Mixing by incompressible flows is a fundamental stabilization mechanism in fluid mechanics, usually associated with the transfer of energy from large to small spatial scales, in a manner that is conservative and reversible for finite times but results in an irreversible loss of information in the long-time limit. In its simplest mathematical setting, this phenomenon is often associated with the study of the long-time dynamics of the transport equation

$$\partial_t \rho + u \cdot \nabla \rho = 0, \tag{1.1}$$

posed in a d-dimensional periodic domain $\mathbb{T}^d$ parametrized by $[0, 2\pi)^d$. The unknown $\rho = \rho(t, x) : [0, \infty) \times \mathbb{T}^d \to \mathbb{R}$ is advected by a specified divergence-free velocity field $u = u(t, x) : [0, \infty) \times \mathbb{T}^d \to \mathbb{R}^d$ and can be assumed, without loss of generality, to be mean-free, namely

$$\int_{\mathbb{T}^d} \rho(t, x)\, \mathrm{d}x = 0, \quad \text{for all } t \geq 0. \tag{1.2}$$

In applications, the unknown ρ denotes the concentration of a scalar (for, e.g., a chemical in water), that passively evolves according to an assigned stirring mechanism. A commonly used measure of the mixing reached by ρ at a given time is in terms of the homogeneous negative Sobolev norms

$$\|\rho(t)\|_{\dot{H}^{-s}}^2 = \sum_{k \in \mathbb{Z}^d \setminus \{0\}} |k|^{-2s} |\rho_k(t)|^2, \tag{1.3}$$

where the ρ_k's are the Fourier coefficients of ρ. This idea was introduced in [42], and then revisited in [38], for different values of the parameter $s > 0$. Decay of (1.3) as $t \to \infty$ implies that the L^2 mass of $\rho(t)$ is being evacuated to high modes, signaling mixing.

Received June 2022; revised March 2023.
MSC2020 subject classifications. Primary 35Q49; secondary 37H05, 37A25, 76F25.
Key words and phrases. Exponential mixing, Lyapunov exponents, random shear flows.

The questions of how fast can the quantity in (1.3) decay (given certain constraints on u), and which velocity fields u can achieve such decay, have attracted considerable attention recently. It is a standard fact that for an incompressible velocity field $u \in L_t^\infty W_x^{1,\infty}$, the quantity (1.3) cannot decay faster than exponentially. For velocity fields in $L_t^\infty W_x^{1,p}$, $p \in (1, \infty)$, exponential lower bounds for (1.3) were established in [22]; see also [32, 50]. The case $p = 1$, corresponding to Bressan's rearrangement cost conjecture [15], is still open.

The main focus of this work is providing *upper* bounds for (1.3). Given a mean-free initial datum $\rho(0, \cdot)$, it was shown in [58] that there exists a velocity field $u \in L_t^\infty W_x^{1,p}$, for some $p \in (2, 3)$, that mixes the corresponding solution exponentially fast. Different regularity constraints on u were then studied in depth in [1]. It is important to notice that in these works, the velocity field u heavily depends on the initial datum. The natural question of finding a universal mixer, namely a velocity field that mixes all mean-free initial data (with a certain regularity at least), was settled in [28]. As in [58], such velocities have only limited space regularity, and they are quite complicated (for instance, they are not time-periodic).

In [8], it was shown that one can construct universal mixers from a variety of stochastic evolution equations on incompressible velocity fields, for example, the Navier–Stokes equations with white-in-time, spatially regular driving. While the techniques of that paper produce spatially smooth mixers, the time regularity is at best C^k for some finite k, and the spatial regularity is not uniformly bounded in time. We note as well that these flows are not periodic-in-time.

In this paper, we establish the existence of space-time smooth, uniformly bounded in time, universal exponential mixers on periodic domains. This manuscript contains:

(a) a robust, dynamics-based framework that allows us to construct smooth, uniformly bounded in time, universal exponential mixers on arbitrary compact boundaryless manifolds (Sections 2, 3 and 4); and

(b) an application of this framework to the classical two-dimensional example of Pierrehumbert [45], consisting of alternating periodic shear flows with randomized phases (Section 5).

Mixing properties of the Pierrehumbert model have been studied extensively in the applied and computational literature (see, e.g., [46, 53, 54] and references therein). Numerical evidence suggests it is a universal exponential mixer [17], although to the authors' best knowledge ours is the first rigorous proof of this assertion.

1.1. *The Pierrehumbert model.* A natural candidate for an exponentially mixing flow on the two-dimensional torus $\mathbb{T}^2$ was introduced by Pierrehumbert in [45]. In this model, the velocity field u alternates, after every time interval of size $\tau > 0$, between two transversal shears with a randomly and independently chosen phase shift. More precisely, let $\{\omega_j = (\omega_j^1, \omega_j^2)\}_{j\in\mathbb{N}} \subset \mathbb{R}^2$ be a sequence of i.i.d. random variables uniformly distributed in $[0, 2\pi)^2$. For $(t, x) \in [0, \infty) \times \mathbb{T}^2$, we define

$$u(t, x_1, x_2) := \begin{pmatrix} \sin(x_2 - \omega_n^1) \\ 0 \end{pmatrix} \tag{1.4}$$

if $t \in [(2n-2)\tau, (2n-1)\tau)$ for some $n \in \mathbb{N}$, and

$$u(t, x_1, x_2) := \begin{pmatrix} 0 \\ \sin(x_1 - \omega_n^2) \end{pmatrix} \tag{1.5}$$

for $t \in [(2n-1)\tau, 2n\tau)$.

Intuitively, we alternate between the two sinusoidal shears, starting without loss of generality with a horizontal shear, while simultaneously, randomly and independently, changing

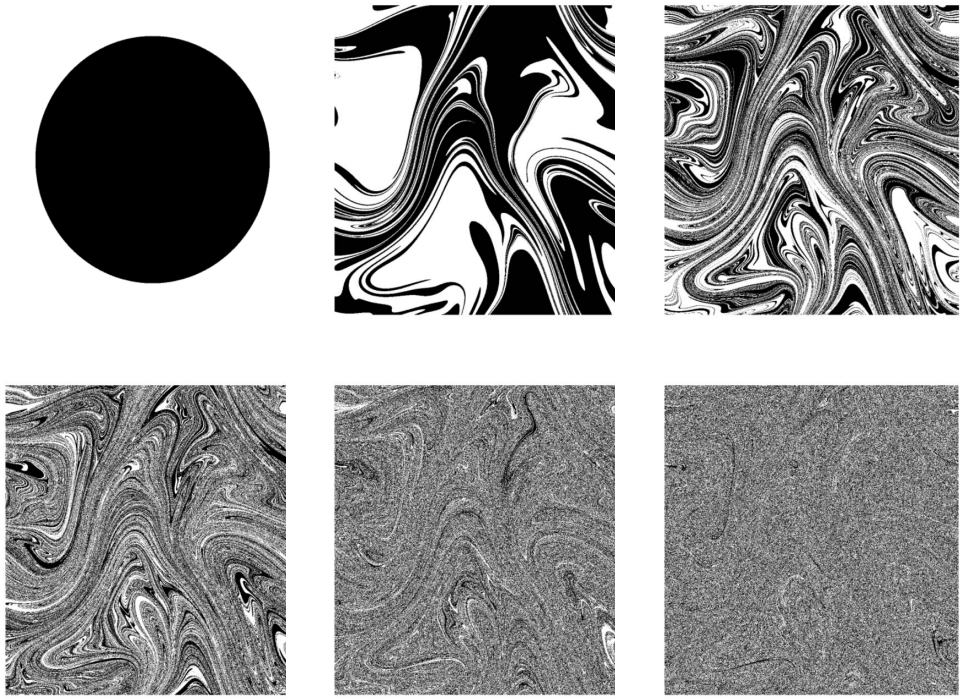

FIG. 1. *Mixing in the Pierrehumbert model with $\tau = 0.05$: snapshots at increasing times of the passive scalar concentration, starting from the initial datum on the top left-hand corner and moving from left to right.*

the phases of each of the shears (as depicted in Fig. 1). The intriguing fact is that, while shear flows on the torus are mixing with at most at an asymptotic rate of $1/\sqrt{t}$ [10], the combination of vertical and horizontal shearing can speed up the mixing considerably. The dynamics and mixing properties of u are best understood at the Lagrangian level. Let $X(t,x) \in \mathbb{T}^2$ denote the position at time $t \geq 0$ of a passive tracer advected by u and started at $X(0,x) = x \in \mathbb{T}^2$. Then it is well known that $X(t,x)$ is a solution to the following ODE:

$$\partial_t X(t,x) = u\big(t, X(t,x)\big), \quad X(0,x) = x. \tag{1.6}$$

For $\beta \in [0, 2\pi)$ and $x \in \mathbb{T}^2$, we define

$$f_\beta^H(x) := \begin{pmatrix} x_1 + \tau \sin(x_2 - \beta) \\ x_2 \end{pmatrix}, \qquad f_\beta^V(x) := \begin{pmatrix} x_1 \\ x_2 + \tau \sin(x_1 - \beta) \end{pmatrix}. \tag{1.7}$$

Then, for a given tuple of random phases $\omega = (\omega^1, \omega^2) \in [0, 2\pi)^2$, the position $X(2\tau, x) \in \mathbb{T}^2$ of the particle at time $t = 2\tau$ is given by $X(2\tau, x) = f_\omega(x)$, where

$$f_\omega(x) := f_{\omega^2}^V \circ f_{\omega^1}^H(x). \tag{1.8}$$

Now, writing any sequence of possible random phase shifts as

$$\underline{\omega} = (\omega_1, \omega_2, \ldots) \in \Omega := \big([0, 2\pi)^2\big)^{\mathbb{N}}, \tag{1.9}$$

we can recover the motion of tracer particle at times $2n\tau$ by just iterating the maps given by f_{ω_i}. That is to say, $X(2\tau n, x) = f_{\underline{\omega}}^n(x)$ where

$$f_{\underline{\omega}}^n(x) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x). \tag{1.10}$$

One of the main results of this article is the following correlation decay estimate for the Pierrehumbert model.

THEOREM 1.1. *Let $q, s > 0$. There exists a function $D : \Omega \to [1, \infty)$ and a constant $\alpha > 0$ such that for all mean-free functions $\varphi, \psi \in H^s(\mathbb{T}^2)$, we have the almost sure estimate*

$$\left| \int_{\mathbb{T}^2} \varphi(x)\psi \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}x \right| \le D(\underline{\omega})\mathrm{e}^{-\alpha n}\|\varphi\|_{H^s}\|\psi\|_{H^s}, \tag{1.11}$$

while the function D satisfies the moment bound $\mathbb{E}|D|^q < \infty$.

The exponential mixing property is encoded in the quenched correlation decay estimate (1.11). By duality, this implies the decay of the negative Sobolev norms (1.3) of the solution ρ to (1.1) with velocity field u whose components are given by (1.4)–(1.5).

REMARK 1.2. While the velocity u as in (1.4)–(1.5) is not continuous in time, it is straightforward to build a continuous or even uniformly smooth-in-time version via an appropriate time-reparametrization, as described in [58], Remark on p. 1914. Therefore, Theorem 1.1 tell us that typical (i.e., almost all) realizations of the Pierrehumbert model are smooth-in-time, uniformly spatially C^∞-in-time universal mixers. To our knowledge, this is the first construction of such a mixer on the periodic box $\mathbb{T}^2$. On the other hand, the mixer so-obtained is not time-periodic. Indeed, the ability to choose i.i.d. phase shifts is essential to our approach below. Establishing universal mixing for the Pierrehumbert model using fixed, deterministic phase shifts are extremely difficult and are related to notorious open problems in smooth ergodic theory; see Section 1.3 below.

1.2. *A general dynamical system framework.* The Pierrehumbert model is just one instance of a general class of random dynamical systems obtained as in (1.10). Given a finite-dimensional compact Riemannian manifold without boundary M, and a probability space of the form $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_0, \mathcal{F}_0, \mathbb{P}_0)^{\mathbb{N}}$, we consider the random compositions of differentiable functions $f_\omega : M \to M$ with $\omega \in \Omega_0$, namely

$$f^n_{\underline{\omega}} = f_{\omega_n} \circ \cdots \circ f_{\omega_1}, \quad n \in \mathbb{N}, \tag{1.12}$$

where $\underline{\omega} = (\omega_1, \omega_2, \ldots) \in \Omega$. Typically in this paper, we will assume Ω_0 is a manifold, for example, $\Omega_0 = \mathbb{R}^k$, $k \ge 1$ and that $\mathbb{P}_0 \ll$ Lebesgue measure. For the sake of this discussion, we will assume that almost surely $f_\omega : M \to M$ preserves Lebesgue measure π on M.

Markov chains arise from random dynamical systems by defining their transition kernels, for any set $A \in \mathrm{Bor}(M)$, as

$$P(x, A) = \mathbb{P}_0\big(f_\omega(x) \in A\big), \qquad P^{n+1}(x, A) = \int P^n(y, A)P(x, \mathrm{d}y), \tag{1.13}$$

and the corresponding transition operator or Markov semigroup, acting on a continuous, bounded function $\varphi : M \to \mathbb{R}$, as

$$P^n\varphi(x) = \int_M \varphi(y)P^n(x, \mathrm{d}y). \tag{1.14}$$

There are several ingredients, which are required for proving exponential mixing for random dynamical systems and they can be generically identified with nondegeneracy conditions on f_ω and P, and various processes associated to them. While we will not discuss them in detail in this Introduction, the following two features are absolutely necessary.

(a) *The kernel P is uniformly geometrically ergodic.* As f_ω preserves Lebesgue measure π almost surely, it follows that π is stationary, that is,

$$\pi(A) = \int P(x, A)\,\mathrm{d}\pi(x), \quad \forall A \in \mathrm{Bor}(M). \tag{1.15}$$

We say that P is uniquely geometrically ergodic if π is the unique stationary measure and is uniformly exponentially attracting, that is, $P^n\varphi(x) \to \int \varphi\,\mathrm{d}\pi$ as $n \to \infty$ exponentially fast for all $\varphi : M \to \mathbb{R}$ continuous and bounded.

(b) *Positivity of the top-Lyapunov exponent.* This implies that almost everywhere in M and with probability 1, nearby particles are pushed away at an exponentially fast rate. This is phrased in terms of the Jacobian $D_x f^n_{\underline{\omega}}$ as the condition that the limit

$$\lambda_1 := \lim_{n\to\infty} \frac{1}{n} \log\left|D_x f^n_{\underline{\omega}}\right| \tag{1.16}$$

exists and is greater than 0 for $\mathbb{P} \times \pi$-a.e. $(\underline{\omega}, x)$. We note here that the limit (1.16) exists and is almost surely constant over $(\underline{\omega}, x)$ under condition (a) (Theorem 2.7 below).

Note that (a) differs from the *almost sure* mixing we wish to prove; instead, (a) refers to mixing of the RDS averaged over all possible noise realizations, that is, *annealed mixing*.

Unfortunately, conditions (a) and (b) mentioned above are far from being sufficient for exponential mixing; in particular, in the method pursued in this manuscript, we will also need unique geometric ergodicity for the so-called projective process

$$\hat{f}_\omega(x, v) := \left(f_\omega(x), \frac{D_x f_\omega v}{|D_x f_\omega v|}\right), \tag{1.17}$$

defined the sphere bundle of M, and the two-point chain with transition kernel

$$P^{(2)}\big((x, y), K\big) = \mathbb{P}_0\big(\big(f_\omega(x), f_\omega(y)\big) \in K\big), \tag{1.18}$$

defined on the product space $M \times M \setminus \Delta$, with $\Delta := \{(x, x) : x \in M\}$. Precise definitions of these processes and the conditions required on them are stated in Section 4. The consequence of the above discussion is the following theorem, which we state here informally.

THEOREM 1.3. *Assume that the dynamical system generated by f_ω is uniquely and uniformly geometrically ergodic, along with the corresponding projective and two-point processes, and possesses a positive Lyapunov exponent. For any $q, s > 0$, there exists a function $D : \Omega \to [1, \infty)$ and a constant $\alpha > 0$ such that for all mean-free functions $\varphi, \psi \in H^s(M)$, we have the almost sure estimate*

$$\left|\int_{\mathbb{T}^2} \varphi(x)\psi \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}x\right| \le D(\underline{\omega})\mathrm{e}^{-\alpha n}\|\varphi\|_{H^s}\|\psi\|_{H^s}, \tag{1.19}$$

while the function D satisfies the moment bound $\mathbb{E}|D|^q < \infty$.

For a summary of the precise conditions under which Theorem 1.3 holds, see Section 4.3.

1.3. *Discussion.*

Mixing for deterministic systems. For a volume-preserving dynamical system, random or deterministic, a positive Lyapunov exponent as in equation (1.16) implies that time-n linearizations exhibit exponential stretching and contracting in tangent space. In the dynamics community, this stretching/contracting is called hyperbolicity, and has long been understood as the primary mechanism responsible for rapid mixing. Systems for which the realization of hyperbolicity is uniform over phase space are called *Anosov maps* or *flows*, and by now they are well-known, checkable conditions for showing that Anosov systems are universal exponential mixers; for discrete-time maps, see, for example, [5, 14, 39], and for the (considerably more subtle) case of continuous-time flows see, for example, [26, 40]. Exponential mixing for hyperbolic systems is a vast subject to which we cannot do justice here.

In applications, the Anosov property is often far too restrictive; for instance, there are strong limitations on which phase spaces admit Anosov maps or flows [30, 51], and those that exist are subject to strong limitations [41]. It is natural instead to study *nonuniformly hyperbolic systems*, that is, those for which the realization of stretching/contracting is nonuniform

over phase space; see, for example, the survey [61]. This motivates the use of the Lyapunov exponent as in (1.16); a system is called *nonuniformly hyperbolic* if the value (1.16) exists and is strictly positive at a positive-volume subset of initial $x \in M$. While existence of the limit follows from standard ergodic theory tools (Theorem 2.7 below), the limit need not be uniform over phase space, hence the term "nonuniform." Nonuniform hyperbolicity is expected to be the mechanism behind typical incompressible mixers [44], for example, via the much-discussed stretch-and-fold mechanism [18].

We note that there exist several sets of theoretical tools for establishing exponential mixing for nonuniformly hyperbolic systems, for example, the framework of Young towers [59, 60]. Unfortunately, checking the conditions of these frameworks for systems of practical interest is notoriously difficult [44]. For instance, it is an outstanding open question [24] to verify nonuniform hyperbolicity for the Chirikov standard map [19], a deterministic, discrete-time toy model of the stretch-and-fold mechanism in Lagrangian flow [23].

Mixing for noisy systems. Establishing nonuniform hyperbolicity is far more tractable in the presence of nondegenerate noise; see, for example, [11] and the discussion and references within. In our setting, we take advantage of Furstenberg's criterion [29], which in its original form gives checkable conditions for the Lyapunov exponent of an i.i.d. product of determinant 1 matrices to be strictly positive. In our setting, an appropriate extension of Furstenberg's criterion implies that under a suitable set of nondegeneracy conditions on the law of a volume-preserving RDS $f_\omega, \omega \sim \mathbb{P}_0$, the Lyapunov exponent is positive (Section 3.2). There is a long literature of results à là Furstenberg," for example, [3, 13, 16, 37, 48, 56]. In this manuscript, we owe much to the approach of [12] in particular.

To go from nonuniform hyperbolicity to correlation decay, the method we present here first appears in [27] (see also [4]), which establishes that with probability 1, the sequences $\varphi \circ f^n_{\underline{\omega}}$ satisfy a CLT when the $f^n_{\underline{\omega}}$ come from the time-n maps of a divergence-free stochastic differential equation satisfying some hypoellipticity conditions. The approach in [27] is based from a large deviations theory for the two-point Markov chain with kernel $P^{(2)}$ defined above, derived in [6].

A version of the methods of this manuscript was employed in [8, 9] to establish almost sure exponential mixing for the Lagrangian flow of stochastic Navier–Stokes on the periodic box with nondegenerate noise. Although many of the basic ideas are the same, there are several significant differences between this and the approach of this manuscript, the most significant being that [8, 9] is geared toward continuous time systems, while in the present manuscript we work with discrete time systems, entailing several technical problems to be overcome (see, e.g., Remark A.3).

Extensions and outlook. There are several ways to introduce randomness into this model, for example, distributing the random phases by a different probabilistic law aside from the uniform distribution used here, or randomizing the switching time τ between the horizontal and vertical shearing. The recent work of Cooperman [20] follows the latter approach, using the tools in the present manuscript while also overcoming the additional challenges created by the presence of almost surely stationary points in the resulting flow.

From a stochastic dynamics perspective, reducing randomness limits the number of degrees of freedom available for establishing controllability and irreducibility of the various Markov processes associated to the Pierrehumbert system. Hence, the fully deterministic case seems to be out of reach at the moment; compare the preceding discussion. An interesting open problem is the case in which the random phase sequence $\underline{\omega}$ is "one-dimensional," namely $\underline{\omega} = (\omega_1, \omega_2, \dots) \in [0, 2\pi)^{\mathbb{N}}$. This amounts to modifying (1.8) to $f_\omega = f^2_\omega \circ f^1_\omega$ for some random $\omega \in [0, 2\pi)$.

The numerical evidence from [17] suggests that randomizing phases and/or times is inessential, and exponential mixing is observed even when no randomness is included in

the model for certain sufficiently large values of τ. Along these lines, a previous version of this paper made the claim that if the shear time τ is chosen to be deterministic and sufficiently long, then the resulting Pierrehumbert model is exponentially mixing. This was shown to be false in a paper of Cooperman [20] published soon after this one; there exist arbitrarily large τ and deterministic choices of phase for which the Pierrehumbert model admits a neutral fixed point, which as it is shown there, is incompatible with exponential mixing.

Organization of the paper. We start in Section 2 with some background and preliminary results for Markov chains and random dynamical systems in an abstract setting, for example, Harris' theorem for verifying geometric ergodicity and Furstenberg's criterion for establishing positivity of Lyapunov exponents. Section 3 specializes to random dynamical systems subjected to absolutely continuous noise, establishing tools for checking the hypotheses of Harris' theorem and Furstenberg's criterion in this setting. Section 4 completes the proof of universal mixing in the abstract setting, and Section 5 applies the abstract result to the Pierrehumbert model.

