Lower bounds on the Lyapunov exponents of stochastic differential equations

Jacob Bedrossian[*] Alex Blumenthal[†] Sam Punshon-Smith[‡]

April 20, 2022

Abstract

In this article, we review our recently introduced methods for obtaining strictly positive lower bounds on the top Lyapunov exponent of high-dimensional, stochastic differential equations such as the weakly-damped Lorenz-96 (L96) model or Galerkin truncations of the 2d Navier-Stokes equations. This hallmark of chaos has long been observed in these models, however, no mathematical proof had been made for either deterministic or stochastic forcing.

The method we proposed combines (A) a new identity connecting the Lyapunov exponents to a Fisher information of the stationary measure of the Markov process tracking tangent directions (the so-called "projective process"); and (B) an L^1-based hypoelliptic regularity estimate to show that this (degenerate) Fisher information is an upper bound on some fractional regularity. For L96 and GNSE, we then further reduce the lower bound of the top Lyapunov exponent to proving that the projective process satisfies Hörmander's condition. We review the recent contributions of the first and third author on the verification of this condition for the 2d Galerkin-Navier-Stokes equations in a rectangular, periodic box of any aspect ratio. Finally, we briefly contrast this work with our earlier work on Lagrangian chaos in the stochastic Navier-Stokes equations. We end the review with a discussion of some open problems.

Contents

2010 *Mathematics subject classification.* Primary: 37H15, 35H10. Secondary: 37D25, 58J65, 35B65

[*]Department of Mathematics, University of Maryland, College Park, MD 20742, USA jacob@math.umd.edu. J.B. was supported by National Science Foundation CAREER grant DMS-1552826, National Science Foundation RNMS #1107444 (Ki-Net)

[†]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA ablumenthal6@gatech.edu. A.B. was supported by National Science Foundation grant DMS-2009431

[‡]School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USAsamuel.punshonsmith@amias.ias.edu. This material was based upon work supported by the National Science Foundation under Award No. DMS-1803481.

1 Lyapunov exponents for stochastic differential equations

Understanding the "generic" long-term dynamics of high (or infinite) dimensional nonlinear systems far from equilibrium remains a daunting task. In physical applications of interest, many such systems are both subject to unpredictable external forcing and observed to be chaotic in the sense of being very sensitive to the initial condition and forcing. Hence, for all practical purposes, the exact dynamics of any specific trajectory cannot be predicted far in advance and any controlled experiments will not be exactly repeatable. Instead of reckoning such systems one trajectory at a time, a common practice is to view initial conditions as *random*, i.e., distributed according to some probabilistic law, and to attempt to understand how this law evolves as it is transported by the dynamics. In this context, the relevant "time-invariant" objects are *equilibrium probabilistic laws* on the phase space of the system, often referred to as *invariant measures* or *stationary measures*.

There is a well-developed abstract theory (smooth ergodic theory) for understanding the invariant measures of chaotic systems, their geometric properties, and how these relate to the asymptotic regimes of trajectories initiated from "typical" initial conditions. On the other hand, it is quite hard to verify mathematically that this abstract program applies to systems of practical interest. There are already extremely challenging open problems for vastly simplified 2d toy models of the kinds of chaotic behavior seen in fluid dynamics, e.g., the Chirikov standard map discussed below in Section 1.1.

It turns out that verifying and understanding chaotic properties is far more tractable for systems subjected to *random* noise. The kinds of systems we have in mind are, for example, hydrodynamical settings such as with wind over a sail, a weather or climate system, or nonlinear wave systems. In these settings it has long been suggested to study the random dynamical system generated by the PDE or ODE subjected to random external forcing, and this is often done in applied mathematics (see e.g. [27, 70] and the references therein). Even with the simplifications coming from the random forcing, and despite considerable efforts, a thorough, mathematically rigorous, understanding of these random systems is still in its infancy, with many basic open questions remaining.

In this article we will review existing work and our recent contributions [16, 21] in proving that a given system of interest modeled by a stochastic differential equation is chaotic as in high sensitivity to initial conditions for trajectories initiated at Lebesgue-typical points in phase space. The specific systems we apply our methods to are the Lorenz-96 system [67] and Galerkin truncations of the 2d Navier-Stokes equations in a rectangular, periodic box (of any aspect ratio), provided they are subjected to sufficiently strong stochastic forcing[1] (equivalently, sufficiently weak damping) and are sufficiently high dimensional. These are the first results of this type for such models, despite overwhelming numerical evidence (see e.g. [27, 54, 70, 74]). Specifically we prove for these models that if the damping parameter is ϵ, then the top Lyapunov exponent (see Sections 1.1 and 1.2 for definition) satisfies

$$\lim_{\epsilon \to 0} \frac{\lambda_1^\epsilon}{\epsilon} = \infty$$

as $\epsilon \to 0$ and in particular, $\exists \epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$, $\lambda_1^\epsilon > 0$.

Outline

In Section 1 we give background on Lyapunov exponents for stochastic differential equations (SDEs). Section 2 concerns formulae of Lyapunov exponents through the stationary statistics of tangent directions and contains both classical results and our recent results from [16] which connects Lyapunov exponents to a certain Fisher information-type quantity. We discuss in Section 3 how to connect the Fisher information to regularity using ideas from hypoellipticity theory (also original work from [16]), and in Section 4 we discuss applications to a class of weakly-driven, weakly-dissipated SDE with bilinear nonlinear drift term (original

[1] The deterministic case remains very far out of reach.

work in [16] for Lorenz-96 and for Galerkin Navier-Stokes in [21]). In Section 5 we briefly discuss our earlier related work on Lagrangian chaos in the (infinite-dimensional) stochastic Navier-Stokes equations [17]. Finally, in Section 6 we discuss some open problems and potential directions for research.

1.1 Lyapunov exponents and their challenges

Let $\Phi^t : \mathbb{R}^n \to \mathbb{R}^n, t \in \mathbb{R}_{\geq 0}$ be a flow (autonomous or not) with differentiable dependence on initial conditions. The *Lyapunov exponent* at $x \in \mathbb{R}^n$, when it exists, is the limit

$$\lambda(x) = \lim_{t\to\infty} \frac{1}{t} \log |D_x \Phi^t| \,,$$

where $D_x\Phi^t$ is the Jacobian of Φ^t at x, i.e. the derivative with respect to the initial condition. Hence, $\lambda(x)$ gives the asymptotic exponential growth rate of the Jacobian as $t \to \infty$.

The exponent $\lambda(x)$ contains information about the divergence of trajectories: heuristically at least, if $d(x, y)$ is small then

$$d(\Phi^t(x), \Phi^t(y)) \approx e^{\lambda(x)t} d(x, y)$$

and hence $\lambda(x) > 0$ implies *exponential sensitivity with respect to initial conditions*, commonly popularized as the "butterfly effect". Morally, a positive Lyapunov exponent at a 'large' proportion of initial conditions $x \in \mathbb{R}^n$ is a hallmark of chaos, the tendency of a dynamical system to exhibit disordered, unpredictable behavior. In this note we refer to a system such that $\lambda(x) > 0$ for Lebesgue a.e. x as *chaotic*[2].

The existence of Lyapunov exponents is usually justified using tools from ergodic theory, and forms a starting point for obtaining more refined dynamical features, such as stable/unstable manifolds in the moving frame along 'typical' trajectories. These ideas form the fundamentals of *smooth ergodic theory*, which aims to study *statistical* properties of chaotic systems, such as decay of correlations, i.e. how $\Phi^t(x), t \gg 1$ can 'forget' the initial $x \in \mathbb{R}^n$, and probabilistic laws such as a strong law of large numbers or central limit theorem for $g \circ \Phi^t(x)$, where $g : \mathbb{R}^n \to \mathbb{R}$ is a suitable *observable* of the system; see e.g. discussions in [8, 11, 66, 83, 85].

A discrete-time example

Unfortunately, estimating $\lambda(x)$ or proving $\lambda(x) > 0$ for specific systems turns out to be extremely challenging. A simple, classical model which shows the challenges is the Chirikov standard map family [30], written here as

$$F_L : \mathbb{T}^2 \circlearrowleft, \quad F_L(x, y) = (2x + L\sin(2\pi x) - y, x),$$

where $\mathbb{T}^2$ is parametrized as $[0, 1)^2$ and both coordinates in $F := F_L$ are taken modulo 1. Here, $L \geq 0$ is a fixed parameter which for purposes of the discussion here will be taken large. The diffeomorphism F is smooth and volume-preserving, and ergodic theory affirms that the Lyapunov exponent $\lambda(x, y) = \lim_n \frac{1}{n} \log |D_{(x,y)}F|$ exists for Leb. a.e. x and satisfies $\lambda(x, y) \geq 0$ where it exists. The Chirikov standard map itself is frequently used as a toy model of more complicated chaotic systems, e.g., the Navier-Stokes equations in transition from laminar flow to turbulence [69].

Observe that when $L \gg 1$ and away from an $O(L^{-1})$ neighborhood of $\{\cos(2\pi x) = 0\}$, the Jacobian $D_{(x,y)}F$ exhibits strong expansion along tangent directions roughly parallel to the x-axis (matched by strong contraction roughly parallel to the y-axis). In view of this, it is widely conjectured that $\{\lambda(x) > 0\}$ has positive Lebesgue measure. Nevertheless this *standard map conjecture* remains wide open [33, 75]. A key obstruction is 'cone twisting': on long timescales, vectors roughly parallel to the x-axis are strongly expanded until the first visit to 'critical strip' near $\{\cos(2\pi x) = 0\}$, where DF is approximately a rotation

[2] We caution the reader that there is no single mathematical definition of "chaos". Some definitions refer to the *existence* of a subset of phase space exhibiting chaotic behavior, e.g., Li-Yorke chaos or the presence of a hyperbolic horseshoe. The results discussed in this note pertain to the long-time behavior of Lebesgue-typical initial conditions.

by 90 degrees. At this point, vectors roughly parallel to the x axis are rotated to be roughly parallel to the y axis, where strong contraction occurs and previously accumulated expansion can be negated. Indeed, an estimate on a Lyapunov exponent requires understanding the asymptotic cancellations in the Jacobian as $t \to \infty$. One manifestation of the subtlety is the wildly tangled coexistence of hyperbolic trajectories [46] and elliptic islands [36].

The problem of estimating Lyapunov exponents for the standard map is far more tractable in the presence of noise / stochastic driving. Let us consider the standard map subjected to small noise: let $\omega_1, \omega_2, \cdots$ be IID random variables uniformly distributed in $[-\epsilon, \epsilon]$ for some $\epsilon > 0$, and consider the random compositions

$$F^n = F_{\omega_n} \circ \cdots \circ F_{\omega_1}\,, \quad F_{\omega_i}(x, y) = F(x + \omega_i, y)\,.$$

One can show show that $\forall \epsilon > 0$, the corresponding Lyapunov exponent $\lambda = \lambda(x, y)$ is *deterministic* (independent of the random samples almost surely) and constant (independent of (x, y)) with probability 1. It is a folklore theorem that $\lambda > 0\ \forall \epsilon > 0$, while for $L \gg 1$ and $\epsilon \gtrsim e^{-L}$, one can show $\lambda \geq \frac{1}{2}\log L$, commensurate with exponential expansion in the x-direction over the bulk of phase space [24]; in a related vein, see also [23, 25, 26, 64, 80].

1.2 Lyapunov exponents for SDE

The topic of this note is to discuss developments in the context of the random dynamical systems generated by stochastic differential equations (SDE), i.e., ODE subjected to Brownian motion driving terms. In this continuous-time framework, numerous additional tools not present in the discrete-time setting become available, e.g., infinitesimal generators, which as we show below, connects the estimation of Lyapunov exponents to regularity estimates (e.g., Sobolev regularity) of solutions to certain (degenerate) elliptic PDE. A highlight of this approach is our application to the Lyapunov exponents of a class of weakly-driven, weakly-forced SDE, including famous models such as Lorenz 96 and Galerkin truncations of the Navier-Stokes equations.

For simplicity, in this note we restrict our attention to SDE on $\mathbb{R}^n$, however, our more general results apply to SDE posed on orientable, geodesically complete, smooth manifolds; see [16]. Let $X_0, X_1, \ldots, X_r : \mathbb{R}^n \to \mathbb{R}^n$ be smooth vector fields on $\mathbb{R}^n$, and let $W_t^1, \ldots, W_t^r$ be a collection of independent, real-valued Brownian motions, with Ω denoting the corresponding canonical space with probability $\mathbf{P}$ and $(\mathcal{F}_t)_{t\geq 0}$ denoting the increasing filtration generated by $\{W_s^k, s \leq t\}_{k=1}^r$. We consider continuous-time processes (x_t) on $\mathbb{R}^n$ solving the SDE

$$\mathrm{d}x_t = X_0(x_t)\,\mathrm{d}t + \sum_{k=1}^{r} X_k(x_t) \circ \mathrm{d}W_t^k, \tag{1.1}$$

for fixed initial data $x_0 \in \mathbb{R}^n$.

Under mild conditions on the vector fields $X_0, \ldots, X_r$ (for example, regularity and the existence of a suitable Lyapunov function to rule out finite time blow-up), global-in-time solutions (x_t) to (1.1) exist, are unique, and have differentiable dependence of x_t on x_0; in particular, for $\mathbf{P}$-a.e. $\omega \in \Omega$ and all $t \geq 0$, there exists a stochastic flow of diffeomorphisms Φ_ω^t such that $\forall x_0 \in \mathbb{R}^n$, the law of the process $(x_t)_{t\geq 0}$ solving (1.1) is the same as that of the process $(\Phi_\omega^t(x_0))_{t\geq 0}$; see e.g. [61] for the details and general theory of SDEs and stochastic flows.

This *stochastic flow of diffeomorphisms* Φ_ω^t is the analogue of the flow Φ^t corresponding to solutions of the initial value problem of an ODE. However, the external stochastic forcing implies a time-inhomogeneity which must be accounted for. One can show that there exists a $\mathbf{P}$-measure preserving semiflow $\theta^t : \Omega \circlearrowleft, t \geq 0$ corresponding to time-shifts on the Brownian paths, i.e., shifting the path $(W_s)_{s\geq 0}$ to $(W_{t+s} - W_t)_{s\geq 0}$. Equipped with this time shift, one has the following with probability 1 and for all $s, t \geq 0$:

$$\Phi_\omega^{s+t} = \Phi_{\theta^s\omega}^t \circ \Phi_\omega^s. \tag{1.2}$$

We now set about summarizing the ergodic theory tools used to study such stochastic flows. First, we note that the trajectories $x_t = \Phi^t_\omega(x_0)$ for fixed initial $x_0 \in \mathbb{R}^n$ form a Markov process adapted to the filtration $(\mathcal{F}_t)$. Moreover, Φ^t_ω has *independent increments*: $\forall s, t \geq 0$, Φ^s_ω and $\Phi^t_{\theta^s \omega}$ are independent.

1.2.1 Stationary measures and long-term statistics

Markov semigroups. We write $P_t(x, A) = \mathbf{P}(\Phi^t_\omega(x) \in A)$ for the time-t transition kernel of (x_t). Let $\mathcal{P}_t$ denote the *Markov semigroup* associated to (x_t), defined for bounded, measurable *observables* $h : \mathbb{R}^n \to \mathbb{R}$ by

$$\mathcal{P}_t h(x) = \mathbf{E}[h(x_t) \,|\, x_0 = x] = \int_{\mathbb{R}^n} h(y) P_t(x, \mathrm{d}y).$$

This semigroup gives the expected value of a given observable given a fixed initial condition. Via the pairing of functions and measures we derive the (formal) dual $\mathcal{P}^*_t$, which gives the evolution of the law of the solution (x_t) given a distribution for the initial condition: for a probability measure $\mu_0 \in \mathcal{P}(\mathbb{R}^n)$ and Borel set $A \subseteq \mathbb{R}^n$,

$$\mathcal{P}^*_t \mu_0(A) = \int_{\mathbb{R}^n} P_t(x, A) \mathrm{d}\mu_0(x) \,.$$

That is, $\mathcal{P}^*_t \mu_0$ is the law of x_t assuming μ_0 is the law of x_0.

