

ALMOST-SURE EXPONENTIAL MIXING OF PASSIVE SCALARS BY THE STOCHASTIC NAVIER–STOKES EQUATIONS

BY JACOB BEDROSSIAN^{1,*}, ALEX BLUMENTHAL^{1,†} AND SAMUEL PUNSHON-SMITH²

¹*Department of Mathematics, University of Maryland, *jacob@math.umd.edu; †ablumenthal6@gatech.edu*

²*Division of Applied Mathematics, Brown University,*

We deduce almost-sure exponentially fast mixing of passive scalars advected by solutions of the stochastically-forced 2D Navier–Stokes equations and 3D hyper-viscous Navier–Stokes equations in $\mathbb{T}^d$ subjected to nondenerate H^σ -regular noise for any σ sufficiently large. That is, for all $s > 0$ there is a deterministic exponential decay rate such that all mean-zero H^s passive scalars decay in H^{-s} at this same rate with probability one. This is equivalent to what is known as *quenched correlation decay* for the Lagrangian flow in the dynamical systems literature. This is a follow-up to our previous work, which establishes a positive Lyapunov exponent for the Lagrangian flow—in general, almost-sure exponential mixing is much stronger than this. Our methods also apply to velocity fields evolving according to finite-dimensional models, for example, Galerkin truncations of Navier–Stokes or the Stokes equations with very degenerate forcing. For all $0 \leq k < \infty$, this exhibits many examples of $C_t^k C_x^\infty$ random velocity fields that are almost-sure exponentially fast mixers.

CONTENTS

1. Introduction	242
1.1. Infinite-dimensional fluid models	243
1.2. Main results and discussion	245
1.3. Finite dimensional evolution and $C_t^k C_x^\infty$ almost-sure exponential mixers	248
2. Outline of the proof	249
2.1. Correlation decay via geometric ergodicity of the two point Lagrangian motion	250
2.2. Conditions for geometric ergodicity	251
2.3. Strong Feller and irreducibility	252
2.4. Construction of the Lyapunov function $\mathcal{V}$	252
2.4.1. Lyapunov functions for the (u_t) process	253
2.4.2. Repulsion from the diagonal $\mathcal{D}$ for the two-point process	253
2.4.3. Verifying the drift condition	255
2.4.4. Outline of the proof of Proposition 2.11	256
2.5. Notation	258
3. Preliminaries, drift conditions and Jacobian estimates	259
3.1. Preliminaries	259
3.2. Fundamental estimates and super-Lyapunov property	261
3.3. Estimates on the Jacobian of the projective process	264
4. Lasota–Yorke bound and spectral gap for the projective process	267
4.1. Geometric ergodicity in C_V	267
4.2. Spectral gap in C_V^1 from a Lasota–Yorke estimate	268
4.3. Lasota–Yorke estimate: Proof of Proposition 4.6	269
4.3.1. Malliavin calculus and preliminaries	269
4.3.2. Defining the control	270

Received May 2020; revised February 2021.

MSC2020 subject classifications. Primary 37A25, 37A30, 37N10, 76F25, 76D06; secondary 37H15, 37A60, 60H15.

Key words and phrases. Exponential mixing, quenched correlation decay, stochastic Navier–Stokes equations, passive scalars, spectral theory of Markov semigroups.

4.3.3. Malliavin derivatives of the Jacobian	272
4.3.4. Nondegeneracy of the Malliavin matrix	273
5. Spectral theory for the twisted Markov semigroup	276
5.1. Proof of Proposition 2.16	276
5.2. Proof of Proposition 2.17	279
5.3. Properties of ψ_p : Proof of Proposition 2.11	281
6. Geometric ergodicity for the two point motion	283
6.1. Strong Feller property: Proof of Proposition 2.6	284
6.1.1. Uniform parabolic Hörmander conditions	284
6.1.2. Sketch of proof of Proposition 2.6	285
6.2. Irreducibility: Proof of Proposition 2.7	287
6.3. Verification of the drift condition: Proof of Proposition 2.14	289
7. Correlation decay: Proof of Theorem 1.4	296
7.1. Correlation bounds in discrete time	297
7.2. Correlation bounds in continuous time	299
7.3. Estimating moments of $D(\omega, u)$	299
Funding	300
References	300

1. Introduction. Passive scalar mixing by fluid motion is an important aspect of many physical phenomena (see, e.g., [63, 69, 71, 72]) and understanding mixing has been a topic of active research in mathematics recently; see, for example, [2, 18, 22, 23, 34, 35, 46, 52, 54, 58, 68] and the references therein (see below for more discussion). In this paper, we will study the mixing of a passive scalar g_t advected by an incompressible velocity field u_t in the absence of diffusivity,

$$(1.1) \quad \partial_t g_t + u_t \cdot \nabla g_t = 0.$$

The previous work in mathematics does not yet consider the case that arises most commonly in physics: velocity fields evolving under the nonlinear dynamics of an ergodic system. The velocity fields we consider are governed by a variety of stochastic fluid models, for example, the stochastically-forced 2D Navier–Stokes equations on $\mathbb{T}^2$:

$$\begin{cases} \partial_t u_t + u_t \cdot \nabla u_t + \nabla p_t = \nu \Delta u_t + Q \dot{W}_t, \\ \operatorname{div} u_t = 0, \\ u_0 =: u, \end{cases}$$

where W_t is a cylindrical Wiener process and Q is a smoothing operator that spatially colors the noise.

We will show that when Q forces all modes directly with a symbol $\hat{Q}(\xi)$ that approximates a power law with both upper and lower bounds $\hat{Q}(\xi) \approx |\xi|^{-\alpha}$ for some $\alpha > \frac{5d}{2}$ (see Assumption 1 below), for instance, we could take $Q = (-\Delta)^{-\alpha/2}$, then u_t is *almost-surely exponentially mixing*. Roughly speaking (see Theorem 1.4 for the rigorous statement and discussion), this means that for all $s > 0$, there is a *deterministic* constant $\hat{\gamma} > 0$ and an *almost-surely finite* random constant $D = D(u, \omega)$ (depending on the initial velocity u and the Brownian path ω), such that the following holds for *all* initial $g_0 = g \in H^s$ with $\int_{\mathbb{T}^2} g \, dx = 0$:

$$(1.2) \quad \|g_t\|_{H^{-s}} = \sup_{\|f\|_{H^s}=1} \int_{\mathbb{T}^2} f(x) g_t(x) \, dx \leq D e^{-\hat{\gamma}t} \|g\|_{H^s},$$

where the supremum is taken over mean-zero functions $f \in H^s$. Alternatively, (1.2) can be formulated in terms of the *Lagrangian flow map* $\phi^t : \mathbb{T}^2 \mapsto \mathbb{T}^2$, solving

$$(1.3) \quad \frac{d}{dt} \phi^t(x) = u_t(\phi^t(x)), \quad \phi^0(x) = x.$$

Indeed, by incompressibility and $g_t = g \circ (\phi^t)^{-1}$, (1.2) is equivalent to

$$(1.4) \quad \left| \int_{\mathbb{T}^2} f(\phi^t(x)) g(x) dx \right| \leq D e^{-\hat{\gamma}t} \|f\|_{H^s} \|g\|_{H^s}.$$