2. Markov chains and random dynamical systems. Some of the concepts used in the sequel are best introduced in a general abstract setting. To fix notation in what follows, X denotes a complete metric space not necessarily compact, and $\mathrm{Bor}(X)$ is its Borel σ-algebra. Let $\mathcal{M}(X)$ denote the space of probability measures on X, endowed with either the weak* topology or the total variation metric

$$d_{\mathrm{TV}}(\mu,\nu) := \sup_{A\in\mathrm{Bor}(X)} \big|\mu(A)-\nu(A)\big|. \tag{2.1}$$

We begin by covering some preliminaries on random dynamical systems in an abstract setting.

2.1. *Markov chain preliminaries.* Assume that P is a Markov transition kernel on X, namely, for each $x \in X$ we have that $P(x,\cdot)$ is a Borel probability on X. For a $n \in \mathbb{N}$, iterates of P are defined inductively by the Chapman–Kolmogorov relation

$$P^{n+1}(x,A) = \int P(x,\mathrm{d}y)P^n(y,A). \tag{2.2}$$

Moreover, we will assume that P has the *Feller* property, that is, if $\varphi : X \to \mathbb{R}$ is a bounded, continuous function on X, then

$$x \mapsto P\varphi(x) := \int P(x,\mathrm{d}y)\varphi(y) \tag{2.3}$$

is continuous as well. That is to say if $x \mapsto P(x,\cdot)$ is continuous in the weak* topology on $\mathcal{M}(X)$, then P has the Feller property. We will also at times refer to the dual action P^* on probability measures. Given $\mu \in \mathcal{M}(X)$, we define

$$P^*\mu(A) := \int P(x,A)\,\mathrm{d}\mu(x), \quad \text{for } A \in \mathrm{Bor}(X). \tag{2.4}$$

A Borel probability measure $\pi \in \mathcal{M}(X)$ is called a *stationary* for P if it is a fixed point of P^*, that is,

$$\pi(A) = \int P(x,A)\,\mathrm{d}\pi(x), \quad \forall A \in \mathrm{Bor}(X). \tag{2.5}$$

A set $A \in \mathrm{Bor}(X)$ is (P,π)-invariant if $P\chi_A = \chi_A$ holds π-almost everywhere, where χ_A is the indicator function of A. We say that π is an *ergodic stationary measure* if all (P,π)-invariant sets have π-measure zero or one. In the setting described above, if π is the unique

stationary measure of a transition kernel P then it is automatically ergodic—this is a standard fact and follows from, for example, [34], Proposition I.2.1.

The following properties of Markov chains are commonly used in the arguments involving mixing in the subsequent sections.

DEFINITION 2.1. Let $n \geq 1$. We say that a set $A \subset X$ is P^n-*small* if there exists a positive measure ν_n on X such that, for all $x \in A$, we have that

$$P^n(x, B) \geq \nu_n(B) \quad \text{for all Borel } B \subset X.$$

We say that $A \subset X$ is small if it is P^n-small for some $n \geq 1$.

Lastly, we recall the following drift-type condition.

DEFINITION 2.2. We say that a function $V : X \to [1, \infty)$ satisfies a *Lyapunov–Foster drift condition* if there exist $\alpha \in (0, 1)$, $b > 0$ and a compact set $C \subset X$ such that

$$PV \leq \alpha V + b\chi_C. \tag{2.6}$$

Given such a V, we define the weighted norm

$$\|\varphi\|_V := \sup_{x \in X} \frac{|\varphi(x)|}{V(x)}, \tag{2.7}$$

and define $M_V(X)$ to be the space of measurable observables $\varphi : X \to \mathbb{R}$ such that $\|\varphi\|_V < \infty$.

2.2. *An abstract Harris' theorem.* For irreducible, aperiodic finite-state Markov chains, the Perron–Frobenius theorem asserts the unique existence of a stationary probability and a quantitative geometric rate of convergence to that stationary measure. An analogue for Markov chains on a complete metric space is Harris' theorem, formulated below in a version suitable for our purposes.

THEOREM 2.3 (Abstract Harris' theorem, [43]). *Let P be a Feller transition kernel and assume the following*:

(a) (*Small sets*) *There exists an* open *small set*.

(b) (*Topological irreducibility*) *For every $x \in X$ and nonempty open set $U \subset X$, there exists $N = N(x, U) \geq 1$ such that $P^N(x, U) > 0$.*

(c) (*Strong aperiodicity*) *There exists $x_* \in X$ such that for all open $U \ni x_*$, we have that $P(x_*, U) > 0$;*

(d) (*Drift condition*) *There exists a function $V : X \to [1, \infty)$ satisfying a Lyapunov–Foster drift condition for P.*

Then P is V-uniformly geometrically ergodic, that is, P admits a unique stationary measure π, and has the property that there exist $D > 0$ and $\gamma \in (0, 1)$ such that for all $x \in X$ and $\varphi \in M_V(X)$, we have

$$\left| P^n\varphi(x) - \int \varphi \, \mathrm{d}\pi \right| \leq DV(x)\|\varphi\|_V \gamma^n. \tag{2.8}$$

The geometric ergodicity statement (2.8) can be equivalently reformulated in terms of the operator norm of P^n, the conclusion being that $\|P^n - \pi\|_V \leq D\gamma^n$, where here we abuse notation and identify π with the operator $\varphi \mapsto \int \varphi d\pi$. While this result is essentially borrowed[1]

[1]While for the sake of consistency we have attempted to use the same terminology as [43], we caution that our definition of "strong aperiodicity" given in Theorem 2.3 above is not the same as that in [43]; see Appendix A for details.

from [43], it is not explicitly stated there, and is instead an assembly of several results found throughout that book along with some additional arguments. For the sake of completeness, a proof is sketched in Appendix A.

In the special case when X is compact, item (d) in Theorem 2.3 is redundant, as a Markov kernel P automatically satisfies a Lyapunov–Foster drift condition for the drift function $V \equiv 1$. In this case, we say that the Markov kernel P is *uniformly geometrically ergodic*.

REMARK 2.4. The version of Harris' theorem presented here does not explicitly indicate the dependence of the constants γ, D. It is entirely reasonable to expect, however, that these constants depend in an explicit way on the various "inputs" to the hypotheses of Theorem 2.3, for example, how large $N(x, U)$ can get depending on the size of the open neighborhood U. Several versions of Harris' theorem do exist, which track dependencies of this kind; see, for example, [7, 31]. On the other hand, these versions do not apply directly to the version of Harris' theorem presented here in Theorem 2.3. A "quantitative" version of Harris' theorem in the above setting will be included in a work in preparation by the first author and Christopher DuPre.

2.3. *Random dynamical systems*. In this manuscript, we will exclusively work with Markov chains derived from *random dynamical systems* (RDS) with independent increments, a framework, which we now describe. Let $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ be a fixed probability space. A continuous RDS with independent increments on X is an assignment to each $\omega \in \Omega_0$ of a continuous mapping $f_\omega : X \to X$, which satisfies the following mild measurability condition: for all $A \subset \mathrm{Bor}(X)$ and $x \in X$, the set $\{\omega \in \Omega_0 : f_\omega(x) \in A\}$ is $\mathcal{F}_0$-measurable. Defining $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_0, \mathcal{F}_0, \mathbb{P})^{\mathbb{N}}$, with elements $\underline{\omega} \in \Omega$ written $\underline{\omega} = (\omega_1, \omega_2, \ldots)$, we consider the random compositions of functions

$$f^n_{\underline{\omega}} = f_{\omega_n} \circ \cdots \circ f_{\omega_1}, \quad n \in \mathbb{N}, \tag{2.9}$$

following the convention that $f^0_{\underline{\omega}}$ is the identity mapping on X. As customary, $\theta : \Omega \to \Omega$ indicates the leftward shift on Ω, that is, if $\underline{\omega} = (\omega_1, \omega_2, \ldots)$ then $\theta\underline{\omega} = (\omega_2, \omega_3, \ldots)$. Note that θ is a measure-preserving transformation on $(\Omega, \mathcal{F}, \mathbb{P})$, that is, $\theta^{-1}\mathcal{F} \subset \mathcal{F}$ and $\mathbb{P} \circ \theta^{-1} = \mathbb{P}$, and moreover, $f^n_{\underline{\omega}}$ satisfies the *cocycle property*:

$$f^{n+m}_{\underline{\omega}} = f^n_{\theta^m \underline{\omega}} \circ f^m_{\underline{\omega}}, \quad \forall m, n \geq 0. \tag{2.10}$$

Continuous RDS in this framework naturally give, rise to Markov chains with kernels

$$P(x, A) = \mathbb{P}_0\big(f_\omega(x) \in A\big). \tag{2.11}$$

Continuity of the f_ω implies automatically that the transition kernel P is Feller. We also note that stationarity of $\pi \in \mathcal{M}(X)$ can be rewritten as

$$\pi(A) = \int \pi \circ f_\omega^{-1}(A)\, \mathrm{d}\mathbb{P}_0(\omega), \quad \forall A \in \mathrm{Bor}(X). \tag{2.12}$$

Informally, π is "f_ω-invariant on average." We will abuse terminology in what follows and describe a measure π as a *stationary measure for an RDS* $f^n_{\underline{\omega}}$ if it is a stationary measure for its corresponding Markov kernel P.

DEFINITION 2.5. We say that a bounded, measurable function $\varphi : X \to \mathbb{R}$ is (P, π)-*invariant* if $P\varphi = \varphi$ holds π-a.e. We say that π is *ergodic* if all (P, π)-invariant functions are π-almost surely constant.

2.3.1. *Linear cocycles over a random dynamical system.* Let $d \geq 2$ and let $GL_d(\mathbb{R})$ denote the space of invertible $d \times d$-matrices with real entries.

DEFINITION 2.6. Let $\mathcal{A} : \Omega_0 \times X \to GL_d(\mathbb{R})$ be a measurable mapping. The *linear cocycle* generated by $\mathcal{A}$ is the composition

$$\mathcal{A}^n_{\underline{\omega},x} := \mathcal{A}_{\omega_n, f^{n-1}_{\underline{\omega}}(x)} \circ \cdots \circ \mathcal{A}_{\omega_1,x} \in GL_d(\mathbb{R}) \tag{2.13}$$

for $n \geq 1$. We call $\mathcal{A}$ a *continuous linear cocycle* if $x \mapsto \mathcal{A}_{\omega,x}$ is continuous for $\mathbb{P}_0$-almost every $\omega \in \Omega_0$.

In many applications and when the $f^n_{\underline{\omega}}$ are smooth, as we will assume in the sequel, $\mathcal{A}^n_{\underline{\omega},x}$ is taken as $D_x f^n_{\underline{\omega}}$, the so-called *derivative cocycle* associated to the random compositions $(f^n_{\underline{\omega}})$. The following version of the multiplicative ergodic theorem (MET) stated below describes the asymptotic behavior of linear cocycles in this setting.

THEOREM 2.7 ([34], Theorem III.1.1). *Let $f^n_{\underline{\omega}}$ be a continuous RDS with ergodic stationary measure $\pi \in \mathcal{M}(X)$. Assume $\mathcal{A}$ satisfies the integrability condition*

$$\int \left(\log^+ |\mathcal{A}_{\omega,x}| + \log^+ |\mathcal{A}^{-1}_{\omega,x}|\right) \mathrm{d}\mathbb{P}_0(\omega)\, \mathrm{d}\pi(x) < \infty. \tag{2.14}$$

Then there exist $r \in \{1, \dots, d\}$, constant values $\lambda_1 > \cdots > \lambda_{r-1} > \lambda_r > -\infty$, and at $\mathbb{P} \times X$-almost every $(\underline{\omega}, x) \in \Omega \times X$, a filtration

$$\mathbb{R}^d =: F_1(\underline{\omega}, x) \supsetneq F_2(\underline{\omega}, x) \supsetneq \cdots \supsetneq F_r(\underline{\omega}, x) \supsetneq F_{r+1}(\underline{\omega}, x) := \{0\} \tag{2.15}$$

of subspaces with the property that

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{A}^n_{\underline{\omega},x} v| \tag{2.16}$$

for all $v \in F_i(\underline{\omega}, x) \setminus F_{i+1}(\underline{\omega}, x)$. Moreover, the assignments $(\underline{\omega}, x) \mapsto F_i(\underline{\omega}, x)$ vary measurably for all i, and $\dim F_i(\underline{\omega}, x)$ *is constant over $\mathbb{P} \times \pi$-typical $(\underline{\omega}, x) \in \Omega \times X$.*

The values λ_i are called *Lyapunov exponents* with *multiplicities*

$$m_i := \dim F_i(\underline{\omega}, x) - \dim F_{i+1}(\underline{\omega}, x), \tag{2.17}$$

while the filtration F_i is referred to as the *Oseledets splitting*.

REMARK 2.8. The proof of Theorem 2.7 due to Ragunathan [47] realizes the Lyapunov exponents as the distinct values among the values

$$\chi_i := \lim_{n\to\infty} \frac{1}{n} \log \sigma_i(\mathcal{A}^n_{\underline{\omega},x}). \tag{2.18}$$

Here, σ_i is the ith singular value of a matrix. The RHS limit exists and is constant (in particular deterministic/nonrandom) for $\mathbb{P} \times \pi$-a.e. $(\underline{\omega}, x)$ by an argument using the subadditive ergodic theorem [35]. In turn, the weights m_i are given by

$$m_i = \#\{1 \leq j \leq d : \chi_j = \lambda_i\}. \tag{2.19}$$

In view of the identity $|\det(B)| = \prod_{i=1}^d \sigma_i(B)$ for $d \times d$ matrices B, the weighted sum

$$\lambda_\Sigma := \sum_{i=1}^r m_i \lambda_i \in (-\infty, \infty) \tag{2.20}$$

is given by the limiting formula

$$\lambda_\Sigma = \lim_{n\to\infty} \frac{1}{n} \log|\det(\mathcal{A}^n_{\underline{\omega},x})| \tag{2.21}$$

for $\mathbb{P} \times \pi$-almost every $(\underline{\omega}, x)$. In particular, if $\det(\mathcal{A}_{\omega,x}) \equiv 1$, as is the case in our applications, then $\lambda_\Sigma = 0$ automatically.

Naturally associated to Lyapunov exponents is the *projectivized dynamics* of the cocycle on tangent directions. Given a linear cocycle over an RDS $f^n_{\underline{\omega}}$ generated by $\mathcal{A}$, we define the *projective RDS* $\widehat{f}^n_{\underline{\omega}}$ on $X \times S^{d-1}$ by

$$\widehat{f}_\omega(x, v) := \big(f_\omega(x), \widehat{\mathcal{A}}_{\omega,x}(v)\big), \qquad \widehat{f}^n_{\underline{\omega}} = \widehat{f}_{\omega_n} \circ \cdots \circ \widehat{f}_{\omega_1}. \tag{2.22}$$

Here and throughout, given an invertible $d \times d$ matrix B and $v \in S^{d-1}$ we write $\widehat{B} : S^{d-1} \to S^{d-1}$ for the mapping

$$\widehat{B}(v) = \frac{Bv}{|Bv|}. \tag{2.23}$$

As a continuous RDS, $\widehat{f}^n_{\underline{\omega}}$ gives rise to a corresponding Markov kernel $\widehat{P}$ on $X \times S^{d-1}$. Stationary measures $\nu \in \mathcal{M}(X \times S^{d-1})$ of $\widehat{P}$ are referred to as *projective stationary measures*, and describe the equilibrium statistics of tangent directions $\mathcal{A}^n_{\underline{\omega},x} v$ for $v \in \mathbb{R}^d$. Note that any projective stationary measure ν projects to a stationary measure for P, that is, the measure $\pi(K) := \nu(K \times S^{d-1})$ is a stationary measure for P.

At times, we will refer to *measurable families of measures* $(\nu_x)_{x\in S}$ for some measurable $S \subset X$. These are collections of measures ν_x on S^{d-1} with the property that $x \mapsto \nu_x(K)$ varies measurably in x for all fixed Borel $K \subset S^{d-1}$. We say that $(\nu_x)_{x\in S}$ is weak* continuous if $x \mapsto \int \psi(v)\,\mathrm{d}\nu_x(v)$ varies continuously in x for all fixed continuous $\psi : S^{d-1} \to \mathbb{R}$.

2.3.2. *Furstenberg's criterion*. Assume $\mathcal{A}^n_{\underline{\omega},x}$ is a linear cocycle over an RDS $f^n_{\underline{\omega}}$ on a complete metric space X and that the assumptions of Theorem 2.7 hold true. It is not hard to check that, without further conditions, we have

$$d\lambda_1 \geq \lambda_\Sigma. \tag{2.24}$$

Equality in this bound, namely $d\lambda_1 = \lambda_\Sigma$, would imply $r = 1$, that is, there is a sole Lyapunov exponent with multiplicity d. Furstenberg's criterion, originally obtained for the Lyapunov exponents of i.i.d. matrices [29], provides a way of ruling out this degenerate scenario. Since then, this circle of ideas has been extended to a variety of settings; see [3, 37] and the references therein. The following version can be derived from [37], Proposition 2, Theorem 3; see also [48], Theorem 2.4.

THEOREM 2.9. *Under the assumptions of Theorem 2.7, if $d\lambda_1 = \lambda_\Sigma$, then there is a measurable family $(\nu_x)_{x\in\mathrm{Supp}(\pi)}$ satisfying the following property: for $\mathbb{P} \times \pi$-almost every $(\underline{\omega}, x)$ and all $n \geq 1$,*

$$\big(\mathcal{A}^n_{\underline{\omega},x}\big)_* \nu_x = \nu_{f^n_{\underline{\omega}} x}. \tag{2.25}$$

Above, we write

$$B_* \nu_x := \nu_x \circ \widehat{B}^{-1} \tag{2.26}$$

for the *pushforward* of ν_x by some measurable $\widehat{B} : S^{d-1} \to S^{d-1}$.

We emphasize that (2.25) is a degenerate scenario, which can be ruled out if the law of $(f^n_{\underline{\omega}}x, \mathcal{A}^n_{\underline{\omega},x})$ in $X \times GL_d(\mathbb{R})$ has "broad enough support." To see this on a formal level, fix $y \in X$ and consider the space of sample paths $\underline{\omega} \in \Omega$ such that $y = f^n_{\underline{\omega}}x$. Observe that the matrix $\mathcal{A}^n_{\underline{\omega},x}$ depends explicitly on the whole sample path $\underline{\omega}$, while $\nu_y = \nu_{f^n_{\underline{\omega}}x}$ depends only on $\underline{\omega}$ implicitly through the fixed value of $y = f^n_{\underline{\omega}}x$. From here, it follows that if (2.25) holds, then the law of $\mathcal{A}^n_{\underline{\omega},x}$ conditioned on $y = f^n_{\underline{\omega}}x$ must have empty interior in $GL_d(\mathbb{R})$ (cf. Lemma 3.6 below). We conclude that if the law of $\mathcal{A}^n_{\underline{\omega},x}$ on $GL_d(\mathbb{R})$ conditioned on the event $f^n_{\underline{\omega}}x = y$ has a nonempty interior, then (2.25) cannot occur, and so $d\lambda_1 > \lambda_\Sigma$ must hold. However, there are some technical hurdles in the above argument: (i) the law of $\mathcal{A}^n_{\underline{\omega},x}$ conditioned on $y = f^n_{\underline{\omega}}x$ can be difficult to work with directly, and (ii) the family $x \mapsto \nu_x$ varies measurably.

The following work-around is inspired by Bougerol [12], and allows us to work instead with a weak* continuous family $\{\nu_x\}$.

PROPOSITION 2.10. *Assume that X is compact and that P is uniformly geometrically ergodic. If $d\lambda_1 = \lambda_\Sigma$, then there is a family $(\nu_x)_{x\in\mathrm{Supp}(\pi)}$ of probability measures on S^{d-1} such that the mapping $x \mapsto \nu_x$ varies* weak* *continuously on* $\mathrm{Supp}(\pi)$, *and so that*

$$\nu_x = \left(\mathcal{A}^n_{\underline{\omega},x}\right)^T_* \nu_{f^n_{\underline{\omega}}x} \tag{2.27}$$

for $\mathbb{P} \times \pi$-almost every $(\underline{\omega}, x)$ and all $n \geq 1$.

Compared to (2.25), $\mathcal{A}^n_{\underline{\omega},x}$ appears on the right-hand side, applied to $\nu_{f^n_{\underline{\omega}}x}$ and under a transpose.

COROLLARY 2.11. *Assume the setting of Proposition 2.10. If $d\lambda_1 = \lambda_\Sigma$, then for all $n \geq 1$ we have*

$$\nu_x = \left(A^T\right)_* \nu_y \tag{2.28}$$

holds for all $x \in \mathrm{Supp}(\pi)$ *and all (y, A) in the support of* $\mathrm{Law}(f^n_{\underline{\omega}}x, \mathcal{A}^n_{\underline{\omega},x})$ *on $X \times GL_d(\mathbb{R})$.*

While Proposition 2.10 and Corollary 2.11 follow from arguments in Bougerol's original paper [12], the results are not stated in the above form. For the sake of completeness, sketches of the proofs of Proposition 2.10 and Corollary 2.11 are given in Appendix B.

REMARK 2.12. Recall that a Markov kernel is called *strong Feller* if $\varphi : X \to \mathbb{R}$ bounded measurable implies $P\varphi : X \to \mathbb{R}$ is continuous. There are several works on the application of Furstenberg-type ideas to the Lypaunov exponents of RDS, for example, to continuous-time RDS satisfying the strong Feller property [6, 9], dealing in particular with the issue of obtaining some regularity of the measurable family $x \mapsto \nu_x$. The version chosen here is particularly well suited to continuous cocycles over RDS, which are uniformly geometrically ergodic but lack the strong Feller property.