Taking a time derivative ∂_t, we (formally) obtain the *backward Kolmogorov equation*

$$\partial_t \mathcal{P}_t h(x) = \mathcal{L} \mathcal{P}_t h(x) \,, \quad \text{where} \quad \mathcal{L} = X_0 + \frac{1}{2} \sum_{i=1}^r X_i^2 \,, \tag{1.3}$$

where, for a given vector field X and $f \in C^\infty$, Xf denotes the derivative of f in the direction X. The differential operator $\mathcal{L}$ is called the *(infinitesimal) generator*. Assuming that the law of x_t has a density p_t with respect to Lebesgue, the formal dual of (1.3) is the *Fokker-Plank equation* (or *Forward Kolmogorov equation*) given by the following PDE

$$\partial_t p_t = \mathcal{L}^* p_t \,, \tag{1.4}$$

where $\mathcal{L}^*$ denotes the formal L^2 adjoint of $\mathcal{L}$. See e.g. [61] for mathematical details.

Stationary measures. We say a measure μ is *stationary* if $\mathcal{P}^*_t \mu = \mu$. That is, if x_0 is distributed with law μ, then x_t is distributed with law[3] μ for all $t > 0$. We say that a set $A \subset \mathbb{R}^n$ is *invariant* if $P_t(x, A) = 1$ for all $x \in A$ and $t \geq 0$, and we say that a stationary measure μ is *ergodic* if all invariant sets have μ-measure 0 or 1. By the pointwise ergodic theorem, ergodic stationary measures determine the long-term statistics of a.e. initial datum in their support [34]: if μ is an ergodic stationary measure, then for any bounded, measurable $\varphi : \mathbb{R}^n \to \mathbb{R}$ and $\mu \times \mathbf{P}$-a.e. (x, ω) we have that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \varphi(\Phi^t_\omega(x)) \, \mathrm{d}t = \int_{\mathbb{R}^n} \varphi(x) \mathrm{d}\mu(x) \,.$$

Unlike for deterministic systems, stationary measures are usually much easier to characterize for SDEs. In particular, it is often possible to show there exists a unique stationary measure and that it has a smooth density with respect to Lebesgue. In such a case, Leb. generic initial conditions all have the same long-term statistics, a property often observed in nature and experiments for the physical systems we are interested in.

[3] It is important to note that (x_t) itself is *not* constant in t; consider e.g. water flowing past a stone in a river.

Existence of stationary measures. If the domain of the Markov process were compact (e.g., $\mathbb{T}^n$ instead of $\mathbb{R}^n$) then existence of stationary measures follows from a standard Krylov-Bogoliubov argument: given an initial probability measure $\mu_0 \in \mathcal{P}(\mathbb{R}^n)$, one considers the time-averaged measures

$$\bar{\mu}_t := \frac{1}{t}\int_0^t \mathcal{P}_s^* \mu_0 \, \mathrm{d}s.$$

Weak-$*$ compactness of probability measures on a compact space ensures the sequence $\{\bar{\mu}_t\}_{t\geq 0}$ has a weak-$*$ limit point μ which by construction must be stationary (assuming some mild well-posedness properties for the original SDE). On a non-compact domain, one must show tightness of the measures $\{\bar{\mu}_t\}_{t\geq 0}$ (this is essentially saying that solutions do not wander off to infinity too often) and use Prokorov's theorem to pass to the limit in the narrow topology. This is often achieved using the method of Lyapunov functions[4]/drift conditions [71], or by using a special structure and the damping in the system (such as the case for e.g. the Navier-Stokes equations [60]).

Uniqueness of stationary measures. The Doob-Khasminskii theorem [34] implies uniqueness is connected to (A) irreducibility and (B) regularization of the Markov semigroups[5] and in particular, one can deduce that any stationary measure is unique if these properties hold in a sufficiently strong sense.

Let us first discuss irreducibility. For a Markov process (x_t) on $\mathbb{R}^n$, we say that (x_t) is *topologically irreducible* if for all open $U \subset \mathbb{R}^n$, $\exists t = t(U, x) \geq 0$ such that

$$P_t(x, U) > 0\,.$$

That is, every initial condition has a positive probability of being in U. This is stronger than necessary to deduce uniqueness, but is sufficient for our discussions.

Regularity is a little more subtle. A sufficient condition is the requirement of *strong Feller*:

$$\forall \varphi : \mathbb{R}^n \to \mathbb{R} \quad \text{bounded, measurable,} \quad \mathcal{P}_t\varphi \in C(\mathbb{R}^n;\mathbb{R}), \quad t > 0.$$

For finite dimensional SDEs, it is reasonably common and there exists machinery to characterize this[6]. When $\mathrm{Span}\{X_i(x), 1 \leq i \leq r\} = \mathbb{R}^n$ at all $x \in \mathbb{R}^n$, $\mathcal{L}$ is elliptic and hence strong Feller follows from classical parabolic regularity theory [65] applied to (1.3) (assuming suitable regularity conditions on the $\{X_j\}$). Absent this direct spanning (e.g., when $r < n$), then $\mathcal{L}$ is only degenerate elliptic. However, nearly sharp sufficient conditions for regularization due to $\mathcal{L}$ were derived by Hörmander [51], who obtained a condition (now called *Hörmander's condition*), in terms of the Lie algebra generated by the vector fields $\{X_i, 0 \leq i \leq r\}$. We will return to this important topic of *hypoellipticity* in Section 3.1.

1.2.2 Lyapunov exponents

We saw that by the ergodic theorem, the long-term behavior of scalar observables is determined by stationary measures. A more sophisticated ergodic theorem connects stationary measures to Lyapunov exponents. Given $x \in \mathbb{R}^n, v \in \mathbb{R}^n\backslash\{0\}$ (v being considered a direction here) and a random sample $\omega \in \Omega$, the Lyapunov exponent at (ω, x, v) is defined as the limit (if it exists)

$$\lambda(\omega, x, v) = \lim_{t\to\infty} \frac{1}{t} \log |D_x\Phi^t_\omega v|\,.$$

The following (truncated) version of Oseledets' *Multiplicative Ergodic Theorem* (MET) addresses the existence of the limit [56, 73, 77].

[4]These are the probabilistic analogue of Lyapunov's 'first method' for ODE, used to ensure convergence to compact attractors. This is not to be confused with Lyapunov exponents, which refer to Lyapunov's 'second method'.

[5]Morally, this is equivalent to how $x \mapsto P_t(x, \cdot)$, i.e., how trajectories with nearby initial conditions have similar statistics.

[6]In infinite dimensions it is much more rare; luckily it is stronger than what is required just to prove uniqueness (see e.g. [48, 59]).

Theorem 1.1. *(Oseledets' multiplicative ergodic theorem [73]) Let μ be an ergodic stationary measure, and assume a mild integrability condition (see, e.g., [56, 73]) then, there exist (deterministic) constants $\lambda_1 > \lambda_2 > \cdots > \lambda_\ell \geq -\infty$ such that for $\mathbf{P} \times \mu$-almost all $(\omega, x) \in \Omega \times \mathbb{R}^n$ and for* all $v \in \mathbb{R}^n \backslash \{0\}$*, the limit defining $\lambda(\omega, x, v)$ exists and takes one of the values $\lambda_i, 1 \leq i \leq \ell$.*

Moreover, there exists a $\mathbf{P} \times \mu$-measurably-varying flag of strictly increasing subspaces

$$\emptyset =: F_{\ell+1}(\omega, x) \subset F_\ell(\omega, x) \subset \ldots \subset F_1(\omega, x) := \mathbb{R}^n$$

such that for $\mathbf{P} \times \mu$-a.e. (ω, x) and $\forall v \in F_j \setminus F_{j+1}$ there holds

$$\lambda_j = \lim_{t \to \infty} \frac{1}{t} \log |D_x \Phi^t_\omega v| = \lambda(\omega, x, v).$$

In particular the top Lyapunov exponent λ_1 is realized at $\mathbf{P} \times \mu$-a.e. (ω, x) and all $v \in \mathbb{R}^n$ outside a positive-codimension subspace $F_2(\omega, x) \subset \mathbb{R}^n$.

We note that under very mild conditions, if the stationary measure μ is unique, it is automatically ergodic; otherwise, each distinct ergodic stationary measure admits its own set of Lyapunov exponents.

The sign of the largest Lyapunov exponent λ_1 is the most relevant to the stability analysis of typical trajectories, in view of the fact that $\lambda(\omega, x, v) = \lambda_1$ for v in an open and dense set. For this reason we frequently refer to λ_1 as "the" Lyapunov exponent. The *sum Lyapunov exponent* also turns out to be crucial:

$$\lambda_\Sigma = \sum_{j=1}^{\ell} m_j \lambda_j = \lim_{t \to \infty} \frac{1}{t} \log |\det D_x \Phi^t_\omega|,$$

which gives the asymptotic exponential expansion/compression of Lebesgue volume under the flow. Here, $m_j = \dim F_j - \dim F_{j+1}$ is the *multiplicity* of the j-th Lyapunov exponent.

2 Formulae for the Lyapunov exponents

Throughout this section, we assume Φ^t_ω is the stochastic flow of diffeomorphisms corresponding to the SDE (1.1) with associated Markov process $x_t = \Phi^t_\omega(x), x \in \mathbb{R}^n$.

2.1 The projective process

As we have seen, Lyapunov exponents are naturally viewed as depending on the tangent direction $v \in \mathbb{R}^n$ at which the derivative $D_x \Phi^t_\omega$ is evaluated. For this reason, to estimate Lyapunov exponents it is natural to consider an auxiliary process on *tangent directions* themselves. To this end, let $\mathbb{SR}^n = \mathbb{R}^n \times \mathbb{S}^{n-1}$ denote the unit tangent bundle of $\mathbb{R}^n$, where $\mathbb{S}^{n-1}$ is the unit sphere in $\mathbb{R}^n$. Given fixed initial $(x, v) \in \mathbb{SR}^n$, we define the process (v_t) on $\mathbb{S}^{n-1}$ by

$$v_t = \frac{D_x \Phi^t_\omega(v)}{|D_x \Phi^t_\omega(v)|}.$$

The full process $z_t = (x_t, v_t)$ on $\mathbb{SR}^n$ is Markovian, and in fact solves an SDE

$$dz_t = \tilde{X}_0(z_t)\, \mathrm{d}t + \sum_{i=1}^{r} \tilde{X}_i(z_t) \circ \mathrm{d}W^{(i)}_t,$$

where the 'lifted' fields $\tilde{X}_i$ are defined as

$$\tilde{X}_i(x, v) := (X_i(x), (I - \Pi_v) \nabla X_i(x) v).$$

Here, we have written $\Pi_v = v \otimes v$ for the orthogonal projection onto the span of $v \in \mathbb{S}^{n-1}$. Below, we denote the corresponding generator by

$$\tilde{\mathcal{L}} := \tilde{X}_0 + \frac{1}{2}\sum_{i=1}^r \tilde{X}_i^2.$$

Lyapunov exponents and stationary measures. Let (x_t, v_t) be a trajectory of the projective process with fixed initial $(x,v) \in \mathbb{SR}^n$, and observe that at integer times $t \in \mathbb{Z}_{>0}$, we have by (1.2)

$$\frac{1}{t}\log|D_x\Phi^t_\omega(v)| = \frac{1}{t}\sum_{i=0}^{t-1}\log|D_{x_i}\Phi^1_{\theta^i\omega}v_i|.$$

Hence, $\log|D_x\Phi^t_\omega|$ is an *additive observable* of (x_t, v_t), i.e., a sum iterated over the trajectory (x_t, v_t). Therefore, the strong law of large numbers for a Markov chain implies the following formula for the Lyapunov exponent:

Proposition 2.1 (See e.g. [56]). *Let ν be an ergodic stationary measure for (x_t, v_t). Assuming the integral is finite, for ν-a.e. initial $(x,v) \in \mathbb{SR}^{n-1}$ and $t \geq 0$ we have*

$$t\lambda(\omega, x, v) = \mathbf{E}\int \log|D_x\Phi^t_\omega v|\,\mathrm{d}\nu(x,v)\,.$$

with probability 1 ($\mathbf{E}$ denotes integration with respect to $d\mathbf{P}(\omega)$).

Moreover, if ν is the unique stationary measure for the (x_t, v_t) process, then for μ-a.e. x, and all $v \in \mathbb{R}^n$, *there holds $\lambda_1 = \lambda(\omega, x, v)$ with probability 1 and*

$$t\lambda_1 = \mathbf{E}\int \log|D_x\Phi^t_\omega v|\,\mathrm{d}\nu(x,v)\,. \tag{2.1}$$

Remark 2.2. This latter statement can be interpreted as saying that the existence of a unique stationary measure for the projective process gives a kind of non-degeneracy of the Oseledets' subspace $F_2(\omega, x)$ with respect to ω [56].

A time-infinitesimal version: the Furstenberg-Khasminskii formula. One of the key benefits of the SDE framework is the ability to take time derivatives, which turns dynamical questions (e.g., estimates of Lyapunov exponents, identification of stationary densities) into functional-analytic ones (e.g., solutions of degenerate elliptic or parabolic equations) for which many tools are available. Taking the time derivative of (2.1) gives what is known as the *Furstenberg-Khasminskii formula* (see e.g. [8,55]):

Proposition 2.3. *Assume (x_t, v_t) admits a unique stationary measure ν on $\mathbb{SR}^n$ projecting to a stationary measure μ on $\mathbb{R}^n$ for (x_t). For $(x,v) \in \mathbb{SR}^n$, define*

$$Q(x) = \operatorname{div} X_0(x) + \frac{1}{2}\sum_{i=1}^r X_i \operatorname{div} X_i(x)\,,$$
$$\tilde{Q}(x,v) = \operatorname{div}\tilde{X}_0(x,v) + \frac{1}{2}\sum_{i=1}^r \tilde{X}_i \operatorname{div}\tilde{X}_i(x,v)\,.$$

Then, provided $Q \in L^1(\mathrm{d}\mu)$ and $\tilde{Q} \in L^1(\mathrm{d}\nu)$, there holds

$$\lambda_\Sigma = \int Q\,\mathrm{d}\mu \quad \textit{and}$$
$$n\lambda_1 - \lambda_\Sigma = \int_{\mathbb{R}^n} Q\,\mathrm{d}\mu - \int_{\mathbb{SR}^n}\tilde{Q}\,\mathrm{d}\nu\,.$$

The first formula expresses $Q(x)$ as the time-infinitesimal rate at which $D_x\Phi^t_\omega$ compresses/expands Lebesgue measure, which in this formula is directly related to the asymptotic exponential volume growth/contraction rate λ_Σ. Similarly, $\tilde{Q}(x,v)$ is the time-infinitesimal rate at which $D_x\Phi^t_\omega$ compresses/expands volume on *the sphere bundle* $\mathbb{SR}^n = \mathbb{R}^n \times \mathbb{S}^{n-1}$. Roughly speaking, contraction of volumes along the $\mathbb{S}^{n-1}$ coordinate is associated with expansion in the Jacobian, while expansion of $\mathbb{S}^{n-1}$-volume is related to contraction in the Jacobian; this reversal is the reason for the minus sign in front of $\tilde{Q}$. For some intuition, observe that $(1,0)$ is a sink and $(0,1)$ is a source for the discrete-time system $v_n = A^n v/|A^n v|$ on S^1, where $A = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$.