We see that (1.4) quantifies an almost sure, exponentially fast decay of correlations. Such estimates are known in the dynamics literature as *quenched correlation decay* [1, 8, 11, 12, 29].

Previously, for random incompressible flows, almost-sure exponential correlation decay as in (1.4) was established in [14, 27] in the case when ϕ^t is the stochastic flow of diffeomorphisms generated by an SDE (in our terminology, when the velocity field is white-in-time). Our study is built mostly on analyzing (1.3) using ideas from random dynamical systems and the geometric ergodicity of Markov processes, including both ideas from [14, 27], and also many new ones, required to handle the fact that the forcing acts only on the velocity field, evolving according to its own nonlinear SPDE.

A key ingredient is the result from our previous work [15] that the flow map ϕ^t has a positive Lyapunov exponent, that is, $D_x \phi^t(x)$ grows exponentially fast for all x with probability 1 (see (2.3) below for rigorous statement). However, quenched correlation decay is a much stronger statement a priori than simply having a positive Lyapunov exponent (see, e.g., [67] and the discussion in Section 1.2).

Let us now set up the rigorous statement of the results. We first state the results for the infinite-dimensional evolution equations, namely 2D Navier–Stokes and 3D hyperviscous Navier–Stokes. Then we specialize the results to finite dimensional evolutions, for example, Galerkin truncations of Navier–Stokes or the Stokes equations subject to very degenerate noise. In the latter case, a more general class of problems can be treated, producing families of random velocity fields that are $C_t^k C_x^\infty$ and almost-surely exponential mixers for any fixed $k < \infty$.

1.1. Infinite-dimensional fluid models. We will consider two (infinite-dimensional) stochastic fluid models on the periodic box $\mathbb{T}^d = (-\pi, \pi]^d$, the 2D Navier–Stokes equations and the 3D hyperviscous Navier–Stokes equations. With a slight abuse of notation, we will define a natural Hilbert space on velocity fields $u : \mathbb{T}^d \rightarrow \mathbb{R}^d$ by

$$\mathbf{L}^2 := \left\{ u \in L^2(\mathbb{T}^d; \mathbb{R}^d) : \int u dx = 0, \operatorname{div} u = 0 \right\}.$$

However, each model has a different natural L^2 -based “energy.” For 2D Navier–Stokes, it is the enstrophy $\|\operatorname{curl} u\|_{\mathbf{L}^2}^2$ and for 3D hyperviscous Navier–Stokes it is the kinetic energy $\|u\|_{\mathbf{L}^2}^2$. In order to take advantage of this energy structure for both the 2D and 3D cases, we will find it notationally convenient to define the following dimension dependent norm:

$$(1.5) \quad \|u\|_{\mathbf{W}} := \begin{cases} \|\operatorname{curl} u\|_{L^2} & \text{if } d = 2, \\ \|u\|_{L^2} & \text{if } d = 3, \end{cases}$$

and let $\mathbf{W}$ be the subspace of $\mathbf{L}^2$ with finite $\|\cdot\|_{\mathbf{W}}$ norm. We will also need to keep track of higher regularity: for $s > 0$, define

$$\mathbf{H}^s = \left\{ u \in H^s(\mathbb{T}^d, \mathbb{R}^d) : \int u dx = 0, \operatorname{div} u = 0 \right\}.$$

Following the convention used in [15, 32], we define a natural real Fourier basis on L^2 divergence free fields by defining for each $m = (k, i) \in \mathbb{K} := \mathbb{Z}_0^d \times \{1, \dots, d-1\}$,

$$e_m(x) = \begin{cases} c_d \gamma_k^i \sin(k \cdot x), & k \in \mathbb{Z}_+^d, \\ c_d \gamma_k^i \cos(k \cdot x), & k \in \mathbb{Z}_-^d, \end{cases}$$

where $\mathbb{Z}_0^d := \mathbb{Z}^d \setminus \{0, \dots, 0\}$, $\mathbb{Z}_+^d = \{k \in \mathbb{Z}_0^d : k^{(d)} > 0\} \cup \{k \in \mathbb{Z}_0^d : k^{(1)} > 0, k^{(d)} = 0\}$ and $\mathbb{Z}_-^d = -\mathbb{Z}_+^d$, and for each $k \in \mathbb{Z}_0^d$, $\{\gamma_k^i\}_{i=1}^{d-1}$ is a set of $d-1$ orthonormal vectors spanning the plane perpendicular to $k \in \mathbb{R}^d$ with the property that $\gamma_{-k}^i = -\gamma_k^i$. The constant $c_d = \sqrt{2}(2\pi)^{-d/2}$ is a normalization factor so that $e_m(x)$ are a complete orthonormal basis on L^2 divergence free vector fields. Note that in dimension $d = 2$ $\mathbb{K} = \mathbb{Z}_0^d$, hence $\gamma_k^1 = \gamma_k$ is just a vector in $\mathbb{R}^2$ perpendicular to k and is therefore given by $\gamma_k = \pm k^\perp / |k|$.

The white-in-time noise that we will force our equations with is $Q\dot{W}_t$, where W_t is a cylindrical Wiener process on $\mathbf{L}^2$ with respect to an associated canonical stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$ and Q is a positive Hilbert–Schmidt operator on $\mathbf{L}^2$, which we assume can be diagonalized with respect to $\{e_m\}$ with eigenvalues $\{q_m\} \in \ell^2(\mathbb{K})$ defined by

$$Qe_m = q_m e_m, \quad m = (k, i) \in \mathbb{K}.$$

In this way, $Q\dot{W}_t$ can be represented in terms of the basis $\{e_m\}$ by

$$Q\dot{W}_t = \sum_{m \in \mathbb{K}} q_m e_m \dot{W}_t^m,$$

where $\{W_t^m\}_{m \in \mathbb{K}}$ are a collection of i.i.d. one-dimensional Wiener processes with respect to $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbf{P})$.

