J. Eur. Math. Soc. 24, 1893–1990 (2022) DOI 10.4171/JEMS/1140

Jacob Bedrossian · Alex Blumenthal · Sam Punshon-Smith

Lagrangian chaos and scalar advection in stochastic fluid mechanics

Received February 23, 2019; revised May 26, 2021

Abstract. We study the Lagrangian flow associated to velocity fields arising from various models of stochastic fluid mechanics. We prove that in many circumstances, these flows are chaotic, that is, the top Lyapunov exponent is strictly positive (*almost surely*, *all* particle trajectories are simultaneously exponentially sensitive with respect to initial conditions). Our main results are for the Navier–Stokes equations on $\mathbb{T}^2$ and the hyper-viscous regularized Navier–Stokes equations on $\mathbb{T}^3$ (at arbitrary fixed Reynolds number and hyper-viscosity parameters), subject to white-in-time, H^s-in-space stochastic forcing which is nondegenerate at high frequencies. Using these results, we further make a mathematically rigorous study of "passive scalar turbulence". To this end, we study statistically stationary solutions to the passive scalar advection-diffusion advected by one of these velocity fields and subjected to a white-in-time random source. We show that the chaotic behavior of Lagrangian dynamics implies a type of anomalous dissipation in the limit of vanishing diffusivity, which in turn, implies Yaglom's law of scalar turbulence – the universal scaling law analogous to the Kolmogorov 4/5 law. Key features of our study are the use of tools from ergodic theory and random dynamical systems in infinite dimensions, namely the Multiplicative Ergodic Theorem and a version of Furstenberg's Criterion, combined with hypoellipticity via Malliavin calculus and approximate control arguments.

Keywords. Lyapunov exponents, Lagrangian chaos, stochastic Navier–Stokes equations, Malliavin calculus, passive scalar turbulence

Contents

Jacob Bedrossian: Department of Mathematics, University of Maryland, College Park, MD, USA; jacob@math.umd.edu

Alex Blumenthal: Department of Mathematics, University of Maryland, College Park, MD, USA; alex123@math.umd.edu

Sam Punshon-Smith (corresponding author): Division of Applied Mathematics, Brown University, Providence, RI, USA; punshs@brown.edu

Mathematics Subject Classification (2020): Primary 60H15, 37H15, 37L55, 35Q30; Secondary 76F20, 60H07, 37A50

1. Introduction and outline

In this paper, we study the stochastic flow of diffeomorphisms $\phi^t : \mathbb{T}^d \to \mathbb{T}^d$, $t \geq 0$, defined by the random ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi^t(x) = u_t(\phi^t(x)), \quad \phi^0(x) = x. \tag{1.1}$$

Here, the random velocity field $u_t : \mathbb{T}^d \to \mathbb{R}^d$ at time $t > 0$ evolves according to one of several stochastically-forced fluid mechanics models, for example, the 2D Navier–Stokes at fixed (but arbitrary) inverse Reynolds number $\nu > 0$ on $\mathbb{T}^2$:

$$\partial_t u_t + u_t \cdot \nabla u_t = -\nabla p_t + \nu \Delta u_t + Q\dot{W}_t, \quad \nabla \cdot u_t = 0,$$

where p_t denotes the pressure at time t and $Q\dot{W}_t$ is a white-in-time, colored-in-space Gaussian process described more precisely below (Section 1.1.1).

It is expected (see for instance [35, 48]) that when u_t evolves according to either the Stokes equations (i.e., zero Reynolds number) or Navier–Stokes at arbitrary Reynolds number, the corresponding Lagrangian flows will generically be chaotic in terms of sensitivity with respect to initial conditions. This phenomenon is sometimes referred to as *Lagrangian chaos*. The primary objective of the present paper is to verify this by proving that the dynamical system defined via (1.1) possesses a strictly positive Lyapunov exponent: that is, there exists a constant $\lambda^+ > 0$, depending on the parameters of the relevant Stokes or Navier–Stokes equation, such that for every $x \in \mathbb{T}^d$ and any initial vector field in the support of the stationary measure μ for the process (u_t), we have that

$$\lim_{t\to\infty} \frac{1}{t} \log |D_x\phi^t| = \lambda^+ > 0 \quad \text{holds with probability 1.}$$

Here, $D_x\phi^t$ refers to the Jacobian matrix of $\phi^t : \mathbb{T}^d \to \mathbb{T}^d$ taken at x. This implies that almost everywhere in $\mathbb{T}^d$ and with probability 1, nearby particles are separated at an exponentially fast rate by the Lagrangian flow ϕ^t.

The study of the statistical behavior of a Lagrangian particle in a random flow has extensively been studied in the mathematics and physics literature, see for instance [35, 48, 49, 79, 81] and references therein. However, for velocities that are continuous in time and solve physical fluid models (like Stokes or Navier–Stokes), there does not appear to be any rigorous results proving Lagrangian chaos.

We further apply our Lagrangian chaos results to the "scalar turbulence" problem in the Batchelor regime of a passive scalar advected by fluid at fixed Reynolds' number in the limit of vanishing molecular diffusivity (see, e.g., [11, 48, 107] and the references therein for physics literature). This problem is both of intrinsic physical interest due to its relevance in numerous applications [107] as well as of mathematical interest, as it seems to represent the simplest physical system in fluid mechanics that displays “turbulent” behavior (see Section 1.2.2 for more precise discussion). In this manuscript, we prove that statistically stationary solutions of the passive scalar advection-diffusion equation (with random velocity fields given by the stochastic fluid models) obey the fundamental *universal* scaling law predicted by Yaglom in 1949 [110] in the vanishing diffusivity limit. Yaglom’s law is the passive scalar analogue of the Kolmogorov 4/5 law – or perhaps more accurately, the closely related 4/3 law; see [55] and the references therein. See Section 1.2 below for rigorous statements.

We also mention the authors’ follow-up works [14–16] building off the results of this manuscript. In summary, these establish a proof of Batchelor’s law for the power spectrum of a statistically stationary passive scalar in the Batchelor regime (for more details, see Remark 1.21). In comparison with hydrodynamic turbulence, this is the Batchelor regime analogue of the Kolmogorov $-5/3$ power law for the power spectral density of a turbulent fluid. To our knowledge, the proof of Yaglom’s law in this manuscript and that of Batchelor’s law in [16] are the first rigorous proof of any universal scaling or power spectrum laws in fluid mechanics for velocities arising from the Stokes or Navier–Stokes equations.

1.1. Setup and assumptions

1.1.1. Probabilistic framework. Let $\mathbb{T}^d = [0,2\pi]^d$ denote the period box. Following the convention used in [44], we define the following real Fourier basis for functions on $\mathbb{T}^d$ by

$$e_k(x) = \begin{cases} \sin(k\cdot x), & k \in \mathbb{Z}^d_+, \\ \cos(k\cdot x), & k\in \mathbb{Z}^d_-, \end{cases}$$

where

$$\mathbb{Z}^d_+ = \{(k_1,k_2,\dots,k_d)\in\mathbb{Z}^d : k_d>0\}\cup\{(k_1,k_2,\dots,k_d)\in\mathbb{Z}^d : k_1>0, k_d=0\}$$

and $\mathbb{Z}^d_- = -\mathbb{Z}^d_+$. We set $\mathbb{Z}^d_0 := \mathbb{Z}^d\setminus\{0,\dots,0\}$ and define $\{\gamma_k\}_{k\in\mathbb{Z}^d_0}$ a collection of full rank $d\times(d-1)$ matrices satisfying $\gamma_k^\top k = 0$, $\gamma_k^\top\gamma_k = \mathrm{Id}$, and $\gamma_{-k} = -\gamma_k$. Note that in dimension $d=2$, γ_k is just a vector in $\mathbb{R}^2$ and is therefore given by $\gamma_k = \pm k^\perp/|k|$. In dimension 3, the matrix γ_k defines a pair of orthogonal vectors γ_k^1,γ_k^2 that span the space perpendicular to k.

Define

$$\mathbf{L}^2 = \left\{u\in L^2(\mathbb{T}^d,\mathbb{R}^d) : \int_{\mathbb{T}^d} u\,\mathrm{d}x = 0,\ \operatorname{div} u = 0\right\}$$

to be the Hilbert space of square integrable, mean-zero, divergence-free vector fields on $\mathbb{T}^d$ and let W_t be a cylindrical Wiener process on $\mathbf{L}^2$ defined by

$$W_t = \sum_{k\in\mathbb{Z}^d_0} e_k\gamma_k W_t^k,$$

where $\{W_t^k\}_{k\in\mathbb{Z}^d_0}$ are a family of independent $(d-1)$-dimensional Wiener processes on a common canonical filtered probability space $(\Omega,\mathscr{F},(\mathscr{F}_t),\mathbf{P})$. Note that W_t is divergence free by the fact that $\gamma_k^\top k = 0$.

Let Q be a compact linear operator on $\mathbf{L}^2$ which is diagonalizable with respect to the $\mathbf{L}^2$ basis $\{e_k \gamma_k\}_{k \in \mathbb{Z}_0^d}$ with singular values $\{q_k\}_{k \in \mathbb{Z}_0^d}$ satisfying the coloring assumption

$$q_k \lesssim |k|^{-\alpha} \tag{1.2}$$

for an arbitrary, fixed $\alpha > \frac{5d}{2}$. Additionally, fix an arbitrary $\sigma > 0$ (depending on α) satisfying

$$\frac{d}{2} + 2 < \alpha - 2(d-1) < \sigma < \alpha - \frac{d}{2} \tag{1.3}$$

and define the Hilbert space

$$\mathbf{H} = \left\{ u \in H^\sigma(\mathbb{T}^d, \mathbb{R}^d) : \int_{\mathbb{T}^d} u \, \mathrm{d}x = 0, \, \operatorname{div} u = 0 \right\},$$

where $H^\sigma(\mathbb{T}^d, \mathbb{R}^d)$ denotes the space of Sobolev regular vector-fields on $\mathbb{T}^d$ (see Section 1.3 for a precise meaning when σ is not an integer). For the entirety of this paper, we will consider a white-in-time stochastic forcing $Q\dot{W}_t$, which takes the form for each $t > 0$ and $x \in \mathbb{T}^d$,

$$Q\dot{W}_t(x) = \sum_{k \in \mathbb{Z}_0^d} q_k e_k(x) \gamma_k \dot{W}_t^k.$$

Remark 1.1. The coloring assumption (1.2) and the upper bound on σ in (1.3) ensure that $\{|k|^\sigma q_k\}$ is square summable over $\mathbb{Z}_0^d$ and therefore QW_t belongs to $\mathbf{H}$ almost surely. See Remark 2.13 for a discussion of the lower bound on σ specified in (1.3).

We will also consider the following nondegeneracy condition on the low modes of the forcing. Define $\mathcal{K}$ to be the set of $k \in \mathbb{Z}_0^d$ such that $q_k \neq 0$.

Assumption 1 (Low mode nondegeneracy). *Assume $k \in \mathcal{K}$ if $|k|_\infty = 1$.*

Above, for $k = (k_i)_{i=1}^d \in \mathbb{Z}^d$ we write $|k|_\infty = \max_i |k_i|$. For several of the finite-dimensional models discussed in this paper, Assumption 1 is actually stronger than needed, i.e., the results we obtain hold with forcing on fewer modes. Sharper sufficient conditions will be specified as we go along.

For the infinite-dimensional models, we will in addition invoke the following nondegeneracy condition on all sufficiently high modes past some arbitrary finite cutoff.

Assumption 2 (High mode nondegeneracy). *There exists an $L > 0$ such that*

$$q_k \gtrsim |k|^{-\alpha} \quad \textit{for } |k|_\infty \geq L,$$

where $\alpha \in (\frac{5d}{2}, \infty)$ is as in the coloring assumption (1.2).

See Remark 1.4 for more discussion on Assumption 2.

1.1.2. Fluid mechanics models. We consider a variety of continuous-time finite-dimensional and infinite-dimensional stochastic fluid mechanical models that govern the behavior of the velocity u_t. These models are listed below as Systems 1–4.

In what follows, we write $\mathbf{H}_\mathcal{K} \subset \mathbf{H}$ for the subspace spanned by the Fourier modes $k \in \mathcal{K}$.

System 1. *We refer to the* Stokes system *in* $\mathbb{T}^d$ *($d = 2, 3$) as the following stochastic PDE for initial* $u_0 \in \mathbf{H}_{\mathcal{K}}$:

$$\begin{cases} \partial_t u_t = -\nabla p_t + \Delta u_t + Q\dot{W}_t, \\ \operatorname{div} u_t = 0, \end{cases} \tag{1.4}$$

where Q satisfies Assumption 1 and $\mathcal{K}$ is finite.

The assumption that $\mathcal{K}$ be finite is both natural (since only a few modes are required by Assumption 1), and expedient, since System 1 is effectively a finite-dimensional Ornstein–Uhlenbeck process. However, the methods of this paper applied to Systems 3 and 4 easily extend to cover System 1 when $\mathcal{K}$ is infinite and Q satisfies Assumption 2. For more details, see Remark 7.5.

System 2. *We refer to the* Galerkin–Navier–Stokes system *in* $\mathbb{T}^d$ *($d = 2, 3$) as the following stochastic ODE for* $u_0 \in \mathbf{H}_N$:

$$\begin{cases} \partial_t u_t + \Pi_N(u_t \cdot \nabla u_t + \nabla p_t) = \nu \Delta u_t + \Pi_N Q\dot{W}_t, \\ \operatorname{div} u_t = 0, \end{cases}$$

where Q satisfies Assumption 1; $N \geq 3$ is an arbitrary fixed integer; Π_N denotes the projection to Fourier modes with $|\cdot|_\infty$ norm $\leq N$; $\mathbf{H}_N$ denotes the projection of $\mathbf{H}$ under Π_N; and $\nu > 0$ is fixed and arbitrary.

Remark 1.2. We emphasize that both Systems 1 and 2 are viewed as evolution equations on the finite-dimensional state spaces $\mathbf{H}_{\mathcal{K}}$ or $\mathbf{H}_N$ and not on the larger infinite-dimensional space $\mathbf{H}$.

System 3. *We refer to the* 2D Navier–Stokes system *as the following stochastic PDE for* $u_0 \in \mathbf{H}$ *on* $\mathbb{T}^2$:

$$\begin{cases} \partial_t u_t + u_t \cdot \nabla u_t = -\nabla p_t + \nu \Delta u_t + Q\dot{W}_t, \\ \operatorname{div} u_t = 0, \end{cases}$$

where Q satisfies Assumptions 1 and 2. Here $\nu > 0$ is arbitrary and fixed.

System 4. *We refer to the* 3D hyper-viscous Navier–Stokes system *as the following stochastic PDE for* $u_0 \in \mathbf{H}$ *on* $\mathbb{T}^3$:

$$\begin{cases} \partial_t u_t + u_t \cdot \nabla u_t = -\nabla p_t - \eta \Delta^2 u_t + Q\dot{W}_t, \\ \operatorname{div} u_t = 0, \end{cases}$$

where Q satisfies Assumptions 1 and 2. Here $\eta > 0$ is arbitrary and fixed.

To unify notation across these systems, from now on we write $\hat{\mathbf{H}}$ for the relevant state space of the (u_t) process: specifically, for System 1 we set $\hat{\mathbf{H}} = \mathbf{H}_{\mathcal{K}}$; for System 2 we set $\hat{\mathbf{H}} = \mathbf{H}_N$; while for Systems 3 and 4 we set $\hat{\mathbf{H}} = \mathbf{H}$.

1.1.3. Well-posedness and stationary measures for Systems 1–4. The following summarizes the well-posedness and ergodicity properties we use for the velocity field process (u_t) in all the cases considered above.

Proposition 1.3. *For each of Systems 1–4, the following holds for any $T > 0$:*

(a) *For all functions $u \in \hat{\mathbf{H}}$ and with probability 1, there exists a unique mild solution $(u_t) \in C([0,T]; \hat{\mathbf{H}})$ with $u_0 = u$. As a function of the noise sample $\omega \in \Omega$, the solution u_t is measurable, $\mathcal{F}_t$-adapted, and belongs to $L^p(\Omega; C([0,T]; \hat{\mathbf{H}}))$ for all $p \geq 1$. Lastly, (u_t) itself is a Feller Markov process on $\hat{\mathbf{H}}$.*

(b) *The Markov process (u_t) admits a unique Borel stationary measure μ in $\hat{\mathbf{H}}$*

See Section A.1 for more details regarding well-posedness as in (a). Existence and uniqueness of the stationary measures as in part (b) can be derived from existing[1] results for each of Systems 1–4. For System 1, uniqueness is classical (it being effectively a finite-dimensional Ornstein–Uhlenbeck process), while uniqueness of the stationary measure for System 2 follows from [44, 99] (in 2D and 3D, respectively), Hörmander's theorem [67], and the Doob–Khasminskii theorem [38]. In the fully nondegenerate case $|q_k| \approx |k|^{-\alpha}$ for all k, uniqueness for System 3 follows from the classical work of Flandoli and Maslowski [52]; the extension to the hyper-viscous 3D case as in System 4 is straightforward. Arguments in [100] can be adapted to cover both Systems 3 and 4 under Assumption 2. We note also that it is not difficult to extend the work of [64] to cover both Systems 3 and 4 in the general case $|q_k| \lesssim |k|^{-\alpha}$ and satisfying the hypoellipticity conditions of [44] in 2D and [99] in 3D.

With the (u_t) process on $\mathbf{H}$, we write ϕ^t for the stochastic flow of diffeomorphisms on $\mathbb{T}^d$ solving (1.1). Note that the flow ϕ^t is a well-defined diffeomorphism since the velocity field u_t belongs to H^σ, where $\sigma > 2 + \frac{d}{2}$ (so it is at least C^2 by Sobolev embedding). This gives rise to an $\mathcal{F}_t$-adapted, Feller Markov process (u_t, x_t) on $\mathbf{H} \times \mathbb{T}^d$ defined by $x_t = \phi^t(x_0)$, where $x_0 = x$ for fixed initial $x \in \mathbb{T}^d$. We refer to (u_t, x_t) as the *Lagrangian flow process* or *Lagrangian process*. A simple check (see Lemma A.5) verifies that $\mu \times \text{Leb}$ is a stationary measure for the Lagrangian process, where Leb stands for Lebesgue measure on $\mathbb{T}^d$. Note that ergodicity of μ does *not* imply ergodicity of $\mu \times \text{Leb}$. Indeed, consider the example $\mathcal{K} = \{(1,0)\}$ with the 2D Stokes equations (1.4): in that case, $\mu \times \text{Leb}$ is not ergodic since the resulting flow is an x-dependent shear and leaves all vertical (y-dependent) lines invariant. One of the purposes of Assumption 1 is to rule out such degeneracies.

Remark 1.4. Our methods currently require some regularity properties that we do not know how to verify without the strong Feller property of the Markov semigroup associated to the (u_t, x_t) process (see Definition 4.1). In particular, the asymptotically strong Feller property [64, 65] is not enough for our purposes. It is for this reason that when treating Systems 3 and 4, we must assume nondegeneracy of the forcing in the high modes as in Assumption 2. As in [45, 52], a straightforward modification of the methods in this paper can be made to prove the strong Feller property when, in Assumption 2, the power laws in the lower and upper bound on $|q_k|$ differ by a small constant < 1.

[1] We note that uniqueness of the stationary measure as in (b) is also a consequence of the arguments in this paper; see Corollary 2.16.

Remark 1.5. Note that the forcing on the (u_t, x_t) process is necessarily degenerate, even if we had completely nondegenerate noise acting on the velocity. This is the main technical challenge in proving the strong Feller property.

1.2. Statement and discussion of results

With the preliminaries now taken care of, we are situated to state our main results on Lagrangian chaos. See Section 2 for a detailed outline of the proof.

Below, $d = 2$ or 3, and the vector field $u_t : \mathbb{T}^d \to \mathbb{R}^d, t > 0$ evolves according to one of Systems 1–4, while the Lagrangian flow $\phi^t : \mathbb{T}^d \to \mathbb{T}^d, t > 0$ is as in (1.1). Throughout, $\hat{\mathbf{H}}$ denotes the relevant vector field space for the system in question, e.g., $\hat{\mathbf{H}} = \mathbf{H}_{\mathcal{K}}$ when working with System 1. As in Proposition 1.3, μ denotes the stationary measure for the (u_t) process on $\hat{\mathbf{H}}$ for each of Systems 1, 2, 3 or 4.

Theorem 1.6 (Positive Lyapunov exponent). *Let (u_t) be governed by any of Systems 1–4. Then, there exists a deterministic constant $\lambda^+ > 0$ such that for* every *initial vector field $u_0 \in \operatorname{supp} \mu$ and $x \in \mathbb{T}^d$, the following limit exists with probability one:*

$$\lambda^+ = \lim_{t\to\infty} \frac{1}{t} \log |D_x\phi^t| > 0.$$

Indeed, as the following corollary states, with probability 1 the Lagrangian flow map ϕ^t expands all vectors at the constant exponential rate $\lambda^+ > 0$ with probability 1.

Corollary 1.7 (Norm growth of the flow map). *Let $\lambda^+ > 0$ be as in Theorem 1.6. For any $\eta \in (0, \lambda^+)$, $(u_0, x) \in \operatorname{supp} \mu \times \mathbb{T}^d$, and any unit vector $v \in \mathbb{R}^d$, there is a (random) constant $\delta = \delta(u_0, x, v, \eta)$ such that $\delta > 0$ almost-surely and for all $t > 0$,*

$$|D_x\phi^t v| \geq \delta e^{t(\lambda^+-\eta)} \quad \textit{with probability } 1.$$

Remark 1.8. Theorem 1.6 and Corollary 1.7 (and the results on scalar advection below) make *fundamental* use of the probabilistic framework. Such results seem hopelessly out of reach for deterministic models of fluid flows commonly observed in nature and many other systems of interest. For a general discussion of the difficulties involved, see, e.g., [97, 111].

A reasonable model for understanding the difficulties involved is the Chirikov Standard map [32], a one-parameter family of deterministic, discrete-time, volume-preserving mappings $\mathbb{T}^2 \to \mathbb{T}^2$ exhibiting the same stretching and folding expected to underly the mixing mechanism of the Lagrangian flow [35]. Although anticipated to be true, it is a decades-old open problem to rigorously verify, for any parameter value, that the standard map is chaotic in the sense of a positive Lyapunov exponent on a positive-volume subset of phase space. Partly explaining the difficulties involved is the fact that very different asymptotic dynamical regimes coexist in phase space: for a topologically "large" subset of parameters, the Standard map has

(1) an abundance of elliptic islands throughout phase space (inhibiting chaos) [43], and

(2) a positive Lyapunov exponent on a set of Hausdorff dimension 2 [61].

The situation is vastly different in the presence of even a small amount of noise: see [21] for positive results confirming chaos for the Standard map subjected to small-amplitude noise.

In this paper, we will apply a principle known as Furstenberg's criterion from random dynamical systems theory: this says, roughly speaking, that $\lambda^+ > 0$ as in Theorem 1.6 if the probabilistic law of the gradient $D_x\phi^t$ is sufficiently nondegenerate. See Section 2 and Section 3 for more discussion.

Remark 1.9. For Systems 1–3, Theorem 1.6 and Corollary 1.7 hold for all initial $u_0 \in \hat{\mathbf{H}}$. For the finite-dimensional System 1 (resp. System 2), the fact that $\operatorname{supp}\mu = \mathbf{H}_{\mathcal{K}}$ (resp. $\operatorname{supp}\mu = \mathbf{H}_N$) follows from hypoellipticity (see, e.g., [44, 99]) and geometric control theory [2,66]. For 2D Navier–Stokes as in System 3, that $\operatorname{supp}\mu = \mathbf{H}$ follows from [1,2]. It is likely that the same is true for 3D hyper-viscous Navier–Stokes as in System 4, but as far as the authors are aware the appropriate controllability theorems do not appear in the literature.

Remark 1.10. The techniques we use currently require well-posed SPDEs, hence the hyper-viscous regularization in System 4. We have included this case to emphasize that our infinite-dimensional methods are not restricted to two-dimensional flow – the treatment of the 3D case (System 4) is only slightly harder than 2D (System 3). In fact, the methods could extend to many settings in which one has an infinite-dimensional model coupled to finitely-many degrees of freedom on a Riemannian manifold.

Remark 1.11. Theorem 1.6 makes two claims:

(1) the Lyapunov exponent λ^+ exists and is constant over a large set of initial conditions with probability 1, and

(2) that $\lambda^+ > 0$ (a priori we only have $\lambda^+ \geq 0$ by incompressibility).

The strong Feller property (and hence our need for nondegenerate noise in the high modes) stems from the proof of (ii), while (i) only really requires the comparatively weaker fact that the (u_t, x_t) process has a unique stationary measure. Indeed, if this is known, then existence and constancy of λ^+ follows by the Furstenberg-Kesten Theorem (Proposition 2.1).

Although the strong Feller property played an important role in early investigations on uniqueness of stationary measures for infinite-dimensional Markov processes [52], it is now well understood that uniqueness holds in settings where the strong Feller property is unknown (and perhaps likely to be false – see, e.g., the discussion in [64]). A relevant example is the very recent work [71], published after this work was completed, which showed that (u_t, x_t) has a unique stationary measure when (u_t) is governed by Navier–Stokes with a spatially smooth, almost-surely bounded noise, and so for this model assertion (i) is known to hold. However, the degeneracy of the noise model considered is too restricting to establish the strong Feller property, and so it remains open whether $\lambda^+ > 0$ for the model in [71].

Remark 1.12. Although Lagrangian chaos as in Theorem 1.6 is only stated above for Systems 1–4, the methods in this paper extend to the Lagrangian flow for a broad class of velocity field processes (u_t). To summarize the general conditions one actually needs, let $\mathbf{H}$ be a space of divergence-free velocity fields which are spatially Sobolev[2] with degree $\alpha > 5d/2$. For the proof to work, we need: (see (1.9) for definitions of auxiliary processes v_t, A_t)

- The one-point process (u_t, x_t), projective process (u_t, x_t, v_t), and matrix process (u_t, x_t, A_t) are all almost-surely globally well posed, Markovian[3], and arise from continuous random dynamical systems on (respectively) $\mathbf{H} \times \mathbb{T}^d$, $\mathbf{H} \times P\mathbb{T}^d$, and $\mathbf{H} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$ (see Section 3 for more details).
- The one-point process is strong Feller and weakly irreducible.
- There exists a unique stationary measure for the projective process on $\mathbf{H} \times P\mathbb{T}^d$.
- The projective and matrix processes satisfy an *approximate control condition*; (property (C') in Definition 4.16).

Our proof uses some of the simple structure of Navier–Stokes on $\mathbb{T}^d$ to verify these properties, but with some additional effort, we believe our methods should extend to a fairly general class of strong Feller, incompressible fluid equations (with enough viscosity to be energy subcritical or critical), such as stochastically forced Boussinesq, magneto-hydrodynamics, Leray-α models, the surface quasi-geostrophic equations, etc. (all considered with enough stochastic forcing to be strong Feller).

As can be seen, approximate controllability plays a crucial role in this argument, and hence future results will benefit greatly from low-mode controllability results, e.g., [1, 2, 92, 105]; see [59] for a detailed review). We note however that due to the relatively simple structure of Navier–Stokes on $\mathbb{T}^d$, the controllability statements needed in the proof of Theorem 1.6 can be obtained “by hand” (Section 7).

Remark 1.13. For 2D Stokes as in System 1, we can prove all our results (above and below) using only the weaker noise condition (see Remark 7.5)

$$\{(1,0), (0,1), (-1,0), (0,-1)\} \subset \mathcal{K}.$$

If these are the only modes, the velocity field is given by the very simple formula

$$u(t,x) = Z_1(t)\begin{pmatrix}\sin y\\ 0\end{pmatrix} + Z_2(t)\begin{pmatrix}\cos y\\ 0\end{pmatrix} + Z_3(t)\begin{pmatrix}0\\ \sin x\end{pmatrix} + Z_4(t)\begin{pmatrix}0\\ \cos x\end{pmatrix},$$

where $Z_j, 1 \le j \le 4$ are independent Ornstein–Uhlenbeck processes (they do not need to be i.i.d., though in that case the flow is statistically homogeneous in space).

We note that Theorem 1.6 and Corollary 1.7 for the finite-dimensional models in Systems 1 and 2 follow from adaptations of previously known criteria [12, 28] (see also [56] and other citations given in Section 2.2) for positive exponents for random dynamical

[2] Although it would require more work, it is probably possible to lower the needed regularity to, e.g., $\alpha > d/2 + 2$.

[3] In fact, these processes themselves do not quite have to be Markov. For example, it suffices to have another variable Z_t such that the system (u_t, Z_t, x_t), (u_t, Z_t, x_t, v_t), (u_t, Z_t, x_t, A_t) satisfy the above conditions in a suitable manner. This is used in our more recent paper [15] to show that, for each k, there are $C_k^t C_x^\infty$ random velocities which display Lagrangian chaos.

systems generated by SDE combined with by-now standard hypoellipticity arguments for Galerkin truncations of Navier–Stokes [44, 99]. Nevertheless, we include them for the following reasons: these results are physically interesting and absent from the literature (to the best of our knowledge); they emphasize that Assumption 2 is not fundamental for Lagrangian chaos; all the ingredients needed for their proof are already required for our results on the infinite-dimensional model in System 3; and, although simpler to work with, they are instructive for the proof in the infinite-dimensional case.

On the contrary, our results for the infinite-dimensional model in Systems 3–4 do not follow from previously existing results, and require a considerable amount of additional work. See Section 2 for an outline.

1.2.1. Scalar advection. Before considering scalar turbulence, consider first the simpler problem of scalar advection without diffusivity

$$\partial_t f_t + u_t \cdot \nabla f_t = 0, \tag{1.5}$$

with (u_t) given by one of System 1–4. Here the initial datum $f_0 : \mathbb{T}^d \to \mathbb{R}$ is in H^1 with $\int f_0 \, \mathrm{d}x = 0$. By the same methods as in Proposition 1.3, the coupled system of (u_t, f_t) has a $\mathbf{P}$-a.s. unique, $\mathscr{F}_t$-adapted mild solution that defines a Feller Markov process on $\mathbf{H} \times H^1$. At times we will call (u_t, f_t) the *scalar process*. Using Theorem 1.6 and some additional work, for the (u_t, f_t) process we prove the following exponential growth of gradients with probability 1 below in Theorem 1.14. Theorem 1.14 is of intrinsic interest and is also the most crucial part of the proof of Theorem 1.16 below.

Theorem 1.14 (Exponential gradient growth without diffusivity). *Consider* (1.5) *with* (u_t) *given by any of Systems* 1–4*. Then there exists a constant* $\lambda > 0$*, depending on the system, with the following property. For any* $\eta \in (0, \lambda)$*, any fixed initial* $f_0 \in H^1 \setminus \{0\}$ *with* $\int f_0 \, \mathrm{d}x = 0$*; and for every fixed initial* $u_0 \in \operatorname{supp} \mu$*, there exists an almost-surely strictly positive random constant* $\delta = \delta(u_0, f_0, \eta) > 0$ *such that for all* $t \geq 0$ *and for* $p \in [1, \infty]$*,*

$$\|\nabla f_t\|_{L^p} \geq \delta e^{(\lambda - \eta)t} \quad \textit{with probability } 1.$$

When $d = 2$*,* $\lambda := \lambda^+$ *as in Theorem* 1.6*.*

Remark 1.15. Recently the question of mixing of scalars, i.e., decay rates in H^{-1} or mixing defined by Bressan in [26], has generated a lot of interest: see, e.g., [3, 4, 46, 70, 89, 104] and the references therein. This refinement will be addressed in future work.

1.2.2. Scalar turbulence in the Batchelor regime. Next, we are interested in studying vanishing diffusivity limits of the stationary measures associated to the following problem:

$$\partial_t g_t + u_t \cdot \nabla g_t = \kappa \Delta g_t + \tilde{Q} \dot{\tilde{W}}_t,$$

with u_t given by one of System 1–4. Here, the initial datum is $g_0 \in H^1$ and has zero mean. The (mean-zero in space) random source $\tilde{Q}\tilde{W}_t$ is of the form

$$\tilde{Q} \dot{\tilde{W}}_t = \sum_{k \in \mathbb{Z}_0^d} \tilde{q}_k e_k(x) \dot{\tilde{W}}_k(t),$$

where $\{\tilde{W}_k\}$ are an additional family of independent one-dimensional canonical Wiener processes also taken on the same filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P})$ and assumed independent of $\{W_k\}$. Define

$$\bar{\varepsilon} := \frac{1}{2} \sum_{k \in \mathbb{Z}_0^d} |\tilde{q}_k|^2 \in (0, \infty).$$

For simplicity we additionally require at least

$$\sum_{k \in \mathbb{Z}_0^d} |k|^2 |\tilde{q}_k|^2 < \infty$$

(though it is likely this condition could be dropped). Note that the random source can be very smooth and degenerate, e.g., compactly supported in frequency. Under these conditions there is a $\mathbf{P}$-a.s. unique, global-in-time, $\mathscr{F}_t$-adapted solution (u_t, g_t) which defines a Feller Markov process on $\mathbf{H} \times H^1$. Moreover, the Krylov–Bogoliubov procedure proves the existence of stationary measures $\{\bar{\mu}^\kappa\}_{\kappa > 0}$ supported on $\mathbf{H} \times H^1$ (note that all such measures satisfy $\bar{\mu}^\kappa(A \times H^1) = \mu(A)$; see Section 8 for more detail). By Itô's lemma, one verifies that statistically stationary solutions g^κ to (1.2.2) satisfy the balance relation

$$\kappa \mathbf{E} \|\nabla g^\kappa\|_{L^2}^2 = \bar{\varepsilon}. \tag{1.6}$$

As above, we are only considering g which satisfy $\int g \, \mathrm{d}x = 0$ (which is conserved due to the mean-zero assumption on $\tilde{Q}$).

The problem (1.2.2) is an idealized model for "scalar turbulence" in the Batchelor regime (see, e.g., [10, 11, 35, 48, 107]), which corresponds to the case when the velocity u is much smoother (in space) than the scalar. Passive scalar turbulence has been the subject of much research in the physics community both because of its intrinsic importance to physical applications and its potential to provide a place to develop analytic methods for understanding other turbulent systems [107]. In Batchelor's original paper [11], he considered a random straining flow as an idealized model for the small scale behavior of a passive scalar. Batchelor used this model to predict the power spectrum of the scalar, now known as *Batchelor's law* (see Remark 1.21 for more details). Later, the *Kraichnan model* was introduced in [80], wherein the velocity field is taken to be a white-in-time Gaussian field with a prescribed correlation function in space. Hence, the random ODE (1.1) is replaced by an SDE with multiplicative noise and the scalar equation (1.5) is replaced with a stochastic transport equation in Stratonovich form. There is an extensive literature on this model in physics; see, e.g., [35, 36, 107] and the references therein. For the Kraichnan model, Theorem 1.6 is proved in [13] using random dynamical systems theory developed in [12].

The questions one is interested in answering about systems such as (1.2.2) are (A) can we develop analytical theories for predicting statistical properties of small scales in the limit $\kappa \to 0$? and (B) to what extent are these statistics *universal*, that is, which properties are independent of detailed information of the system? The predictions for (A) often come in the form of quantities such as *structure functions*, for example

$$\mathbf{E}(\delta_\ell g^\kappa)^p \sim C_p |\ell|^{\zeta_p}, \quad \ell_D \lesssim |\ell| \lesssim \ell_I,$$

where

$$\delta_\ell g(x) := g(x+\ell) - g(x),$$

(where the meaning of $\sim$ is left informal for now) for a range of scales (ℓ_D, ℓ_I) known as the *inertial range* (here ℓ_D is for *dissipative* and ℓ_I is for *integral*) assumed to satisfy $\lim_{\kappa \to 0} \ell_D(\kappa) = 0$ and ℓ_I much smaller than the length-scales of the large scale forcing in the system (but independent of κ). For (B), the corresponding question is then to answer: For which p are the quantities ζ_p, C_p and/or ℓ_D universal? The first predictions of this general type were due to Kolmogorov [76–78] in 1941, who studied the 3D Navier–Stokes equations as $\nu \to 0$. One of the cornerstones of the theory is that in 3D Navier–Stokes, the kinetic energy, a conserved quantity for $\nu = 0$, is transferred to higher and higher frequencies as $\nu \to 0$ (called a *direct cascade* [55]), where it is then dissipated by viscosity so that the dissipation rate is non-vanishing as $\nu \to 0$. This property is called *anomalous dissipation*. In statistically stationary solutions, the flux of kinetic energy through the inertial range is predicted to be constant (as a function of ℓ and asymptotically, as a function of ν), resulting in the celebrated prediction now known as the Kolmogorov 4/5 law:

$$\mathbf{E}\left(\delta_\ell u \cdot \frac{\ell}{|\ell|}\right)^3 \sim -\frac{4}{5}\varepsilon|\ell|, \quad \ell \in (\ell_D, \ell_I),$$

where ε is the net input of energy per unit time and volume (see [17,55] and the references therein for more discussions). Indeed, the quantity appearing on the left-hand side can be related to the energy flux through scale ℓ. The 4/5 law is equivalent to the similar 4/3 law:

$$\mathbf{E}\left(|\delta_\ell u|^2 \delta_\ell u \cdot \frac{\ell}{|\ell|}\right) \sim -\frac{4}{3}\varepsilon|\ell|, \quad \ell \in (\ell_D, \ell_I).$$

As they are equivalent we will not distinguish these laws and simply refer to the pair together as the "4/5 law". The 4/5 law is very well matched by experiments, is considered the only exact law of 3D turbulence, and is almost completely universal[4]. See [55] for a detailed introduction to turbulence in the 3D Navier–Stokes equations.

Since Kolmogorov's work, analogous (to varying degrees) "turbulent" dynamics have been studied in a variety of physical systems, such as 2D Navier–Stokes turbulence [22], the many varied regimes of magneto-hydrodynamic turbulence [20], wave turbulence in dispersive equations [91, 113], and passive scalar turbulence [48, 107]. The anomalous dissipation of a quantity that is conserved in the limit of zero dissipative effects is one of the defining characteristics of such turbulent systems. Accordingly, in all such examples, the flux of these conserved quantities through the inertial range converges to a constant, and so for each anomalously dissipated conservation law one expects a universal scaling law analogous to the 4/5 law. In 1949, Yaglom [110] found this law for the "enstrophy" flux[5] of the passive scalar

$$\mathbf{E}\left(|\delta_\ell g^\kappa|^2 \delta_\ell u \cdot \frac{\ell}{|\ell|}\right) \sim -\frac{4}{d}\bar{\varepsilon}|\ell|, \quad \ell \in (\ell_D, \ell_I).$$

[4] Both the constant and the exponent are universal; it is not clear whether ℓ_D is universal.

[5] For simplicity, we will refer to the L^2 density $\|g\|^2$ as the "enstrophy" in analogy with the vorticity form of the 2D Navier–Stokes equations.

This is the law we confirm for (1.2.2) (in a spherically averaged sense); see Theorem 1.16 below for the rigorous meaning of $\sim$ in this statement.

In [17], it is proved that the Kolmogorov 4/5 law follows for statistically stationary solutions of the 3D Navier–Stokes using that $\lim_{\nu\to 0} \nu \mathbf{E}\|u^\nu\|_{L^2}^2 = 0$. This property is referred to therein as "weak anomalous dissipation"[6], and is a natural form of anomalous dissipation for statistically stationary solutions (see [17] for more discussion). In this work, we use Theorem 1.14 to prove the analogous statement here ((1.7) below) by adapting arguments from [17]; see Section 8 for details. Then Yaglom's law, as stated in (1.8), follows from a straightforward variation of the argument in [17]. The limit (1.7) cannot hold if solutions to (1.2.2) remain concentrated in low frequencies in the limit $\kappa \to 0$; indeed, in this case it is easy to check that $\kappa \mathbf{E}\|g^\kappa\|_{L^2}^2 \gtrsim 1$ (see also Remark 1.18 below). For (1.7) to hold, the fluid needs to induce a direct cascade of enstrophy to successively smaller scales where it is more efficiently dissipated by the $\kappa \Delta g^\kappa$ term, resulting in a much-enhanced dissipation rate. It is Theorem 1.14 that ultimately implies the Lagrangian flow-map creates the requisite small scales everywhere in the domain with probability 1. See also the earlier work using norm growth in the inviscid passive scalar problem to obtain "enhanced dissipation" effects for $\kappa > 0$ models [33, 114] and the recent related work [34].

We remark that the idea that Lagrangian chaos and scalar turbulence scaling laws should be intimately related has long been expected by the physics community; see, e.g., [6, 7, 107, 112] and the references therein for more information.

Theorem 1.16 (Scalar turbulence in the Batchelor regime). *Let $\{u, g^\kappa\}_{\kappa>0}$ be a sequence of statistically stationary solutions to (1.2.2) with (u_t) given by any of Systems 1–4. Then:*

(i) *the Weak Anomalous Dissipation property holds:*

$$\lim_{\kappa\to 0} \kappa \mathbf{E}\|g^\kappa\|_{L^2}^2 = 0. \tag{1.7}$$

(ii) *Yaglom's law holds over a suitable inertial range: that is, for all $\kappa > 0$ small, there exists an $\ell_D(\kappa) > 0$ with $\lim_{\kappa\to 0} \ell_D(\kappa) = 0$ such that*

$$\lim_{\ell_I\to 0} \limsup_{\kappa\to 0} \sup_{\ell\in[\ell_D,\ell_I]} \left| \frac{1}{\ell}\mathbf{E} ⨍_{\mathbb{T}^d} ⨍_{\mathbb{S}^{d-1}} |\delta_{\ell n} g^\kappa|^2 \delta_{\ell n} u \cdot n \,\mathrm{d}S(n)\,\mathrm{d}x + \frac{4}{d}\bar{\varepsilon} \right| = 0. \tag{1.8}$$

Here n denotes a unit vector in $\mathbb{S}^{d-1}$ and $⨍$ denotes the average, i.e.,

$$⨍_{\mathbb{T}^d} \cdot \,\mathrm{d}x = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \cdot \,\mathrm{d}x \quad \textit{and} \quad ⨍_{\mathbb{S}^{d-1}} \cdot \,\mathrm{d}S(n) = \frac{1}{\omega_{d-1}} \int_{\mathbb{T}^d} \cdot \,\mathrm{d}x,$$

where ω_{d-1} denotes the surface area of the $\mathbb{S}^{d-1}$ (2π in 2D and 4π in 3D).

Remark 1.17. Note that by time stationarity, (1.8) is the same as asserting the expected value of arbitrary length time averages follow Yaglom's law. Further, as in [17], if one

[6] We remark that this property is equivalent to the assertion that the Taylor microscale goes to zero as Reynolds number goes to infinity; see [17] for details.

assumes Q and $\tilde{Q}$ are spatially homogeneous, then there exists spatially homogeneous statistically stationary solutions to the system (u_t, g_t) and one can remove the x average from (1.8), that is, (1.8) holds a.e. in x (existence of spatially homogeneous statistical stationary solutions is classical in this case; see, e.g., [86, Corollary 3.3.4], but the statements in Theorem 1.16 were not previously known).

Remark 1.18. Note that by the balance (1.6), the weak anomalous dissipation property (1.7), and Sobolev interpolation, there holds $\lim_{\kappa \to 0} \kappa \mathbf{E}\|g^\kappa\|_{H^\gamma}^2 = 0$ for all $\gamma \in (0, 1)$ and $\lim_{\kappa \to 0} \kappa \mathbf{E}\|g^\kappa\|_{H^\gamma}^2 = +\infty$ for all $\gamma > 1$.

Remark 1.19. There exist a number of works providing sufficient conditions to deduce turbulence scaling laws and related results, see for example [17, 18, 51] in the stochastic case and [31, 42, 47, 93] in deterministic cases. See the references therein for earlier work in the physics literature. However, no works prior to Theorem 1.16 have provided a proof of such laws without sufficient conditions that remain unverified as of writing. This is natural, as Yaglom's law in this relatively simple, fixed Reynolds number, Batchelor regime is likely to be by far the simplest assertion (to prove) of all statements in statistical theories of turbulence in fluid mechanics.

Remark 1.20. The proof of Theorem 1.16 is based on first replacing

$$f^\kappa(t, x) = \sqrt{\kappa} g(t, x)$$

so that f^κ now solves a passive scalar equation with fluctuation-dissipation-type scaling: $\partial_t f_t^\kappa + u_t \cdot \nabla f_t^\kappa = \kappa \Delta f_t^\kappa + \sqrt{\kappa} \tilde{Q} d\tilde{W}_t$ and studying the $\kappa \to 0$ limit. This fluctuation-dissipation-type scaling is long-understood to be important in many contexts in PDEs and has been studied numerous times; see, e.g., [17, 53, 69, 83–86, 106].

Remark 1.21. Building off the results in this manuscript in a series of follow-up works [14, 15] culminating in [16], we have proved Batchelor's prediction on the cumulative power spectrum of the passive scalar in the same setting as [11]. This assertion amounts to: there exist N_0, C_0 both independent of κ such that for all $N_0 < N < \kappa^{-1/2}$, there holds

$$\frac{1}{C_0} \log N < \sum_{|k| \le N} \mathbf{E}|\hat{g}(k)|^2 \le C_0 \log N.$$

Note that Yaglom's law only requires Theorem 1.6, and roughly speaking, the corresponding fact that chaotic Lagrangian flow pushes at least some of the L^2 mass of a passive scalar from large to small scales (Theorem 1.14). In comparison, Batchelor's prediction for the power spectrum is more refined, and requires that *all large-scale L^2 mass* be "evacuated" to higher frequencies. This additional statement does not follow just from Theorem 1.6, instead requiring the arguments in the authors' follow-up works [14–16] (all completed after this work).