2.2 Sign-definite formulas for Lyapunov exponents

The Furstenberg-Khasminskii formula is highly remarkable in that it reduces the problem of estimating Lyapunov exponents to computing the ensemble average of a single *deterministic* observable, $\tilde{Q}$, with respect to the stationary measure of (x_t, v_t). On the other hand, the formula itself is sign-indefinite, as $\tilde{Q}(x,v)$ takes on both positive and negative values as (x,v) is varied. This is reflective of the cancellation problem mentioned earlier in the estimation of Lyapunov exponents: tangent growth previously accumulated can be 'canceled out' by rotation into contracting directions later on in the trajectory. Hence, without a very precise characterization of ν, it would be very challenging to obtain any useful quantitative estimates on λ_1 from this formula.

Given the above, it makes sense to seek a *sign-definite* formula for the Lyapunov exponent. Below, given measures $\lambda, \eta, \eta \ll \lambda$ on measurable space X, the *relative entropy* $H(\eta|\lambda)$ of η given λ is defined by

$$H(\eta \,|\, \lambda) = \int_X \log\left(\frac{\mathrm{d}\eta}{\mathrm{d}\lambda}\right) \mathrm{d}\eta\,.$$

Observe that $H(\eta \,|\, \lambda) \geq 0$, while by strict convexity of $\log$ and Jensen's inequality, we have $H(\eta|\lambda) = 0$ iff $\eta = \lambda$. We also write $\hat{\Phi}^t_\omega : \mathbb{SR}^d \circlearrowleft$ for the stochastic flow associated to full lifted process (x_t, v_t) on $\mathbb{SR}^n$; that is, $\hat{\Phi}^t_\omega(x_0, v_0) = (x_t, v_t)$. Lastly, given a diffeomorphism Φ of a Riemannian manifold M and a density g on M, we define $\Phi_* g$ to be the density

$$\Phi_* g(x) = g \circ \Phi^{-1}(x) |\det D_x \Phi^{-1}|\,,$$

noting that if x is distributed like $g\,\mathrm{dVol}_M$, then $\Phi(x)$ is distributed like $\Phi_* g\,\mathrm{dVol}_M$.

The following deep formula has its roots in Furstenberg's seminal paper [41] and ideas à la Furstenberg have been developed by a variety of authors (e.g., [13, 29, 63, 79, 82]), and can be stated as follows: if $\nu \in \mathcal{P}(\mathbb{SR}^n)$ is a stationary probability measure for the projective process (x_t, v_t) and $\mathrm{d}\nu(x,v) = \mathrm{d}\nu_x(v)\mathrm{d}\mu(x)$ the disintegration of ν, then for all $t > 0$ there holds the following identity (often an inequality in more general settings).

Proposition 2.4 (See e.g. [13]). *Assume (x_t, v_t) admits a unique stationary measure ν with density $f = \frac{d\nu}{dq}$, where $\mathrm{d}q = \mathrm{dVol}_{\mathbb{SR}^n}$ is Riemannian volume measure on $\mathbb{SR}^n = \mathbb{R}^n \times \mathbb{S}^{n-1}$. Let μ be the corresponding stationary measure for (x_t) with density $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}x}$. Writing*

$$f_t := (\hat{\Phi}^t_\omega)_* f\,, \quad \rho_t := (\Phi^t_\omega)_* \rho\,,$$

we have (under the same integrability condition as Theorem 1.1)

$$\mathbf{E} H(\rho_t|\rho) = -t\lambda_\Sigma\,, \quad \textit{and} \quad \mathbf{E} H(f_t|f) = t(n\lambda_1 - 2\lambda_\Sigma)\,.$$

At least in simple settings, such as for SDEs with a unique stationary measure for the projective process, the formula follows from a slightly more subtle analysis of volume compression/expansion on $\mathbb{SR}^n$

suitably combined with ergodic theory. Furstenberg [41] was the first to relate relative entropy to Lyapunov exponents; at the generality above, the proof is due to Baxendale [13].

To explore the consequences of Proposition 2.4, let us re-write it into a more suggestive form. Let $f_x(v) = f(x,v)/\rho(x)$, $f_{t,x}(v) = f_t(x,v)/\rho_t(x)$ denote the *conditional densities* of f and f_t along the fiber $\mathbb{S}_x\mathbb{R}^n \simeq \mathbb{S}^{n-1}$. One can then combine the above formulae into the identity

$$\mathbf{E}H(f_t|f) - \mathbf{E}H(\rho_t|\rho) = \mathbf{E}\int_{\mathbb{R}^n} H(f_{t,x}|f_x)\mathrm{d}\mu(x) = t(n\lambda_1 - \lambda_\Sigma)\,. \tag{2.2}$$

The left-hand side of this identity is the expectation of a positive quantity, while the right-hand side is non-negative due to the general inequality $n\lambda_1 \geq \lambda_\Sigma$. By the strict convexity, we have

$$n\lambda_1 = \lambda_\Sigma \qquad \Longleftrightarrow \qquad f_{t,x} \equiv f_x \text{ with probability 1 for all } t \geq 0 \text{ and } \mu \text{ almost every } x.$$

Unraveling the definitions, $f_{t,x} \equiv f_x$ means that

$$(D_x\Phi^t_\omega)_* f_x = f_{\Phi^t_\omega(x)}\,,$$

i.e., the matrices $D_x\Phi^t_\omega$, viewed as acting on $\mathbb{S}^{n-1}$ embedded in $\mathbb{R}^n$, transform the conditional density f_x into the density $f_{\Phi^t_\omega(x)}$ of tangent directions at $\Phi^t_\omega(x)$. This is a very rigid condition in view of the fact that given any two (absolutely continuous) densities h, h' on $\mathbb{S}^{n-1}$,

$$\{A \in GL_n(\mathbb{R}) : A_* h = h'\}$$

has empty interior in the space of $n \times n$ matrices. One can obtain the following beautiful dichotomy by a more detailed analysis of the rigidity in a group of matrices in SL_n that preserve a given probability measure; see, e.g., [13, 41, 63].

Theorem 2.5 (Furstenberg Criterion)**.** *Suppose the same setting as Proposition 2.4. If $n\lambda_1 = \lambda_\Sigma$, then one of the following holds:*

(a) There is a continuously-varying family of inner products $x \mapsto \langle\cdot,\cdot\rangle_x$ with the property that $D_x\Phi^t_\omega$ is an isometry from $\langle\cdot,\cdot\rangle_x$ to $\langle\cdot,\cdot\rangle_{\Phi^t_\omega(x)}$ with probability 1 for all $t \geq 0$.

(b) There is a (locally) continuously-varying family of proper subspaces $x \mapsto L^i_x \subset \mathbb{R}^d$ with the property that $D_x\Phi^t_\omega(\cup_i L^i_x) = \cup_i L^i_{\Phi^t_\omega(x)}$ with probability 1 for all $t \geq 0$.

Remark 2.6. Note that in the above, the inner products and the L^i are *deterministic*, which is highly rigid for many random systems. Note that they are also *continuously-varying*.

However, if one is interested in deducing $\lambda_1 > 0$, this criterion is really only useful if $\lambda_\Sigma = 0$, i.e. the system is volume preserving, otherwise one only obtains the non-degeneracy $n\lambda_1 > \lambda_\Sigma$. Moreover, Theorem 2.5 lacks any quantitative information, and so it cannot be used to obtain concrete estimates with respect to parameters. Hence, it generally cannot be applied to dissipative systems, even weakly dissipative.

In the volume preserving case however, criteria à la Furstenberg can be a very powerful tool. In our previous work [17], we used a suitable (partially) infinite-dimensional extension of Theorem 2.5 to show that the Lagrangian flow map (i.e. the trajectories of particles in a fluid) is chaotic when the fluid evolves by the stochastically forced 2D Navier-Stokes equations (called *Lagrangian chaos* in the fluid mechanics literature). See Section 5 for more information.

2.3 The best of both worlds: sign-definite and time-infinitesimal

Proposition 2.4 is, on its face, a quantitative and sign-definite formula for Lyapunov exponents, and this leads to a strong and relatively easy-to-rule-out dichotomy for the degenerate scenario $n\lambda_1 = \lambda_\Sigma$. On the other hand, the formula itself is not straightforward to work with, requiring both the stationary density f for (x_t, v_t) as well as the time-t flow Φ^t_ω and its derivative $D_x\Phi^t_\omega$ as ω varies. In particular, it is unclear how to glean *quantitative* information beyond the "soft" inequality $n\lambda_1 > \lambda_\Sigma$, as would be relevant for a damped system (i.e., $\lambda_\Sigma < 0$).

In view of the sign-indefinite formula (2.1) and its time-infinitesimal version, the Furstenberg-Khasminskii formula, it is reasonable to hope that a time-infinitesimal version of Proposition 2.4 might exist. The authors establish such a formula in our recent work [16].

Proposition 2.7 (Theorem A in [16])**.** *Assume (x_t, v_t) has a unique stationary measure ν with density $f = \frac{\mathrm{d}\nu}{\mathrm{d}q}$ on $\mathbb{S}\mathbb{R}^n$. Let μ denote the corresponding stationary measure for (x_t) on $\mathbb{R}^n$ with density $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}x}$. Define the modified* Fisher informations

$$FI(f) = \frac{1}{2}\sum_{i=1}^{r}\int_{\mathbb{S}\mathbb{R}^n}\frac{|\tilde{X}_i^* f|^2}{f}\,\mathrm{d}q\,, \quad FI(\rho) = \frac{1}{2}\sum_{i=1}^{r}\int_{\mathbb{R}^n}\frac{|X_i^*\rho|^2}{\rho}\,\mathrm{d}x\,.$$

Under a mild moment criterion (see [16]), we have

$$FI(\rho) = -\lambda_\Sigma\,, \quad \text{and} \quad FI(f) = n\lambda_1 - 2\lambda_\Sigma\,.$$

Recall that $\tilde{X}_i^*$ denotes the adjoint of $\tilde{X}_i$ viewed as an operator on $L^2(dq)$.

Remark 2.8. One can show that $FI(f) - FI(\rho)$ corresponds to an analogous Fisher information on the conditional densities $\hat{f}_x(v)$, providing the exact time-infinitesimal analogue of (2.2) (see [16]).

These *Fisher-information*-type formulas for Lyapunov exponents enjoy many of the best qualities of the previous formulas: (A) they are sign-definite, like those in Proposition 2.4, and (B) are also time-infinitesimal like those in Proposition 2.3, and so are inherently simpler, requiring only the stationary density f for (x_t, v_t) and how it is acted on by the first-order differential operators $\tilde{X}_i^*$.

A key feature of Proposition 2.7 is that a lower bound on $FI(f)$ implies a lower bound on $n\lambda_1 - 2\lambda_\Sigma$. The $FI(f)$ itself has the connotation of a *partial regularity* of f along the forcing directions $\tilde{X}_i$. This is reminiscent of techniques in Hörmander's theory of hypoelliptic operators, where partial regularity along forcing directions implies regularity in *all* directions under an appropriate Lie algebra spanning condition involving the drift X_0. This connection is explored in the next section.

3 Quantitative lower bounds by the Fisher information

Let us now set about obtaining quantitative estimates on Lyapunov exponents using the Fisher information as in Proposition 2.7. For this, it will be most useful to consider the weakly-forced system

$$\mathrm{d}x_t = X_0^\epsilon(x_t)\,\mathrm{d}t + \sqrt{\epsilon}\sum_{k=1}^{r} X_k^\epsilon(x_t)\circ \mathrm{d}W_t^k, \tag{3.1}$$

where we have also allowed ϵ dependence in the vector fields X_j^ϵ. In this case, Proposition 2.7 gives the following Fisher information formula on the stationary density f^ϵ of the projective process associated to (3.1)

$$\frac{\epsilon}{2}\sum_{j=1}^{r}\int\frac{|\widetilde{X}_j^* f^\epsilon|^2}{f^\epsilon}\mathrm{d}q = n\lambda_1 - 2\lambda_\Sigma.$$

If $\widetilde{X}_j$ has a bounded divergence[7], by Cauchy-Schwarz, $\exists C > 0$ such that

$$\sum_{j=1}^{r} \|\widetilde{X}_j f^\epsilon\|_{L^1}^2 \leq C + FI(f^\epsilon) = \left(C + \frac{n\lambda_1 - 2\lambda_\Sigma}{\epsilon}\right).$$

Hence, we have related L^1-type directional regularity in the forcing directions to the Lyapunov exponents. *If* the lifted forcing directions $\{\widetilde{X}_j\}_{j=1}^r$ spanned the entire tangent space $T_w\mathbb{S}\mathbb{R}^n$ everywhere, then we would obtain a lower bound of the Lyapunov exponents of the type

$$\|f^\epsilon\|_{\dot{W}^{1,1}}^2 \lesssim \left(1 + \frac{n\lambda_1 - 2\lambda_\Sigma}{\epsilon}\right), \tag{3.2}$$

and so we find a straightforward lower bound on $n\lambda_1 - 2\lambda_\Sigma$ in terms of the regularity of f^ϵ. This kind of lower bound is clearly most useful if λ_Σ is small, especially $O(\epsilon)$, but crucially, it does not have to be *exactly* zero. In this manner, we can treat systems which are close to volume preserving, but not necessarily exactly volume preserving. This is at the crux of why we can treat systems like Lorenz-96 and Galerkin-Navier-Stokes whereas traditional à la Furstenberg methods based on e.g. Theorem 2.5, cannot.

3.1 Hypoellipticity

It is not usually the case that $\{\widetilde{X}_j\}_{j=1}^r$ spans $T_w\mathbb{S}\mathbb{R}^n$ and so the lower bound (3.2) is generally false. For example, for additive noise, the lifts satisfy $\widetilde{X}_j = (X_j, 0)$ and so clearly this fails to span $T_w\mathbb{S}\mathbb{R}^n$, regardless of whether or not $\{X_j\}_{j=1}^r$ spans $T_x\mathbb{R}^n$. Hence, in general, the Fisher information connects regularity in the lifted forcing directions to the Lyapunov exponents, but a priori, not any other directions in $T_w\mathbb{S}\mathbb{R}^n$. For this, we need a concept known as *hypoellipticity*, by which solutions to Kolmogorov equations such as (1.3) or (1.4), can be smooth even when $\mathcal{L}$ is degenerate, i.e. even when the forcing directions do not span the tangent space. This effect was studied first by Kolmogorov [58] in 1934, however clarity on the effect was not fully obtained until Hörmander's 1967 work [51].

Let us discuss Hörmander's main insights from [51]. It will make sense to quantify fractional regularity along a vector field X using the group e^{tX} and the L^p Hölder-type semi-norm (brushing aside minor technical details)

$$|h|_{X,s} := \sup_{t\in(-1,1)} |t|^{-s}\|e^{tX}h - h\|_{L^p}.$$

Hörmander's original work was based in L^2; our work will be based in L^1. For now, we set $p = 2$.