We will assume that Q satisfies the following regularity and nondegeneracy assumption.

ASSUMPTION 1. There exists α satisfying $\alpha > \frac{5d}{2}$ and a constant C such that

$$\frac{1}{C} \|(-\Delta)^{-\alpha/2} u\|_{\mathbf{L}^2} \leq \|Qu\|_{\mathbf{L}^2} \leq C \|(-\Delta)^{-\alpha/2} u\|_{\mathbf{L}^2}.$$

Equivalently,

$$(1.6) \quad |q_m| \approx |k|^{-\alpha}.$$

REMARK 1.1. Assumption 1 essentially says that the forcing is QW_t has high spatial regularity but cannot be C^∞ . The nondegeneracy requirement on Q can be weakened to a more mild nondegeneracy at only high frequencies, that is, enforcing (1.6) only when $|k| \geq L$ for some fixed $L \gg 1$ (see [15] for more details), but fully nondegenerate noise simplifies some arguments. We are currently unable to treat noise, which is degenerate at high frequencies as in [42, 44]. See Remark 2.20 below for a more precise discussion of the two places we fundamentally depend on Assumption 1 at high frequencies.

We will define our primary phase space of interest to be the following:

$$\mathbf{H} = \mathbf{H}^\sigma, \quad \text{for some fixed } \sigma \in \left(\alpha - 2(d-1), \alpha - \frac{d}{2}\right).$$

Note that we have chosen α sufficiently large to ensure that $\sigma > \frac{d}{2} + 3$ so that we have the embedding $\mathbf{H} \hookrightarrow C^3$. We will consider a stochastic evolution (u_t) in $\mathbf{H}$, which we refer to as the *velocity process*, solving one of the two following stochastic PDEs:

SYSTEM 1 (2D Navier–Stokes equations).

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

where $u_0 = u \in \mathbf{H}$ and the viscosity parameter $\nu > 0$ is fixed.

SYSTEM 2 (3D hyperviscous Navier–Stokes).

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

where $u_0 = u \in \mathbf{H}$. Here, the hyperviscosity parameter $\nu > 0$ is fixed.

REMARK 1.2. It is well known that System 2 is globally well posed when the bi-Laplacian Δ^2 is replaced by $-(-\Delta)^\alpha$ for any $\alpha > 5/4$ (by a straightforward generalization of the proof in 2D). We leave it to a future work to check that the proof here of Theorem 1.4 below is valid for these alternative hyperviscosity terms.

The following well-posedness theorem is classical (See Section 3.1).

PROPOSITION 1.3. *For both Systems 1, 2 and all initial data $u \in \mathbf{H}$, there exists a $\mathbf{P}$ -a.s. unique, global-in-time, $\mathcal{F}_t$ -adapted mild solution (u_t) satisfying $u_0 = u$. Moreover, (u_t) defines a Feller–Markov process and the corresponding Markov semigroup has a unique stationary probability measure μ on $\mathbf{H}$.*

With the (u_t) process on $\mathbf{H}$ as in Proposition 1.3, we write $\phi^t = \phi_u^t$ for the stochastic flow of diffeomorphisms on $\mathbb{T}^d$ solving (1.3) with initial velocity field u . This gives rise to an $\mathcal{F}_t$ -adapted, Feller–Markov process (u_t, x_t) on $\mathbf{H} \times \mathbb{T}^d$, which we refer to as the *Lagrangian process* defined by $x_t = \phi^t(x)$, where $x_0 = x \in \mathbb{T}^d$. Uniqueness of the stationary measure $\mu \times \text{Leb}$ for the process (u_t, x_t) for both of Systems 1 and 2 was proved in [15].

1.2. *Main results and discussion.* We are now ready to state our main results, for which we give further context and discussion afterwards.

THEOREM 1.4. *Let (u_t) be as in any of Systems 1–2, initiated at μ -generic $u \in \mathbf{H}$, with μ as in Proposition 1.3. Fix $s > 0$ and $p \geq 1$. Then there exists a (deterministic) constant $\hat{\gamma} = \hat{\gamma}(s, p) > 0$, depending only on s, p and the parameters of the system (e.g., Q, ν , etc.), and a random constant $D = D(\omega, u) : \Omega \times \mathbf{H} \rightarrow [1, \infty)$ depending on the same parameters as $\hat{\gamma}$ and additionally on (ω, u) , the sample path and the initial data, which satisfy the following properties:*

(i) *For any $f, g \in H^s(\mathbb{T}^d)$, satisfying $\int f \, dx = \int g \, dx = 0$ the following holds for all $t \geq 0$:*

$$(1.7) \quad \left| \int f(x) g(\phi_u^t(x)) \, dx \right| \leq D \|f\|_{H^s} \|g\|_{H^s} e^{-\hat{\gamma}t}.$$

(ii) *For u fixed, $D(\cdot, u)$ is $\mathbf{P}$ -a.e. finite and moreover satisfies the following: $\exists \beta \geq 2$ (independent of u, p, s) such that $\forall \eta > 0$ there holds*

$$(1.8) \quad \mathbf{E} D^p \lesssim_{p, \eta, s} (1 + \|u\|_{\mathbf{H}}^2)^{p\beta} \exp(\eta \|u\|_{\mathbf{W}}^2) < \infty.$$

For $s \leq 1$ and $p \geq 1$, one can take $\hat{\gamma} \gtrsim \frac{s}{p}$.

As discussed above (1.4), an immediate corollary of Theorem 1.4 is H^{-s} decay for passively advected scalars in the absence of dissipation and with no sources as in (1.1) above.

COROLLARY 1.5. *In the setting of Theorem 1.4, fix $s > 0$, $p \geq 1$, and let $\hat{\gamma} = \hat{\gamma}(s, p)$ as in Theorem 1.4. Then, for any mean-zero initial $g \in H^s$, the solution g_t to the passive scalar problem (1.1) satisfies the estimate*

$$(1.9) \qquad \|g_t\|_{H^{-s}} \leq D e^{-\hat{\gamma}t} \|g\|_{H^s},$$

where D is the random constant appearing in Theorem 1.4.