1.3. A guide to notation

We include here for the convenience of the reader a guide to commonly used notation in the paper.

- We use the notation $f \lesssim g$ if there exists a constant $C > 0$ such that $f \le Cg$, where C is independent of the parameters of interest. Sometimes we use the notation
$$f \lesssim_{a,b,c,\dots} g$$
to emphasize the dependence of the implicit constant on the parameters, for example, $C = C(a, b, c, \dots)$. We denote $f \approx g$ if $f \lesssim g$ and $g \lesssim f$.
- Throughout, $\mathbb{R}^d$ is endowed with the standard Euclidean inner product $(\cdot, \cdot)$ and corresponding norm $|\cdot|$. We continue to write $|\cdot|$ for the corresponding matrix norm. We use $|k|_p$ to denote the ℓ^p norms.
- When the domain of the L^p space is omitted, it is always understood to be $\mathbb{T}^d$:
$$\|f\|_{L^p} = \|f\|_{L^p(\mathbb{T}^d)}.$$
We use the notations $\mathbf{E}X = \int_\Omega X(\omega)\mathbf{P}(\mathrm{d}\omega)$ and $\|X\|_{L^p(\Omega)} = (\mathbf{E}|X|^p)^{\frac{1}{p}}$. We use the notation
$$\|f\|_{H^s}^2 = \sum_{k\in\mathbb{Z}^d} |k|^{2s}|\hat{f}(k)|^2$$
(denoting $\hat{f}(k) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{T}^d} e^{-ik\cdot x} f(x)dx$ the usual complex Fourier transform).
- If M is a Riemannian manifold, we write Leb_M for the Lebesgue volume on M. For short, we write Leb for the normalized Lebesgue measure on $\mathbb{T}^d$.
- For $d \ge 1$, we write $M_{d\times d}(\mathbb{R})$ for the space of real $d \times d$ matrices, and $\mathrm{SL}_d(\mathbb{R})$ for the subgroup of matrices of determinant 1.
- We write $P^{d-1} = P(\mathbb{R}^d)$ for the real projective space of $\mathbb{R}^d$, i.e., the manifold of equivalence classes of vectors in $\mathbb{R}^d \setminus \{0\}$ up to scaling. When it is clear from context, we will abuse notation and intentionally confuse an element $v \in P^{d-1}$ with a unit vector representative $v \in \mathbb{R}^d$, and vice versa. Likewise $\mathbb{S}^{d-1}$ denotes the unit sphere in $\mathbb{R}^d$.
- Given a matrix $B \in M_{d\times d}(\mathbb{R})$, we use the same symbol $B : P^{d-1} \to P^{d-1}$ to denote the corresponding map on projective space. If ν is a probability measure on P^{d-1}, we write $B_*\nu := \nu \circ B^{-1}$ for the pushforward of ν by B.
- We denote $\mathbf{L}^2$ the Hilbert space of square integrable, divergence-free, mean zero functions.
- For $\sigma \in (\alpha - 2(d-1), \alpha - \frac{d}{2})$ fixed, we write $\mathbf{H}$ for the subspace of H^σ divergence-free, mean-zero vector fields on $\mathbb{T}^d$, $d = 2$ or 3. Given $N \ge 1$ as in System 2, we write $\mathbf{H}_N \subset \mathbf{H}$ for the span of all Fourier modes k with $|k|_\infty \le N$. Given $\mathcal{K} \subset \mathbb{Z}^d$ as in Assumption 1, we write $\mathbf{H}_\mathcal{K} \subset \mathbf{H}$ for the span of all Fourier modes in $\mathcal{K}$.
- Given the vector field process (u_t) on $\hat{\mathbf{H}}$ governed by Systems 1, 2, 3 or 4, we write (u_t, x_t) for the *Lagrangian process* on $\hat{\mathbf{H}} \times \mathbb{T}^d$ as defined by $x_t = \phi^t(x_0)$, ϕ^t as in (1.1), where $\hat{\mathbf{H}}$ is the appropriate space of vector fields as above. We write $\Theta^t_\omega : \hat{\mathbf{H}} \times \mathbb{T}^d \to \hat{\mathbf{H}} \times \mathbb{T}^d$, $t \ge 0$, for the corresponding RDS as defined in Section 2.1. We write (u_t, x_t, v_t) for the *projective process* on $\hat{\mathbf{H}} \times \mathbb{T}^d \times P^{d-1}$ as defined in Section 2.5, and (u_t, x_t, A_t) for the *matrix process* on $\hat{\mathbf{H}} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$ as defined

in Section 2.3. These processes are governed by (u_t) as in Systems 1–4 and the random ODE

$$\partial_t x_t = u_t(x_t), \tag{1.9a}$$
$$\partial_t v_t = \Pi_{v_t} Du_t(x_t) v_t, \tag{1.9b}$$
$$\partial_t \check{v}_t = -\Pi_{\check{v}_t} Du_t(x_t)^\top \check{v}_t, \tag{1.9c}$$
$$\partial_t A_t = Du_t(x_t) A_t, \tag{1.9d}$$

where $\Pi_v = \mathrm{Id} - v \otimes v$ is the orthogonal projection from $\mathbb{R}^d$ onto the tangent space of $\mathbb{S}^{d-1}$ (viewing v as a unit vector in $\mathbb{R}^d$).

- We denote by $B(u,u) = (I + \nabla(-\Delta)^{-1}\nabla\cdot)(u \cdot \nabla u)$ the Euler nonlinearity in both 2D and 3D. We similarly denote $A = -\nu\Delta + \eta\Delta^2$ in 3D and $A = -\nu\Delta$ in 2D.
- In Section 6, we will frequently denote $w_t = (u_t, x_t, v_t, z_t) \in \mathbf{H} \times \mathcal{M}$ the augmented projective process with z_t being a standard Wiener process on $\mathbb{R}^{2d}$. Here, $\mathcal{M}$ is the finite-dimensional manifold $\mathcal{M} = \mathbb{T}^d \times P^{d-1} \times \mathbb{R}^{2d}$. We will also denote w_t^ρ the ρ cut-off version (sometimes dropping the ρ when we do not care to distinguish between the cut-off and non-cut-off versions.)
- We denote $T_v\mathcal{M}$ the tangent space of $\mathcal{M}$ at (x,v,z) (suppressing the dependence on (x,z) since they are flat. For a given direction in $h \in \mathbf{H} \times T_v$ we write $Dw_t h$ for the derivative of the process with respect to the initial data in the direction h.
- In Section 6, for a fixed $T > 0$, write $\Omega_T = C([0,T]; \mathbf{H}^{-\varepsilon} \times \mathbb{R}^{2d})$ for the path space for the noise driving the process w_t, where $\varepsilon > 0$ is such that $Q\mathbf{H}^{-\varepsilon} \subseteq \mathbf{H}$. Correspondingly, we denote $\mathcal{D}_g$ the direction Malliavin derivative in a "Cameron–Martin" direction $g \in L^2([0,T]; \mathbf{L}^2 \times \mathbb{R}^{2d})$. More generally, we denote by $\mathcal{D}$ the Malliavin derivative operator, and for a separable Hilbert space $\mathcal{H}$, we write $\mathbb{W}^{1,2}(\Omega_T; \mathcal{H})$ for the associated domain of $\mathcal{H}$-valued Malliavin differentiable random variables with norm
$$\|X\|^2_{\mathbb{W}^{1,2}(\Omega_T;\mathcal{H})} = \mathbf{E}\|X\|^2_{\mathcal{H}} + \mathbf{E}\|\mathcal{D}X\|^2_{L^2([0,T];\mathbf{L}^2\times\mathbb{R}^{2d})\otimes\mathcal{H}}.$$
The adjoint operator $\mathcal{D}^*$ (the Skorokhod integral) is denoted by
$$\int_0^T \langle g, \delta W_t\rangle_{\mathbf{L}^2} = \mathcal{D}^* g$$
for each $g \in \mathbb{W}^{1,2}(\Omega_T; L^2([0,T]; \mathbf{L}^2 \times \mathbb{R}^{2d}))$.

2. Outline of the proofs

Let us now give a somewhat detailed outline for the proofs of the main results of this paper, starting with Theorem 1.6.

The basic structure of the proof can be summarized in two main points:

(1) The Multiplicative Ergodic Theorem and a variant of Furstenberg's criterion show that, given suitable ergodic properties of the dynamics, the Lyapunov exponent is strictly positive unless there is a certain *almost surely* invariant structure in the motion of $x_t = \phi^t(x_0)$ and the gradient $D_{x_0}\phi^t$.

(2) Hypoellipticity and approximate controllability arguments show that (A) the dynamics satisfy suitable ergodic properties and that (B) a rich range of motions of x_t and $D_{x_0}\phi^t$ are realized. This will rule out the invariant structure and allow us to deduce a positive Lyapunov exponent as in Theorem 1.6.

As we will see below, both are significantly harder in the infinite-dimensional case (Systems 3–4).

2.1. *The RDS framework and the Multiplicative Ergodic Theorem*

Theorem 1.6 makes two assertions:

(i) that the limit defining the Lyapunov exponent λ^+ exists and is constant almost surely, and

(ii) that this exponent satisfies $\lambda^+ > 0$.

Let us first outline how to prove assertion (i) using tools from random dynamical systems theory.

To start, we must formulate the Lagrangian process (u_t, x_t) as a stochastic flow or *random dynamical system* (RDS) on $\hat{\mathbf{H}} \times \mathbb{T}^d$ (here, $\hat{\mathbf{H}}$ is as in the beginning of Section 1.2). That is, given a random noise path $\omega \in \Omega$ and a fixed initial $(u_0, x_0) \in \hat{\mathbf{H}} \times \mathbb{T}^d$, the assignment $(u_0, x_0) \mapsto (u_t, x_t)$ is realized as

$$(u_t, x_t) = \Theta^t_\omega(u_0, x_0),$$

where $\Theta^t_\omega : \hat{\mathbf{H}} \times \mathbb{T}^d \to \hat{\mathbf{H}} \times \mathbb{T}^d$ is a continuous mapping depending measurably on the noise parameter ω (see Section 3.1.1 for details). In our setting, Θ^t_ω is of the form

$$\Theta^t_\omega(u, x) = (\mathcal{U}^t_\omega(u), \phi^t_{\omega,u}(x)),$$

where $\mathcal{U}^t_\omega : \hat{\mathbf{H}} \to \hat{\mathbf{H}}$ is the time-t mapping associated to the equation governing (u_t) (any of Systems 1–4), i.e., the map sending $u_0 \mapsto u_t$, and $\phi^t_{\omega,u} = \phi^t : \mathbb{T}^d \to \mathbb{T}^d$ is the time-t Lagrangian flow map associated to the noise parameter ω and the initial vector field $u \in \hat{\mathbf{H}}$ as in (1.1), i.e., the diffeomorphism on $\mathbb{T}^d$ sending $x_0 \mapsto x_t$. In the context of RDS, the matrix-valued mapping $\Omega \times \hat{\mathbf{H}} \times \mathbb{T}^d \to M_{d\times d}(\mathbb{R})$ sending $(\omega, u, x) \mapsto D_x\phi^t_{\omega,u}$ for fixed $t > 0$ is an object known as a *linear cocycle* over the RDS Θ^t_ω.

For more background on random dynamics and a precise enumeration of the assumptions involved, see Sections 3.1 and 3.2, where the relevant theory and assumptions are spelled out for an abstract RDS $\mathcal{T}$ acting on a metric space Z and a linear cocycle $\mathcal{A}$ over $\mathcal{T}$. Throughout Section 3 we intend to apply this with $\mathcal{T}$ replaced by the Lagrangian flow Θ acting on $Z = \hat{\mathbf{H}} \times \mathbb{T}^d$ with $\mathcal{A}$ replaced by the gradient cocycle $D_x\phi^t$. It is straightforward to verify the assumptions made in Sections 3.1–3.2 for Θ and $D_x\phi^t$; this is carried out in the Appendix (Section A.1).

A fundamental result pertaining to linear cocycles is the Multiplicative Ergodic Theorem, stated in full in Section 3.2.2 as Theorem 3.13. For the purposes of this discussion, we state below the following consequence, often referred to as the Furstenberg–Kesten Theorem [57].

Proposition 2.1. *The limit*

$$\lambda^+(\omega, u, x) := \lim_{t\to\infty} \frac{1}{t} \log |D_x \phi^t_{\omega,u}|$$

exists for $\mathbf{P}$*-a.e.* ω *and* $\mu \times \mathrm{Leb}$*-a.e.* $(u, x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$*, where* μ *is the stationary measure for the* (u_t) *process as in Proposition 1.3. Moreover, if* $\mu \times \mathrm{Leb}$ *is an ergodic stationary measure (Definition 3.9) for the Lagrangian process* (u_t, x_t)*, then the limiting value* λ^+ *does not depend on* (ω, u, x)*.*

Ergodicity of $\mu \times \mathrm{Leb}$ as a stationary measure for the Lagrangian process (u_t, x_t) is a necessary ingredient for Theorem 1.6. See Section 2.7 below for a discussion of the ergodic properties of the (u_t, x_t) process.

Remark 2.2. Note that in Theorem 1.6, the Lyapunov exponent λ^+ is asserted to exist with probability 1 at *every* initial $(u, x) \in \operatorname{supp} \mu \times \mathbb{T}^d$, as opposed to $\mu \times \mathrm{Leb}$-almost every (u, x) as in Proposition 2.1. The strong Feller property (Definition 4.1) for the (u_t, x_t) process allows us to pass between these formulations: see Lemma 4.2 (b) in Section 4.

2.2. Determining positive Lyapunov exponents: Furstenberg's criterion

An entirely separate matter is to verify that λ^+ as in Proposition 2.1 is strictly positive. This problem is notoriously difficult (see Remark 1.8 above). Aiding us, however, is the fact that the cocycle $(\omega, u, x) \mapsto D_x \phi^t_{\omega,u}$ is subjected to some noise. For such cocycles, a powerful tool known as Furstenberg's criterion implies $\lambda^+ > 0$ under suitable non-degeneracy conditions described in detail below. The criterion was originally obtained in [57] for IID products of matrices, and extended in scope by various authors in the ensuing years: see, e.g., [9,12,60,62,88], and also the citations of Chapter 1 of [25] for a more complete bibliography.

Ignoring for now the requisite quantifiers and other details, the relevant version of Furstenberg's criterion can be stated as follows. Proposition 2.3 below is a version of the criterion given in [88], and will be stated in full as Theorem 3.18 in Section 3.3. Below, $P^{d-1} = P(\mathbb{R}^d)$ denotes the manifold of one-dimensional subspaces of $\mathbb{R}^d$.

Proposition 2.3 (Informal Furstenberg criterion). *Assume* $\mu \times \mathrm{Leb}$ *is an ergodic stationary measure for the Lagrangian process* (u_t, x_t)*. If* $\lambda^+ = 0$*, then to each* $(\mu \times \mathrm{Leb})$*-generic* (u, x)*, there is associated a deterministic (i.e.,* ω*-independent) probability measure* $\nu_{u,x}$ *on* P^{d-1} *with the property that*

$$(D_x \phi^t_{\omega,u})_* \nu_{u,x} = \nu_{\Theta^t_\omega(u,x)} \tag{2.1}$$

for all $t > 0$ *and* $\mathbf{P} \times \mu \times \mathrm{Leb}$*-almost all* $(\omega, u, x) \in \Omega \times \mathbf{H} \times \mathbb{T}^d$*.*

To prove $\lambda^+ > 0$, then, it suffices to obtain a contradiction from the conclusions of Proposition 2.3.

Conceptually, the measures $\nu_{u,x}$ should be thought of as deterministic "configurations" of vectors on $\mathbb{R}^d$, and relation (2.1) says that this (u, x)-dependent family $(\nu_{u,x})$

of deterministic "configurations" is left invariant by the Jacobian matrices $D_x\phi^t_{\omega,u}$ with probability 1. As such, relation (2.1) has the connotation of a *degeneracy* in the probabilistic law of the matrices $D_x\phi^t_{\omega,u}$ with ω distributed as $\mathbf{P}$.

2.3. *Ruling out Furstenberg's criterion: Finite-dimensional models*

Let us show how one can rule out (2.1) in Furstenberg's criterion. To start, we recall the fact that given a fixed pair of probability measures ν, ν' on P^{d-1}, the set of matrices $M \in \mathrm{SL}_d(\mathbb{R})$ for which $M_*\nu = \nu'$ has empty interior (Lemma 3.19). Thus, (2.1) can be ruled out if we check that for a positive measure set of $(u,x), (u',x') \in \hat{\mathbf{H}} \times \mathbb{T}^d$, the support of the probabilistic law of the process $A_t := D_{x_0}\phi^t_{\omega,u_0}$ conditioned on the event $(u_0,x_0) = (u,x)$, $(u_t,x_t) = (u',x')$ has a nonempty interior in $\mathrm{SL}_d(\mathbb{R})$ (e.g., when the conditional law is absolutely continuous).

For the finite-dimensional models in Systems 1 and 2, we can compute this conditional law explicitly. The matrix-valued process $A_t := D_x\phi^t_{\omega,u}$ is a component of the Markov process (u_t, x_t, A_t) generated by the (u_t) together with (1.1) and

$$\partial_t A_t = Du_t(x_t)A_t \tag{2.2}$$

on the finite-dimensional manifold $\mathcal{M} := \hat{\mathbf{H}} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$.

Under suitable nondegeneracy conditions on the SDE governing (u_t, x_t, A_t), for instance, Hörmander's condition as described in Section 2.7, the law $Q_t((u,x,\mathrm{Id}),\cdot)$ of (u_t,x_t,A_t) conditioned on $(u_0,x_0,A_0) = (u,x,\mathrm{Id})$ admits an everywhere-positive smooth density

$$\rho = \rho_{(u,x)} : \hat{\mathbf{H}} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R}) \to (0,\infty)$$

for all initial $(u,x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$. It follows that for any pair $(u,x), (u',x') \in \hat{\mathbf{H}} \times \mathbb{T}^d$ and any $t > 0$, the probabilistic law of A_t conditioned on $(u_0,x_0) = (u,x)$, $(u_t,x_t) = (u',x')$ admits a smooth, everywhere-positive density $\hat\rho = \hat\rho_{(u,x),(u',x')}$, given for $M \in \mathrm{SL}_d(\mathbb{R})$ by

$$\hat\rho(M) = \frac{\rho(u',x',M)}{\int_{\mathrm{SL}_d(\mathbb{R})} \rho(u',x',M')\,\mathrm{dLeb}_{\mathrm{SL}_d(\mathbb{R})}(M')}.$$

We conclude that (2.1) is impossible, hence $\lambda^+ > 0$, when Hörmander's condition for the matrix process (u_t,x_t,A_t) is satisfied. See Proposition 2.10 in Section 2.7 below for a precise statement of Hörmander's condition, and see condition (C) in Section 3.3.2 for a more detailed version of this argument.

We note that the technique of using Hörmander's condition for the matrix process (u_t,x_t,A_t) to rule out Furstenberg's criterion is well known; see, e.g., [12, 28].

2.4. *Furstenberg's criterion: Infinite-dimensional models*

For the infinite-dimensional models, Systems 3–4, we are not aware of any means by which one can prove a positive density for the conditional law of $A_t = D_x\phi^t_{\omega,u}$ as was possible for the finite-dimensional models.

Instead, we are able to prove a certain "approximate controllability" statement, which we describe below. To articulate this, we define the *projective process* (u_t, x_t, v_t) on $\mathbf{H} \times \mathbb{T}^d \times P^{d-1}$, where (v_t) is defined for initial v_0 by setting v_t to be the projective representative of $D_{x_0}\phi^t_{\omega,u_0} v_0$. Equivalently, (v_t) is generated by (u_t), (1.1) and

$$\partial_t v_t = \Pi_{v_t} Du(x_t) v_t. \tag{2.3}$$

Here, Π_{v_t} denotes the projection onto the orthogonal complement of (a unit vector representative of) v_t.

Proposition 2.4. *Consider the Markov processes (u_t, x_t, v_t) and (u_t, x_t, A_t) generated by either of System 3 or 4, together with (1.1), (2.2), and (2.3). Then, for any $x, x' \in \mathbb{T}^d$ and $t > 0$, we have the following:*

(a) *For any $\varepsilon, M > 0$, we have that*

$$\mathbf{P}((u_t, x_t) \in B_\varepsilon(0) \times B_\varepsilon(x'),\ |A_t| > M \mid u_0 = 0,\ x_0 = x,\ A_0 = \mathrm{Id}) > 0.$$

(b) *For any $\varepsilon > 0$, $v \in P^{d-1}$ and open $V \subset P^{d-1}$, we have*

$$\mathbf{P}((u_t, x_t) \in B_\varepsilon(0) \times B_\varepsilon(x'),\ v_t \in V \mid u_0 = 0,\ x_0 = x,\ v_0 = v) > 0.$$

Condition (a) says, roughly, that gradient norms can be made arbitrarily large while "approximately conditioning" on the time 0 and time t values of the Lagrangian process, while condition (b) says that we can rotate vectors arbitrarily in projective space. We see that this is weaker than obtaining information on the conditional law, but is clearly closely related. Our proof of Proposition 2.4 for Systems 3 and 4 is very physically intuitive; see Section 2.7 for more discussion.

Furstenberg's criterion as in Proposition 2.3 cannot be applied directly to the "softer" nondegeneracy condition in Proposition 2.4. Possible issues include (1) that the family of measures $\{\nu_{u,x}\}_{(u,x)\in\mathbf{H}\times\mathbb{T}^d}$ in Proposition 2.3 might, a priori, be discontinuous in space, and (2) that the individual measures $\nu_{u,x}$ could be quite pathological, e.g., singular continuous w.r.t. Lebesgue on P^{d-1}. To address this, we obtain the following classification of all possible demeanors of the measure family $\nu_{u,x}$.

Proposition 2.5. *Assume $\mu \times \mathrm{Leb}$ is an ergodic stationary measure for the Lagrangian process (u_t, x_t), and moreover, assume that the Lagrangian process (u_t, x_t) satisfies the strong Feller property (Definition 4.1). If $\lambda^+ = 0$, then one of the following alternatives holds:*

(a) *There is a continuously-varying family $\{\langle\cdot,\cdot\rangle_{u,x}\}_{(u,x)\in\mathbf{H}\times\mathbb{T}^d}$ of inner products on $\mathbb{R}^d$ such that*

$$\langle D_x\phi^t_{\omega,u} v, D_x\phi^t_{\omega,u} w\rangle_{\Theta^t_\omega(u,x)} = \langle v, w\rangle_{u,x} \quad \textit{with probability } 1$$

for all $v, w \in \mathbb{R}^d$, $t > 0$ and $(u, x) \in \mathbf{H} \times \mathbb{T}^d$.

(b) *There are $p \geq 1$ families $\{E^i_{(u,x)}\}_{(u,x)\in\mathbf{H}\times\mathbb{T}^d}$, $1 \leq i \leq p$, of proper linear subspaces of $\mathbb{R}^d$ such that*

   (i) *$(u, x) \mapsto E^i_{(u,x)}$ is* locally continuous up to relabeling *(see Theorem 4.7* (b) *for details),*

(ii) *for all* $(u,x) \in \mathbf{H} \times \mathbb{T}^d$ *and* $1 \le i \le p$,

$$D_x\phi^t_{\omega,u}(E^i_{u,x}) = E^{\pi(i)}_{\Theta^t_\omega(u,x)} \quad \textit{with probability } 1.$$

Here, $\pi = \pi_{\omega,u,x}$ *is a permutation of* $\{1,\dots,p\}$.

Note that the Strong Feller property of the Lagrangian process is explicitly required; see Remark 2.6 below for more discussion. We discuss proving the strong Feller property in Section 2.7 below. Roughly speaking, Proposition 2.5 follows from the strong Feller property as well as certain rigid geometric properties of $\mathrm{SL}_d(\mathbb{R})$ (Lemma 4.6) imposed by the condition of leaving a projective measure invariant (in the sense of Furstenberg's criterion as in Proposition 2.3).

Proposition 2.5 is the analogue of Theorem 6.8 in Baxendale's paper [12], a similar classification-type theorem for the derivative cocycle of an SDE on a finite-dimensional manifold. The analogue we obtain (stated as Theorem 4.7 and proved in Section 4.2) is considerably more general and applies to linear cocycles over continuous-time RDS on possibly infinite-dimensional Polish spaces. Our more general setting entails numerous complications not addressed in [12]; see Remark 4.15 for a more thorough discussion of these.

Alternatives (a)–(b) in Proposition 2.5 can now be ruled out by straightforward continuity arguments and approximate controllability as in Proposition 2.4; see Section 4.3 for more details. Once this has been carried out, the proof of Theorem 1.6 for Systems 3 and 4 is complete.

Remark 2.6. As far as the authors are aware, the strong Feller property of the Lagrangian process (u_t, x_t) is required for Proposition 2.5. Specifically, the strong Feller property is used to verify that the "configurations" appearing in alternatives (a), (b) of Proposition 2.5 are continuously-varying in an appropriate sense. We emphasize that this continuity is critical to the argument for ruling out (a), (b) using the approximate controllability condition in Proposition 2.4.

In particular, this is precisely the step we are not able to execute for 2D Navier–Stokes with "truly hypoelliptic" forcing (that is, forcing only a handful of low modes as in Assumption 1 and forgoing forcing all sufficiently high modes as in Assumption 2). In this regime, the strong Feller property is likely to be false for Systems 3–4 (see [64]).

2.5. Expansion in all directions: Proof of Corollary 1.7

For both the finite and infinite-dimensional systems considered in this paper, Corollary 1.7 does not follow immediately from Theorem 1.6. Indeed, a priori it is possible that given $(u,x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$, there are some $v \in \mathbb{R}^d$ for which $\limsup_{t\to\infty} \frac{1}{t} \log |D_x\phi^t_{\omega,u} v| < \lambda^+$ holds with probability 1.

We can rule this out using the ergodic theory of the projective process (u_t, x_t, v_t) as in equation (2.3) above. There is a well-known correspondence between the stationary probability measures ν on $\hat{\mathbf{H}} \times \mathbb{T}^d \times P^{d-1}$ and the asymptotic exponential growth rates $\lim_{t\to\infty} \frac{1}{t} \log |D_x\phi^t_{\omega,u} v|$ realized "with probability 1" as v varies in $\mathbb{R}^d \setminus \{0\}$. The correspondence is given by the so-called Random Multiplicative Ergodic Theorem (see [73, Theorem III.1.2]). We will not state the full result here, except to note the following relevant consequence.

Proposition 2.7. *Assume that there is a* unique *stationary measure* ν *for the projective process* (u_t, x_t, v_t)*. Then, for* $(\mu \times \mathrm{Leb})$*-almost every* $(u, x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$ *and* every $v \in \mathbb{R}^d \setminus \{0\}$*, we have that*

$$\lim_{t\to\infty} \frac{1}{t} \log |D_x \phi^t_{\omega,u} v| = \lambda^+ \quad \textit{with probability } 1.$$

Proposition 2.7 is formulated in a more general way as Proposition 3.16 in Section 3.2.3, to which we refer the reader for more details. The expansion estimate appearing in Corollary 1.7 now follows from a straightforward argument.

Added to our growing list of ingredients is uniqueness of the stationary measure ν for the projective process, to which we refer the reader to Section 2.7 for more information.

2.6. Gradient growth: Proof of Theorem 1.14

Given an initial $u_0 = u \in \hat{\mathbf{H}}$, an initial scalar $f_0 = f \in H^1$, $\int f \,\mathrm{d}x = 0$, and a noise parameter $\omega \in \Omega$, the corresponding solution (f_t) for the passive advection equation (1.5) is given by

$$f_t(x) = f \circ (\phi^t_{\omega,u})^{-1}(x).$$

By incompressibility, we have (recall $-\top$ is standard shorthand for the inverse transpose)

$$\|\nabla f_t\|_{L^1} = \int |\nabla f_t(x)| \,\mathrm{d}x = \int |(D_x \phi^t_{\omega,u})^{-\top} \nabla f_0(x)| \,\mathrm{d}x.$$

The object $(D_x \phi^t_{\omega,u})^{-\top}$ defines a cocycle over the RDS Θ^t_ω on $\hat{\mathbf{H}} \times \mathbb{T}^d$ in the same manner as $D_x \phi^t_{\omega,u}$. To complete the proof of Theorem 1.14, it suffices to obtain the following analogue of Corollary 1.7 for this new cocycle.

Proposition 2.8. *There is a constant* $\lambda > 0$ *with the following property. For any* $\eta > 0$*,* $\eta \ll \lambda$*,* $\mu \times \mathrm{Leb}$*-almost every* $(u, x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$*, and every unit vector* $v \in \mathbb{R}^d$*, there is a (random) constant* $\hat{\delta} = \hat{\delta}_\omega(u, x, v, \eta)$ *(i.e., depending on the noise parameter* $\omega \in \Omega$*) such that with probability* 1*,* $\hat{\delta} > 0$ *and*

$$|(D_x \phi^t_{\omega,u})^{-\top} v| \geq \hat{\delta} e^{t(\lambda-\eta)}.$$

When $d = 2$*, we have* $\lambda = \lambda^+$*.*

Setting $v = \nabla f_0(x)/|\nabla f_0(x)|$ and integrating over $\{x \in \mathbb{T}^d : \nabla f_0 \neq 0\}$, we obtain Theorem 1.14 for $p = 1$. The estimate for the remaining L^p spaces follows from

$$\|\nabla f_t\|_{L^1} \lesssim \|\nabla f_t\|_{L^p}$$

for all $p \in [1, \infty]$.

To prove Proposition 2.8, we will prove Theorem 1.6 and Corollary 1.7 with the $(-\top)$-cocycle $(D_x \phi^t)^{-\top}$ replacing the usual $D_x \phi^t$. Let us summarize briefly how this will be done. For Theorem 1.6 we have the following.

Proposition 2.9. *The following statements hold:*

(a) *For* $\mu \times \mathrm{Leb}$*-almost every* $(u, x) \in \hat{\mathbf{H}} \times \mathbb{T}^d$*, the growth rate*

$$\check{\lambda}^+(\omega, u, x) = \lim_{t\to\infty} \frac{1}{t} \log |(D_x\phi^t_{\omega,u})^{-\top}|$$

exists with probability 1*. Moreover, if* $\mu \times \mathrm{Leb}$ *is the unique (hence ergodic) stationary measure for the* (u_t, x_t) *process, then* $\check{\lambda}^+$ *is independent of* ω, u, x*.*

(b) *Let* λ^+ *be as in Proposition* 2.1*. Then* $\lambda^+ > 0$ *iff* $\check{\lambda}^+ > 0$*. Indeed,* $\lambda^+ = \check{\lambda}^+$ *if* $d = 2$*.*

Item (a) is merely a repetition of Proposition 2.1 for the $(-\top)$-cocycle and is a consequence of the Multiplicative Ergodic Theorem; see Theorem 3.13 for details. As in Theorem 1.6, passing between “almost every” and “every” is done using the Strong Feller property; see Remark 2.2. Item (b) is a consequence of a general relationship between the Lyapunov exponents of $D_x\phi^t$ and $(D_x\phi^t)^{-\top}$; see Section 3.2.5 for details. In particular, note that the relation $\lambda^+ = \check{\lambda}^+$ is exclusive to $d = 2$; the authors are unaware of any reason to expect it to hold in dimension $d = 3$.

Having shown (Theorem 1.6) that $\lambda^+ > 0$, we conclude $\check{\lambda}^+ > 0$. To prove the analogue of Corollary 1.7 for the $(-\top)$-cocycle will require, as in Proposition 2.7, for us to study the so-called $(-\top)$-*projective process* $(u_t, x_t, \check{v}_t)$ on $\hat{\mathbf{H}} \times \mathbb{T}^d \times P^{d-1}$, defined for initial $\check{v}_0 \in P^{d-1}$ by setting $\check{v}_t$ to be the projective representative of $(D_x\phi^t_{\omega,u})^{-\top}\check{v}_0$. Equivalently, the $(\check{v}_t)$ process is governed by (u_t), (1.1), and

$$\partial_t \check{v}_t = -\Pi_{\check{v}_t}(Du_t(x_t))^\top \check{v}_t.$$

Repeating Proposition 2.7 verbatim with $D_x\phi^t$ replaced by $(D_x\phi^t)^{-\top}$, we see that Proposition 2.8 follows immediately from the existence of a unique (hence ergodic) stationary measure $\check{\nu}$ for the $(-\top)$-projective process $(u_t, x_t, \check{v}_t)$.

2.7. Hypoellipticity

The previous discussion of the proofs of Theorems 1.6 and 1.14 requires a number of ingredients pertaining to the properties of the various stochastic processes (Lagrangian, projective, $(-\top)$-projective, and matrix) mentioned so far. Specifically, we need:

(a) Uniqueness of the stationary measure for the
   - (i) Lagrangian,
   - (ii) projective,
   - (iii) $(-\top)$-projective processes.

(b) For the infinite-dimensional Systems 3–4, the Strong Feller property (Definition 4.1) for the Lagrangian process (u_t, x_t).

(c) For the matrix process (u_t, x_t, A_t) and projective process (u_t, x_t, v_t), either:
   - (i) Hörmander’s condition for the SDE defining (u_t, x_t, A_t) for the finite-dimensional Systems 1–2, or
   - (ii) approximate controllability condition in Proposition 2.4 for the infinite-dimensional Systems 3–4.

Let us recall briefly where each of these is used. First, ingredient (a) (i) was used to deduce the almost-sure constancy of the exponential growth rates $\lambda^+, \check{\lambda}^+$ as in Proposition 2.1 and Proposition 2.9 (a), respectively. Meanwhile, (a) (ii) was used to deduce almost sure growth for $(D_x\phi^t)v$ in Corollary 1.7 (see Proposition 2.7); analogously, (a) (iii) was used to deduce growth of the $(D_x\phi^t)^{-\top}v$ in Proposition 2.8. On the other hand, (b) is used to justify the refinement of Furstenberg's criterion (Proposition 2.5) used for Systems 3–4. For the finite-dimensional Systems 1, 2, ingredient (c) (i) was used to rule out Furstenberg's criterion (Proposition 2.3); see the discussion in Section 2.3. Lastly, ingredient (c) (ii) was used to rule out the refinement of Furstenberg's criterion in Proposition 2.5 for Systems 3–4.

All of items (a)–(c) require us to understand how the noise in the low modes of u_t spread to the degrees of freedom associated with the Lagrangian flow. Note the additional degrees of freedom $(x_t, v_t, \check{v}_t, A_t)$ solve a series of random ODEs (collected below in equation (1.9)). Since these unknowns are not directly forced by any noise, the corresponding SDEs are degenerate and we need to depend on hypoellipticity to show (a)–(c).

2.7.1. Finite dimensions: Systems 1 and 2. Let us discuss how the ingredients for the finite-dimensional Systems 1 and 2 are obtained. For these models, all relevant stochastic processes as above are given by an SDE on a finite-dimensional manifold. Provided that one can show the algebra formed by taking successive Lie brackets of vector fields associated to the drift and the noise directions $e_k\gamma_k^i$ span the tangent space at every point, a condition known as *Hörmander's condition* (see Definition 5.1 for a precise definition and Remark 2.11 for a conceptual discussion), we may apply Hörmander's Theorem (see [67,68] and the discussions in [37,63]) to deduce that the Markov transition kernels for the Lagrangian, projective, $(-\top)$-projective and matrix processes have a smooth positive density. Assumption 1 ultimately ensures that Hörmander's condition is satisfied. Specifically we prove the following proposition in Section 5:

Proposition 2.10. *Assume (u_t) is governed by either of the finite-dimensional Systems 1 or 2. For each of*

(i) *the Lagrangian process (u_t, x_t),*

(ii) *the projective process (u_t, x_t, v_t),*

(iii) *the matrix process (u_t, x_t, A_t), and*

(iv) *the $(-\top)$-projective process $(u_t, x_t, \check{v}_t)$,*

the SDE governing the relevant process satisfies Hörmander's condition.

By standard arguments (see, e.g., [38]), uniqueness of the stationary measures then follows for the Lagrangian, projective and $(-\top)$-projective processes [37], thereby fulfilling ingredients (a) (i)–(iii) above as well as (b). Likewise (c) (i) is immediately satisfied for the matrix process.

Remark 2.11. Physically, one may view Hörmander's condition as an infinitesimal controllability statement. When it is satisfied for the (u_t, x_t, v_t, A_t) process, one can infinitesimally move each component of this process independently of the others using special choices of noise paths. Hence, all possible infinitesimal deformations of the flow map are realized with nonzero probability.

2.7.2. Infinite dimensions: Systems 3–4. In infinite dimensions, Hörmander's condition is not applicable and so we must work harder to verify ingredients (a) (i)–(iii). There have been a number of works proving uniqueness of the stationary measure for the Navier–Stokes equations under degenerate and nondegenerate noise (for instance [50, 52, 64, 65]). A standard approach is to apply the Doob–Khasminskii Theorem [41, 72], the fact that distinct ergodic stationary measures for strong Feller processes (Definition 4.1) have disjoint supports, and then to check that there exists a point which belongs to the support of every invariant measure (a.k.a. weak irreducibility). Following this strategy, in Section 6 we prove the strong Feller property for the Lagrangian, projective and $(-\top)$-projective processes.

Proposition 2.12 (Strong Feller). *For Systems 3–4, the Markov semigroups associated with the Lagrangian process (u_t, x_t) and the projective processes (u_t, x_t, v_t), $(u_t, x_t, \check{v}_t)$ are all strong Feller in $\mathbf{H} \times \mathbb{T}^d \times P^{d-1}$.*

Remark 2.13. This proposition is where we need the lower bound $\sigma > \alpha - 2(d-1)$ as in (1.3).

Remark 2.14. If the noise is suitably nondegenerate, then the strong Feller property for the Navier–Stokes equations can be proved by the Bismut–Elworthy–Li formula (see for instance [52] and [29]). However, if the noise is too degenerate, then it is not known whether the strong Feller property even holds. Indeed, to get around this difficulty, Hairer and Mattingly [64, 65] introduced a weaker notion, *the asymptotic strong Feller property*, which when combined with weak irreducibility, gives a generalization of the Doob–Khasminskii Theorem, still giving uniqueness of the stationary measure. While the asymptotic strong Feller property is clearly good enough obtain ingredients (a) (i)–(iii), it *does not* appear to be enough to prove the refinement of Furstenberg's criterion (Proposition 2.5), which requires that (u_t, x_t) be strong Feller (ingredient (b)). It is precisely this strong Feller requirement for Furstenberg's criterion that dictates our nondegeneracy Assumption 2.

To conclude uniqueness of the stationary measures as in (a) (i)–(iii), it suffices to prove the following weak irreducibility properties, proved in Section 7 below.

Proposition 2.15. *For Systems 3–4, we have the following:*

(1) *The support of any stationary measure for the Lagrangian process (u_t, x_t) on $\mathbf{H} \times \mathbb{T}^d$ must contain the set $\{0\} \times \mathbb{T}^d$.*
(2) *The support of any stationary measure for the projective processes (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ on $\mathbf{H} \times \mathbb{T}^d \times P^{d-1}$ must contain $\{0\} \times \mathbb{T}^d \times P^{d-1}$.*

Uniqueness of the stationary measures now follow.

Corollary 2.16. *The processes (u_t), (u_t, x_t), (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ all have unique stationary measures.*

Additionally, it remains to address ingredient (c) (ii), the approximate controllability condition in Proposition 2.4. Once Propositions 2.15 and 2.4 are completed, the proof of Theorem 1.6 for Systems 3-4 is complete.

2.7.3. Strong Feller. Our proof of Proposition 2.12 is inspired by methods of Eckmann and Hairer [45]. In [45], the authors prove strong Feller for the complex Ginzburg–Landau equations with forcing that satisfies Assumption 2, using a cut-off technique and a high-low frequency splitting. This cut-off approach has since been extended to Markov selections of the 3D Navier–Stokes equations in [100]. Similar results to [100] were proved in [5] using the infinite-dimensional Kolmogorov equation. Our proof of strong Feller is closer to [45] and [100], but differs in our choice of the cut-off process, the use of non-adapted controls, estimates on Skorokhod integrals, and an interpolation inequality introduced in [65] used to circumvent some technicalities with applying Norris's Lemma in $L^2([0,1])$.

Similarly to [45,52,100], it does not seem possible to obtain an estimate on the derivative of the Markov semigroup of the projective process (u_t, x_t, v_t). The strategy is to show that such an estimate is available for a "cut-off" or "regularized" process. In our setting, we will find it convenient to augment the projective process (u_t, x_t, v_t) by a Brownian motion (z_t) on $\mathbb{R}^{2d}$ (likewise for the $(-\top)$ projective process). The augmented process $w_t = (u_t, x_t, v_t, z_t)$ solves an abstract evolution equation

$$\partial_t w_t = F(w_t) - A w_t + Q\dot{W}_t$$

on $\mathbf{H} \times \mathcal{M}$, where $\mathcal{M}$ is a smooth finite-dimensional manifold. Let $\hat{P}_t$ be the Markov semigroup associated to w_t, then our goal is to find a regularized process w_t^ρ such that $\mathbf{P}((w_t)_{t\in[0,T]} \neq (w_t^\rho)_{t\in[0,T]})$ is vanishingly small as $\rho \to \infty$ but for which one can obtain a derivative estimate on the associated semigroup $\hat{P}_t^\rho$.

Remark 2.17. It is important to note that our choice of cut-off process w_t^ρ is different from that used in [45] and [100] and uses the augmentation by z_t to introduce new sources of noise while avoiding technical difficulties with multiplicative white noise (see Section 6 for more details on the cut-off process).

Our main effort is then to prove that the cut-off semigroup $\hat{P}_t^\rho$ satisfies the following gradient estimate (Proposition 6.1):

$$|D\hat{P}_t^\rho \phi(w)h| \lesssim_\rho t^{-a_*}(1 + \|w\|_{\mathbf{H}}^{b_*})\|\phi\|_{L^\infty}\|h\|_{\mathbf{H}\times T_v\mathcal{M}} \tag{2.4}$$

for all bounded measurable ϕ on $\mathbf{H} \times \mathcal{M}$, $h \in \mathbf{H} \times T_v\mathcal{M}$, and sufficiently small t, and a_* and b_* are certain positive constants. We show in the proof of Proposition 2.12 in Section 6 this estimate on $\hat{P}_t^\rho$ implies that $\hat{P}_t$ is strong Feller, albeit without an estimate on the derivative.

The fundamental tool for proving (2.4) is Malliavin calculus. This involves taking derivatives of the solution with respect to the noise. Well-posedness of the cutoff process implies that for each ρ and initial data $w \in \mathbf{H} \times \mathcal{M}$, the solution w_t^ρ at time $t > 0$ is a continuous function of the noise path $W \in C(\mathbb{R}_+, \mathbf{H}^{-\varepsilon})$, where $\varepsilon > 0$ is such that $Q(\mathbf{H}^{-\varepsilon}) \subseteq \mathbf{H}$. Specifically, we have the Itô map $W|_{[0,t]} \mapsto w_t^\rho[W|_{[0,t]}]$ is a continuous mapping from $C([0,t],\mathbf{H}^{-\varepsilon})$ to $\mathbf{H}\times\mathcal{M}$ for each $t>0$. In fact, it is straightforward to show that $W|_{[0,t]}\mapsto w_t^\rho[W|_{[0,t]}]$ is actually Fréchet differentiable over the Banach space $C([0,t];\mathbf{H}^{-\varepsilon})$ (see for instance [65, Proposition 4.1]). Indeed, for any process $g=(g_t)$ (not necessarily adapted to $\mathcal{F}_t$) that belongs almost surely to $L^2(\mathbb{R}_+,\mathbf{L}^2)$, the Malliavin derivative $\mathcal{D}_g w_t^\rho$ of w_t^ρ in the direction of g, defined by

$$\mathcal{D}_g w_t^\rho=\frac{\mathrm{d}}{\mathrm{d}h}w_t^\rho[W+hG]\big|_{h=0},\quad G=\int_0^\cdot g_s\,\mathrm{d}s, \tag{2.5}$$

exists for each $t>0$. We will often refer to g as a *control*. A key feature of the Malliavin derivative is the celebrated Malliavin integration by parts formula, which states that for each $\phi\in C_b^1(\mathbf{H}\times\mathcal{M})$ and a suitably regular g (see Proposition 6.4 for the precise conditions) one has the identity

$$\mathbf{E}(D\phi(w_t^\rho)\mathcal{D}_g w_t^\rho)=\mathbf{E}\mathcal{D}_g\phi(w_t^\rho)=\mathbf{E}\left(\phi(w_t^\rho)\int_0^t\langle g_s,\delta W_s\rangle_{\mathbf{L}^2}\right), \tag{2.6}$$

where the stochastic integral $\int_0^t\langle g_s,\delta W_s\rangle_{\mathbf{L}^2}$ above denotes the Skorokhod integral (see, e.g., [95, Definition 1.3.1] or [37, Section 11.3]). If g is adapted to the filtration $\mathcal{F}_t$, then the Skorokhod integral coincides with the usual Itô integral. Formula (2.6) can be used to obtain smoothing estimates on the semigroup $\hat{P}_t^\rho$. Indeed, if for every $h\in\mathbf{H}\times T_v\mathcal{M}$, one could find a "nice enough" control g such that $\mathcal{D}_g w_t^\rho=Dw_t^\rho h$, where $Dw_t^\rho h$ denotes the direction derivative of w_t^ρ in the direction h with respect to the initial data, then an estimate on $D\hat{P}_t^\rho$ follows from (2.6) as long as one can bound the Skorokhod integral term (see (2.8) below for more details). However, in our setting we are unable to find such a control g due to subtleties in infinite dimensions. Instead we opt to find a control g such that for each fixed $0<T<1$, we have

$$\mathcal{D}_g w_T^\rho=Dw_T^\rho h+r_T, \tag{2.7}$$

where r_T is a remainder which will be small when T is small, and consequently the Skorokhod integral $E|\int_0^t\langle g_s,\delta W_s\rangle|$ will be singular as T approaches 0 (see Lemma 6.5 for the exact estimates). The (non-adapted) control g is chosen with an elaboration of the high-low splitting used in [45]. At high frequencies it is chosen such that the contribution to the Malliavin integration by parts formula reduces to the Bismut–Elworthy–Li formula, while at lower frequencies, the control is set by inverting a finite-dimensional approximation of the Malliavin matrix (the partial Malliavin matrix) while attempting to minimize the amount by which the low frequency control perturbs the higher frequencies. The invertibility of the partial Malliavin matrix can be deduced from the fact that the projective process associated to finite-dimensional approximations of the Navier–Stokes equations satisfy Hörmander's condition (shown in Section 5).