There are two key ideas in [51]. The first, and simpler idea, comes from the Campbell-Baker-Hausdorff formula, which implies for any two vector fields X, Y that (essentially, the Zassenhaus formula):

$$e^{-tX}e^{-tY}e^{tX}e^{tY} = e^{t^2[X,Y]+O(t^3)},$$

where here $[X, Y]$ is the Lie bracket, i.e. the commutator (see [50] and [51]). In particular, marching forward and then backward by two vector fields X, Y does not quite get back to where it started (unless X, Y commute). Therefore we have (using that the e^{tX} are bounded on L^p),

$$\|e^{t^2[X,Y]+O(t^3)} - I\|_{L^p} \lesssim \|e^{tX} - I\|_{L^p} + \|e^{tY} - I\|_{L^p} + \|e^{-tX} - I\|_{L^p} + \|e^{-tY} - I\|_{L^p},$$

which suggests the remarkable property that any fractional regularity of a function h in directions X, Y, i.e. $|h|_{X,s} + |h|_{Y,s} < \infty$, implies h also has (a little less) fractional regularity in the commutator direction $[X, Y]$. Another version of Campbell-Baker-Hausdorff (see [51]) gives

$$e^{t(X+Y)} = e^{tX}e^{tY}e^{t^2[X,Y]}...,$$

[7] This is not the case for our examples, but this will not be important as we will eventually work only locally.

where the $\ldots$ corresponds to a formal product expansion of higher commutators of tX and tY (and thus higher powers in t). Combined with the previous formal discussion, this suggests that regularity in directions X, Y should also supply regularity in the direction $X+Y$ (and indeed, any linear combination). By iterating these heuristics, we get the suggestion that a priori regularity along any set of vector fields $\{Z_0, ...Z_r\}$ should imply that there should also be some regularity in *any* direction $Z \in \mathrm{Lie}(Z_0, ..., Z_r)$, where the *Lie algebra* is given by the span of all possible combinations of commutators

$$\mathrm{Lie}(Z_0, \ldots, Z_r) := \mathrm{span}\{\mathrm{ad}(Y_m) \ldots \mathrm{ad}(Y_1) Y_0 : Y_j \in \{Z_0, Z_1, ..., Z_r\}\, m \geq 0\},$$

and where $\mathrm{ad}(X)Y := [X, Y]$. In [51], these heuristics are made rigorous with the following functional inequality: Suppose that $\forall z \in \mathbb{R}^n$, $\mathrm{Lie}_z(Z_0, ..., Z_r) = \{Z(z) : Z \in \mathrm{Lie}(Z_0, ..., Z_r)\} = T_z\mathbb{R}^n$. Then $\forall s_j \in (0, 1)$, $\exists s_\star$ such that for all $0 < s < s_\star$, $\forall R > 0$, and $\forall h \in C_c^\infty(B(0, R))$ there holds

$$\|h\|_{H^s} \lesssim_R \|h\|_{L^2} + \sum_{j=0}^{r} |h|_{Z_j, s_j}. \tag{3.3}$$

In particular, this inequality holds a priori for any $h \in C_c^\infty(B_R(0))$ and it has nothing to do directly with solutions to any PDE. Making this rigorous requires dealing with the errors in the CBH formulas used above. At any step of the argument, these errors are lower regularity but in new directions, and so dealing with this requires a little finesse and interpolations to close the argument.

Inequality (3.3) is already an interesting observation that can expand the directions of regularity. In particular, one can use an L^1 analogue of (3.3) to provide a lower bound on the Fisher information based on regularity in any direction contained in the Lie algebra of the forcing directions $\{\tilde{X}_1, ..., \tilde{X}_r\}$. However, Hörmander was *far* from done. Indeed, this is clearly unsatisfying to some degree as this will not even depend on the underlying deterministic dynamical system under consideration, encoded in the drift vector field $\widetilde{X}_0$. Moreover, for additive forcing, (3.3) fails to add anything at all. For Hörmander's second main insight, consider the backward Kolmogorov equation

$$\mathcal{L}g = Z_0 g + \frac{1}{2}\sum_{j=1}^{r} Z_j^2 g = F. \tag{3.4}$$

Assuming $\{Z_j\}_{j=0}^r$ have bounded divergence[8] one obtains the standard L^2 "energy" estimate:

$$\sum_{j=1}^{r} \|Z_j g\|_{L^2}^2 \lesssim \|g\|_{L^2}^2 + \|F\|_{L^2}^2.$$

After applying a smooth cutoff $\chi_R(x) = \chi(x/R)$ where $\chi \in C_c^\infty(B_2(0))$, $0 \leq \chi \leq 1$ and $\chi(x) = 1$ for $|x| \leq 1$ and dealing with the commutators as in a Caccioppoli estimate, the functional inequality (3.3) combined with this estimate implies that *if* $\mathrm{Lie}_z(Z_1, ..., Z_r) = T_z\mathbb{R}^n$ at all z, then we would obtain an estimate like

$$\|\chi_R g\|_{H^s} \lesssim_R \|g\|_{L^2(B_{2R}(0))} + \|F\|_{L^2(B_{2R}(0))}.$$

However, as discussed above, this condition on the vector fields is often too strong to be useful for us here.

However, another natural a priori estimate on g is available from (3.4). Indeed, pairing (3.4) with a test function φ we obtain

$$\left|\int \varphi Z_0 f \,\mathrm{d}q\right| \leq \frac{1}{2}\sum_{j=1}^{r} \|Z_j^* \varphi\|_{L^2} \|Z_j g\|_{L^2} \lesssim \frac{1}{2}\sum_{j=1}^{r} \left(\|\varphi\|_{L^2} + \|Z_j \varphi\|_{L^2}\right)\left(\|g\|_{L^2} + \|F\|_{L^2}\right).$$

[8] Alternatively, one can consider the estimates suitably localized.

This simple observation shows that for solutions of $\mathcal{L}g = F$, H^1-type regularity in the forcing directions automatically provides a corresponding dual H^{-1}-type regularity on $Z_0 g$. The cornerstone of [51] is the following functional inequality (i.e. again, not directly related to solutions of any PDEs): if $\mathrm{Lie}_z(Z_0, Z_1, ...Z_r) = \mathbb{R}^n$ everywhere, then $\exists s \in (0,1)$ such that if $R > 0$ and $h \in C_c^\infty(B_R(0))$, then

$$\|h\|_{H^s} \lesssim \|h\|_{L^2} + \sup_{\varphi : \|\varphi\|_{L^2} + \sum_{j=1}^r \|Z_j \varphi\|_{L^2} \le 1} \left| \int \varphi Z_0 h \, \mathrm{d}q \right| + \sum_{j=1}^r \|Z_j h\|_{L^2} =: \|h\|_{H^1_{\mathrm{hyp}}}. \tag{3.5}$$

The key heuristic behind this functional inequality is the following observation

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|e^{tZ_0}h - h\|_{L^2}^2 &= \left\langle e^{tZ_0}h - h, Z_0 e^{tZ_0} h \right\rangle \\ &\le \left(\|h\|_{L^2} + \sum_{j=1}^r \|Z_j e^{tZ_0^*}(e^{tZ_0}h - h)\|_{L^2} \right) \|g\|_{H^1_{\mathrm{hyp}}}. \end{aligned}$$

Therefore, *if* we had something like

$$\sum_{j=1}^r \|Z_j e^{tZ_0^*}(e^{tZ_0}h - h)\|_{L^2} \lesssim \sum_{j=1}^r \|Z_j h\|_{L^2}, \tag{3.6}$$

then we could combine the L^2 estimate on $\{Z_j\}_{j=1}^r$ with the corresponding dual negative regularity in the Z_0 direction to obtain some positive fractional regularity in the Z_0 direction, specifically we would have $1/2$ regularity from

$$\|e^{tZ_0}h - h\|_{L^2}^2 \lesssim t\|h\|_{H^1_{\mathrm{hyp}}}^2.$$

Unfortunately (3.6) doesn't generally hold[9], and Hörmander uses a rather ingenious regularization argument to turn this heuristic into reality. We shall henceforth call functional inequalities of the type (3.5) *Hörmander inequalities*.

The gain in regularity from (3.5) combines with the Kolmogorov equation to get the estimate

$$\|g\chi_R\|_{H^s} \lesssim \|g\|_{L^2(B_{2R}(0))} + \|F\|_{L^2(B_{2R}(0))},$$

and so provides an analogue of the gain of regularity when studying elliptic equations (though only fractional regularity). As in that theory, this regularity gain can be iterated to imply that any L^2 solution of $\mathcal{L}g = F$ is C^∞ if $F \in C^\infty$ [51].

3.2 Uniform hypoellipticity

Next, we want to make the arguments which are quantitative with respect to parameters, and hence we will introduce the notion of *uniform* hypoellipticity. Let us formalize the definition of Hörmander's condition for elliptic-type and parabolic-type equations. For a manifold M, we denote $\mathfrak{X}(M)$ the set of smooth vector fields on M.

Definition 3.1 (Hörmander's condition)**.** Given a manifold $\mathcal{M}$ and a collection of vector fields $\{Z_0, Z_1, \ldots, Z_r\} \subset \mathfrak{X}(\mathcal{M})$, we define collections of vector fields $\mathscr{X}_0 \subseteq \mathscr{X}_1 \subseteq \ldots$ recursively by

$$\begin{aligned} &\mathscr{X}_0 = \{Z_j \,:\, j \ge 1\}, \\ &\mathscr{X}_{k+1} = \mathscr{X}_k \cup \{[Z_j, Z] \,:\, Z \in \mathscr{X}_k, \quad j \ge 0\}. \end{aligned}$$

[9] As in the easier inequalities above, the heuristic (3.6) neglects the creation of higher order commutators, in fact one requires regularity in many other directions in $\mathrm{Lie}(Z_0, ..., Z_r)$ as a result.

We say that $\{Z_i\}_{i=0}^r$ satisfies the *parabolic Hörmander condition* if there exists k such that for all $w \in \mathcal{M}$,

$$\text{span}\,\{Z(w)\,:\,Z \in \mathscr{X}_k\} = T_w\mathcal{M}.$$

We say that $\{Z_i\}_{i=0}^r$ satisfies the *(elliptic) Hörmander condition* if this holds with $\mathscr{X}_0 = \{Z_j : j \geq 0\}$.

Note that the parabolic Hörmander condition is slightly stronger than the elliptic Hörmander condition.

Definition 3.2 (Uniform Hörmander's condition)**.** Let $\mathcal{M}$ be a manifold, and let $\{Z_0^\epsilon, Z_1^\epsilon, ..., Z_r^\epsilon\} \subset \mathfrak{X}(\mathcal{M})$ be a set of vector fields parameterized by $\epsilon \in (0, 1]$. With $\mathscr{X}_k$ defined as in Definition 3.1 in the parabolic case (resp. elliptic), we say $\{Z_0^\epsilon, Z_1^\epsilon, ..., Z_r^\epsilon\}$ satisfies the uniform parabolic (resp. elliptic) Hörmander condition on $\mathcal{M}$ if $\exists k \in \mathbb{N}$, such that for any open, bounded set $U \subseteq \mathcal{M}$ there exist constants $\{K_n\}_{n=0}^\infty$, such that for all $\epsilon \in (0, 1]$ and all $x \in U$, there is a finite subset $V(x) \subset \mathscr{X}_k$ such that $\forall \xi \in T_x\mathcal{M}$,

$$|\xi| \leq K_0 \sum_{Z \in V(x)} |Z(x) \cdot \xi| \qquad \sum_{Z \in V(x)} \|Z\|_{C^n} \leq K_n.$$

This definition stipulates that any ϵ dependence is *locally* (on the manifold) uniform in terms of both regularity and spanning. Now we are ready to state the uniform L^1-type Hörmander inequality suitable for use with the Fisher information, proved in [16]. There are many works extending Hörmander's theory in various ways see e.g., [2, 5, 20, 45, 57, 62, 72] and the references therein. However, as far as the authors are aware, there are no works in the $L^1 - L^\infty$ framework. We also need to consider the forward Kolmogorov equation $\tilde{\mathcal{L}}^* f = 0$, as opposed to the case of the backward Kolmogorov equation considered by Hörmander [51]; this changes some details but little of significant consequence is different.

Theorem 3.3 (L^1-type uniform Hörmander inequality; Theorem 4.2 [16])**.** *let $\{X_0^\epsilon, X_1^\epsilon, ..., X_r^\epsilon\}$ be a collection of vector fields on $\mathbb{S}\mathbb{R}^n$ satisfying the uniform elliptic Hörmander condition as in Definition 3.2. Then, $\exists s_\star \in (0, 1)$ such that if $B_R(x_0) \subset \mathbb{R}^n$ is an open ball and $h \in C_c^\infty(B_R(x_0) \times \mathbb{S}^{n-1})$, then for all $0 < s < s_\star$, $\exists C = C(R, x_0, s)$ such that $\forall \epsilon \in (0, 1)$ there holds the following fractional regularity*[10] *estimate* uniformly in ϵ

$$\|h\|_{W^{s,1}} \leq C \left(\|h\|_{L^1} + \sup_{\varphi : \|\varphi\|_{L^\infty} + \sum_{j=1}^r \|X_j^\epsilon \varphi\|_{L^\infty} \leq 1} \left| \int \varphi (X_0^\epsilon)^* h \, \mathrm{d}q \right| + \sum_{j=1}^r \|(X_j^\epsilon)^* h\|_{L^1} \right).$$

In particular, applying a smooth cutoff $\chi_R := \chi(x/R)$ for some $\chi \in C_c^\infty(B_2(0))$ with $0 \leq \chi \leq 1$ and $\chi \equiv 1$ if $|x| \leq 1$ to the Kolmogorov equation $\tilde{\mathcal{L}}^* f^\epsilon = 0$ (assuming also $\|f^\epsilon\|_{L^1} = 1$) and suitably estimating the commutators, we obtain

$$\|\chi_R f^\epsilon\|_{W^{s,1}}^2 \lesssim_R 1 + FI(f^\epsilon). \tag{3.7}$$

Remark 3.4. Hypoellipticity plays a classical role in the theory of SDEs. In particular, the parabolic Hörmander condition of Definition 3.1 is exactly the condition most often used to deduce that the Markov semigroup $\mathcal{P}_t$ is strong Feller (the exposition of [47] is especially intuitive). The parabolic Hörmander condition also often plays a role in proving irreducibility via geometric control theory (see discussions in [43, 49, 53] and specifically in [16] in regards to the projective process). For many applications, it is likely that the parabolic Hörmander's condition will be used to prove that there exists a unique stationary measure

[10] For $s \in (0, 1)$ we may define $W^{s,1}$ on a geodesically complete, n-dimensional Riemannian manifold with bounded geometry $\mathcal{M}$ as

$$\|w\|_{W^{s,1}} = \|w\|_{L^1} + \left(\int_{\mathcal{M}} \int_{h \in T_x\mathcal{M} : |h| < \delta_0} \frac{|w(\exp_x h) - w(x)|}{|h|^{s+n}} \mathrm{d}h \mathrm{d}q(x) \right),$$

where $\exp_x : T_x\mathcal{M} \to \mathcal{M}$ is the exponential map on $\mathcal{M}$ and $\mathrm{d}q$ is the Riemannian volume measure. See e.g. [81] for more details.

ν for the projective process (via Doob-Khasminskii [34]), as required to apply Proposition 2.7. Hence the condition of uniformity-in-ϵ in Definition 3.2 will usually be the only additional information required to apply Theorem 3.3.

Remark 3.5. Quantitative arguments based on L^2 Hörmander inequalities can be found in [3,20] (completed concurrently with or after [16]). Thinking about hypoellipticity in terms of functional inequalities, rather than qualitative statements about regularity of solutions to PDEs, has other important advantages as well, for example, it is easier to adapt classical elliptic and parabolic PDE methods, such as De Giorgi or Moser iterations, into hypoelliptic equations [20,45,72].

Obtaining the above Theorem 3.3 follows an argument generally based on Hörmander's original paper [51], however, the L^1-L^∞ framework, as opposed to the self-dual L^2 framework in [51], necessitates a more complicated regularization argument than that used [51] (which was already quite delicate!). Moreover, as we are always interested in sphere bundles here, one cannot avoid working on smooth manifolds, which at least under the assumption of geodesic completeness, only adds some technical complexity rather than fundamental difficulties.