REMARK 1.6. By time reversibility of the advection equation, no pointwise in time decay, such as in (1.9), can hold without some regularity assumption on g . Indeed, it is well known that nonuniformity of correlation decay in this sense is ubiquitous among mixing systems; see, for example, Section 2.5 of [61]. On the other hand, by a density argument, for μ -generic u and all mean-zero $g \in L^2$, we have $g_t \rightarrow 0$ in L^2 as $t \rightarrow \infty$ $\mathbf{P}$ -a.e. at a rate which is uniform over compact sets in L^2 . See [34] for more discussion.

REMARK 1.7. It is known that uniformly $W^{1,p}$ velocity fields for $p > 1$ can mix scalars at-most exponentially fast (see [46, 68]). On the other hand, the stochastic models in Systems 1, 2 are never uniformly bounded in time in any Sobolev space due to inevitable large fluctuations in the Brownian paths. Nevertheless, an exponential decay rate is still effectively optimal in this setting, due to control on large deviations of the velocity fields (see Lemma 3.10). From this, it is not hard to show the following: $\forall s > 0$, $p \geq 1$, $\exists \gamma' = \gamma'(s, p) > 0$ (deterministic) and a random constant $D' = D'(\omega, u) : \Omega \times \mathbf{H} \rightarrow [1, \infty)$ satisfying (1.8) such that for all $g \in H^{-s}$,

$$\|g_t\|_{H^{-s}} \geq (D')^{-1} e^{-\gamma't} \|g\|_{H^{-s}}.$$

REMARK 1.8. One always has $\|g_t\|_{H^{-s}} \leq \|g\|_{L^2}$, hence D quantifies an upper bound on the random time-scale one will have to wait in order to see the exponential decay $\hat{\gamma}$. If one normalizes $\|g\|_{L^2} = 1$, then mixing occurs for times $t \gtrsim \tau = \frac{1}{\hat{\gamma}} \log(D\|g\|_{H^s})$ and estimates such as (1.8) shows that $\exists \delta > 0$ such that the following exponential moment holds: $\exists \beta'$ large so that $\forall \eta > 0$,

$$\mathbf{E} e^{\delta \tau} \lesssim_{\eta} (1 + \|u\|_{\mathbf{H}}^2)^{\beta'} \exp(\eta \|u\|_{\mathbf{W}}^2) \|g\|_{H^s}.$$

REMARK 1.9. After the completion of this work, we have since produced two follow-up papers: first [17], which uses the main results herein to extend to the advection-diffusion equation uniformly in diffusivity; and second, [16], which uses the results therein to provide a rigorous proof of the Batchelor power spectrum in passive scalar turbulence (see, e.g., [3, 13, 28, 69] and the references therein for more background).

It is well known that sufficiently regular velocity fields can mix at most exponentially fast. Refining exactly the relation between regularity and mixing rate is the content of Bressan’s conjecture [18] and has been studied in several works, for example [46, 68], and the references therein. In the other direction, the construction of exponential mixers has also proved challenging. Roughly speaking, deterministic, time-autonomous flows possess a coherent flow direction along which no mixing can occur, presenting significant difficulties in the construction of such mixers. Classical results on exponential mixing for Axiom A (uniformly hyperbolic) flows include [26, 53]; see also [66] for a discussion of challenges in establish correlation decay in this setting.

Exponential mixers which are Hölder continuous and smooth away from a finite set of hyperplanes were constructed in [34] (the fields studied in Theorem 1.4 are not C_x^∞ but for all $k < \infty$, can be chosen C_x^k by choosing α large enough). See also the work [2] where the

velocity is chosen depending on the scalar, and [73] where the decay is on long but finite time-intervals.

Informally, the local/infinitesimal mechanism responsible for mixing is hyperbolicity, or stretching and contracting in various directions in phase space. Given a Lagrangian flow ϕ^t , hyperbolicity measured by the Lyapunov exponent, that is, the asymptotic exponential growth rate

$$\lambda := \lim_t \frac{1}{t} \log \|D_x \phi^t\|$$

at which spatial gradients grow for typical x (when this limit exists); note that when ϕ^t is incompressible, $\lambda \geq 0$. A positive value ($\lambda > 0$) indicates exponential-in-time separation of nearby Lagrangian trajectories, hence hyperbolic stretching / contraction. Checking this positivity, even for seemingly innocent toy models of incompressible flow such as the Chirikov standard map [21], is a notoriously challenging (and largely open) problem due to the convoluted coexistence of hyperbolic trajectories and coherent structures such as elliptic islands along which no hyperbolicity takes place [30, 40, 60, 74].

The introduction of randomness is a natural way to rule out the formation of permanent coherent structures. The classical work of Pierrehumbert [62] studied passive scalar advection under alternating sine shear flows with random phase, obtaining numerical evidence of almost-sure exponential mixing for this model.¹ Following the seminal work of Furstenberg [37] on i.i.d. compositions of random matrices, a variety of authors (see, e.g., [20, 51] and references therein) developed techniques to prove that under nondegeneracy conditions on the noise (ruling out, in particular, the almost-sure preservation of permanent coherent structures), incompressible SDE generate stochastic flows of diffeomorphisms with positive Lyapunov exponents. Building on this, for such models, almost-sure exponential mixing for a class of nondegenerate SDEs was proved in [14, 27].

Note that incompressible SDE correspond to white-in-time velocity fields, that is, Lagrangian particles are directly driven by noise. Our contribution is to adapt this circle of ideas to the more general setting where the Lagrangian flow is *indirectly* forced by noise acting on a velocity field, itself evolving according to a stochastic fluid model in an infinite-dimensional phase space. This presents fundamental difficulties requiring many new ideas to treat; see Section 2 below.

This manuscript and the proof of Theorem 1.4 builds off our previous work [15], where we established a positive Lyapunov exponent for ϕ^t for all initial $x \in \mathbb{T}^d$ and velocity fields u . Going from this to exponential mixing often requires more: hyperbolic expansion is *local* in x , and is therefore much weaker than the *global* statement of Theorem 1.4. Indeed, a positive Lyapunov exponent is basically equivalent to an (almost-sure, eventual) exponential growth of $\|\nabla g_t\|_{L^2}$ for any nonzero $g \in H^1$ (see the discussion in [15]), which is much weaker than (1.7). One can construct many examples of dynamical systems with a positive Lyapunov exponent but arbitrarily slow (e.g., polynomial or logarithmic) decay of correlations, for example, Pommeau–Manneville maps (see, e.g., [67]).