The fact that we can have a remainder in (2.7) and can still prove a smoothing estimate depends heavily on the precise dependence of the bounds on r_T and the Skorokhod integral. The key idea, inspired by [45] and [29] involves using the semigroup property and the integration by parts formula (2.6) to write

$$\begin{aligned}D\hat{P}_{2T}^\rho\phi(w)h&=\mathbf{E}\big(D\hat{P}_T^\rho\phi(w_T^\rho)Dw_T^\rho h\big)\\&=\mathbf{E}\left(\hat{P}_T^\rho\phi(w_T^\rho)\int_0^T\langle g_t,\delta W(t)\rangle_{\mathbf{L}^2}\right)-\mathbf{E}\big(D\hat{P}_T^\rho\phi(w_T^\rho)r_T\big).\end{aligned} \tag{2.8}$$

Using the estimates on r_T and the Skorokhod integral, one can close estimates on $D\hat{P}_t^\rho \phi$ for sufficiently short times. The details of this argument can be found in the proof of Proposition 6.1.

2.7.4. Weak irreducibility and approximate control. Let us first discuss Proposition 2.15. For simplicity, here we only discuss the 2D case, System 3. Weak irreducibility for (u_t) is a consequence of the energy/enstrophy dissipation (see Section 7 and, e.g., [44]), which shows that 0 is in the support of all stationary measures for the (u_t) processes. Using a stability argument and the positivity of the Wiener measures, the main content of the irreducibility in Proposition 2.15 is the study of the control problem

$$\partial_t u_t = -B(u_t, u_t) - Au_t + Qg(t), \tag{2.9}$$

where $g \in C^\infty(\mathbb{R}_+, \mathbf{L}^2)$ is a smooth control,

$$A = -\Delta,$$
$$B(u,u) = (I + \nabla(-\Delta)^{-1}\nabla\cdot)(u \cdot \nabla u).$$

Here, $x_t, v_t, \check{v}_t$ and A_t are implicitly controlled through (u_t). First, we prove that for all $(x,v), (x',v') \in \mathbb{T}^d \times P^{d-1}$, there exist smooth controls g such that

$$(u_0, x_0, v_0) = (0, x, v), \quad (u_1, x_1, v_1) = (0, x', v')$$

(and analogously for the $(u_t, x_t, \check{v}_t)$ process). We note that it suffices to control near $u_t \approx 0$ precisely because 0 is in the support of the stationary measure μ. To solve this control problem we use that the following flows are *exact* solutions (for arbitrary a, b) of the steady Euler equation $B(u,u) = 0$ as well as eigenfunctions of A:

$$u(y_1, y_2) = \begin{pmatrix} \cos(y_2 - b) \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \cos(y_1 - a) \end{pmatrix}, \begin{pmatrix} \sin(y_2 - b) \\ -\sin(y_1 - a) \end{pmatrix}. \tag{2.10}$$

The first two are shear flows whereas the last flow is a cellular flow with separatrices aligned along the diagonals. The first two flows are used to move the particle x_t whereas the latter flow is used to move v_t without moving the particle. Once these flows can be formed, it is not difficult to verify the necessary controllability of system (2.9); see Lemma 7.1 for details. Note that Assumption 1 is slightly stronger than what is necessary to form the flows (2.10), which is why, for example, Remark 1.13 holds (see Lemma 7.1 and Remark 7.5). Similarly, for the case of Systems 3 and 4, one can prove Theorem 1.6 (and all of the other main results) using only Assumption 2; see Remark 7.6.

The nondegeneracy of the (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ processes needed to prove Proposition 2.15 and condition (b) in Proposition 2.4 then follow from the controllability and suitable stability estimates (see Section 7 for details). In order to satisfy condition (a) in Proposition 2.4, we also need to demonstrate arbitrarily large growth of A_t in the (u_t, x_t, A_t) process (under similar constraints as for the projective control statements). This is done by applying the cellular flow as above, but shifted so that the hyperbolic fixed point causes exponential growth of A_t without moving the particle x_t; see Proposition 7.4 for details.

2.8. *Proof of Yaglom's Law* (1.8) *as in Theorem* 1.16 (ii)

Next, we summarize the proof of Theorem 1.16 (see Section 8 for details). First, we prove the limit (1.7). This result follows from a straightforward adaptation of the compactness method of [17], originally applied to passive scalars with deterministic, constant-in-time velocity fields. The first step is to renormalize $f_t^\kappa = \sqrt{\kappa} g_t$ to obtain

$$\partial_t f_t^\kappa + u_t \cdot \nabla f_t^\kappa = \kappa \Delta f_t^\kappa + \sqrt{\kappa} \tilde{Q} \dot{\tilde{W}}_t. \tag{2.11}$$

The balance (1.6) then becomes, for statistically stationary solutions,

$$\mathbf{E} \|\nabla f^\kappa\|_{L^2}^2 = \bar{\varepsilon}. \tag{2.12}$$

Denote by $\{\bar{\mu}^\kappa\}_{\kappa>0}$ a sequence of stationary measures to (2.11) supported on $\hat{\mathbf{H}} \times H^1$. Equality (2.12) is sufficient to obtain tightness of $\{\bar{\mu}^\kappa\}_{\kappa>0}$ to pass to the limit and deduce the existence of a stationary measure $\bar{\mu}^0$ of the problem (2.11) with $\kappa = 0$ supported on $\hat{\mathbf{H}} \times H^1$. Theorem 1.14 is then applied to prove by contradiction that necessarily $\bar{\mu}^0 = \mu \times \delta_0$, where δ_0 denotes the Dirac delta centered at zero and μ is the stationary measure for (u_t). The limit (1.7) then follows from additional moment bounds in L^2; see Section 8 for more details.

In order to prove (1.8) we in turn adapt the method of [17] . One of the basic identities used in [17] is a version of the classical Kármán–Horvath–Monin relation [40, 55, 90] which is a refinement of the L^2 energy balance. Here, we apply a similar identity, now a refinement of the L^2 balance for g_t (see Proposition 8.4 below). This identity implies a differential equation (in weak form) for the quantity (see (8.5)),

$$\bar{\mathfrak{D}}(\ell) = \mathbf{E} \fint_{\mathbb{T}^d} \fint_{\mathbb{S}^{d-1}} \delta_{\ell n} u \cdot n |\delta_{\ell n} g|^2 \,\mathrm{d}S(n)\,\mathrm{d}x.$$

Solving the ODE (8.5) in terms of the source and dissipation, we apply (1.7) to show that the effect of the diffusivity on the balance vanishes over an appropriate range of scales $[\ell_D(\kappa), \ell_I]$ satisfying $\lim_{\kappa\to 0} \ell_D(\kappa) = 0$. This then yields (1.8).

3. Random dynamical systems preliminaries

In this section we will present necessary background from random dynamical systems theory. This section is mostly an exposition of material drawn from various sources in the dynamics literature. General references include the books of Arnold [8], Kifer [73], and Kuksin and Shirikyan [86].

The plan for Section 3 is as follows. We begin in Section 3.1 with some essential ergodic-theoretical background: the definition and standard axioms we use for random dynamical systems (RDS) and some elementary results. Section 3.2 introduces the notion of *linear cocycle* over a given RDS and formulates the Multiplicative Ergodic Theorem (MET), allowing us to define the Lyapunov exponent λ^+ appearing in Theorem 1.6. In Section 3.3 we turn our attention to the problem of how to prove $\lambda^+ > 0$ using Furstenberg's criterion (Theorem 3.18).

3.1. Elements of ergodic theory of random dynamical systems

3.1.1. Basic setup for random dynamics. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and let (θ^t) be a measure-preserving semiflow on Ω, i.e., $\theta : [0, \infty) \times \Omega \to \Omega, (t, \omega) \mapsto \theta^t\omega$ is a measurable mapping satisfying

(i) $\theta^0\omega \equiv \omega$ for all $\omega \in \Omega$,

(ii) $\theta^t \circ \theta^s = \theta^{t+s}$ for all $s, t \geq 0$, and

(iii) $\mathbf{P} \circ (\theta^t)^{-1} = \mathbf{P}$ for all $t \geq 0$.

At times (which we will specify), it will be useful to assume that Ω has some topological structure. If so, we will assume additionally that Ω is a Borel subset of a Polish space, and $\mathcal{F}$ is the set of Borel subsets of Ω.

Let (Z, d) be a separable, complete metric space. A *random dynamical system* or RDS on Z is an assignment to each $\omega \in \Omega$ of a mapping $\mathcal{T}_\omega : [0, \infty) \times Z \to Z$ satisfying the following basic properties:

(i) (Measurability) The mapping $\mathcal{T} : [0, \infty) \times \Omega \times Z \to Z$, $(t, \omega, z) \mapsto \mathcal{T}^t_\omega z$, is measurable with respect to $\mathrm{Bor}([0, \infty)) \otimes \mathcal{F} \otimes \mathrm{Bor}(Z)$ and $\mathrm{Bor}(Z)$.

(ii) (Cocycle property) For all $\omega \in \Omega$, we have $\mathcal{T}^0_\omega = \mathrm{Id}_Z$ (the identity mapping on Z), and for $s, t \geq 0$, we have $\mathcal{T}^{s+t}_\omega = \mathcal{T}^t_{\theta^s\omega} \circ \mathcal{T}^s_\omega$.

(iii) (Continuity) For all elements $\omega \in \Omega$, the mapping $\mathcal{T}_\omega : [0, \infty) \times Z \to Z$ belongs to $C_{u,b}([0, \infty) \times Z, Z)$.

Here, for metric spaces V, W, the space $C_{u,b}(V, W) \subset C(V, W)$ is defined as follows:

Definition 3.1. We define[7] $C_{u,b}(V, W)$ to be the space of continuous maps $F : V \to W$ for which the following holds for each bounded $U \subset V$:

(a) The restriction $F|_U$ is uniformly continuous.

(b) the image $F(U)$ is a bounded subset of W.

We endow $C_{u,b}(V, W)$ with the topology of uniform convergence on bounded sets (abbreviated UCBS). It is a simple exercise to check that if $(F_n)_n$ is a sequence in $C_{u,b}(V, W)$ converging to some $F : V \to W$ in the UCBS mode, then $F \in C_{u,b}(V, W)$ holds. Moreover, it is a simple exercise to check in this setting that $C_{u,b}(V, W)$ is metrizable.

Note that automatically, condition (iii) implies that $\mathcal{T}^t_\omega \in C_{u,b}(Z, Z)$ for all $t \geq 0$, $\omega \in \Omega$. Indeed, by condition (iii), for any $\omega \in \Omega$, $T > 0$ and bounded $U \subset Z$, the family $\{\mathcal{T}^t_\omega|_U : U \to Z\}_{t\in[0,T]}$ is equicontinuous since $\mathcal{T}_\omega : [0, T] \times U \to Z$ is uniformly continuous.

Definition 3.2. We refer to $\mathcal{T}$ satisfying (i)–(iii) above as a *continuous RDS* on Z.

In addition to (i)–(iii) above, we will almost always assume that the RDS $\mathcal{T}$ satisfies the following *independent increments assumption*.

---

[7] We use the slightly nonstandard topology $C_{u,b}(Z, Z)$ to accommodate for the situation when Z is not locally compact. The regularity of $C_{u,b}$-topology is used in several places, especially in Section 4, and so will be assumed from this point on.

(H1) For all $s, t > 0$, we have that $\mathcal{T}^t_\omega$ is independent of $\mathcal{T}^s_{\theta^t\omega}$. That is, the σ-subalgebra $\sigma(\mathcal{T}^t_\cdot) \subset \mathcal{F}$ generated by the $C_{u,b}(Z, Z)$-valued random variable $\omega \mapsto \mathcal{T}^t_\omega$ is independent of the σ-subalgebra $\sigma(\mathcal{T}^s_{\theta^t\cdot})$ generated by $\omega \mapsto T^s_{\theta^t\omega}$.

Example 3.3. Let $n \geq 1$ and let $Y_0, Y_1, \ldots, Y_m$ be smooth, globally Lipschitz vector fields on $\mathbb{R}^n$. Let $W^1_t, \ldots, W^m_t$ be independent standard Brownian motions. Then the stochastic differential equation

$$\mathrm{d}X_t = Y_0(X_t)\,\mathrm{d}t + \sum_{i=1}^{m} Y_i(X_t)\mathrm{d}W^i_t$$

defines a random dynamical system on $Z = \mathbb{R}^n$, where $\Omega = C_0([0,\infty), \mathbb{R})^{\otimes m}$ is the k-fold product of Canonical Spaces equipped with the k-fold product Borel σ-algebra and Wiener measure $\mathbf{P}$, and $\theta^t : \Omega \to \Omega$ is the shift of increments by $t \geq 0$, written $\theta^t\omega(s) := \omega(t+s) - \omega(t)$. The resulting RDS satisfies the measurability and continuity conditions (i)–(iii). The independent increments condition (H1) follows from the independence of the Brownian increments $W^i_{s+t} - W^i_t$ and W^j_t for all $s, t > 0$ and each $1 \leq i, j \leq k$. See, e.g., [8, 87] for more details.

3.1.2. Markov chain formulation and stationary measures. For fixed $z \in Z$, consider the stochastic process $(z_t)_{t\geq 0}$ given by $z_t = \mathcal{T}^t_\omega z_0, z_0 := z$.

Lemma 3.4. *Let $\mathcal{T}$ be a continuous RDS as in Section 3.1.1 satisfying the independent increments condition (H1). Then, the process $(z_t)_{t\geq 0}$ as above is* Markovian.

For a proof of Lemma 3.4, see, e.g., Kuksin and Shirikyan [86], where the Markov property is proved under a somewhat weaker hypothesis than (H1).

For $t > 0, z \in Z$ and $K \in \mathrm{Bor}(Z)$, we define the *Markov kernel*

$$P_t(z, K) := \mathbf{P}(z_t \in K \mid z_0 = z).$$

The Markov kernel $P_t(z, K)$ has a natural action on any bounded measurable observable $h : Z \to \mathbb{R}$

$$P_t h(z) := \int h(z') P_t(z, \mathrm{d}z').$$

The Markov property of (z_t) implies the semigroup relation $P_{s+t} = P_t \circ P_s$. We refer to the operators $(P_t)_{t>0}$ as the *Markov semigroup* associated to (z_t).

The proof of the following proposition is straightforward and omitted for brevity.

Proposition 3.5. *Assume the setting of Lemma 3.4.*

(a) *The semigroup (P_t) has the* Feller property, *i.e., for any $t \geq 0$ and any $h : Z \to \mathbb{R}$ be continuous and bounded, we have that $P_t h$ is defined and is a continuous function $Z \to \mathbb{R}$.*

(b) *The semigroup (P_t) has the following C^0-continuity property: for any bounded $h \in C_{u,b}(Z, \mathbb{R})$, we have that*

(1) *$P_t h \in C_{u,b}(Z, \mathbb{R})$, with $\|P_t h\|_{L^\infty} \leq \|h\|_{L^\infty}$, for all $t > 0$,*

(2) *the mapping $t \mapsto P_t h, t \geq 0$ is continuous in the topology on $C_{u,b}(Z, \mathbb{R})$.*

We regard the (formal) dual $(P_t)^*$ of the operator P_t as acting on the space of finite signed Borel measures on Z. Given a finite signed Borel μ on Z, $(P_t)^*\mu$ is defined for Borel $A \subset Z$ by

$$(P_t)^*\mu(A) = \int P_t(z, A)\, \mathrm{d}\mu(z).$$

If μ is a Borel probability on Z for which $(P_t)^*\mu = \mu$ for all $t \geq 0$, we call μ *stationary*.

The following lemma is a consequence of a standard Krylov–Bogoliubov argument. Recall that a collection $\mathcal{S}$ of Borel probabilities on a Polish space X is called *tight* if for all $\varepsilon > 0$ there is some compact subset $K = K_\varepsilon \subset X$ such that for all $\nu \in \mathcal{S}$, we have $\nu(K_\varepsilon) \geq 1 - \varepsilon$.

Lemma 3.6. *Assume the setting of Lemma 3.4. Then the Markov semigroup (P_t) admits at least one stationary measure μ in either of the following circumstances:*

(a) *The space Z is compact.*

(b) *There exists a Borel probability μ_0 for which the sequence $\{(P_t)^*\mu_0\}_{t\geq 0}$ is tight.*

Note that (a) implies (b).

3.1.3. Skew product formulation and invariant measures. The material in Section 3.1.3 is mostly taken from [73, Chapter I].

The Markov chain formulation given above is useful in that it identifies "time-invariant" statistics on Z for the RDS, namely, its stationary measures. On the other hand, the Markov kernel loses some structure of the RDS, in the sense that the same Markov kernel can arise from qualitatively different RDS. See, e.g., [73, Example I.1.1] for an extreme example of this.

The following *skew product* formulation, unlike the Markov chain, encodes the entire RDS.

Definition 3.7. The *skew product* associated to the above random dynamics is the mapping $\tau : [0, \infty) \times \Omega \times Z \to \Omega \times Z$ given by

$$\tau(t, \omega, z) = \tau^t(\omega, z) = (\theta^t\omega, \mathcal{T}^t_\omega z).$$

We regard τ as a single "deterministic", measurable semiflow on the augmented space $\Omega \times Z$. In particular, this provides us a connection between "standard" ergodic theory, i.e., the theory of invariant measures for individual mappings of a measurable space, and our present setting of random dynamical systems. The following lemma makes this connection explicit.

Recall that a probability measure η on $\Omega \times Z$ is *invariant* for the semiflow τ if $\eta \circ (\tau^t)^{-1} = \eta$ for all $t \geq 0$.

Lemma 3.8 ([73, Lemma I.2.3]). *Assume $\mathcal{T}$ is a continuous RDS as in Section 3.1.1 satisfying (H1) and generating the Markov semigroup (P_t) as in Lemma 3.4. Let μ be a Borel probability measure on Z. Then the following are equivalent:*

(a) *The measure $\mathbf{P} \times \mu$ is invariant for the skew product (τ^t).*

(b) *The measure μ is stationary for the Markov semigroup (P_t).*

A similar correspondence exists between the *ergodic* stationary measures of the semigroup (P_t) and the ergodic invariant measures of the skew product (τ^t).

Recall the following standard definition from ergodic theory (see, for example, [109]): a (τ^t)-invariant measure η is *ergodic* if, for any bounded measurable $h : \Omega \times Z \to \mathbb{R}$ for which $h \circ \tau^t = h$ holds η-almost-surely for all $t \geq 0$, we have that h is constant η-almost surely. For stationary measures μ of the Markov semigroup (P_t), we use the following definitions:

Definition 3.9 ([73, p. 19]). Let $h : Z \to \mathbb{R}$ be bounded and Borel measurable. Given a stationary μ, we say that ϕ is (P_t, μ)-*invariant* if $P_t\phi = \phi$ holds μ-almost surely for all $t \geq 0$. We say that a set $K \subset Z$ is (P_t, μ) invariant if its characteristic function χ_K is (P_t, μ)-invariant in the above sense.

We call a stationary measure μ *ergodic* if the only (P_t, μ)-invariant functions are μ-almost-surely constant.

Proposition 3.10 ([73, Theorem I.2.1]). *Assume the setting of Lemma 3.8. Let μ be a stationary measure for (P_t), noting that $\mathbf{P} \times \mu$ is an invariant measure for (τ^t) by Lemma 3.8. Then the following are equivalent:*

(a) *The invariant measure $\mathbf{P} \times \mu$ is ergodic for the skew product (τ^t).*

(b) *The stationary measure μ is ergodic for the Markov semigroup (P_t).*

3.2. Linear cocycles over RDS and the Multiplicative Ergodic Theorem

We start by defining and motivating the concept of a linear cocycle over a random dynamical system in Section 3.2.1. Next, in Section 3.2.2 we state precisely the Multiplicative Ergodic Theorem (Theorem 3.13). The remainder of Section 3.2 is devoted to establishing useful corollaries and refinements of Theorem 3.13.

3.2.1. Basic setting: Linear cocycles over RDS. Fix a positive integer d. Roughly speaking, a linear cocycle over a given "base" dynamical system is a composition of *time-dependent* $d \times d$-matrices driven by the dynamics on the base. More precisely, in our setting we have the following definition.

Definition 3.11. Let $\mathcal{T}$ be a continuous RDS as in Section 3.1.1, referred to below as the *base RDS*, and let (τ^t) be its associated skew product as in Section 3.1.3. A d-dimensional linear cocycle $\mathcal{A}$ over the base RDS $\mathcal{T}$ is a mapping $\mathcal{A} : \Omega \to C_{u,b}([0,\infty) \times Z, M_{d\times d}(\mathbb{R}))$ with the following properties:

(i) The evaluation mapping $\Omega \times [0,\infty) \times Z \to M_{d\times d}(\mathbb{R})$ sending $(\omega, t, z) \mapsto \mathcal{A}^t_{\omega,z}$ is $\mathcal{F} \otimes \mathrm{Bor}([0,\infty)) \otimes \mathrm{Bor}(Z)$-measurable.

(ii) The mapping $\mathcal{A}$ satisfies the *cocycle property*: for any $z \in Z, \omega \in \Omega$ we have $\mathcal{A}^0_{\omega,z} = \mathrm{Id}_{\mathbb{R}^d}$, the $d \times d$ identity matrix, and for $s, t \geq 0$ we have

$$\mathcal{A}^{s+t}_{\omega,z} = \mathcal{A}^s_{\tau^t(\omega,z)} \circ \mathcal{A}^t_{\omega,z}. \tag{3.1}$$

To motivate this definition, consider the following example.

Example 3.12. Let Z be a Riemannian manifold and assume that for each element $\omega \in \Omega$, $\mathcal{T}^t_\omega : Z \to Z$ is a C^1 mapping on Z (e.g., the RDS defined in Example 3.3). The cocycle $\mathcal{A}^t_{\omega,z} := D_z \mathcal{T}^t_\omega, z \in Z, t \geq 0$, is often referred to as the *derivative cocycle* for $\mathcal{T}$. The cocycle property (3.1) is a manifestation of the Chain Rule from standard calculus and the cocycle property (ii) in Section 3.1 for the RDS $\mathcal{T}$. For more information, see, e.g., [8, 73].

3.2.2. The Multiplicative Ergodic Theorem (MET). It is of natural interest, in the setting described above, to study the *asymptotic exponential growth rate*

$$\lim_{t\to\infty} \frac{1}{t} \log |\mathcal{A}^t_{\omega,z} v| , \tag{3.2}$$

at $z \in Z, v \in \mathbb{R}^d$. When it exists, the quantity in (3.2) is the *Lyapunov exponent* at z in the direction v. For systems such as those in Example 3.12, the existence and *positivity* of the limit (3.2) implies that the orbit of z is *sensitive with respect to initial conditions*, a possible symptom of an asymptotically chaotic regime for $\mathcal{T}$.

However, there is a priori no guarantee that the limits (3.2) even exist in the first place. As it turns out, the most successful approach to the problem of the existence of the limits (3.2) is through ergodic theory: the limits (3.2) exist for all $v \in \mathbb{R}^d$, $\mathbf{P}$-almost all $\omega \in \Omega$, and for points $z \in Z$ *generic with respect to stationary measures* for the RDS $\mathcal{T}$, modulo a condition ensuring $|\mathcal{A}^t_{\omega,z}|$ does not get too large too fast as $t \to \infty$ for “most” $(\omega, z) \in \Omega \times Z$. This is the content of the MET, which we will now state precisely.

Let μ be a stationary measure for the RDS $\mathcal{T}$ satisfying the independent increments condition (H1). Let $\mathcal{A}$ be a linear cocycle as above. Throughout, we will assume the following *integrability condition* for the cocycle $\mathcal{A}$.

(H2) The triple $(\mathcal{T}, \mathcal{A}, \mu)$ has the property that $\mathcal{A}^t_{\omega,z}$ is an invertible matrix for all $\omega \in \Omega$, $t \in [0, \infty)$, $z \in Z$, and[8]

$$\mathbf{E}\int \Big(\sup_{0\leq t\leq 1} \log^+ |\mathcal{A}^t_{\omega,z}|\Big)\, \mathrm{d}\mu(z), \mathbf{E}\int \Big(\sup_{0\leq t\leq 1} \log^+ |(\mathcal{A}^t_{\omega,z})^{-1}|\Big)\, \mathrm{d}\mu(z) < \infty. \tag{3.3}$$

These conditions are standard for the derivative cocycles of stochastic flows generated by SDE; see, e.g., [74].

Theorem 3.13 (Multiplicative Ergodic Theorem; Theorem 3.4.1 in [8]). *Let $\mathcal{T}$ be a continuous RDS as in Section 3.1.1 satisfying condition (H1). Let μ be an* ergodic *stationary measure associated to $\mathcal{T}$ and assume that $\mathcal{A}$ is a linear cocycle over $\mathcal{T}$ for which the integrability condition (H2) holds. Then there exist r distinct deterministic real numbers*

$$\lambda_1 > \cdots > \lambda_r,$$

$r \in \{1, \ldots, d\}$, a (τ^t)-invariant[9] set $\Gamma \subset \Omega \times Z$ of full $\mathbf{P} \times \mu$-measure, and for each

[8] Here, $\log^+(a) := \max\{0, \log a\}$ for $a > 0$.

[9] That is, $\mathcal{T}^t(\Gamma) \subset \Gamma$ with probability 1 for all $t \geq 0$.

$(\omega, z) \in \Gamma$, a flag of subspaces

$$\mathbb{R}^d =: F_1 \supset F_2(\omega, z) \supset \cdots \supset F_r(\omega, z) \supset F_{r+1} := \{0\},$$

with $\dim F_i \equiv m_i$ *for constants* $m_i \in \{1, \ldots, d\}$, $1 \le i \le r$, *for which the following holds. For any* $1 \le i \le r$ *and* $v \in F_i(\omega, z) \setminus F_{i+1}(\omega, z)$, *we have*

$$\lim_{t \to \infty} \frac{1}{t} \log |\mathcal{A}^t_{\omega,z} v| = \lambda_i. \tag{3.4}$$

Moreover, the assignment $(\omega, z) \mapsto F_i(\omega, z)$ *varies measurably.*[10]

Note that automatically, for any $(\omega, z) \in \Gamma$ and $t > 0$ we have that

$$\mathcal{A}^t_{\omega,z} F_i(\omega, z) = F_i(\tau^t(\omega, z))$$

for each $i = 1, \ldots, r$. For instance, at $i = 2$, the forward growth rates of v and $\mathcal{A}^s_{\omega,z} v$ as in (3.4) must coincide for any $s > 0$, hence $\mathcal{A}^s_{\omega,z}$ must map $F_2(\omega, z)$ to $F_2(\tau^s(\omega, z))$. A similar argument applies for $i > 2$; see, e.g., [108] for more details.

The MET as above is originally due to Oseledets [96]; since then many proofs of the MET have been recorded, each providing a different perspective on this seminal result. One perspective useful to us in this study is that given by the proof-technique of Ragunathan [98] and Ruelle [101, 102]. For future use, we record the following intermediate step in this proof.

Below, for a $d \times d$-matrix A and for $1 \le i \le d$, we write $\sigma_i(A)$ for the i-th singular value of A.

Lemma 3.14. *Let* λ_i *and* $(\omega, z) \mapsto F_i(\omega, z)$, $1 \le i \le r$ *be as in Theorem 3.13.*

(i) *For any* $1 \le i \le d$, *the limits*

$$\chi_i = \lim_{t \to \infty} \frac{1}{t} \log \sigma_i(\mathcal{A}^t_{\omega,z})$$

exist and are constant for $\mathbf{P} \times \mu$*-almost every* $(\omega, z) \in \Omega \times Z$. *Moreover, the Lyapunov exponents* λ_i, $1 \le i \le r$ *are precisely the distinct values among the* χ_i, $1 \le i \le d$.

(ii) *For* $\mathbf{P} \times \mu$*-almost every* $(\omega, z) \in \Omega \times Z$, *the limit*

$$\Lambda_{\omega,z} := \lim_{t \to \infty} \frac{1}{t} \log\big((\mathcal{A}^t_{\omega,z})^\top \mathcal{A}^t_{\omega,z}\big)$$

exists. The matrix $\Lambda_{\omega,z}$ *is symmetric with distinct eigenvalues* λ_i, $1 \le i \le r$ *and corresponding eigenspaces* $E_1(\omega, z), \ldots, E_r(\omega, z)$. *Moreover, for each* $1 \le i \le r$ *we have*

$$F_i(\omega, z) = \bigoplus_{j=i}^{r} E_j(\omega, z).$$

[10] We view the codomain of this mapping as the Grassmannian manifold $\mathrm{Gr}(\mathbb{R}^d)$ of subspaces of $\mathbb{R}^d$ equipped with its corresponding Borel σ-algebra.

Lemma 3.14 (i) is often proved using the Kingman Subadditive Ergodic Theorem [75]. Item (ii) follows from item (i) and a linear algebra argument; see [98,101] for more details.

Note that from Lemma 3.14 (i), we have that $\lambda_1 = \lambda^+$ and $\lambda_r = \lambda^-$, where

$$\lambda^+ = \lim_{t\to\infty} \frac{1}{t} \log |\mathcal{A}^t_{\omega,z}|, \quad \lambda^- = \lim_{t\to\infty} -\frac{1}{t} \log |(\mathcal{A}^t_{\omega,z})^{-1}|, \tag{3.5}$$

since for any invertible matrix $A \in M_{d\times d}(\mathbb{R})$ we have $\sigma_1(A) = |A|$, $\sigma_d(A) = |A^{-1}|^{-1}$. In particular, $r > 1$ (i.e., there exist at least two distinct Lyapunov exponents) if and only if $\lambda^+ > \lambda^-$. Of course, the problem of verifying that $\lambda^+ > \lambda^-$ for concrete systems is often extremely challenging: this is precisely the subject of Sections 3.3 and 4.

For the remainder of Section 3.2 we will continue our discussion of linear cocycles and the MET by introducing several auxiliary processes associated to a linear cocycle $\mathcal{A}$, namely, the *projective process* (Section 3.2.3) and *matrix processes* (Section 3.2.4), as well as the $(-\top)$-*cocycle* $\check{\mathcal{A}}$ associated to $\mathcal{A}$ (Section 3.2.5).

3.2.3. Projective RDS associated to the cocycle $\mathcal{A}$. Let us write P^{d-1} for the projective space associated to $\mathbb{R}^d$. The action of an invertible matrix $A \in M_{d\times d}(\mathbb{R})$ on $\mathbb{R}^d$ descends to a well-defined action $A : P^{d-1} \to P^{d-1}$.

With this understanding, we can think of the cocycle $\mathcal{A}$ as giving rise to an RDS on the product $Z \times P^{d-1}$, i.e., that given for $\omega \in \Omega$ by

$$(t, z, v) \mapsto (\mathcal{T}^t_\omega z, \mathcal{A}^t_{\omega,z} v), \quad (z, v) \in Z \times P^{d-1}, t \in [0, \infty).$$

We refer to the RDS on $Z \times P^{d-1}$ as the *projective RDS* or *projective process*. As one can easily check, this is a continuous RDS in the sense of Section 3.1.1 with $Z \times P^{d-1}$ replacing Z. Correspondingly, we will assume in what follows that the following independent increments condition, analogous to (H1), is satisfied:

(H3) For all $s, t > 0$, we have that the $C_{u,b}(Z, Z) \times C_{u,b}(Z, M_{d\times d}(\mathbb{R}))$-valued random variables $(\mathcal{T}^t_\cdot, \mathcal{A}^t_{\cdot,\cdot})$ and $(\mathcal{T}^s_{\theta^t \cdot}, \mathcal{A}^s_{\theta^t \cdot,\cdot})$ on $(\Omega, \mathcal{F}, \mathbf{P})$ are independent.

Assumption (H3) ensures (Lemma 3.4) that associated to the RDS on $Z \times P^{d-1}$ is a Markov process $(z_t, v_t)_{t\geq 0}$ on $Z \times P^{d-1}$ with transition kernel

$$\begin{aligned}\hat{P}_t((z, v), K) &= \mathbf{P}((z_t, v_t) \in K \mid (z_0, v_0) = (z, v)) \\ &= \mathbf{P}\{(\mathcal{T}^t_\omega z, \mathcal{A}^t_{\omega,z} v) \in K\}\end{aligned}$$

defined for $(z, v) \in Z \times P^{d-1}$, $K \subset Z \times P^{d-1}$ Borel. In addition, we can consider the associated skew product semiflow $\hat{\tau}^t : \Omega \times Z \times P^{d-1} \to \Omega \times Z \times P^{d-1}$, $t \in [0, \infty)$, as in Section 3.1.3.

We now turn our attention to the relationship between the ergodic theory of the projective process and the MET. It is not hard to see that any stationary measure ν for $(\hat{P}_t)$ must project to some (P_t)-stationary measure μ on the Z-factor. Conversely, by Lemma 3.6 we have the following.

Lemma 3.15. *Given a stationary measure μ for (P_t), there exists at least one stationary measure ν for the projective semigroup $(\hat{P}_t)$ such that $\nu(A \times P^{d-1}) = \mu(A)$.*

If ν as above is the *unique* stationary measure with marginal μ, then we obtain the following refinement of the MET.

Proposition 3.16. *Assume that there is only one stationary measure ν for the projective RDS projecting to μ on the Z-factor. Then we have the following: for μ-almost every $z \in Z$ and any $v \in \mathbb{R}^d \setminus \{0\}$, we have*

$$\lim_{t\to\infty} \frac{1}{t} \log |\mathcal{A}^t_{\omega,z} v| = \lambda_1$$

with $\mathbf{P}$*-probability* 1.

Note that Proposition 3.16 is actually a corollary of the more general Random Multiplicative Ergodic Theorem, discovered independently by Kifer ([73, Theorem III.1.2]) and Carverhill ([27]), describing the situation when several stationary measures ν project to a single stationary μ. Since we do not use this more general formulation here, we omit it and refer the interested reader to the references above for more information.

3.2.4. Matrix RDS associated to the cocycle $\mathcal{A}$. The cocycle $\mathcal{A}$ also gives rise to an RDS on the product space $Z \times M_{d\times d}(\mathbb{R})$; for $\omega \in \Omega$, the time-t mapping applied to $(z, B) \in Z \times M_{d\times d}(\mathbb{R})$ is given by

$$(z, B) \mapsto (\mathcal{T}^n_\omega z, \mathcal{A}^n_{\omega,z} B).$$

Like before, this RDS on $Z \times M_{d\times d}(\mathbb{R})$ falls into the framework given in Section 3.1.1 with $Z \times M_{d\times d}(\mathbb{R})$ replacing Z.

Similarly, under the independent increments hypothesis (H3) we can associate to this RDS a Markov process (z_t, A_t) on $Z \times M_{d\times d}(\mathbb{R})$ with transition kernel $Q_t((z, A), K)$. Note that if the matrix $A \in M_{d\times d}(\mathbb{R})$ is invertible and $K = K_1 \times K_2$, where $K_1 \subset Z$, $K_2 \subset M_{d\times d}(\mathbb{R})$, then

$$Q_t((z, A), K) = Q_t((z, \mathrm{Id}), K_1 \times (K_2 A^{-1})),$$

where $\mathrm{Id} = \mathrm{Id}_{\mathbb{R}^d}$. Thus, frequently we are only interested in the Markov kernel (Q_t) evaluated at (z, Id).

3.2.5. The MET for the $(-\top)$-cocycle $\check{\mathcal{A}}$. In this paper we will also need to consider what we call the $(-\top)$-*cocycle* $\check{\mathcal{A}}$, defined for $z \in Z, \omega \in \Omega, t \geq 0$ by

$$\check{\mathcal{A}}^t_{\omega,z} = (\mathcal{A}^t_{\omega,z})^{-\top}.$$

Here, "$(-\top)$" refers to the inverse-transpose of a $(d \times d)$-matrix. As one can easily check, $\check{\mathcal{A}}$ is a linear cocycle over the RDS $\mathcal{T}$; when (H2) and (H3) for the original cocycle $\mathcal{A}$ are assumed, the same hold for the $(-\top)$-cocycle $\check{\mathcal{A}}$. Therefore the MET (Theorem 3.13) and all the aforementioned material applies, yielding Lyapunov exponents $\check{\lambda}_1 > \cdots > \check{\lambda}_{\check{r}}$ and associated subspaces $\check{F}_2(\omega, z), \ldots, \check{F}_{\check{r}}(\omega, z)$.

These objects can be directly represented in terms of the exponents and subspaces of the original cocycle $\mathcal{A}$.

Proposition 3.17 ([8, Theorem 5.1.1]). *We have $\check{r} = r$, and for each $1 \le i \le r$, we have*

$$\check{\lambda}_i = -\lambda_{r-(i-1)}, \tag{3.6}$$

$$\check{F}_i(\omega, z) = \big(F_{r-(i-1)+1}(\omega, z)\big)^{\perp} \quad \textit{for almost all } (\omega, z) \in \Omega \times Z. \tag{3.7}$$

Proof. This follows on applying Lemma 3.14 to the cocycle $\check{\mathcal{A}}$ and noting that

$$\log\big((\check{\mathcal{A}}^t_{\omega,z})^\top \check{\mathcal{A}}^t_{\omega,z}\big) = -\log\big((\mathcal{A}^t_{\omega,z})^\top \mathcal{A}^t_{\omega,z}\big)$$

holds for all $(\omega, z) \in \Omega \times Z$ and $t \ge 0$. ■

Under assumption (H3), the cocycle $\check{\mathcal{A}}$ induces the $(-\top)$-*projective process* $(z_t, \check{v}_t)$ on $Z \times P^{d-1}$ defined for fixed initial $z_0 \in Z, \check{v}_0 \in P^{d-1}$ by setting $\check{v}_t$ to be the projective representative of $\check{\mathcal{A}}^t_{\omega,z_0}\check{v}_0$. Then all the material from Section 3.2.3 applies with $\check{\mathcal{A}}$ replacing $\mathcal{A}$ and $(z_t, \check{v}_t)$ replacing (z_t, v_t).

In particular, the conclusions of Proposition 3.16 hold with $\check{\mathcal{A}}$ replacing $\mathcal{A}$ when the stationary measure for $(z_t, \check{v}_t)$ projecting to μ on the Z factor is unique.

3.3. *The MET in the random setting: Furstenberg's criterion*

Furstenberg's criterion was originally discovered by Furstenberg in his seminal 1968 paper, *Noncommuting random products* [56]. It has since been refined and extended over the subsequent years by a variety of authors; see Section 2.2 for some citations.

In Section 3.3.1 we will state Furstenberg's criterion precisely in the setup of Sections 3.1 and 3.2. In Section 3.3.2 we provide a condition for checking Furstenberg's criterion which is most useful when $\mathcal{T}$ and $\mathcal{A}$ are generated by finite-dimensional SDE. In Section 4 we will consider conditions for checking Furstenberg's criterion which are amenable to the situation when the phase space for $\mathcal{T}$ is more general and, possibly, infinite-dimensional.

For the remainder of Section 3 we assume the setting of Sections 3.1 and 3.2. Specifically, $\mathcal{T}$ is a continuous RDS on the metric space Z as in Section 3.1.1 satisfying (H1) and admitting an ergodic stationary measure μ, while the cocycle $\mathcal{A}$ over $\mathcal{T}$ satisfies the conditions of Section 3.2.1 as well as the integrability condition (H2) and the independent increments condition (H3).

3.3.1. Furstenberg's criterion in the RDS setting. Furstenberg's criterion revolves around a central theme: if $\lambda^+ = \lambda^-$ as above, then there is a *deterministic*, i.e., ω-independent, structure preserved by the cocycle $\mathcal{A}$ with probability one. Below, given an invertible matrix M and a probability measure ν on P^{d-1}, we write $M_*\nu = \nu \circ M^{-1}$ for the *pushforward* of ν by M. Here and throughout, we abuse notation and think of invertible matrices M as mappings $P^{d-1} \to P^{d-1}$.

Theorem 3.18. *If $\lambda^+ = \lambda^-$, then for each $z \in Z$ there a is Borel measure ν_z on P^{d-1} such that* (i) *the assignment $z \mapsto \nu_z$ is measurable*[11] *and* (ii) *for each $t \in [0, \infty)$ and*

[11] To wit, for any Borel $K \subset P^{d-1}$, the function $z \mapsto \nu_z(K)$ is Borel measurable. Equivalently, $z \mapsto \nu_z$ is Borel measurable in the weak* topology on finite Borel measures on P^{d-1}.

$(\mathbf{P} \times \mu)$-almost all $(\omega, z) \in \Omega \times Z$ (perhaps depending on t), we have that

$$(\mathcal{A}^t_{\omega,z})_* \nu_z = \nu_{\mathcal{T}^t_\omega z}. \tag{3.8}$$

Theorem 3.18 as above is a consequence of [88, Proposition 2 and Theorem 3]. Deducing the version given above requires passing from the discrete-time setting of [88] to our present continuous-time setting, and is the reason why the $(\mathbf{P} \times \mu)$-almost sure set may depend on t. Further details are left to the reader.

Note that automatically, if $\lambda^+ = \lambda^-$, then the measure ν on $Z \times P^{d-1}$ defined by

$$\mathrm{d}\nu(z, v) = \mathrm{d}\mu(z)\mathrm{d}\nu_z(v), \quad (z, v) \in Z \times P^{d-1},$$

is a stationary measure for the Markov semigroup $(\hat{P}^t)$ associated to the projective RDS on $Z \times P^{d-1}$.

We conclude that $\lambda^+ > \lambda^-$ if, from the conclusions of Theorem 3.18, we can derive a contradiction. Our goal in the remainder of Section 3 is to identify criteria for the cocycle $\mathcal{A}$ under which a contradiction can be derived.

Before continuing, let us establish some useful vocabulary. Any measurable family (ν_z) of probability measures on P^{d-1} will be referred to as a *family of fiber measures*, while for $z \in Z$ the individual measure ν_z will be called the *fiber measure at z*. If the family of fiber measures (ν_z) satisfies (3.8) for all $t \geq 0$ and $\mathbf{P} \times \mu$-almost every $(\omega, z) \in \Omega \times Z$ (the almost-sure set perhaps depending on t), we call (ν_z) an *invariant fiber measure family*.

3.3.2. Nondegeneracy of conditional laws. For simplicity, and since our primary application in this paper falls in this special case, let us restrict our attention to the case when $\mathcal{A}$ is an $\mathrm{SL}_d(\mathbb{R})$ cocycle. That is, $\det \mathcal{A}^t_{\omega,z} \equiv 1$ for all $t \geq 0, z \in Z, \omega \in \Omega$.

Our starting point is the following observation.

Lemma 3.19. *Let ν, ν' be Borel probability measures on P^{d-1}. Then the set*

$$\{A \in \mathrm{SL}_d(\mathbb{R}) : A_* \nu = \nu'\} \subset \mathrm{SL}_d(\mathbb{R})$$

has empty interior.

The proof is straightforward and is omitted.

In relation to condition (3.8), Lemma 3.19 says that if for some $t_0 > 0$ we can somehow fix both z and the image $z' = \mathcal{T}^{t_0}_\omega z$, then the set of matrices mapping the measure $\nu = \nu_z$ to $\nu' = \nu_{z'}$ is "small" in the topological sense.

We can make sense of this using regular conditional probabilities. Let us consider the measure $Q_{t_0}((z, \mathrm{Id}), \cdot)$ on $Z \times M_{d\times d}(\mathbb{R})$ and disintegrate it according to the value z_{t_0} attained by the (z_t) process, conditioned on $z_0 = z$. To wit, fix $t_0 > 0$; for $P_{t_0}(z, \cdot)$-generic $z' \in Z$, we intend to define the *regular conditional probability*

$$Q^{t_0}_{z,z'}(K) := \mathbf{P}(\mathcal{A}^{t_0}_{\omega,z} \in K \mid \mathcal{T}^{t_0}_\omega z = z'), \quad K \in \mathrm{Bor}(\mathrm{SL}_d(\mathbb{R})).$$

This is justified rigorously below.