Let us briefly see, heuristically, how one would approach the proof of Theorem 3.3. Motivated by the above discussion regarding [51], the main challenge is to obtain $1/2$ of a derivative of L^1 Hölder-type regularity in the $\tilde{X}_0^*$ "direction". By a bootstrap-type argument, we may assume that we have corresponding regularity along all of the other vector fields in $\mathrm{Lie}_z(\tilde{X}_0, \tilde{X}_1, ..., \tilde{X}_r)$ (see [16] for details). Let S_t be a (carefully designed) regularization operator $S_t : L^p \to L^p$. We obtain for any $w \in C_c^\infty(B_R(0) \times \mathbb{S}^{n-1})$,

$$\|e^{t\tilde{X}_0^*} w - w\|_{L^1} \leq \|e^{t\tilde{X}_0^*}\left(S_\tau^* w - w\right)\|_{L^1} + \|S_\tau^* w - w\|_{L^1} + \|e^{t\tilde{X}_0^*} S_\tau^* w - S_\tau^* w\|_{L^1}.$$

We eventually set $\tau \sim \sqrt{t}$ and the regularization operator will be designed so that the first two terms are $O(\tau)$ so we need mainly to work on the latter term, which by duality is estimated by

$$\|e^{t\tilde{X}_0^*} S_\tau^* w - S_\tau^* w\|_{L^1} \leq \sup_{\|v\|_{L^\infty} \leq 1} \left| \int_0^t \int_{\mathbb{S}\mathbb{R}^n} \left(e^{s\tilde{X}_0} v\right) X_0^* S_\tau^* w \, \mathrm{d}q \mathrm{d}s \right|,$$

and for any fixed $v \in L^\infty$ we have

$$\begin{aligned} \left| \int_{\mathbb{S}\mathbb{R}^n} (e^{s\tilde{X}_0} v) X_0^* S_\tau^* w \, \mathrm{d}q \right| &\leq \left| \int_{\mathbb{S}\mathbb{R}^n} (e^{s\tilde{X}_0} v) [\tilde{X}_0, S_\tau]^* w \, \mathrm{d}q \right| + \left| \int_{\mathbb{S}\mathbb{R}^n} (S_\tau e^{s\tilde{X}_0} v) \tilde{X}_0^* w \, \mathrm{d}q \right| \\ &\leq \|e^{sX_0} v\|_{L^\infty} \|[\tilde{X}_0, S_\tau]^* w\|_{L^1} + \left(\|S_\tau e^{s\tilde{X}_0} v\|_\infty + \sum_{j=1}^r \|X_j S_\tau e^{s\tilde{X}_0} v\|_{L^\infty} \right) \mathfrak{D}(w), \end{aligned}$$

where

$$\mathfrak{D}(h) := \sup_{\varphi : \|\varphi\|_{L^\infty} + \sum_{j=1}^r \|X_j^\epsilon \varphi\|_{L^\infty} \leq 1} \left| \int_{\mathbb{S}\mathbb{R}^n} \varphi (\tilde{X}_0^\epsilon)^* h \, \mathrm{d}q \right|.$$

Hence, the challenge is designing a regularizer such that the commutator $[\tilde{X}_0, S_\tau]^*$ loses only $O(\tau^{-1})$ using no a priori regularity in the $\tilde{X}_0$ direction and similarly that S_τ regularizes the forcing fields $\tilde{X}_j$ like $O(\tau^{-1})$. To do this, we let S_τ be a modified version of Hörmander's regularizer, which averages the function along directions in $\mathrm{Lie}_z(\tilde{X}_0, ...\tilde{X}_r)$ a corresponding amount (higher commutators corresponding to less regularization) in a carefully ordered way. Specifically, because these 'directional mollifiers' do not commute, the order in which they are applied is very important. Hörmander regularized with S_τ, whereas we are fundamentally regularizing with its adjoint S_τ^*, which reverses the delicate ordering. Despite the added difficulty, this turns out to be an important choice for our framework.

4 Chaos for 2d Galerkin-Navier-Stokes and related models

In this section, we outline how to apply the above ideas to prove a positive Lyapunov exponent for Galerkin truncations of the stochastic 2d Navier-Stokes. A general class of models with similar bilinear drift term, which we call *Euler-like* systems, are given by the following SDE:

$$\mathrm{d}x_t^\epsilon = (B(x_t^\epsilon, x_t^\epsilon) - \epsilon A x_t^\epsilon)\mathrm{d}t + \sum_{k=1}^{r} X_k \mathrm{d}W_t^k\,. \tag{4.1}$$

Here, $\{X_k\}_{k=1}^r$ is a collection of *constant* (x-independent) forcing vector fields (i.e. additive forcing) while $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is a nontrivial (not identically zero) bilinear drift that satisfies

$$\operatorname{div} B = 0\,, \quad x \cdot B(x,x) = 0\,,$$

so in particular the unforced $\epsilon = 0$ dynamics preserve the norm[11], given by $\frac{1}{2}\|x\|^2$, and volume in $\mathbb{R}^n$ (i.e. the Liouville property). The term $-\epsilon A$ provides weak linear damping, where A is assumed to be a symmetric, positive-definite $n \times n$ matrix. Stochastically forced versions of the Lorenz 96 model (L96) [67], Galerkin truncations of 2d and 3d Navier-Stokes on a torus (of arbitrary aspect ratio) [21, 37, 78] and truncations of commonly used shell models for turbulence [35, 44, 68, 84] can be cast in this form. The 2D stochastic Galerkin-Navier-Stokes equations will be described in more detail in Section 4.3 below.

The bilinearity of B implies that solutions can be naturally rescaled into a weakly-damped, weakly-driven system, and the two scalings are equivalent as far as Lyapunov exponents are concerned. Indeed, while the scaling (4.1) is common among models of complex real-world systems, the stationary measure μ has characteristic energy $\int |x|^2 \mathrm{d}\mu(x) \approx \epsilon^{-1}$. Since we are concerned with the regime $\epsilon \ll 1$, it is natural to rescale and consider a weakly-damped, weakly-driven system. Hence, it is more natural to re-scale so that the long-time behavior remains bounded and non-vanishing as $\epsilon \to 0$. By rescaling $x_t^\epsilon \mapsto \sqrt{\epsilon}x^\epsilon_{\sqrt{\epsilon}t}$, replacing $\epsilon \mapsto \epsilon^{3/2}$, and using the self-similarity of Brownian motion is equivalent in law to the weakly-driven, weakly damped form

$$\mathrm{d}x_t^\epsilon = (B(x_t^\epsilon, x_t^\epsilon) - \epsilon A x_t^\epsilon)\mathrm{d}t + \sqrt{\epsilon}\sum_{k=1}^{r} X_k \mathrm{d}W_t^k. \tag{4.2}$$

Most importantly, this rescaling does not affect our results on Lyapunov exponents, since upon setting $\hat{\epsilon} = \epsilon^{3/2}$, the Lyapunov exponent $\hat{\lambda}_1^{\hat{\epsilon}}$ of (4.2) with parameter $\hat{\epsilon}$ is related to the Lyapunov exponent λ_1^ϵ of (4.1) by the identity $\frac{\hat{\lambda}_1^{\hat{\epsilon}}}{\hat{\epsilon}} = \frac{\lambda_1^\epsilon}{\epsilon}$. This kind of scaling is sometimes called *fluctuation-dissipation* due to the balance between the forcing and the dissipation.

For this class of systems (4.1), our result below gives a sufficient condition for a positive Lyapunov exponent in terms of projective hypoellipticity, i.e., if the lifted vector fields $\{\tilde{X}_0^\epsilon, \tilde{X}_1, \ldots \tilde{X}_r\}$ corresponding to the projective process $(x_t^\epsilon, v_t^\epsilon)$ (denoting $X_0^\epsilon(x) = B(x,x) - \epsilon Ax$) satisfy Hörmander's condition on $\mathbb{SR}^n$.

Theorem 4.1 (Theorem C; [16])**.** *Assume that*

(i) $\{\tilde{X}_0^\epsilon, \tilde{X}_1, \ldots, \tilde{X}_r\}$ *satisfy the elliptic Hörmander's condition uniformly in* $\epsilon \in (0,1)$ *as in Definition 3.2;*

(ii) the bilinear term B *is nontrivial, i.e.,* $B(x,x) \neq 0$ *for some* $x \in \mathbb{R}^n$*; and*

(iii) the process $(x_t^\epsilon, v_t^\epsilon)$ *admits a unique stationary density* f^ϵ.

[11] In the case of the vorticity form of the 2D Navier-Stokes equations that we will be studying below, this quantity is the *enstrophy*.

Then, the limit defining the Lyapunov exponent λ_1^ϵ *of* (4.1) *exists, and satisfies*

$$\lim_{\epsilon \to 0} \frac{\lambda_1^\epsilon}{\epsilon} = \infty\,.$$

In particular, $\exists \epsilon_0 > 0$ *such that for all* $\epsilon \in (0, \epsilon_0)$ *there holds* $\lambda_1^\epsilon > 0$.

A sketch of the proof of Theorem 4.1 is given in Section 4.1 below. The most difficult part of applying this result to a concrete system, e.g., Galerkin-Navier-Stokes, is to prove the parabolic Hörmander condition for the projective process: general comments on this problem are given in Section 4.2, while the issue of affirming this for Galerkin-Navier-Stokes is taken up in Section 4.3.

Given parabolic Hörmander's condition, unique existence of f^ϵ follows, via the Doob-Khasminskii theorem, from topological irreducibility of $(x_t^\epsilon, v_t^\epsilon)$, i.e., the ability to approximately control random trajectories by controlling noise paths. For Euler-like models such as (4.1), this follows from geometric control theory arguments and the following well-known cancellation condition on $B(x, x)$ (known to hold for many models such as Galerkin Navier-Stokes, c.f. [43, 49]): there exists a collection of vectors $\{e_1, \dots e_s\} \subset \mathbb{R}^n$ with

$$\operatorname{span}\{e_1, \dots e_s\} = \operatorname{span}\{X_1, \dots, X_r\}$$

such that for each $1 \le k \le s$, $B(e_k, e_k) = 0$. For more details see Section 5.3 of [16].

Remark 4.2. The inverse Lyapunov exponent $(\lambda_1^\epsilon)^{-1}$ is sometimes called the Lyapunov time, and is the "typical" length of time one must wait for tangent vectors to grow by a factor of e. Thus, the estimate $\lambda_1^\epsilon \gg \epsilon$ implies that the Lyapunov time is $\ll \epsilon^{-1}$. On the other hand, ϵ^{-1} is the typical amount of time it takes for the Brownian motion $\sqrt{\epsilon} W_t$ to reach an $O(1)$ magnitude; for this reason it is reasonable to refer to ϵ^{-1} as a kind of "diffusion timescale". So, stated differently, our results indicate that as $\epsilon \to 0$, arbitrarily many Lyapunov times elapse before a single "diffusion time" has elapsed, indicating a remarkable sensitivity of the Lyapunov exponent to the presence of noise.

Based on these ideas, one would like to assert that the scaling $\lambda_1^\epsilon \gg \epsilon$ implies that the deterministic dynamics are "close" to positive Lyapunov exponent dynamics, agnostic as to whether the zero-noise system has a positive exponent on a positive area set. However, this assertion does not follow from the scaling $\lambda_1^\epsilon \gg \epsilon$ alone: even if the Brownian motion itself is small, there could already be a substantial difference between random and corresponding deterministic (zero-noise) trajectories well before time ϵ^{-1}, e.g., if there is already strong vector growth in the deterministic dynamics. For more on this, see the open problems in Section 6.

4.1 Zero-noise limit and rigidity: Proof sketch of Theorem 4.1

Applying the Fisher information identity (Proposition 2.7) to the Euler-like system (4.2) and using that $\lambda_\Sigma^\epsilon = -\epsilon \operatorname{tr} A$, we obtain

$$FI(f^\epsilon) = \frac{n\lambda_1^\epsilon}{\epsilon} + 2 \operatorname{tr} A.$$

By the regularity lower bound (3.7), this implies that for each open ball $B_R(0)$ we have the lower bound

$$\|\chi_R f^\epsilon\|_{W^{s,1}}^2 \lesssim_R 1 + \frac{\lambda_1^\epsilon}{\epsilon},$$

where the regularity $s \in (0, 1)$ and the implicit constant $C = C_R$ are independent of ϵ.

From this, we see that if $\liminf_\epsilon \epsilon^{-1} \lambda_1^\epsilon$ were to remain bounded, then f^ϵ would be bounded in $W^{s,1}_{\rm loc}$ uniformly in ϵ. As $W^{s,1}$ is locally compactly embedded in L^1 and f^ϵ naturally satisfies certain uniform-in-ϵ moment bounds, one can deduce, by sending $\epsilon \to 0$, that at least one of the following must hold true (see Proposition 6.1, [16] for details):

(a) either $\lim_{\epsilon\to 0} \frac{\lambda_1^\epsilon}{\epsilon} = \infty$; or

(b) the zero-noise flow (x_t^0, v_t^0) admits a stationary density $f^0 \in L^1(\mathbb{SR}^n)$.

Let us consider alternative (b). While it is natural and common for the projective processes of SDE to admit stationary densities, the existence of an absolutely continuous invariant measure $f^0 \mathrm{d}q$ for the projective process of the $\epsilon = 0$ problem

$$\dot{x}_t = B(x_t, x_t), \tag{4.3}$$

is quite rigid. Indeed, in view of the fact that vector growth implies concentration of Lebesgue measure in projective space (c.f. the discussion in Section 2.1 after Proposition 2.3), the existence of an invariant density essentially rules out *any* vector growth for the $\epsilon = 0$ projective process (x_t^0, v_t^0). Precisely, a generalization of Theorem 2.32 in [9] (see [16] for details) implies that there is a measurably varying Riemannian metric $x \mapsto g_x$ such that Φ^t is an *isometry* with respect to g_x, namely

$$g_x(D_x\Phi^t v, D_x\Phi^t w) = g_{\Phi^t(x)}(v, w), \quad v, w \in T_x\mathbb{R}^n\,,$$

where $\Phi^t : \mathbb{R}^n \to \mathbb{R}^n$ is the flow associate to the $\epsilon = 0$ dynamics (4.3). Hence, we see that if $\liminf_{\epsilon\to 0} \epsilon^{-1}\lambda_1 < \infty$, then the deterministic, measure-preserving $\epsilon = 0$ dynamics must be in a situation analogous to possibility (a) in Theorem 2.5.

In our setting, we show that there is necessarily *some* norm growth as $t \to \infty$ for the $\epsilon = 0$ dynamics due to *shearing between conserved energy shells* $\{x \in \mathbb{R}^n \,:\, |x|^2 = E\}$. This is straightforward to check: due to the scaling symmetry $\Phi^t(\alpha x) = \alpha\Phi^{\alpha t}(x)$, $\alpha > 0$, we have the following orthogonal decomposition of the linearization $D_x\Phi^t$ in the direction $x \in \mathbb{R}^n$:

$$D_x\Phi^t x = \Phi^t(x) + tB(\Phi^t(x), \Phi^t(x))\,,$$

noting that $y \cdot B(y, y) \equiv 0$ for all $y \in \mathbb{R}^n$. Hence, one obtains the lower bound

$$|D_x\Phi^t| \geq t\frac{|B(\Phi^t(x), \Phi^t(x))|}{|x|}$$

for each $x \in \mathbb{R}^n\backslash\{0\}$ and each $t > 0$. This contradicts the existence of the Riemannian metric g_x via a Poincaré recurrence argument and the fact that the set of stationary points $\{x \in \mathbb{R}^n \,:\, B(x,x) = 0,\, |x|^2 \leq R\}$ is a zero volume set. This is summarized in the following proposition (a proof of which is given in [16]).