3. Random dynamical systems with absolutely continuous noise. In this section, we begin to specialize to the class of random dynamical systems pertaining to our main results, that is, those posed on finite-dimensional compact manifolds and forced by noise drawn from a smooth manifold and absolutely continuous with respect to the Lebesgue measure. We focus on obtaining checkable sufficient conditions for several of the properties we will use later on, namely the existence of open small sets (Proposition 3.1), strong aperiodicity (Lemma 3.2) and for ruling out Furstenberg's criterion (Proposition 3.3).

In what follows, we will consider classes of continuous RDS f_ω acting on a compact manifold M without boundary. Throughout this section, we assume the following.

(R) We have that Ω_0 is a smooth, complete Riemannian manifold. The law $\mathbb{P}_0$ on Ω_0 admits a density ρ_0 with respect to Lebesgue measure $\mathrm{d}\omega$ on Ω_0. Additionally, the mapping $\Omega_0 \times M \to M$ given by

$$(\omega, x) \mapsto f_\omega(x) \tag{3.1}$$

is C^2.

It is sometimes useful to view the RDS $\{f^n_{\underline{\omega}}\}$ as a function of finitely many of the coordinates ω_i of the sequence $\underline{\omega} = \{\omega_i\}_{i\geq 1}$. To this end, we adopt the notation $\underline{\omega}^n = (\omega_1, \dots, \omega_n)$ for elements of the space Ω_0^n, the Cartesian product of n copies of Ω_0. With a slight abuse of notation, we will use the symbol ρ_0 for the density of the product law $\mathbb{P}_0^n$ on Ω_0^n, defined as

$$\rho_0(\underline{\omega}^n) = \rho_0(\omega_1) \cdots \rho_0(\omega_n). \tag{3.2}$$

3.1. *Sufficient conditions for small sets and aperiodicity.* First, we give a checkable sufficient condition for RDS satisfying assumption (R) to admit open small sets (Definition 2.1).

Our condition is framed for fixed $x \in M$, in terms of the mapping $\Phi_x : \Omega_0^n \to M$ defined for $\underline{\omega}^n = (\omega_1, \dots, \omega_n)$ by

$$\Phi_x(\underline{\omega}^n) := f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x). \tag{3.3}$$

PROPOSITION 3.1. *Assume the RDS f_ω satisfies assumption (R), and there exist $n \geq 1$ and $(\underline{\omega}^n_\star, x_\star) \in \mathrm{Supp}(\rho_0) \times \mathrm{Supp}(\pi)$, which satisfy the following properties*:

(i) *There exist some $c > 0$, $\varepsilon > 0$ such that $\rho_0(\underline{\omega}^n) \geq c > 0$ if $|\underline{\omega}^n - \underline{\omega}^n_\star| < \varepsilon$.*
(ii) *$\Phi_{x_\star}$ is a submersion at $\underline{\omega}^n = \underline{\omega}^n_\star$.*

Then P admits an open, P^n-small set with corresponding measure $\nu_n \ll \mathrm{Leb}_M$.

PROOF. It suffices to consider the case when $n = 1$; the general proof is similar. By the constant rank theorem, submersion implies there are C^2 coordinate charts near $\omega_\star := \underline{\omega}^1_\star$ in Ω_0 and near $f_{\omega_\star}(x_\star)$ in M for which $\omega_\star \mapsto \Phi_{x_\star}(\omega_\star)$ is an orthogonal projection. It is immediate that the pushforward $(\Phi_{x_\star})_*\mathbb{P}_0$ is absolutely continuous with respect to the Lebesgue measure Leb on M and has density bounded below by a constant $c' > 0$ in an ε'-neighborhood of $f_{\omega_\star}(x_\star)$ for some $\varepsilon' > 0$.

Since having full rank is an open property in the space of matrices, it follows that Φ_y is a submersion for all y in a small neighborhood of $x_\star$, and the same argument yields that $(\Phi_y)_*\mathbb{P}_0$ is absolutely continuous with respect to Leb on M with density bounded below by $c'/2$ in a small neighborhood of $f_{\omega_\star}(x_\star)$. Note we have used here the fact that for $r \geq 1$, the C^r data of the charts supplied by the constant rank theorem are controlled by the C^r data of $(\omega, x) \mapsto f_\omega(x)$.

All in all, we conclude

$$P(x, K) \geq \frac{c'}{2} \mathrm{Leb}(K \cap U') \tag{3.4}$$

for all x in an open neighborhood V of $x_\star$, where U' is a small open neighborhood of $\Phi_{x_\star}(\omega_\star)$. We conclude V is a ν-small set with $\nu(K) := \frac{c'}{2}\mathrm{Leb}(K \cap U')$. $\square$

We now give a simple condition for the Markov kernel P to be strongly aperiodic in the sense of Theorem 2.3.

LEMMA 3.2. *Assume there exist $\omega_\star \in \mathrm{Supp}(\mathbb{P}_0)$ and $x_\star \in M$ such that $f_{\omega_\star}(x_\star) = x_\star$. Then, for any open $U \subset M$ containing $x_\star$ we have that $P(x_\star, U) > 0$.*

The proof follows from the continuity of $\omega \mapsto f_\omega(x)$ and is omitted for brevity.

3.2. *Ruling out Furstenberg's criterion.* We now turn to the treatment of the derivatives $D_x f^n_{\underline{\omega}}$ as a linear cocycle over f_ω, and will provide sufficient conditions under which the degeneracy $d\lambda_1 = \lambda_\Sigma$ can be ruled out using Furstenberg's criterion (Proposition 2.10). To start, the following condition ensures that Lyapunov exponents exist as in Theorem 2.7.

(I) For a.e. ω, we have that f_ω is a local diffeomorphism, and admits an ergodic stationary measure π on M with the property that

$$\int \left(\log^+ |D_x f_\omega| + \log^+ |(D_x f_\omega)^{-1}|\right) \mathrm{d}\pi(x)\, \mathrm{d}\mathbb{P}_0(\omega) < \infty \tag{3.5}$$

Assuming conditions (R) and (I), Theorem 2.7 applies to the (continuous) linear cocycle $\mathcal{A}^n_{\underline{\omega},x} = D_x f^n_{\underline{\omega}}$ over the RDS $f^n_{\underline{\omega}}$ equipped with its stationary measure π.

We now turn to the task of ruling out the degeneracy $d\lambda_1 = \lambda_\Sigma$. For the sake of simplicity, we state our results below in the special case $M = \mathbb{T}^d$, $d \geq 1$, where as usual the periodic box $\mathbb{T}^d$ is parametrized by $[0, 2\pi)^d$. Note that in this case, the tangent bundle $TM = T\mathbb{T}^d$ is parallelizable, that is, it is diffeomorphic to the Cartesian product $\mathbb{T}^d \times \mathbb{R}^d$.

Below, for $n \geq 1$ and $x \in \mathbb{T}^d$, the mapping $\Phi_x : \Omega_0^n \to \mathbb{T}^d$ is as defined previously. Additionally, we define $\widehat{\Phi}_x : \Omega_0^n \to SL_d(\mathbb{R})$ by

$$\widehat{\Phi}_x(\underline{\omega}^n) := \frac{1}{|\det D_x f^n_{\underline{\omega}}|^{1/d}} D_x f^n_{\underline{\omega}}, \tag{3.6}$$

where here we follow the slightly nonstandard convention that $SL_d(\mathbb{R})$ denotes the space of $d \times d$ matrices with determinant ± 1. Since f_ω is a local diffeomorphism, $|\det D_x f^n_{\underline{\omega}}| \neq 0$, and hence $\widehat{\Phi}_x$ is well-defined. Observe that $\widehat{\Phi}_x : \Omega_0^n \to SL_d(\mathbb{R})$ is a C^1 mapping.

For $\underline{\omega}^n \in \Omega_0^n$ and $x \in \mathbb{T}^d$, define

$$\Sigma_{x;\underline{\omega}^n} := \ker D_{\underline{\omega}^n} \Phi_x \subset T_{\underline{\omega}^n} \Omega_0^n. \tag{3.7}$$

PROPOSITION 3.3. *Assume that f_ω satisfies conditions (R) and (I), and that the corresponding Markov kernel P is uniformly geometrically ergodic with stationary measure π. Moreover, assume there exist $n \geq 1$ and $(\underline{\omega}^n_\star, x_\star) \in \mathrm{Supp}(\rho_0) \times \mathrm{Supp}(\pi)$, which satisfy the following properties*:

(i) *There are $c, \varepsilon > 0$ such that $\rho_0(\underline{\omega}^n) \geq c > 0$ if $|\underline{\omega}^n_\star - \underline{\omega}^n| < \varepsilon$.*
(ii) *The mapping $\Phi_{x_\star}$ is a submersion at $\underline{\omega}^n = \underline{\omega}^n_\star$.*
(iii) *The restriction of $D_{\underline{\omega}^n_\star} \widehat{\Phi}_{x_\star}$ to $\Sigma_{x_\star;\underline{\omega}^n}$ is surjective as a linear operator onto the tangent space $T_{\widehat{\Phi}_{x_\star}(\underline{\omega}^n_\star)} SL_d(\mathbb{R})$.*

Then $d\lambda_1 > \lambda_\Sigma$.

REMARK 3.4. In our primary applications in this paper, the RDS $f^n_{\underline{\omega}}$ is volume-preserving, and so $\det(D_x f^n_{\underline{\omega}}) \equiv 1$. As a result, $\lambda_\Sigma = 0$ and the conditions of Proposition 3.3 imply $\lambda_1 > 0$.

The proof of Proposition 3.3 uses the following two lemmas. The first is a consequence of the constant rank theorem.

LEMMA 3.5. *Let $U \subset \mathbb{R}^a$, $V \subset \mathbb{R}^b$, $W \subset \mathbb{R}^c$ be open neighborhoods of the origin, with $a, b, c \in \mathbb{N}$. Let $F : U \to V$, $G : U \to W$ be C^2 mappings and let $\rho : U \to (0, \infty)$ be a strictly positive continuous function. Let $\mathrm{d}u$, $\mathrm{d}v$, $\mathrm{d}w$ denote Lebesgue measure on U, V, W, respectively. Lastly, assume*

(a) *$D_0 F$ is a surjection*;
(b) *$D_0 G|_{\ker D_0 F}$ is a surjection.*

Then there is a nonnegative continuous function $\widehat{\rho} : V \times W \to [0, \infty)$ *such that* $\widehat{\rho}\,\mathrm{d}v\,\mathrm{d}w = (F, G)_*[\rho\,\mathrm{d}u]$ *and* $\widehat{\rho} > 0$ *on a neighborhood of the origin.*

We also require the following.

LEMMA 3.6. *Let* η, η' *be measures on* S^{d-1}. *Then*

$$M_{\eta,\eta'} = \{A \in SL_d(\mathbb{R}) : A_*\eta = \eta'\} \tag{3.8}$$

has empty interior in $SL_d(\mathbb{R})$.

Below we provide a sketch with references for the proof of this standard fact.

SKETCHED PROOF OF LEMMA 3.6. In the case $\eta = \eta'$, we have that $M_\eta = M_{\eta,\eta}$ is a subgroup of $SL_d(\mathbb{R})$, and by, for example, [9], Lemma 4.6, we have the following cases:

(a) M_η is compact, in which case there is some fixed inner product $\langle\cdot,\cdot\rangle'$ on $\mathbb{R}^d$ with respect to which each element $A \in M_\eta$ is an isometry (cf. [6], Lemma 6.7(ii)); and

(b) M_η is noncompact, in which case there is a collection of proper subspaces $V_1, \dots, V_k$ of $\mathbb{R}^d$ such that for all $A \in M_\eta$, we have $A(\bigcup_i V_i) = \bigcup_i V_i$ (cf. [29], Theorem 8.6).

In either case, it is straightforward to show by a perturbation argument that M_η has an empty interior.

When $\eta \neq \eta'$, consider the mapping $\varphi : SL_d(\mathbb{R}) \times SL_d(\mathbb{R}) \to SL_d(\mathbb{R})$ given by $\varphi(A, B) = A^{-1}B$. It is not hard to show that φ is smooth, and that its derivative has constant rank $d^2 - 1$. In particular, φ is an open mapping, that is, the image of an open set is open. Moreover, we have that $\varphi(M_{\eta,\eta'} \times M_{\eta,\eta'}) \subset M_\eta$. This all implies that if $M_{\eta,\eta'}$ contained an open set, then M_η would also contain an open set, a contradiction. Therefore, $M_{\eta,\eta'}$ has an empty interior. □

PROOF OF PROPOSITION 3.3. Below we present the proof when $n = 1$, writing $\omega_\star = \underline{\omega}^1_\star$. The general proof is essentially the same. By Corollary 2.11 and Lemma 3.6, it suffices to show that

$$\big(\{f_{\omega_\star}(x_\star)\} \times SL_d(\mathbb{R})\big) \cap \mathrm{Supp}\big(\mathrm{Law}\big(\Phi_{x_\star}(\cdot), \widehat{\Phi}_{x_\star}(\cdot)\big)\big) \tag{3.9}$$

contains a set of the form $\{f_{\omega_\star}(x_\star)\} \times U$ where $U \subset SL_d(\mathbb{R})$ is an open neighborhood. This follows from the hypotheses of Proposition 3.3, and on passing to local smooth charts, Lemma 3.5 applied to $F = \Phi_{x_\star}$, $G = \widehat{\Phi}_{x_\star}$. □

REMARK 3.7. A version of Proposition 3.3 can also be written down in the case when M is a general compact Riemannian manifold. To do this, one should work with the fiber bundle $SL(M)$ over M with fibers

$$SL_p(M) = \{A : \mathbb{R}^n \to T_pM \text{ linear such that } \det(A) = \pm 1\}, \tag{3.10}$$

regarded as a principle bundle over M with structure group $SL_d(\mathbb{R})$; see [36]. Note that det is defined using the standard inner product on $\mathbb{R}^d$ in the domain and the Riemannian inner product on T_pM in the codomain. Fixing $x \in M$ and a determinant 1 isomorphism $\mathbb{R}^d \to T_xM$, we can now view the fiber $SL_p(M)$ over $p \in M$ as the space of determinant 1 mappings $A : T_xM \to T_pM$. With this convention, one can view $\widehat{\Phi}_x$ as a mapping from Ω_0^n to $SL(M)$, with

$$\widehat{\Phi}_x(\underline{\omega}^n) = \left(\Phi_x(\underline{\omega}^n), \frac{1}{|\det D_x f^n_{\underline{\omega}}|^{1/d}} D_x f^n_{\underline{\omega}}\right). \tag{3.11}$$

With these identifications, the proof of Proposition 3.3 can now be carried out in charts and follows *mutatis mutandi*.

4. Almost sure mixing for incompressible RDS with absolutely continuous noise. Equipped with tools for estimating Lyapunov exponents, we turn now to conditions ensuring almost sure correlation decay. In the following informal discussion, we will assume that f_ω is a differentiable RDS on a compact manifold M in the setting of Section 3, that is, assuming conditions (R) and (I). Additionally, we will assume the following:

(U) f_ω admits stationary measure π, which is almost surely invariant, that is, $(f_\omega)_*\pi = \pi$ for $\mathbb{P}_0$-almost every $\omega \in \Omega_0$.

Assuming (U), almost sure mixing asserts decay of the correlation terms

$$\mathrm{Cor}_n(\varphi, \psi) := \left| \int \varphi(x)\psi \circ f^n_{\underline{\omega}}(x)\, \mathrm{d}\pi(x) \right| \quad \text{with probability 1} \tag{4.1}$$

as $n \to \infty$ for sufficiently regular observables φ, ψ with π-mean zero. The term $\mathrm{Cor}_n(\varphi, \psi)$ depends on the random sample $\underline{\omega}$; we can estimate the probability that this quantity is not too large by using Chebyshev's inequality, that is,

$$\mathbb{P}\{\mathrm{Cor}_n(\varphi, \psi) > \varepsilon\} \le \varepsilon^{-2}\mathbb{E}_{\mathbb{P}} \left| \int \varphi\psi \circ f^n_{\underline{\omega}}\, \mathrm{d}\pi \right|^2. \tag{4.2}$$

The numerator on the right-hand side is equal to

$$\mathbb{E}_{\mathbb{P}} \left| \int \varphi\psi \circ f^n_{\underline{\omega}}\, \mathrm{d}\pi \right|^2 = \mathbb{E}_{\mathbb{P}} \int \varphi(x)\varphi(y)\psi \circ f^n_{\underline{\omega}}(x)\psi \circ f^n_{\underline{\omega}}(y)\, \mathrm{d}\pi(x)\, \mathrm{d}\pi(y), \tag{4.3}$$

having introduced the dummy integration variable $y \in M$. We can rewrite this expression using the two-point process (x_n, y_n) on $M \times M$ given for fixed initial x_0, y_0 by

$$(x_n, y_n) = \big(f^n_{\underline{\omega}}(x_0), f^n_{\underline{\omega}}(y_0)\big). \tag{4.4}$$

In this notation, the expression in (4.3) is given by

$$\mathbb{E}_{\mathbb{P}} \left| \int \varphi\psi \circ f^n_{\underline{\omega}}\, \mathrm{d}\pi \right|^2 = \int \psi^{(2)} \big(P^{(2)}\big)^n \varphi^{(2)}\, \mathrm{d}\pi^{(2)}. \tag{4.5}$$

Here, $\varphi^{(2)}(x, y) := \varphi(x)\varphi(y)$, with $\psi^{(2)}$ similarly defined, while $\mathrm{d}\pi^{(2)}(x, y) = \mathrm{d}\pi(x)\, \mathrm{d}\pi(y)$, and $P^{(2)}$ denotes the Markov semigroup for (x_n, y_n). We conclude from this calculation the following general principle:

Geometric ergodicity of the two-point Markov semigroup $P^{(2)}$ implies exponential mixing with probability 1 for f_ω.

This principle and the argument presented above was known to the Sinai School in the early 2000s; see, for example, the paper [27] or the more recent work [8].

On the other hand, we emphasize that $P^{(2)}$ cannot possibly be uniformly geometrically ergodic with respect to $\pi^{(2)}$, since the diagonal

$$\Delta = \{(x, x) : x \in M\} \subset M \times M \tag{4.6}$$

is almost surely invariant under the action of $(x, y) \mapsto (f_\omega(x), f_\omega(y))$, and so supports at least one additional stationary measure for $P^{(2)}$ not equal to $\pi^{(2)}$. Instead, we must work with the noncompact phase space

$$M^{(2)} := M \times M \setminus \Delta, \tag{4.7}$$

which necessitates the use of a Lyapunov function $V : M^{(2)} \to [1, \infty)$ as in Harris' theorem (Theorem 2.3) to ensure positive recurrence of $P^{(2)}$ away from Δ.

The construction of such a V is the primary goal of this section. In Section 4.1, we collect some preliminary results, which connect Lyapunov exponents for f_ω to growth of moments of $|D_x f^n_{\underline{\omega}}|$. This is useful to prove repulsion from Δ; if x, y are sufficiently close, then

$$f^n_{\underline{\omega}}(y) \approx f^n_{\underline{\omega}}(x) + D_x f^n_{\underline{\omega}}(y-x), \tag{4.8}$$

and so growth of $|D_x f^n_{\underline{\omega}}|$ implies some repulsion from Δ, at least at the level of the linearization. This is promoted to *nonlinear* repulsion in Section 4.2, where we define the drift function V used in our proof. Finally, in Section 4.3, we apply Harris' theorem to $P^{(2)}$ and fill in the details needed to establish quenched mixing for f_ω in Proposition 4.6.

4.1. *Projective and twisted semigroups.* In this subsection, we define several auxiliary Markov semigroups and affiliated linear operators for use in our construction of the desired Lyapunov function V.

Conceptually, the form of the function V depends most on its values near the diagonal Δ. Here, we approximate the two-point motion by its linearization. This motivates the use of the *linear semigroup* TP acting on TM, defined for bounded, measurable $\varphi : TM \to \mathbb{R}$ by

$$TP\varphi(x,u) := \mathbb{E}_{\mathbb{P}_0}\varphi\big(f_\omega(x), D_x f_\omega u\big) \quad x \in M, u \in T_x M. \tag{4.9}$$

Of particular use are functions on TM of the form

$$\varphi(x,u) = |u|^{-q}\psi\big(x, u/|u|\big), \quad q > 0, \tag{4.10}$$

where $\psi : SM \to \mathbb{R}$ is measurable and bounded. Here and throughout, SM denotes the unit tangent bundle of M consisting of elements $(x,v) \in TM$ such that $|v| = 1$. Restricting TP to such observables, we see that

$$TP\varphi(x,u) = \widehat{P}_q\psi\big(x, u/|u|\big), \tag{4.11}$$

where $\widehat{P}_q$ is the *twisted projective semigroup* defined for bounded, measurable observables $\psi : SM \to \mathbb{R}$ by

$$\widehat{P}_q\psi(x,v) = \mathbb{E}_{\mathbb{P}_0}|D_x f_\omega v|^{-q}\psi\big(f_\omega(x), D_x f_\omega v/|D_x f_\omega v|\big). \tag{4.12}$$

The closely related $\widehat{P} = \widehat{P}_0$ is the *projective semigroup*, a Markov semigroup on SM corresponding to the Markov chain (x_n, v_n) defined for fixed initial $(x_0, v_0) \in SM$ by

$$(x_n, v_n) = \big(f^n_{\underline{\omega}}(x_0), D_{x_0} f^n_{\underline{\omega}}(v_0)/\big|D_{x_0} f^n_{\underline{\omega}}(v_0)\big|\big). \tag{4.13}$$

We now turn to the spectral theory of $\widehat{P}$, $\widehat{P}_q$, which we will later use to construct Lyapunov functions for $P^{(2)}$. For a $d \times d$ invertible matrix A, we write $m(A) = |A^{-1}|^{-1}$. The following assumption is useful and simplifies arguments.

(B) There is a constant $C_0 > 0$ such that

$$m(D_x f_\omega) \ge C_0^{-1} > 0 \quad \text{and} \quad |D_x f_\omega| \le C_0 \tag{4.14}$$

with probability 1. Moreover, $\|f_\omega\|_{C^2}$ is $\mathbb{P}_0$-essentially bounded.