REMARK 1.10 (The role of nonlinearity). If one drops the nonlinearity from Systems 1 or 2, one is left with the Stokes equation, and the velocity field is simply

$$u_t(x) = \sum_{m \in \mathbb{K}} \beta_t^m e_m(x),$$

¹To our knowledge, at time of writing there is no proof in the literature that this model provides an exponential mixer with probability 1.

where $\{\beta_t^m\}_{m \in \mathbb{K}}$ is a family of independent Ornstein–Uhlenbeck processes with variances $\frac{q_m^2}{2\nu|k|^2}$ (with $m = (k, i)$). Heuristically, one can expect the Lagrangian flow to behave similar to successively applying i.i.d. random shear flows of all orientations, which would resemble a hyperbolic toral automorphism. In accordance with this intuition, the Stokes equations is by far the easiest case to treat.

In our work here and in [15], the nonlinearity in the Navier–Stokes equations is (a priori) an *enemy*. The modes become correlated through the nonlinearity which can introduce structure that could inhibit hyperbolicity of the flow map. For example, in two dimensions, the Navier–Stokes equations form coherent vortices inside of which there is no hyperbolicity. See, for example, [9, 10, 63, 64] and the references therein for discussions on this and its implications for passive scalar dynamics. The Furstenberg criteria arguments in [15], for example, prove that such vortices cannot permanently entrain any part of the Lagrangian flow. Theorem 1.4 gives exponential tail control on any slow down of hyperbolicity caused by transient coherent vortices.

The Navier–Stokes equations require the creation of significantly more robust tools than those that existed previously in the literature (see Section 2). Finally, it bears remarking that the case of Navier–Stokes for $0 < \nu \ll 1$ is the model of primary physical relevance for Batchelor-regime passive scalar turbulence (see [16] and the references therein). Other models are significantly less physical (i.e., toy models) or apply only to a very small set of experimental settings. Hence, it is crucial to develop methods capable of treating Navier–Stokes at arbitrary Reynolds number.

1.3. Finite dimensional evolution and $C_t^k C_x^\infty$ almost-sure exponential mixers. When considering finite dimensional evolutions for the velocity fields, our methods significantly simplify it suffices to impose much weaker nondegeneracy conditions on the noise.

ASSUMPTION 2 (Low mode nondegeneracy). Define $\mathcal{K}_0 \subset \mathbb{K}$ to be the set of $m \in \mathbb{K}$ such that $q_m \neq 0$. Assume $m \in \mathcal{K}_0$ if $|m|_\infty \leq 2$ (for $m = (k, i)$, $k = (k_i)_{i=1}^d \in \mathbb{Z}^d$ we write $|m|_\infty = \max_i |k_i|$).

We write $\mathbf{H}_{\mathcal{K}_0} \subset \mathbf{H}$ for the subspace spanned by the Fourier modes $m \in \mathcal{K}_0$ and $\mathbf{H}_N \subset \mathbf{H}$ for the subspace spanned by the Fourier modes satisfying $|m|_\infty \leq N$. Consider the Stokes system (with very degenerate forcing) and Galerkin–Navier–Stokes systems defined as the following.

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

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

where Q satisfies Assumption 2 and $\mathcal{K}_0$ is finite.

SYSTEM 4. We refer to the *Galerkin–Navier–Stokes system* in $\mathbb{T}^d$ ($d = 2, 3$) as the following, for $u_0 = u \in \mathbf{H}_N$:

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

where Q satisfies Assumption 2; $N \geq 3$ is an integer; Π_N denotes the projection to Fourier modes with $|\cdot|_\infty$ norm $\leq N$; $\mathbf{H}_N$ denotes the span of the first N Fourier modes; and $\nu > 0$ is fixed and arbitrary.

In addition to Systems 3 and 4, which feature fluid models subjected to white-in-time forcing, the methods easily extend to treat fluid systems subjected to certain types of forcing that we refer to as “OU tower noise.” This is basically an external forcing that is a projection of an Ornstein–Uhlenbeck process on $\mathbb{R}^M$. Note that the force can be C_t^k , $k \geq 1$ (see Remark 1.13 below). In particular, we consider the following set of systems (stated in a general Navier–Stokes-like setting).

SYSTEM 5. We refer to the (*generalized*) *Galerkin–Navier–Stokes system with OU tower noise* in $\mathbb{T}^d$ ($d = 2, 3$) as the following stochastic ODE for $u_0 \in \mathbf{H}_N$:

$$\partial_t u_t + X(u, u) = \nu \Delta u_t + Q Z_t,$$

$$\partial_t Z_t = -\mathcal{A} Z_t + \Gamma \dot{W}_t,$$

where $Z_t \in \mathbf{H}_M$, the operator $\mathcal{A} : \mathbf{H}_M \rightarrow \mathbf{H}_M$ is diagonalizable and has a strictly positive spectrum, and the bilinear term $X(u, u) : \mathbf{H}_N \times \mathbf{H}_N \rightarrow \mathbf{H}_N$ satisfies $u \cdot X(u, u) = 0$ and $\forall j$, $X(e_j, e_j) = 0$. Note that (u_t) is not Markov, but (u_t, Z_t) is Markov.

Theorem 1.4 extends to all of Systems 3, 4 and 5.

THEOREM 1.11. *Consider any of Systems 3–5. Assume that Q satisfies Assumption 2 and that the parabolic Hörmander condition is satisfied for (u_t) or (u_t, Z_t) (see, e.g., [41]). Then:*

- (i) *the Lagrangian flow (1.3) has a strictly positive Lyapunov exponent in the same sense as described in [15];*
- (ii) *all of the results of Theorem 1.4 hold.*

REMARK 1.12. We have chosen to include Theorem 1.11 to emphasize that our methods do not fundamentally require nonspatially smooth velocity fields, nor do they require velocity fields that are directly subjected to white-in-time forcing. The difficulty in extending Theorem 1.4 to include $C_t^k C_x^\infty$ velocity fields is the lack of sufficiently strong hypoellipticity results in infinite-dimensions, that is, the lack of a sufficiently strong replacement for Hörmander’s theorem. See Remark 2.20 below.