Lemma 3.20 ([30]). *Assume $\mathcal{A}$ is an $\mathrm{SL}_d(\mathbb{R})$ cocycle and that Ω is a Borel subset of a Polish space equipped with the σ-algebra $\mathcal{F}$ of Borel subsets of Ω. Fix $z \in Z$. Then* *there is a mapping* $Z \times \mathrm{Bor}(\mathrm{SL}_d(\mathbb{R})) \mapsto [0,1]$, $(z', K) \mapsto Q^{t_0}_{z,z'}(K)$, *with the following properties:*

(1) *For each* $K \in \mathrm{Bor}(\mathrm{SL}_d(\mathbb{R}))$, *the mapping* $z' \mapsto Q^{t_0}_{z,z'}(K)$ *is Borel measurable.*

(2) *For* $P_{t_0}(z,\cdot)$*-almost all* $z' \in Z$, *the set function* $Q^{t_0}_{z,z'}(\cdot) : \mathrm{Bor}(\mathrm{SL}_d(\mathbb{R})) \to [0,1]$ *is a Borel probability measure on* $\mathrm{SL}_d(\mathbb{R})$.

(3) *For any bounded measurable function* $h : Z \times \mathrm{SL}_d(\mathbb{R}) \to \mathbb{R}$, *we have that*

$$\int h(z', A')\, Q^{t_0}_{z,z'}(\mathrm{d}A') P_{t_0}(z, \mathrm{d}z') = \int h(z', A') Q_{t_0}((z, \mathrm{Id}), \mathrm{d}(z', A')).$$

Definition 3.21. Let $\mathcal{A}$ be an $\mathrm{SL}_d(\mathbb{R})$-cocycle and assume $(\Omega, \mathcal{F})$ is as in Lemma 3.20. We say that $\mathcal{A}$ satisfies condition (C) if there is a $t_0 > 0$ and a set $S \subset Z$ of positive μ-measure with the following property: for each $z \in S$, there is a $P_{t_0}(z,\cdot)$-positive measure set $S_z \subset Z$ such that $Q^{t_0}_{z,z'}(\cdot)$ is defined and is absolutely continuous with respect to Lebesgue measure on $\mathrm{SL}_d(\mathbb{R})$.

Note that if (C) holds and $z \in S$, $z' \in S_z$, then the support of $Q^{t_0}_{z,z'}(\cdot)$ has nonempty interior in $\mathrm{SL}_d(\mathbb{R})$. Therefore by Theorem 3.18 and Lemma 3.19 we conclude the following.

Corollary 3.22. *If the* $\mathrm{SL}_d(\mathbb{R})$*-cocycle* $\mathcal{A}$ *satisfies condition* (C), *then* $\lambda^+ > \lambda^-$. *In particular,* $\lambda_1 > 0$.

Proof. By Theorem 3.18, $\lambda^+ > \lambda^-$. Since $\mathcal{A}$ is an $\mathrm{SL}_d(\mathbb{R})$ cocycle, it follows from basic linear algebra that

$$1 = \det(\mathcal{A}^t_{\omega,z}) = \prod_{i=1}^{d} \sigma_i(\mathcal{A}^t_{\omega,z})$$

for all $z \in Z, \omega \in \Omega, t \geq 0$. Thus from Lemma 3.14 we have that $\sum_{i=1}^d \chi_i = 0$, (χ_i) as in Lemma 3.14 (i). Since $\lambda_1 = \lambda^+ = \chi_1, \lambda_r = \lambda^- = \chi_d$, we conclude from $\lambda^+ > \lambda^-$ that $\chi_1 > 0$ and $\chi_d < 0$. ■

Condition (C) holds for a large class of systems for which the process (z_t, A_t) is governed by a finite-dimensional SDE on $Z \times \mathrm{SL}_d(\mathbb{R})$; see Section 2.7. We note that condition (C) is a straightforward adaptation of a condition given in [27] for the Lyapunov exponent of a divergence-free SDE be positive.

4. Positive Lyapunov exponents for cocycles over infinite-dimensional RDS

For stochastic processes on infinite-dimensional spaces there is no corresponding analogue of Hörmander's Theorem. As a result it is frequently quite difficult in applications to verify the condition (C) (Definition 3.21).

Thankfully, condition (C) is far from necessary to rule out the criterion in Theorem 3.18. In this section we prove a sufficient condition, weaker than (C), which is better suited for infinite-dimensional RDS. To the best of our knowledge, this result appears to be new. The proof is carried out in several steps.

First, in Section 4.1 we will establish the *continuous dependence* of an invariant fiber measure family (ν_z) on the base point z under the assumption that the Markov semigroup P_t associated to the RDS $\mathcal{T}$ has the strong Feller property (Definition 4.1 below). Leveraging this continuity result, in Section 4.2 we will take advantage of algebraic properties of $\mathrm{SL}_d(\mathbb{R})$ to obtain a classification (Theorem 4.7) for the family (ν_z) under the assumption that $\lambda^+ = \lambda^-$ as in Furstenberg's criterion (Theorem 3.18). Finally, in Section 4.3 we will state a weakening (C') (Definition 4.16) ruling out each alternative in the classification we obtain.

For the entirety of Section 4, we assume the setting given at the beginning of Section 3.3.

4.1. *From measurable to topological*

The goal of Section 4.1 is to turn the measurable information contained in Theorem 3.18, namely, that the invariant measure family (ν_z) satisfies (3.8) for $(\mathbf{P} \times \mu)$-almost all (ω, z), into topological information concerning "all" ω, in a suitable sense, and all z in a closed set. This will be accomplished in two phases: First, the family (ν_z) will be replaced with a *μ-almost sure version* $(\bar{\nu}_z)_{z \in \operatorname{supp}\mu}$ which is weak* continuous as z varies in $\operatorname{supp}\mu$ (Proposition 4.3). Second, the $\mathbf{P} \times \mu$-almost sure relation (3.8) for the family (ν_z) will be turned into a corresponding relation among the family $(\bar{\nu}_z)$ for all $z \in \operatorname{supp}\mu$ and "$\mathbf{P}$-almost-all" replaced by "all", in a sense to be made precise (Lemma 4.4).

The material in Section 4.1 is analogous to [12, Proposition 6.3 and Lemma 6.5]. For a summary of the differences between the latter and our results in this setting, see Remark 4.15 below.

Going forward, we will require an additional regularity assumption on the Markov semigroup (P_t) associated to the RDS $\mathcal{T}$, which we now spell out here.

Definition 4.1. We say that the Markov semigroup (P_t) has the *strong Feller property* if for all bounded, measurable $h : Z \to \mathbb{R}$, and for all $t > 0$, the function $P_t h : Z \to \mathbb{R}$ is bounded and continuous.

At times it will be helpful to use the following well-known result regarding strong Feller semigroups.

Lemma 4.2. *Assume Z is a Polish space.*

(a) *If the Markov semigroup (P_t) on Z has the strong Feller property, then it is automatically* ultra Feller, *i.e., for all $t > 0$ the mapping $z \mapsto P_t(z, \cdot)$ is continuous in the total variation distance*[12] *$\|\cdot\|_{tv}$ on the space of finite signed measures on Z.*

(b) *Let μ be a stationary measure for (P_t) and let $K \subset Z$ be a Borel set of full μ measure. Then $P_t(z, K) = 1$ for all $t > 0$ and $z \in \operatorname{supp}\mu$.*

[12]Given two finite signed measures η_1 and η_2 on the same measurable space $(X, \mathfrak{F})$, the total variation distance is defined by $\|\eta_1 - \eta_2\|_{tv} = \sup_{K \in \mathfrak{F}} |\eta_1(K) - \eta_2(K)|$.

Proof. Item (a) is proved in [103].

For (b), one checks that for all $t \geq 0$, the set $\{z \in Z : P_t(z, K) = 1\}$ is dense in $\operatorname{supp} \mu$. Item (b) now follows from continuity in total variation as in (a). ■

With these preparations out of the way, we can now state precisely the first step in our program, a continuity result for the invariant measure family $(\nu_z)_{z \in Z}$.

Proposition 4.3. *Assume (P_t) is strong Feller, and let (ν_z) be an invariant fiber measure family on Z as in Section 3.3.1. Then, there exists an invariant fiber measure family $(\bar{\nu}_z)$, defined for $z \in \operatorname{supp} \mu \subset Z$, with the following properties:*

(a) *The family $(\bar{\nu}_z)$ is a* μ-almost sure version *of the original family (ν_z), that is, for μ-almost every $z \in \operatorname{supp} \mu$, we have $\nu_z = \bar{\nu}_z$.*
(b) *The family $(\bar{\nu}_z)$ is continuously varying in the weak* topology on P^{d-1}.*

That is, by Proposition 4.3 we can replace the possibly discontinuous invariant measure family (ν_z) with a continuously-varying invariant measure family $(\bar{\nu}_z)$ defined at each $z \in \operatorname{supp} \mu$, at the expense of modifying (ν_z) on a set of μ-measure zero. So as not to interrupt the flow of ideas, Proposition 4.3 is proved at the end of Section 4.1.

Let us now describe the second step in our program, namely, turning the $\mathbf{P} \times \mu$-almost sure relation (3.8) into an analogous relation holding "surely" – roughly speaking, holding for all $(\omega, z) \in \Omega \times \operatorname{supp} \mu$ and for all $t \geq 0$, in a sense we make precise below.

To begin, some notation: let us write

$$\mathcal{C} = C_{u,b}(Z, Z) \times C_{u,b}(Z, M_{d \times d}(\mathbb{R}))$$

equipped with the product topology, where the $C_{u,b}$ spaces are as in Definition 3.1. Elements of $\mathcal{C}$ are written (T, A), where

$$T : Z \to Z, \quad z \mapsto Tz \in Z$$

and

$$A : Z \to M_{d \times d}(\mathbb{R}), \quad z \mapsto A_z \in M_{d \times d}(\mathbb{R}).$$

Given $t \geq 0$, let us write $\mathcal{S}_t$ for the topological support of the $\mathcal{C}$-valued random variable $(\mathcal{T}^t_\omega, \mathcal{A}^t_\omega)$ where ω is distributed as $\mathbf{P}$. We set

$$\mathcal{S} = \overline{\bigcup_{t \geq 0} \mathcal{S}_t}$$

for the closure of the union of the $\mathcal{S}_t$ in $\mathcal{C}$.

Lemma 4.4. *Assume the setting, notation and conclusions of Proposition 4.3. Then, for any $z \in \operatorname{supp} \mu$ and $(T, A) \in \mathcal{S}$, we have that $Tz \in \operatorname{supp} \mu$, and*

$$A_z \bar{\nu}_z = \bar{\nu}_{Tz}. \tag{4.1}$$

Relation (4.1) for all $(T, A) \in \mathcal{S}$ is analogous to the "measure-theoretical" relation (3.8); in contrast to the latter, (4.1) holds identically for all (T, A) in the *closed* subset $\mathcal{S} \subset \mathcal{C}$. For this reason we regard (4.1) as a "topological" statement, as opposed to a measure-theoretic one.

We now turn to the proofs of Proposition 4.3 and Lemma 4.4.

Proof of Proposition 4.3. Fix a continuous function $g : P^{d-1} \to \mathbb{R}$. Define $G : Z \to \mathbb{R}$ by

$$G(z) = \int g(v)\,\mathrm{d}\nu_z(v).$$

We begin by making the following claim.

Claim 4.5. *There is a full μ-measure subset $\tilde{Z} \subset \operatorname{supp}\mu$ with the following property. Let $G : Z \to \mathbb{R}$ be as above. Then $G|_{\tilde{Z}}$ has the property that for any Cauchy sequence $\{z^m\}_{m\geq 1} \subset \tilde{Z}$, we have that the sequence $\{G(z^m)\}_{m\geq 1}$ is Cauchy.*

Assuming the claim, let us define the family $(\bar{\nu}_z)$. To start, for $z \in \tilde{Z}$ we set $\bar{\nu}_z := \nu_z$. Note that this ensures $(\bar{\nu}_z)$ is a version of (ν_z) as in item (a) above.

Next, for $z \in \operatorname{supp}\mu \setminus \tilde{Z}$, we define $\bar{\nu}_z$ as follows. Since $\tilde{Z}$ is dense in $\operatorname{supp}\mu$, we can find a sequence $\{z^m\}_{m\geq 1} \subset \tilde{Z}$ converging to z. We now define $\bar{\nu}_z$ to be any weak* limit of the $\bar{\nu}_{z^m}$ (at least one exists by Prokhorov's Theorem since P^{d-1} is compact [19]).

Indeed, the weak* limit $\lim_{m\to\infty} \bar{\nu}_{z^m}$ actually exists: to see this, fix any $g : P^{d-1} \to \mathbb{R}$ continuous and observe that the sequence $\{G(z^m) = \int g(v)\,\mathrm{d}\bar{\nu}_{z^m}(v)\}_{m\geq 1}$ is Cauchy by the Claim; this implies weak* convergence. Moreover, this same argument implies that the definition of $\bar{\nu}_z, z \in \operatorname{supp}\mu \setminus \tilde{Z}$ is independent of the approximating sequence $\{z^m\}_{m\geq 1}$ in $\tilde{Z}$.

To show continuity as in item (b), fix a continuous $g : P^{d-1} \to \mathbb{R}$; we will check that $\bar{G}(z) := \int g(u)\mathrm{d}\bar{\nu}_z(u)$ is a continuous real-valued function. For this, fix $z \in \operatorname{supp}\mu$ and let $\{z^m\}_{m\geq 1} \subset \operatorname{supp}\mu$ be a sequence converging to z. For each m, fix $\check{z}^m \in \tilde{Z}$ such that $d(\check{z}^m, z^m) < 1/m$ and $|G(\check{z}^m) - \bar{G}(z)| < 1/m$. Then

$$|\bar{G}(z^m) - \bar{G}(z)| \leq |\bar{G}(z^m) - G(\check{z}^m)| + |G(\check{z}^m) - \bar{G}(z)| \leq \frac{1}{m} + |G(\check{z}^m) - \bar{G}(z)|.$$

The claim and our definition of $\bar{\nu}_z$ imply that the second right-hand-side-term goes to zero. This completes the proof of continuity as in item (b). It remains to prove the claim.

Proof of Claim 4.5. It is straightforward to construct a full μ-measure subset $\tilde{Z} \subset \operatorname{supp}\mu$ with the property that for all $z \in \tilde{Z}$ and rational t, we have with probability 1 that $\mathcal{T}^t_\omega z \in \tilde{Z}$ and that (3.8) holds. For such $z \in \tilde{Z}$, on integrating the left- and right-hand sides of (3.8) with respect to $d\mathbf{P}(\omega)$, we obtain that

$$\int (\hat{P}_t g)(z, v)\,\mathrm{d}\bar{\nu}_z(v) = P_t G(z),$$

where $\hat{P}_t$ denotes the Markov semigroup associated to the projective process as defined in Section 3.2.3.

Now, fix a Cauchy sequence $\{z^m\}_{m\geq 1} \subset \tilde{Z}$ converging to some $z \in Z$. Fix $\varepsilon > 0$ and fix a neighborhood U of z; without loss, $\{z^m\}_{m\geq 1} \subset U$. Since $\hat{P}_t g \to g$ uniformly on bounded subsets of $Z \times P^{d-1}$ (Proposition 3.5 (b)), we have $\hat{P}_t g \to g$ uniformly on $U \times P^{d-1}$. Fix $t = t_\varepsilon$ for which $|\hat{P}_s g - g| < \varepsilon$ on all of $U \times P^{d-1}$ for all $s \in [0, t_\varepsilon]$.

Fix a rational $t_* \in [0, t_\varepsilon]$. Given integers $m, m' \geq 1$, we estimate (note that the function $G(z) = \int g(v)\,\mathrm{d}\nu_z(v)$ is measurable and bounded from above):

$$\begin{aligned}|G(z^m) - G(z^{m'})| &= \left| \int g(v)\,\mathrm{d}\nu_{z^m}(v) - \int g(v)\,\mathrm{d}\nu_{z^{m'}}(v) \right| \\ &\leq \int |g(u) - P_{t_*} g(v)|\,\mathrm{d}\nu_{z^m}(v) \\ &\quad + \left| \int P_{t_*} g(v)\,\mathrm{d}\nu_{z^m}(v) - \int P_{t_*} g(v)\,\mathrm{d}\nu_{z^{m'}}(v) \right| \\ &\quad + \int |g(v) - P_{t_*} g(v)|\,\mathrm{d}\nu_{z^{m'}}(v) \\ &\leq 2\varepsilon + |P_{t_*} G(z^m) - P_{t_*} G(z^{m'})|.\end{aligned}$$

Now, $P_{t_*} G$ is a continuous function by the strong Feller property, and so $\{P_{t_*} G(z^m)\}_{m \geq 1}$ is a Cauchy sequence. The Cauchy property for $\{G(z^m)\}_{m \geq 1}$ now follows. ■

Proof of Lemma 4.4. We begin by verifying that $Tz \in \operatorname{supp} \mu$ for any $z \in \operatorname{supp} \mu$ and $(T, A) \in \mathcal{S}$. To start, observe that since $\operatorname{supp} \mu$ has full μ-measure, we have from stationarity that $P_t(z, \operatorname{supp} \mu) = 1$ for all $t > 0$ and for μ-almost all $z \in Z$. As one can easily check, for continuous RDS $\mathcal{T}$ as in Section 3.1.1 satisfying (H1), the mapping $z \mapsto P_t(z, \cdot)$ is weak* continuous (irrespective of the strong Feller property). Thus, by the Portmanteau Theorem and the density of μ-almost sure sets in $\operatorname{supp} \mu$, we conclude that $P_t(z, \operatorname{supp} \mu) = 1$ for all $z \in \operatorname{supp} \mu$.

So, for any fixed $z \in Z$, we have for all $t \geq 0$ that $\mathcal{T}_\omega^t z \in \operatorname{supp} \mu$ with probability 1. In particular, any $(T, A) \in \mathcal{S}_t$ is the limit (in the topology on $\mathcal{C}$) of elements $(T^m, A^m) \in \mathcal{S}_t$ for which $T^m z \in \operatorname{supp} \mu$ for all m. Therefore $Tz \in \operatorname{supp} \mu$ holds by the closedness of $\operatorname{supp} \mu$ for any $(T, A) \in \mathcal{S}_t$. A similar argument implies $Tz \in \operatorname{supp} \mu$ for any $(T, A) \in \mathcal{S}$.

Let us now move on to verifying relation (4.1). For $z \in \operatorname{supp} \mu$, we define

$$G_z = \{(T, A) \in \mathcal{C} : (A_z)_* \bar{\nu}_z = \bar{\nu}_{Tz}\}.$$

Note that by the argument in the previous two paragraphs, $\bar{\nu}_{Tz}$ is defined for all $z \in \operatorname{supp} \mu$ and $(T, A) \in \mathcal{S}$. To complete the proof of Lemma 4.4 it will suffice to show that $G_z \supset \mathcal{S}$ for all $z \in \operatorname{supp} \mu$.

To start, one checks that G_z is closed in $\mathcal{C}$ by the closedness of $\operatorname{supp} \mu$ and the fact that $z \mapsto \bar{\nu}_z$ is weak* continuous. Next, let $\tilde{Z}$ be as constructed in the proof of Claim 4.5. It follows that for $z \in \tilde{Z}$ and all rational t that

$$\mathbf{P}((\mathcal{T}_\omega^t, \mathcal{A}_\omega^t) \in G_z) = 1.$$

So, for all rational $t \geq 0$ we deduce that G_z is dense in $\mathcal{S}_t$, hence $G_z \supset \mathcal{S}_t$ since $G_z, \mathcal{S}_t$ are closed in $\mathcal{C}$. Moreover, for irrational $t \geq 0$, each $(T, A) \in \mathcal{S}_t$ is a limit of elements $(T^n, A^n) \in \mathcal{S}_{t_n}$ in $\mathcal{C}$, where $\{t_n\}$ is a sequence of rationals for which $t_n \to t$ as $n \to \infty$. Again by closedness of G_z we deduce that $G_z \supset \mathcal{S}_t$ for all $t \geq 0$. We conclude $G_z \supset \mathcal{S}$ for all $z \in \tilde{Z}$.

To conclude for $z \in \operatorname{supp} \mu \setminus \tilde{Z}$, let $z^m \to z$ be a convergent sequence, $z^m \in \tilde{Z}$, and fix $(T, A) \in \mathcal{S}$. That $(T, A) \in G_z$ now follows from the fact that $(T, A) \in G_{z^m}$ for all m from above and from the continuity of $z \mapsto \bar{\nu}_z$. This completes the proof of Lemma 4.4.

4.2. A refinement of Furstenberg's criterion

The refinement of Furstenberg's criterion we present here is effectively a *classification* of the fiber measures $\nu_z, z \in \operatorname{supp} \mu$ comprising a family satisfying the "topological" relation (4.1). For the sake of brevity, and because it serves our purposes in this paper, we prove this classification when d, the dimension of the cocycle $\mathcal{A}$, is less than or equal to 3, although it is likely to hold in higher dimensions (see Remark 4.14).

This classification, Theorem 4.7 below, is the analogue in our setting of [12, Theorem 6.8]. Our situation is significantly more general and entails several subtleties unique to our setting; see Remarks 4.12 and 4.15 for more discussion.

The germ of this idea comes from the geometry of $\mathrm{SL}_d(\mathbb{R})$ and the restrictions placed on the subgroup of matrices preserving a single projective measure. To wit, we have the following (for any dimension $d \geq 1$):

Lemma 4.6. *Let $d \geq 1$. Let η be a Borel measure on P^{d-1}. Define $H = H_\eta \subset \mathrm{SL}_d(\mathbb{R})$ to be the subgroup of matrices $A \in \mathrm{SL}_d(\mathbb{R})$ for which $A_*\eta = \eta$. Then H is closed, and moreover we have the following dichotomy:*

(a) *If H is compact, then there is an inner product $\langle \cdot, \cdot \rangle'$ on $\mathbb{R}^d$, with corresponding norm $\| \cdot \|'$, with respect to which every $A \in H$ is an isometry.*

(b) *If H is noncompact, then there exist distinct, proper, nontrivial linear subspaces $E^1, \dots, E^p \subset \mathbb{R}^d$, $p \geq 1$, with the following properties:*
   - (i) *We have $\eta(\bigcup_i E^i) = 1$.*
   - (ii) *For all $A \in H$, we have $AE^i = E^{\pi(i)}$ for all $1 \leq i \leq p$, where $\pi = \pi_A$ is a permutation on $\{1, \dots, p\}$.*
   - (iii) *For each $1 \leq i \leq p$ there is an inner product $\langle \cdot, \cdot \rangle^i$ on E^i such that for all $A \in H$, we have that $A|_{E^i}$ is conformal with respect to the inner products $\langle \cdot, \cdot \rangle^i, \langle \cdot, \cdot \rangle^{\pi(i)}$, respectively.*

Lemma 4.6 (a) can be found in [12, Proposition 6.7 (ii)], while the argument for Lemma 4.6 (b) is an extension of arguments appearing in [56, proof of Theorem 8.6]. Since Lemma 4.6 is crucial to our approach and contains strictly more information than what the authors can find in the literature, we will provide a proof sketch later on in Section 4.2.

Building off Lemma 4.6, we give below a corresponding classification of the linear cocycles $\mathcal{A}$ preserving the invariant measure family $(\bar{\nu}_z)$ as in (4.1).

Theorem 4.7 (Classification of invariant fiber measure families). *Assume $d \leq 3$, and assume the setting, notation and conclusions of Proposition 4.3 and Lemma 4.4. Let $(\bar{\nu}_z)_{z \in \operatorname{supp} \mu}$ denote the invariant measure family so-obtained. Then one of the following alternatives holds.*

(a) *There is a continuously-varying assignment to each $z \in \operatorname{supp}\mu$ of an inner product $\langle\cdot,\cdot\rangle_z$ on $\mathbb{R}^d$ with the property that for all $(T, A) \in \mathcal{S}$ and $z \in \operatorname{supp}\mu$, we have that $A_z : (\mathbb{R}^d, \langle\cdot,\cdot\rangle_z) \to (\mathbb{R}^d, \langle\cdot,\cdot\rangle_{Tz})$ is an isometry.*

(b) *For some $p \geq 1$, the following holds. There are p measurably-varying assignments to each element $z \in \operatorname{supp}\mu$ of a proper, distinct, nontrivial linear subspace $E_z^i \subsetneq \mathbb{R}^d$, $1 \leq i \leq p$, with the property that for each $z \in \operatorname{supp}\mu$ and $(T, A) \in \mathcal{S}$, we have $A_z E_z^i = E_{Tz}^{\pi(i)}$ for all $1 \leq i \leq p$, where $\pi = \pi_{(T,A)}$ is a permutation on $\{1, \ldots, p\}$. Moreover, $\bar{\nu}_z(\bigcup_{i=1}^p E_z^i) = 1$. Finally, the collection (E_z^i) is* locally continuous up to re-labelling*: for every $z \in \operatorname{supp}\mu$ there is an open neighborhood $U \subset Z$ and a labelling of the subspaces $E_z^i, z \in U \cap \operatorname{supp}\mu$ with the property that $z \mapsto E_z^i$, $z \in U \cap \operatorname{supp}\mu$ is continuously varying.*

The proof of Theorem 4.7 deviates significantly from that in [12, Theorem 6.8], particularly where it is proved that the objects in alternatives (a) and (b) above are continuously varying. See Remark 4.12 below for a discussion of the subtleties involved.

For the remainder of Section 4.2 we will prove Lemma 4.6 and Theorem 4.7.

Proof of Lemma 4.6

We will prove Lemma 4.6 for any value of the dimension d. Let us first dispense with the relatively easier proof of part (a), i.e., the case when $H = H_\eta$ is a compact subgroup of $\mathrm{SL}_d(\mathbb{R})$. If H is compact, then it admits a right-invariant Haar probability measure γ ([54, Proposition 11.4]). That is, γ is a Borel probability measure on H with the property that for any $A \in H$ and Borel $K \subset H$, we have $\gamma(KA) = \gamma(K)$. With $(\cdot,\cdot)$ the standard inner product on $\mathbb{R}^d$, we define $\langle\cdot,\cdot\rangle'$ on $\mathbb{R}^d$ for $v, w \in \mathbb{R}^d$ by

$$\langle v, w\rangle' := \int_H (A'v, A'w)\, \mathrm{d}\gamma(A').$$

Using right-invariance of γ, one easily checks that $\langle Av, Aw\rangle' = \langle v, w\rangle'$ for all $v, w \in \mathbb{R}^d$ and $A \in H$. This completes the proof of Lemma 4.6 in case (a).

Before proceeding to case (b), let us state and prove the following useful claim.

Claim 4.8. *Let $k \geq 1$ and let (M_n) be a sequence of determinant 1 matrices in $M_{k\times k}(\mathbb{R})$ for which $|M_n| \to \infty$ as $n \to \infty$. Then, on refining to a subsequence $(M_{n'})$, there exist proper linear subspaces $V^1, V^2 \subset \mathbb{R}^k$ for which $\operatorname{dist}(M_{n'}v, V^2) \to 0$ as $n' \to \infty$ for all $v \notin V^1$.*

Proof of Claim 4.8. Using the fact that $\det M_n \equiv 1$ for all n, we can, without loss, pass to a subsequence with the property that for some fixed $1 \leq l < k$, we have

$$\frac{\sigma_l(M_n)}{\sigma_{l+1}(M_n)} \to \infty. \tag{4.2}$$

Applying the Singular Value Decomposition to each matrix M_n, let V_n^1 be the unique $(k-l)$-dimensional subspace for which $|M_n|_{V_n^1}| = \sigma_{l+1}(M_n)$, and let V_n^2 be the unique l-dimensional subspace for which $|M_n^{-1}|_{V_n^2}| = (\sigma_l(M_n))^{-1}$. Passing to a further subsequence, we can assume that the subspaces V_n^1, V_n^2 converge to subspaces V^1, V^2, respectively. It now follows from (4.2) that for all $v \notin V^1$, $\lim_{n\to\infty} \operatorname{dist}(M_n v, V^2) = 0$, as desired. ■

We now proceed to case (b), which we prove in a series of lemmas. Assume $H = H_\eta$ is noncompact, and consider the set $\mathscr{G}$ of finite tuples of proper, nontrivial, distinct subspaces $(E^i)_{i=1}^p$ of $\mathbb{R}^d$ for which $\eta(\bigcup_i E^i) = 1$. Applying Claim 4.8 to a sequence $\{M_n\}$ in H with $|M_n| \to \infty$, note that the pair $(V^i)_{i=1}^2$ so-obtained is such a tuple. If $(E^i)_{i=1}^p$ and $(\check{E}^i)_{i=1}^{\check{p}}$ are two such tuples, let us write

$$(E^i) \le (\check{E}^i) \quad \text{if} \bigcup_i E_i \subset \bigcup_i \check{E}_i .$$

Note that $\le$ is a partial order on $\mathscr{G}$. We say that two tuples $(E^i)_{i=1}^p, (\check{E}^j)_{j=1}^{\check{p}}$ in $\mathscr{G}$ are *equivalent up to relabeling* if $p = \check{p}$ and there is some permutation π on $\{1, \dots, p\}$ for which $\check{E}^j = E^{\pi(j)}$ for all $1 \le j \le p$.

Lemma 4.9. *Let η, H_η be as in the setting of Lemma 4.6 and assume H_η is noncompact (case* (b)*). Then there is a unique tuple $(E^i)_{i=1}^p$ (up to relabeling) of distinct, proper and nontrivial linear subspaces of $\mathbb{R}^d$ minimal with respect to the partial order $\le$ on $\mathscr{G}$. This tuple has the property that for each $A \in H_\eta$, there is a permutation $\pi = \pi_A$ of $\{1, \dots, p\}$ for which $AE^i = E^{\pi(i)}$ for all $1 \le i \le p$.*

Lemma 4.9 is straightforward and left to the reader (see Theorem 8.6 in [56] for more detail). The minimal tuple (E^i) therefore satisfies conditions (i)–(ii) in Lemma 4.6. Item (iii) is verified below.

Lemma 4.10. *For each $1 \le i \le p$, there is an inner product $\langle \cdot, \cdot \rangle^i$ on E^i with the property that for each $A \in H$, we have that $A : (E^i, \langle \cdot, \cdot \rangle^i) \to (E^{\pi(i)}, \langle \cdot, \cdot \rangle^{\pi(i)})$, $\pi = \pi_A$, is a conformal mapping.*

Proof. For (ii), form the subgroup $\tilde{H} = \tilde{H}_\eta = \{A \in H_\eta : AE^i = E^i \text{ for all } 1 \le i \le p\}$. As one can check, $\tilde{H} \subset H$ is a closed, normal subgroup of finite index. The quotient group $H/\tilde{H}$ is naturally isomorphic to a subgroup $\mathfrak{S}$ of the group of permutations on p symbols. Let us assume for the moment that $\mathfrak{S}$ acts transitively[13] on $\{1, \dots, p\}$; we will remove this restriction at the end of the proof.

Fix an arbitrary $i \in \{1, \dots, p\}$ and form

$$\check{H}^{(i)} = \{(\det(A|_{E^i}))^{-\frac{1}{\dim E^i}} A|_{E^i} : A \in \tilde{H}\} .$$

Note that linear operators in $\check{H}^{(i)}$ preserve the measure $\eta|_{E^i}$. Since any $A \in \tilde{H}$ maps E^i into itself, we can think of $\check{H}^{(i)}$ as a subgroup of $\mathrm{SL}_{\dim E^i}(\mathbb{R})$ on identifying E^i with $\mathbb{R}^{\dim E^i}$. We claim that $\check{H}^{(i)}$ is compact. If not, then by Claim 4.8 there are proper linear subspaces $\check{V}^1, \check{V}^2 \subset E^i$ for which $\eta(\check{V}^1 \cup \check{V}^2) = \eta(E^i)$. This contradicts minimality of $(E^i)_{i=1}^p$ as in item (i). Thus $\check{H}^{(i)}$ is compact; it now follows from Lemma 4.6 (a) that

[13]Let $S \subset \{1, \dots, p\}$ and assume $\mathfrak{S}S = S$. We say that $\mathfrak{S}$ acts *transitively* on S if for all $i, j \in S$ there is some $\pi \in \mathfrak{S}$ for which $\pi(i) = j$.

there exists an inner product $\langle\cdot,\cdot\rangle^i$ on E^i with respect to which $\check{H}^{(i)}$ acts isometrically. Equivalently, linear operators of the form $A|_{E^i}$, $A \in \tilde{H}$ act conformally with respect to $\langle\cdot,\cdot\rangle^i$.

We now define $\langle\cdot,\cdot\rangle^j$, $1 \le j \le p$, $j \ne i$, as follows: for each such j, fix an $M \in H$ for which $ME^i = E^j$ (such an M exists since $\mathfrak{S} = H/\tilde{H}$ acts transitively on $\{1,\dots,p\}$ by assumption) and define

$$\langle v, w\rangle^j = \langle M^{-1}v, M^{-1}w\rangle^i, \quad v, w \in E^j. \tag{4.3}$$

This definition is independent of M: if $M'E^i = E^j$ for some other $M' \in H$, then we have

$$\langle (M')^{-1}v, (M')^{-1}w\rangle^i = \langle M^{-1}v, M^{-1}w\rangle^i = \langle v, w\rangle^j$$

for all $v, w \in E^j$. By a similar computation, one checks that if $A \in H$ maps $AE^i = E^j$, then A is conformal with respect to the inner products $\langle\cdot,\cdot\rangle^i$, $\langle\cdot,\cdot\rangle^j$, respectively. This completes the proof when $\mathfrak{S} \cong H/\tilde{H}$ acts transitively on $\{1,\dots,p\}$.

Let us now address the situation when $\mathfrak{S}$ does not act transitively on $\{1,\dots,p\}$. In this case, by a standard argument there is a unique partition of $\{1,\dots,p\}$ into disjoint sets $\mathcal{P}_l$, $1 \le l \le k$, such that for each partition atom $\mathcal{P}_l$, we have (1) $\mathfrak{S}\mathcal{P}_l = \mathcal{P}_l$, and (2) $\mathfrak{S}$ acts transitively on $\mathcal{P}_l$. For each $\mathcal{P}_l$, repeat the construction of $\langle\cdot,\cdot\rangle^i$ for some fixed arbitrary $i \in \mathcal{P}_l$, and then define $\langle\cdot,\cdot\rangle^j$, $j \in \mathcal{P}_l$, $j \ne i$, as in (4.3) for some arbitrary $M \in H$ sending $ME^i = E^j$ (such an M exists since $\mathfrak{S}$ acts transitively on $\mathcal{P}_l$ by construction). Lemma 4.10 now follows from the previous arguments, since for all $A \in H$, we can have $AE^i = E^j$ only if i, j belong to the same $\mathcal{P}_l$ for some $1 \le l \le k$. ∎

Proof of Theorem 4.7. We first give the following preliminary lemma. For $z \in \operatorname{supp}\mu$, define

$$O_z = \{Tz : (T, A) \in \mathcal{S}\}.$$

Note that $O_z \subset \operatorname{supp}\mu$ holds by Lemma 4.4. Using ergodicity of μ and the strong Feller property, we get the following.

Lemma 4.11. *For all $z \in \operatorname{supp}\mu$, we have $\mu(O_z) = 1$.*

Proof. First, let us check that O_z is a (P_t, μ)-invariant set in the sense of Definition 3.9. Fix $t > 0$ and let $y \in O_z$. Then $y = Tz$ for some $(T, A) \in \mathcal{S}$. Now, fix a **P**-generic $\omega \in \Omega$ and set $T' = T^t_\omega$, $A' = \mathcal{A}^t_\omega$. Noting $(T', A') \in \mathcal{S}$ with probability 1, we see that $T'y = T' \circ Tz$, hence $T'y \in O_z$. Since $y \in O_z$ was arbitrary, we conclude $\mathcal{T}^t_\omega y \in O_z$ for any $t \ge 0$ with probability 1, hence O_z is (P_t, μ)-invariant.

It follows from ergodicity for μ (Definition 3.9) that O_z has zero or full μ-measure. To check $\mu(O_z) > 0$, assume otherwise and observe that by stationarity, $P_t(y, O_z) = 0$ for μ-almost all $y \in Z$. From the ultra-Feller property for the semigroup (P_t) as in Lemma 4.2, we conclude $P_t(z, O_z) = 0$, a contradiction (note $\{y \in Z : P_t(y, O_z) = 0\}$ must be dense in $\operatorname{supp}\mu$). We conclude $\mu(O_z) > 0$, hence $\mu(O_z) = 1$. ∎

Proof of Theorem 4.7. Fix $z_0 \in \operatorname{supp}\mu$, thought of as a reference point, and consider the $\mathrm{SL}_2(\mathbb{R})$ subgroup

$$H_{z_0} = \{A \in \mathrm{SL}_2(\mathbb{R}) : A_*\bar{\nu}_{z_0} = \bar{\nu}_{z_0}\}.$$

Note that H_{z_0} is closed by Lemma 4.6. We claim that if H_{z_0} is compact we are in case (a), while if H_{z_0} is noncompact then we are in case (b). Crucially, this distinction does not depend on the choice of reference point $z_0 \in Z$; see Remark 4.13 below for a discussion of this point.

Case (a)*: H_{z_0} is compact.* By Lemma 4.6 there is an inner product $\langle \cdot, \cdot \rangle_{z_0}$ with respect to which all matrices in H_{z_0} act as isometries. We define the family $\{\langle \cdot, \cdot \rangle_z\}_{z \in \operatorname{supp}\mu}$ as follows. For each $z \in \operatorname{supp}\mu$, fix $y \in O_{z_0} \cap O_z$ (such a point exists since $\mu(O_{z_0} \cap O_z) = 1$ by Lemma 4.11) and let $(T, A), (T', A') \in \mathcal{S}$ be such that $Tz_0 = y, T'z = y$.

For $v, w \in \mathbb{R}^d$ we define

$$\langle v, w \rangle_z = \langle A_{z_0}^{-1} \circ A'_z v, A_{z_0}^{-1} \circ A'_z w \rangle_{z_0} .$$

Let us check this definition does not depend on the exact choice of $(T, A), (T', A')$. If $(\bar{T}, \bar{A}), (\bar{T}', \bar{A}') \in \mathcal{S}$ are any other elements for which $\bar{T}z_0 = y, \bar{T}'z = y$, then Lemma 4.4 implies

$$(\bar{A}_{z_0})^{-1} \bar{A}'_z (A'_z)^{-1} A_{z_0} \in H_{z_0},$$

and so

$$\langle A_{z_0}^{-1} \circ A'_z v, A_{z_0}^{-1} \circ A'_z w \rangle_{z_0} = \langle \bar{A}_{z_0}^{-1} \circ \bar{A}'_z v, \bar{A}_{z_0}^{-1} \circ \bar{A}'_z w \rangle_{z_0}$$

holds by Lemma 4.6 (a). By a similar proof, one checks that for each $(T, A) \in \mathcal{S}$ and $z \in \operatorname{supp}\mu$, we have that $A_z : (\mathbb{R}^d, \langle \cdot, \cdot \rangle_z) \to (\mathbb{R}^d, \langle \cdot, \cdot \rangle_{Tz})$ is an isometry.

To prove continuity of $z \mapsto \langle \cdot, \cdot \rangle_z$, we do the following. For each $z \in \operatorname{supp}\mu$, the inner product $\langle \cdot, \cdot \rangle_z$ gives rise to a Euclidean volume on $\mathbb{R}^d$ and an induced volume $\tilde{\nu}_z$ on P^{d-1}. By the isometry property, it follows that for all $(T, A) \in \mathcal{S}$, we have that $(A_z)_* \tilde{\nu}_z = \tilde{\nu}_{Tz}$ for all $z \in \operatorname{supp}\mu$. Thus $(\tilde{\nu}_z)_{z \in \operatorname{supp}\mu}$ defines an invariant measure family on $\operatorname{supp}\mu$. Repeating the proof of Proposition 4.3 for this new invariant measure family, we conclude $(\tilde{\nu}_y)$ is continuously varying in the weak* topology.

From the weak* continuity of $z \mapsto \tilde{\nu}_z$ and the fact that $\tilde{\nu}_z \ll \mathrm{Leb}_{P^{d-1}}$ for all z, we conclude that the densities

$$\rho_z := \frac{d\tilde{\nu}_z}{d\,\mathrm{Leb}_{P^{d-1}}}, \quad z \mapsto \rho_z : P^{d-1} \to \mathbb{R},$$

vary continuously in the uniform norm on $C(P^{d-1}, \mathbb{R})$. It is now straightforward to check that the corresponding inner products $z \mapsto \langle \cdot, \cdot \rangle_z$ vary continuously.

Remark 4.12. It is a subtle point in the proof of Theorem 4.7 (a) above that the original invariant measure family $(\bar{\nu}_z)_{z \in \operatorname{supp}\mu}$ need not coincide with the measure family $(\tilde{\nu}_z)_{z \in \operatorname{supp}\mu}$. Indeed, we do not rule out the possibility that the $(\bar{\nu}_z)$ consist of some combination of atomic, singular continuous and absolutely continuous measures. As such, one cannot deduce continuity of the resulting inner products $\langle \cdot, \cdot \rangle_z$, $z \in \operatorname{supp}\mu$ directly from the $(\bar{\nu}_z)$. As we will see below, the proof of Theorem 4.7(b) has a similar complication which must be addressed.

By comparison, [12, Theorem 6.8] avoids this subtlety for two reasons:

(1) in that framework, under a nondegeneracy condition it follows that the fiber measures $\bar{\nu}_z$ are automatically absolutely continuous with respect to the volume on P^{d-1},

(2) [12, Theorem 6.8] invokes an additional hypothesis that we are not able to justify either at the level of generality of Theorem 4.7 or for the Lagrangian flow corresponding to the infinite-dimensional Systems 3 and 4.

Case (b)*: H_{z_0} is noncompact.* Let $E^i_{z_0} = E^i$, $1 \le i \le p$, be as in Lemma 4.6 (b) applied to $H = H_{z_0}$. For each i, let $\langle \cdot , \cdot \rangle^i$ denote the corresponding inner product on $E^i = E^i_{z_0}$. For $z \in \operatorname{supp} \mu$ we define E^i_z as follows. Fix $y \in O_{z_0} \cap O_z$, as in the proof for case (a), and fix $(T, A), (T', A')$ for which $Tz_0 = y, T'z = y$. We define

$$E^i_z = (A'_z)^{-1} \circ A_{z_0}(E^i_{z_0}).$$

We also define the inner products $\langle \cdot , \cdot \rangle^i_z$ on E^i_z by setting, for $v, w \in E^i_z$,

$$\langle v, w\rangle^i_z = \langle (A_{z_0})^{-1} A'_z v, (A_{z_0})^{-1} A'_z w\rangle^i .$$

As in the proof of case (a), one checks that the above definitions do not depend on the exact choices of $y \in O_z \cap O_{z_0}$ or $(T, A), (T', A') \in \mathcal{S}$. By a similar check, the invariance property for the $E^i_z, z \in \operatorname{supp} \mu$ similarly holds, and moreover, for $(T, A) \in \mathcal{S}$ and $z \in \operatorname{supp} \mu$, we have that $A_z : (E^i_z, \langle \cdot , \cdot \rangle^i_z) \to (E^{\pi(i)}_{Tz}, \langle \cdot , \cdot \rangle^{\pi(i)}_{Tz})$ is conformal.

Let us now prove the continuity statement. Observe that since $d \le 3$, there are two cases: either $\dim E^i_z \equiv 1$ for all i or $\dim E^i_z = 2$ for some i, z. If the former, local continuity of $z \mapsto E^i_z$ up to relabeling follows immediately from the fact that $\bar{\nu}_z|_{E^i_z}$ is a delta mass supported on the projective point corresponding to E^i_z. If the latter, then by Claim 4.8 we must have that $p \le 2$ and that at most one of the E^i_z is two-dimensional for each $z \in \operatorname{supp} \mu$. We focus on the case $p = 1$; essentially the same proof applies when $p = 2$. Hereafter let us write $E_z := E^1_z$. Note that this can only occur when $d = 3$, which hereafter we assume.

In analogy with the proof of Theorem 4.7 (a), consider for each element $z \in \operatorname{supp} \mu$, the Euclidean volume m_z on $E_z \subset \mathbb{R}^3$ induced by the inner product $\langle \cdot , \cdot \rangle_z := \langle \cdot , \cdot \rangle^1_z$. This induces a normalized volume $\tilde{\nu}_z$ on the projectivization of E_z in P^2. As in the proof for case (a), the fiber measure family $(\tilde{\nu}_z)_{z \in \operatorname{supp} \mu}$ is invariant as in (3.8). This follows from the conformality property for the inner product $\langle \cdot , \cdot \rangle_z$. As in the proof of case (a), we can repeat the arguments of Proposition 4.3, from which we obtain that the family $(\tilde{\nu}_z)$ is weak* continuous. Continuity of $z \mapsto E_z$ now follows. ■

We conclude Section 4.2 with several remarks.

Remark 4.13. The determination between case (a) and (b) made at the beginning of the proof of Theorem 4.7 does not depend on the reference point $z_0 \in \operatorname{supp} \mu$. Indeed, given $z, z' \in \operatorname{supp} \mu$ one can obtain a group isomorphism $H_z \to H_{z'}$ as follows: fix $y \in O_z \cap O_{z'}$ and let $(T, A), (T', A') \in \mathcal{S}$ be such that $Tz = y, T'z = y$. Then the mapping $H_z \to H_{z'}$ given by $H_z \ni M \mapsto (A'_{z'})^{-1} A_z M A_z^{-1} A'_{z'} \in H_{z'}$ is an isomorphism from H_z to $H_{z'}$.

Remark 4.14. The restriction to $d \le 3$ is only relevant in case (b) of Theorem 4.7. For $d \ge 4$ the result is likely to be true, but the proof is lengthier due to the fact that among the E^i_z, $z \in \operatorname{supp} \mu$, $1 \le i \le p$, there may be arbitrarily many subspaces of dimension ≥ 2.