Note that condition (B) implies condition (I) from Section 3. Below, the semigroups $\widehat{P}$, $\widehat{P}_q$ are regarded as bounded linear operators on $C^0(SM)$, with $\|\cdot\|$ denoting the uniform norm. Their spectrum will be denoted by $\sigma(\widehat{P})$, $\sigma(\widehat{P}_p)$, respectively.

LEMMA 4.1. *Assume conditions (R) and (B) hold true. Moreover, assume $\widehat{P}$ is uniformly geometrically ergodic with respect to a (unique) stationary measure $\widehat{\pi}$ on SM. Then there exists $q_0 > 0$ such that for all $q \in [0, q_0]$, it holds that $\widehat{P}_q$ admits a simple dominant eigenvalue $r(q) > 0$ such that $\sigma(\widehat{P}_q) \setminus \{r(q)\}$ is contained in a ball of radius less than $r(q)$. Consequently, for all such q there is a unique dominant eigenfunction ψ_q such that*

$$\lim_{n\to\infty}\|\psi_q - r(q)^{-n}\widehat{P}_q^n \mathbf{1}\| = 0. \tag{4.15}$$

Here, we have written $\mathbf{1}$ for the constant function identically equal to 1. The value $\Lambda(q) = \log r(q)$ is often referred to as the *moment Lyapunov exponent*, and is related to large deviation estimates in the convergence of Lyapunov exponents [2].

PROOF. As $\widehat{P}$ is uniformly geometrically ergodic, the eigenvalue 1 is simple and has the property that $\sigma(\widehat{P}) \setminus \{1\}$ is contained in a ball of radius $\delta < 1$; see [43], Chapter 20.1. To complete the proof, it suffices by classical spectral theory [33], to check that $\widehat{P}_q \to \widehat{P}$ in norm as $q \to 0$. For $\psi : SM \to \mathbb{R}$, we have

$$\left|\widehat{P}_q \psi(x, v) - \widehat{P}\psi(x, v)\right| \le \|\psi\| \cdot \mathbb{E}_{\mathbb{P}_0}\left||D_x f_\omega v|^q - 1\right|. \tag{4.16}$$

To estimate this, we observe that for $a > 0$ and $q \in \mathbb{R}$, we have

$$\left|a^q - 1\right| \le \left|q(a-1)\right| \cdot \max\{1, a^{q-1}\}. \tag{4.17}$$

By (B), it follows that

$$\left||D_x f_\omega v|^q - 1\right| \le |q|(C_0 + 1)C_0^{1+|q|}, \tag{4.18}$$

on handling each of the cases $q > 1$, $q < 1$ separately. The conclusion of the lemma now follows. □

The following is used to control the value of $r(q)$.

LEMMA 4.2. *Under the hypotheses of Lemma 4.1, we have that $q \mapsto r(q)$ is continuously differentiable over $q \in [0, q_0)$ and satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}q} r(0) = -\lambda_1, \tag{4.19}$$

where λ_1 is the Lyapunov exponent of the derivative cocycle $D_x f^n_{\underline{\omega}}$. Consequently, if $\lambda_1 > 0$ then $r(q) < 1$ for all sufficiently small $q > 0$.

The proof is straightforward; see, for example, [2] or [8], Lemma 5.10, for the proof in a similar setting. In the construction to come, we will use the following additional properties of ψ_p.

COROLLARY 4.3. *Assume the hypotheses of Lemma 4.1.*

(i) *For any $\varepsilon > 0$, we have that $\psi_q = \psi_q^{C^0} + \psi_q^{C^1}$, where $\|\psi_q^{C^0}\| \le \varepsilon$ while $\psi_q^{C^1}$ is continuously differentiable and satisfies $\|\psi_q^{C^1}\| + \|D\psi_q^{C^1}\| \lesssim_\varepsilon 1$.*

(ii) *If, in addition, the Markov chain (x_n, v_n) is topologically irreducible, then ψ_q is strictly positive on SM.*

PROOF. Item (i) is an immediate consequence of the C^0-convergence $r(q)^{-n}\widehat{P}_q^n \mathbf{1} \to \psi_q$ and the fact that $\widehat{P}_q^n \mathbf{1}$ is a C^1 function for all $n \ge 1$ (this uses assumption (R)).

For item (ii), there is an open neighborhood $U \subset SM$ such that $\psi_q|_U \ge c > 0$ for some constant $c > 0$. Now, given $(x, v) \in SM$, by assumption (B) we have

$$\psi_q(x, v) \ge C_0^{-n|q|} r(q)^{-n} \mathbb{E}_{(x,v)} \psi_q(x_n, v_n) \ge c C_0^{-n|q|} r(q)^{-n} \widehat{P}^n\big((x, v), U\big) \tag{4.20}$$

for all $n \ge 1$. By topological irreducibility, there is some n such that $P^n((x, v), U) > 0$, completing the proof. □

4.2. *Drift condition for the two-point process.* We now set about constructing the drift function V for the two-point process. Throughout this discussion, we take on all the hypotheses of Lemma 4.1 as well as topological irreducibility for (x_n, v_n) as in Corollary 4.3(ii).

Notation. First, for $s > 0$, we define

$$\Delta(s) := \{(x, y) \in M^{(2)} : d(x, y) < s\}. \tag{4.21}$$

Let $s_0 > 0$ denote the minimal injectivity radius of the exponential map $\exp_x, x \in M$ (see, e.g., [25], Chapter 3), and $w : \Delta(r) \to TM$, $w(x, y) := \exp_x^{-1}(y)$, noting that if $d(x, y) < s_0$ then $\exp_x^{-1}(y) \in T_x M$ is uniquely defined. Set $\widehat{w}(x, y)$ to be $w(x, y)/|w(x, y)|$.

Fix a parameter $p \in (0, q_0)$ so that $r(p) < 1$ (which is possible due to Lemma 4.2), and fix a value $s < s_0$, to be determined as we proceed. Our drift function will be of the form

$$V(x, y) = \chi(x, y) d(x, y)^{-p} \psi_p(x, \widehat{w}(x, y)) + \hat{c}(1 - \chi(x, y)), \tag{4.22}$$

where χ is a bump function supported near the diagonal, so that $\chi(x, y) \equiv 1$ on $\Delta(s/2)$ and $\chi(x, y) \equiv 0$ on $\Delta(3s/4)^c$. Here, $\hat{c} > 0$ is a constant chosen so that $V(x, y) \geq 1$ for all $(x, y) \in M^{(2)}$.

We now set about proving the desired drift condition. For this, we will use the following approximation lemma applied to the function $\widehat{V} : TM \to \mathbb{R}$ defined by

$$\varepsilon \widehat{V}(x, v) := |v|^{-p} \psi_p(x, v). \tag{4.23}$$

LEMMA 4.4. *Assume the setting at the beginning of Section 4.2. Assume s is sufficiently small, depending only on the $\mathbb{P}_0$-essential supremum of $\|f_\omega\|_{C^2}$, as in condition (B). Then, for any $\varepsilon > 0$, there is a constant $C = C_\varepsilon$ such that*

$$|TP\widehat{V}(x, w(x, y)) - P^{(2)} V(x, y)| \lesssim \varepsilon d(x, y)^{-p} + C_\varepsilon d(x, y)^{1-p}, \tag{4.24}$$

for all $(x, y) \in \Delta(s/(2C_0))$, uniformly in ε. Here, C_0 is the constant appearing in condition (B).

PROOF. Given $\varepsilon > 0$, let $\psi_p = \psi_p^{C^0} + \psi_p^{C^1}$ be as in Corollary 4.3(i), noting that $\|\psi_p^{C^0}\|_{C^0} \leq \varepsilon$ and $\|\psi_p^{C^1}\|_{C^1} \leq C_\varepsilon < \infty$. We now collect the estimates we use. From the first-order Taylor expansion on manifolds and condition (B), we have that

$$d_M(f_\omega(y), \exp_{f_\omega(x)} D_x f_\omega(w(x, y))) \lesssim |w(x, y)|^2, \tag{4.25}$$

for all $x, y \in \Delta(s)$, with s taken sufficiently small in terms of EssSup $\|f_\omega\|_{C^2}$. It follows that

$$|w(f_\omega(x), f_\omega(y)) - D_x f_\omega(w(x, y))| \lesssim |w(x, y)|^2, \tag{4.26}$$

which in conjunction with the mean value theorem for $a \mapsto a^{-p}$ implies

$$\big||w(f_\omega(x), f_\omega(y))|^{-p} - |D_x f_\omega(w)|^{-p}\big| \lesssim p|w|^{-p-1}|w|^2 \leq |w|^{1-p}, \tag{4.27}$$

for $w = w(x, y)$. Lastly, using (B) once more, we estimate

$$\left|\widehat{w}(f_\omega x, f_\omega y) - \frac{D_x f_\omega(w)}{|D_x f_\omega(w)|}\right| \lesssim \frac{|w|^2}{|w|} = |w|, \tag{4.28}$$

using the following general inequality for vectors v, v':

$$\left|\frac{v}{|v|} - \frac{v'}{|v'|}\right| \leq 2\frac{|v - v'|}{\min\{|v|, |v'|\}}. \tag{4.29}$$

For the main estimate, we start by noting that if $(x, y) \in \Delta(s/(2C_0))$, then $(f_\omega(x), f_\omega(y)) \in \Delta(s/2)$, and so

$$\begin{aligned} P^{(2)}V(x,y) &= \mathbb{E}\big|w\big(f_\omega(x), f_\omega(y)\big)\big|^{-p}\psi_p\big(x, \widehat{w}\big(f_\omega(x), f_\omega(y)\big)\big) \\ &= \mathbb{E}\big|w\big(f_\omega(x), f_\omega(y)\big)\big|^{-p}\psi_p^{C^1}\big(x, \widehat{w}\big(f_\omega(x), f_\omega(y)\big)\big) \\ &\quad + O\big(\varepsilon d(x,y)^{-p}\big). \end{aligned} \tag{4.30}$$

Meanwhile, we have that

$$\begin{aligned} TP\widehat{V}(x,w) &= \mathbb{E}\big|D_x f_\omega(w)\big|^{-p}\psi_p\big(x, D_x f_\omega(w)/\big|D_x f_\omega(w)\big|\big) \\ &= \mathbb{E}\big|D_x f_\omega(w)\big|^{-p}\psi_p^{C^1}\big(x, D_x f_\omega(w)/\big|D_x f_\omega(w)\big|\big) \\ &\quad + O\big(\varepsilon d(x,y)^{-p}\big). \end{aligned} \tag{4.31}$$

Combining the above estimates, we obtain

$$\begin{aligned} &P^{(2)}V(x,y) - TP\widehat{V}(x,w) \\ &\quad = O\big(\varepsilon d(x,y)^{-p} + \|\psi_p^{C^1}\|_{C^0} d(x,y)^{1-p} + \|\psi_p^{C^1}\|_{C^1} d(x,y)^{1-p}\big) \\ &\quad = O\big(\varepsilon d(x,y)^{-p} + C_\varepsilon d(x,y)^{1-p}\big), \end{aligned} \tag{4.32}$$

as desired. □

We collect this into the following drift-type estimate.

PROPOSITION 4.5. *Assume the setting at the beginning of Section 4.2, and moreover, that the Lyapunov exponent λ_1 of $(f_{\underline{\omega}}^n)$ is positive. Let p be sufficiently small so that $r(p) < 1$ (Lemma 4.2). Then there exist $\gamma < 1$ and $s = s_* > 0$ so that*

$$P^{(2)}V(x,y) < \gamma V(x,y) \quad \forall (x,y) \in \Delta(s_*). \tag{4.33}$$

Note that since $\Delta(s_*)^c \subset M^{(2)}$ is compact, the above estimate implies that V satisfies a drift condition for $P^{(2)}$ as in Definition 2.2.

PROOF. Let $\varepsilon > 0$ be sufficiently small, constraints on which will be made as we go along. Apply Lemma 4.4, and observe that for $(x, y) \in \Delta(s/(2C_0))$ we have that

$$P^{(2)}V(x,y) \le TP\widehat{V}\big(x, w(x,y)\big) + K\big(\varepsilon d(x,y)^{-p} + C_\varepsilon d(x,y)^{1-p}\big), \tag{4.34}$$

where $K > 0$ is a constant not depending on ε. By our construction, $\widehat{V}(x, w(x,y))$ is an eigenfunction for TP and satisfies the relation

$$\begin{aligned} TP\widehat{V}\big(x, w(x,y)\big) &= r(p)\widehat{V}\big(x, w(x,y)\big) = r(p)V(x,y) \\ &= r(p)d(x,y)^{-p}\psi_p\big(x, \widehat{w}(x,y)\big). \end{aligned} \tag{4.35}$$

Plugging this in and assuming $(x, y) \in \Delta(s_*)$, for some $s_* > 0$ to be determined,

$$P^{(2)}V(x,y) \le d(x,y)^{-p}\big(r(p)\psi_p\big(x, \widehat{w}(x,y)\big) + K\varepsilon + KC_\varepsilon s_*\big). \tag{4.36}$$

We now specify the constants ε, s_*, γ required to complete the proof. To start, set

$$\varepsilon = \frac{1}{100K}\big(1 - r(p)\big)\inf_{(x,v)\in SM}\psi_p(x,v), \tag{4.37}$$

which is > 0 by Corollary 4.3(ii). With this value of ε fixed, define $s_* = \varepsilon / C_\varepsilon$. Plugging these choices into the above bound for $P^{(2)}V$, we see that

$$P^{(2)}V(x,y) \le d(x,y)^{-p}\Big(r(p) + \frac{1}{50}(1-r(p))\Big)\psi_p(x, \widehat{w}(x,y)) = \gamma V(x,y), \tag{4.38}$$

for $(x,y) \in \Delta(s_*)$, where γ, defined to be the parenthetical term, is automatically less than 1. □

4.3. *Deducing scalar mixing.* The following connects geometric ergodicity of the two-point process to almost sure correlation decay for the one-point process.

PROPOSITION 4.6. *Assume f_ω is a continuous RDS on a compact, orientable, Riemannian manifold M without boundary satisfying condition (U). Assume that the two-point process with kernel $P^{(2)}$ on $M^{(2)}$ is V-geometrically ergodic, where $V : M^{(2)} \to [0,\infty)$ is integrable with respect to $\pi^{(2)} = \pi \times \pi$. Let $q, s > 0$ be fixed.*

Then there is a function $D = D_{q,s} : \Omega \to [1,\infty)$ and a constant $\alpha = \alpha_{q,s} > 0$ such that for all H^s functions $\varphi, \psi : M \to \mathbb{R}$ of π-mean zero, we have that

$$\left|\int \varphi(x)\psi \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}\pi(x)\right| \le D(\underline{\omega})\mathrm{e}^{-\alpha n}\|\varphi\|_{H^s}\|\psi\|_{H^s}, \tag{4.39}$$

while the function D satisfies $\mathbb{E}_{\mathbb{P}} D^q < \infty$.

The proof in the case $M = \mathbb{T}^d$ is essentially contained in [8], Section 7, which we sketch here briefly for the sake of completeness. See Remark 4.7 for comments on the case when M is a general compact manifold without boundary.

PROOF SKETCH OF PROPOSITION 4.6 WHEN $M = \mathbb{T}^d$. We present a sketch here assuming that the density $\frac{\mathrm{d}\pi}{\mathrm{d}\,\mathrm{Leb}} \equiv 1$; the general case in condition (U) is handled in essentially the same way.

Let $\{e_k, k \in \mathbb{Z}^d\}$ denote the orthogonal basis $e_k(x) = \mathrm{e}^{ik\cdot x}$ for $L^2(\mathbb{T}^d)$ and let $\varphi, \psi : \mathbb{T}^d \to \mathbb{R}$ be smooth, mean-zero functions with Fourier expansions $\varphi = \sum_{k\in\mathbb{Z}^d_0} \varphi_k e_k$, $\psi = \sum_{k\in\mathbb{Z}^d_0} \psi_k e_k$. Here, $\mathbb{Z}^d_0 := \mathbb{Z}^d \setminus \{0\}$, noting that φ, ψ are assumed to be mean-zero. Lastly, we will assume that

$$\|(P^{(2)})^n - 1 \otimes \pi^{(2)}\|_V \le C\mathrm{e}^{-\beta n}, \quad \beta > 0, \tag{4.40}$$

with notation as in Harris' theorem (Theorem 2.3).

To start, fix $\zeta > 0$, to be specified as we go along, and for $k, k' \in \mathbb{Z}^d_0$ define

$$N_{k,k'} = \max\left\{n \ge 0 : \left|\int e_k(x)e_{k'} \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}x\right| > \mathrm{e}^{-\zeta n}\right\}, \tag{4.41}$$

noting by the Chebyshev inequality argument presented at the beginning of Section 4 that

$$\mathbb{P}(N_{k,k'} > \ell) \lesssim \mathrm{e}^{\ell(2\zeta-\beta)}. \tag{4.42}$$

Note that in this step we used that $V \in L^1(\pi^{(2)})$. Hereafter, we assume $\zeta < \beta/2$, so in particular $N_{k,k'}$ is almost surely finite, and observe that we have the estimate

$$\left|\int e_k(x)e_{k'} \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}x\right| \le \mathrm{e}^{\zeta N_{k,k'}}\mathrm{e}^{-\zeta n}, \tag{4.43}$$

hence

$$\left|\int \varphi(x)\psi \circ f^n_{\underline{\omega}}(x)\,\mathrm{d}x\right| \le \mathrm{e}^{-\zeta n}\sum_{k,k'} |\varphi_k||\psi_{k'}|\mathrm{e}^{\zeta N_{k,k'}}. \tag{4.44}$$

Define the random variable

$$K = \max\{|k| \vee |k'| : \mathrm{e}^{\zeta N_{k,k'}} > |k||k'|\} \tag{4.45}$$

where $a \vee b = \max\{a, b\}$ for $a, b \in \mathbb{R}$, and observe that

$$\mathbb{P}(K > \ell) \leq 2 \sum_{\substack{k,k' \in \mathbb{Z}_0^d \\ |k| > \ell}} \mathbb{P}(\mathrm{e}^{\zeta N_{k,k'}} > |k||k'|) \lesssim \sum_{\substack{k \in \mathbb{Z}_0^d \\ |k| > \ell}} |k|^{\frac{2\zeta - \beta}{\zeta}} \lesssim \ell^{d + \frac{2\zeta - \beta}{\zeta}}, \tag{4.46}$$

assuming, as we shall going forward, that ζ is small enough so that $d + \frac{2\zeta - \beta}{\zeta} < 0$. The random variable K is almost surely finite, as is the random variable

$$\widehat{D} = \max_{|k|, |k'| \leq K} \mathrm{e}^{\zeta N_{k,k'}}. \tag{4.47}$$

Noting that $\mathrm{e}^{\zeta N_{k,k'}} \leq \widehat{D}|k||k'|$ unconditionally, we conclude from (4.44) that

$$\begin{aligned} \left| \int \varphi(x) \psi \circ f_{\underline{\omega}}^n(x) \, \mathrm{d}x \right| &\leq \widehat{D} \mathrm{e}^{-\zeta n} \left(\sum_k |k||\varphi_k| \right) \left(\sum_{k'} |k'||\psi_{k'}| \right) \\ &\leq \widehat{D} \mathrm{e}^{-\zeta n} \|\varphi\|_{H^{\frac{d}{2}+2}} \|\psi\|_{H^{\frac{d}{2}+2}}. \end{aligned}$$

This is our desired estimate, except for the fact that on the right-hand side we bound in terms of $H^{\frac{d}{2}+2}$. Going from here to H^s regularity for arbitrary $s > 0$ follows from an approximation argument and requires shrinking the mixing rate ζ; see, for example, [8], Lemma 7.1, or [21], Lemma 4.2. □

REMARK 4.7. To handle the case when M is a general compact manifold without boundary, let $U_1, \dots, U_m$ be a finite cover of M by smooth charts (ϕ_i, U_i) with $\phi_i : U_i \to \mathbb{R}^d$, $d = \dim M$. Assume, as we may, that the ϕ_i have bounded range in $\mathbb{R}^d$, and let (χ_i) be a smooth partition of unity with χ_i supported on each U_i. Now, [55], Theorem 7.5.1, implies that the expression

$$\left(\sum_i \|\underbrace{\chi_i \varphi \circ \phi_i^{-1}}_{=:\varphi_i}\|_{H^s}^2 \right)^{1/2} \tag{4.48}$$

is an equivalent norm for the H^s norm of φ on M. From here, one can embed each U_i in the periodic box $\mathbb{T}^d$, obtaining sequences of functions $e_{i,k}$, $k \in \mathbb{Z}_0^d$ such that under the expansion $\varphi = \sum_{i,k} \varphi_{i,k} e_{i,k}$ with coefficients $\varphi_{i,k} \in \mathbb{C}$, we have

$$\|\varphi\|_{H^s}^2 \approx \sum_{i,k} |k|^{2s} |\varphi_{i,k}|^2. \tag{4.49}$$

The proof of Proposition 4.6 now translates to this setting, throughout working with the sequences $\{e_{i,k}\}$ over indices i, k instead of $\{e_k\}$ over k. We omit any further details.

4.4. *Summary of sufficient conditions for almost sure exponential mixing.* The paper thus far culminates in Proposition 4.6, which provides a sufficient condition for an RDS to be exponentially mixing with probability 1. The following is a summary of how, for a given RDS, one can check these sufficient conditions using the tools developed up until this point. In the following section, these conditions will be checked for the Pierrehumbert model introduced in Section 1.