REMARK 1.13. Theorem 1.11 contains many examples of $C_t^k C_x^\infty$ velocity fields for any $k \geq 0$. For example, consider the following:

$$u_t(x) = \sum_{m \in \mathbb{K} : |m|_\infty \leq 2} \hat{u}_t^m e_m(x),$$

where the coefficients are given by the system

$$\begin{aligned} \partial_t \hat{u}_t^m &= -\hat{u}_t^m + Z_t^{m,0}, \\ \partial_t Z_t^{m,\ell} &= -(\ell + 1) Z_t^{m,\ell} + Z_t^{j,\ell-1} \quad 1 \leq \ell \leq n, \\ \partial_t Z_t^{m,n} &= -Z_t^{m,n} + \dot{W}_t^m. \end{aligned}$$

By indexing correctly, one can rewrite this in the form stated in System 5 with $X \equiv 0$. One can check that the parabolic Hörmander condition is satisfied. This example also explains the terminology “OU tower.”

2. Outline of the proof. In this section, we give the main steps for the proof of Theorem 1.4; details will be given in Sections 3–7. The exposition we give here focuses on the infinite-dimensional Systems 1, 2, with the finite-dimensional systems described in Section 1.3 addressed in a series of remarks as we go along.

2.1. *Correlation decay via geometric ergodicity of the two point Lagrangian motion.* Our primary tool for investigating quenched correlation decay as in Theorem 1.4 is to study the following Markov process.

DEFINITION 2.1. We define the *two-point Lagrangian process* (u_t, x_t, y_t) on $\mathbf{H} \times \mathbb{T}^d \times \mathbb{T}^d$ by

$$x_t = \phi_u^t(x), \quad y_t = \phi_u^t(y)$$

for fixed initial $(u, x, y) \in \mathbf{H} \times \mathbb{T}^d \times \mathbb{T}^d$.

The two-point process (u_t, x_t, y_t) simultaneously tracks the velocity field u_t as well as two separate Lagrangian flow trajectories x_t, y_t . Throughout, we assume $x \neq y$, hence $x_t \neq y_t$ for all $t > 0$, since the diagonal

$$\mathcal{D} = \{(x, x) : x \in \mathbb{T}^d\} \subset \mathbb{T}^d \times \mathbb{T}^d$$

and its complement $\mathcal{D}^c$ are invariant sets.

In what follows, we write $P_t^{(2)}$ for the Markov semigroup associated to the Markov process (u_t, x_t, y_t) on $\mathbf{H} \times \mathcal{D}^c$, defined on observables $\psi : \mathbf{H} \times \mathcal{D}^c \rightarrow \mathbb{R}$ by

$$P_t^{(2)}\psi(u, x, y) := \mathbf{E}_{(u,x,y)}\psi(u_t, x_t, y_t)$$

whenever the right-hand side is defined. Note that $\mu \times \text{Leb} \times \text{Leb}$ is automatically a stationary measure for the two-point process. We will deduce Theorem 1.4 from geometric ergodicity with respect to this measure.

THEOREM 2.2. *There exist $\hat{\alpha} > 0$ and a measurable function $\mathcal{V} : \mathbf{H} \times \mathcal{D}^c \rightarrow [1, \infty)$, with $\mathcal{V} \in L^1(\mu \times \text{Leb} \times \text{Leb})$ such that for each measurable bounded $\psi : \mathbf{H} \times \mathbb{T}^d \times \mathbb{T}^d$ with $\iint \psi \, d\mu \, dx \, dy = 0$ and each $(u, x, y) \in \mathbf{H} \times \mathcal{D}^c$, we have that*

$$|P_t^{(2)}\psi(u, x, y)| \leq \mathcal{V}(u, x, y)e^{-\hat{\alpha}t} \|\psi\|_{L^\infty}$$

for all $t \geq 0$.

REMARK 2.3. The idea of using geometric ergodicity of a two-point process to deduce quenched correlation decay is known to experts in random dynamical systems, although it does not appear to be generally well known. To the best of our knowledge, this idea first appears in the literature in [27] on quenched correlation decay for SDE on compact manifolds (see also [8]).

The bulk of the work in this paper is aimed at proving Theorem 2.2. Before proceeding to describe the proof, we first give an indication of how Theorem 2.2 will be used to deduce Theorem 1.4. Fix mean-zero $f, g \in L^\infty(\mathbb{T}^d)$ and $\hat{\gamma} \in (0, \frac{\hat{\alpha}}{2})$. For all $n \in \mathbb{Z}_{\geq 0}$, $u \in \mathbf{H}$,

$$\begin{aligned} \mathbf{P}\left\{\left|\int f(g \circ \phi_u^n) \, dx\right| > e^{-\hat{\gamma}n}\right\} &\leq e^{2\hat{\gamma}n} \mathbf{E}\left(\int f(g \circ \phi_u^n) \, dx\right)^2 \\ &= e^{2\hat{\gamma}n} \int \tilde{f}(x, y)(P_n^{(2)}\tilde{g})(u, x, y) \, dx \, dy, \end{aligned}$$

where $\tilde{f}(x, y) = f(x)f(y)$, $\tilde{g}(x, y) = g(x)g(y)$. By Theorem 2.2, the above right-hand side is bounded $\lesssim_{u,f,g} e^{(2\hat{\gamma}-\hat{\alpha})n}$. We conclude by the Borel–Cantelli lemma that

$$\left|\int f(g \circ \phi_u^n) \, dx\right| \leq \tilde{D}e^{-\hat{\gamma}n},$$

for all $n \geq 1$, where $\tilde{D} = \tilde{D}(f, g, \omega, u)$ is a random constant depending on f, g and the initial velocity field u . Additional work is needed to determine the dependence of the random constant $\tilde{D}$ on f, g and u , as well as to pass from discrete to continuous time. These arguments are carried out in detail in Section 7.

2.2. Conditions for geometric ergodicity. A prevailing strategy for proving correlation decay for Markov chains on noncompact spaces is to verify conditions guaranteeing that the Markov process visits a given “bounded” region of phase space with a positive asymptotic frequency. Many criteria of this kind have been developed (sometimes called Harris theorems): for a detailed account see, for example, the reference [57]. Note that when considering the 2-point process, the diagonal $x = y$ must be considered “at infinity” and the primary difficulty of the proof will be localizing the process away from this degenerate set in phase space.