Thus, the trick applied in case (b) above must be applied to the projectivization of the Euclidean volume on each E_z^i separately, and continuity derived in this way. Since the case $d \leq 3$ suits the purpose of our main application in this paper, we leave off the $d \geq 4$ case to a future work.

Remark 4.15. Let us summarize the differences between [12, Theorem 6.8] and the analogue pursued here in Section 4.2. To start, [12, Theorem 6.8] proves the classification in Theorem 4.7 above in the special case when Z is a locally compact Riemannian manifold, $\mathcal{T}$ is the stochastic flow of diffeomorphisms generated by a hypoelliptic SDE satisfying suitable nondegeneracy properties, and $\mathcal{A}$ is its corresponding derivative cocycle.

In comparison, Theorem 4.7 does not require that $\mathcal{A}$ be the derivative cocycle of $\mathcal{T}$. This requires that we work with the product space $\mathcal{C}$ of pairs of mappings and cocycles, as is done in Lemma 4.4. Moreover, and arguably of greater consequence, is the fact that the base RDS $\mathcal{T}$ is not necessarily invertible, nor is its phase space Z locally compact. These differences are emblematic of dynamics on infinite-dimensional spaces and are exemplified by our intended application to the Navier–Stokes equations and more generally to regularizing semilinear parabolic problems. This raises numerous issues which we have dealt with over the course of Section 3, e.g., the definition of the topology on observables with respect to which (P^t) is a C^0-semigroup (Proposition 3.5).

Finally, [12, Theorem 6.8], of which the main result Theorem 4.7 is an analogue, invokes an additional hypothesis to get continuity of the obtained invariant inner products in case (a) (resp., finite union of proper linear subspaces in case (b)). This additional hypothesis is not accessible in our setting. This brings up a significant subtlety (Remark 4.12), unique to our setting, which our argument addresses.

4.3. Sufficient condition for $\lambda_1 > 0$: Approximate controllability criteria

We will now state a weaker version of the criterion (C) in Section 3.3 which can be used to rule out the alternatives (a) and (b) in Theorem 4.7.

Definition 4.16. We say that the cocycle $\mathcal{A}$ satisfies the *approximate controllability condition* (C') if there exist $z, z' \in \operatorname{supp} \mu$ such that z' belongs to the support of the measure $P_{t_0}(z, \cdot)$ for some $t_0 > 0$, and we have each of the following.

(a) We have $Q_{t_0}((z, \mathrm{Id}), B_\varepsilon(z') \times \{A \in \mathrm{SL}_d(\mathbb{R}) : |A| > M\}) > 0$ for any $\varepsilon, M > 0$.

(b) For any $v \in P^{d-1}$, open $V \subset P^{d-1}$ and $\varepsilon > 0$, we have $\hat{P}_{t_0}((z, v), B_\varepsilon(z') \times V) > 0$.

We can now prove the following.

Proposition 4.17. *Let $d \leq 3$. Let $\mathcal{A}$ be an $\mathrm{SL}_d(\mathbb{R})$ linear cocycle as in Section 3.2.1 over a continuous RDS $\mathcal{T}$ as in Section 3.1.1 satisfying (H1)–(H3) for which the Markov semigroup (P_t) has the strong Feller property. Let μ be an ergodic stationary measure for which the approximate controllability condition (C') holds. Then, $\lambda^+ > \lambda^-$, and in particular $\lambda_1 > 0$, in the MET (Theorem 3.13).*

Proof. If $\lambda^+ = \lambda^-$, then Theorem 3.18 applies, and so either case (a) or case (b) holds in Theorem 4.7.

We start by ruling out (a). For $y \in \operatorname{supp} \mu$, write $|\cdot|_y$ for the norm corresponding to the inner product $\langle \cdot, \cdot \rangle_y$. Let

$$\kappa = \max\left\{ \max_{v \in \mathbb{R}^d \setminus \{0\}} \frac{|v|_z}{|v|}, \max_{v \in \mathbb{R}^d \setminus \{0\}} \frac{|v|_{z'}}{|v|} \right\}.$$

Fix $\varepsilon > 0$ so that $\frac{1}{2}|\cdot|_y \le |\cdot|_{z'} \le 2|\cdot|_y$ for all $y \in B_\varepsilon(z')$.

Now, condition (C') (a) says that there is a $\mathbf{P}$-positive measure set $E \subset \Omega$ such that $\mathcal{T}_\omega^{t_0} z \in B_\varepsilon(z')$ and $|\mathcal{A}_{\omega,z}^{t_0}| > 2\kappa^2$ for all $\omega \in E$. Without loss we can assume

$$\{(\mathcal{T}_\omega^{t_0}, \mathcal{A}_{\omega,\cdot}^{t_0}) : \omega \in E\} \subset \mathcal{S}_{t_0},$$

perhaps on paring off an $\mathbf{P}$-measure zero set from E. By Theorem 4.7 (a), for all $\omega \in E$ we deduce

$$|\mathcal{A}_{\omega,z}^{t_0}|_{z,y} = 1,$$

where $|\cdot|_{z,y}$ is the matrix norm induced by the norms $|\cdot|_z$ at z and $|\cdot|_y$ at $y = \mathcal{T}_\omega^{t_0} z$. From this we obtain the estimate $|A_{\omega,z}^{t_0}| \le 2\kappa^2$ in the matrix norm induced from $|\cdot|$. This is a contradiction to (C') (a).

Turning to case (b), take $\varepsilon > 0$ sufficiently small so that

(i) a labelling of the E_y^i, $y \in B_\varepsilon(z')$ exists for which the mapping $y \mapsto E_y^i$ is continuous for $1 \le i \le p$, and

(ii) there is an open set $V \subset P^{d-1}$ for which $V \cap (\bigcup_i E_y^i) = \emptyset$ for all $y \in B_\varepsilon(z')$.

Fix an arbitrary $1 \le i \le p$ and $v \in E_z^i \setminus (\bigcup_{j \ne i} E_z^j)$. Condition (C') (b) implies that there is a $\mathbf{P}$-positive measure set $E \subset \Omega$ such that for all $\omega \in E$, we have $\mathcal{T}_\omega^{t_0} z \in B_\varepsilon(z')$ and $\mathcal{A}_{\omega,z}^{t_0} v \in V$. As before, on paring off a $\mathbf{P}$-measure zero set we can assume $(\mathcal{T}_\omega^{t_0}, \mathcal{A}_\omega^{t_0}) \in \mathcal{S}_{t_0}$ for all $\omega \in E$, from which we deduce (Theorem 4.7 (b)) that

$$\mathcal{A}_{\omega,z}^{t_0} E_z^j = E_{\mathcal{T}_\omega^{t_0} z}^{\pi_\omega(j)}$$

for all $\omega \in E$ and $1 \le j \le p$, where π_ω is some permutation on $\{1, \dots, p\}$. But at $j = i$ this is a contradiction, since $v \in E_z^i$ yet $\mathcal{A}_{\omega,z}^{t_0} v \notin E_{\mathcal{T}_\omega^{t_0} z}^l$ for any $l \in \{1, \dots, p\}$ by construction. ∎

5. Lie brackets and Hörmander's condition

The main goal of this section is to explore how noise in the low modes of a fluid model spreads to other variables coupled to the flow. Specifically, for (u_t) given by Systems 1 and 2, we will show that the projective processes (u_t, x_t, v_t), $(u_t, x_t, \check{v}_t)$, and the matrix process (u_t, x_t, A_t) are all generated by vector fields satisfying the parabolic Hörmander condition in both 2 and 3 dimensions (Definition 5.1). Using the a priori estimates on (u_t) and that $\mathbb{T}^d \times P^{d-1}$ is compact, Hörmander's theorem (see, e.g., [67, 68] and [37, 63]) then implies (u_t, x_t), (u_t, x_t, v_t), $(u_t, x_t, \check{v}_t)$ have absolutely continuous Markov kernels (with respect to Lebesgue measures) and unique stationary measures. Similarly, (u_t, x_t, A_t) also has an absolutely continuous Markov kernel and therefore the arguments given in Section 2.3 are validated. Theorem 1.6 hence follows for Systems 1 and 2.

In what follows it is technically more convenient to deal with the space $\mathbb{S}^{d-1}$ in place of P^{d-1} while still denoting v_t and $\check{v}_t$ the corresponding versions in $\mathbb{S}^{d-1}$. Since P^{d-1} and $\mathbb{S}^{d-1}$ are locally diffeomorphic, proving Hörmander's condition on $\mathbb{S}^{d-1}$ implies Hörmander's condition for P^{d-1}.

5.1. Preliminaries

Recall the orthogonal $L^2(\mathbb{T}^d)$ basis $\{e_k\}_{k\in\mathbb{Z}_0^d}$ and the family of $d\times(d-1)$ matrices $\{\gamma_k\}_{k\in\mathbb{Z}_0^d}$ introduced in Section 1.1.1 satisfying

$$\gamma_k^\top k = 0 \quad \text{and} \quad \gamma_k^\top \gamma_k = \mathrm{Id}.$$

We will denote for each $k\in\mathbb{Z}_0^d$ the column vectors $\{\gamma_k^1,\dots,\gamma_k^{d-1}\}$ of the matrix γ_k. These vectors consequently form an orthonormal basis for the subspace of vectors in $\mathbb{R}^d$ perpendicular to k. Note that for each $k\in\mathbb{Z}_0^d$ and $i\in\{1,\dots,d-1\}$, $e_k\gamma_k^i$ is a divergence-free, mean-zero vector field on $\mathbb{T}^d$ and the collection

$$\{e_k\gamma_k^i : k\in\mathbb{Z}_0^d,\ i=\{1,\dots,d-1\}\}$$

forms an orthogonal basis for $\mathbf{L}^2$ with respect to the inner product

$$\langle u^1,u^2\rangle_{\mathbf{L}^2} = \int_{\mathbb{T}^d} u^1(x)\cdot u^2(x)\,\mathrm{d}x.$$

This means that given a $u\in\mathbf{L}^2$, we can write

$$u = \sum_{i=1}^{d-1}\sum_{k\in\mathbb{Z}_0^d}(u)_k^i e_k\gamma_k^i, \quad \text{where } (u)_k^i = \frac{1}{\pi(2\pi)^{d-1}}\langle u, e_k\gamma_k^i\rangle_{\mathbf{L}^2}.$$

It follows that, given (u_t) solving any of Systems 1 or 2, we can write the equations for (x_t,v_t) in $\mathbb{T}^d\times\mathbb{S}^{d-1}$ as

$$\frac{\mathrm{d}}{\mathrm{d}t}x_t = \sum_{i=1}^{d-1}\sum_{k\in\mathbb{Z}_0^d}(u_t)_k^i e_k(x_t)\gamma_k^i, \tag{5.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}v_t = \sum_{i=1}^{d-1}\sum_{k\in\mathbb{Z}_0^d}(u_t)_k^i (k\cdot v_t)e_{-k}(x_t)\Pi_{v_t}\gamma_k^i. \tag{5.2}$$

Likewise the inverse transpose projective process $(\check{v}_t)$ in $\mathbb{S}^{d-1}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\check{v}_t = -\sum_{i=1}^{d-1}\sum_{k\in\mathbb{Z}_0^d}(u_t)_k^i (\gamma_k^i\cdot\check{v}_t)e_{-k}(x_t)\Pi_{\check{v}_t}k,$$

and the matrix process (A_t) in $\mathrm{SL}^d(\mathbb{R})$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}A_t = \sum_{i=1}^{d-1}\sum_{k\in\mathbb{Z}_0^d}(u_t)_k^i e_{-k}(x)(\gamma_k^i\otimes k)A_t. \tag{5.3}$$

We are interested in studying the hypoellipticity of the processes (u_t, x_t, v_t) and (u_t, x_t, A_t) , when (u_t) is governed by System 1 or 2. Recall that Systems 1 and 2 both live in a finite-dimensional subspace $\hat{\mathbf{H}}$ of $\mathbf{H}$ (see Section 1.3). In both cases the process of interest will take the form of an abstract degenerate SDE

$$\mathrm{d}y_t = X_0(y_t)\,\mathrm{d}t + \sum_{j=1}^{M} X_j \mathrm{d}W_t^j \tag{5.4}$$

on $\hat{\mathbf{H}} \times \mathcal{M}$, where $\mathcal{M}$ is a finite-dimensional Riemannian manifold (either $\mathbb{T}^d \times \mathbb{S}^{d-1}$ or $\mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$). Here X_0 is a vector field on $\hat{\mathbf{H}} \times \mathcal{M}$ associated to the drift, while $\{X_j\}_{j=1}^M$ is an enumeration of the vectors $\{q_k e_k \gamma_k^i : k \in \mathcal{K},\ i = 1, \ldots, d-1\}$ in $\hat{\mathbf{H}}$.

Recall the *Lie bracket* (or commutator) of two vector fields X and Y on a smooth manifold $\mathcal{Y}$ is defined for each $y \in \mathcal{Y}$ by

$$[X, Y](y) = D_X Y(y) - D_Y X(y),$$

where D_X and D_Y denote the directional derivatives in the direction X and Y respectively. The Hörmander condition is now stated as follows:

Definition 5.1 (Parabolic Hörmander Condition). A family of vector fields $\{X_k\}_{k=0}^M$ on a smooth manifold $\mathcal{M}$ is said to satisfy the *parabolic Hörmander condition* if for each $y \in \mathcal{Y}$ the vectors

$$\begin{array}{lll} X_k(y), & k = 1, \ldots, M & \\ [X_k, X_j](y), & k = 1, \ldots, M, & j = 0, \ldots, M, \\ \big[X_k, [X_j, X_\ell]\big](y), & k = 1, \ldots, M, & j, \ell = 0, \ldots, M, \\ \vdots & \vdots & \end{array}$$

span $T_y\mathcal{M}$.

Theorem 5.2 ([67]; see also [37, 63, 68]). *Let $P_t(y, A) = \mathbf{P}(y_t \in A \mid y_0 = y)$ be the Markov kernel associated to the finite-dimensional SDE (5.4). If Definition 5.1 is satisfied, then $P_t(y, \cdot)$ is absolutely continuous with respect to* $\mathrm{Leb}_{\hat{\mathbf{H}} \times \mathcal{M}}$.

5.2. *Lie brackets for the projective process*

In this subsection we study the spanning properties of Lie brackets for the process (x_t, v_t) in $\mathbb{T}^d \times \mathbb{S}^{d-1}$. The equations (5.1) and (5.2) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_t \\ v_t \end{pmatrix} = V(u_t, x_t, v_t),$$

where $V(u, x, v)$ is the vector field defined for each $(u, x, v) \in \mathbf{H} \times \mathbb{T}^d \times \mathbb{S}^{d-1}$ by

$$V(u, x, v) = \sum_{i=1}^{d-1} \sum_{k \in \mathbb{Z}_0^d} \begin{pmatrix} (u)_k^i e_k(x) \gamma_k^i \\ (u)_k^i (k \cdot v) e_{-k}(x) (\Pi_v \gamma_k^i) \end{pmatrix} \in T_x\mathbb{T}^d \times T_v\mathbb{S}^{d-1}.$$

Note that $V(u, x, v)$ is linear in u and therefore the Lie-bracket $[e_k \gamma_k^i, V]$ does not depend on u and is readily seen to be given by

$$[e_k \gamma_k^i, V](x, v) = \begin{pmatrix} e_k(x)\gamma_k^i \\ (k \cdot v)e_{-k}(x)(\Pi_v \gamma_k^i) \end{pmatrix}.$$

The following lemma gives sufficient conditions for $[e_k \gamma_k^i, V]$ to span $T_x\mathbb{T}^d \times T_v\mathbb{S}^{d-1}$.

Lemma 5.3. *Let $k^1, \dots, k^d$ be d linearly independent elements of $\mathbb{Z}_0^d$ and define*

$$K = \{k^1, \dots, k^d\} \cup \{-k^1, \dots, -k^d\} \subseteq \mathbb{Z}_0^d.$$

Then at each point $(x, v) \in \mathbb{T}^d \times \mathbb{S}^{d-1}$, we have

$$\operatorname{span}\{[e_k \gamma_k^i, V](x, v) : k \in K,\ i = 1, \dots, d-1\} = T_x\mathbb{T}^d \times T_v\mathbb{S}^{d-1}.$$

Proof. Let $k \in K$. Using the identity $e_k^2 + e_{-k}^2 = 1$ and the fact that $-k \in K$, we find that for each $(x, v) \in \mathbb{T}^d \times \mathbb{S}^{d-1}$ (recall the symmetry $\gamma_{-k} = -\gamma_k$)

$$e_k(x)[e_k \gamma_k^i, V](x, v) - e_{-k}(x)[e_{-k}\gamma_{-k}^i, V](x, v) = \begin{pmatrix} \gamma_k^i \\ 0 \end{pmatrix}$$

and

$$e_{-k}(x)[e_k \gamma_k^i, V](x, v) + e_k(x)[e_{-k}\gamma_{-k}^i, V](x, v) = \begin{pmatrix} 0 \\ (k \cdot v)(\Pi_v \gamma_k^i) \end{pmatrix}.$$

Therefore it suffices to show that

$$\operatorname{span}\{\gamma_k^i : k \in K,\ i \in \{1, \dots, d-1\}\} = \mathbb{R}^d, \tag{5.5}$$

and for each $v \in \mathbb{S}^{d-1}$

$$\operatorname{span}\{(k \cdot v)(\Pi_v \gamma_k^i) : k \in K,\ i \in \{1, \dots, d-1\}\} = T_v\mathbb{S}^{d-1}. \tag{5.6}$$

Condition (5.5) follows from the linear independence of k^1 and k^2 and the fact that $\{\gamma_k^i\}_{i=1}^{d-1}$ spans the space perpendicular to k. Condition (5.6) follows from the fact that by linear independence of $k^1, \dots, k^d$, that for each $v \in \mathbb{S}^{d-1}$, there exists a $k \in K$ such that $v \cdot k \neq 0$ and therefore, since $\{\gamma_k^i\}_{i=1}^{d-1}$ spans the space perpendicular to k, the vectors $\{\Pi_v \gamma_k^i\}_{i=1}^{d-1}$ span $T_v\mathbb{S}^{d-1}$. ∎

Remark 5.4. As discussed in Section 2.3, for System 1, this is the only place where Assumption 1 is used.

Remark 5.5. It is not difficult to see that we may replace v_t with $\check{v}_t$ in the above lemma, without changing the proof much. The only difference being that condition (5.6) is now replaced with

$$\operatorname{span}\{(\gamma_k^i \cdot v)(\Pi_v k) : k \in K,\ i \in \{1, \dots, d-1\}\} = T_v\mathbb{S}^{d-1},$$

which can be deduced from the fact that by linear independence of $k_1, \dots, k^d$, there exists at least $d-1$ linearly independent elements $\hat{k}^1, \dots, \hat{k}^{d-1}$ of K such $\gamma_{\hat{k}^j}^i \cdot v \neq 0$ for some $i = 1, \dots, d-1$ and such that the sets $\{\Pi_v \hat{k}^j : j = 1, \dots, d-1\}$ spans $T_v\mathbb{S}^{d-1}$.

5.3. Lie brackets for the matrix process

We would also like to study the spanning properties of Lie brackets for the process (x_t, A_t) in $\mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$. Similarly to the (x_t, v_t) process, equations (5.1) and (5.3) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_t \\ A_t \end{pmatrix} = G(u_t, x_t, A_t),$$

where for each $(u, x, A) \in \mathbf{H} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$,

$$G(u, x, A) = \sum_{i=1}^{d-1} \sum_{k \in \mathbb{Z}_0^d} \begin{pmatrix} (u)_k^i e_k(x) \gamma_k^i \\ (u)_k^i e_{-k}(x)(\gamma_k^i \otimes k)A \end{pmatrix} \in T_x \mathbb{T}^d \times T_A \,\mathrm{SL}_d(\mathbb{R}).$$

Again, $G(u, x, A)$ is linear in u and so the Lie-bracket $[e_k \gamma_k^i, G]$ does not depend on u.

Lemma 5.6. *Let $k^1, \dots, k^{d+1}$ be $d+1$ elements of $\mathbb{Z}_0^d$ given by $k^1 = (0, 1)$, $k^2 = (1, 0)$, $k^3 = (1, 1)$ for $d = 2$ and $k^1 = (0, 0, 1)$, $k_2 = (0, 1, 0)$, $k^3 = (0, 0, 1)$, $k^4 = (1, 1, 1)$ for $d = 3$. Define $K = \{k^1, \dots, k^{d+1}\} \cup \{-k^1, \dots, -k^{d+1}\} \subseteq \mathbb{Z}_0^d$. Then at each point $(x, A) \in \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$, we have*

$$\mathrm{span}\{[e_k \gamma_k^i, G](x, A) : k \in K,\ i \in \{1, \dots, d-1\}\} = T_x \mathbb{T}^d \times T_A \,\mathrm{SL}_d(\mathbb{R}).$$

Proof. Following the same proof strategy as in the proof of Lemma 5.3, we may conclude that it suffices to show that

$$\mathrm{span}\{(\gamma_k^i \otimes k)A : k \in K,\ i \in \{1, \dots, d-1\}\} = T_A \,\mathrm{SL}_d(\mathbb{R}).$$

Using that the Lie algebra $\mathrm{sl}_d(\mathbb{R})$ of traceless $d \times d$ matrices is linearly isomorphic to $T_A \mathrm{SL}_d(\mathbb{R})$ by right (or left) multiplication by A, the above spanning condition is equivalent to showing that

$$\mathrm{span}\{(\gamma_k^i \otimes k) : k \in K,\ i \in \{1, \dots, d-1\}\} = \mathrm{sl}_d(\mathbb{R}). \tag{5.7}$$

Condition (5.7) follows from the fact that for the vectors $k^1, \dots, k^{d+1}$ given, the $d^2 - 1$ matrices $\{(\gamma_k^i \otimes k) : k = \{k^1, \dots, k^{d+1}\},\ i \in \{1, \dots, d-1\}\}$ are all linearly independent in $\mathrm{sl}_d(\mathbb{R})$. Since $\mathrm{sl}_d(\mathbb{R})$ is $(d^2 - 1)$-dimensional, condition (5.7) must hold. ∎

5.4. Hörmander condition for Stokes and Galerkin–Navier–Stokes systems

We now turn to study the hypoellipticity of the projective process (u_t, x_t, v_t) and matrix process (u_t, x_t, A_t) when (u_t) satisfies either Systems 1 or 2. We will define the vector field U^S on $\mathbf{H}_{\mathcal{K}}$ associated with the Stokes System 1 by

$$U^S(u) := -\sum_{i=1}^{d-1} \sum_{k \in \mathcal{K}} |k|^2 (u)_k^i e_k \gamma_k^i,$$

and the vector field U^{NS} on $\mathbf{H}_N$ associated with the Galerkin–Navier–Stokes System 2 by

$$U^{NS}(u) := -\sum_{i=1}^{d-1} \sum_{|k|_\infty \le N} \left(B_k^i(u, u) + |k|^2 (u)_k^i\right) e_k \gamma_k^i,$$

where for each $u \in \mathbf{H}_N$ (recall the definition of B from Section 1.3),

$$B_k^i(u,u) := \frac{1}{\pi(2\pi)^{d-1}} \langle B(u,u), e_k \gamma_k^i \rangle_{\mathbf{L}^2}.$$

The following lemma gives sufficient conditions for (u_t, x_t, v_t) to satisfy the parabolic Hörmander condition:

Lemma 5.7. *Let $\{X_j\}_{j=1}^M$ denote an enumeration of the vectors*

$$\{q_k e_k \gamma_k^i : k \in \mathcal{K},\ i = 1, \dots, d-1\}$$

and let X_0 be a vector field on $\hat{\mathbf{H}} \times \mathbb{T}^d \times \mathbb{S}^{d-1}$ of the form

$$X_0(u,x,v) = U(u) + V(u,x,v).$$

The following holds:

(1) *If $U(u) = U^S(u)$ and $\mathcal{K}$ contains the elements $(1,0), (0,1)$ and their inversions for $d = 2$ and the elements $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ and their inversions for $d = 3$, then $\{X_j\}_{j=0}^M$ satisfies the parabolic Hörmander condition.*

(2) *If $U(u) = U^{NS}(u)$ and $\mathcal{K}$ contains the elements $(1,0)$ and $(1,1)$ and their inversions for $d = 2$ and the elements $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ and their inversions for $d = 3$, then $\{X_j\}_{j=0}^M$ satisfies the parabolic Hörmander condition.*

Proof. We will consider only the Galerkin–Navier–Stokes case, since the Stokes case is even simpler. Fix $(u,x,v) \in \mathbf{H}_N \times \mathbb{T}^d \times \mathbb{S}^{d-1}$ and denote $\mathcal{V}(u,x,v)$ the span of the iterated Lie brackets of $\{X_j\}_{j=0}^M$. We have for each $k \in \mathcal{K}$ and $i = 1, \dots, d-1$,

$$[e_k \gamma_k^i, X_0] = [e_k \gamma_k^i, U^{NS}] + [e_k \gamma_k^i, V],$$

and because of the linear dependence of the vector field V on u, we obtain

$$\big[e_j \gamma_j^i, [e_k \gamma_k^i, X_0]\big] = \big[e_j \gamma_j^i, [e_k \gamma_k^i, U^{NS}]\big].$$

We will find it useful to use the following result adapted from [44] and [99].

Lemma 5.8. *Suppose that $K \subseteq \mathbb{Z}_0^d$ satisfies $K = -K$. Then at each $u \in \mathbf{H}_N$ and for each $i = 1, \dots, d-1$,*

$$\begin{aligned} &\operatorname{span}\{[e_k \gamma_k^i, [e_j \gamma_j^i, U^{NS}]] : j,k \in K\} \\ &\quad = \operatorname{span}\{e_{j+k}\gamma_{j+k}^i, e_{j-k}\gamma_{j-k}^i, e_{k-j}\gamma_{k-j}^i, e_{-j-k}\gamma_{-j-k}^i : j,k \in K\}. \end{aligned}$$

Using the fact that $(1,0)$ and $(1,1)$ and $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ are generators for the groups $(\mathbb{Z}^2, +)$ and $(\mathbb{Z}^3, +)$, respectively, we can iterate Lemma 5.8 for fixed i, taking further Lie brackets with of these new directions. Then repeating the same argument for each $i = 1, \dots, d-1$ to obtain all directions in $\mathbf{H}_N$ and conclude that

$$\mathbf{H}_N \subseteq \mathcal{V}(u,x,v).$$

This means that in order for $\{X_j\}_{j=0}^M$ to satisfy the parabolic Hörmander condition, it suffices to show that

$$\operatorname{span}\{[e_k \gamma_k^i, V] : k \in \mathcal{K}, i \in \{1, \dots, d-1\}\} = T_v \mathbb{S}^{d-1}.$$

This follows from Lemma 5.3. ∎

Analogously we have sufficient conditions for (u_t, x_t, A_t) to satisfy the parabolic Hörmander condition. The proof is almost exactly the same as the proof of Lemma 5.7, with V replaced with G. We omit the proof.

Lemma 5.9. *Let $\{X_j\}_{j=1}^M$ denote an enumeration of the vectors*

$$\{q_k e_k \gamma_k^i : k \in \mathcal{K}, i = 1, \dots, d-1\}$$

and let X_0 be a vector field on $\hat{\mathbf{H}} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$ given by

$$X_0(u, x, A) = U(u) + G(u, x, A).$$

The following holds:

(1) *If $U(u) = U^S(u)$ and $\mathcal{K}$ contains the elements $(1,0), (0,1), (1,1)$ and their inversions for $d = 2$ and the elements $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(1,1,1)$ and their inversions for $d = 3$, then $\{X_j\}_{j=0}^M$ satisfies the parabolic Hörmander condition.*

(2) *If $U(u) = U^{NS}(u)$ and $\mathcal{K}$ contains the elements $(1,0)$ and $(1,1)$ and their inversions for $d = 2$ and the elements $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and their inversions for $d = 3$, then $\{X_j\}_{j=0}^M$ satisfies the parabolic Hörmander condition.*

6. Strong Feller for the Lagrangian and projective processes

In Section 6 we will prove Proposition 2.12. We show the proof for the (u_t, x_t, v_t) process; the $(u_t, x_t, \check{v}_t)$ process is the same. Note that strong Feller for (u_t, x_t, v_t) implies the same for (u_t) and (u_t, x_t) due to the structure of the coupling.

6.1. The cutoff process

As described in Section 2.7.3 the main strategy involves proving gradient estimates on a suitable cut-off process w_t^ρ. To begin, define the following augmented system:

$$\begin{aligned}
\partial_t u_t &= -B(u_t, u_t) - A u_t + Q \dot{W}_t^u, \\
\partial_t x_t &= u_t(x_t), \\
\partial_t v_t &= \Pi_{v_t} D u_t(x_t) v_t, \\
\partial_t z_t &= \dot{W}_t^z,
\end{aligned}$$

where W_t^u is a cylindrical Wiener process on $\mathbf{L}^2$ and $W_t^z \in \mathbb{R}^{2d}$ is a finite-dimensional Wiener process independent from W_t^u. We denote this augmented process by

$$w_t = (u_t, x_t, v_t, z_t) \in \mathbf{H} \times \mathcal{M},$$

where $\mathcal{M} = \mathbb{T}^d \times \mathbb{S}^{d-1} \times \mathbb{R}^{2d}$, which satisfies the abstract SPDE

$$\partial_t w_t = \hat{F}(w_t) - A w_t + \hat{Q}\dot{W}_t, \tag{6.1}$$

where $\hat{F}$ and $\hat{Q}\dot{W}$ are given by

$$\hat{F}(u,x,v,z) = \begin{pmatrix} -B(u,u) \\ u(x) \\ \Pi_v Du(x)v \\ 0 \end{pmatrix}, \quad \hat{Q}\dot{W} = \begin{pmatrix} Q\dot{W}^u \\ 0 \\ 0 \\ \dot{W}^z \end{pmatrix}$$

(with extended definitions $Aw = (\nu(-\Delta)u, 0, 0, 0)$ in $d = 2$ and $Aw = (\nu(-\Delta)u + \eta\Delta^2 u, 0, 0, 0)$ in $d = 3$). For the remainder of this section, we will refer to the initial data of the process simply as

$$w_0 =: w.$$

Our goal will be to prove strong Feller for the augmented process (6.1). As z_t is completely uncoupled from (u_t, x_t, v_t), by restricting the class of test functions, this implies strong Feller for the original (u_t, x_t, v_t) process. Further, note that by restricting the class of test functions, strong Feller for the process defined with $v_t \in \mathbb{S}^{d-1}$ implies strong Feller for the process defined directly with $v_t \in P^{d-1}$ by relating elements in P^{d-1} to representatives in $\mathbb{S}^{d-1}$.

To define w^ρ, we will couple z_t to the x_t and v_t variables to regularize the dynamics. Specifically, as in [45], define a smooth, nonnegative cutoff function χ satisfying

$$\chi(z) = \begin{cases} 0, & z < 1, \\ 1, & z > 2, \end{cases}$$

and let $\chi_\rho(x) = \chi(x/\rho)$ for $\rho > 0$. We then define a regularized drift $F_\rho(w)$ by

$$F_\rho(u,x,v,z) = (1 - \chi_{3\rho}(\|u\|_{\mathbf{H}}))\hat{F}(u,x,v,z) + \chi_\rho(\|u\|_{\mathbf{H}})H(v,z),$$

where $H(v,z)$ is a bounded vector-field on $\mathbf{H} \times \mathcal{M}$ given by

$$H(v,z) = \begin{pmatrix} 0 \\ \sum_{j=1}^d \hat{e}_j \frac{z_j}{(1+|z_j|^2)^{\frac{1}{2}}} \\ \Pi_v \sum_{j=1}^d \hat{e}_j \frac{z_{d+j}}{(1+|z_{d+j}|^2)^{\frac{1}{2}}} \\ 0 \end{pmatrix},$$

and where we are denoting $\{\hat{e}_j\}_{j=1}^d$ the canonical basis elements in $\mathbb{R}^d$, and we are using that for each $v \in \mathbb{S}^{d-1}$, the elements $\{\Pi_v e_j\}_{j=1}^d$ span $T_v\mathbb{S}^{d-1}$. The cutoff/regularized process $w_t^\rho = (u_t^\rho, x_t^\rho, v_t^\rho, z_t)$ then satisfies the SPDE (replacing $\hat{Q} \mapsto Q$ for notational simplicity),

$$\partial_t w_t^\rho = F_\rho(w_t^\rho) - A w_t^\rho + Q\dot{W}_t. \tag{6.2}$$

It is for this process we will prove a gradient estimate on the Markov semigroup. As in [45, 100], the purpose of the cutoff is to regularize the nonlinearity so that the flow is globally Lipschitz, which is very convenient for the Malliavin calculus and high/low frequency splitting methods employed below. However, when the nonlinearity is turned off, the hypoellipticity disappears. Recovering the hypoelliptic effect is the purpose of the additional noise coming from the coupling with z_t. In [45, 100], this role is played by multiplicative white noise. This is too singular to carry out directly on the Navier–Stokes equations; in [100] it is dealt with by further mollifying the nonlinearity. One can view the use of z_t as providing a suitable regularization of the multiplicative white noise.

In what follows, we denote (via a slight abuse of notation) for H^γ, $\mathbf{L}^2$, and $\mathbf{H}$,

$$\|w_t\|_{\mathbf{H}} := \|u_t\|_{\mathbf{H}} + |z_t|,$$

where the z_t factor is treated implicitly and the manifold-valued nature of w_t is disregarded. This "norm" is not really a norm since w_t does not below to a linear space. We will also denote $T_v\mathcal{M}$ the tangent space of $\mathcal{M}$ at (x, v, z); note that the tangent space only depends on v since the other components are flat.

We are now ready to begin the proof of Proposition 2.12. The proof requires a number of estimates on w_t^ρ, its Jacobian (Fréchet derivative with respect to the initial data), various approximate Jacobians and approximate inverse Jacobians, and the Malliavin derivatives thereof. These are outlined in Section 6.5 below after the main bulk of the proof. Finally, we emphasize that for the rest of the section, the implicit constants are *always independent of t, T, $\|h\|_{\mathbf{H}\times T_v\mathcal{M}}$, and $\|w_t\|_{\mathbf{H}}$ unless specifically indicated otherwise*. Moreover, we are always assuming $T \le 1$.

The main effort in the proof of Proposition 2.12 is to obtain the following derivative estimate on the cutoff process, the proof of which comprises the rest of Section 6.

Proposition 6.1. *There exists $a_*, b_* > 0$ such that for all ρ sufficiently large, there exists a $T^* > 0$ such that for all $\phi \in C_b^2(\mathbf{H}\times\mathcal{M})$ and for $t < T^*$ the mapping $w \mapsto \hat{P}_t^\rho\phi(w)$ is differentiable and for each $w \in \mathbf{H}\times\mathcal{M}$ the derivative $D\hat{P}_\rho^t\phi(w)$ is a bounded linear operator on $\mathbf{H}\times T_v\mathcal{M}$ and satisfies for each $h \in \mathbf{H}\times T_v\mathcal{M}$,*

$$\left|D\hat{P}_t^\rho\phi(w)h\right| \lesssim_\rho t^{-a_*}(1+\|w\|_{\mathbf{H}}^{b_*})\|\phi\|_{L^\infty}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}. \tag{6.3}$$

Indeed, we do not expect that such a gradient estimate (6.3) is available for $\hat{P}_t$. Nonetheless, estimate (6.3) *is* enough to prove the strong Feller property for w_t, $\hat{P}_t$.

Proof of Proposition 2.12. Let ϕ be a bounded, measurable observable on $\mathbf{H}\times\mathcal{M}$. Let $t \le 1$ be chosen small shortly. Let $w^1, w^2 \in \mathbf{H}\times\mathcal{M}$ be such that $d(w^1, w^2) \le 1$. Naturally, we estimate the non-cut-off process by approximation and an $\varepsilon/3$ argument:

$$\begin{aligned}|\hat{P}_t\phi(w^1) - \hat{P}_t\phi(w^2)| \le |\hat{P}_t\phi(w^1) - \hat{P}_t^\rho\phi(w^1)| + |\hat{P}_t\phi(w^2) - \hat{P}_t^\rho\phi(w^2)|\\ + |\hat{P}_t^\rho\phi(w^1) - \hat{P}_t^\rho\phi(w^2)|.\end{aligned} \tag{6.4}$$

We will show that for all $\varepsilon > 0$, we may choose ρ sufficiently large and $d(w^1, w^2)$ sufficiently small such that each term is $< \varepsilon$ in size. For the first two terms in (6.4), note that

$$\begin{aligned}|\hat{P}_t\phi(w^i) - \hat{P}_t^\rho\phi(w^i)| &= |\mathbf{E}\phi(w_t(w^i) - \mathbf{E}\phi(w_t^\rho(w^i))|\\ &\le \|\phi\|_{L^\infty}\mathbf{P}\Big(\sup_{s\in(0,t)}\|w_s(w^i)\|_{\mathbf{H}} > \rho\Big),\end{aligned}$$

where $i = 1, 2$. Then, by the moment bounds in Proposition A.3, this gives the following (with implicit constant independent of t):

$$|\hat{P}_t\phi(w^i) - \hat{P}_t^\rho\phi(w^i)| \lesssim \rho^{-1}\|\phi\|_{L^\infty}(1 + \|w^i\|_{\mathbf{H}}).$$

We may now choose ρ sufficiently large depending only on $\|\phi\|_{L^\infty}$, $\|w^i\|_{\mathbf{H}}$, and ε such that

$$|\hat{P}_t\phi(w^1) - \hat{P}_t\phi(w^2)| \le |\hat{P}_t^\rho\phi(w^1) - \hat{P}_t^\rho\phi(w^2)| + 2\varepsilon.$$

Once we have fixed ρ, we may now fix $t < T_*$ such that (6.3) holds for the cutoff process. By an adaptation of [38, Lemma 7.1.5], we see that Proposition 6.1 implies

$$|\hat{P}_t^\rho\phi(w^1) - \hat{P}_t^\rho\phi(w^2)| \lesssim t^{-a_*}\|\phi\|_{L^\infty}(1 + \|w^1\|_{\mathbf{H}}^{b_*})d(w^1, w^2), \tag{6.5}$$

where for $w^i = (u^i, x^i, v^i, z^i) \in \mathbf{H} \times \mathcal{M}$, we denote

$$d(w^1, w^2) = \|u^1 - u^2\|_{\mathbf{H}} + d_{\mathcal{M}}((x^1, v^1, z^1), (x^2, v^2, z^2)),$$

where $d_{\mathcal{M}}$ is the geodesic distance on $\mathcal{M}$. Therefore, for the third term in (6.4), we may apply (6.5) and choose $d(w^1, w^2)$ sufficiently small such that

$$|\hat{P}_t\phi(w^1) - \hat{P}_t\phi(w^2)| < 3\varepsilon.$$

Hence, $\hat{P}_t$ is strong Feller. ∎

In what follows, we will drop the ρ superscripts and w_t will denote the solution to the cut-off equation (6.2).

6.2. *Derivative estimate for cutoff process via Malliavin calculus*

6.2.1. Malliavin Calculus preliminaries. First, let us recall some basics on Malliavin calculus. For the interested reader, we recommend taking a look at the monographs [23, 24, 37, 95] for a more in-depth coverage of some of these tools and techniques.

Denote $(\Omega_T, \mathbb{H}_T, \mathbf{P}_T)$ the classical Wiener space associated with a cylindrical Wiener process $(W_t)_{t\in[0,T]}$ taking values in a separable Hilbert space $\mathcal{W}$ for some finite time horizon $T > 0$ which takes values in

$$\Omega_T = C_0([0, T]; \mathcal{W}^-) = \{W \in C([0, T]; \mathcal{W}^-) : W_0 = 0\},$$

where $\mathcal{W}^-$ is any Hilbert extension of $\mathcal{W}$ such that the inclusion is Hilbert–Schmidt (since a cylindrical Wiener process W_t does not actually take values in $\mathcal{W}$). The Gaussian measure $\mathbf{P}_T$ on Ω_T is uniquely determined by its Cameron-Martin space

$$\mathbb{H}_T := \left\{G = \int_0^\cdot g_s\,\mathrm{d}s : g \in L^2([0, T]; \mathcal{W})\right\},$$

which compactly and densely embeds into $C_0([0, T]; \mathcal{W}^-)$ (see, e.g., [24, Section 2.4]).

For most of this section we will be dealing with Fréchet differentiable random variables $X : \Omega_T \to \mathcal{H} \times \mathfrak{M}$, where $\mathcal{H}$ is a separable Hilbert space and $\mathfrak{M}$ is a smooth finite-dimensional Riemannian manifold. For a given Fréchet differentiable random variable X, on $\Omega_T = C([0,T]; \mathcal{W}^-)$, we define the Malliavin derivative $\mathcal{D}_g X$ of X in direction $g \in L^2([0,T], \mathcal{W})$ by

$$\mathcal{D}_g X := \frac{\mathrm{d}}{\mathrm{d}\varepsilon} X(W + \varepsilon G)|_{h=0}, \quad G = \int_0^\cdot g_s \,\mathrm{d}s.$$

The Fréchet differentiability of X implies that $\mathcal{D}_g X$ admits a representation of the form

$$\mathcal{D}_g X = \int_0^T \mathcal{D}_s X g_s \,\mathrm{d}s, \tag{6.6}$$

where for almost every $s \in [0,T]$, we think of $\mathcal{D}_s X$ as a random, bounded linear operator from $\mathcal{W}$ to $\mathcal{H} \times T_m\mathfrak{M}$, where m denotes the projection of X onto $\mathfrak{M}$. Equivalently, we can identify $\mathcal{D}_s X$ as an element of the tensor bundle $\mathcal{W} \otimes (\mathcal{H} \times T\mathfrak{M})$. We will commonly use the following norm of $\mathcal{D}_s X$:

$$\|\mathcal{D}_s X\|_{\mathcal{W}\to\mathcal{H}\times T_m\mathfrak{M}} := \sup_{\substack{f\in\mathcal{W} \\ \|f\|_{\mathcal{W}}=1}} \|\mathcal{D}_s X f\|_{\mathcal{H}\times T_m\mathfrak{M}},$$

where $\mathcal{D}_s X f$ denotes the action of $\mathcal{D}_s X$ as a linear operator on an element $f \in \mathcal{W}$. Note that (up to isomorphism) we can identify the quantity $\mathcal{D}X = (\mathcal{D}_s X)_{s\in[0,T]}$ as a random curve in the tensor bundle $\mathcal{W} \otimes (\mathcal{H} \times T\mathfrak{M})$.