(0) *Basic assumptions*: Let M be a compact Riemannian manifold without boundary and $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ a probability space, $(\Omega, \mathcal{F}, \mathbb{P}) := (\Omega_0, \mathcal{F}_0, \mathbb{P}_0)^{\mathbb{Z}_{\geq 1}}$. Let $\omega \mapsto f_\omega$ be a measurable assignment to each $\omega \in \Omega_0$ of a continuous mapping $f_\omega : M \to M$, with compositions

$$f^n_{\underline{\omega}} = f_{\omega_n} \circ \cdots \circ f_{\omega_1} \tag{4.50}$$

$\underline{\omega} = (\omega_1, \omega_2, \ldots) \in \Omega$. In addition, f_ω will be assumed to satisfy the following:

(R) The i.i.d. noise parameters $\omega_i, i \geq 1$ come from a probability space $(\Omega_0, \mathbb{P}_0) = (\mathbb{R}^k, \rho_0 \,\mathrm{d}\,\mathrm{Leb})$ for some density $\rho_0 : \mathbb{R}^k \to \mathbb{R}_{\geq 0}$, and the mapping $(\omega, x) \mapsto f_\omega(x)$ is C^2 differentiable;

(B) There is a constant $C_0 > 0$ such that

$$m(D_x f_\omega) \geq C_0^{-1} > 0 \quad \text{and} \quad |D_x f_\omega| \leq C_0, \tag{4.51}$$

almost surely, and $\|f_\omega\|_{C^2}$ is $\mathbb{P}_0$-essentially bounded;

and the following mild strengthening of (U):

(U') For $\mathbb{P}_0$-a.e. ω, we have that f_ω preserves $\pi = \mathrm{Leb}$ on M.

(1) *Conditions for the one-point chain*: The one-point chain is the Markov chain on M with transition kernel $P(x, K) := \mathbb{P}_0(f_\omega(x) \in K)$. We will assume

(UE-P) The kernel P is uniformly geometrically ergodic with unique stationary measure π.

By Theorem 2.3 and compactness of M, this follows on showing that P is topologically irreducible, admits an open small set (Proposition 3.1) and is strongly aperiodic (Lemma 3.2).

(2) *Positive Lyapunov exponent*: By *(0)* and *(1)*, Theorem 2.7 applied to the cocycle $\mathcal{A}^n_{\underline{\omega},x} = D_x f^n_{\underline{\omega}}$ implies that the asymptotic exponential growth rate

$$\lambda_1 := \lim_{n\to\infty} \frac{1}{n} \log|D_x f^n_{\underline{\omega}}|$$

exists and is constant over $\mathbb{P} \times \pi$-a.e. $(\underline{\omega}, x)$. We will assume

(LE) $\lambda_1 > 0$.

Proposition 3.3 gives a sufficient condition for $d\lambda_1 > \lambda_\Sigma$, where $d = \dim M$, while almost sure volume preservation as in (U') implies $\lambda_\Sigma = 0$ (see Remark 2.8).

(3) *Conditions for the projective chain*: The projective chain is the Markov chain on the sphere bundle SM with transition kernel $\widehat{P}((x, v), K) = \mathbb{P}_0(\widehat{f}_\omega(x, v) \in K)$, where for $(x, v) \in SM$,

$$\widehat{f}_\omega(x, v) := \left(f_\omega(x), \frac{D_x f_\omega v}{|D_x f_\omega v|} \right). \tag{4.52}$$

We will assume

(UE-$\widehat{P}$) The kernel $\widehat{P}$ is uniformly geometrically ergodic with unique stationary measure $\widehat{\pi}$.

By Theorem 2.3 and compactness of SM, this follows on showing that $\widehat{P}$ is topologically irreducible, admits an open small set (Proposition 3.1) and is strongly aperiodic (Lemma 3.2).

(4) *Conditions for the two-point chain*: The two-point chain is the Markov chain on $M^{(2)} := M \times M \setminus \Delta$, $\Delta := \{(x, x) : x \in M\} \subset M \times M$ with transition kernel,

$$P^{(2)}\big((x, y), K\big) = \mathbb{P}_0\big(\big(f_\omega(x), f_\omega(y)\big) \in K\big). \tag{4.53}$$

By condition (U'), the measure $\pi^{(2)} := \pi \times \pi$ is almost surely invariant under $f_\omega \times f_\omega$, hence stationary for $P^{(2)}$. We will assume:

(UE-$P^{(2)}$) The kernel $P^{(2)}$ is V-uniformly geometrically ergodic, where $V : M^{(2)} \to \mathbb{R}$ is of the form

$$V(x, y) = d(x, y)^{-p}\psi(x, y), \tag{4.54}$$

for some $p > 0$ small, where $\psi : M^{(2)} \to (0, \infty)$ is continuous, bounded from above and bounded from below by a constant $c > 0$ on a small neighborhood of Δ.

As before, we must check in Theorem 2.3 that $P^{(2)}$ is topologically irreducible, admits an open small set (Proposition 3.1) and is strongly aperiodic (Lemma 3.2). Additionally, since $M^{(2)}$ is noncompact it is necessary to build V as above satisfying a drift condition (Definition 2.2). This is done using the constructions of Section 4.2, which relies on conditions (R), (B), (U'), (UE-P), (LE) and (UE-$\widehat{P}$), as well as topological irreducibility for $\widehat{P}$ (cf. Corollary 4.3(ii)).

Finally, if (UE-$P^{(2)}$) holds, then Proposition 4.6 applies, noting that V as in (4.54) is automatically integrable with respect to $\pi^{(2)}$ on $M^{(2)}$ if $p < 1$ and $\pi = \text{Leb}$ as assumed in (U'). We conclude almost sure exponential mixing for f_ω as in (4.39).

5. Application to the Pierrehumbert model. This section concerns the Pierrehumbert model presented in Section 1.1. Section 5.1 establishes topological irreducibility for the one-point, projective and two-point Markov chains. Section 5.2 treats positivity of the Lyapunov exponent for the Pierrehumbert model. Section 5.3 gives conditions for the existence of open small sets and aperiodicity for the three Markov chains. Section 5.4 summarizes the proof of Theorem 1.1 on exponential mixing for the Pierrehumbert model.

Notation. Throughout, $\mathbb{T}^2$ is parametrized by $[0, 2\pi)^2$. We slightly abuse notation and write $x + y$ for the sum of $x, y \in \mathbb{T}^2$ modulo 2π in both coordinates. For $x = (x_1, x_2) \in \mathbb{T}^2$, we write $[x]_i = x_i \in [0, 2\pi)$ for the ith coordinate of x, regarded as a real number modulo 2π.

The real number $\tau > 0$ is fixed, and $\Omega_0 = [0, 2\pi]^2$ denotes the space of possible noise parameters. Given $\omega = (\omega^1, \omega^2) \in \Omega_0$, we define

$$f_\omega = f^V_{\omega^2} \circ f^H_{\omega^1}, \tag{5.1}$$

$$f^H_\beta(x) = \begin{pmatrix} x_1 + \tau \sin(x_2 - \beta) \\ x_2 \end{pmatrix}, \qquad f^V_\beta(x) = \begin{pmatrix} x_1 \\ x_2 + \tau \sin(x_1 - \beta) \end{pmatrix}, \tag{5.2}$$

for $x = (x_1, x_2) \in \mathbb{T}^2$. Given $\underline{\omega} = (\omega_1, \omega_2, \ldots) \in \Omega := \Omega_0^{\mathbb{N}}$, we write $f^n_{\underline{\omega}} = f_{\omega_n} \circ \cdots \circ f_{\omega_1}$ for the corresponding composition of maps. At times, if $\omega_1, \ldots, \omega_n$ have been specified, we will abuse notation and write $x_n = f_{\omega_n} \circ \cdots \circ f_{\omega_1}(x)$.

5.1. *Topological irreducibility.*

5.1.1. *Irreducibility of the one-point Markov chain.* The random dynamical system $f^n_{\underline{\omega}}$ corresponding to the one-point process on $\mathbb{T}^2$ is exactly controllable, as we show below.

LEMMA 5.1 (Exact controllability of the one-point chain). *Given $x, y \in \mathbb{T}^2$, there exists $N = N(\tau) \in \mathbb{N}$ and $\underline{\omega} = (\omega_1, \ldots, \omega_N) \in \Omega_0^N$ such that*

$$f^N_{\underline{\omega}}(x) = y.$$

In particular, the one-point Markov kernel P is topologically irreducible.

PROOF. To start, set $N = N(\tau) = \lceil \frac{4\pi}{\tau} \rceil$ and observe that

$$\bar{\tau}_i := \frac{[x]_i - [y]_i}{N}$$

belongs to $[-\tau, \tau]$ for $i = 1, 2$. We now define $\underline{\omega}^N = (\omega_1, \dots, \omega_N) \in \Omega_0^N$ inductively as follows: given $\underline{\omega}^n := (\omega_1, \dots, \omega_n)$, choose $\omega_{n+1} = (\omega_{n+1}^1, \omega_{n+1}^2)$ such that

$$\tau \sin\big([f_{\underline{\omega}^n}^n(x)]_2 - \omega_{n+1}^1\big) = \bar{\tau}_1, \qquad \tau \sin\big([f_{\omega_{n+1}^1}^H \circ f_{\underline{\omega}^n}^n(x)]_1 - \omega_{n+1}^2\big) = \bar{\tau}_2,$$

noting that for each $1 \le n \le N$, we have

$$f_{\underline{\omega}^n}^n(x) = x + n(\bar{\tau}_1, \bar{\tau}_2).$$

By our definitions, this ensures $f_{\underline{\omega}^N}^N(x) = y$, as desired. $\square$

5.1.2. *Irreducibility of the projective Markov chain.* We now turn attention to the linear cocycle associated to the Jacobians $\mathcal{A}_{\underline{\omega},x}^n := D_x f_{\underline{\omega}}^n$, regarding $\mathcal{A}^n, n \ge 1$ as a mapping $\Omega \times \mathbb{T}^2 \to GL_2(\mathbb{R})$. We define the projective dynamics $\widehat{f}_\omega : \mathbb{T}^2 \times S^1 \to \mathbb{T}^2 \times S^1$, $S^1 \subset \mathbb{R}^2$ the unit circle, by

$$\widehat{f}_\omega(x, v) = \left(f_\omega(x), \frac{D_x f_\omega(v)}{|D_x f_\omega(v)|} \right),$$

and write $\widehat{f}_{\underline{\omega}}^n = \widehat{f}_{\omega_n} \circ \cdots \circ \widehat{f}_{\omega_1}$ for the time-n composition. As we will show, the RDS $\widehat{f}_{\underline{\omega}}^n$ is approximately controllable.

PROPOSITION 5.2. *For any $\varepsilon > 0$ and $(x, v), (x_\star, v_\star) \in \mathbb{T}^2 \times S^1$, there exists $N \in \mathbb{N}$ and $\underline{\omega}^N$ such that*

$$d\big(\widehat{f}_{\underline{\omega}^N}^N(x, v), (x_\star, v_\star)\big) < \varepsilon.$$

Moreover, N is bounded uniformly from above in terms of ε and τ alone. In particular, the Markov kernel $\widehat{P}$ associated to $\widehat{f}_{\underline{\omega}}^n$ is topologically irreducible.

We begin the proof with the following straightforward computation and some notation: the individual time-one Jacobians $\mathcal{A}_{\omega,x} := D_x f_\omega$ are of the form

$$\mathcal{A}_{\omega,x} = \begin{pmatrix} 1 & C^H \\ C^V & 1 + C^H C^V \end{pmatrix}, \tag{5.3}$$

where

$$C^H = C^H(\omega, x) = \tau \cos\big([x]_2 - \omega^1\big), \tag{5.4}$$

$$C^V = C^V(\omega, x) = \tau \cos\big([f_{\omega^1}^H(x)]_1 - \omega^2\big). \tag{5.5}$$

Given $x \in \mathbb{T}^2$, $v \in S^1$ and assuming $\omega_1, \dots, \omega_n$ have been specified, let us write

$$x_n := f_{\underline{\omega}}^n(x), \qquad w_n := D_x f_{\underline{\omega}}^n(v), \qquad v_n := \frac{D_x f_{\underline{\omega}}^n(v)}{|D_x f_{\underline{\omega}}^n(v)|},$$

so that in particular w_n and v_n are parallel for all n, with $v_n \cdot w_n > 0$, while $(x_n, v_n) = \widehat{f}_{\omega_n} \circ \cdots \circ \widehat{f}_{\omega_1}(x, v)$. Lastly, given $u \in \mathbb{R}^2$ we write $[u]_i, i = 1, 2$ for the ith coordinate of u.

Next, we establish the following intermediate lemma allowing to approximately control to points of the form $(\bar{x}, (1/\sqrt{2}, 1/\sqrt{2})) \in \mathbb{T}^2 \times S^1$.

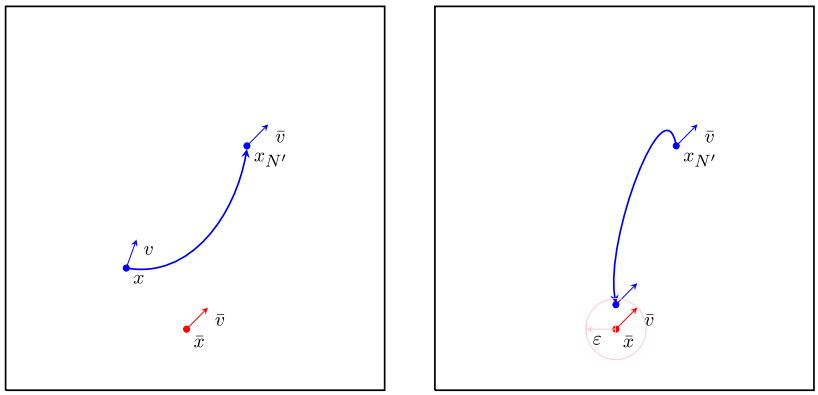

FIG. 2. *A visual explanation of the proof of approximate controllability for the projective chain: We first achieve the correct alignment in projective space, that is,* $\bar{v} = (1/\sqrt{2}, 1/\sqrt{2})$ *after which we perform rigid motions to bring the point within an* ε*-neighborhood of the target point* $\bar{x}$.

LEMMA 5.3. *Let* $(x, v) \in \mathbb{T}^2 \times S^1$, $\bar{x} \in \mathbb{T}^2$ *and* $\varepsilon > 0$. *Then there exists* $N_1 \in \mathbb{N}$ *and* $\underline{\omega}^{N_1}$ *such that*

$$d_{\mathbb{T}^2 \times S^1}\big((x_{N_1}, v_{N_1}), (\bar{x}, (1/\sqrt{2}, 1/\sqrt{2}))\big) < \varepsilon.$$

PROOF OF LEMMA 5.3. We divide the proof into two steps:

(1) *Achieving correct angle*. There exists $N' \in \mathbb{N}$ and $\omega_1, \dots, \omega_{N'} \in \Omega_0$ such that $v_{N'} = (1/\sqrt{2}, 1/\sqrt{2})$.

(2) *Rigid motions*. There exists $N'' \in \mathbb{N}$ and $\omega_{N'+1}, \dots, \omega_{N_1} \in \Omega_0$, $N_1 := N' + N''$, such that $v_{N_1} = (1/\sqrt{2}, 1/\sqrt{2})$ and $d_{\mathbb{T}^2}(x_{N_1}, \bar{x}) < \varepsilon$.

We refer the reader to schematic in Figure 2 where we have provided a visual explanation of the proof of approximate controllability for the projective chain.

Step (1). Since v is a unit vector, we must have that $[v]_i \geq 1/\sqrt{2}$ for one of $i = 1, 2$. For now, let us assume that $[v]_2 \geq 1/\sqrt{2}$; if not, the same proof works with minor modifications, details omitted.

Observe that for any $y \in \mathbb{T}^2$, by varying $\eta = (\eta^1, \eta^2) \in \Omega_0$ we can specify the pair $(C^H(\eta, y), C^V(\eta, y))$ to take any value in $[-\tau, \tau]^2$. In particular, if $C^V(\eta, y) = 0$, then

$$D_y f_\eta(u) = \begin{pmatrix} 1 & C^H(\eta, y) \\ 0 & 1 \end{pmatrix} u = \begin{pmatrix} [u]_1 + C^H(\eta, y)[u]_2 \\ [u]_2 \end{pmatrix}$$

for any $u \in \mathbb{R}^2$. With this in mind, we will choose $\omega_i = (\omega_i^1, \omega_i^2)$ for each i so that

$$C^V(\omega_i, x_{i-1}) = 0, \qquad C^H(\omega_i, x_{i-1}) = \frac{[w]_2 - [w]_1}{N'}.$$

where $N' \in \mathbb{N}$ is the smallest natural number so that $\frac{[w]_2 - [w]_1}{N'} \in [-\tau, \tau]$. This ensures that at time N', we have that

$$w_{N'} = \begin{pmatrix} [w]_1 + [w]_2 \displaystyle\sum_{i=1}^{N'} C^H(\omega_i, x_{i-1}) \\ [w]_2 \end{pmatrix} = \begin{pmatrix} [w]_2 \\ [w]_2 \end{pmatrix}.$$

In particular, $v_{N'} = (1/\sqrt{2}, 1/\sqrt{2})$, as desired.

Step (2). We will choose $\omega_i, i \geq N' + 1$ subject to two constraints. The first is that the relation

$$C^H(\omega_i, x_{i-1}) = \frac{C^V(\omega_i, x_{i-1})}{1 - C^V(\omega_i, x_{i-1})}$$

holds for each $i \geq N' + 1$, and ensures (by (5.3)) that $v_i = v_{N'} = (1/\sqrt{2}, 1/\sqrt{2})$ for all $i \geq N' + 1$.

For the second constraint, let $\zeta = (\zeta^1, \zeta^2) \in [0, \tau/2\pi]^2$, to be specified momentarily. Our second constraint will be that for all $i \geq N' + 1$, we have that

$$C^H(\omega_i, x_{i-1}) = \sqrt{\tau^2 - (2\pi\zeta^1)^2}, \qquad C^V(\omega_i, x_{i-1}) = \sqrt{\tau^2 - (2\pi\zeta^2)^2}.$$

That both the first and second constraints are satisfied requires that ζ^1, ζ^2 satisfy the relation

$$\sqrt{\tau^2 - (2\pi\zeta^1)^2} = \frac{\sqrt{\tau^2 - (2\pi\zeta^2)^2}}{1 - \sqrt{\tau^2 - (2\pi\zeta^2)^2}}. \tag{5.6}$$

With the first and second constraints in place, observe that for $i \geq N' + 1$ we will have that

$$(x_i, v_i) = \left(x_{i-1} + \begin{pmatrix} \sqrt{\tau^2 - (C^H(\omega_i, x_{i-1}))^2} \\ \sqrt{\tau^2 - (C^V(\omega_i, x_{i-1}))^2} \end{pmatrix}, v_{i-1} \right) = \left(x_{i-1} + \begin{pmatrix} 2\pi\zeta^1 \\ 2\pi\zeta^2 \end{pmatrix}, v_{i-1} \right).$$

CLAIM 5.4. There exists $(\zeta^1, \zeta^2) \in [0, \tau/(2\pi)]^2$ such that (5.6) holds, and moreover, ζ^1 and ζ^2 are *rationally independent*: there are no nonzero solutions $a_1, a_2 \in \mathbb{Q}$, $b \in \mathbb{Z}$ to the equation $a_1\zeta^1 + a_2\zeta^2 = b$.

Assuming the claim and fixing ζ^1, ζ^2, the Weyl equidistribution theorem in dimension $d = 2$ (see Corollary C.3 in Appendix C) implies there exists $N'' \in \mathbb{N}$, uniformly bounded from above in terms of ε, such that

$$x_{N''} = x_{N'} + N'' \begin{pmatrix} 2\pi\zeta^1 \\ 2\pi\zeta^2 \end{pmatrix}$$

is within distance ε of $\bar{x}$. $\square$

PROOF OF CLAIM. It can be checked using standard computer algebra software, for example, Mathematica, that the set of solutions $\mathcal{S} = \{(\zeta^1, \zeta^2)\}$ in $[0, \tau/2\pi]^2$ to the identity (5.6) is of the form $\mathcal{S} = \{(x, g(x)) : x \in [0, \tau/2\pi]\}$ where $g = g_\tau : [0, \tau/2\pi] \to [0, \tau/2\pi]$ is of the form

$$\begin{aligned} g(x) = {} & \frac{1}{2\pi(\tau^2 - 4\pi^2x^2 - 1)^2} \times \big(\tau^6 - 3\tau^4 + 2\tau^2\sqrt{\tau^2 - 4\pi^2x^2} + 16\pi^4(\tau^2 - 1)x^4 \\ & + 4\pi^2x^2\big(-2\tau^4 + 4\tau^2 - 2\sqrt{\tau^2 - 4\pi^2x^2} + 1\big)\big)^{1/2} \end{aligned}$$

By inspection, g is real-analytic on the open interval $(0, \tau/2\pi)$.

The set $\mathcal{RD}$ of rationally dependent ζ is of the form

$$\mathcal{RD} = \left(\bigcup_{a \in \mathbb{Q}} \{(a, y) : y \in \mathbb{R}\} \cup \bigcup_{a_1, a_2 \in \mathbb{Q}} \{(x, a_1x + a_2) : x \in \mathbb{R}\} \right) \cap [0, \tau/2\pi]^2,$$

and lies on a countable union of affine lines. The intersection of any real-analytic graph and an affine line has at finitely many points, and so $\mathcal{S} \cap \mathcal{RD}$ is at-most countable. Since $\mathcal{S}$ is uncountable, the proof is complete. $\square$

PROOF OF PROPOSITION 5.2. Let $\varepsilon > 0$ and $(x, v), (x_\star, v_\star) \in \mathbb{T}^2$ be fixed. Using again the fact that $f_\omega^{-1} = f^H_{\omega^1+\pi} \circ f^V_{\omega^2+\pi}$, a slight modification of Step 1 in the proof of Lemma 5.3 implies that with N' chosen as in the proof of Lemma 5.3, we have that there exists $\hat{\omega}^{N'} \in \Omega_0^{N'}$ such that

$$D_{x_\star}(f^{N'}_{\underline{\hat{\omega}}^{N'}})^{-1}(v_\star)$$

is parallel to $(1/\sqrt{2}, 1/\sqrt{2})$. Define $\bar{x} := (f^{N'}_{\underline{\hat{\omega}}^{N'}})^{-1}(x_\star)$. Let L_0 be an $\underline{\omega}$-uniform upper bound for $\mathrm{Lip}(\widehat{f}^{N'}_{\underline{\omega}})$. Set $\varepsilon_0 = \varepsilon/2L_0$ and let $N_1 \in \mathbb{N}$, $\underline{\omega}^{N_1}$ be as in Lemma 5.3 so that $d_{\mathbb{T}^2}(x_{N_1}, \bar{x}) < \varepsilon_0$ and $v_{N_1} = (1/\sqrt{2}, 1/\sqrt{2})$. Setting $\omega_{N_1+i} = \hat{\omega}_i$, $1 \le i \le N'$, $N := N_1 + N'$, we conclude that

$$d_{\mathbb{T}^2\times S^1}\big((x_N, v_N), (x_\star, v_\star)\big) < \varepsilon,$$

as desired. □

5.1.3. *Irreducibility of the two-point Markov chain.* Recall that $\Delta \subset \mathbb{T}^2 \times \mathbb{T}^2$ denotes the diagonal, $\Delta = \{(x, x) : x \in \mathbb{T}^2\}$.