Proposition 4.3 (Proposition 6.2 [16])**.** *Assume that the bilinear mapping B is not identically 0. Let ν be any invariant probability measure for $\hat{\Phi}^t$ (the flow corresponding to the (deterministic) $\epsilon = 0$ projective process) with the property that $\nu(A \times \mathbb{S}^{n-1}) = \mu(A)$, where $\mu \ll \mathrm{Leb}_{\mathbb{R}^n}$. Then, ν is singular with respect to volume measure $\mathrm{d}q$ on $\mathbb{SR}^n$.*

4.2 Verifying projective hypoellipticity: A sufficient condition

We address here the challenge of verifying the parabolic Hörmander condition on the sphere bundle $\mathbb{SR}^n$. Recall that given a smooth vector field X on $\mathbb{R}^n$ we define its lift $\tilde{X}$ to the sphere bundle $\mathbb{SR}^n$ by

$$\tilde{X}(x, v) = (X(x), \nabla X(x)v - v\langle v, \nabla X(x)v\rangle)\,,$$

where $\nabla X(x)$ denotes the (covariant) derivative of X at x and is viewed as a linear endomorphism on $T_x\mathbb{R}^n$. Many of the following general observations about the lifted fields were made in [13]; see also [16] for detailed discussions.

An important property is that the lifting operation can be seen to be a Lie algebra isomorphism onto its range with respect to the Lie bracket, i.e., $[\tilde{X}, \tilde{Y}] = [X, Y]\tilde{}$. Using this observation, the parabolic

Hörmander condition (see Definition 3.1) on $\mathbb{SR}^n$ for the lifts of a collection of vector fields $\{X_0, X_1, \ldots, X_r\} \subset \mathfrak{X}(\mathbb{R}^n)$ can be related to non-degeneracy properties of the Lie sub-algebra $\mathfrak{m}_x(X_0; X_1, \ldots, X_r)$ of $\mathfrak{sl}(T_x\mathbb{R}^n)$ defined by

$$\mathfrak{m}_x(X_0; X_1, \ldots, X_r) := \left\{ \nabla X(x) - \tfrac{1}{n} \operatorname{div} X(x) \operatorname{Id} \,:\, X \in \operatorname{Lie}(X_0; X_1, \ldots, X_r)\,,\ X(x) = 0 \right\},$$

where

$$\operatorname{Lie}(X_0; X_1, \ldots, X_r) := \operatorname{Lie}(X_1, \ldots, X_r, [X_0, X_1], \ldots, [X_0, X_r]),$$

is the *zero-time ideal* generated by $\{X_0, X_1, \ldots, X_r\}$, with X_0 a distinguished "drift" vector field (recall that $\mathfrak{sl}_n(T_x\mathbb{R}^n)$ is the Lie algebra of traceless linear endomorphisms of $T_x\mathbb{R}^n$).

Particularly, if for each $x \in \mathbb{R}^n$, $\mathfrak{m}_x(X_0; X_1, \ldots, X_r)$ acts *transitively* on $\mathbb{S}^{n-1}$ in the sense that for each $(x, v) \in \mathbb{SR}^n$ one has

$$\{Av - v\langle v, Av\rangle \,:\, A \in \mathfrak{m}_x(X_0; X_1, \ldots, X_r)\} = T_v\mathbb{S}^{n-1}, \tag{4.4}$$

then the parabolic Hörmander condition for $\{X_0, X_1, \ldots, X_r\}$ on $\mathbb{R}^n$ is equivalent to the parabolic Hörmander condition for the lifts $\{\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_r\}$ on $\mathbb{SR}^n$. Moreover, the uniform parabolic Hörmander condition is satisfied on $\mathbb{SR}^n$ if and only if it is satisfied on $\mathbb{R}^n$ and (4.4) holds uniformly in the same sense as Definition 3.2. Since $\mathfrak{sl}(\mathbb{R}^n)$ acts transitively on $\mathbb{R}^n \backslash \{0\}$ (see for instance [28]), a sufficient condition for transitivity on $\mathbb{S}^{n-1}$ is

$$\mathfrak{m}_x(X_0; X_1, \ldots X_r) = \mathfrak{sl}(T_x\mathbb{R}^n).$$

In the specific case of Euler-like models (4.2) with $X_0^\epsilon(x) = B(x, x) - \epsilon A$ and $\{X_k\}_{k=1}^r$ as in (4.2), the situation can be simplified if $\operatorname{Lie}(X_0; X_1, \ldots, X_r)$ contains the constant vector fields $\{\partial_{x_k}\}_{k=1}^n$. In this case, the family of x and ϵ-independent endomorphisms

$$H_k := \nabla[\partial_{x_k}, X_0^\epsilon] = \nabla[\partial_{x_k}, B], \quad k = 1, \ldots n,$$

generate the Lie algebra $\mathfrak{m}_x(X_0^\epsilon; X_1, \ldots, X_r)$ at all $x \in \mathbb{R}^n$. This argument implies the following sufficient condition for projective spanning.

Corollary 4.4 (See [16])**.** *Consider the bilinear Euler-like models* (4.2)*. If* $\operatorname{Lie}(X_0; X_1, \ldots, X_r)$ *contains* $\{\partial_{x_k}\}_{k=1}^n$*, then* $\{\tilde{X}_0^\epsilon, \tilde{X}_1, \ldots \tilde{X}_r\}$ *satisfy the uniform parabolic Hörmander condition (in the sense of Definition 3.2) on* $\mathbb{SR}^n$ *if*

$$\operatorname{Lie}(H^1, \ldots, H^n) = \mathfrak{sl}(\mathbb{R}^n).$$

This criterion is highly useful, having reduced projective spanning to a question about a single Lie algebra of trace-free matrices.

In [16], we verified this condition directly for the Lorenz 96 system [67], which is defined for n unknowns in a periodic array by the nonlinearity B given by

$$B_\ell(x, x) = x_{\ell+1}x_{\ell-1} - x_{\ell-2}x_{\ell-1}. \tag{4.5}$$

The traditional case is $n = 40$, but it can be considered in any finite dimension. In particular we proved the following.

Corollary 4.5 (Corollary D; [16])**.** *Consider the L96 system given by* (4.2) *with the nonlinearity* (4.5) *and* $X_k = q_k e_k$ *for* $k \in \{1, .., r\}$*,* $q_k \in \mathbb{R}$*, and* e_k *the canonical unit vectors. If* $q_1, q_2 \neq 0$ *and* $n \geq 7$*, then*

$$\lim_{\epsilon \to 0} \frac{\lambda_1^\epsilon}{\epsilon} = \infty\,.$$

In particular $\exists \epsilon_0 > 0$ *such that* $\lambda_1^\epsilon > 0$ *if* $\epsilon \in (0, \epsilon_0)$.

4.3 Projective hypoellipticity for 2d Galerkin-Navier-Stokes

Let's now see how we can go about verifying the projective hypoellipticity condition for a high-dimensional model of physical importance, namely Galerkin truncations of the 2d stochastic Navier-Stokes equations on the torus of arbitrary side-length ratio $\mathbb{T}^2_r = [0, 2\pi) \times [0, \frac{2\pi}{r})$ (periodized) for $r > 0$. Recall that the Navier-Stokes equations on $\mathbb{T}^2_r$ in vorticity form are given by

$$\partial_t w + u \cdot \nabla w - \epsilon \Delta w = \sqrt{\epsilon}\dot{W}_t,$$

where u is the divergence free velocity field satisfying the Biot-Savart law $u = \nabla^{\perp}(-\Delta)^{-1}w$ and $\dot{W}_t$ is a white-in time, colored-in-space Gaussian forcing which we will take to be diagonalizable with respect to the Fourier basis with Fourier transform supported on a small number of modes.

In the work [21] by the first and last authors of this note, we consider a Galerkin truncation of the 2d stochastic Navier-Stokes equations at an arbitrary frequency $N \geq 1$ in Fourier space by projecting onto the Fourier modes in the truncated lattice

$$\mathbb{Z}^2_{0,N} := \{(k_1, k_2) \in \mathbb{Z}^2 \backslash \{0\} \,:\, \max\{|k_1|, |k_2|\} \leq N\} \subseteq \mathbb{Z}^2,$$

giving rise to a $n = |\mathbb{Z}^2_{0,N}| = (2N + 1)^2 - 1$ dimensional stochastic differential equation with the reality constraint $w_{-k} = \overline{w}_k$ for $w = (w_k) \in \mathbb{C}^{\mathbb{Z}^2_{0,N}}$ (that is, the vector is indexed over $\mathbb{Z}^2_{0,N}$) governed by

$$\mathrm{d}w_k = (B_k(w, w) - \epsilon|k|^2_r w_k)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}W^k, \tag{4.6}$$

where $|k|^2_r := k_1^2 + r^2 k_2^2$, and $W^k_t = \alpha_k W^{a,k}_t + i\beta_k W^{b,k}_t$ are independent complex Wiener processes satisfying $W^k_t = \overline{W}^{-k}_t$ ($W^{a,k}_t, W^{b,k}_t$ are standard iid Wiener processes) with α_k, β_k arbitrary such that $\alpha_k = 0 \Leftrightarrow \beta_k = 0$. The symmetrized non-linearity $B_k(w, w)$ is given by

$$B_k(w, w) := \frac{1}{2} \sum_{j+\ell=k} c_{j,\ell} w_j w_\ell, \quad c_{j,\ell} := \langle j^{\perp}, \ell \rangle_r \left(\frac{1}{|\ell|^2_r} - \frac{1}{|j|^2_r} \right)$$

where the sum runs over all $j, \ell \in Z^2_{0,N}$ such that $j + \ell = k$ and we are using the notation $\langle j^{\perp}, \ell \rangle_r := r(j_2\ell_1 - j_1\ell_2)$. In what follows the coefficient $c_{j,\ell}$ always depends on r but we suppress the dependence for notational simplicity.

We will regard the configuration space $\mathbb{C}^{\mathbb{Z}^2_{0,N}}$ as a complex manifold with complexified tangent space spanned by the complex basis vectors $\{\partial_{w_k} \,:\, k \in \mathbb{Z}^2_{0,N}\}$ (Wirtinger derivatives) satisfying $\partial_{w_{-k}} = \bar{\partial}_{w_k}$. See [52] for the notion of complexified tangent space and [21] for discussion on how to use this complex framework for checking Hörmander's condition. In this basis, we can formulate the SDE (4.6) in the canonical form

$$\mathrm{d}w_t = X^{\epsilon}_0(w_t) + \sum_{k \in \mathcal{Z}^0} \sqrt{\epsilon}\partial_{w_k}\mathrm{d}W^k_t,$$

where the drift vector field X^{ϵ}_0 is given by $X^{\epsilon}_0(w) := \sum_{k \in \mathbb{Z}^2_{0,N}} (B_k(w, w) - \epsilon|k|^2_r w_k)\partial_{w_k}$ and the set of driving modes $\mathcal{Z}^0$ is given by $\mathcal{Z}^0 := \{k \in \mathbb{Z}^2_{0,N} \,:\, \alpha_k, \beta_k \neq 0\}$.

As in the setting of [37, 48], we consider very degenerate forcing and study how it spreads throughout the system via the nonlinearity $B_\ell(w, w)$. Specifically, we define the sets

$$\mathcal{Z}^n = \{\ell \in \mathbb{Z}^2_{0,N} \,:\, \ell = j + k,\, j \in \mathcal{Z}^0,\, k \in \mathcal{Z}^{n-1}, c_{j,k} \neq 0\}, \quad n \geq 0$$

and assume that the driving modes $\mathcal{Z}^0$ satisfy $\bigcup_{n \geq 0} \mathcal{Z}^n = \mathbb{Z}^2_{0,N}$. Under this assumption on $\mathcal{Z}^0$ it can be shown (see [16] Proposition 3.6 or [37, 48]) that the complexified Lie algebra $\mathrm{Lie}(X^{\epsilon}_0; \{\partial_{w_k} \,:\, k \in \mathcal{Z}^0\})$ contains the constant vector fields $\{\partial_{w_k} \,:\, k \in \mathbb{Z}^2_{0,N}\}$ and therefore satisfies the uniform parabolic Hörmander condition on $\mathbb{C}^{\mathbb{Z}^2_{0,N}}$.

4.3.1 A distinctness condition on a diagonal subalgebra

As discussed in Section 4.2, in order to verify projective hypoellipticity for the vector fields $X_0^\epsilon; \{\partial_{w_k} \,:\, k \in \mathcal{Z}^0\}$, it suffices to study the generating properties of a suitable matrix Lie algebra. In [21], we show this can be reformulated to a condition on the constant, *real valued* matrices $H^k = \nabla[\partial_{w_k}, B]$, $k \in \mathbb{Z}^2_{0,N}$, represented in $\{\partial_{w_k}\}$ coordinates by $(H^k)_{\ell,j} = \partial_{w_j}\partial_{w_k} B_\ell(w,w) = c_{j,k}\delta_{\ell=j+k}$. After obtaining this reformulation, the main result of [21] is the following non-degeneracy property of the matrices $\{H^k\}$.

Theorem 4.6 (Theorem 2.13, [21] (see also Proposition 3.11))**.** *Consider the 2d stochastic Galerkin Navier-Stokes equations with frequency truncation N on $\mathbb{T}^2_r$ and suppose that $N \geq 392$. Then, the following holds:*

$$\mathrm{Lie}(\{H^k : k \in \mathbb{Z}^2_{0,N}\}) = \mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R})\,, \tag{4.7}$$

where $\mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R})$ denotes the Lie algebra of real-valued traceless matrices indexed by the truncated lattice $\mathbb{Z}^2_{0,N}$. Therefore projective hypoellipticity holds for (4.2) and by Theorem 4.1 the top Lyapunov exponent satisfies $\lim_{\epsilon\to 0}\epsilon^{-1}\lambda_1^\epsilon = \infty$.

Remark 4.7. Verifying the Lie algebra generating condition (4.7) can be quite challenging due to the fact that there there are $n = |\mathbb{Z}^2_{0,N}|$ matrices and n^2-1 degrees of freedom to span. The matrices are also banded in the sense that for each k, $(H^k)_{\ell,j}$ couples most of the lattice values ℓ, j along the band $k = \ell - j$ and therefore it is extremely challenging to isolate elementary matrices (matrices with only one non-zero entry) as one can do in "local in frequency" models like L96 (4.5) (see [16]). Moreover, brute force computational approaches that successively generate Lie bracket generations and count the rank by Gaussian elimination (such as the Lie-Tree algorithm in [38]) are only available for fixed $r \in \mathbb{R}_+$ and $N \in \mathbb{Z}_+$, and can be subject to numerical error (for instance if r is chosen irrational) which destroy the validity of the proof.

In order to show that (4.7) holds, in [21] we take an approach inspired by the root-space decomposition of semi-simple Lie algebras and study genericity properties of the following diagonal sub-algebra of $\mathrm{Lie}(\{H^k\})$

$$\mathfrak{h} := \mathrm{span}\{\mathbb{D}^k \,:\, k \in \mathbb{Z}^2_{0,N}\},$$

where $\mathbb{D}^k = [H^k, H^{-k}]$ are a family of diagonal matrices with diagonal elements $\mathbb{D}^k_i = (\mathbb{D}^k)_{ii}$ given by

$$\mathbb{D}^k_i = c_{i,k}c_{i+k,k}\mathbb{1}_{\mathbb{Z}^2_{0,N}}(i+k) - c_{i,k}c_{i-k,k}\mathbb{1}_{\mathbb{Z}^2_{0,N}}(i-k).$$

Using that, for a given diagonal matrix $\mathbb{D} \in \mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R})$, the adjoint action $\mathrm{ad}(\mathbb{D}) : \mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R}) \to \mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R})$, where $\mathrm{ad}(\mathbb{D})H = [\mathbb{D}, H]$, has eigenvectors given by the elementary matrices $E^{i,j}$ (i.e. a matrix with a one in the ith row and jth column and zero elsewhere) $\mathrm{ad}(\mathbb{D})E^{i,j} = (\mathbb{D}_i - \mathbb{D}_j)E^{i,j}$, means that $\mathrm{ad}(\mathbb{D})$ has a simple spectrum if the diagonal entries of $\mathbb{D}$ have *distinct differences*, $\mathbb{D}_i - \mathbb{D}_j \neq \mathbb{D}_{i'} - \mathbb{D}_{j'}$, $(i,j) \neq (i',j')$. This implies that if H is a matrix with non-zero non-diagonal entries and $\mathbb{D}$ has distinct differences, then for $M = n^2 - n$, the Krylov subspace

$$\mathrm{span}\{H, \mathrm{ad}(\mathbb{D})H, \mathrm{ad}(\mathbb{D})^2 H, \ldots, \mathrm{ad}(\mathbb{D})^{M-1}H\}$$

contains the set $\{E^{i,j} \,:\, i,j \in \mathbb{Z}^2_{0,N},\, i \neq j\}$, which is easily seen to generate $\mathfrak{sl}_{\mathbb{Z}^2_{0,N}}(\mathbb{R})$.