For each initial $w \in \mathbf{H} \times \mathcal{M}$ we may view the solution w_t to (6.1) as an Itô map $\Phi_t^w : C([0,t], \mathbf{H}^{-\varepsilon} \times \mathbb{R}^{2d}) \to \mathbf{H} \times \mathcal{M}$ for an appropriate choice of $\varepsilon > 0$ such that

$$Q\mathbf{H}^{-\varepsilon} \subseteq \mathbf{H} := \mathbf{H}^\sigma.$$

We have the following well-known result (cf. [65, Proposition 4.1]):

Proposition 6.2. *For each initial data $w \in \mathbf{H} \times \mathcal{M}$, let Φ_t^w be the Itô map associated to the abstract SPDE (6.1). Then Φ_t^w is Fréchet differentiable and its Fréchet derivative $D\Phi_t^w : \Omega_t \to \mathcal{H} \times T\mathfrak{M}$ for each $G \in C(\mathbb{R}_+; \mathbf{H}^{-\varepsilon} \times \mathbb{R}^{2d})$ satisfies*

$$\mathrm{d}D\Phi_t^w G = D\hat{F}(\Phi_t^w) D\Phi_t^w G \,\mathrm{d}t - AD\Phi_t^w G \,\mathrm{d}t + Q\,\mathrm{d}G_t.$$

Since the solution w_t to (6.1) is adapted to the filtration $\mathcal{F}_t$ generated by W_t, we have $\mathcal{D}_s w_t = 0$ if $s \geq t$. Indeed, from Proposition 6.2 we see that for $g \in L^2([0,T]; \mathbf{L}^2 \times \mathbb{R}^{2d})$ we have an exact formula for $\mathcal{D}_g w_t$,

$$\partial_t \mathcal{D}_g w_t = DF(w_t)\mathcal{D}_g w_t - A\mathcal{D}_g w_t + Qg_t, \quad \mathcal{D}_g w_0 = 0.$$

Then, if one defines for $0 \leq s \leq t$ the Jacobian $J_{s,t}$ (viewed as a bounded linear operator from $\mathbf{H} \times T_{v_s}\mathcal{M}$ to $\mathbf{H} \times T_{v_t}\mathcal{M}$) as the solution to the equation

$$\partial_t J_{s,t} = DF(w_t)J_{s,t} - AJ_{s,t}, \quad J_{s,s} = \mathrm{Id},$$

Duhamel's formula implies that

$$\mathcal{D}_g w_t = \int_0^t J_{s,t} Q g_s \,\mathrm{d}s,$$

consequently, by equation (6.6), this implies the following formula for $\mathcal{D}_s w_t$:

$$\mathcal{D}_s w_t = \begin{cases} J_{s,t} Q, & s < t, \\ 0, & s > t. \end{cases}$$

If the random variable $X : \Omega_T \to \mathcal{H}$ takes values in a separable Hilbert space $\mathcal{H}$, then for Fréchet differentiable X which belong to $L^2(\Omega_T; \mathcal{H})$ the Malliavin derivative $\mathcal{D}X$ belongs to the space $L^2(\Omega_T; L^2([0,T]; \mathcal{W} \times \mathcal{H}))$. It is well known (see [95, Chapter 1]) that the operator $X \to \mathcal{D}X$ is closable unbounded linear operator from $L^2(\Omega_T; \mathcal{H})$ to $L^2(\Omega_T; L^2([0,T]; \mathcal{W} \times \mathcal{H}))$ and that the closure $\mathcal{D}$ has dense domain

$$\mathrm{Dom}(\mathcal{D}) = \mathbb{W}^{1,2}(\Omega_T; \mathcal{H}) \subseteq L^2(\Omega_T; \mathcal{H}),$$

which is a Hilbert space of random variables closed with respect to the "Sobolev" norm

$$\|X\|^2_{\mathbb{W}^{1,2}(\Omega_T;\mathcal{H})} := \mathbf{E}\|X\|^2_{\mathcal{H}} + \mathbf{E}\|\mathcal{D}X\|^2_{L^2([0,T];\mathcal{W}\otimes\mathcal{H})}.$$

It is important to note that, in light of the isomorphisms

$$L^2(\Omega_T; \mathcal{H}) \simeq L^2(\Omega_T) \otimes \mathcal{H}$$

and

$$L^2(\Omega_T; L^2([0,T]; \mathcal{W} \otimes \mathcal{H})) \simeq L^2(\Omega_T; L^2([0,T]; \mathcal{W})) \otimes \mathcal{H},$$

the operator $\mathcal{D}$ simply acts a the identity map on the $\mathcal{H}$ factor and therefore only needs to be described via its action on $\mathbb{R}$-valued random variables

$$\mathcal{D} : \mathbb{W}^{1,2}(\Omega_T) \subseteq L^2(\Omega_T) \to L^2(\Omega_T; L^2([0,T]; \mathcal{W})).$$

Working instead with just real-valued random variables $X \in L^2(\Omega_T)$, we denote the adjoint operator $\mathcal{D}^*$ by

$$\mathcal{D}^* : \mathrm{Dom}(\mathcal{D}^*) \subseteq L^2(\Omega_T; L^2([0,T]; \mathcal{W})) \to L^2(\Omega_T)$$

and referred to as the *Skorokhod integral*. We will denote its action on $g \in \mathrm{Dom}(\mathcal{D}^*)$ by

$$\int_0^T \langle g_t, \delta W_t \rangle_{\mathcal{W}} := \mathcal{D}^* g.$$

In general, $\mathcal{D}^*$ is also an unbounded linear operator with a dense domain. The Skorokhod integral can be viewed as an extension of the usual Itô integral to non-adapted integrands. Indeed, when $g \in L^2(\Omega_T; L^2([0,T], \mathcal{W}))$ is adapted to the filtration $\mathcal{F}_t$ generated by W_t, then $\int_0^T \langle g_t, \delta W_t \rangle_{\mathbf{L}^2}$ coincides with the usual Itô integral $\int_0^T \langle g_t, \mathrm{d}W_t \rangle_{\mathbf{L}^2}$, and is therefore a bounded operator on adapted integrands via the Ito isometry (see [95, Section 1.3]). More generally, there is an extension of the Itô isometry for non-adapted integrands:

Proposition 6.3 ([95, Proposition 1.3.1]). *Let $g \in \mathrm{Dom}(\mathcal{D}^*)$. Then the following identity holds:*

$$\mathbf{E}\left(\int_0^T \langle g_t, \delta W_t \rangle_{\mathcal{W}} \right)^2 = \mathbf{E} \int_0^T \|g_s\|^2_{\mathcal{W}} \, \mathrm{d}s + \mathbf{E} \int_0^T \int_0^T \mathrm{tr}_{\mathcal{W}}(\mathcal{D}_s g_t \mathcal{D}_t g_s) \, \mathrm{d}s \, \mathrm{d}t.$$

Note that when g is adapted, we automatically have $\mathcal{D}_s g_t \mathcal{D}_t g_s = 0$, and therefore (6.3) recovers the standard Itô formula and therefore g belongs to $\mathrm{Dom}(\mathcal{D}^*)$. Moreover, if $g \in \mathbb{W}^{1,2}(\Omega_T; L^2([0,T];\mathcal{W}))$ is not necessarily adapted, then we have

$$\begin{aligned}
&\mathbf{E}\int_0^T \|g_s\|_{\mathcal{W}}^2\,\mathrm{d}s + \mathbf{E}\int_0^T\int_0^T \mathrm{tr}_{\mathcal{W}}(\mathcal{D}_s g_t \mathcal{D}_t g_s)\,\mathrm{d}s\,\mathrm{d}t\\
&\quad\le \mathbf{E}\int_0^T \|g_s\|_{\mathcal{W}}^2\,\mathrm{d}s + \mathbf{E}\int_0^T\int_0^T \|\mathcal{D}_t g_s\|_{\mathcal{W}\to\mathcal{W}}^2\,\mathrm{d}s\,\mathrm{d}t\\
&\quad= \|g\|^2_{\mathbb{W}^{1,2}(\Omega_T;L^2([0,T];\mathcal{W}))},
\end{aligned}$$

so that in general we have that $\mathbb{W}^{1,2}(\Omega_T; L^2([0,T];\mathcal{W})) \subseteq \mathrm{Dom}(\mathcal{D}^*)$.

One of the valuable features of the Skorokhod integral is the role it plays in the so-called *Malliavin integration by parts formula*, which plays a crucial role in the regularity theory for the law of the stochastic process w_t. We state the formula in the following form, which follows immediately from the definition of the Skorokhod integral and the Malliavin derivative on real-valued random variables.

Proposition 6.4. *Let ϕ be a bounded differentiable function on $\mathbf{H}\times\mathcal{M}$ with bounded derivatives and $g \in \mathbb{W}^{1,2}(\Omega_T; L^2([0,T];\mathbf{L}^2\times\mathbb{R}^{2d}))$. Then the following relation holds for each $t\in[0,T]$:*

$$\mathbf{E}\mathcal{D}_g\phi(w_t) = \mathbf{E}\langle \mathcal{D}\phi(w_t), g\rangle_{L^2([0,t],\mathbf{L}^2\times\mathbb{R}^{2d})} = \mathbf{E}\left(\phi(w_t)\int_0^t \langle g_s, \delta W_s\rangle_{\mathbf{L}^2}\right).$$

6.2.2. Derivative estimate. As discussed in Section 2.7, the integration by parts formula (6.4) can be used to obtain a gradient estimate on the Markov semigroup if for any $h\in\mathbf{H}\times T_v\mathcal{M}$ and some $T>0$, one can obtain a control

$$g\in\mathbb{W}^{1,2}(\Omega_T; L^2([0,T];\mathbf{L}^2\times\mathbb{R}^{2d}))$$

(depending on h) such that we have the equality

$$\mathcal{D}_g w_T = Dw_T h,$$

where $Dw_T : \mathbf{H}\times T\mathcal{M}\to\mathbf{H}\times T\mathcal{M}$ denotes the Fréchet derivative with respect to the initial data. This however, does not appear to be possible to do in general. We must instead find a control $g\in L^2([0,T];\mathcal{W})$ which satisfies this approximately, so that for some time $T>0$ we have

$$\mathcal{D}_g w_T = Dw_T h + r_T,$$

where r_T is a remainder that satisfies $r_T\to 0$ as $T\to 0$.

Indeed, most of the work of this section is to prove the following key lemma.

Lemma 6.5. *For all $\rho>0$, there exist constants $a_*, b_*>0$ such that for T sufficiently small and $h\in\mathbf{H}\times T_v\mathcal{M}$ there exists a corresponding control $g=(g_t)_{t\in[0,T]}$ satisfying*

$$\begin{aligned}
&\mathbf{E}\int_0^T \|g_t\|_{\mathbf{L}^2}^2\,\mathrm{d}t + \mathbf{E}\int_0^T\int_0^T \|\mathcal{D}_s g_t\|^2_{\mathbf{L}^2\to\mathbf{L}^2}\,\mathrm{d}s\,\mathrm{d}t\\
&\quad\lesssim_\rho T^{-2a_*}(1+\|w\|_{\mathbf{H}})^{2b_*}\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}},
\end{aligned} \tag{6.7}$$

such that for $r_T = \mathcal{D}_g w_T - Dw_T h$, *we have*

$$\mathbf{E}\|r_T\|^2_{\mathbf{H}\times T_{v_T}\mathcal{M}} \lesssim_\rho T\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}}. \tag{6.8}$$

We remark that it would suffice to prove $\mathbf{E}\|r_T\|^2_{\mathbf{H}\times T_{v_T}\mathcal{M}} \lesssim_\rho q(T)\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}}$ for some q satisfying $\lim_{T\to 0} q(T) = 0$.

Lemma 6.5 is indeed enough to prove Proposition 6.1.

Proof of Proposition 6.1. Using the control from Lemma 6.5, we can now estimate the derivative of the semigroup in direction h at time $2T$ for $\phi \in C^2$,

$$\begin{aligned} D\hat{P}_{2T}\phi(w)h &= \mathbf{E}(D\hat{P}_T\phi(w_T)Dw_T h) \\ &= \mathbf{E}(D\hat{P}_T\phi(w_T)\mathcal{D}_g w_T) - \mathbf{E}(D\hat{P}_T\phi(w_T)r_T). \end{aligned}$$

Using the chain rule $\mathcal{D}_g \hat{P}_T\phi(w_T) = D\hat{P}_T\phi(w_T)\mathcal{D}_g w_T$, the Malliavin integration by parts formula (Proposition 6.4) gives

$$D\hat{P}_{2T}\phi(w)h = \mathbf{E}\Big(\hat{P}_T\phi(w_T)\int_0^T \langle g_t, \delta W_t\rangle_{\mathbf{L}^2}\Big) - \mathbf{E}(D\hat{P}_T\phi(w_T)r_T), \tag{6.9}$$

where the stochastic integral is interpreted as a Skorokhod integral, since the control is not adapted. By identity (6.3) and estimate (6.7) we have

$$\begin{aligned} \mathbf{E}\Big(\int_0^T \langle g_t, \delta W_t\rangle_{\mathbf{L}^2}\Big)^2 &\le \mathbf{E}\int_0^T \|g_t\|^2_{\mathbf{L}^2}\,\mathrm{d}t + \mathbf{E}\int_0^T\int_0^T \|\mathcal{D}_s g_t\|^2_{\mathbf{L}^2\to\mathbf{L}^2}\,\mathrm{d}s\,\mathrm{d}t \\ &\lesssim_\rho T^{-2a_*}(1+\|w\|_{\mathbf{H}})^{2b_*}\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}}. \end{aligned} \tag{6.10}$$

To finish the proof, introduce the seminorm $\|\cdot\|_{a_*,b_*,T_*}$ on $C([0,T_*];C^1(\mathbf{H}\times\mathcal{M}))$, for $a_*, b_* > 0$ and $0 < T_* \le 1$ by

$$\|f\|_{a_*,b_*,T_*} = \sup_{\substack{t\in[0,T_*]\\ w\in\mathbf{H}\times\mathcal{M}\\ h\in\mathbf{H}\times T_v\mathcal{M},\, h\neq 0}} \frac{t^{a_*}|Df_t(w)h|}{\|h\|_{\mathbf{H}\times T_v\mathcal{M}}(1+\|w\|_{\mathbf{H}})^{b_*}}.$$

Then it follows from (6.8) and (6.10) that for $2T < T_*$,

$$\begin{aligned} |DP_{2T}\phi(w)h| &\lesssim \|\phi\|_{L^\infty}T^{-a_*}(1+\|w\|^{b_*}_{\mathbf{H}})\|h\|_{\mathbf{H}\times T_v\mathcal{M}} \\ &\quad + \|P\phi\|_{a_*,b_*,T_*}T^{-a_*}\sqrt{\mathbf{E}(1+\|w_T\|_{\mathbf{H}})^{2b_*}}\sqrt{\mathbf{E}\|r_T\|^2_{\mathbf{H}\times T_{v_T}\mathcal{M}}} \\ &\lesssim \|\phi\|_{L^\infty}T^{-a_*}(1+\|w\|_{\mathbf{H}})^{b_*}\|h\|_{\mathbf{H}\times T_v\mathcal{M}} \\ &\quad + \|P\phi\|_{a_*,b_*,T_*}T^{-a_*+\frac{1}{2}}(1+\|w\|_{\mathbf{H}})^{b_*}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}, \end{aligned}$$

and therefore

$$\|P\phi\|_{a_*,b_*,T_*} \lesssim \|\phi\|_{L^\infty} + T_*^{\frac{1}{2}}\|P\phi\|_{a_*,b_*,T_*},$$

then by taking T_* small enough, we obtain

$$\|P\phi\|_{a_*,b_*,T_*} \lesssim \|\phi\|_{L^\infty}.$$

This is the a priori estimate stated in (6.3). ∎

6.3. Construction of control and estimates of remainder

The rest of the section is dedicated to proving Lemma 6.5. First, we implement a splitting into high and low frequencies similar to that of [45, 100]. This will allow us to build a control that works differently on the high and low frequencies. To this, denote the set $K_L \subseteq \mathbb{Z}_0^d$ of low modes by

$$K_L = \{k \in \mathbb{Z}_0^d : |k|_\infty \leq L\},$$

where L is as in Assumption 2. Let $\Pi_L : \mathbf{H} \to \mathbf{H}$ denote the corresponding orthogonal projection onto the "low modes" belonging to K_L and let $\Pi_H = I - \Pi_L$ be the complementary projection onto the "high modes" belonging to $\mathbb{Z}_0^d \setminus K_L$. Let $\mathbf{H}_L$ and $\mathbf{H}_H$ denote the ranges of Π_L and Π_H, respectively, so that we have the orthogonal decomposition

$$\mathbf{H} = \mathbf{H}_L \oplus \mathbf{H}_H.$$

Given $w = (u, x, v, z) \in \mathbf{H} \times \mathcal{M}$, we will extend the definition of Π_L and Π_H to $\mathbf{H} \times \mathcal{M}$ so that $\mathcal{M}$ is included with the low modes by

$$w^L = \Pi_L w = (u^L, x, v, z) \quad \text{and} \quad w^H = \Pi_H w = u^H.$$

Naturally this defines low and high processes w_t^L and w_t^H, which satisfy (note of course they are coupled)

$$\begin{aligned}\partial_t w_t^L &= F_L(w_t) - A_L w_t^L + Q_L \dot{W}_t^L,\\ \partial_t w_t^H &= F_H(w_t) - A_H w_t^H + Q_H \dot{W}_t^H,\end{aligned}$$

where

$$\begin{aligned}F_L(w) = \Pi_L F(w), \quad F_H(w) = \Pi_H F(w), \quad A_H w = \Pi_H A w,\\ Q_L = \Pi_L Q, \qquad\qquad Q_H = \Pi_H Q.\end{aligned}$$

We also define the finite-dimensional matrix $U_{s,t}^L$ which we view a linear operator from $\mathbf{H}_L \times T_{v_s}\mathcal{M}$ to $\mathbf{H}_L \times T_{v_t}\mathcal{M}$ as well as the bounded linear operator $U_{s,t}^H$ from $\mathbf{H}_H$ to $\mathbf{H}_H$ by

$$\partial_t U_{s,t}^L = -A_L U_{s,t}^L + D_L F_L(w_t) U_{s,t}^L, \quad U_{s,s}^L = \mathrm{Id},$$

and for $0 \leq s \leq t$,

$$\partial_t U_{s,t}^H = -A_H U_{s,t}^H + D_H F_H(w_t) U_{s,t}^H, \quad U_{s,s}^H = \mathrm{Id}.$$

Both $U_{s,t}^L$ and $U_{s,t}^H$ serve as approximations for the full Jacobian $J_{s,t}$ of the flow $w \mapsto w_t$ projected onto the low and high-modes when t is small. We see that $U_{s,t}^L$ is an invertible operator: denote its inverse by

$$V_{s,t}^L = (U_{s,t}^L)^{-1}.$$

When $s = 0$, we write $U_t^L = U_{0,t}^L$ and $V_t^L = V_{0,t}^L$. Using the fact that U_t^L is invertible, we can write $U_{s,t}^L = U_t^L V_s^L$.

Definition 6.6. Define the *partial Malliavin matrix* $\mathcal{C}_t^L : \mathbf{H}_L \times T_v\mathcal{M} \to \mathbf{H}_L \times T_v\mathcal{M}$ by

$$\mathcal{C}_t^L := \int_0^t V_s^L Q_L (V_s^L Q_L)^\top \, \mathrm{d}s.$$

Remark 6.7. The matrix $\mathcal{C}_t^L$ is the analogue of the reduced Malliavin matrix, introduced by Norris [94], in order to simplify Malliavin's proof of Hörmander's theorem. The name *partial Malliavin matrix* comes from [45], and indicates that it is a finite-dimensional Malliavin matrix associated to the low modes. The reduced Malliavin matrix is convenient primarily because, unlike the original Malliavin matrix, $\mathcal{C}_t^L$ is adapted which greatly simplifies calculations, for example, Itô's lemma giving rise to the brackets in a straightforward manner. Note the reduction here is only possible because U_t^L is invertible due to finite dimensionality; see [64, 65] for more discussion. This invertibility motivates the frequency splitting between high and low frequencies.

One of main results of Section 6 is the nondegeneracy of $\mathcal{C}_t^L$, which allows us to build the low frequencies part of the control g_t. That is, we have the following; the proof is involved and is carried out in Section 6.4 below.

Lemma 6.8. *The matrix $\mathcal{C}_T^L$ is almost surely invertible on $\mathbf{H}_L \times T_v\mathcal{M}$. Furthermore, there exists constants a, b such that for all $p \geq 1$,*

$$\mathbf{E}|(\mathcal{C}_T^L)^{-1}|^p \lesssim_{\rho,p} T^{-ap}(1+|z|)^{bp}.$$

Using Lemma 6.8, we can now construct the control. Specifically, fix an $h \in \mathbf{H} \times T_v\mathcal{M}$, a $T \in (0,1)$, a frequency cut-off N chosen as $N := T^{-2a}(1+|z|)^{2b}$ (a and b as in Lemma 6.8) and define $t \mapsto g_t \in \mathbf{L}^2$ by

$$g_t^L = (V_t^L Q_L)^\top (\mathcal{C}_T^L)^{-1} V_T^L D w_T^L h, \tag{6.11}$$

$$g_t^H = -Q_H^{-1} \Pi_{\leq N} D_L F_H(w_t)\zeta_t + 2T^{-1} Q_H^{-1} U_{0,t}^H h_H \mathbb{1}_{[T/4,3T/4]}(t), \tag{6.12}$$

where $\Pi_{\leq N}$ is a projection onto frequencies less than N and (ζ_t) is a process belonging for each $t \in [0,T]$ to $\mathbf{H}_L \times \mathbb{T}_{v_t}\mathcal{M}$ and solving the following system:

$$\begin{cases} \dot{\zeta}_t = -A_L\zeta_t + D_L F_L(w_t)\zeta_t + Q_L g_t^L + D_H F_L(w_t)\xi_t, \\ \dot{\xi}_t = -A_H\xi_t + D_H F_H(w_t)\xi_t + \Pi_{>N} D_L F_H(w_t)\zeta_t \\ \qquad + 2T^{-1} U_{0,t}^H h_H \mathbb{1}_{[T/4,3T/4]}(t), \end{cases} \tag{6.13}$$

with $\xi_0 = 0$ and $\zeta_0 = 0$. If one assumes that a solution to (6.13) exists and is unique (this is proved in Lemma 6.10 below), then we find that the choice of control is made specifically so that the remainder r_T assumes a nice form. The intuitive choice is to use the much simpler control, as was done in [45],

$$\tilde{g}_t^L = (V_t^L Q_L)^\top (\mathcal{C}_T^L)^{-1} V_T^L D w_T^L h,$$

$$\tilde{g}_t^H = 2T^{-1} Q_H^{-1} U_{0,t}^H h_H \mathbb{1}_{[T/4,3T/4]}(t),$$

however, this leads to errors which are too large at high frequencies in the case of Navier–Stokes, specifically, the low frequencies create too large of an effect on the high frequencies. The choice (6.12) adjusts the control to account for this perturbation at high frequencies, however, this procedure is not straightforward as this would require, in some sense, choosing the control based on the Malliavin derivative itself, which is why we must solve the entire system (6.13). Indeed, by uniqueness we see below a self-consistency:$\mathcal{D}_g w_t^L = \zeta_t$ and $\mathcal{D}_g w_t^H = \xi_t$.

In what follows the implicit constant is always independent of N unless otherwise indicated.

Lemma 6.9. *Assume that g is defined as above and that exists a unique solution (ζ_t, ξ_t) to system (6.13) in the space $L^2(\Omega; C([0,T]; \mathbf{H} \times T_{v_t}\mathcal{M}))$, then the remainder*

$$r_T = \mathcal{D}_g w_T - D w_T h$$

satisfies

$$r_T^L = \int_0^T U_{t,T}^L D_H F_L(w_t)\xi_t \,\mathrm{d}t, \tag{6.14}$$

$$\begin{aligned} r_T^H &= \int_0^T U_{t,T}^H \Pi_{>N} D_L F_H(w_t)\zeta_t \,\mathrm{d}t \\ &\quad - \int_0^T U_{t,T}^H D_L F_H(w_t) D_H w_t^L h_H \,\mathrm{d}t - D_L w_T^H h_L. \end{aligned} \tag{6.15}$$

Proof. Using (6.13), Proposition 6.2, and uniqueness for the system

$$\begin{cases} \mathcal{D}_g w_t^L = \int_0^t U_t^L V_\tau^L D_H F_L(w_t)(\mathcal{D}_g w_\tau^H - \xi_\tau)\,\mathrm{d}\tau, \\ \mathcal{D}_g w_t^H = \int_0^t U_{\tau,t}^H D_L F_H(w_t)(\mathcal{D}_g w_\tau^L - \zeta_\tau)\,\mathrm{d}\tau, \end{cases}$$

in $L^2(\Omega; C([0,T]; \mathbf{H} \times T_{v_t}\mathcal{M}))$, we obtain that in fact $\mathcal{D}_g w_t^L = \zeta_t$ and $\mathcal{D}_g w_t^H = \xi_t$ and we obtain the following formulas for the Malliavin derivatives at time T:

$$\begin{aligned} \mathcal{D}_g w_T^L &= D w_T^L h + U_T^L \int_0^T V_t^L D_H F_L(w_t)\xi_t \,\mathrm{d}t, \\ \mathcal{D}_g w_T^H &= U_{0,T}^H h_H + \int_0^T U_{t,T}^H \Pi_{>N} D_L F_H(w_t)\zeta_t \,\mathrm{d}t. \end{aligned}$$

Note that $\mathcal{D}_g w_T^L$ is equal to $D w_T^L h$ plus remainders, while $\mathcal{D}_g w_T^H$ is a perturbation of $U_{0,T}^H h_H$, that is,

$$D w_t^H h_H = U_{0,t}^H h_H + \int_0^t U_{s,t}^H D_L F_H(w_s) D_H w_s^L h_H \,\mathrm{d}s.$$

Using this relation, we now write

$$\begin{cases} \mathcal{D}_g w_T^L = D w_T^L h + r_T^L. \\ \mathcal{D}_g w_T^H = D w_T^H h + r_T^H. \end{cases} \tag{6.16}$$

where r_T^L and r_T^H are given by (6.14) and (6.15). ∎

Next, we construct a unique solution to (6.13) and provide the necessary quantitative estimates. These in turn will imply the existence of a suitable control g_t.

Lemma 6.10. *For all $T > 0$ sufficiently small (depending only on ρ), and all $p \geq 2$, there exists a unique solution $\eta_t = (\zeta_t, \xi_t) \in \mathbf{H} \times T_{v_t}\mathcal{M}$ on $[0, T]$ to system (6.13) in $L^p(\Omega; C([0, T]; \mathbf{H} \times T_{v_t}\mathcal{M}))$ satisfying*

$$\Big(\mathbf{E} \sup_{t\in[0,T]} \|\eta_t\|^p_{\mathbf{H}\times T_{v_t}\mathcal{M}}\Big)^{\frac{1}{p}} + \Big(\mathbf{E} \sup_{s,t\in[0,T]} \|\mathcal{D}_s\eta_t\|^p_{\mathbf{L}^2\to\mathbf{H}\times T_{v_t}\mathcal{M}}\Big)^{\frac{1}{p}} \lesssim T^{-2a}(1 + |z|)^{2b}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}.$$

Note that η_t is not adapted to the filtration $(\mathcal{F}_t)$.

Proof. Formally we may re-write a solution to (6.13) as

$$\zeta_t = \int_0^t U^L_{s,t} Q_L g^L_s \,\mathrm{d}s + \int_0^t U^L_{s,t} D_H F_L(w_s)\xi_s \,\mathrm{d}s, \tag{6.17}$$

$$\xi_t = \frac{2}{T}|[0, t] \cap [\tfrac{T}{4}, \tfrac{3T}{4}]| U^H_{0,t} h_H + \int_0^t U^H_{s,t} \Pi_{>N} D_L F_H(w_s)\zeta_s \,\mathrm{d}s. \tag{6.18}$$

The lemma is proved via a fixed point for the pair $\eta = \{(\zeta_t, \xi_t), t \in [0, T]\}$ in the Banach space $\mathbb{X}_T$ defined by the following norm:

$$\|\eta\|_{\mathbb{X}_T} := \Big(\mathbf{E} \sup_{t\in[0,T]} \|\eta_t\|^p_{\mathbf{H}\times T_{v_t}\mathcal{M}}\Big)^{\frac{1}{p}} + \Big(\mathbf{E} \sup_{s,t\in[0,T]} \|\mathcal{D}_s\eta_t\|^p_{\mathbf{L}^2\to\mathbf{H}\times T_{v_t}\mathcal{M}}\Big)^{\frac{1}{p}}.$$

Note that equations (6.17) and (6.18) are linear and can be written more compactly on $\mathbb{X}_T$ as

$$\eta = L_T\eta + F_T,$$

where L_T and F_T are given by

$$(L_T\eta)_t = \begin{pmatrix} \int_0^t U^L_{s,t} D_H F_L(w_s)\xi_s \,\mathrm{d}s \\ \int_0^t U^H_{s,t} \Pi_{>N} D_L F_H(w_s)\zeta_s \,\mathrm{d}s \end{pmatrix}$$

and

$$(F_T)_t = \begin{pmatrix} \int_0^t U^L_{s,t} Q_L g^L_s \,\mathrm{d}s \\ \frac{2}{T}|[0, t] \cap [\frac{T}{4}, \frac{3T}{4}]| U^H_{0,t} h_H \end{pmatrix}.$$

Our goal will be to estimate $L_T\eta$ and F_T in $\mathbb{X}_T$. Specifically, we will show that

$$\|L_T\eta\|_{\mathbb{X}_T} \lesssim_\rho T^{\frac{1}{2}}\|\eta\|_{\mathbb{X}_T}, \tag{6.19}$$

$$\|F_T\|_{\mathbb{X}_T} \lesssim_\rho T^{-2a}(1 + |z|)^{2b}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}. \tag{6.20}$$

This implies that for small enough T (depending only on ρ), the mapping $\eta \mapsto L_T\eta + F_T$ is a contraction and maps the ball $\mathbb{B}_T = \{\eta \in \mathbb{X}_T : \|\eta\|_{\mathbb{X}_T} \leq 2\|F_T\|_{\mathbb{X}_T}\}$ into itself. By the Contraction Mapping Theorem this implies the existence of a unique solution to $\eta = L_T\eta + F_T$ satisfying

$$\|\eta\|_{\mathbb{X}_T} \leq 2\|F_T\|_{\mathbb{X}_T} \lesssim_\rho T^{-2a}(1 + |z|)^{2b}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}.$$

To estimate $L_T\eta$ and F_T in $\mathbb{X}_T$, we need to compute the Malliavin derivatives. We find for each $f \in \mathbf{L}^2$,

$$(\mathcal{D}_s F_T f)_t = \begin{pmatrix} \int_0^t [\mathcal{D}_s U^L_{r,t} f] Q_L g^L_r \,\mathrm{d}r + \int_0^t U^L_{r,t} Q_L \mathcal{D}_s g^L_r f \,\mathrm{d}r \\ [\mathcal{D}_s U^H_{0,t} f] h_H \frac{2}{T} |[0,t] \cap [\frac{T}{4}, \frac{3T}{4}]| \end{pmatrix},$$

and for each $\eta \in \mathbb{X}_T$ using the chain rule

$$\mathcal{D}_s(L_T\eta) f = [\mathcal{D}_s L_T f]\eta + L_T \mathcal{D}_s \eta f,$$

where $([\mathcal{D}_s L_T f]\eta)_t$ is

$$\begin{pmatrix} \int_0^t [\mathcal{D}_s U^L_{r,t} f] D_H F_L(w_r)\xi_r \,\mathrm{d}r + \int_0^t U^L_{r,t} Q_L D^2 F_L[\xi_r, J_{s,r} Q f] \,\mathrm{d}r \\ \int_0^t [\mathcal{D}_s U^H_{r,t} f] \Pi_{>N} D_L F_H(w_r)\zeta_r \,\mathrm{d}r + \int_0^t U^H_{r,t} \Pi_{>N} D^2 F_H[\zeta_r, J_{s,r} Q f] \end{pmatrix}.$$

We observe by Lemmas 6.8, 6.21, and 6.19, that

$$\mathbf{E} \sup_{s\in[0,T]} |g^L_s|^p \lesssim T^{-ap}(1+|z|)^{bp} \|h\|_{\mathbf{H}\times T_v\mathcal{M}},$$

and by the product rule, Lemmas 6.21, 6.23, and 6.24, there holds

$$\mathbf{E} \sup_{s,t\in[0,T]} \|\mathcal{D}_s g^L_t\|^p_{\mathbf{L}^2\to\mathbf{L}^2_L} \lesssim T^{-(2a+1)p}(1+|z|)^{2bp} \|h\|^p_{\mathbf{H}\times T_v\mathcal{M}}.$$

Using the bounds on $U^L_{r,t}, U^H_{r,t}, \mathcal{D}_s U^L_{r,t}$ and $\mathcal{D}_s U^H_{r,t}$, in Lemmas 6.21 and 6.23, we moreover have

$$\|F_T\|_{\mathbb{X}_T} \lesssim_\rho T\|g^L\|_{\mathbb{X}_T} + \|h_H\|_{\mathbf{H}_H} \lesssim T^{-2a}(1+|z|)^{2b}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}.$$

To estimate $L_T\eta$, we use the bounds on $U^L_{r,t}, U^H_{r,t}$ (from Lemma 6.21) to obtain the almost sure bounds

$$\begin{aligned} \sup_{t\in[0,T]} \|(L_T\eta)_t\|_{\mathbf{H}\times T_{v_t}\mathcal{M}} &\lesssim_\rho T \sup_{t\in[0,T]} \|\xi_t\|_{\mathbf{H}_H} + T^{\frac{1}{2}} \sup_{t\in[0,T]} |\zeta_t| \\ &\lesssim_\rho T^{\frac{1}{2}} \sup_{t\in[0,T]} \|\eta_t\|_{\mathbf{H}\times T_{v_t}\mathcal{M}}. \end{aligned} \tag{6.21}$$

Additionally, using bounds on $J_{s,t}, \mathcal{D}_s U^L_{r,t}$ and $\mathcal{D}_s U^H_{r,t}$ (from Lemmas 6.19 and 6.23), we also find

$$\sup_{s,t\in[0,T]} \|([\mathcal{D}_s L_T]\eta)_t\|_{\mathbf{L}^2\to\mathbf{H}\times T_{v_t}\mathcal{M}} \lesssim_\rho T^{\frac{1}{2}} \sup_{t\in[0,T]} \|\eta_t\|_{\mathbf{H}\times T_{v_t}\mathcal{M}},$$

and therefore by estimate (6.21) applied to $\mathcal{D}_s\eta$ instead of η, we find

$$\begin{aligned} &\sup_{s,t\in[0,T]} \|\mathcal{D}_s(L_T\eta)_t\|_{\mathbf{L}^2\to\mathbf{H}\times T_{v_t}\mathcal{M}} \\ &\quad\lesssim T^{\frac{1}{2}}\Big(\sup_{t\in[0,T]} \|\eta_t\|_{\mathbf{H}\times T_{v_t}\mathcal{M}} + \sup_{s,t\in[0,T]} \|\mathcal{D}_s\eta_t\|_{\mathbf{L}^2\to\mathbf{H}\times T_{v_t}\mathcal{M}}\Big). \end{aligned} \tag{6.22}$$

Putting (6.21) and (6.22) together and taking the $L^p(\Omega)$ norm gives estimate (6.19). ■

We are now ready to prove Lemma 6.5.

Proof of Lemma 6.5. First we prove estimate (6.8) on the remainder r_T. It is here where we will need to set the choice of N depending on T and $|z|$. To begin, we note that from equation (6.18), using the cut-off $\Pi_{>N}$, we obtain the following improved estimate on ξ_t:

$$\Big(\mathbf{E}\sup_{t\in[0,T]}\|\xi_t\|^2_{\mathbf{L}^2_H}\Big)^{\frac{1}{2}} \lesssim \|h\|_{\mathbf{H}\times T_v\mathcal{M}} + N^{1-\sigma}T^{-2a}(1+|z|)^{2b}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}.$$

Therefore, since $\sigma-1>1$ and the definition $N=T^{-2a}(1+|z|)^{2b}$, we obtain the T independent bound

$$\Big(\mathbf{E}\sup_{t\in[0,T]}\|\xi_t\|^2_{\mathbf{L}^2_H}\Big)^{\frac{1}{2}} \lesssim \|h\|_{\mathbf{H}\times T_v\mathcal{M}}.$$

Recall the definition of the remainders (6.14) and (6.15). We estimate r^L first. We find (noting $|D_H F_L(w_t)\xi_t| \lesssim \chi_{3\rho}(\|u_t\|_{\mathbf{H}})\|u_t\|_{H^\gamma}\|\xi_t\|_{\mathbf{L}^2_H}$ for any $\gamma>\frac{d}{2}+1$ due to the frequency projection)

$$|r^L_T| = \left|U^L_T\int_0^T V^L_t D_H F_L(w_t)\xi_t\,\mathrm{d}t\right| \lesssim_\rho T\sup_{t\in[0,T]}\big(|U^L_t||V^L_t|\|\xi_t\|_{\mathbf{L}^2_H}\big),$$

and therefore using almost sure bounds on U^L_t and V^L_t from Lemma 6.21,

$$\mathbf{E}|r^L_T|^2 \lesssim_\rho T^2\mathbf{E}\sup_{t\in[0,T]}\|\xi_t\|^2_{\mathbf{L}^2_H} \lesssim T^2\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}}.$$

Hence, r^L_T satisfies the estimate required for (6.8).

Turn next to estimating r^H_t. We again use the frequency truncation $\Pi_{\le N}$ and the choice $N=T^{-2a}(1+|z|)^{2b}$ to find

$$\begin{aligned}\|r^H_T\|_{\mathbf{H}_H} &\lesssim \int_0^T \frac{1}{N(T-t)^{\frac{1}{2}}}|\zeta_t|\,\mathrm{d}t + \int_0^T\frac{1}{(T-t)^{\frac{1}{2}}}|D_H w^L_t h_H|\,\mathrm{d}t + \|D_L w^H_T h_L\|_{\mathbf{H}_H}\\ &\lesssim_\rho T^{\frac{1}{2}+2a}(1+|z|)^{2b}\sup_{t\in[0,T]}|\zeta_t| + T^{\frac{1}{2}}\sup_{t\in[0,T]}|D_H w^L_t h_H| + \|D_L w^H_T h_L\|_{\mathbf{H}_H}.\end{aligned}$$

Using that Lemma 6.10 gives

$$T^{4a}(1+|z|)^{4b}\mathbf{E}\sup_{t\in[0,T]}|\zeta_t|^2 \lesssim 1,$$

along with Lemma 6.22 for $D_H w^L$ and $D_L w^H$, we conclude that r^H_T satisfies the estimate required for (6.8).

Next we show estimate (6.7) on the control g. Recall from the proof of Lemma 6.10 that we can use the bounds on the partial Malliavin matrix $\mathcal{C}^L_t$ to get the following estimate on g^L:

$$\mathbf{E}\sup_{0<t\le T}|g^L_t|^2 + \mathbf{E}\sup_{s,t\in[0,T]}\|\mathcal{D}_s g^L_t\|^2_{\mathbf{L}^2\to\mathbf{L}^2_L} \lesssim T^{-4a}(1+|z|)^{4b}\|h\|^2_{\mathbf{H}\times T_v\mathcal{M}}.$$

It remains to estimate g^H. Recall the following formulas for g^H_t and $\mathcal{D}_s g^H_t$:

$$g^H_t = -Q_H^{-1}\Pi_{\le N}D_L F_H(w_t)\zeta_t + 2T^{-1}Q_H^{-1}U^H_{0,t}h_H\mathbb{1}_{[T/4,3T/4]}(t)$$

and

$$\mathcal{D}_s g_t^H f = Q_H^{-1}\Pi_{\le N} D^2 F_H(w_t)[\zeta_t, J_{s,t} Q f] + Q_H^{-1}\Pi_{\le N} D_L F_H(w_t)\mathcal{D}_s \zeta_t f \\ + 2T^{-1} Q_H^{-1}\mathcal{D}_s U_{0,t}^H f h_H \mathbb{1}_{[T/4,3T/4]}(t).$$

Using the cut-off $\Pi_{\le N}$ and the *lower bound* in Assumption 2,

$$\begin{aligned}\mathbf{E}\int_0^T \|g_t^H\|_{\mathbf{L}_H^2}^2 \,\mathrm{d}t &\lesssim \mathbf{E}\int_0^T \|Q_H^{-1}\Pi_{\le N} D_L F_H(w_t)\zeta_t\|_{\mathbf{L}_H^2}^2 \\ &\quad + T^{-1}\mathbf{E}\int_{T/4}^{3T/4} \|Q_H^{-1} U_{0,t}^H h_H\|_{\mathbf{L}_H^2}^2 \,\mathrm{d}t \\ &\lesssim \mathbf{E}\int_0^T \|\Pi_{\le N} D_L F_H(w_t)\zeta_t\|_{H^\alpha}^2 \,\mathrm{d}t \\ &\quad + T^{-1}\mathbf{E}\int_{T/4}^{3T/4} \|U_{0,t}^H h_H\|_{H^\alpha}^2 \,\mathrm{d}t \\ &\lesssim_\rho N^4 T \sup_{t\in[0,T]} |\zeta_t|^2 + T^{-2}(1+\|w\|_{\mathbf{H}}^2)\|h_H\|_{\mathbf{H}_H}^2,\end{aligned}$$

where in the last line we used (6.36) on $U_{0,t}^H$ with $\gamma = \alpha - 1$. This is where we use the requirement $\sigma \in (\alpha - 2(d-1), \alpha - \frac{d}{2})$. A similar calculation for $\mathcal{D}_s g_t^H$ yields and

$$\begin{aligned}&\mathbf{E}\int_0^T\int_0^T \|\mathcal{D}_s g_t^H\|_{\mathbf{L}^2\to\mathbf{L}_H^2}^2 \,\mathrm{d}t\,\mathrm{d}s \\ &\quad\lesssim_\rho N^4 T^2 \mathbf{E}\Big(\sup_{t\in[0,T]}|\zeta_t| + \sup_{s,t\in[0,T]} \|\mathcal{D}_s\zeta_t\|_{\mathbf{L}^2\to\mathbf{H}_L\times T_{v_t}\mathcal{M}}^2\Big) \\ &\qquad + T^{-2}(1+\|w\|_{\mathbf{H}}^2)\|h_H\|_{\mathbf{H}_H}^2.\end{aligned}$$

Using the estimate on ξ_t from Lemma 6.10 and our choice of $N = T^{-2a}(1+|z|)^{2b}$, we find

$$\begin{aligned}&\mathbf{E}\int_0^T \|g_t^H\|_{\mathbf{L}_H^2}^2 \,\mathrm{d}t + \mathbf{E}\int_0^T\int_0^T \|\mathcal{D}_s g_t^H\|_{\mathbf{L}^2\to\mathbf{L}_H^2}^2 \,\mathrm{d}t\,\mathrm{d}s \\ &\quad\lesssim_\rho T^{-8a}(1+\|w\|_{\mathbf{H}})^{8b+2}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}^2.\end{aligned}$$

Therefore we have the desired estimate (6.7) on g_t. ∎

6.4. *Nondegeneracy of the partial Malliavin matrix*

For simplicity of presentation and brevity, we will only detail the proof in the case of nondegenerate noise on the Navier–Stokes equations (i.e., $L = 1$), that is,

$$|q_k| \approx |k|^{-\alpha} \quad \text{for all } k \in \mathbb{Z}_0^d.$$

Once one has the hypoellipticity deduced in Section 5, the adaptation to the weaker Assumption 2 is a well-understood extension using methods from previous works [44, 45, 63, 99, 100]. This is discussed in more detail in Remark 6.16 below.

Define the set

$$\mathbb{K} = (\mathbb{Z}_0^d \times \{1, \ldots, d-1\}) \cup \{1, 2, \ldots, 2d\}.$$

Note that each element $m \in \mathbb{K}$ is either a pair $(k, i) \in \mathbb{Z}_0^d \times \{1, \ldots, d-1\}$ or an integer $j \in \{1, \ldots, 2d\}$. We will also denote the set $\mathbb{K}_L$ in a similar way with $\mathbb{Z}_0^d$ replaced by K_L and define $\mathbb{K}_H = \mathbb{K} \setminus \mathbb{K}_L$. The operator $\hat{Q}$ on $\mathbf{L}^2 \times \mathbb{R}^{2d}$ gives rise to a family of vector fields $\{Q^m\}_{m \in \mathbb{K}}$ on $\mathbf{H} \times \mathcal{M}$ defined by

$$Q^m = \begin{cases} q_k e_k \gamma_k^i & \text{if } m = (k, i) \in \mathbb{Z}_0^d \times \{1, \ldots, d-1\}, \\ \hat{e}_j^z & \text{if } m = j \in \{1, \ldots, 2d\}, \end{cases}$$

where we are denoting $\{\hat{e}_j^z\}$ the canonical basis on $\mathbb{R}^{2d}$. The pivotal lemma is the following nondegeneracy of the partial Malliavin matrix $\mathcal{C}_t^L$.

Lemma 6.11. *For all $p \geq 1$, $t < 1$, $\varepsilon > 0$ and $w \in \mathbf{H} \times \mathcal{M}$, there exist constants $a, b > 1$ such that*

$$\sup_{\substack{h \in \mathbf{H}_L \times T_v \mathcal{M} \\ |h|=1}} \mathbf{P}\left(\sum_{m \in \mathbb{K}_L} \int_0^t \langle V_s^L Q^m, h \rangle_L^2 \, \mathrm{d}s < \varepsilon \right) \lesssim_{p,\rho} t^{-ap} (1 + |z|)^{bp} \varepsilon^p,$$

where the constant is independent of ε and the initial data.

Above $\langle \cdot, \cdot \rangle_L$ denotes the Riemannian metric on $\mathbf{H} \times \mathcal{M}$. We omit the dependence on $v \in \mathbb{S}^{d-1}$.

Note that

$$\sum_{m \in \mathbb{K}_L} \int_0^t \langle V_s^L Q^m(w_s), h \rangle_L^2 \, \mathrm{d}s = \langle h, \mathcal{C}_t^L h \rangle_L,$$

so that Lemma 6.11 is really about nondegeneracy of $\mathcal{C}_t^L$. It is a standard fact in the theory of Malliavin calculus that Lemma 6.8 is sufficient to deduce the moment bounds on $(\mathcal{C}_t^L)^{-1}$ stated in Lemma 6.11.

To begin, we will need the following lemma that relates time-derivatives of certain quantities to appropriate Lie brackets.