PROPOSITION 5.5. *Given $(x, y), (x_\star, y_\star) \in \mathbb{T}^2 \times \mathbb{T}^2 \setminus \Delta$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ and some $\underline{\omega}^N \in \Omega_0^N$ such that*

$$\mathrm{dist}\big(f^N_{\underline{\omega}^N}(x), x_\star\big) + \mathrm{dist}\big(f^N_{\underline{\omega}^N}(y), y_\star\big) < \varepsilon.$$

In particular, the two-point Markov kernel $P^{(2)}$ is topologically irreducible.

To prove Proposition 5.5, we start with following preliminary computation.

LEMMA 5.6. *Given $x, y \in \mathbb{T}^2$, $\gamma^i \in [0, 2\pi)$, $i = 1, 2$, define*

$$\beta^1 = \frac{[x]_2 + [y]_2}{2} - \gamma^1, \qquad \beta^2 = \frac{[x]_1 + [y]_1}{2} - \gamma^2.$$

Then

$$\big[f^H_{\beta^1}(x) - f^H_{\beta^1}(y)\big]_1 = [x - y]_1 + 2\tau \sin\left(\frac{[x - y]_2}{2}\right)\cos(\gamma^1) \tag{5.7}$$

and

$$\big[f^V_{\beta^2}(x) - f^V_{\beta^2}(y)\big]_2 = [x - y]_2 + 2\tau \sin\left(\frac{[x - y]_1}{2}\right)\cos(\gamma^2). \tag{5.8}$$

In particular, setting $\gamma^i = \frac{\pi}{2}$, one has that $[f^H_{\beta^1}(x) - f^H_{\beta^1}(y)]_1 = [x - y]_1$ and $[f^V_{\beta^2}(x) - f^V_{\beta^2}(y)]_2 = [x - y]_2$.

The following intermediate lemma establishes approximate controllability to configurations in $\mathbb{T}^2 \times \mathbb{T}^2 \setminus \Delta$ of a special form.

LEMMA 5.7. *For any $x, y \in \mathbb{T}^2 \times \mathbb{T}^2 \setminus \Delta$, $\delta_1 > 0$, $\varepsilon > 0$ and $\bar{x} \in \mathbb{T}^2$ such that $\frac{\tau\cos(\delta_1/2)}{2\pi}$ is irrational, there exists $N \in \mathbb{N}$ such that $x_N = f^N_{\underline{\omega}^N}(x)$, $y_N = f^N_{\underline{\omega}^N}(y)$ satisfy*

$$x_N - y_N = (\delta_1, 0) \quad \text{and} \quad d_{\mathbb{T}^2}(x_N, \bar{x}) < \varepsilon. \tag{5.9}$$

Moreover, we have that $N \le \hat{N}$, where $\hat{N} = \hat{N}(d_{\mathbb{T}^2}(x, y), \delta_1, \varepsilon, \tau)$; in particular, N is bounded uniformly over $(x, y) \in \mathbb{T}^2 \setminus \Delta$ with $d_{\mathbb{T}^2}(x, y) \ge \eta$ for each fixed $\eta > 0$.

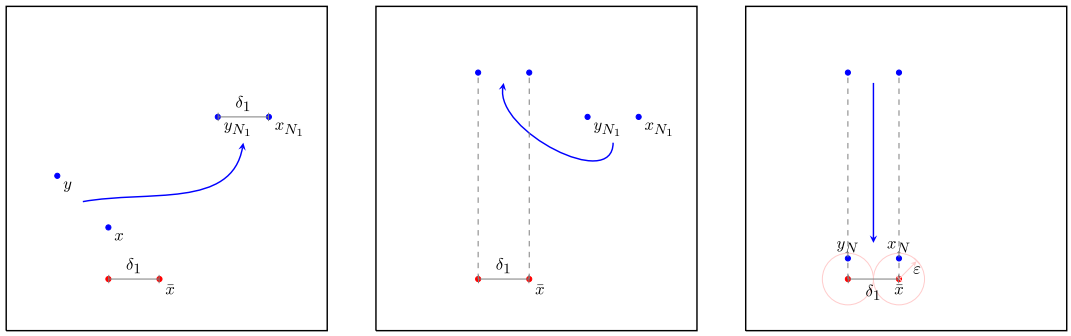

FIG. 3. *A visual explanation of the proof of approximate controllability for the two-point chain: we first achieve the correct separation between the two points after which we perform rigid motions to align the points with the same first-coordinate as the target point, and then bring the process within an ε-neighborhood of the target points.*

PROOF. Throughout, when $\underline{\omega}^n \in \Omega_0^n$ has been specified, we will write $x_n = f^n_{\underline{\omega}^n}(x)$, $y_n = f^n_{\underline{\omega}^n}(y)$. We split the proof into two steps:

(1) *Achieving correct separation.* There exist $N_1 \in \mathbb{N}$ and $\underline{\omega}^{N_1}$ such that $[x_{N_1} - y_{N_1}]_1 = \delta_1$ and $[x_{N_1}]_2 = [y_{N_1}]_2$.
(2) *Rigid motions.* There exist $N_2 \in \mathbb{N}$ and $\underline{\omega}^{N_2}$ such that $x_{N_1+N_2} - y_{N_1+N_2} = x_{N_1} - y_{N_1} = (\delta_1, 0)$ and $\mathrm{dist}_{\mathbb{T}^2}(x_{N_1+N_2}, \bar{x}) < \varepsilon$.

As we will see, the value N_1 will be bounded uniformly from above depending only on $\mathrm{dist}_{\mathbb{T}^2}(x, y)$, τ and δ_1, while N_2 will depend only on δ_1, τ and ε. We refer the reader to schematic in Figure 3 where we have provided a visual explanation of the proof of approximate controllability.

Step (1). Assume $[x]_2 \neq [y]_2$; we will remove this constraint momentarily. To start, we will find an appropriate $N' = N'(\tau, [x-y]_2)$ such that $[x_{N'} - y_{N'}]_2 = [x-y]_2$ and $[x_{N'} - y_{N'}]_1 = \delta_1$. In each step, we are free to choose the phase shift ω_n^2 so that $[x_{n+1} - y_{n+1}]_2 = [x_n - y_n]_2 = \cdots = [x-y]_2$ (Lemma 5.6), and so it suffices to specify only the sequence $\omega_1^1, \ldots, \omega_{N'}^1$.

For this, set

$$N' = \left\lceil \frac{4\pi}{2\tau |\sin(\frac{[x-y]_2}{2})|} \right\rceil$$

and for each $1 \le i \le N'$, define $\omega_i^1 = \frac{[x_{i-1}+y_{i-1}]_2}{2} - \gamma$, where γ is chosen so that

$$2\tau \sin\left(\frac{[x-y]_2}{2}\right) \cos\gamma = \frac{1}{N'}(\delta_1 - [x-y]_1).$$

With this assignment, it is straightforward to check that

$$[x_i - y_i]_1 = \frac{1}{N'}(\delta_1 - [x-y]_1) + [x_{i-1} - y_{i-1}]_1 = \cdots = \frac{i}{N'}(\delta_1 - [x-y]_1) + [x-y]_1,$$

hence $[x_{N'} - y_{N'}]_1 = \delta_1$.

Now, we will define phase shifts $\omega_{N'+1}, \ldots, \omega_{N'+N''}$, $N_1 := N' + N''$ so that $[x_{N_1} - y_{N_1}]_1 = \delta_1$ and $[x_{N_1}]_2 = [y_{N_1}]_2$, completing Step (1). Using Lemma 5.6 as earlier, we can always find $\omega^1_{N'+1}, \ldots, \omega^1_{N'+N''}$ so that $[x_{N_1} - y_{N_1}]_1 = \cdots = [x_{N'} - y_{N'}]_1 = \delta_1$; thus it suffices to specify $\omega^2_{N'+1}, \ldots, \omega^2_{N'+N''}$ to achieve the desired vertical displacement. Since $[x_{N'} - y_{N'}]_1 = \delta_1 \neq 0$, we can do this by repeating the above argument on exchanging the roles played by the horizontal and vertical axes, noting that the resulting value of N'' depends only on the horizontal separation δ_1 and τ; further details are omitted.

If $[x-y]_2 = 0$, then $[x-y]_1 \neq 0$ must hold, and it is straightforward to check that $\hat{x} = f_{(0,0)}(x)$, $\hat{y} = f_{(0,0)}(y)$ have the property that $[\hat{x}-\hat{y}]_2 \neq 0$, with $[\hat{x}-\hat{y}]_2$ bounded uniformly away from 0 depending only on $[x-y]_1$ and τ. One can now apply the preceding argument with $(\hat{x}, \hat{y})$ replacing (x, y).

Step (2). To start, we will specify $M' \in \mathbb{N}$ depending only on τ and $\omega_{N_1+1}, \dots, \omega_{N_1+M'}$ so that $x_{N_1+M'} - y_{N_1+M'} = x_{N_1} - y_{N_1} = (\delta_1, 0)$ and $[x_{N_1+M'}]_1 = [\bar{x}]_1$. Again using Lemma 5.6, we can always choose $\omega^2_{N_1+1}, \dots, \omega^2_{N_1+M'}$ to leave $[x_i - y_i]_2$ unchanged for $i = N_1, N_1 + 1, \dots, N_1 + M'$, and so it suffices to specify $\omega^1_{N_1+1}, \dots, \omega^1_{N_1+M'}$. To do this, we can repeat the construction in the proof of exact controllability of the 1-point process (Lemma 5.1): set $M' = \lceil \frac{4\pi}{\tau} \rceil$ and for each $1 \le i \le M'$, choose $\omega^1_{N_1+i}$ so that

$$\tau \sin\big([x_{N_1+i-1}]_2 - \omega^1_{N_1+i}\big) = \frac{[\bar{x}]_1 - [x_{N_1}]_1}{M'}.$$

As before, $[x_{N_1+i}]_1 = [x_{N_1}]_1 + \frac{i}{M'}([\bar{x}]_1 - [x_{N_1}]_1)$, and so $[x_{N_1+M'}]_1 = [\bar{x}]_1$ as desired. Since $[x_{N_1+i}]_2 = [y_{N_1+i}]_2$ by construction for each $1 \le i \le M'$, it follows[2] that $[x_{N_1+M'} - y_{N_1+M'}]_1 = [x_{N_1} - y_{N_1}]_1 = \delta_1$, hence $[y_{N_1+M'}]_1 = [\bar{x}]_1 + \delta_1$.

Now, given $\varepsilon > 0$, we will find $M'' = M''(\tau, \delta_1, \varepsilon) \in \mathbb{N}$ and $\omega_{N_1+M'+1}, \dots, \omega_{N_1+N_2}$, $N_2 := M' + M''$, such that $\mathrm{dist}_{\mathbb{T}^1}([x_{N_1+N_2}]_2, [\bar{x}]_2) < \varepsilon$ while preserving the constraints $x_{N_1+M'+i} - y_{N_1+M'+i} = (\delta_1, 0)$ and $[x_{N_1+M'+i}]_1 = [\bar{x}]_1$ for all $1 \le i \le M''$. To do this, at each $1 \le i \le M''$, the choice

$$\omega^1_{N_1+M'+i} = [x_{N_1+M'+i-1}]_2 = [y_{N_1+M'+i-1}]_2$$

will ensure $[x_{N_1+M'+i}]_1 = [x_{N_1+M'+i-1}]_1 = \cdots = [\bar{x}]_1$ and $[y_{N_1+M'+i}]_1 = [\bar{x}]_1 + \delta_1$. Meanwhile,

$$\omega^2_{N_1+M'+i} = \frac{[x_{N_1+M'+i-1}]_1 + [y_{N_1+M'+i-1}]_1}{2} - \frac{\pi}{2} = \frac{2[\bar{x}]_1 + \delta_1}{2} - \frac{\pi}{2}$$

will ensure $[x_{N_1+M'+i}]_2 = [y_{N_2+M'+i}]_2$ (Lemma 5.6); indeed, we have that

$$[x_{N_1+M'+i}]_2 = [y_{N_2+M'+i}]_2 = [x_{N_1+M'+i-1}]_2 + \tau \cos\frac{\delta_1}{2}.$$

By irrationality of $\frac{\tau\cos(\delta_1/2)}{2\pi}$, the Weyl equidistribution theorem (see Corollary C.3 in Appendix C) implies that there exists M'' (depending only on ε, τ and δ_1) such that

$$\mathrm{dist}_{\mathbb{T}^1}\big(M''\tau\cos(\delta_1/2), [\bar{x}]_2\big) < \frac{\varepsilon}{2}.$$

Note that here, $\mathbb{T}^1$ is parametrized by $[0, 2\pi)$. This choice of M'' implies

$$\mathrm{dist}_{\mathbb{T}^2}(x_{N_1+N_2}, \bar{x}) + \mathrm{dist}_{\mathbb{T}^2}\big(y_{N_1+N_2}, \bar{x} + (\delta_1, 0)\big) < \varepsilon,$$

as desired. □

PROOF OF PROPOSITION 5.5. Fix $\varepsilon > 0$ and $(x, y), (x_\star, y_\star) \in \mathbb{T}^2 \times \mathbb{T}^2 \setminus \Delta$. Let $\delta_1 > 0$ be such that $\frac{\tau\cos(\delta_1/2)}{2\pi}$ is irrational. Observe that

$$f^{-1}_{(\beta^1,\beta^2)} = f^H_{\beta^1+\pi} \circ f^V_{\beta^2+\pi}.$$

[2]Along each horizontal line, the horizontal shear f^H_β acts rigidly for all β. That is, if $[x]_2 = [y]_2$, then $[f^H_\beta(x) - f^H_\beta(y)]_1 = [x-y]_1 = a - c$.

In particular, we can (with cosmetic changes) apply Step 1 in the proof of Lemma 5.7 to obtain $N_0 = N_0(\delta_1, \tau, d_{\mathbb{T}^2}(x_\star, y_\star))$ and $\underline{\hat{\omega}}^{N_0}$ such that $\hat{x} := (f^{N_0}_{\underline{\hat{\omega}}^{N_0}})^{-1}(x_\star)$ and $\hat{y} = (f^{N_0}_{\underline{\hat{\omega}}^{N_0}})^{-1}(y_\star)$ satisfy

$$\hat{x} - \hat{y} = (\delta_1, 0).$$

Let L_0 be a uniform (ε-independent) upper bound for $\mathrm{Lip}(f^{N_0}_{\underline{\omega}})$ and fix $\varepsilon_0 = \varepsilon/2L_0$. Let $N_0' \in \mathbb{N}, \underline{\omega}^{N_0'}$ be as in Lemma 5.7 so that (5.9) holds with the replacements $\bar{x} \mapsto \hat{x}$, $\varepsilon \mapsto \varepsilon_0$, $N \mapsto N_0'$. Define now $N = N_0 + N_0'$ and let $\underline{\omega}^N$ be the concatenation $\underline{\hat{\omega}}^{N_0}\underline{\omega}^{N_0'}$, that is,

$$\underline{\omega}^N = (\omega_1, \dots, \omega_{N_0'}, \hat{\omega}_1, \hat{\omega}_2, \dots, \hat{\omega}_{N_0}),$$

observing that by our construction, $d_{\mathbb{T}^2}(f^N_{\underline{\omega}^N}(x), x_\star) + d_{\mathbb{T}^2}(f^N_{\underline{\omega}^N}(y), y_\star) < \varepsilon$, as desired. $\square$

5.2. *Positivity of Lyapunov exponents.* The main goal is to verify the assumptions of Proposition 3.3 for the Pierrehumbert mappings $(f^n_{\underline{\omega}})$. In this setting, it suffices to show that $\exists n \geq 1$ and $\underline{\omega}^n_\star \in \Omega^n_0$ such that (i) the mapping

$$\Phi_{x_\star}(\underline{\omega}^n) := f^n_{\underline{\omega}^n}(x_\star), \quad \Phi_{x_\star} : \Omega^n_0 \to \mathbb{T}^2$$

is a submersion at $\underline{\omega}^n_\star$; and (ii) that the mapping

$$\widehat{\Phi}_{x_\star}(\underline{\omega}^n) := D_{x_\star} f^n_{\underline{\omega}^n}, \quad \widehat{\Phi}_{x_\star} : \Omega^n_0 \to SL_2(\mathbb{R})$$

has the property that $D_{\underline{\omega}^n_\star}\widehat{\Phi}_{x_\star}$ maps $\Sigma_* := \ker D_{\underline{\omega}^n_\star}\Phi_{x_\star}$ surjectively onto $T_{\widehat{\Phi}_{x_\star}(\underline{\omega}^n_\star)}SL_2(\mathbb{R})$.

We note that if property (i) holds at some $(\underline{\omega}^n_\star, x_\star)$, observe that $\Sigma_\star$ must be at most $2n-2$ dimensional, since each time increment injects two new real noise parameters. As the target space for $D_{\underline{\omega}^n_\star}\widehat{\Phi}_{x_\star}$ is three-dimensional ($SL_2(\mathbb{R})$ is 3-dimensional), property (ii) enforces the constraint $2n - 2 \geq 3$, hence $n \geq 3$ since n is an integer.

We have verified properties (i) and (ii) directly for $n = 3$ at the following values of $x_\star$ and $\underline{\omega}^3_\star$:

$$\begin{aligned} &x_\star = (\pi/2, \pi), \\ &\underline{\omega}^3_\star = (\omega^1_\star, \omega^2_\star, \omega^3_\star) = \big((0,0), (3\pi/2+1, \pi/2-1), (3\pi/2+1, 5\pi/2-2)\big). \end{aligned} \tag{5.10}$$

The computation itself is lengthy and only summarized briefly below. To start, we have

$$D_{\underline{\omega}^3_\star}\Phi_{x_\star} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 & -1 \end{pmatrix}. \tag{5.11}$$

In particular, property (i) is satisfied and $\Sigma_\star = \ker D_{\underline{\omega}^3_\star}\Phi_{x_\star}$ is 4-dimensional and spanned by the columns of the matrix

$$K = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{5.12}$$

At this $x_\star$, $\underline{\omega}^3_\star$, we have moreover that

$$M := D_{\underline{\omega}^3_\star}\widehat{\Phi}_{x_\star} = \begin{pmatrix} 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 1 & -1 & -2 & 0 & 0 & 0 \end{pmatrix}, \tag{5.13}$$

having identified the space of 2×2 real matrices with $\mathbb{R}^4$ via the parametrization

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Now, to prove $M|_{\Sigma_\star}$ surjects, it suffices to show its rank as a linear operator is at least 3; for this, it suffices to notice that

$$MK = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 0 & 0 & -2 & -1 \end{pmatrix} \tag{5.14}$$

has rank 3. Thus, (i) and (ii) hold at this choice of $(x_\star, \underline{\omega}_\star^3)$, and Proposition 3.3 applies.

PROPOSITION 5.8. *The dynamical system generated by the Pierrehumbert model* (1.7)–(1.8) *has a positive Lyapunov exponent* $\lambda_1 > 0$.

REMARK 5.9. In general, the number n of iterates necessary to use the sufficient condition in Proposition 3.3 depends on the dimension d of the phase space M and the number ℓ of degrees of freedom in the noise space Ω_0 at each time increment. To work out the precise relationship, notice that $SL_d(\mathbb{R})$ has dimension $d^2 - 1$, while $\Sigma_\star$ will be at most $\ell n - d$-dimensional, hence the inequality

$$n \geq (d^2 + d - 1)/\ell \tag{5.15}$$

must be satisfied. In particular, lowering the number of degrees of freedom ℓ in Ω_0 and/or increasing the dimension d of the phase space M forces n to rise, making this condition progressively harder to check.

5.3. *Small set property and aperiodicity.* We now check the existence of open small sets and aperiodicity for each of the one-point, projective and two-point Markov chains using Proposition 3.1 and Lemma 3.2.

5.3.1. *The one-point process.* We take $n = 1$ in Proposition 3.1; it suffices to find $x_\star$ and $\omega_\star$ such that $\Phi_{x_\star} : \Omega_0 \to \mathbb{T}^2$ is a submersion at $\omega_\star$. We set

$$x_\star = (0, 0), \qquad \underline{\omega}_\star^1 = (0, 0), \tag{5.16}$$

so that

$$D_{\underline{\omega}_\star^1} \Phi_{x_\star} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \tag{5.17}$$

has full rank. Moreover, $x_\star$ is a fixed point of $f_{\omega_\star}$, so that Lemma 3.2 implies aperiodicity, as desired.

5.3.2. *The projective process.* We apply Proposition 3.1 with $n = 2$ (notice that $n = 1$ cannot work). We verify the submersion condition for $\Phi_{(x_\star, v_\star)}(\underline{\omega}^2) = \widehat{f}_{\underline{\omega}^2}(x_\star, v_\star)$ at $\underline{\omega}_\star^2$ for

$$x_\star = (0, 0), \qquad v_\star = (1, 0), \qquad \underline{\omega}_\star^2 = (\omega_\star^1, \omega_\star^2) = ((0, 0), (\pi/2, \pi - 1)). \tag{5.18}$$

The corresponding Jacobian is given by

$$D_{\underline{\omega}^2_\star}\Phi_{(x_\star,v_\star)} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \tag{5.19}$$

where here we treat $T_{(x,v)}\mathbb{T}^2 \times S^1$ for $(x, v) \in \mathbb{T}^2 \times S^1$ as a subspace of $\mathbb{R}^4$. As one can check, this matrix has rank 3 and therefore satisfies the submersion condition.