However in our setting the diagonal matrices $\mathbb{D}^k$ have an inversion symmetry $\mathbb{D}^k_{-i} = -\mathbb{D}^k_i$ and therefore there *cannot* be a matrix in $\mathfrak{h}$ with all differences distinct. Moreover, we do not have a matrix with *all* off diagonal entries non-zero due to the degeneracies present in $c_{j,k}$ and the presence of the Galerkin cut-off. Nevertheless, in [21] we are able to deduce the following sufficient condition on the family $\{\mathbb{D}^k\}$ that ensures (4.7) holds:

Proposition 4.8 (Corollary 4.9 and Lemma 5.2 [21]). *Let $N \geq 8$. If for each $(i, j, \ell, m) \in (\mathbb{Z}^2_{0,N})^4$ satisfying $i + j + \ell + m = 0$ and $(i + j, \ell + m) \neq 0$, $(i + \ell, j + m) \neq 0$, $(i + m, j + \ell) \neq 0$, there exists a $k \in \mathbb{Z}^2_{0,N}$ such that*

$$\mathbb{D}^k_i + \mathbb{D}^k_j + \mathbb{D}^k_\ell + \mathbb{D}^k_m \neq 0, \tag{4.8}$$

then (4.7) *holds.*

The proof of Proposition 4.8 is not straightforward. However, its proof uses some similar ideas as the proof of (4.8) but is otherwise significantly easier, so we only discuss the latter.

4.3.2 Verifying the distinctness condition using computational algebraic geometry

The distinctness condition (4.8) is not a simple one to verify. Indeed, ignoring the Galerkin cut-off N for now, $\mathbb{D}^k_i$ are rational algebraic expressions in the variables (i, k, r) (being comprised of products and sums of the coefficients $c_{j,k}$), and therefore proving (4.8) amounts to showing that the family of *Diophantine equations*[12]

$$\mathbb{D}^k_i + \mathbb{D}^k_j + \mathbb{D}^k_\ell + \mathbb{D}^k_m = 0, \quad \text{for each} \quad k \in \mathbb{Z}^2_{0,N} \tag{4.9}$$

have *no solutions* (i, j, ℓ, m, r) satisfying the constraints of Proposition 4.8. Due to the complexity of the expression for $\mathbb{D}^k_i$, there is little hope to verify such a result by hand (the resulting polynomials are degree 16 in 9 variables). However, if one extends each of the 9 variables $(i, j, \ell, m, r) = (i_1, i_2, j_1, j_2, \ell_1, \ell_2, m_1, m_2, r)$ to the algebraically closed field $\mathbb{C}$, then (4.9) along with $i + j + \ell + m = 0$ defines a polynomial ideal I with an associated algebraic variety $\mathbb{V}(I)$ in $\mathbb{C}^9$. Such a high dimensional variety is rather complicated due to the inherent symmetries of the rational equation in (4.9), however its analysis is nonetheless amenable to techniques from algebraic geometry, particularly the strong Nullstellensatz and computer algorithms for computing Gröbner bases (see [31] for a review of the algebraic geometry concepts). Indeed, without the Galerkin cut-off (the formal infinite dimensional limit), in [21] we proved, by computing Gröbner bases in rational arithmetic using the F4 algorithm [39] implemented in the computer algebra system Maple [1], that the identity $\mathbb{V}(I) = \mathbb{V}(g)$ holds, where g is the following "saturating" polynomial

$$g(i, j, \ell, m, r) = r^2 |i|^2_r |j|^2_r |\ell|^2_r |m|^2_r (|i + j|^2_r + |\ell + m|^2_r)(|i + \ell|^2_r + |j + m|^2_r)(|i + m|^2_r + |j + \ell|^2_r)$$

whose non-vanishing encodes the constraints in Proposition 4.8, thereby showing that (4.8) holds.

Dealing with the Galerkin truncation adds significant difficulties to the proof as the associated rational system (4.9) is instead piecewise defined (depending on k and N) and therefore doesn't easily reduce to a problem about polynomial inconsistency. Nonetheless, by considering 34 different polynomial ideals associated to different possible algebraic forms, in [21] we were able to show that if N is taken large enough (bigger than 392 to be precise) then (4.8) still holds with the Galerkin truncation present and therefore Theorem 4.6 holds.

Finally, it is worth remarking that even without the Galerkin cut-off, the system of rational equations (4.9) is complex enough to become computationally intractable (even for modern computer algebra algorithms) without some carefully chosen simplifications, variable orderings, choice of saturating polynomial g and sheer luck; see [21] for more details.

5 Lagrangian chaos in stochastic Navier-Stokes

At present, the results above based on Proposition 2.7 are restricted to finite dimensional problems. Indeed, even while the Fisher information can potentially be extended to infinite dimensions under certain conditions[13], for any parabolic SPDE problem, we will always have $\lambda_\Sigma = -\infty$. The existence of positive

[12] At least considering $r = 1$ or another fixed, rational number.

[13] If $X^*\nu \ll \nu$ and we define $\beta^\nu_{X^*} := \frac{\mathrm{d}\tilde{X}^*\nu}{\mathrm{d}\nu}$, then $FI(f) = \frac{1}{2}\sum_k \|\beta^\nu_{\tilde{X}^*}\|^2_{L^2(\nu)}$, and there is no explicit dependence on any reference measure or Riemannian metric; see [16] for more details.

Lyapunov exponents for the infinite-dimensional, stochastic Navier-Stokes equations remains open as of the writing of this note.

However, there is another important problem in fluid mechanics where we have been able to make progress. Consider the (infinite-dimensional) 2d Navier-Stokes equations[14] in $\mathbb{T}^2$,

$$\partial_t u_t + (u_t \cdot \nabla u_t + \nabla p - \nu \Delta u_t) = \sum_k q_k e_k \dot{W}_t^k, \quad \operatorname{div} u_t = 0, \tag{5.1}$$

where the $q_k \in \mathbb{R}$ and e_k are eigenfunctions of the Stokes operator. The *Lagrangian flow map* $\varphi^t_{\omega,u} : \mathbb{T}^2 \mapsto \mathbb{T}^2$ is defined by the trajectories of particles moving with the fluid

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi^t_{\omega,u}(x) = u_t(\varphi^t_{\omega,u}(x)), \quad \varphi^0_{\omega,u}(x) = x,$$

where note that the diffeomorphism $\varphi^t_{\omega,u}$ depends on the initial velocity u and the noise path ω and is therefore a co-cycle over the skew product $\Theta_t : \Omega \times H^s \circlearrowleft$, where $\Theta_t(\omega, u) = (\theta_t\omega, \Psi^t_\omega(u))$ and $\Psi^t_\omega : H^s \circlearrowleft$ is the 2d Navier Stokes flow on H^s associated with (5.1). One can naturally ask whether or not (u_t) is chaotic, as we have done in previous sections, or if the motion of particles immersed in the fluid is chaotic, e.g., if the Lagrangian Lyapunov exponent is strictly positive. The latter is known as *Lagrangian chaos* [4, 6, 10, 27, 32, 42, 86] (to distinguish it from chaos of (u_t) itself, which is sometimes called *Eulerian chaos*). While both are expected to be observed in turbulent flows, Lagrangian chaos is not incompatible with Eulerian "order", i.e., a negative exponent for the (u_t) process.

In [17] we proved, under the condition that $|q_k| \approx |k|^{-\alpha}$ for some $\alpha > 10$, that $\exists \lambda_1 > 0$ deterministic and independent of initial x and initial velocity u such that the following limit holds almost-surely:

$$\lim_{t\to\infty} \frac{1}{t} \log |D_x \varphi^t_{\omega,u}| = \lambda_1 > 0. \tag{5.2}$$

This Lagrangian chaos was later upgraded in [18, 19] to the much stronger property of uniform-in-diffusivity, almost-sure exponential mixing. To formulate this notion, we consider (g_t) a passive scalar solving the (random) advection-diffusion equation

$$\partial_t g_t + u_t \cdot \nabla g_t = \kappa \Delta g_t, \quad g_0 = g,$$

for $\kappa \in [0, 1]$ and a fixed, mean-zero scalar $g \in L^2(\mathbb{T}^2)$. In [18, 19], we proved that there exists a (deterministic) constant $\mu > 0$ such that for all $\kappa \in [0, 1]$ and initial divergence free $u \in H^s$ (for some sufficiently large s), there exists a random constant $D = D(\omega, \kappa, u)$ such that for all $g \in H^1$ (mean-zero)

$$\|g_t\|_{H^{-1}} \le D e^{-\mu t} \|g\|_{H^1}$$

where D is almost-surely finite and satisfies the *uniform-in-κ* moment bound (for some fixed constant q and for any $\eta > 0$),

$$\mathbf{E} D^2 \lesssim_\eta (1 + \|u\|_{H^s})^q e^{\eta \|u\|^2_{H^1}}.$$

One can show that this result is essentially optimal up to getting sharper quantitative estimates on μ and D, at least if $\kappa = 0$ [18, 19]. This uniform, exponential mixing plays the key role in obtaining a proof of Batchelor's power spectrum [12] of passive scalar turbulence in some regimes [15].

Let us simply comment on the Lagrangian chaos statement (5.2), as it is most closely related to the rest of this note. The main step is to deduce an analogue of Theorem 2.5 for the Lagrangian flow map, using that while the Lagrangian flow map depends on an infinite dimensional Markov process, the Jacobian $D_x \varphi^t_{\omega,u}$

[14] The 3D Navier-Stokes equations can be treated provided the $-\nu\Delta u_t$ is replaced with the hyperviscous damping $\nu\Delta^2 u_t$.

itself is finite dimensional. This is done in our work [17] by extending Furstenberg's criterion to handle general linear cocycles over infinite dimensional processes in the same way that $D_x\varphi^t_{\omega,u}$ depends on the sample paths (u_t).

The Lagrangian flow is divergence-free, and thus the Lagrangian Lyapunov exponents satisfy $\lambda_\Sigma = 0$ and $\lambda_1 \geq 0$, so ruling out the degenerate situations in Theorem 2.5 would immediately imply $\lambda_1 > 0$. A key difficulty in this infinite-dimensional context is to ensure that the rigid invariant structures (now functions of the fluid velocity field u and the Lagrangian tracer position x) in our analogue of Theorem 2.5 vary *continuously* as functions of u and x. It is at this step that we require the non-degeneracy type condition on the noise $|q_k| \gtrsim |k|^{-\alpha}$, which is used to ensure that the Markov process $(u_t, \varphi^t(x))$ is strong Feller.

At the time of writing, it remains an interesting open problem to extend our works [17–19] to degenerate noise such as that used in [48] or [59]. It bears remarking that the methods of [48] apply to the one-point process $(u_t, \varphi^t(x))$ (this is used in our work [18]), however, it is nevertheless unclear how to prove Lagrangian chaos without a sufficiently strong analogue of Theorem 2.5, and it is unclear how to obtain such a theorem without the use of the strong Feller property.

6 Looking forward

The work we reviewed here raises a number of potential research directions.

Tighter hypoelliptic regularity estimates. The scaling $\lambda_1^\epsilon \gg \epsilon$ that naturally follows from our above analysis is surely suboptimal – even if the deterministic problem were to be completely integrable, the scaling is likely to be $O(\epsilon^\gamma)$ for some $\gamma < 1$ depending on dimension (see e.g. [14, 76]). To begin with, one may attempt to strengthen the hypoelliptic regularity estimate by refining the ϵ scaling to something like

$$\|f^\epsilon\|^2_{W^{s,1}} \lesssim 1 + \frac{n\lambda_1^\epsilon - 2\lambda_\Sigma^\epsilon}{\epsilon^\gamma},$$

for some constant $0 < \gamma < 1$. If such an estimate were true, the same compactness-rigidity argument of Theorem 4.1 would imply a scaling like $\lambda_1^\epsilon \gtrsim \epsilon^\gamma$. An improvement of this type seems plausible given the proof of Theorem 3.3. It might be necessary, in general, to use a more specialized norm on the left-hand side, but local weak L^1 compactness, i.e. equi-integrability, is all that is really required for the compactness-rigidity argument to apply.

Beyond compactness-rigidity. Compactness-rigidity arguments may remain limited in their ability to obtain optimal or nearly optimal scalings for λ_1, regardless of the ways one can improve Theorem 3.3. Another approach is to find some way to work more directly on $\epsilon > 0$. This was essentially the approach of works [14, 76], however, the method of these papers only applies if one has a nearly-complete understanding of the pathwise random dynamics. We are unlikely to ever obtain such a complete understanding of the dynamics of models such as L96 or Galerkin-Navier-Stokes, but there may be hope that partial information, such as the isolation of robust, finite-time exponential growth mechanisms, could be used to obtain better lower bounds on $\|f^\epsilon\|_{W^{s_*,1}}$. An approach with a vaguely related flavor for random perturbations of discrete-time systems, including the Chirikov standard map, was carried out in the previous work [24].

Finer dynamical information: moment Lyapunov exponents. Lyapunov exponents provide asymptotic exponential growth rates of the Jacobian, but they provide no quantitative information on how long it takes for this growth to be realized with high probability. One tool to analyze this is the study of large deviations of the convergence of the sequences $\frac{1}{t}\log|D_x\Phi^t_\omega v|$. The associated rate function is the Legendre transform of the *moment Lyapunov exponent function* $p \mapsto \Lambda(p) := \lim_{t\to\infty}\frac{1}{t}\log \mathbf{E}|D_x\Phi^t_\omega v|^p$ (the limit defining $\Lambda(p)$ exists and is independent of (x, v, ω) under fairly general conditions [7]). It would be highly interesting to see if the quantitative estimates obtained by e.g. Theorem 4.1 extend also to quantitative estimates on the moment Lyapunov exponents. We remark that the moment Lagrangian Lyapunov exponents play a key role in our works [18, 19].

Lyapunov times of small-noise perturbations of completely integrable systems The phase space of a completely integrable Hamiltonian flow is foliated by invariant torii along which the dynamics is a translation flow– such systems are highly ordered and non-chaotic. On the other hand, small perturbations of the Hamiltonian are known to break the most "resonant" of these torii, while torii with sufficiently "non-resonant" frequencies persist due to KAM theory. It is an interesting and highly challenging open problem to prove that this 'breakage' results in the formation of a positive-volume set admitting a positive Lyapunov exponent. For the most part such problems are wide open, and related to the standard map conjecture discussed in Section 1.1. The recent work of Berger and Turaev [22] established a renormalization technique for proving the *existence* of smooth perturbations resulting in a positive Lyapunov exponent, but it remains open to affirm how 'generic' such perturbations actually are.