Proposition 6.12. *Let G be a bounded vector field on $\mathbf{H} \times \mathcal{M}$ whose range belongs to $\mathbf{H}_L \times T\mathcal{M}$ and with two bounded derivatives. Then the following formula holds:*

$$\begin{aligned} V_t^L G(w_t) = G(w) &+ \int_0^t V_L^s([F, G]_L(w_s) - [A, G]_L(w_s)) \, \mathrm{d}s \\ &+ \frac{1}{2} \sum_{m \in \mathbb{K}} \int_0^t V_s^L D^2 G(w_s)[Q^m, Q^m] \, \mathrm{d}s \\ &+ \int_0^t V_s^L DG(w_s) Q \mathrm{d}W_s, \end{aligned} \tag{6.23}$$

and for any two differentiable vector fields F, G over $\mathbf{H} \times \mathcal{M}$, we denote

$$[F, G]_L \equiv \Pi_L [F, G](w) = (DG_L)(w) F(w) - (DF_L)(w) G(w)$$

and

$$[A, G]_L(w) \equiv D_L G_L(w) A_L w - A_L G(w).$$

Proof. The proof follows from Itô's formula on $G(w_t)$ and the fact that V_t^L satisfies

$$V_t^L = \mathrm{Id} - \int_0^t V_s^L(D_L F_L(w_s) - A_L)\,\mathrm{d}s. \qquad \blacksquare$$

Remark 6.13. Note that since we assume that $\operatorname{Ran} G(w) \subseteq \mathbf{H}_L \times T_v\mathcal{M}$ and that the vector fields $\{Q^m\}_{m\in\mathbb{K}}$ have the property that $\operatorname{Ran} Q^m \subseteq \mathbf{H}_L \times T_v\mathcal{M}$ if $m \in \mathbb{K}_L$ and $\operatorname{Ran} Q^m \subseteq \mathbf{H}_H$ if $m \in \mathbb{K}_H$, then the sum above converges by the fact that the noise is of Hilbert–Schmidt type and therefore the sum over high frequencies can be bounded

$$\sum_{m\in\mathbb{K}_H} \|D^2 G(w)[Q^m, Q^m]\|_{\mathbf{H}_L\times T_v\mathcal{M}} \le \|D_H^2 G(w)\|_{\mathbf{H}_H\otimes\mathbf{H}_H\to\mathbf{H}_L\times T_v\mathcal{M}} \sum_{k\in K_H} q_k^2 < \infty.$$

For convenience, we define the following operator Λ_L that maps smooth vector fields on $\mathbf{H}\times\mathcal{M}$ to smooth vector fields on $\mathbf{H}\times\mathcal{M}$ with range in $\mathbf{H}_L \times T\mathcal{M}$, defined by

$$\Lambda_L G := [F, G]_L - [A, G]_L + \frac{1}{2}\sum_{m\in\mathbb{K}} D^2 G[Q^m, Q^m].$$

Lemma 6.14. *The following estimates hold for each $m \in \mathbb{K}_L$:*

$$|\Lambda_L Q^m|(w) \lesssim_\rho 1, \quad |\Lambda_L^2 Q^m|(w) \lesssim_\rho 1, \quad \sum_{j\in\mathbb{K}} |[Q^j, \Lambda_L Q^m]_L|^2(w) \lesssim_\rho 1.$$

Proof. The proof follows from the fact that below the cut-off $\|u\|_{\mathbf{H}} \le 6\rho$, we can bound

$$|[F, Q^m]_L| + |[A, Q^m]_L| \lesssim_\rho (1 + \|u\|_{\mathbf{H}}^2) \lesssim_\rho 1.$$

When $\|u\|_{\mathbf{H}} > 6\rho$, the Navier–Stokes nonlinearity is turned off and the above nonlinear term does not contribute, so we can just use the estimate $|[A, e_k\gamma_k^i]_L| \lesssim 1$. There are also terms which are nonlinear in z, however they are bounded and have bounded derivatives, so that $|[F, \hat{e}_j^z]_L| \lesssim 1$. The only other subtlety involves ensuring that the infinite sum in $m \in \mathbb{K}$ converges. However, this is due to the fact that the $m \in \mathbb{K}_L$ and the noise is Hilbert–Schmidt. $\blacksquare$

Lemma 6.15. *The uniform lower-bound*

$$\max\left\{|\langle Q^m, h\rangle_L|, |\langle \Lambda_L Q^m, h\rangle_L| : m \in \mathbb{K}_L\right\} \gtrsim_\rho \frac{|h|}{(1+|z|)^3} \tag{6.24}$$

holds for every initial data $w = (u, x, v, z) \in \mathbf{H}\times\mathcal{M}$, and $h \in \mathbf{H}_L \times T_v\mathcal{M}$.

Proof. To show (6.24), we must consider the different behaviors of

$$\langle \Lambda_L Q^m, h\rangle_L = \langle [F, Q^m]_L, h\rangle_L - \langle [A, Q^m]_L, h\rangle_L,$$

for different values of the initial data $w \in \mathbf{H}\times\mathcal{M}$ due to the presence of the cut-off. We divide the proof into two cases using a parameter $\delta \in (0, 1)$, which will be determined later.

Case 1. We first consider the case where $\chi_\rho(\|u\|_{\mathbf{H}}) > \delta$. This case is the easiest, since we can use the z process to help span the (x, v) directions. Indeed, notice that if we choose an $m \in \mathbb{K}_L$ so that $m = j \in \{1, \dots, 2d\}$, then $Q^m = \hat{e}^z_j$, then one easily computes for $j = 1, \dots, d$

$$|\langle \Lambda_L Q^m(w), h\rangle| = \frac{\chi_\rho(\|u\|_{\mathbf{H}})}{(1 + |z^j|^2)^{3/2}} |\langle \hat{e}_j, h\rangle_L| \geq \frac{\delta}{(1 + |z|)^3} |\langle \hat{e}_j, h\rangle_L|,$$

where $\{\hat{e}_j\}_{j=1}^d$ is the canonical basis for $\mathbb{R}^d$, taken here to be elements of $T_x\mathbb{T}^d \subseteq \mathbf{H}_L \times T_v\mathcal{M}$. Similarly for $j = d+1, \dots, 2d$, we have

$$|\langle \Lambda_L Q^k(w), h\rangle| \geq \frac{\delta}{(1 + |z|)^3} |\langle \Pi_v \hat{e}_{j-d}, h\rangle_L|$$

and $\{\Pi_v \hat{e}_j\}_{j=1}^d$ is a spanning set for $T_v\mathbb{S}^{d-1} \subseteq \mathbf{H}_L \times T_v\mathcal{M}$. Therefore we can easily conclude the lower bound

$$\max\bigl\{|\langle Q^m, h\rangle_L|, |\langle \Lambda_L Q^m, h\rangle_L| : m \in \mathbb{K}_L\bigr\} \gtrsim \delta \frac{|h|}{(1 + |z|)^3}.$$

Case 2. We now consider the case $\chi_\rho(\|u\|_{\mathbf{H}}) \leq \delta$. Here, we cannot rely on the regularization introduced by the z process since we are in a region where its coupling with x and v may be turned off or very small. Here, the drift is fully turned on and if we choose $m \in \mathbb{K}_L$ so that $m = (k, i)$ and $Q^m = q_k e_k \gamma_k^i$, we obtain

$$\begin{aligned}\Lambda_L Q^m(w) = {} & q_k[V_0(w), e_k\gamma_k^i] - q_k[B(u, u), e_k\gamma_k^i]_L - q_k[A, e_k\gamma_k^i]_L \\ & - q_k \frac{1}{\rho} \chi'(\|u\|_{\mathbf{H}}/\rho) \frac{u_k}{\|u\|_{\mathbf{H}}} H(v, z).\end{aligned}$$

Using the fact that we are in the region $\|u\|_{\mathbf{H}} \leq 2\rho$, we have

$$|\langle [A, e_k\gamma_k^i]_L, h\rangle_L| + |\langle [B(u, u), e_k\gamma_k^i]_L, h\rangle_L| \lesssim_\rho \sum_{i=1}^{d-1} \sum_{k \in K_L} |\langle e_k\gamma_k^i, h\rangle_L|, \tag{6.25}$$

additionally, since $\chi_\rho(\|u\|_{\mathbf{H}}) \leq \delta$,

$$\frac{1}{\rho} \chi'(\|u\|_{\mathbf{H}}/\rho) \frac{|u_k|}{\|u\|_{\mathbf{H}}} |\langle H(v, z), h\rangle_L| \lesssim_\rho \delta |h|.$$

This implies that

$$\delta|h| + |\langle \Lambda_L Q^m, h\rangle_L| + \sum_{i=1}^{d-1} \sum_{k \in K_L} q_k |\langle e_k\gamma_k^i, h\rangle_L| \gtrsim_\rho |\langle [V, e_k\gamma_k^i], h\rangle_L|,$$

which, in turn, implies that

$$\begin{aligned}& \delta|h| + \max\bigl\{|\langle Q^m, h\rangle_L|, |\langle \Lambda_L Q^m, h\rangle_L| : m \in \mathbb{K}_L\bigr\} \\ & \quad \gtrsim_\rho \max\bigl\{|\langle [V, e_k\gamma_k^i], h\rangle_L|, |\langle e_k\gamma_k^i, h\rangle_L| : k \in K_L,\ i \in \{1, \dots, d-1\}\bigr\}.\end{aligned}$$

Finally, an easy modification of Lemma 5.3 gives

$$\max\left\{|\langle [V, e_k\gamma_k^i], h\rangle_L|, |\langle e_k\gamma_k^i, h\rangle_L| : k \in K_L,\ i \in \{1, \dots, d-1\}\right\} \gtrsim |h|,$$

so that taking δ small enough (depending on ρ) we obtain the desired lower bound. ∎

We are now equipped to prove Lemma 6.11.

Proof of Lemma 6.11. Fix initial data $w \in \mathbf{H} \times \mathcal{M}$ and let $h \in \mathbf{H}_L \times T_v\mathcal{M}$ with $|h| = 1$, fix $t \in (0, 1)$. Denote for each $m \in \mathbb{K}_L$,

$$X_s^m \equiv \langle V_s^L Q^m, h\rangle_L.$$

It is sufficient to show that

$$\mathbf{P}\left(\bigcap_{m\in\mathbb{K}_L}\{\|X^m\|^2_{L^2([0,t])} < \varepsilon\}\right) \lesssim_{p,\rho} t^{-ap}(1+|z|)^{bp}\varepsilon^p, \tag{6.26}$$

where the constant does not depend on h or the initial data. Using Proposition 6.12, as well as Lemmas A.7, 6.14 and 6.21, we find that we have the almost-sure bound

$$[X^m]^2_{C^1([0,1])} \le C_\rho, \tag{6.27}$$

where $C_\rho \ge 1$ is a deterministic constant depending only on ρ. Applying the interpolation Lemma A.6 with $f = \int_0^\cdot X_s\,\mathrm{d}s$ and $\alpha = 1$, and then applying Cauchy–Schwarz, we arrive at the inequality

$$\|X^m\|_{L^\infty([0,t])} \le 4t^{-\frac{1}{2}}\|X^m\|^{\frac{1}{2}}_{L^2([0,t])} \cdot \max\left\{\|X^m\|^{\frac{1}{2}}_{L^2([0,t])}, [X^m]^{\frac{1}{2}}_{C^1([0,1])}\right\}. \tag{6.28}$$

Therefore, we can deduce

$$\mathbf{P}\left(\bigcap_{m\in\mathbb{K}_L}\{\|X^m\|^2_{L^2([0,t])} < \varepsilon\}\right) \le \mathbf{P}\left(\bigcap_{m\in\mathbb{K}_L}\{\|X^m\|_{L^\infty([0,t])} < 4C_\rho t^{-\frac{1}{2}}\varepsilon^{\frac{1}{4}}\}\right).$$

Next, using Proposition 6.12, we write

$$X_s^m = X_0^m + \int_0^s B_r^m\,\mathrm{d}r,$$

where B_s^m is the $\mathbb{R}$-valued adapted process defined by

$$B_s^m := \langle V_s^L \Lambda_L Q^m(w_s), h\rangle_L.$$

This means that when $\|X^m\|_{L^\infty(0,t)} < 4C_\rho t^{-1/2}\varepsilon^{1/4}$, then one also has

$$\left|\int_0^s B_r^m\,\mathrm{d}r\right| \le 8C_\rho t^{-\frac{1}{2}}\varepsilon^{\frac{1}{4}}.$$

Applying Lemma A.6 again with $f = \int_0^\cdot B_s^m\,\mathrm{d}s$ and $\alpha = \frac{1}{3}$, we find

$$\|B^m\|_{L^\infty([0,t])} \le 4t^{-1}\|f\|^{1/4}_{L^\infty([0,t])} \times \max\left\{\|f\|^{3/4}_{L^\infty([0,t])}, [B^m]^{3/4}_{C^{1/3}([0,1])}\right\}. \tag{6.29}$$

Again an application of Proposition 6.12 to $f(s)$ along with the bounds given by Lemmas 6.14 and 6.21 easily gives, by Lemma A.7, the moment estimate on the Hölder seminorm of B^m_s for each $p \geq 1$,

$$\mathbf{E}[B^m]^p_{C^{1/3}([0,1])} \lesssim_{p,\rho} 1. \tag{6.30}$$

Since estimate (6.27) implies that for each $p \in (1,\infty)$ and every $\varepsilon \in (0,1)$,

$$\mathbf{P}\big([B^m]_{C^{1/3}([0,1])} \geq 8C_\rho t^{-\frac{1}{2}} \varepsilon^{-\frac{1}{204}}\big) \lesssim_{p,\rho} \varepsilon^p,$$

we can with overwhelming probability restrict ourselves to the event

$$\bigcap_{m\in\mathbb{K}_L} \big\{[X^m]_{C^{1/3}([0,1])} < 8C_\rho t^{-\frac{1}{2}} \varepsilon^{-\frac{1}{204}}\big\}.$$

The choice of the exact power for $\varepsilon^{-1/204}$ above is somewhat arbitrary and is chosen simply to give rise to the power of $\varepsilon^{1/18}$ in the final inequality (6.32). It is certainly possible to use other powers on ε without changing the essence of the proof.

Using inequality (6.29), we conclude that for every $p \geq 1$,

$$\begin{aligned}
&\mathbf{P}\Big(\bigcap_{m\in\mathbb{K}_L} \big\{\|X^m\|^2_{L^2([0,t])} < \varepsilon\big\}\Big) \\
&\quad \lesssim_{p,\rho} \mathbf{P}\Big(\bigcap_{m\in\mathbb{K}_L} \big\{\|X^m\|_{L^\infty([0,t])} < 4C_\rho t^{-\frac{1}{2}} \varepsilon^{\frac{1}{4}}\big\} \\
&\qquad\qquad \cap \big\{\|B^m\|_{L^\infty([0,t])} < 32C_\rho t^{-\frac{3}{2}} \varepsilon^{\frac{1}{17}}\big\}\Big) + \varepsilon^p.
\end{aligned} \tag{6.31}$$

By choosing $\varepsilon \lesssim t^a$ small enough for a large enough constant $a > 1$, we can remove the factor of $t^{-1/2}$ and $t^{-3/2}$ above at the expense of a slightly worse power on ε. To remove this t-dependent restriction on ε, we can treat the case $t^a \lesssim \varepsilon$ by simply using the fact that probabilities are bounded by 1 and that $1 \lesssim t^{-ap}\varepsilon^p$ to deduce that for all $\varepsilon \in (0,1)$ and $p \geq 1$,

$$\begin{aligned}
&\mathbf{P}\Big(\bigcap_{m\in\mathbb{K}_L} \big\{\|X^m\|^2_{L^2([0,t])} < \varepsilon\big\}\Big) \\
&\quad \lesssim_{p,\rho} \mathbf{P}\Big(\bigcap_{m\in\mathbb{K}_L} \big\{\|X^m\|_{L^\infty([0,t])} < \varepsilon^{\frac{1}{5}}\big\} \\
&\qquad\qquad \cap \big\{\|B^m\|_{L^\infty([0,t])} < \varepsilon^{\frac{1}{18}}\big\}\Big) + t^{-ap}\varepsilon^p.
\end{aligned} \tag{6.32}$$

Next, we show that for small enough ε, and each initial data $w \in \mathbf{H} \times \mathcal{M}$,

$$\bigcap_{m\in\mathbb{K}_L} \big\{|X^m_0| \leq \varepsilon\big\} \cap \big\{|B^m_0| < \varepsilon^{r_*}\big\} = \emptyset, \tag{6.33}$$

where r^* is some number less than 1. That is, at time $t = 0$ for small enough ε, it is not possible for all the $\{X^m\}$ and all the $\{B^m\}$ to be small. Indeed, since $X^m_0 = \langle Q^m, h\rangle_L$ and $B^m_0 = \langle \Lambda_L Q^m, h\rangle_L$, this follows from Lemma 6.15 since the estimates $|\langle Q^m, h\rangle_L| < \varepsilon$

and $|\langle \Lambda_L Q^m, h\rangle_L| < \varepsilon^{r^*}$ imply by (6.24) that

$$1 \lesssim_\rho (1+|z|)^3 \varepsilon^{r_*}.$$

Therefore choosing ε small enough so that $\varepsilon \lesssim_\rho (1+|z|)^{-b}$ for a sufficiently large constant $b > 0$ we deduce a contradiction and conclude that (6.33) must hold. Again, to remove the z-dependent restriction on ε we can replace ε by $(1+|z|)^b\varepsilon$ on the right-hand side of estimate (6.32), giving our desired estimate (6.26). ■

Remark 6.16. In order to treat noise as in Assumption 2, one needs to adjust the above proof in two ways. First, in the definition of the cutoff process (6.2), one needs to add additional Brownian motions to the modes k in (u_t) for which $k \notin \mathcal{K}$, in the same manner as was done for the Lagrangian flow, that is, $\chi_\rho(\|u\|_{\mathbf{H}}) e_k \gamma_k^i z_{k,i}/(1+|z_{k,i}|^2)^{1/2}$ for $k \notin \mathcal{K}$. Then in the proof of Lemma 6.11, for $\chi_\rho(\|u\|_{\mathbf{H}}) < \delta$, one needs to use Lie brackets of the Navier–Stokes nonlinearity to fill the missing degrees of freedom in Navier–Stokes (these brackets are computed for 2D and 3D, respectively, in [44, 99]; see also Section 5). This requires taking one more time derivative in the proof of Lemma 6.11 (allowing noise from the high frequencies to propagate to the lower modes), which in turn, requires the use of a version of Norris' Lemma [94] (in addition to Lemma A.6), as described in, e.g., [63]. Analogous to [45, 100], one needs to slightly refine the statement found in, e.g., [63] to handle the singularity for short-times but this is a straightforward calculation.

6.5. Basic estimates on Jacobians and Malliavin derivatives

The proofs of the following lemmas are standard and are omitted for brevity (see [38]).

Lemma 6.17. *The statements of Proposition A.3 hold for the (w_t) process. We record the quantitative estimates here for the readers' convenience. For all $\gamma < \alpha - \frac{d}{2}$, $T \le 1$, and $p \in [2,\infty)$ there holds*

$$\mathbf{E} \sup_{t\in[0,T]} \|w_t\|_{H^\gamma}^p \lesssim_{p,\gamma,\rho} 1 + \|w_0\|_{H^\gamma}^p,$$

$$\mathbf{E}\int_0^T \|w_s\|_{H^{\gamma+(d-1)}}^2 \,\mathrm{d}s \lesssim_{\gamma,\rho} 1 + \|w_0\|_{H^\gamma}^2.$$

We also need the following improved short-time regularization estimates. Specifically, for regularities all the way up to $\gamma < \sigma + (d-1)$. This is crucial for dealing with the high frequencies of the control.

Lemma 6.18. *For all $\gamma \in (\sigma, \sigma+(d-1))$, $p \in [2,\infty)$, and $T \le 1$ there holds for all $\delta > 0$,*

$$\mathbf{E}\Big(\sup_{t\in[0,T]} t^{\frac{\gamma-\sigma}{2(d-1)}} \|w_t\|_{H^\gamma}\Big)^p \lesssim_p 1 + \|w_0\|_{H^\sigma}^p, \tag{6.34}$$

$$\mathbf{E}\int_t^T \|w_s\|_{H^{\gamma+(d-1)}}^2 \,\mathrm{d}s \lesssim_\delta 1 + t^{-2r}\|w_0\|_{H^\sigma}^4, \tag{6.35}$$

where

$$r = \frac{\sigma - (\gamma + 2 - d + 2(d-1)\delta)}{2(d-1)} > 0.$$

Lemma 6.19. *The following properties are satisfied for $J_{s,t}$ and $U^H_{s,t}$ for $0 < s < t < T \le 1$:*

(i) *There holds for $\gamma \le \sigma$, (almost surely)*

$$\|J_{s,t}h\|_{H^\gamma \times T_{v_t}\mathcal{M}} + \|U^H_{s,t}h^H\|_{H^\gamma} \lesssim_\rho \|h\|_{H^\gamma \times T_{v_s}\mathcal{M}},$$
$$\int_0^T \|J_{s,t}h\|^2_{H^{\gamma+(d-1)}\times T_{v_t}\mathcal{M}}\,\mathrm{d}t + \int_0^T \|U^H_{s,t}h^H\|^2_{H^{\gamma+(d-1)}}\,\mathrm{d}t \lesssim_\rho \|h\|^2_{H^\gamma \times T_{v_s}\mathcal{M}}.$$

(ii) *For all $\gamma \in (\sigma, \sigma+(d-1))$ there holds (almost surely),*

$$(t-s)^{\frac{\gamma-\sigma}{2(d-1)}}\|J_{s,t}h\|_{H^\gamma\times T_{v_t}\mathcal{M}} + (t-s)^{\frac{\gamma-\sigma}{2(d-1)}}\|U^H_{s,t}h^H\|_{H^\gamma} \lesssim_{\rho,T,\delta} \|h\|_{H^\sigma\times T_{v_s}\mathcal{M}}.$$

(iii) *For all $\gamma \in (\sigma, \sigma+(d-1))$ and all δ sufficiently small*

$$\begin{aligned}&\mathbf{E}\int_{s+s'}^T \|J_{s,t}h\|^2_{H^{\gamma+2(d-1)}\times T_{v_t}\mathcal{M}}\,\mathrm{d}t + \mathbf{E}\int_{s+s'}^T \|U^H_{s,t}h^H\|^2_{H^{\gamma+2(d-1)}}\,\mathrm{d}t \\ &\lesssim_\delta (s')^{-2r}(1+\|w_0\|_{H^\sigma})^2\|h\|^2_{H^\sigma\times T_{v_s}\mathcal{M}},\end{aligned} \tag{6.36}$$

where

$$r = \frac{\sigma - (\gamma + 2 - d + 2(d-1)\delta)}{2(d-1)} > 0.$$

Remark 6.20. Note that the above estimates all hold almost surely and are independent of w_0 except for (6.36). This is because only (6.36) requires regularities above σ on the (linearization of) the nonlinear term.

Lemma 6.21. *For each $p \ge 1$ an $T \le 1$, the processes U^L_t and V^L_t satisfy the following bounds:*

$$\sup_{t\in[0,T]}(|U^L_t| + |V^L_t|) \lesssim_{\rho,p} 1,$$

and the constants do not *depend on the initial data for w_t.*

We also require the following estimates on the Jacobian, as in [45], which control the effect of low frequencies on high frequencies and vice-versa.

Lemma 6.22. *For each $T \le 1$, $h^L \in \mathbf{H}_L \times T_v\mathcal{M}$ and $h^H \in \mathbf{H}_H$ we have the almost sure bounds*

$$\begin{aligned}\sup_{0<t<T}\|D_H w^H_t h^L\|_{\mathbf{H}_H} &\lesssim_\rho T^{\frac{1}{2}}|h_L|,\\ \sup_{0<t<T}|D_H w^L_t h^H| &\lesssim_\rho T\|h_H\|_{\mathbf{H}_H},\end{aligned} \tag{6.37}$$

where the constants do not depend on the initial data w.

Proof. Consider the case of $D_L w^H$. In this case we have

$$\partial_t(D_L w^H_t h^L) = D_H F_H(w_t) D_L w^H_t h^L + D_L F_H(w_t) D_L w^L_t h^L - A_H(D_L w^H_t h^L)$$

and $D_L w^H_0 h^L = 0$. Therefore

$$D_L w^H_t h^L = \int_0^t U^H_{s,t} D_L F_H(w_s) D_L w^L_s h^L\,\mathrm{d}s.$$

By Lemma 6.19,

$$\begin{aligned}\|D_L w_t^H h^L\|_{\mathbf{H}_L} &\lesssim \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}} \|D_L F_H(w_s) D_L w_s^L h^L\|_{H^{\sigma-1}}\,\mathrm{d}s \\ &\lesssim_\rho \int_0^t \frac{1}{(t-s)^{\frac{1}{2}}}\,\mathrm{d}s \Big(\sup_{0<s<t} \|J_{0,s} h^L\|_{\mathbf{H}\times T_{v_s}\mathcal{M}} \Big) \\ &\lesssim \sqrt{t}|h^L|.\end{aligned}$$

The estimate on (6.37) follows similarly (except no smoothing is necessary). ■

Next, we compute and estimate the Malliavin derivatives of the necessary quantities. First, we compute

$$\begin{aligned}\mathcal{D}_s w_t f &= J_{s,t} Q f, \\ \mathcal{D}_s(U_{r,t}^L h) f &= \int_r^t U_{l,t}^L \bar{D}^2 F_L(w_l)[U_{r,l}^L h, J_{s,l} Q f]\,\mathrm{d}l, \\ \mathcal{D}_s(U_{r,t}^H h) f &= \int_r^t U_{l,t}^H \bar{D}^2 F_H(w_l)[U_{r,l}^H h, J_{s,l} Q f]\,\mathrm{d}l,\end{aligned}$$

where $\bar{D}^2 F$ denotes the full second variation of F extended to the linear space $\mathbf{H}_L \times \mathbb{R}^{4d}$. We further have

$$\begin{aligned}\mathcal{D}_s D w_t h f &= \int_0^t J_{r,t} \bar{D}^2 F(w_r)[\mathcal{D}_s w_r f, J_{0,r} h]\,\mathrm{d}r \\ &= \int_s^t J_{r,t} \bar{D}^2 F(w_r)[J_{s,r} Q f, J_{0,r} h]\,\mathrm{d}r.\end{aligned}$$

Furthermore, one has the following for the derivatives of the inverse Malliavin matrix and V_t^L:

$$\mathcal{D}_s(\mathcal{C}_T^L)^{-1} f = -(\mathcal{C}_T^L)^{-1}[\mathcal{D}_s \mathcal{C}_T^L f](\mathcal{C}_T^L)^{-1}, \quad \mathcal{D}_s V_t^L f = -V_t^L[\mathcal{D}_s U_t^L f] V_t^L.$$

Lemma 6.23. *The following estimates hold almost surely for $T \le 1$ (and are independent of $\|w_0\|_{\mathbf{H}}$):*

$$\begin{aligned}\sup_{0<r<t<T} |\mathcal{D}_s U_{r,t}^L h^L|_{\mathbf{L}^2 \to \mathbf{H}_L \times T_{v_t}\mathcal{M}} &\lesssim_\rho t \|h^L\|_{\mathbf{H}_L \times T_{v_r}\mathcal{M}}, \\ \sup_{0<r<t<T} |\mathcal{D}_s V_{r,t}^L h^L|_{\mathbf{L}^2 \to \mathbf{H}_L \times T_{v_t}\mathcal{M}} &\lesssim_\rho t \|h^L\|_{\mathbf{H}_L \times T_{v_r}\mathcal{M}}, \\ \sup_{0<r<t<T} \|\mathcal{D}_s U_{r,t}^H h^H\|_{\mathbf{L}^2 \to \mathbf{H}_H} &\lesssim_\rho t^{\frac{1}{2}} \|h^H\|_{\mathbf{H}_L \times T_{v_r}\mathcal{M}}, \\ \sup_{0<r<t<T} \|\mathcal{D}_s J_{r,t} h\|_{\mathbf{L}^2 \to \mathbf{H} \times T_{v_t}\mathcal{M}} &\lesssim_\rho t^{\frac{1}{2}} \|h\|_{\mathbf{H}_L \times T_{v_r}\mathcal{M}}.\end{aligned}$$

Proof. By using the formula above, the case of $\mathcal{D}_s U_{r,t}^L$ follows immediately from Lemma 6.21. The case of U^H follows from (noting that $\sigma < \alpha - \frac{d}{2}$ and that $Q : \mathbf{L}^2 \to \mathbf{H} \times \mathcal{M}$

is bounded)

$$\|\mathcal{D}_s(U^H_{r,t}h)f\|_{\mathbf{H}_H} \lesssim \int_r^t \frac{1}{(t-l)^{\frac{1}{2}}}\|U^H_{r,l}h\|_{\mathbf{H}_H}\|J_{s,l}Qf\|_{\mathbf{H}\times T_{v_l}\mathcal{M}}\,\mathrm{d}l$$
$$\lesssim \sqrt{t}\|h\|_{\mathbf{H}\times T_v\mathcal{M}}\|f\|_{\mathbf{L}^2}.$$

Consider next estimating $\mathcal{D}_s Dw_t hf$. For this we get (almost surely due to the cutoff)

$$\|\mathcal{D}_s(Dw_t h)f\|_{\mathbf{H}_H} \lesssim_\rho \int_s^t \frac{1}{(t-r)^{\frac{1}{2}}}\|J_{s,r}Qf\|_{\mathbf{H}\times T_{v_r}\mathcal{M}}\|J_{0,r}h\|_{H^\sigma}\,dr \lesssim t^{\frac{1}{2}}. \qquad \blacksquare$$

Lemma 6.24. *The following holds for all $s < T$ and $1 \le p < \infty$ (the constants a, b are from Lemma 6.8):*

$$\mathbf{E}\|\mathcal{D}_s(\mathcal{C}^L_T)^{-1}\|^p_{\mathbf{L}^2\to\mathbf{H}_L\times T_v\mathcal{M}} \lesssim_p (T^{-2a+1}(1+|z|)^{2b})^p.$$

Proof. Follows by Lemma 6.21 and Lemma 6.8. ∎

7. Weak irreducibility and approximate control

First, we prove Proposition 2.15, hence deducing the weak irreducibility of the stationary measures for the Markov processes (u_t, x_t), (u_t, x_t, v_t), $(u_t, x_t, \check{v}_t)$. Combined with the strong Feller property, this yields unique stationary measures for these processes by the Doob–Khasminskii Theorem [41,72].

Lemma 7.1. *Recall the control problem (2.9) for Systems 3–4. Suppose that $\mathcal{K}$ is symmetric and $(1,0), (0,1) \in \mathcal{K}$ in 2D and $(1,0,0), (0,1,0), (0,0,1) \in \mathcal{K}$ in 3D. Let (x, v), (x', v') be arbitrary points in $\mathbb{T}^d \times \mathbb{S}^{d-1}$. Then there exists a smooth control Qg such that*

$$(u_0, x_0, v_0) = (0, x, v), \quad (u_1, x_1, v_1) = (0, x', v').$$

Furthermore, g can be chosen to depend smoothly on x, x', v, v' and supported only in frequencies $|k|_\infty \le 1$. All of the above holds also for the $(u_t, x_t, \check{v}_t)$ process.

Remark 7.2. By choosing arbitrary representatives on $\mathbb{S}^{d-1}$, it is clear that controlling the (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ processes, regarding $v_t, \check{v}_t$ as elements on $\mathbb{S}^{d-1}$, implies controllability of the processes when considered on P^{d-1}.

Proof. First, let us consider the two-dimensional case. Let $x = (a_0, b_0)$ and $x' = (a_1, b_1)$. For $t \in (0, 1/4)$, suppose the velocity field is given by the shear flow

$$u_t(y_1, y_2) = f_a(t)\begin{pmatrix}\cos(y_2 - b_0)\\ 0\end{pmatrix},$$

such that $f_a \in C_c^\infty(0, 1/4)$ and $\int_0^{1/4} f_a(t)\,\mathrm{d}t = a_1 - a_0$. Similarly, for $t \in (1/4, 1/2)$, suppose the velocity field was the shear flow

$$u_t(y_1, y_2) = f_b(t)\begin{pmatrix}0\\ \cos(y_1 - a_1)\end{pmatrix},$$

such that $f_b \in C_c^\infty((1/4, 1/2))$ and $\int_{1/4}^{1/2} f_b(t)\,\mathrm{d}t = b_1 - b_0$. It follows that the solution to the ODE (1.9a) satisfies $x_1 = (a_1, b_1)$.

Next, we explain how to set g in order to produce these flows. Notice that the shear flows $(\cos(y - b_0), 0)$ and $(0, \cos(x - a_1))$ are stationary solutions of 2D Euler: the nonlinearity vanishes on these flows. Hence, it suffices to control the Stokes flow, which gives the following control:

$$Qg(t) = (f_a'(t) + f_a(t)) \begin{pmatrix} \cos(y_2 - b_0) \\ 0 \end{pmatrix} + (f_b'(t) + f_b(t)) \begin{pmatrix} 0 \\ \cos(y_1 - a_1) \end{pmatrix}.$$

By the angle-difference formula and the assumptions on $\mathcal{K}$, g satisfies the requisite properties.

Next, we augment the previous control also to deal with v_t; the treatment for $\check{v}_t$ is analogous and is omitted for brevity. During this time we have moved v_t some amount, let $v_{1/2}$ be the new value. Suppose that the velocity field were given by the cellular flow

$$u(t, y_1, y_2) = f_v(t) \begin{pmatrix} -\sin(y_2 - b_1) \\ \sin(y_1 - a_1) \end{pmatrix} \tag{7.1}$$

such that $f_v \in C_c^\infty((1/2, 1))$ with

$$\int_{1/2}^{1} f_v(t)\,\mathrm{d}t = \angle(v', v_{1/2}),$$

where $\angle(v', v_{1/2}) = \cos^{-1} \langle v', v_{1/2} \rangle$ denotes the angle between v' and $v_{1/2}$. This induces a rotation of v_t by the angle $\angle(v', v_{1/2})$ via (1.9b) into the desired final point without moving x_t.

As above, the cellular flow is both a stationary solution of the 2D Euler equations and an eigenfunction of the Stokes operator. Therefore, it suffices to set g on $t \in (1/2, 1)$ to be such that

$$Qg(t) = (f_v'(t) + f_v(t)) \begin{pmatrix} -\sin(y_2 - b_1) \\ \sin(y_1 - a_1) \end{pmatrix}.$$

This completes the proof in 2D.

Next, consider the 3D argument. It is clear that a similar proof applies to the (u_t, x_t) process by utilizing 2D shear flows aligned with any of the three Cartesian directions. For the (u_t, x_t, v_t) process, we consider the problem of controlling the v_t process (as an element of $\mathbb{S}^2$) from one arbitrary position $v \in \mathbb{S}^2$ to another $v' \in \mathbb{S}^2$ without moving x_t using 2D cellular flows aligned with any of the three Cartesian directions. Each of these flows induces rotation along curves of constant “latitude” aligned with one of the three Cartesian directions. Note that no flow gives lines of constant longitude in any direction. Arbitrarily, set the x, y plane to be the equatorial plane relative to which we assign latitude and longitude. Using the cellular flow that is constant in z, adjust the longitude of v_t so that $v_{1/3}$ lies in the y, z plane. Then, using a cellular flow that is constant in x, adjust the latitude so that $v_{2/3}$ lies at the latitude of v'. Finally, by re-applying the cellular flow that is constant in z, adjust the longitude so that $v' = v_1$. ■

The controllability provided in Lemma 7.1 implies the following nondegeneracy of the Markov transition kernels.

Lemma 7.3. *For all $t > 0$ and $\varepsilon > 0$, there exists $\varepsilon' > 0$ such that for all $(x, v), (x', v') \in \mathbb{T}^d \times \mathbb{S}^{d-1}$ and all $u \in B_{\varepsilon'}(0)$,*

$$\mathbf{P}((u_t, x_t) \in B_\varepsilon(0) \times B_\varepsilon(x') | (u_0, x_0) = (u, x)) > 0,$$
$$\mathbf{P}((u_t, x_t, v_t) \in B_\varepsilon(0) \times B_\varepsilon(x') \times B_\varepsilon(v') \mid (u_0, x_0, v_0) = (u, x, v)) > 0,$$
$$\mathbf{P}((u_t, x_t, \check{v}_t) \in B_\varepsilon(0) \times B_\varepsilon(x') \times B_\varepsilon(v') \mid (u_0, x_0, v_0) = (u, x, v)) > 0.$$

Proof. Such nondegeneracy properties normally follow from standard perturbation arguments. However, one must be somewhat careful with the regularity, as we require that $\sigma \in (\alpha - 2(d-1), \alpha - \frac{d}{2})$ (i.e., close to the highest available regularity). Let us treat the (u_t, x_t) process; the (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ processes are the same. Let Qg be a control given as in Lemma 7.1 corresponding to the desired endpoints x, x'. Let u_t^c be the controlled solution from Lemma 7.1. The first step is to prove that for all ε, there holds

$$\mathbf{P}(\|u_t - u_t^c\|_{L^\infty([0,1];\mathbf{H})} \lesssim \varepsilon) > 0. \tag{7.2}$$

Note that the control is built from only $\Pi_{\leq 1} Qg$. By the regularity of the stochastic convolution Γ_t (see Lemma A.4) and positivity of the Wiener measure, for all $\varepsilon > 0$,

$$\mathbf{P}\left(\sup_{t \in (0,1)} \|\Gamma_t - \int_0^t e^{-(t-s)A} Qg_s \, \mathrm{d}s\|_{L^\infty([0,1];\mathbf{H})} < \varepsilon \right) > 0. \tag{7.3}$$

Let u_t be a solution to the stochastic Navier–Stokes with a sample path ω such that the event in (7.3) holds. Then from the mild form

$$\begin{aligned} u_t - u_t^c = e^{-tA} u_0 &+ \int_0^t e^{-(t-s)A} (B(u_s, u_s) - B(u_s^c, u_s^c)) \, \mathrm{d}s \\ &+ \Gamma_t - \int_0^t e^{-(t-s)A} Qg_s \, \mathrm{d}s \end{aligned}$$

(actually by our choice of control $B(u_s^c, u_s^c) = 0$). By a generalized Grönwall's inequality [82, Lemma A.2] and parabolic smoothing, we have

$$\|u_t - u_t^c\|_{L^\infty([0,1];\mathbf{H})} \leq K'\varepsilon$$

for a universal constant K' depending only on σ, α (provided that $\|u_0\|_{\mathbf{H}} \leq \varepsilon$). Therefore, we have (7.2). For the x_t process, we similarly let x_t and x_t^c be the trajectories associated with the controlled system and that of the sample path ω (respectively). Then (viewing x_t, x_t^c as elements in $\mathbb{R}^d$)

$$\frac{\mathrm{d}}{\mathrm{d}t}(x_t^c - x_t) = u_t^c(x_t^c) - u_t(x_t) = (u_t^c(x_t^c) - u_t^c(x_t)) + (u_t^c(x_t) - u_t(x_t)).$$

We then obtain by the stability of the (u_t) process (by potentially adjusting K' and using $\sigma > \frac{d}{2} + 1$ to apply Sobolev embedding to ∇u),

$$\mathbf{P}\big(\{\|u_1\|_{\mathbf{H}} \leq K'\varepsilon\} \cap \{d(x_1, x') < K'\varepsilon\}\big) > 0.$$

The desired nondegeneracy for the Markov transition kernel then follows. ∎

Proof of Proposition 2.15. We prove this in the case of (u_t, x_t); the processes including P^{d-1} are the same. First, we verify irreducibility of stationary measures of the (u_t) process in H^σ. In the case L^2 this is well known; see, e.g., [44]. This can be proved by observing that if there were no forcing we have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u_t\|_{L^2}^2 \leq -\|\nabla u_t\|_{L^2}^2 \lesssim -\|u_t\|_{L^2}^2.$$

At the same time, in the absence of forcing, standard energy estimates give the uniform bound with $\delta > 0$, $\|u_t\|_{H^{\sigma+\delta}} \lesssim_\delta \|u_0\|_{H^{\sigma+\delta}}$ with an implicit constant that is *independent of time*. Hence, Sobolev interpolation gives $\|u_t\|_{H^\sigma} \lesssim \|u_0\|_{H^{\sigma+\delta}} e^{-ct}$, for some constant c depending only on σ, δ.

Let $\tilde{\mu}$ be an arbitrary stationary measure supported on $\mathbf{H} \times \mathbb{T}^d$. By the parabolic smoothing (see, e.g., (A.4)) and stationarity, $\tilde{\mu}$ is also supported on $H^{\sigma+\delta}$ for $0 < \delta < \alpha - \frac{d}{2} - \sigma$. Therefore, there exists a $C > 0$ such that

$$\tilde{\mu}(\{\|u\|_{H^{\sigma+\delta}} \leq C\} \times \mathbb{T}^d) > \frac{1}{2}.$$

Denote the set $\mathcal{B} = \{u \in \mathbf{H} : \|u\|_{H^{\sigma+\delta}} \leq C\} \times \mathbb{T}^d \subset \mathbf{H} \times \mathbb{T}^d$. The stability argument applied in Lemma 7.3 (with $g \equiv 0$) gives the desired uniform decay: for all γ, there exists a T_γ such that for all $(u, x) \in \mathcal{B}$,

$$\mathbf{P}\big((u_{T_\gamma}, x_{T_\gamma}) \in B_\gamma(0) \times B_\gamma(x') \mid (u_0, x_0) = (u, x)\big) > 0.$$

Next, it follows from Lemma 7.3 that for γ' sufficiently small, there exists a γ (depending only γ') such that for any $x' \in \mathbb{T}^d$, and all $(u, x) \in \mathcal{B}$,

$$\mathbf{P}\big((u_{T_\gamma+1}, x_{T_\gamma+1}) \in B_{\gamma'}(0) \times B_{\gamma'}(x') \mid (u_0, x_0) = (u, x)\big) > 0.$$

Since this implies that

$$\tilde{\mu}(B_{\gamma'}(0) \times B_{\gamma'}(x')) \geq \int_{\mathcal{B}} P_{T_\gamma+1}((u, x), B_{\gamma'}(0) \times B_{\gamma'}(x'))\tilde{\mu}(\mathrm{d}u, \mathrm{d}x) > 0,$$

it follows that $(0, x')$ is in the support of the stationary measure. ■

Next, in order to complete the proof of Theorem 1.6 in the case of Systems 3–4, it suffices to prove the following, which shows that arbitrarily large gradient growth can be obtained on the unit time interval.

Proposition 7.4. *For all $M > 0$ and $\varepsilon > 0$,*

$$\begin{aligned}\mathbf{P}\big((u_1, x_1, A_1) \in B_\varepsilon(0) \times B_\varepsilon(0) \times \{A \in \mathrm{SL}_d(\mathbb{R}) : |A| > M\} \mid \\ (u_0, x_0, A_0) = (0, 0, \mathrm{Id})\big) > 0.\end{aligned} \tag{7.4}$$

Together with Lemma 7.3, this implies that Systems 3–4 satisfy Definition 4.16 and hence Proposition 4.17 applies and the proof of Theorem 1.6 is completed.

Proof. The control step is proved as in Lemma 7.1, except now we apply the cellular flow translated so that the hyperbolic point is at the origin:

$$u(t) = f_+ \begin{pmatrix} \sin(y_2 - b) \\ \sin(y_1 - a) \end{pmatrix},$$

with $\int_0^1 f_+(s)\,\mathrm{d}s = \log M$. Then set g analogous to the choices in Lemma 7.1 (the size of g now depends on M). The stability step proceeds as in Lemma 7.3. ∎

Remark 7.5. All of the above controllability arguments also apply to the System 1 in $\mathbb{T}^2$ with only the condition: $\mathcal{K}$ symmetric and $(1,0),(0,1) \in \mathcal{K}$. This condition is not enough to guarantee that the (u_t, x_t, A_t) process satisfies Hörmander's condition. We can still verify Definition 4.16 in this case, and hence it is sufficient to deduce Theorem 1.6. The claim in Remark 1.13 follows. Further, our arguments on Navier–Stokes similarly apply to the System 1 in $\mathbb{T}^d$ with infinitely many modes forced, under Assumption 2.

Remark 7.6. For Systems 3–4, using higher frequency shear flows and cellular flows, one can make all the same arguments in this section if we only take Assumption 2. Hence, by also Remark 6.16, we can prove Theorem 1.6 (and all our other results) for Systems 3–4 using *only* Assumption 2.

8. Applications to scalar turbulence

In this section we prove Theorem 1.16. First, we prove the weak anomalous dissipation property (1.7), Theorem 1.16, part (i). For this, we adapt the compactness-contradiction method of [17]. Hence, it is easiest to begin by defining $f^\kappa = \sqrt{\kappa} g$ as in (2.11) and recall the re-scaled balance relation (2.12). Next, we are interested in studying the limits of stationary measures $\bar{\mu}^\kappa$ to the problem (2.11) coupled with any of Systems 1–4. It is standard that this (one-way) coupled system is well posed in the sense of Proposition 1.3 and defines an $\mathcal{F}_t$-adapted, Feller Markov process; see, e.g., [86]. Similarly, the Krylov–Bogoliubov method implies the following:

Lemma 8.1. *For all $\kappa > 0$, there exists a stationary probability measure $\bar{\mu}^\kappa$ for the Markov process (u_t, f_t^κ) supported on $\mathbf{H} \times H^1$. Furthermore, the measure satisfies the following for all $p \geq 2$ (with implicit constant independent of κ),*

$$\int_{\mathbf{H}\times H^1} \|\nabla f\|_{L^2}^2 \,\mathrm{d}\bar{\mu}^\kappa(u,f) = \bar{\varepsilon}, \tag{8.1}$$

$$\int_{\mathbf{H}\times H^1} \|f\|_{L^2}^p \,\mathrm{d}\bar{\mu}^\kappa(u,f) \lesssim_p \bar{\varepsilon}^{\frac{p}{2}}. \tag{8.2}$$

The following lemma is a straightforward adaptation of arguments in [17, 83, 86]. Unlike in [17], the velocity field is not bounded a.s., however, the situation is not significantly different (using Proposition 1.3); indeed, the original arguments of Kuksin [83] were specifically on the Navier–Stokes equations (see also [85, 86]).

Lemma 8.2. *Let $\{\bar{\mu}^\kappa\}_{\kappa>0}$ be a family of stationary probability measure of problem (2.11) as in Lemma 8.1, indexed by the diffusivity parameter κ, and (u_t) given by one of Systems 1–4. Then, the measures $\{\bar{\mu}^\kappa\}_{\kappa>0}$ are tight on $\mathbf{H} \times L^2$ as $\kappa \to 0$ and the subsequential weak limit $\bar{\mu}^0$ is a stationary measure of the inviscid problem (1.5) with*

$$\mu(A) = \bar{\mu}^0(A \times H^1)$$

and $\bar{\mu}^0$ *satisfies*

$$\begin{aligned} \int_{\mathbf{H}\times H^1} \|\nabla f\|_{L^2}^2 \mathrm{d}\bar{\mu}^0(u,f) &\leq \bar{\varepsilon}, \\ \int_{\mathbf{H}\times H^1} \|f\|_{L^2}^p \mathrm{d}\bar{\mu}^0(u,f) &\lesssim_p \bar{\varepsilon}^{\frac{p}{2}}. \end{aligned} \tag{8.3}$$

Proof. Tightness follows from equation (8.1) (and the corresponding balance on u) and Prokhorov's theorem. The estimates follow from (8.1) and lower semicontinuity. Finally, that $\bar{\mu}^0$ is a stationary measure of the inviscid problem (1.5) follows as in the corresponding statements in [17, 83] and is omitted for the sake of brevity. ■

Analogous to the arguments in [17], we deduce that necessarily $\bar{\mu}^0 = \mu \times \delta_0$ via Theorem 1.14.