For Lemma 3.2, we check directly that

$$x_\star = (0,0), \qquad v_\star = \frac{1}{\sqrt{10}}(\sqrt{5-\sqrt{5}}, \sqrt{5+\sqrt{5}}) \tag{5.20}$$

has the property that $\widehat{f}_{\omega_\star}(x_\star, v_\star) = (x_\star, v_\star)$.

5.3.3. *The two-point process.* We apply Proposition 3.1 with $n = 2$ and

$$x_\star = (\pi,\pi), \qquad y_\star = (0,0), \qquad \underline{\omega}^2_\star = (\omega^1_\star, \omega^2_\star) = ((0,0),(0,0)). \tag{5.21}$$

Identifying $T_{(x,y)}\mathbb{T}^2 \times \mathbb{T}^2 \cong \mathbb{R}^2 \times \mathbb{R}^2$ for $(x, y) \in \mathbb{T}^2 \times \mathbb{T}^2$, the corresponding Jacobian is

$$D_{\underline{\omega}^2_\star}\Phi_{(x_\star,y_\star)} = \begin{pmatrix} 2 & -1 & 1 & 0 \\ -3 & 2 & -1 & 1 \\ -2 & -1 & -1 & 0 \\ -3 & -2 & -1 & -1 \end{pmatrix}. \tag{5.22}$$

This matrix is invertible, hence surjective, as desired. Checking aperiodicity with Lemma 3.2, the choice

$$\omega_\star = (0,0), \qquad x_\star = (\pi,\pi), \qquad y_\star = (0,0) \tag{5.23}$$

has the property that $f_{\omega_\star}(x_\star) = x_\star$, $f_{\omega_\star}(y_\star) = y_\star$.

5.4. *Almost sure exponential mixing for the Pierrehumbert model.* Let us now summarize the proof of Theorem 1.1 in terms of conditions *(0)*–*(4)* described in Section 4.4.

- The basic assumptions in *(0)* are evident for the Pierrehumbert model $f^n_{\underline{\omega}}$ as defined in (1.7).
- Uniform geometric ergodicity of the one-point kernel P as in condition *(1)* (resp., projective kernel $\widehat{P}$ as in condition *(3)*) follows from Harris' theorem (Theorem 2.3), having checked topological irreducibility (Lemma 5.1, resp., Proposition 5.2), and the existence of open small sets and strong aperiodicity (Section 5.3.1, resp., Section 5.3.2).
- A positive Lyapunov exponent for $f^n_{\underline{\omega}}$ was checked in Proposition 5.8, so condition *(2)* is met.
- Geometric ergodicity for the two-point process as in *(4)* follows from Harris' Theorem 2.3: irreducibility was checked in Proposition 5.5 and the small set and aperiodicity properties in Section 5.3.3, while the drift condition follows from *(0)*– *(3)*.

With *(0)*–*(4)* in place, Proposition 4.6 applies, completing the proof of Theorem 1.1.

APPENDIX A: PROOF OF THEOREM 2.3

Below we sketch how Theorem 2.3 can be reduced from results in the book [43]. The following is a brief sketch of the necessary definitions and basic facts, interspersed with supplementary arguments not found in [43]. Throughout, P is the transition kernel of a Markov chain on a complete metric space X, not necessarily Feller unless otherwise stated. Recall that $\mathcal{M}(X)$ is the space of Borel probability measures on X.

A.1. T-chain property. We begin from the following definitions.

DEFINITION A.1. Let $a = \{a(n)\}_{n\geq 1}$ be a *sample distribution*, that is, $a(n) \geq 0$ for all n and $\sum_1^\infty a(n) = 1$. The *sample Markov kernel* $K_a(x, \cdot)$ is defined by

$$K_a(x, \cdot) = \sum_{n=1}^{\infty} a(n) P^n(x, \cdot), \quad x \in X. \tag{A.1}$$

Observe that with P as above, we have that $K_a(x, \cdot) \in \mathcal{M}(X)$ for all $x \in X$.

DEFINITION A.2. We say that P is a *T-chain* if there is an assignment to each $x \in X$ of a finite measure $T(x, \cdot)$ on X such that:

(i) for all $A \in \mathrm{Bor}(X)$, we have $x \mapsto T(x, A)$ is lower semicontinuous, that is, if $x_n \to x$ then $\liminf_n T(x_n, A) \geq T(x, A)$;
(ii) there exists a sample distribution a so that $K_a(x, \cdot) \geq T(x, \cdot)$ for all $x \in X$; and
(iii) $T(x, X) > 0$ for all $x \in X$.

REMARK A.3. The T-chain property is a weakening of the strong Feller property, which is equivalent to the continuity of $x \mapsto P(x, \cdot)$ in the TV norm on $\mathcal{M}(X)$ [49]. The additional flexibility of the T-chain property is quite useful in applications. For instance, it is straightforward to check from the definitions that for any $n \geq 1$, we have that if the iterated kernel P^n is a T-chain, then P is a T-chain. The same is not true for the strong Feller property.

We use the following sufficient condition for the T-chain property.

PROPOSITION A.4. *Assume P is Feller, topologically irreducible and admits an open small set. Then P is a T-chain.*

PROOF. The following is adapted from Propositions 6.2.3 and 6.2.4 in [43]. Assume $A \subset X$ is open and P^n-small for some $n \geq 1$ and measure ν_n on X. Fix the sample distribution $a(n) = 2^{-n}$, $n \geq 1$, and for $x \in X$, define

$$T(x, B) = \nu_n(B) \sum_{m=1}^{\infty} a(m) P^{m-n}(x, A), \tag{A.2}$$

using the convention that $P^0(x, \cdot) = \delta_x(\cdot)$, the unit Dirac mass at x. By the small set property, $T(x, B) \leq K_a(x, B)$ for all $B \subset X$ measurable, while $T(x, X) > 0$ for all $x \in X$ by topological irreducibility, which ensures $P^N(x, A) > 0$ for some $N = N(x, A) \in \mathbb{N}$.

Lastly, we check that $x \mapsto T(x, K)$ is lower semicontinuous for all measurable $K \subset X$. For this, it suffices to check that $x \mapsto P^N(x, A)$ is lower semicontinuous for all N. Note that $P^N(x, A) = P^N \chi_A(x)$, where χ_A is the indicator function for A. Since A is open, χ_A is lower semicontinuous, and so is the pointwise limit of an increasing sequence of continuous functions $\varphi_n : X \to \mathbb{R}$ (Exercise 4(g), p. 132 of [52]); without loss, the φ_n may be taken to be nonnegative. By the monotone convergence theorem, we have that $P^N \chi_A$ is the pointwise increasing limit of $P^N \varphi_n$ as $n \to \infty$. By the Feller property, $P^N \varphi_n$ is continuous for each n, and so we conclude that $P^N \chi_A$ is the pointwise limit of an increasing sequence of continuous functions, hence lower semicontinuous. $\square$

A.2. ψ-irreducibility. Let ϕ be a measure on X. The kernel P is said to be *ϕ-irreducible* if, whenever $\phi(K) > 0$ for some measurable $K \subset X$, we have that for all $x \in X$ there exists some $n \geq 1$ such that $P^n(x, K) > 0$.

Now, consider the sample distribution $a(n) = 2^{-n}$, $n \geq 1$. By [43], Proposition 4.2.2, if P is ϕ-irreducible then the measure

$$\psi(A) = \int_X \phi(\mathrm{d}x) K_a(x, A) \tag{A.3}$$

is such that P is ψ-irreducible, and moreover, it is *maximal* in the sense that:

(i) For any other measure ϕ' on X, P is ϕ' irreducible iff $\phi' \ll \psi$; and
(ii) If $\psi(A) = 0$, then $\psi\{x \in X : P^n(x, A) > 0 \text{ for some } n\} = 0$.

From here on, we adopt the convention in [43] of referring to P as ψ-irreducible when ψ satisfies the properties (i), (ii) above. The T-chain property allows us to check ψ-irreducibility as follows.

LEMMA A.5. *If P is both T-chain and topologically irreducible,*

(a) *it is ψ-irreducible; and*
(b) *ψ is locally positive, that is, $\psi(U) > 0$ for all nonempty open $U \subset X$.*

PROOF. Item (a) is Proposition 6.2.2 of [43]. Indeed, the proof given there implies P is ϕ-irreducible with $\phi(\cdot) = T(x, \cdot)$ for any $x \in X$, and ψ can be taken to be given by (A.3). Item (b) follows from topological irreducibility. □

The following corollary is useful in our arguments to come on periodicity.

COROLLARY A.6. *Suppose P is Feller, topologically irreducible and admits an open small set. Moreover, assume that for some $d \geq 1$ we that P^d is topologically irreducible. Then P is ψ-irreducible and P^d is ψ^d-irreducible, and $\psi^d \ll \psi$.*

PROOF. Fix an open, ν_n-small set A and let $n' \geq 0$ such that $n + n' = md$ for some integer $m \geq 1$. Observe that A is also a ν_{md}-small set with $\nu_{md}(\cdot) := (P^{n'})^* \nu_n = \int P^{n'}(x, \cdot) d\nu_n(x)$. Apply Proposition A.4 and Lemma A.5 to both P and P^d to conclude each is ψ (resp., ψ^d)-irreducible for some maximal irreducibility measure ψ (resp., ψ^d). Fixing an arbitrary $x_0 \in X$, using (A.2) to build the continuous component T out of the ν_{md}-small set A, and using (A.3) to identify a maximal irreducibility measure, we can take

$$\psi(\cdot) = 2^{md-1} \sum_{k=md}^{\infty} \sum_{\ell=1}^{\infty} 2^{-k-\ell} P^{k-md}(x_0, A) \int \nu_{md}(\mathrm{d}x) P^{\ell}(x, \cdot)$$

and

$$\psi^d(\cdot) := 2^{m-1} \sum_{k=m}^{\infty} \sum_{\ell=1}^{\infty} 2^{-k-\ell} P^{(k-m)d}(x_0, A) \int \nu_{md}(\mathrm{d}x) P^{\ell d}(x, \cdot).$$

Note that $\psi^d \ll \psi$, as desired. Note that by maximality, it follows that $\hat{\psi}^d \ll \hat{\psi}$ holds for all maximal irreducibility measures $\hat{\psi}$ for P and $\hat{\psi}^d$ for P^d. □

A.3. Petite sets. Let a be a sample distribution and let ν_a be a nontrivial measure on X. We say that a set $A \subset X$ is *ν_a-petite* if

$$K_a(x, B) \geq \nu_a(B)$$

for all measurable $B \subset X$ and for all $x \in A$. The T-chain property can be used to check petiteness.

LEMMA A.7 (Theorem 6.2.5(ii) in [43]). *If P is a ψ-irreducible T-chain, then every compact set is petite.*

We note that a small set is petite, but the converse is not true in general.

A.4. Periodicity. The following result characterizes *periodicity* of ψ-irreducible Markov chains.

LEMMA A.8 (Theorem 5.4.4 in [43]). *Assume P is ψ-irreducible and there exists a ν_n-small set $C \subset X$ with $\psi(C) > 0$. Then there exists an integer $d \geq 1$ and a disjoint collection of measurable sets $D_1, \dots, D_d$ (a "d-cycle") such that*:

(a) *For all $x \in D_i$, $i \in \{1, \dots, d\}$, we have $P(x, D_{i+1}) = 1$. Here, we follow the convention that $D_{i+kd} := D_i$ for all $k \in \mathbb{Z}$.*
(b) *We have $\psi(X \setminus \bigcup_{i=1}^d D_i) = 0$.*

Moreover, the collection $\{D_i\}$ is maximal *in the sense that if $d' \geq 1$, $\{D_i'\}_{i=1}^{d'}$ is another collection satisfying (a) and (b) above, then d' cuts d, while if $d = d'$ then up to a reordering of the i, we have that $D_i = D_i'$ up to ψ-null sets.*

When $d = 1$, we call P *aperiodic*, while if $d > 1$ then d is called the *period* of P. The following provides a useful relationship between d-cycles and small sets.

LEMMA A.9. *Suppose P is ψ-irreducible and admits a ν_n-small set A for which $\psi(A) > 0$. Assume P is periodic of period $d \geq 1$ and let $\{D_i\}$ be a d-cycle. Then there exists $i \in \{1, \dots, d\}$ such that $\psi(A \setminus D_i) = 0$, that is, $A \subset D_i$ up to a ψ-null set.*

PROOF. Suppose, for the sake of contradiction, that $\psi(A \cap D_\ell) \neq 0$ for $\ell = i$, $j \in \{1, \dots, d\}$, $i \neq j$. Then, for all $x \in A \cap D_j$, we have $P^n(x, D_{i+n}) = 0$, hence $\nu_n(D_{i+n}) = 0$. On the other hand, for all $x \in A \cap D_i$, we have

$$1 = P^n(x, D_{i+n}) = P^n(x, D_{i+n}) - \nu_n(D_{i+n}) \leq 1 - \nu_n(X) < 1. \tag{A.4}$$

This is a contradiction. Therefore, $\psi(A \cap D_i) \neq 0$ for at most one $i \in \{1, \dots, d\}$. As $\psi(X \setminus \bigcup_i D_i) = 0$, the conclusion follows. $\square$

The following can be used to rule out periodicity.

LEMMA A.10. *Assume P is Feller, topologically irreducible, admits an open small set, and that there exists $x_\star \in X$ such that $P(x_\star, U) > 0$ for all open $U \ni x_\star$ (what we call "strong aperiodiciy" in the statement of Theorem 2.3). Then P is aperiodic.*

The proof uses the following claim which we prove first.

CLAIM A.11. Under the assumptions of Lemma A.10, we have that P^n is topologically irreducible for all $n \geq 1$.

PROOF OF CLAIM. Fix $n > 1$, an open $U \subset X$ and a point $x_0 \in X$. To start, let $N = N(x_\star, U)$ be such that $P^N(x_\star, U) > 0$. Since P^N is Feller and U is open, the mapping $x \mapsto P^N(x, U)$ is lower semicontinuous (cf. the proof of Proposition A.4) and so there is an open neighborhood $U_\star$ containing $x_\star$ such that $P^N(x, U) \geq \frac{1}{2}P^N(x_\star, U) > 0$ for all $x \in U_\star$. Similarly, as $P^N(x_\star, U_\star) > 0$ by hypothesis, there is an open neighborhood $U'_\star \subset U_\star$ containing $x_\star$ such that $P^i(x, U_\star) \geq \frac{1}{2}P^i(x_\star, U_\star) > 0$ for all $i \in \{0, \dots, n-1\}$. Finally, let $N_\star = N(x_0, U'_\star)$ be such that $P^{N_\star}(x_0, U'_\star) > 0$ and observe that

$$P^{N+N_\star+i}(x_0, U) \geq \int_{\substack{y \in U_\star \\ y' \in U'_\star}} P^{N_\star}(x_0, dy')P^i(y', dy)P^N(y, U) > 0.$$

To conclude, note that there is some $i \in \{0, \dots, n-1\}$ such that $n|(N + N_\star + i)$. $\square$

PROOF OF LEMMA A.10. By Proposition A.4, P is a T-chain, hence ψ-irreducible by Lemma A.5, with $\psi(U) > 0$ for all open $U \subset X$. For the sake of contradiction, assume the period d is > 1. By Corollary A.6 and the claim, P^d is likewise ψ^d-irreducible with $\psi^d \ll \psi$.

Let $\{D_1, \dots, D_d\}$ denote a d-cycle for P (Lemma A.8). By Lemma A.9, we have $\psi(A \setminus D_i) = 0$ for some $i \in \{1, \dots, d\}$, hence $\psi^d(A \setminus D_i) = 0$. From this, we obtain $\psi^d(A \cap D_i) = \psi^d(A) > 0$. As ψ_d is an irreducibility measure for P^d, we have that for all $x \in X$ there exists $k \geq 1$ with $P^{kd}(x, D_i \cap A) > 0$. On the other hand, by periodicity we have that $P^{kd}(x, D_j \cap A) \leq P^{kd}(x, D_j) = 0$ for all $x \in D_{j-1}$, $k \geq 1$ and $j \in \{1, \dots, d\}$; this is a contradiction at $j = i$. $\square$

A.5. Harris' theorem. Our version of Harris' theorem as stated in Theorem 2.3 will be deduced from the following.

PROPOSITION A.12 (Theorem 16.1.2 of [43]). *Assume that P is ψ-irreducible and aperiodic, and moreover, that there is some petite set $C \subset X$ and a function $V : X \to [0, \infty)$ satisfying the drift condition*

$$PV \leq \alpha V + b\chi_C$$

for some $\alpha \in (0, 1)$, $b > 0$. Then P is V-uniformly geometrically ergodic.

PROOF OF THEOREM 2.3. Recall the hypotheses of Theorem 2.3: P is Feller, and moreover, (a) P admits an open small set; (b) is topologically irreducible; (c) there exists $x_\star \in X$ such that $P(x_\star, U) > 0$ for all $U \ni x_\star$ open and (d) the drift condition $PV \leq \alpha V + b\chi_C$ holds for some $V : X \to [1, \infty)$, where $\alpha \in (0, 1)$, $b > 0$ and $C \subset X$ is compact.

By (a) and (b), P is a T-chain (Proposition A.4). By Lemma A.5, the T-chain property and topological irreducibility imply P is ψ-irreducible. Lemma A.10 implies P is aperiodic. By Lemma A.7, every compact set is petite. So, the drift condition as stated in Theorem 2.3(d) implies that in Proposition A.12. $\square$

APPENDIX B: PROOFS OF PROPOSITION 2.10 AND COROLLARY 2.11

Recall that $f^n_{\underline{\omega}}$ is a continuous RDS over a compact metric space X with transition kernel P, assumed to be uniformly geometrically ergodic with stationary measure π. Moreover, $x \mapsto \mathcal{A}_{\omega,x}$ is a continuous linear cocycle of $d \times d$-matrices satisfying the integrability assumption of Theorem 2.7. Write $\lambda_1(\mathcal{A}) > \cdots > \lambda_r(\mathcal{A})$ for the Lyapunov exponents of $\mathcal{A}^n_{\underline{\omega},x}$ and $\lambda_\Sigma(\mathcal{A})$ for the summed Lyapunov exponent.

B.1. Proof of Proposition 2.10. We seek to show that if

$$d\lambda_1(\mathcal{A}) = \lambda_\Sigma(\mathcal{A}),$$

then there is a weak* continuous family $(\nu_x)_{x\in\mathrm{Supp}(\pi)}$ with the property that for $\mathbb{P}\times\pi$-almost every $(\omega, x) \in \Omega_0 \times X$, we have that

$$\nu_x = \left(\mathcal{A}^n_{\omega,x}\right)^T \nu_{f_\omega x}. \tag{B.1}$$

Equivalently, we can consider the cocycle $\mathcal{B}^n_{\underline{\omega},x} = \mathcal{B}_{\omega_n, f^{n-1}_{\underline{\omega}}x} \circ \cdots \circ \mathcal{B}_{\omega_1,x}$ generated by $\mathcal{B} : \Omega_0 \times X \to GL_d(\mathbb{R})$,

$$\mathcal{B}_{\omega,x} = (\mathcal{A}_{\omega,x})^{-T}, \qquad \mathcal{B}^n_{\underline{\omega},x} = \left(\mathcal{A}^n_{\underline{\omega},x}\right)^{-T}.$$

With this notation, it suffices to find a weak* continuous family (ν_x) such that

$$\mathcal{B}^n_{\underline{\omega},x}\nu_x = \nu_{f^n_{\underline{\omega}}x} \tag{B.2}$$

for $\mathbb{P}\times\pi$-a.e. $(\underline{\omega}, x)$ and for all $n \geq 1$. Theorem 2.7 applies to the cocycle $\mathcal{B}^n_{\underline{\omega},x}$, yielding Lyapunov exponents $\lambda_i(\mathcal{B})$, $1 \leq i \leq r'$ and a summed Lyapunov exponent $\lambda_\Sigma(\mathcal{B})$.

LEMMA B.1. *We have that $r = r'$ and*

$$\lambda_i(\mathcal{B}) = \lambda_{r-i+1}(\mathcal{A})$$

for all $1 \leq i \leq r$. Additionally, $\lambda_\Sigma(\mathcal{B}) = \lambda_\Sigma(\mathcal{A})$.

PROOF. By Remark 2.8, the Lyapunov exponents $\lambda_i(\mathcal{A})$ are the distinct values among the quantities $\chi_i(\mathcal{A})$,

$$\chi_i(\mathcal{A}) := \lim_{n\to\infty} \frac{1}{n} \log \sigma_i\left(\mathcal{A}^n_{\underline{\omega},x}\right),$$

while the multiplicity $m_i(\mathcal{A})$ of the ith Lyapunov exponent $\lambda_i(\mathcal{A})$ is given by

$$m_i = \#\{1 \leq j \leq d : \chi_j(\mathcal{A}) = \lambda_i(\mathcal{A})\}.$$

The same holds for the limits $\chi_i(\mathcal{B})$, exponents $\lambda_i(\mathcal{B})$ and multiplicities $m_i(\mathcal{B})$. Indeed, we have that

$$\chi_i(\mathcal{B}) = \lim_{n\to\infty} \frac{1}{n} \log \sigma_i\left(\mathcal{B}^n_{\underline{\omega},x}\right) = -\lim_{n\to\infty} \frac{1}{n} \log \sigma_{d-i+1}\left(\mathcal{A}^n_{\underline{\omega},x}\right)$$

on using that for an invertible matrix A, we have that

$$\sigma_i\left(A^{-T}\right) = \frac{1}{\sigma_{d-i+1}(A)}.$$

The desired conclusion follows. □

In view of Lemma B.1, we see that $d\lambda_1(\mathcal{A}) = \lambda_\Sigma(\mathcal{A})$ holds iff $d\lambda_1(\mathcal{B}) = \lambda_\Sigma(\mathcal{B})$. The latter implies the existence of a *measurable* family (ν_x) such that (B.2) holds for $\mathbb{P}\times\pi$-almost every $(\underline{\omega}, x)$ and for all $n \geq 1$. It remains to show that we can construct a weak*-continuous family of measures $(\hat{\nu}_x)$ agreeing with (ν_x) π-almost everywhere. For this, we will prove the following slightly stronger result, which holds irrespective of whether or not $d\lambda_1(\mathcal{A}) = \lambda_\Sigma(\mathcal{A})$.