The following is a closely related *stochastic* dynamics problem: starting from a completely integrable system and adding a small amount of noise, how many Lyapunov times elapse for the random dynamics before the "stochastic divergence" timescale when the deterministic flow and the stochastic flow differ by $O(1)$? Estimating the stochastic divergence timescale is essentially a large deviations problem, and has already been carried out for small random perturbations of completely integrable systems; see, e.g., [40]. On the other hand, estimating Lyapunov times beyond the crude $(\lambda_1^\epsilon)^{-1}$ estimate is a large deviations estimate for the convergence of finite-time Lyapunov exponents to their asymptotic value λ_1^ϵ. The associated rate function in this case is the Legendre transform of the moment Lyapunov exponent $\Lambda(p)$ mentioned earlier; a positive result for the program described above would require quantitative-in-ϵ estimates on $\Lambda(p)$.

More general noise models. One simple potential extension is Theorem 4.1 to different types of multiplicative noise. Another important extension would be to noise models which are not white-in-time, for example noise of the type used in [59], which is challenging because our work is deeply tied to the elliptic nature of the generator $\tilde{\mathcal{L}}^*$. A simpler example of non-white forcing can be constructed from 'towers' of coupled Ornstein-Uhlenbeck processes, which can be built to be C^k in time for any $k \geq 0$ (see, e.g., [17, 18] for details).

Lagrangian chaos. There are several directions of research to extend our results in [17–19], such as studying degenerate noise as in [48, 59], extending to more realistic physical settings such as bounded domains with stochastic boundary driving, and extending Proposition 2.7 to the Lagrangian flow map in a variety of settings, which would help to facilitate quantitative estimates (note one will have to use the conditional density version so that one does not see the effect of the λ_Σ associated to the Navier-Stokes equations themselves).

References

[1] *Maple (2020). maplesoft, a division of waterloo maple inc., waterloo, ontario.*

[2] F. Abedin and G. Tralli, *Harnack inequality for a class of Kolmogorov–Fokker–Planck equations in non-divergence form*, Archive for Rational Mechanics and Analysis **233** (2019), no. 2, 867–900.

[3] D. Albritton, R. Beekie, and M. Novack, *Enhanced dissipation and Hörmander's hypoellipticity*, arXiv preprint arXiv:2105.12308 (2021).

[4] C. H Amon, A. M Guzmán, and B. Morel, *Lagrangian chaos, Eulerian chaos, and mixing enhancement in converging–diverging channel flows*, Physics of Fluids **8** (1996), no. 5, 1192–1206.

[5] F. Anceschi, S. Polidoro, and M. A. Ragusa, *Moser's estimates for degenerate Kolmogorov equations with non-negative divergence lower order coefficients*, Nonlinear Analysis **189** (2019), 111568.

[6] T. M Antonsen Jr, Z. Fan, E. Ott, and E Garcia-Lopez, *The role of chaotic orbits in the determination of power spectra of passive scalars*, Physics of Fluids **8** (1996), no. 11, 3094–3104.

[7] L. Arnold, *A formula connecting sample and moment stability of linear stochastic systems*, SIAM Journal on Applied Mathematics **44** (1984), no. 4, 793–802.

[8] L. Arnold, *Random dynamical systems*, Dynamical systems, 1995, pp. 1–43.

[9] L. Arnold, D. C. Nguyen, and V. Oseledets, *Jordan normal form for linear cocycles*, Random Operators and Stochastic Equations **7** (1999), no. 4, 303–358.

[10] E Balkovsky and A Fouxon, *Universal long-time properties of Lagrangian statistics in the Batchelor regime and their application to the passive scalar problem*, Physical Review E **60** (1999), no. 4, 4164.

[11] L. Barreira and Y. B Pesin, *Lyapunov exponents and smooth ergodic theory*, Vol. 23, American Mathematical Soc., 2002.

[12] G. K Batchelor, *Small-scale variation of convected quantities like temperature in turbulent fluid part 1. general discussion and the case of small conductivity*, Journal of Fluid Mechanics **5** (1959), no. 1, 113–133.

[13] P. H Baxendale, *Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms*, Probability Theory and Related Fields **81** (1989), no. 4, 521–554.

[14] P. H Baxendale and L. Goukasian, *Lyapunov exponents for small random perturbations of Hamiltonian systems*, Annals of probability (2002), 101–134.

[15] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *The Batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics at fixed reynolds number*, To appear in Comm. Pure Appl. Math.. arXiv:1911.11014 (2019).

[16] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *A regularity method for lower bounds on the Lyapunov exponent for stochastic differential equations*, To appear in Invent. Math.. arXiv:2007.15827 (2020).

[17] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *Lagrangian chaos and scalar advection in stochastic fluid mechanics*, To appear in J. Euro. Math. Soc.. arXiv:1809.06484 (2018).

[18] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *Almost-sure exponential mixing of passive scalars by the stochastic Navier-Stokes equations*, To appear in Ann. of Prob.. arXiv:1905.03869 (2019).

[19] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith, *Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic Navier-Stokes*, Prob. Theory Rel. Fields **179** (2021), no. 3, 777–834.

[20] J. Bedrossian and K. Liss, *Quantitative spectral gaps and uniform lower bounds in the small noise limit for Markov semigroups generated by hypoelliptic stochastic differential equations*, To appear in Prob. Math. Phys. (2020).

[21] J. Bedrossian and S. Punshon-Smith, *Chaos in stochastic 2d Galerkin-Navier-Stokes*, Preprint, arXiv:2106.13748 (2021).

[22] P. Berger and D. Turaev, *On Herman's positive entropy conjecture*, Advances in Mathematics **349** (2019), 1234–1288.

[23] A. Blumenthal, J. Xue, and Y. Yang, *Lyapunov exponents for random perturbations of coupled standard maps*, arXiv preprint arXiv:2004.10626 (2020).

[24] A. Blumenthal, J. Xue, and L.-S. Young, *Lyapunov exponents for random perturbations of some area-preserving maps including the standard map*, Annals of Mathematics (2017), 285–310.

[25] A. Blumenthal, J. Xue, and L.-S. Young, *Lyapunov exponents and correlation decay for random perturbations of some prototypical 2d maps*, Communications in Mathematical Physics **359** (2018), no. 1, 347–373.

[26] A. Blumenthal and Y. Yang, *Positive Lyapunov exponent for random perturbations of predominantly expanding multimodal circle maps*, arXiv preprint arXiv:1805.09219 (2018).

[27] T. Bohr, M. H Jensen, G. Paladin, and A. Vulpiani, *Dynamical systems approach to turbulence*, Cambridge University Press, 2005.

[28] W. M Boothby and E. N Wilson, *Determination of the transitivity of bilinear systems*, SIAM journal on control and optimization **17** (1979), no. 2, 212–221.

[29] A. Carverhill, *Furstenberg's theorem for nonlinear stochastic systems*, Probability theory and related fields **74** (1987), no. 4, 529–534.

[30] B. V Chirikov et al., *A universal instability of many-dimensional oscillator systems*, Physics reports **52** (1979), no. 5, 263–379.

[31] D. Cox, J. Little, and D. OShea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, Springer Science & Business Media, 2013.

[32] A. Crisanti, M. Falcioni, A. Vulpiani, and G. Paladin, *Lagrangian chaos: transport, mixing and diffusion in fluids*, La Rivista del Nuovo Cimento (1978-1999) **14** (1991), no. 12, 1–80.

[33] S. Crovisier and S. Senti, *Un problème pour le xxi(i)ème siècle*, La Gazette des mathé§'aticiens **156** (2018).

[34] G. Da Prato, J. Zabczyk, and J Zabczyk, *Ergodicity for infinite dimensional systems*, Vol. 229, Cambridge University Press, 1996.

[35] P. D Ditlevsen, *Turbulence and shell models*, Cambridge University Press, 2010.

[36] P. Duarte, *Plenty of elliptic islands for the standard family of area preserving maps*, Annales de l'institut henri poincare (c) non linear analysis, 1994, pp. 359–409.

[37] W. E and J. C Mattingly, *Ergodicity for the Navier-Stokes equation with degenerate random forcing: Finite-dimensional approximation*, Commun. Pure Appl. Math. **54** (Nov. 2001), no. 11, 1386–1402.

[38] D. Elliott, *Bilinear control systems: Matrices in action*, Springer, Dordrecht, 2009.

[39] J.-C. Faugere, *A new efficient algorithm for computing gröbner bases (f4)*, Journal of pure and applied algebra **139** (1999), no. 1-3, 61–88.

[40] M. I Freidlin and A. D Wentzell, *Random perturbations of Hamiltonian systems*, Vol. 523, American Mathematical Soc., 1994.

[41] H. Furstenberg, *Noncommuting random products*, Transactions of the American Mathematical Society **108** (1963), no. 3, 377–428.

[42] S. Galluccio and A Vulpiani, *Stretching of material lines and surfaces in systems with lagrangian chaos*, Physica A: Statistical Mechanics and its Applications **212** (1994), no. 1-2, 75–98.

[43] N. E Glatt-Holtz, D. P Herzog, and J. C Mattingly, *Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations*, Ann. PDE **4** (Dec. 2018), no. 2 (en).

[44] E. Gledzer, *Hydrodynamic-type system admitting two quadratic integrals of motion*, Dokl. akad. nauk sssr, 1973, pp. 1046–1048.

[45] F Golse, C Imbert, C Mouhot, and A Vasseur, *Harnack inequality for kinetic fokker-planck equations with rough coefficients and application to the landau equation*, to appear in Annali della Scuola Normale Superiore di Pisa (2016).

[46] A. Gorodetski, *On the stochastic sea of the standard map*, Communications in Mathematical Physics **309** (2012), no. 1, 155–192.

[47] M. Hairer, *On Malliavin's proof of Hörmander's theorem*, Bulletin des sciences mathematiques **135** (2011), no. 6-7, 650–666.

[48] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, Ann. of Math. **164** (2006), no. 3, 993–1032.

[49] D. P Herzog and J. C Mattingly, *A practical criterion for positivity of transition densities*, Nonlinearity **28** (July 2015), no. 8, 2823 (en).

[50] G. Hochschild, *The structure of lie groups*, Holden-day, 1965.

[51] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Mathematica **119** (1967), no. 1, 147–171.

[52] D. Huybrechts, *Complex geometry: an introduction*, Springer Science & Business Media, 2005.

[53] V. Jurdjevic, *Geometric control theory*, Cambridge university press, 1997.

[54] A. Karimi and M. R Paul, *Extensive chaos in the Lorenz-96 model*, Chaos: An interdisciplinary journal of nonlinear science **20** (2010), no. 4, 043105.

[55] R. Khasminskii, *Stochastic stability of differential equations*, Vol. 66, Springer Science & Business Media, 2011.

[56] Y. Kifer, *Ergodic theory of random transformations*, Vol. 10, Springer Science & Business Media, 2012.

[57] A. E. Kogoj and S. Polidoro, *Harnack inequality for hypoelliptic second order partial differential operators*, Potential Anal. **45** (20164), no. 14, 545–555.

[58] A. Kolmogorov, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2) **35** (1934), no. 1, 116–117. MR1503147

[59] S. Kuksin, V. Nersesyan, and A. Shirikyan, *Exponential mixing for a class of dissipative pdes with bounded degenerate noise*, Geometric and Functional Analysis (2020), 1–62.

[60] S. Kuksin and A. Shirikyan, *Mathematics of two-dimensional turbulence*, Vol. 194, Cambridge University Press, 2012.

[61] H. Kunita, *Stochastic flows and stochastic differential equations*, Vol. 24, Cambridge university press, 1997.

[62] A. Lanconelli, A. Pascucci, and S. Polidoro, *Gaussian lower bounds for non-homogeneous kolmogorov equations with measurable coefficients*, Journal of Evolution Equations (2020), 1–19.

[63] F. Ledrappier, *Positivity of the exponent for stationary sequences of matrices*, Lyapunov exponents, 1986, pp. 56–73.

[64] Z. Lian and M. Stenlund, *Positive Lyapunov exponent by a random perturbation*, Dynamical Systems **27** (2012), no. 2, 239–252.

[65] G. M Lieberman, *Second order parabolic differential equations*, World scientific, 1996.

[66] P.-D. Liu and M. Qian, *Smooth ergodic theory of random dynamical systems*, Springer, 2006.

[67] E. N Lorenz, *Predictability: A problem partly solved*, Proc. seminar on predictability, 1996.

[68] V. S L'vov, E. Podivilov, A. Pomyalov, I. Procaccia, and D. Vandembroucq, *An optimal shell model of turbulence*, Physical Review E **58** (1998), no. 2, 1811.

[69] R. MacKay, *An appraisal of the ruelle-takens route to turbulence*, The global geometry of turbulence, 1991, pp. 233–246.

[70] A. J Majda, *Introduction to turbulent dynamical systems in complex systems*, Springer, 2016.

[71] S. P Meyn and R. L Tweedie, *Markov chains and stochastic stability*, Springer Science & Business Media, 2012.

[72] C. Mouhot, *De Giorgi–Nash–Moser and Hörmander theories: new interplays*, Proceedings of the international congress of mathematicians rio de janeiro, 2018, pp. 2467–2493.

[73] V. I. Oseledets, *A multiplicative ergodic theorem. characteristic Ljapunov exponents of dynamical systems*, Trudy Moskovskogo Matematicheskogo Obshchestva **19** (1968), 179–210.

[74] E. Ott, B. R Hunt, I. Szunyogh, A. V Zimin, E. J Kostelich, M. Corazza, E. Kalnay, D. Patil, and J. A Yorke, *A local ensemble Kalman filter for atmospheric data assimilation*, Tellus A: Dynamic Meteorology and Oceanography **56** (2004), no. 5, 415–428.

[75] Y. Pesin and V. Climenhaga, *Open problems in the theory of non-uniform hyperbolicity*, Discrete Contin. Dyn. Syst **27** (2010), no. 2, 589–607.

[76] M. A Pinsky and V. Wihstutz, *Lyapunov exponents of nilpotent Itô systems*, Stochastics: An International Journal of Probability and Stochastic Processes **25** (1988), no. 1, 43–57.

[77] M. S Raghunathan, *A proof of Oseledec's multiplicative ergodic theorem*, Israel Journal of Mathematics **32** (1979), no. 4, 356–362.

[78] M. Romito and L. Xu, *Ergodicity of the 3D stochastic Navier–Stokes equations driven by mildly degenerate noise*, Stochastic Process. Appl. **121** (Apr. 2011), no. 4, 673–700.

[79] G. Royer, *Croissance exponentielle de produits Markoviens de matrices aléatoires*, Annales de l'ihp probabilités et statistiques, 1980, pp. 49–62.

[80] M. Shamis and T. Spencer, *Bounds on the lyapunov exponent via crude estimates on the density of states*, Communications in Mathematical Physics **338** (2015), no. 2, 705–720.

[81] H. Triebel, *Theory of function spaces II*, Birkhauser, 1992.

[82] A. Virtser, *On products of random matrices and operators*, Theory of Probability & Its Applications **24** (1980), no. 2, 367–377.

[83] A. Wilkinson, *What are Lyapunov exponents, and why are they interesting?*, Bulletin of the American Mathematical Society **54** (2017), no. 1, 79–105.

[84] M. Yamada and K. Ohkitani, *Lyapunov spectrum of a chaotic model of three-dimensional turbulence*, Journal of the Physical Society of Japan **56** (1987), no. 12, 4210–4213.

[85] L.-S. Young, *Mathematical theory of Lyapunov exponents*, Journal of Physics A: Mathematical and Theoretical **46** (2013), no. 25, 254001.

[86] G.-C. Yuan, K. Nam, T. M Antonsen Jr, E. Ott, and P. N Guzdar, *Power spectrum of passive scalars in two dimensional chaotic flows*, Chaos: An Interdisciplinary Journal of Nonlinear Science **10** (2000), no. 1, 39–49.