For obtaining geometric ergodicity of the 2-point process, we will implement a series of criteria developed by Goldys and Maslowski [39], particularly useful for semigroups generated by SPDE, for checking the conditions of the abstract Harris-type theorems (cf. [57]; see also, for example, [19, 47–49, 56] for other approaches to geometric ergodicity for SPDE). A simplified version of their framework is as follows. Let $\mathcal{Z}$ be a Polish space and let $(Z_t)_{t \geq 0}$ be a continuous-time Markov process on $\mathcal{Z}$ with transition kernels $P_t(z, K) = \mathbf{P}_z(Z_t \in K)$. As usual we define the Markov semigroup on observables $\psi : \mathcal{Z} \rightarrow \mathbb{R}$ by

$$P_t \psi(z) = \mathbf{E}_z(\psi(Z_t)) = \int_{\mathcal{Z}} \psi(z') P_t(z, dz').$$

CONDITION 1 (Strong Feller). We say that a Markov process (Z_t) is *strong Feller* if for all $t > 0$ and bounded measurable $\psi : \mathcal{Z} \rightarrow \mathbb{R}$, we have that $z \mapsto P_t \psi(z)$ is continuous on $\mathcal{Z}$ for all $t > 0$.

CONDITION 2 (Topological irreducibility). We say that a Markov process (Z_t) is *topologically irreducible* if for all open $U \subset \mathcal{Z}$, we have that $P_t(z, U) > 0$ for all $t > 0, z \in \mathcal{Z}$.

The next two conditions refer to a given measurable function $\mathcal{V} : \mathcal{Z} \rightarrow [1, \infty)$.

CONDITION 3 (Uniform lower bounds). For each $r > 1$, there exists a compact set $K \subset \mathcal{Z}$ and a time $t_0 = t_0(r) > 0$ such that

$$\inf_{\{\mathcal{V}(z) \leq r\}} P_{t_0}(z, K) > 0.$$

CONDITION 4 (Drift condition). We say $\mathcal{V} : \mathcal{Z} \rightarrow [1, \infty)$ satisfies a *drift condition* (aka Lyapunov function) if there are constants $k, \kappa, c > 0$ such that

$$P_t \mathcal{V} \leq k e^{-\kappa t} \mathcal{V} + c$$

holds pointwise.

Below, given $\mathcal{V} : \mathcal{Z} \rightarrow [1, \infty)$, we write $C_{\mathcal{V}}$ for the Banach space of continuous observables $\psi : \mathcal{Z} \rightarrow \mathbb{R}$ for which the norm

$$(2.1) \quad \|\psi\|_{C_{\mathcal{V}}} := \sup_{z \in \mathcal{Z}} \frac{|\psi(z)|}{\mathcal{V}(z)}$$

is finite.

THEOREM 2.4 (Follows from Theorem 3.1 and Lemma 3.2 in [39]). *Suppose that the Markov process (Z_t) and the function $\mathcal{V} : \mathcal{Z} \rightarrow [1, \infty)$ satisfy Conditions 1, 2, 3 and 4. Then the Markov process (Z_t) admits a unique stationary measure m , with respect to which (Z_t) is geometrically ergodic in $C_{\mathcal{V}}$. That is, for all $\psi \in C_{\mathcal{V}}$, we have that*

$$\left| P_t \psi(z) - \int \psi \, dm \right| \leq C \mathcal{V}(z) e^{-\beta t} \|\psi\|_{C_{\mathcal{V}}} \quad \text{for all } t > 0,$$

where $C > 0$, $\beta > 0$ are constants.

REMARK 2.5. Harris-type theorems typically have two sets of assumptions: a *minorization condition* satisfied by a subset S of phase space, which guarantees that orbits initiated from S couple with some uniform probability > 0 , and a *drift condition* (Condition 4), which controls excursions from S . Conditions 1–3 are sufficient for verifying a suitable minorization condition for the sublevel sets $\{\mathcal{V} \leq r\}$, $r > 0$. See [45, 57] for more discussion on abstract Harris theorems.

We will apply Theorem 2.4 to the two-point Markov process (u_t, x_t, y_t) on the state space $\mathbf{H} \times \mathcal{D}^c$. Along the way, in Section 4.1 we also apply Theorem 2.4 to another related Markov process.

2.3. Strong Feller and irreducibility. Let us now begin sketching how to verify the conditions of Theorem 2.4 for the two-point process (u_t, x_t, y_t) on $\mathbf{H} \times \mathcal{D}^c$. First, we will prove the strong Feller property as in Condition 1 on a scale of Sobolev spaces—this refinement is useful in verifying Condition 3 later on. Note that the evolution of (x_t, y_t) in Definition 2.1 is not subject to noise. After verifying the requisite uniform Hörmander conditions (Section 6.1.1), the proof follows by methods used previously in [15]. A brief sketch is included in Section 6.1 for completeness.

PROPOSITION 2.6. *For any $\sigma' \in (\alpha - 2(d - 1), \alpha - \frac{d}{2})$ the two-point Markov process (u_t, x_t, y_t) on $\mathbf{H}^{\sigma'} \times \mathcal{D}^c$ is strong Feller.*

Next, we verify topological irreducibility as in Condition 2. This follows by a relatively simple approximate controllability statement; see Section 6.2 for the details.

PROPOSITION 2.7. *For any $\sigma' \in (\frac{d}{2} + 2, \alpha - \frac{d}{2})$, the two-point process (u_t, x_t, y_t) , regarded as a process on $\mathbf{H}^{\sigma'} \times \mathcal{D}^c$, is topologically irreducible.*

REMARK 2.8. Checking Conditions 1–4 for the finite-dimensional systems in Section 1.3 is considerably easier. Conditions 1 and 2 follow immediately from Hörmander’s theorem. By similar arguments, Condition 3 follows immediately as long as the function $\mathcal{V}$ has the property that its sublevel sets $\{\mathcal{V} \leq r\} \subset \mathbf{H} \times \mathcal{D}^c$ are bounded (also away from $\mathcal{D}$), hence compact, for all $r > 1$. We return to the construction of $\mathcal{V}$ and Condition 4 for these systems in Remark 2.19 below.

2.4. Construction of the Lyapunov function $\mathcal{V}$. Let us now turn to the most difficult task: constructing a Lyapunov function $\mathcal{V}$ for the two-point process which satisfies Conditions 3 and 4.