Corollary 8.3. *The only stationary measure for the process* (u_t, f_t^0) *is the measure* $\mu \times \delta_0$.

Proof. Let us use the notation $f_{t,u,f}$ to denote the scalar process f_t^0 associated with initial conditions $(u_0, f_0) = (u, f) \in \mathbf{H} \times H^1$. Let $\bar{\mu}$ be any ergodic stationary measure for the process; by stationarity we have

$$\mathbf{E} \int_{\mathbf{H}\times H^1} \left(\int_{\mathbb{T}^d} |\nabla f_{t,u,f}|^2 \,\mathrm{d}x \right) \mathrm{d}\bar{\mu}(u,f) = \int_{\mathbf{H}\times H^1} \left(\int_{\mathbb{T}^d} |\nabla f|^2 \,\mathrm{d}x \right) \mathrm{d}\bar{\mu}(u,f)$$

at all times $t \geq 0$. On the other hand, if $\bar{\mu}$ is not of the form $\mu \times \delta_0$, then by Theorem 1.14 there is a positive $\bar{\mu}$-measure set $\mathcal{A} \subset \mathbf{H} \times H^1 \setminus \{0\}$ with the property that for all $(u, f) \in \mathcal{A}$, we have $\mathbf{E}(\int_{\mathbb{T}^d} |\nabla f_{t,u,f}|^2 \,\mathrm{d}x) \to \infty$ as $t \to \infty$. This implies a contradiction. ■

Theorem 1.16*, part* (i). Follows from Lemma 8.2 together with Corollary 8.3 and (8.3) (with $p > 2$). ■

Next, a variant of arguments in [17] gives Yaglom's law (1.8).

Proof of Theorem 1.16*, part* (ii). To adapt the arguments of [17], the first step is to derive the analogue of the Kármán–Howarth–Monin (KHM) relation [40, 55, 90] for the passive scalar. In what follows u and g denote statistically stationary solutions to (1.2.2). Define the scalar two point correlation

$$\mathfrak{G}(y) = \mathbf{E} \fint_{\mathbb{T}^d} g(x) g(x+y) \,\mathrm{d}x,$$

and the vector

$$\mathfrak{D}(y) = \mathbf{E} \fint_{\mathbb{T}^d} |\delta_y g(x)|^2 \delta_y u \,\mathrm{d}x.$$

Similarly, denote the two point covariance of the noise

$$\mathfrak{a}(y) = \frac{1}{2} \sum_{k\in\mathbb{Z}_0^d} \fint_{\mathbb{T}^d} |\tilde{q}_k|^2 e_k(x) \otimes e_k(x+y) \,\mathrm{d}x.$$

Note that $\mathfrak{a}(0) = \bar{\varepsilon}$. The KHM relation is the manifestation of the L^2 balance on the two point correlation $\mathfrak{G}$; it is significantly simpler for scalars than for the 3D Navier–Stokes equations. Hence, the proof is omitted for brevity; see [17] for details.

Proposition 8.4 (Scalar KHM relation). *Let (u_t, g_t) be a statistically stationary solution to (1.2.2) coupled to one of Systems 1–4. Then, for any $\eta = \eta(y)$ a smooth, compactly supported test function on $\mathbb{R}^d$, there holds*

$$\frac{1}{2}\int_{\mathbb{R}^d} \nabla\eta(y)\cdot\mathfrak{D}(y)\,\mathrm{d}y = 2\kappa\int_{\mathbb{R}^d}\Delta\eta(y)\mathfrak{G}(y)\,\mathrm{d}y + 2\int_{\mathbb{R}^d}\eta(y)\mathfrak{a}(y)\,\mathrm{d}y. \tag{8.4}$$

Define (suppressing the time-dependence as anyway, the time-dependence vanishes after expectations due to stationarity)

$$\bar{\mathfrak{D}}(\ell) = \mathbf{E}\fint_{\mathbb{T}^d}\fint_{\mathbb{S}^{d-1}} |\delta_{\ell n} g|^2 \delta_{\ell n} u\cdot n\,\mathrm{d}S(n)\,\mathrm{d}x.$$

Equipped with Proposition 8.4, we may proceed as in [17] by testing (8.4) with a radially symmetric test function $\eta(h) = \phi(|h|)$. Hence, we obtain the following ODE for $\bar{\mathfrak{D}}$ in the distributional sense:

$$\frac{\mathrm{d}}{\mathrm{d}\ell}(\ell^{d-1}\bar{\mathfrak{D}}) = -4\kappa\frac{\mathrm{d}}{\mathrm{d}\ell}\left(\ell^{d-1}\frac{\mathrm{d}}{\mathrm{d}\ell}\bar{\mathfrak{G}}\right) - 4\ell^{d-1}\bar{\mathfrak{a}}, \tag{8.5}$$

where we denote the spherically averaged quantities

$$\bar{\mathfrak{G}}(\ell) = \fint_{\mathbb{S}^{d-1}} \mathfrak{G}(\ell n)\,\mathrm{d}S(n),$$
$$\bar{\mathfrak{a}}(\ell) = \fint_{\mathbb{S}^{d-1}} \mathfrak{a}(\ell n)\,\mathrm{d}S(n).$$

From here, the proof proceeds as in the proof of the 4/3 law in [17]. Specifically, one first integrates (8.5) yielding,

$$\frac{\bar{\mathfrak{D}}(\ell)}{\ell} = -4\kappa\frac{\bar{\mathfrak{G}}'(\ell)}{\ell} - 4\ell^{-d}\int_0^\ell r^{d-1}\bar{\mathfrak{a}}(r)\,\mathrm{d}r.$$

Then the weak anomalous dissipation (1.7) is used to eliminate the term involving κ as $\kappa\to 0$ over an appropriate range of scales $[\ell_D, \ell_I]$ with $\ell_D(\kappa)^2 = o(\kappa\mathbf{E}\|g\|^2)$ as $\kappa\to 0$, to conclude that

$$\lim_{\kappa\to 0}\sup_{[\ell_D(\kappa),\ell_I]} 4\kappa\left|\frac{\bar{\mathfrak{G}}'(\ell)}{\ell}\right| = 0.$$

Finally, regularity of $\bar{\mathfrak{a}}(\ell)$ near $\ell = 0$ is used to deduce that

$$4\ell^{-d}\int_0^\ell r^{d-1}\bar{\mathfrak{a}}(r)\,\mathrm{d}r \to \frac{4}{d}\bar{\varepsilon}\quad \text{as } \ell\to 0.$$

The resulting estimate for $\frac{\bar{\mathfrak{D}}(\ell)}{\ell}$ is then asymptotically $-\frac{4}{d}\bar{\varepsilon}$ as $\ell_I\to 0$. ∎

Appendix A. Some background lemmas

A.1. Well-posedness and the RDS framework

In this section we will confirm that the various processes considered in this paper, e.g., the Eulerian process (u_t) and the Lagrangian process (u_t, x_t), arise as random dynamical systems in the framework of Section 3.

To start, without loss of generality, we may regard our probability space Ω as a countable product of canonical spaces $(C([0,\infty),\mathbb{R}))^{\otimes\mathbb{N}}$ with the product topology; likewise, $\mathscr{F}$ is the corresponding Borel sigma algebra and $\mathbf{P}$ the countable product of Wiener measures.

For each of Systems 1–4, we follow the standard procedure of defining the (u_t) process to be a solution of the corresponding equation in the mild sense [38, 86], i.e.,

$$u_t = e^{-tA}u_0 + \Gamma_t + \int_0^t e^{-(t-s)A}B(u_s, u_s)\,\mathrm{d}s, \tag{A.1}$$

where $\Gamma_t = \int_0^t e^{-(t-s)A}Q\,\mathrm{d}W(s)$ is the pertinent stochastic convolution for our additive noise. Recall that for all systems in consideration, A denotes the dominant linear term and B the pertinent bilinear nonlinearity (note $B \equiv 0$ for System 1).

Proposition A.1 ([38]). *For System 1 we have the following. For $\mathbf{P}$-almost every $\omega \in \Omega$; all $u_0 \in \mathbf{H}_{\mathcal{K}}$ and all $T > 0$, $p \geq 1$, there is a unique solution $(u_t) \in C([0,T];\mathbf{H}_{\mathcal{K}})$ to (A.1). Moreover, the process (u_t) is $\mathscr{F}_t$-adapted, with $u \in L^p(\Omega; C([0,T];\mathbf{H}_{\mathcal{K}}))$. Additionally suppose there holds*

$$\lim_{n\to\infty} \|QW_n - QW\|_{L^\infty(0,T;\mathbf{H}_{\mathcal{K}})} = 0,$$
$$\lim_{n\to\infty} \|u_0^{(n)} - u_0\|_{\mathbf{H}_{\mathcal{K}}} = 0.$$

Then the corresponding solutions $u_t^{(n)}$ satisfy

$$\lim_{n\to\infty} \|u_t^{(n)} - u_t\|_{L^\infty(0,T;\mathbf{H}_{\mathcal{K}})} = 0.$$

Moreover, this convergence is uniform over bounded sets, e.g., one has $\|u_0\|_{\mathbf{H}_{\mathcal{K}}} \leq C$ and $\|QW\|_{L^\infty(0,T;\mathbf{H}_{\mathcal{K}})} \leq C$ for $C < \infty$.

Proposition A.2 ([38, 86]). *For System 2 we have the following. For $\mathbf{P}$-almost every $\omega \in \Omega$; all $u_0 \in \mathbf{H}_N$ and all $T > 0$, $p \geq 1$, we have that there exists a unique solution $(u_t) \in C([0,T];\mathbf{H}_N)$ to equation (A.1). Moreover, the process (u_t) is $\mathscr{F}_t$-adapted, with $u \in L^p(\Omega; C([0,T];\mathbf{H}_N))$. Additionally suppose there holds*

$$\lim_{n\to\infty} \|QW_n - QW\|_{L^\infty(0,T;\mathbf{H}_N)} = 0,$$
$$\lim_{n\to\infty} \|u_0^{(n)} - u_0\|_{\mathbf{H}_N} = 0.$$

Then the corresponding solutions $u_t^{(n)}$ satisfy

$$\lim_{n\to\infty} \|u_t^{(n)} - u_t\|_{L^\infty(0,T;\mathbf{H}_N)} = 0.$$

Moreover, this convergence is uniform over bounded sets, e.g., one has $\|u_0\|_{\mathbf{H}_N} \leq C$ *and* $\|QW\|_{L^\infty(0,T;\mathbf{H}_N)} \leq C$ *for* $C < \infty$.

Proposition A.3 ([38, 86]). *For Systems 3 and 4 we have the following with* $d = 2$ *if System 3 and* $d = 3$ *if System 4. For* $\mathbf{P}$*-almost every* $\omega \in \Omega$, *for all* $u_0 \in \mathbf{H} \cap H^\gamma$ *with* $\gamma < \alpha - \frac{d}{2}$ *and for all* $T > 0$, $p \geq 1$, *we have that there exists a unique solution* $(u_t) \in C([0,T];\mathbf{H} \cap H^\gamma)$ *to equation (A.1). Moreover, the process* (u_t) *is* $\mathcal{F}_t$*-adapted, with* $u \in L^p(\Omega; C([0,T];\mathbf{H} \cap H^\gamma)) \cap L^2(\Omega; L^2(0,T;H^{\gamma+(d-1)}))$. *Additionally:*

(i) *For all* $p \geq 1$ *and* $\gamma < \gamma' < \alpha - \frac{d}{2}$,

$$\mathbf{E} \sup_{t \in [0,T]} \|u_t\|_{H^\gamma}^p \lesssim_{T,p,\gamma} 1 + \|u_0\|_{\mathbf{H} \cap H^\gamma}^p, \tag{A.2}$$

$$\mathbf{E} \int_0^T \|u_s\|_{H^{\gamma+(d-1)}}^2 \, \mathrm{d}s \lesssim_{T,\delta} 1 + \|u_0\|_{H^\gamma}^2, \tag{A.3}$$

$$\mathbf{E} \sup_{t \in [0,T]} \left(t^{\frac{\gamma'-\gamma}{2(d-1)}} \|u_t\|_{H^{\gamma'}} \right)^p \lesssim_{p,T,\gamma,\gamma'} 1 + \|u_0\|_{H^\gamma}^p. \tag{A.4}$$

(ii) *Suppose for* $\gamma, \delta > 0$ *arbitrary satisfying* $\gamma + \delta < \alpha - \frac{d}{2}$, *there holds*

$$\lim_{n \to \infty} \|QW_n - QW\|_{L^\infty(0,T;H^{\gamma+\delta})} = 0,$$

$$\lim_{n \to \infty} \|u_0^{(n)} - u_0\|_{H^\gamma} = 0.$$

Then the corresponding solutions $u_t^{(n)}$ *satisfy*

$$\lim_{n \to \infty} \|u_t^{(n)} - u_t\|_{L^\infty(0,T;H^\gamma)} = 0.$$

Moreover, this convergence is uniform over bounded sets, e.g., $\|u_0\|_{H^\gamma} \leq C$ *and* $\|QW\|_{L^\infty(0,T;H^{\gamma+\delta})} \leq C$ *for* $C < \infty$.

Proof. Item (i) is a consequence of standard arguments (see, e.g., [86]) combined with the following estimates on the stochastic convolution Γ_t:

Lemma A.4. *Let*

$$\Gamma_t = \int_0^t e^{-(t-s)A} Q \mathrm{d}W(s).$$

Then for all $T > 0$, $p \in [1,\infty)$, *and* $\gamma < \alpha + \frac{d}{2} - 1$,

$$\mathbf{E} \sup_{t \in [0,T]} \|\Gamma_t\|_{H^\gamma}^p \lesssim_{p,T,\gamma} 1,$$

$$\mathbf{E} \int_0^T \|\Gamma_t\|_{H^{\gamma+(d-1)}}^2 \, \mathrm{d}t \lesssim_{p,T,\gamma} 1.$$

Lemma A.4 follows from the Factorization Lemma, the Burkholder–Davis–Gundy inequality, and the smoothing properties of the heat semigroup (see, e.g., [38]).

Proposition A.3 (ii) can be proved by essentially the same stability argument as that in the proof of Lemma 7.3, to which we refer the reader for details. ∎

Let

$$\mathcal{U} : [0,\infty) \times \Omega \times \hat{\mathbf{H}} \to \hat{\mathbf{H}}, \quad (t,\omega,u) \mapsto \mathcal{U}^t_\omega(u)$$

denote the mapping sending, for a given $t \geq 0$ and $\mathbf{P}$-generic $\omega \in \Omega$, a given $u \in \hat{\mathbf{H}}$ to the time-t vector field u_t conditioned on $u_0 = u$. We conclude from Proposition A.3 that $\mathcal{U}$ is a continuous RDS in the sense of Section 3.1.1 on the space $Z = \hat{\mathbf{H}}$ satisfying condition (H1). Similarly, the random ODE (1.1) defining the auxiliary process $x_t = \phi^t_{\omega,u_0} x_0$ is well posed, and we conclude as before that the corresponding mapping $\Theta : [0,\infty) \times \Omega \times \hat{\mathbf{H}} \times \mathbb{T}^d \to \hat{\mathbf{H}} \times \mathbb{T}^d$ for the Lagrangian flow process (u_t, x_t) is a continuous RDS satisfying (H1) on the space $Z = \hat{\mathbf{H}} \times \mathbb{T}^d$. We leave it to the reader to confirm that the same is true for each of the processes (u_t, x_t, v_t) and $(u_t, x_t, \check{v}_t)$ on $Z = \hat{\mathbf{H}} \times \mathbb{T}^d \times P^{d-1}$ and (u_t, x_t, A_t) on $Z = \hat{\mathbf{H}} \times \mathbb{T}^d \times \mathrm{SL}_d(\mathbb{R})$, defined by the random ODE in (1.9).

In addition, in this paper we consider the linear cocycles

$$\mathcal{A}, \check{\mathcal{A}} : [0,\infty) \times \Omega \times \hat{\mathbf{H}} \times \mathbb{T}^d \to M_{d\times d}(\mathbb{R})$$

defined by

$$\mathcal{A}^t_{\omega,u,x} = D_x \phi^t_{\omega,u} \quad \text{and} \quad \check{\mathcal{A}}^t_{\omega,u,x} = (\mathcal{A}^t_{\omega,u,x})^{-\top}.$$

The integrability condition (H2) in Section 3.2.2 for each of these processes follows from (A.4) above, while the independent increments condition (H3) is equivalent to condition (H1) for the (u_t, x_t, A_t) process.

We close this part with a brief check that $\mu \times \mathrm{Leb}$ is a stationary measure for the (u_t, x_t) process.

Lemma A.5. *For any of Systems 1–4, we have that $\mu \times \mathrm{Leb}$ is a stationary measure for the Markov process (u_t, x_t).*

Proof. It suffices to show that for all bounded measurable $\psi : \hat{\mathbf{H}} \times \mathbb{T}^d \to \mathbb{R}$, we have that

$$\mathbf{E} \int \psi(u_t, x_t)\, \mathrm{d}\mu(u_0)\, \mathrm{d}x_0 = \int \psi(u,x)\, \mathrm{d}\mu(u)\, \mathrm{d}x.$$

To see this, define $\hat{\psi}(u) := \int \psi(u,x)\, \mathrm{d}x$ and observe that by Fubini, the above left-hand side coincides with

$$(*) := \mathbf{E} \int \left(\int \psi(u_t, \phi^t_{\omega,u_0}(x_0))\, \mathrm{d}x_0 \right) \mathrm{d}\mu(u_0) = \mathbf{E} \int \hat{\psi}(u_t)\, \mathrm{d}\mu(u_0),$$

having used that $x \mapsto \phi^t_{\omega,u_0}(x)$ preserves Leb on $\mathbb{T}^d$. By stationarity of μ, the expression $(*)$ equals

$$\int \hat{\psi}(u)\, \mathrm{d}\mu(u) = \int \psi(u,x)\, \mathrm{d}\mu(u)\, \mathrm{d}x. \qquad \blacksquare$$

A.2. Hölder estimates and interpolation inequalities

The following interpolation lemma is very useful for proving invertibility of the Malliavin matrix and is taken from [65, Lemma 6.14].

Lemma A.6. *Let f be a C^1 function on $[0,1]$ and let $\alpha \in (0,1]$. Then the following inequality holds for all $t \in (0,1)$:*

$$\|\partial_t f\|_{L^\infty([0,t])} \leq \frac{4}{t}\|f\|_{L^\infty([0,t])}^{\frac{\alpha}{\alpha+1}} \max\Big\{\|f\|_{L^\infty([0,t])}^{\frac{1}{1+\alpha}}, [\partial_t f]_{C^\alpha([0,t])}^{\frac{1}{1+\alpha}}\Big\},$$

where $[\,\cdot\,]_{C^\alpha([0,t])}$ denotes the α-Hölder seminorm on $[0,t]$.

The following estimate on the Hölder norms of a process in a general Hilbert space is also useful for verifying the Hölder estimates required to use the previous lemma.

Lemma A.7. *Let $\mathcal{H}$ and $\mathcal{W}$ be separable Hilbert spaces and let Y_t, $t \in [0,1]$, be an $\mathcal{H}$-valued process given by*

$$Y_t = Y_0 + \int_0^t B_s \, \mathrm{d}s + \int_0^t Q_s \, \mathrm{d}W_s,$$

where W_t is a cylindrical Wiener process on $\mathcal{W}$, and B_t, Q_t are predictable processes taking values in $\mathcal{H}$ and $\mathcal{L}^2(\mathcal{W},\mathcal{H})$, the space of bounded Hilbert–Schmidt operators from $\mathcal{W}$ to $\mathcal{H}$. Assume that B_t and Q_t satisfy, for every $p \geq 2$,

$$R_p := \mathbf{E}\big(\|B\|^p_{L^\infty([0,1];\mathcal{H})} + \|Q\|^p_{L^\infty([0,1];\mathcal{L}^2(\mathcal{W},\mathcal{H}))}\big) < \infty,$$

then for every $p \geq 2$, and $\gamma \in (0, 1/2 - 1/p)$ we have the estimate

$$\mathbf{E}\|Y\|^p_{C^\gamma([0,1];\mathcal{H})} \lesssim_p R_p.$$

Proof. The proof is a more or less classical. It involves the Burkholder–Davis–Gundy (BDG) inequality. Indeed, the BDG inequality for the stochastic integral above (see, e.g., [39, Theorem 4.37]) implies for $t \neq s$, $t, s \in [0,1]$ that for $p \geq 2$,

$$\mathbf{E}\|Y_t - Y_s\|^p_{\mathcal{H}} \lesssim_p |t-s|^p \mathbf{E}\|B\|^p_{L^\infty([0,1];\mathcal{H})} + |t-s|^{p/2}\mathbf{E}\|Q\|^p_{L^\infty([0,1];\mathcal{L}^2(\mathcal{W},\mathcal{H}))}.$$

It follows that for $t, s \in [0,1]$, and $1/p < r < 1/2$,

$$\frac{\mathbf{E}\|Y_t - Y_s\|^p_{\mathcal{H}}}{|t-s|^{rp+1}} \lesssim_p R_p |t-s|^{p(1/2-r)-1}.$$

Since

$$\int_0^1 \int_0^1 |t-s|^{p(1/2-r)-1} \, \mathrm{d}s \, \mathrm{d}t < \infty$$

for $r < 1/2$, we find after applying Fubini

$$\mathbf{E}\int_0^1 \int_0^1 \frac{\|Y_t - Y_s\|^p_{\mathcal{H}}}{|t-s|^{rp+1}} \, \mathrm{d}s \, \mathrm{d}t \lesssim_p R_p.$$

By a classic application of the Garcia–Rodemich–Rumsey–Rosenblatt inequality (see [58]) we obtain for $\gamma = r - 1/p > 0$ that

$$\mathbf{E}\|Y_t\|^p_{C^\gamma([0,1];\mathcal{H})} \lesssim \mathbf{E}\int_0^1 \int_0^1 \frac{\|Y_t - Y_s\|^p_{\mathcal{H}}}{|t-s|^{rp+1}} \, \mathrm{d}s \, \mathrm{d}t \lesssim_p R_p. \qquad \blacksquare$$

Funding. Jacob Bedrossian was supported by NSF CAREER grant DMS-1552826 and NSF RNMS #1107444 (Ki-Net). This material is based upon work supported by the National Science Foundation under Award No. DMS-1604805 (Alex Blumenthal) and Award No. DMS-1803481 (Sam Punshon-Smith).

References

[1] Agrachev, A. A., Sarychev, A. V.: Navier–Stokes equations: Controllability by means of low modes forcing. J. Math. Fluid Mech. **7**, 108–152 (2005) Zbl 1075.93014 MR 2127744

[2] Agrachev, A. A., Sarychev, A. V.: Controllability of 2D Euler and Navier–Stokes equations by degenerate forcing. Comm. Math. Phys. **265**, 673–697 (2006) Zbl 1105.93008 MR 2231685

[3] Alberti, G., Crippa, G., Mazzucato, A. L.: Exponential self-similar mixing and loss of regularity for continuity equations. C. R. Math. Acad. Sci. Paris **352**, 901–906 (2014) Zbl 1302.35241 MR 3268760

[4] Alberti, G., Crippa, G., Mazzucato, A. L.: Exponential self-similar mixing by incompressible flows. J. Amer. Math. Soc. **32**, 445–490 (2019) Zbl 1456.35158 MR 3904158

[5] Albeverio, S., Debussche, A., Xu, L.: Exponential mixing of the 3D stochastic Navier–Stokes equations driven by mildly degenerate noises. Appl. Math. Optim. **66**, 273–308 (2012) Zbl 1303.60051 MR 2965267

[6] Antonsen Jr, T. M., Fan, Z., Ott, E., Garcia-Lopez, E.: The role of chaotic orbits in the determination of power spectra of passive scalars. Phys. Fluids **8**, 3094–3104 (1996) Zbl 1027.76662

[7] Antonsen Jr, T. M., Ott, E.: Multifractal power spectra of passive scalars convected by chaotic fluid flows. Phys. Rev. A **44**, art. 851 (1991)

[8] Arnold, L.: Random Dynamical Systems. Springer Monogr. Math., Springer, Berlin (1995) Zbl 0906.34001 MR 1723992

[9] Avila, A., Viana, M.: Extremal Lyapunov exponents: An invariance principle and applications. Invent. Math. **181**, 115–189 (2010) Zbl 1196.37054 MR 2651382

[10] Balkovsky, E., Fouxon, A.: Universal long-time properties of Lagrangian statistics in the Batchelor regime and their application to the passive scalar problem. Phys. Rev. E (3) **60**, 4164–4174 (1999) MR 1717022

[11] Batchelor, G. K.: Small-scale variation of convected quantities like temperature in turbulent fluid. I. General discussion and the case of small conductivity. J. Fluid Mech. **5**, 113–133 (1959) Zbl 0085.39701 MR 0100457

[12] Baxendale, P. H.: Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms. Probab. Theory Related Fields **81**, 521–554 (1989) Zbl 0645.58043 MR 0995809

[13] Baxendale, P. H., Rozovskiĭ, B. L.: Kinematic dynamo and intermittence in a turbulent flow. Geophys. Astrophy. Fluid Dyn. **73**, 33–60 (1993) MR 1289018

[14] Bedrossian, J., Blumenthal, A., Punshon-Smith, S.: Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic Navier–Stokes. Probab. Theory Related Fields **179**, 777–834 (2021) Zbl 07334598 MR 4242626

[15] Bedrossian, J., Blumenthal, A., Punshon-Smith, S.: Almost-sure exponential mixing of passive scalars by the stochastic Navier–Stokes equations. arXiv:1905.03869 (2019); to appear in Ann. Probab.

[16] Bedrossian, J., Blumenthal, A., Punshon-Smith, S.: The Batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics at fixed Reynolds number. arXiv:1911.11014 (2019); to appear in Comm. Pure Appl. Math.

[17] Bedrossian, J., Coti Zelati, M., Glatt-Holtz, N.: Invariant measures for passive scalars in the small noise inviscid limit. Comm. Math. Phys. **348**, 101–127 (2016) Zbl 1351.35123 MR 3551262

[18] Bedrossian, J., Coti Zelati, M., Punshon-Smith, S., Weber, F.: Sufficient conditions for dual cascade flux laws in the stochastic 2d Navier–Stokes equations. Arch. Ration. Mech. Anal. **237**, 103–145 (2020) Zbl 07201254 MR 4090463

[19] Billingsley, P.: Convergence of Probability Measures. John Wiley & Sons, New York (2013) Zbl 0172.21201 MR 0233396

[20] Biskamp, D.: Magnetohydrodynamic Turbulence. Cambridge University, Cambridge (2003) Zbl 1145.76300

[21] Blumenthal, A., Xue, J., Young, L.-S.: Lyapunov exponents for random perturbations of some area-preserving maps including the standard map. Ann. of Math. (2) **185**, 285–310 (2017) Zbl 1365.37039 MR 3583355

[22] Boffetta, G., Ecke, R. E.: Two-dimensional turbulence. Annu. Rev. Fluid Mech. **44**, 427–451 (2012) Zbl 1350.76022 MR 2882604

[23] Bogachev, V. I.: Differentiable measures and the Malliavin calculus. **87**, 3577–3731 (1997) Zbl 0929.58015 MR 1613197

[24] Bogachev, V. I.: Gaussian Measures. Math. Surveys Monogr. 62, American Mathematical Society, Providence (1998) Zbl 0913.60035 MR 1642391

[25] Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrödinger Operators. Progr. Probab. Stat. 8, Birkhäuser, Boston (1985) MR 0886674

[26] Bressan, A.: A lemma and a conjecture on the cost of rearrangements. Rend. Semin. Mat. Univ. Padova **110**, 97–102 (2003) Zbl 1114.05002 MR 2033002

[27] Carverhill, A.: Flows of stochastic dynamical systems: Ergodic theory. Stochastics **14**, 273–317 (1985) Zbl 0536.58019 MR 805125

[28] Carverhill, A.: Furstenberg's theorem for nonlinear stochastic systems. Probab. Theory Related Fields **74**, 529–534 (1987) Zbl 0592.60043 MR 0876254

[29] Cerrai, S.: Ergodicity for stochastic reaction-diffusion systems with polynomial coefficients. Stochastics Stochastics Rep. **67**, 17–51 (1999) Zbl 0935.60049 MR 1717811

[30] Chang, J. T., Pollard, D.: Conditioning as disintegration. Stat. Neerl. **51**, 287–317 (1997) Zbl 0889.62003 MR 1484954

[31] Cheskidov, A., Shvydkoy, R.: Euler equations and turbulence: Analytical approach to intermittency. SIAM J. Math. Anal. **46**, 353–374 (2014) Zbl 1295.76010 MR 3152734

[32] Chirikov, B. V.: A universal instability of many-dimensional oscillator systems. Phys. Rep. **52**, 264–379 (1979) MR 0536429

[33] Constantin, P., Kiselev, A., Ryzhik, L., Zlatoš, A.: Diffusion and mixing in fluid flow. Ann. of Math. (2) **168**, 643–674 (2008) Zbl 1180.35084 MR 2434887

[34] Coti Zelati, M., Delgadino, M. G., Elgindi, T. M.: On the relation between enhanced dissipation timescales and mixing rates. Comm. Pure Appl. Math. **73**, 1205–1244 (2020) Zbl 1444.76098 MR 4156602

[35] Crisanti, A., Falcioni, M., Vulpiani, A., Paladin, G.: Lagrangian chaos: Transport, mixing and diffusion in fluids. Riv. Nuovo Cimento (3) **14**, art. 80 (1991) MR 1162096

[36] Crisanti, A., Paladin, G., Vulpiani, A.: Products of Random Matrices in Statistical Physics. Springer Ser. Solid-State Sci. 104, Springer, Berlin (1993) Zbl 0784.58003 MR 1278483

[37] Da Prato, G.: Introduction to Stochastic Analysis and Malliavin Calculus. Appunti. Sc. Norm. Super. Pisa (N. S.) 13, Edizioni della Normale, Pisa, 3rd ed. (2014) Zbl 1416.60001 MR 3186829

[38] Da Prato, G., Zabczyk, J.: Ergodicity for Infinite-Dimensional Systems. London Math. Soc. Lecture Note Ser. 229, Cambridge University, Cambridge (1996) Zbl 0849.60052 MR 1417491

[39] Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Encyclopedia Math. Appl. 152, Cambridge University, Cambridge, 2nd ed. (2014) Zbl 1317.60077 MR 3236753

[40] De Karman, T., Howarth, L.: On the statistical theory of isotropic turbulence. Proc. Roy. Soc. A **164**, 192–215 (1938) JFM 64.1453.03

[41] Doob, J. L.: Asymptotic properties of Markoff transition prababilities. Trans. Amer. Math. Soc. **63**, 393–421 (1948) Zbl 0041.45406 MR 0025097

[42] Drivas, T. D.: Turbulent cascade direction and Lagrangian time-asymmetry. J. Nonlinear Sci. **29**, 65–88 (2019) Zbl 1420.35207 MR 3908861

[43] Duarte, P.: Plenty of elliptic islands for the standard family of area preserving maps. Ann. Inst. H. Poincaré Anal. Non Linéaire **11**, 359–409 (1994) Zbl 0838.58030 MR 1287238

[44] E, W., Mattingly, J. C.: Ergodicity for the Navier–Stokes equation with degenerate random forcing: Finite-dimensional approximation. Comm. Pure Appl. Math. **54**, 1386–1402 (2001) Zbl 1024.76012 MR 1846802

[45] Eckmann, J.-P., Hairer, M.: Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. Comm. Math. Phys. **219**, 523–565 (2001) Zbl 0983.60058 MR 1838749

[46] Elgindi, T. M., Zlatoš, A.: Universal mixers in all dimensions. Adv. Math. **356**, art. 106807, 33 pp. (2019) Zbl 1428.35357 MR 4008523

[47] Eyink, G. L.: Local 4/5-law and energy dissipation anomaly in turbulence. Nonlinearity **16**, 137–145 (2003) Zbl 1138.76358 MR 1950780

[48] Falkovich, G., Gawędzki, K., Vergassola, M.: Particles and fields in fluid turbulence. Rev. Modern Phys. **73**, 913–975 (2001) Zbl 1205.76133 MR 1878800

[49] Fannjiang, A., Papanicolaou, G.: Convection enhanced diffusion for periodic flows. SIAM J. Appl. Math. **54**, 333–408 (1994) Zbl 0796.76084 MR 1265233

[50] Flandoli, F., Gątarek, D.: Martingale and stationary solutions for stochastic Navier–Stokes equations. Probab. Theory Related Fields **102**, 367–391 (1995) Zbl 0831.60072 MR 1339739

[51] Flandoli, F., Gubinelli, M., Hairer, M., Romito, M.: Rigorous remarks about scaling laws in turbulent fluids. Comm. Math. Phys. **278**, 1–29 (2008) Zbl 1140.76011 MR 2367196

[52] Flandoli, F., Maslowski, B.: Ergodicity of the 2-D Navier–Stokes equation under random perturbations. Comm. Math. Phys. **172**, 119–141 (1995) Zbl 0845.35080 MR 1346374

[53] Földes, J., Sy, M.: Invariant measures and global well posedness for the SQG equation. Arch. Ration. Mech. Anal. **241**, 187–230 (2021) Zbl 07364833 MR 4271958

[54] Folland, G. B.: Real Analysis. Pure Appl. Math. (New York), John Wiley & Sons, New York (1999) Zbl 0549.28001 MR 0767633

[55] Frisch, U.: Turbulence. Cambridge University, Cambridge (1995) Zbl 0832.76001 MR 1428905

[56] Furstenberg, H.: Noncommuting random products. Trans. Amer. Math. Soc. **108**, 377–428 (1963) Zbl 0203.19102 MR 0163345

[57] Furstenberg, H., Kesten, H.: Products of random matrices. Ann. Math. Statist. **31**, 457–469 (1960) Zbl 0137.35501 MR 0121828

[58] Garsia, A. M., Rodemich, E., Rumsey, H., Jr.: A real variable lemma and the continuity of paths of some Gaussian processes. Indiana Univ. Math. J. **20**, 565–578 (1970/71) Zbl 0252.60020 MR 0267632

[59] Glatt-Holtz, N. E., Herzog, D. P., Mattingly, J. C.: Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations. Ann. PDE **4**, art. 16, 103 pp. (2018) Zbl 1410.35118 MR 3862850

[60] Gol'dsheĭd, I. Y., Margulis, G. A.: Lyapunov indices of a product of random matrices. Russian Math. Surveys **44**, 11–71 (1989) Zbl 0687.60008 MR 1040268

[61] Gorodetski, A.: On stochastic sea of the standard map. Comm. Math. Phys. **309**, 155–192 (2012) Zbl 1347.37049 MR 2864790

[62] Guivarc'h, Y., Raugi, A.: Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. Z. Wahrsch. Verw. Gebiete **69**, 187–242 (1985) Zbl 0558.60009 MR 0779457

[63] Hairer, M.: On Malliavin's proof of Hörmander's theorem. Bull. Sci. Math. **135**, 650–666 (2011) Zbl 1242.60085 MR 2838095

[64] Hairer, M., Mattingly, J. C.: Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. Ann. of Math. (2) **164**, 993–1032 (2006) Zbl 1130.37038 MR 2259251

[65] Hairer, M., Mattingly, J. C.: A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. Electron. J. Probab. **16**, no. 23, 658–738 (2011) Zbl 1228.60072 MR 2786645

[66] Herzog, D. P., Mattingly, J. C.: A practical criterion for positivity of transition densities. Nonlinearity **28**, 2823–2845 (2015) Zbl 1329.60280 MR 3382587

[67] Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. **119**, 147–171 (1967) Zbl 0156.10701 MR 0222474

[68] Hörmander, L.: The Analysis of Linear Partial Differential Operators. I–IV. Grundlehren Math. Wiss. 256, 257, 274, 275, Springer, Berlin (1983–1985) Zbl 0521.35001 (Vol. I), Zbl 0521.35002 (Vol. II), Zbl 0601.35001 (Vol. III), Zbl 0612.35001 (Vol. IV) MR 1065993 (Vol. I), MR 2108588 (Vol. II), MR 1313500 (Vol. III), MR 1481433 (Vol. IV)

[69] Huang, G., Kuksin, S.: On the energy transfer to high frequencies in the damped/driven nonlinear Schrödinger equation. Stoch. Partial Differ. Equ. Anal. Comput. **9**, 867–891 (2021)

[70] Iyer, G., Kiselev, A., Xu, X.: Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. Nonlinearity **27**, 973–985 (2014) Zbl 1293.35248 MR 3207161

[71] Jakšić, V., Nersesyan, V., Pillet, C.-A., Shirikyan, A.: Large deviations and entropy production in viscous fluid flows. Arch. Ration. Mech. Anal. **240**, 1675–1725 (2021) Zbl 07357868 MR 4264955

[72] Khasminskii, R. Z.: Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. Teor. Verojatnost. i Primenen. **5**, 196–214 (1960) MR 0133871

[73] Kifer, Y.: Ergodic Theory of Random Transformations. Progr. Probab. Stat. 10, Birkhäuser, Boston (1986) Zbl 0604.28014 MR 0884892

[74] Kifer, Y.: A note on integrability of C^r-norms of stochastic flows and applications. In: Stochastic Mechanics and Stochastic Processes, Springer, Berlin, 125–131 (1988) Zbl 0653.60054 MR 0960163

[75] Kingman, J. F. C.: Subadditive ergodic theory. Ann. Probab. **1**, 883–909 (1973) Zbl 0311.60018 MR 0356192

[76] Kolmogoroff, A. N.: Dissipation of energy in the locally isotropic turbulence. C. R. (Dokl.) Acad. Sci. URSS (N.S.) **32**, 16–18 (1941) Zbl 0063.03292 MR 0005851

[77] Kolmogoroff, A. N.: On degeneration of isotropic turbulence in an incompressible viscous liquid. C. R. (Dokl.) Acad. Sci. URSS (N. S.) **31**, 538–540 (1941) Zbl 0026.17001 MR 0004568

[78] Kolmogoroff, A.: The local structure of turbulence in incompressible viscous fluid for very large Reynold's numbers. C. R. (Dokl.) Acad. Sci. URSS (N.S.) **30**, 301–305 (1941) Zbl 67.0850.06 MR 0004146

[79] Komorowski, T., Peszat, S., Szarek, T.: Passive tracer in a flow corresponding to two-dimensional stochastic Navier–Stokes equations. Nonlinearity **26**, 1999–2026 (2013) Zbl 1273.76083 MR 3071450

[80] Kraichnan, R. H.: Small-scale structure of a scalar field convected by turbulence. Phys. Fluids **11**, 945–953 (1968) Zbl 0164.28904

[81] Kraichnan, R. H.: Diffusion by a random velocity field. Phys. Fluids **13**, 22–31 (1970) Zbl 0193.27106

[82] Kruse, R.: Strong and weak approximation of semilinear stochastic evolution equations. Lecture Notes in Math. 2093, Springer, Cham (2014) Zbl 1285.60002 MR 3154916

[83] Kuksin, S. B.: The Eulerian limit for 2D statistical hydrodynamics. J. Stat. Phys. **115**, 469–492 (2004) Zbl 1157.76319 MR 2070104

[84] Kuksin, S. B.: On distribution of energy and vorticity for solutions of 2D Navier–Stokes equation with small viscosity. Comm. Math. Phys. **284**, 407–424 (2008) Zbl 1168.35034 MR 2448135

[85] Kuksin, S., Shirikyan, A.: Randomly forced CGL equation: Stationary measures and the inviscid limit. J. Phys. A **37**, 3805–3822 (2004) Zbl 1047.35061 MR 2039838

[86] Kuksin, S., Shirikyan, A.: Mathematics of Two-Dimensional Turbulence. Cambridge Tracts in Math. 194, Cambridge University, Cambridge (2012) Zbl 1333.76003 MR 3443633

[87] Kunita, H.: Stochastic differential equations with jumps and stochastic flows of diffeomorphisms. In: Itô's Stochastic Calculus and Probability Theory, Springer, Tokyo, 197–211 (1996) Zbl 0899.60048 MR 1439526

[88] Ledrappier, F.: Positivity of the exponent for stationary sequences of matrices. In: Lyapunov Exponents (Bremen 1984), Lecture Notes in Math. 1186, Springer, Berlin, 56–73 (1986) Zbl 0591.60036 MR 0850070

[89] Lin, Z., Thiffeault, J.-L., Doering, C. R.: Optimal stirring strategies for passive scalar mixing. J. Fluid Mech. **675**, 465–476 (2011) Zbl 1241.76361 MR 2801050

[90] Monin, A. S., Yaglom, A. M.: Statistical Fluid Mechanics: Mechanics of Turbulence. Vol. II. Dover, Mineola (2007) Zbl 1140.76004 MR 2406668

[91] Nazarenko, S.: Wave Turbulence. Lecture Notes in Phys. 825, Springer, Heidelberg (2011) Zbl 1220.76006 MR 3014432

[92] Nersesyan, V.: Approximate controllability of Lagrangian trajectories of the 3D Navier–Stokes system by a finite-dimensional force. Nonlinearity **28**, 825–848 (2015) Zbl 1308.35171 MR 3319385

[93] Nie, Q., Tanveer, S.: A note on third-order structure functions in turbulence. Roy. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. **455**, 1615–1635 (1999) Zbl 0952.76026 MR 1701547

[94] Norris, J.: Simplified Malliavin calculus. In: Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math. 1204, Springer, Berlin, 101–130 (1986) Zbl 0609.60066 MR 0942019

[95] Nualart, D.: The Malliavin Calculus and Related Topics. Probab. Appl. (New York), Springer, New York, 2nd ed. (2006) Zbl 1099.60003 MR 2200233

[96] Oseledec, V. I.: A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. Trudy Moskov. Mat. Obšč. **19**, 179–210 (1968) MR 0240280

[97] Pesin, Y., Climenhaga, V.: Open problems in the theory of non-uniform hyperbolicity. Discrete Contin. Dyn. Syst. **27**, 589–607 (2010) Zbl 1200.37001 MR 2600681

[98] Raghunathan, M. S.: A proof of Oseledec's multiplicative ergodic theorem. Israel J. Math. **32**, 356–362 (1979) Zbl 0415.28013 MR 571089

[99] Romito, M.: Ergodicity of the finite-dimensional approximation of the 3D Navier–Stokes equations forced by a degenerate noise. J. Statist. Phys. **114**, 155–177 (2004) Zbl 1060.76027 MR 2032128

[100] Romito, M., Xu, L.: Ergodicity of the 3D stochastic Navier–Stokes equations driven by mildly degenerate noise. Stochastic Process. Appl. **121**, 673–700 (2011) Zbl 1369.76048 MR 2770903

[101] Ruelle, D.: Ergodic theory of differentiable dynamical systems. Publ. Math. Inst. Hautes Études Sci. **50**, 27–58 (1979) Zbl 0426.58014 MR 0556581

[102] Ruelle, D.: Characteristic exponents and invariant manifolds in Hilbert space. Ann. of Math. (2) **115**, 243–290 (1982) Zbl 0493.58015 MR 647807

[103] Seidler, J.: A note on the strong Feller property. Unpublished lecture notes **98** (2001)

[104] Seis, C.: Maximal mixing by incompressible fluid flows. Nonlinearity **26**, 3279–3289 (2013) Zbl 1396.76042 MR 3141856

[105] Shirikyan, A.: Contrôlabilité exacte en projections pour les équations de Navier–Stokes tridimensionnelles. Ann. Inst. H. Poincaré Anal. Non Linéaire **24**, 521–537 (2007) Zbl 1119.93021 MR 2334990

[106] Shirikyan, A.: Local times for solutions of the complex Ginzburg–Landau equation and the inviscid limit. J. Math. Anal. Appl. **384**, 130–137 (2011) Zbl 1248.60075 MR 2822855

[107] Shraiman, B. I., Siggia, E. D.: Scalar turbulence. Nature **405**, art. 639 (2000)

[108] Viana, M.: Lectures on Lyapunov Exponents. Cambridge Stud. Adv. Math. 145, Cambridge University, Cambridge (2014) Zbl 1309.37001 MR 3289050

[109] Walters, P.: An Introduction to Ergodic Theory. Grad. Texts in Math. 79, Springer, New York (2000) Zbl 0958.28011

[110] Yaglom, A. M.: On the local structure of the temperature field in a turbulent fluid. Dokl. Akad. Nauk SSSR (N.S.) **69**, 743–746 (1949) Zbl 0036.25701 MR 0034165

[111] Young, L.-S.: Mathematical theory of Lyapunov exponents. J. Phys. A **46**, art. 254001, 17 pp. (2013) Zbl 1351.37135 MR 3067028

[112] Yuan, G.-C., Nam, K., Antonsen, T. M., Jr., Ott, E., Guzdar, P. N.: Power spectrum of passive scalars in two-dimensional chaotic flows. Chaos **10**, 39–49 (2000) Zbl 0983.76034 MR 1759761

[113] Zakharov, V. E., L'vov, V. S., Falkovich, G.: Kolmogorov Spectra of Turbulence I: Wave Turbulence. Springer, Berlin (1992) Zbl 0786.76002

[114] Zlatoš, A.: Diffusion in fluid flow: Dissipation enhancement by flows in 2D. Comm. Partial Differential Equations **35**, 496–534 (2010) Zbl 1201.35106 MR 2748635