LEMMA B.2. *Assume that (ν_x) is a measurable family of measures on S^{d-1} and that the relation*

$$\nu_x = \mathbb{E}(\mathcal{A}^n_{\underline{\omega},x})^T_* \nu_{f^n_{\underline{\omega}}x} \tag{B.3}$$

holds for π-almost all $x \in X$ and for all $n \geq 1$. Then there exists a weak* *measurable family $(\hat{\nu}_x)_{x\in \mathrm{Supp}(\pi)}$ satisfying the invariance relation* (B.3) *and for which*

$$\nu_x = \hat{\nu}_x \quad \textit{for } \pi\textit{-a.e. } x \in X.$$

To complete the proof of Proposition 2.10, we observe that if $d\lambda_1(\mathcal{A}) = \lambda_\Sigma(\mathcal{A})$, then (B.3) holds pointwise almost surely in $\underline{\omega}$ without the expectation $\mathbb{E}$ for the measurable family (ν_x), hence (B.3) is immediate and weak* continuity follows.

PROOF OF LEMMA B.2. Some notation: given $\varphi \in C(P^{d-1}, \mathbb{R})$ let $G_\varphi : X \times GL_d(\mathbb{R}) \to \mathbb{R}$ denote the function $G_\varphi(x, A) = \int \varphi(Av)\,\mathrm{d}\nu_x$. Given $G : X \times GL_d(\mathbb{R}) \to \mathbb{R}$, define the function $R^n G : X \to \mathbb{R}$, $n \geq 1$, by

$$R^n G(x) := \mathbb{E}G\big(f^n_{\underline{\omega}}x, (\mathcal{A}^n_{\underline{\omega},x})^T\big).$$

Lastly, we define $g_\varphi(x) := \int \varphi\,\mathrm{d}\nu_x$.

By a standard density argument, to prove Lemma B.2 it suffices to prove the following.

CLAIM B.3. For each $\varphi \in C(X, S^{d-1})$, there is a continuous function $\bar{g}_\varphi : \mathrm{Supp}\,\pi \to \mathbb{R}$ such that $g_\varphi = \bar{g}_\varphi$ holds π-almost everywhere.

To prove the claim, fix $\varphi \in C(X, S^{d-1})$. To start, by Lusin's theorem and the weak* compactness of the space of probability measures on S^{d-1}, there is an increasing sequence of compact subsets $C_1 \subset C_2 \subset \cdots$ of X such that for each n, (i) we have that $x \mapsto \nu_x$ is continuous along $x \in C_n$ and (ii) we have $\pi(C_n) \geq 1 - 1/n$.

Observe that for $G := G_\varphi$, we have that $G|_{C_n \times GL_d(\mathbb{R})}$ is continuous. By the Tietze extension theorem, for each n there is a function $G_n : X \times GL_d(\mathbb{R}) \to \mathbb{R}$ such that

$$G|_{C_n \times GL_d(\mathbb{R})} = G_n|_{C_n \times GL_d(\mathbb{R})}$$

and such that $\|G_n\|_\infty \leq \|G\|_\infty$. Lastly, define r_n to be the indicator function of $C_n \times GL_d(\mathbb{R})$ and observe that $r_n G = r_n G_n$ holds.

Now, let $g = g_\varphi$. Let $X' \subset X$ be the π-full measure set along which (B.3) holds, and observe that for $x \in X'$, we have

$$\begin{aligned}
\big|g(x) - R^n G_n(x, \mathrm{Id})\big| &= \big|R^n G(x) - R^n G_n(x)\big| \\
&\leq \big|R^n[r_n G](x) - R^n[r_n G_n](x)\big| + 2\|G\|_\infty \big|R^n(1 - r_n)(x)\big| \\
&= 2\|G\|_\infty \big|R^n(1 - r_n)(x)\big|.
\end{aligned}$$

Above, we have abused notation somewhat and treated r_n as a function on $X \times GL_d(\mathbb{R})$. Since r_n has no dependence on the $GL_d(\mathbb{R})$ coordinate, we see that

$$R^n(1 - r_n)(x) = P^n \chi_{C_n^c}(x)$$

and that by geometric ergodicity, $P^n \chi_{C_n^c} \leq (1/n) + C\gamma^n$. Therefore, $g|_{X'}$ is a uniform limit of continuous functions $R^n G_n|_{X'}$. Since $X' \subset \mathrm{Supp}\,\pi$ is dense, it follows that there is a continuous version $\bar{g} : \mathrm{Supp}\,\pi \to \mathbb{R}$ of g agreeing up to π-null sets. $\square$

B.2. Proof of Corollary 2.11. Let $(\nu_x)_{x\in\mathrm{Supp}(\pi)}$ denote the weak* continuous family of measures on S^{d-1} such that (B.1) holds. We seek to show now that for all $x \in \mathrm{Supp}(\pi)$, $n \geq 1$ and for all (y, A) in the support of the law of $(f^n_{\underline{\omega}}x, \mathcal{A}^n_{\underline{\omega},x})$, we have that

$$\nu_x = \left(A^T\right)_* \nu_y. \tag{B.4}$$

Fix such a pair (y, A). By the definition of topological support, there are sequences $\{\underline{\omega}^{(m)}\} \subset \Omega$, $\{x^{(m)}\} \subset X$ such that $(f^n_{\underline{\omega}^{(m)}}x^{(m)}, \mathcal{A}^n_{\underline{\omega}^{(m)},x}) \to (y, A)$ as $m \to \infty$ (in particular, $x^{(m)} \to x$). As (B.1) holds with full probability, without loss we may choose our sequence $\{\underline{\omega}^{(m)}\}$ so that (B.1) holds for each $(\underline{\omega}^{(m)}, x^{(m)})$. Equation (B.4) now follows on taking the limit $m \to \infty$ in the relation

$$\nu_{x^{(m)}} = \left(\mathcal{A}^n_{\underline{\omega}^{(m)},x^{(m)}}\right)^T_* \nu_{f^n_{\underline{\omega}^{(m)}}x^{(m)}}$$

and using that $x \mapsto \nu_x$ is weak* continuous.

APPENDIX C: THE WEYL EQUIDISTRIBUTION THEOREM

We use at several points the following "uniform" version of the Weyl equidistribution theorem in $d \geq 1$ dimensions. Below, $\mathbb{T}^d$ is parametrized by $[0, 2\pi)^d$ with addition defined modulo 2π.

DEFINITION C.1. We say that $\zeta = (\zeta^1, \dots, \zeta^d) \in \mathbb{R}^d$ is *rationally independent* if the only solution $a = (a_1, \dots, a_d) \in \mathbb{Q}^d$, $b \in \mathbb{Z}$ to the equation

$$a_1\zeta^1 + \cdots + a_d\zeta^d = b$$

is the trivial solution $a = (0, \dots, 0)$, $b = 0$.

Given $\zeta \in \mathbb{R}^d$, we define the mapping $T : \mathbb{T}^d \circlearrowleft$, $t \in \mathbb{R}$, by

$$T_\zeta(x) = x + 2\pi\zeta,$$

where as usual the right-hand side is considered modulo 2π in each coordinate.

THEOREM C.2 (Theorem 6.18 in [57]). *If ζ is rationally independent, then T_ζ is uniquely ergodic with invariant measure $m = \mathrm{Leb}_{\mathbb{T}^d}$.*

COROLLARY C.3 ("Uniform" Weyl equidistribution theorem). *For any $\zeta \in \mathbb{R}^d$ rationally independent and for any $\varepsilon > 0$, there exists $N = N(\zeta, \varepsilon)$ such that the following holds: for any $x, y \in \mathbb{T}^d$, there exists $n = n(x, y, \zeta, \varepsilon) \leq N(\zeta, \varepsilon)$ such that*

$$d_{\mathbb{T}^d}\left(T^n_\zeta(x), y\right) < \varepsilon.$$

PROOF. By Theorem 6.19(i) in [57], unique ergodicity of T_ζ implies that for all continuous $h : \mathbb{T}^d \to \mathbb{R}$, we have that

$$S_n h := \frac{1}{n}\sum_{i=0}^{n-1} h \circ T^i_\zeta \to \int h \,\mathrm{d}x$$

converges uniformly on $\mathbb{T}^d$. Fix $\varepsilon > 0$ and let $h_1, \dots, h_M$ be a smooth, nonnegative partition of unity with $\mathrm{Supp}(h_i) \subset B_{\varepsilon/2}(x_i)$ for some collection of centers $\{x_1, \dots, x_M\} \subset \mathbb{T}^d$. Let $N = N(\zeta, \varepsilon)$ be sufficiently large so that for all $1 \leq i \leq M$, we have that $|S_n h_i - \int h_i \,\mathrm{d}x| > \frac{1}{2}\int h_i \,\mathrm{d}x$ uniformly on $\mathbb{T}^d$.

Fix now $x, y \in \mathbb{T}^d$. Observe that $y \in B_{\varepsilon/2}(x_i)$ for some $i \in \{1, \dots, M\}$, and that $S_N h_i(x) > \frac{1}{2}\int h_i \,\mathrm{d}x > 0$. In particular, $h_i \circ T^n_\zeta(x) > 0$ for some $n = n(x, y, \zeta, \varepsilon)$. Since both y and $T^n_\zeta(x)$ lie in $B_{\varepsilon/2}(x_i)$, we conclude $d_{\mathbb{T}^d}(T^n_\zeta(x), y) < \varepsilon$. $\square$

Acknowledgments. MCZ would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programs *Mathematical aspects of turbulence* and *Frontiers in kinetic theory* where work on this paper was undertaken.

Funding. AB was supported by National Science Foundation grant DMS-2009431.

MCZ acknowledges funding from the Royal Society through a University Research Fellowship (URF\R1\191492).

REFERENCES

[1] ALBERTI, G., CRIPPA, G. and MAZZUCATO, A. L. (2019). Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.* **32** 445–490. MR3904158 https://doi.org/10.1090/jams/913

[2] ARNOLD, L. (1984). A formula connecting sample and moment stability of linear stochastic systems. *SIAM J. Appl. Math.* **44** 793–802. MR0750951 https://doi.org/10.1137/0144057

[3] AVILA, A. and VIANA, M. (2010). Extremal Lyapunov exponents: An invariance principle and applications. *Invent. Math.* **181** 115–189. MR2651382 https://doi.org/10.1007/s00222-010-0243-1

[4] AYYER, A., LIVERANI, C. and STENLUND, M. (2009). Quenched CLT for random toral automorphism. *Discrete Contin. Dyn. Syst.* **24** 331–348. MR2486579 https://doi.org/10.3934/dcds.2009.24.331

[5] BALADI, V. (2000). *Positive Transfer Operators and Decay of Correlations. Advanced Series in Nonlinear Dynamics* **16**. World Scientific, River Edge, NJ. MR1793194 https://doi.org/10.1142/9789812813633

[6] BAXENDALE, P. H. (1989). Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms. *Probab. Theory Related Fields* **81** 521–554. MR0995809 https://doi.org/10.1007/BF00367301

[7] BAXENDALE, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.* **15** 700–738. MR2114987 https://doi.org/10.1214/105051604000000710

[8] BEDROSSIAN, J., BLUMENTHAL, A. and PUNSHON-SMITH, S. (2022). Almost-sure exponential mixing of passive scalars by the stochastic Navier–Stokes equations. *Ann. Probab.* **50** 241–303. MR4385127 https://doi.org/10.1214/21-aop1533

[9] BEDROSSIAN, J., BLUMENTHAL, A. and PUNSHON-SMITH, S. (2022). Lagrangian chaos and scalar advection in stochastic fluid mechanics. *J. Eur. Math. Soc.* (*JEMS*) **24** 1893–1990. MR4404792 https://doi.org/10.4171/jems/1140

[10] BEDROSSIAN, J. and COTI ZELATI, M. (2017). Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Arch. Ration. Mech. Anal.* **224** 1161–1204. MR3621820 https://doi.org/10.1007/s00205-017-1099-y

[11] BLUMENTHAL, A., XUE, J. and YOUNG, L.-S. (2017). Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. *Ann. of Math.* (2) **185** 285–310. MR3583355 https://doi.org/10.4007/annals.2017.185.1.5

[12] BOUGEROL, P. (1988). Comparaison des exposants de Lyapounov des processus markoviens multiplicatifs. *Ann. Inst. Henri Poincaré Probab. Stat.* **24** 439–489. MR0978021

[13] BOUGEROL, P. and LACROIX, J. (1985). *Products of Random Matrices with Applications to Schrödinger Operators. Progress in Probability and Statistics* **8**. Birkhäuser, Boston, MA. MR0886674 https://doi.org/10.1007/978-1-4684-9172-2

[14] BOWEN, R. (2008). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, revised ed. *Lecture Notes in Math.* **470**. Springer, Berlin. MR2423393

[15] BRESSAN, A. (2003). A lemma and a conjecture on the cost of rearrangements. *Rend. Semin. Mat. Univ. Padova* **110** 97–102. MR2033002

[16] CARVERHILL, A. (1987). Furstenberg's theorem for nonlinear stochastic systems. *Probab. Theory Related Fields* **74** 529–534. MR0876254 https://doi.org/10.1007/BF00363514

[17] CHENG, L.-T., RAJASEKARAN, F., WONG, K. Y. J. and ZLATOŠ, A. (2023). Numerical evidence of exponential mixing by alternating shear flows. *Commun. Math. Sci.* **21** 529–541. MR4545838 https://doi.org/10.4310/cms.2023.v21.n2.a10

[18] CHILDRESS, S. and GILBERT, A. D. (1995). *Stretch, Twist, Fold: The Fast Dynamo* **37**. Springer, Berlin.

[19] CHIRIKOV, B. V. (1979). A universal instability of many-dimensional oscillator systems. *Phys. Rep.* **52** 264–379. MR0536429 https://doi.org/10.1016/0370-1573(79)90023-1

[20] COOPERMAN, W. (2022). Exponential mixing by shear flows. arXiv e-prints.

[21] COTI ZELATI, M., DELGADINO, M. G. and ELGINDI, T. M. (2020). On the relation between enhanced dissipation timescales and mixing rates. *Comm. Pure Appl. Math.* **73** 1205–1244. MR4156602 https://doi.org/10.1002/cpa.21831

[22] CRIPPA, G. and DE LELLIS, C. (2008). Estimates and regularity results for the DiPerna–Lions flow. *J. Reine Angew. Math.* **616** 15–46. MR2369485 https://doi.org/10.1515/CRELLE.2008.016

[23] CRISANTI, A., FALCIONI, M., VULPIANI, A. and PALADIN, G. (1991). Lagrangian chaos: Transport, mixing and diffusion in fluids. *Riv. Nuovo Cimento* (3) **14** 80. MR1162096 https://doi.org/10.1007/BF02811193

[24] CROVISIER, S. and SENTI, S. (2019). A problem for the 21st/22nd century. *Eur. Math. Soc. Newsl.* **114** 8–13. MR4448566

[25] DO CARMO, M. P. (1992). *Riemannian Geometry*. *Mathematics*: *Theory & Applications*. Birkhäuser, Boston, MA. MR1138207 https://doi.org/10.1007/978-1-4757-2201-7

[26] DOLGOPYAT, D. (1998). On decay of correlations in Anosov flows. *Ann. of Math.* (2) **147** 357–390. MR1626749 https://doi.org/10.2307/121012

[27] DOLGOPYAT, D., KALOSHIN, V. and KORALOV, L. (2004). Sample path properties of the stochastic flows. *Ann. Probab.* **32** 1–27. MR2040774 https://doi.org/10.1214/aop/1078415827

[28] ELGINDI, T. M. and ZLATOŠ, A. (2019). Universal mixers in all dimensions. *Adv. Math.* **356** 106807, 33. MR4008523 https://doi.org/10.1016/j.aim.2019.106807

[29] FURSTENBERG, H. (1963). Noncommuting random products. *Trans. Amer. Math. Soc.* **108** 377–428. MR0163345 https://doi.org/10.2307/1993589

[30] GROMOV, M. (1981). Groups of polynomial growth and expanding maps. *Publ. Math. Inst. Hautes Études Sci.* **53** 53–73. MR0623534

[31] HAIRER, M. and MATTINGLY, J. C. (2011). Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis*, *Random Fields and Applications VI*. *Progress in Probability* **63** 109–117. Birkhäuser/Springer Basel AG, Basel. MR2857021 https://doi.org/10.1007/978-3-0348-0021-1_7

[32] IYER, G., KISELEV, A. and XU, X. (2014). Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. *Nonlinearity* **27** 973–985. MR3207161 https://doi.org/10.1088/0951-7715/27/5/973

[33] KATO, T. (1995). *Perturbation Theory for Linear Operators*. *Classics in Mathematics*. Springer, Berlin. MR1335452

[34] KIFER, Y. (1986). *Ergodic Theory of Random Transformations*. *Progress in Probability and Statistics* **10**. Birkhäuser, Boston, MA. MR0884892 https://doi.org/10.1007/978-1-4684-9175-3

[35] KINGMAN, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* **1** 883–909. MR0356192 https://doi.org/10.1214/aop/1176996798

[36] KOBAYASHI, S. and NOMIZU, K. (1996). *Foundations of Differential Geometry. Vol. II. Wiley Classics Library*. Wiley, New York. MR1393941

[37] LEDRAPPIER, F. (1986). Positivity of the exponent for stationary sequences of matrices. In *Lyapunov Exponents* (*Bremen*, 1984). *Lecture Notes in Math.* **1186** 56–73. Springer, Berlin. MR0850070 https://doi.org/10.1007/BFb0076833

[38] LIN, Z., THIFFEAULT, J.-L. and DOERING, C. R. (2011). Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.* **675** 465–476. MR2801050 https://doi.org/10.1017/S0022112011000292

[39] LIVERANI, C. (1995). Decay of correlations. *Ann. of Math.* (2) **142** 239–301. MR1343323 https://doi.org/10.2307/2118636

[40] LIVERANI, C. (2004). On contact Anosov flows. *Ann. of Math.* (2) **159** 1275–1312. MR2113022 https://doi.org/10.4007/annals.2004.159.1275

[41] MANNING, A. (1974). There are no new Anosov diffeomorphisms on tori. *Amer. J. Math.* **96** 422–429. MR0358865 https://doi.org/10.2307/2373551

[42] MATHEW, G., MEZIĆ, I. and PETZOLD, L. (2005). A multiscale measure for mixing. *Phys. D* **211** 23–46. MR2200439 https://doi.org/10.1016/j.physd.2005.07.017

[43] MEYN, S. and TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge Univ. Press, Cambridge. MR2509253 https://doi.org/10.1017/CBO9780511626630

[44] PESIN, Y. and CLIMENHAGA, V. (2010). Open problems in the theory of non-uniform hyperbolicity. *Discrete Contin. Dyn. Syst.* **27** 589–607. MR2600681 https://doi.org/10.3934/dcds.2010.27.589

[45] PIERREHUMBERT, R. (1994). Tracer microstructure in the large-eddy dominated regime. *Chaos Solitons Fractals* **4** 1091–1110.

[46] PLASTING, S. C. and YOUNG, W. R. (2006). A bound on scalar variance for the advection-diffusion equation. *J. Fluid Mech.* **552** 289–298. MR2262118 https://doi.org/10.1017/S0022112006008639

[47] RAGHUNATHAN, M. S. (1979). A proof of Oseledec's multiplicative ergodic theorem. *Israel J. Math.* **32** 356–362. MR0571089 https://doi.org/10.1007/BF02760464

[48] ROYER, G. (1980). Croissance exponentielle de produits markoviens de matrices aléatoires. *Ann. Inst. Henri Poincaré Sect. B* (*N.S.*) **16** 49–62. MR0575176

[49] SEIDLER, J. (2001). A note on the strong Feller property. Unpublished Lecture Notes.

[50] SEIS, C. (2013). Maximal mixing by incompressible fluid flows. *Nonlinearity* **26** 3279–3289. MR3141856 https://doi.org/10.1088/0951-7715/26/12/3279

[51] SMALE, S. (1967). Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** 747–817. MR0228014 https://doi.org/10.1090/S0002-9904-1967-11798-1

[52] STROMBERG, K. R. (1981). *An Introduction to Classical Real Analysis*. AMS Chelsea Publishing, Providence, RI. MR3444105 https://doi.org/10.1090/chel/376

[53] THIFFEAULT, J.-L. (2012). Using multiscale norms to quantify mixing and transport. *Nonlinearity* **25** R1–R44. MR2876867 https://doi.org/10.1088/0951-7715/25/2/R1

[54] THIFFEAULT, J.-L., DOERING, C. R. and GIBBON, J. D. (2004). A bound on mixing efficiency for the advection-diffusion equation. *J. Fluid Mech.* **521** 105–114. MR2260450 https://doi.org/10.1017/S0022112004001739

[55] TRIEBEL, H. (1992). *Theory of Function Spaces. II. Monographs in Mathematics* **84**. Birkhäuser, Basel. MR1163193 https://doi.org/10.1007/978-3-0346-0419-2

[56] VIRCER, A. D. (1979). Matrix and operator random products. *Teor. Veroyatn. Primen.* **24** 361–370. MR0532448

[57] WALTERS, P. (1982). *An Introduction to Ergodic Theory. Graduate Texts in Mathematics* **79**. Springer, New York–Berlin. MR0648108

[58] YAO, Y. and ZLATOŠ, A. (2017). Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc.* (*JEMS*) **19** 1911–1948. MR3656475 https://doi.org/10.4171/JEMS/709

[59] YOUNG, L.-S. (1998). Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* (2) **147** 585–650. MR1637655 https://doi.org/10.2307/120960

[60] YOUNG, L.-S. (1999). Recurrence times and rates of mixing. *Israel J. Math.* **110** 153–188. MR1750438 https://doi.org/10.1007/BF02808180

[61] YOUNG, L.-S. (2013). Mathematical theory of Lyapunov exponents. *J. Phys. A* **46** 254001, 17. MR3067028 https://doi.org/10.1088/1751-8113/46/25/254001