2.4.1. *Lyapunov functions for the (u_t) process.* The first step is to find a suitable Lyapunov functional for the (u_t) process. Unlike previous works we are aware of treating geometric ergodicity of stochastic Navier–Stokes (see, e.g., [39, 43]), we need the Lyapunov functions to control (u_t) in $\mathbf{H}$ regularity. To our knowledge, these have not previously appeared in the literature and a somewhat nontrivial additional effort is needed to deduce them (but see closely related results in Section 3.5.3 of [50]). The resulting geometric ergodicity statements are of some independent interest.

To simplify notation, we will denote each corresponding Lyapunov function by the same symbol $V_{\beta, \eta}$. Lemma 2.9 is proved below in Section 3 (see Lemma 3.7).

LEMMA 2.9. *If $d = 2$, define $\mathcal{Q} = \sup_{m=(k,i) \in \mathbb{K}} |k| |q_m|$, and if $d = 3$ define $\mathcal{Q} = \sup_{m=(k,i) \in \mathbb{K}} |q_m|$. Let $0 < \eta < \eta^* = \nu/64\mathcal{Q}$, $\beta \geq 0$, and define*

$$(2.2) \quad V_{\beta, \eta}(u) = (1 + \|u\|_{\mathbf{H}}^2)^\beta \exp(\eta \|u\|_{\mathbf{W}}^2),$$

where $\|\cdot\|_{\mathbf{W}}$ is as in (1.5). Then (2.2) satisfies Condition 4 for the (u_t) process.

REMARK 2.10. Naturally, for Systems 3 and 4 we need only to consider the case $\beta = 0$ in Lemma 2.9. For the System 5, V must also depend on Z ((u_t, Z_t) is now the relevant Markov process). Let $\mathcal{A} = \Xi D \Xi^{-1}$ be the diagonalization of $\mathcal{A}$. Then, for η_1, η_2 chosen sufficiently small (depending on $D, \Xi, \mathcal{Q}, \Gamma, \nu$) it suffices to take the following for V :

$$V(u, Z) = \exp(\eta_1 \|u\|^2 + \eta_2 \|\Xi^{-1} Z\|^2).$$

Indeed, note that $Y_t = \Xi^{-1} Z_t$ solves $dY_t = -DY_t + \Xi^{-1} \Gamma \dot{W}_t$, which is compatible with the $\Xi^{-1} Z$ factor in V ; the $\mathcal{Q} Z_t$ term in the (u_t) evolution is absorbed by $-\|\nabla u_t\|^2$ and $-\|D^{1/2} Y_t\|^2$ (cf. Section 3.2).

2.4.2. *Repulsion from the diagonal $\mathcal{D}$ for the two-point process.* The families of Lyapunov functions defined in Lemma 2.9 for the (u_t) process capture repulsion from parts of phase space where $\|u_t\|_{\mathbf{H}}$ is unboundedly large. For the two-point process, however, we additionally require repulsion from the diagonal $\mathcal{D}$, which is considered part of “infinity” for the two point motion (x_t, y_t) . For this, we will crucially use the fact that near the diagonal the positive Lyapunov exponent for the Lagrangian flow ϕ^t (as established in [15]) causes x_t and y_t to diverge from each other at an exponential rate with probability 1.

To make this more precise, suppose x, y are close together and consider the coordinate change $(x, y) \mapsto (x, w)$, where $w = w(x, y)$ is the minimal displacement vector from x to y . This induces a Markov process (u_t, x_t, w_t) , with $w_t = w(x_t, y_t) \in \mathbb{R}^d$, which is continuous in time as long as x_t, y_t remain close together. Note that $\mathbb{T}^d \times \{0\}$ plays the same role for the linearized process (x_t, w_t^*) as the diagonal $\mathcal{D}$ does for the two-point process (x_t, y_t) . Near the diagonal, we can approximate w_t by the linearized process (w_t^*) on $\mathbb{R}^d \setminus \{0\}$,

$$w_t \approx w_t^* := D_x \phi^t w, \quad w = w(x, y).$$

In [15], we proved that there exists $\lambda_1 > 0$ (deterministic) such that

$$(2.3) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \log |w_t^*| = \lambda_1 > 0,$$

with probability 1 for all initial $(u, x) \in \mathbf{H} \times \mathbb{T}^d$ and $w \in \mathbb{R}^d \setminus \{0\}$. Hence, one anticipates that on average, $|w_t^*|^{-p} \approx e^{-p\lambda_1 t} |w|^{-p}$ for $p > 0$. Thus, to capture repulsion from the diagonal it is natural to consider candidate Lyapunov functions of the form

$$(2.4) \quad f_p(u, x, w) = |w|^{-p} \psi_p(u, x, w/|w|),$$

where $\psi_p : \mathbf{H} \times \mathbb{T}^d \times \mathbb{S}^{d-1} \rightarrow \mathbb{R}$ is nonnegative (actually, it is natural to enforce $\psi_p(u, x, -w) = \psi_p(u, x, w)$, so ψ_p will be thought of as a function on the projective bundle² $\mathbf{H} \times P\mathbb{T}^d$).

Let us write TP_t for the Markov semigroup on $\mathbf{H} \times \mathbb{T}^d \times \mathbb{R}^d$ corresponding to (u_t, x_t, w_t^*) . Repulsion as in (2.3) suggests to look for an f_p that is an eigenvector of TP_t with eigenvalue < 1 . Then it is straightforward to see that we would have the following drift condition for TP_t :

$$TP_t f_p = e^{-\Lambda(p)t} f_p,$$

for some $\Lambda(p) > 0$.

More precisely, let $V = V_{\beta,\eta}$ be as in Lemma 2.9. We write C_V for the space of continuous functions on $\mathbf{H} \times P\mathbb{T}^d$ with bounded $\|\cdot\|_{C_V}$ norm and define C_V^1 to be the space of continuously differentiable functions on $\mathbf{H} \times P\mathbb{T}^d$ for which

$$\|\psi\|_{C_V^1} := \|\psi\|_{C_V} + \sup_{(u,x,v) \in \mathbf{H} \times P\mathbb{T}^d} \frac{\|D\psi(u, x, v)\|_{\mathbf{H}^*}}{V(u)}$$

is finite. As discussed in [43], since C_V and C_V^1 are not separable, it is sometimes necessary to work with the spaces $\mathring{C}_V$ and $\mathring{C}_V^1$, which are the C_V -closure and C_V^1 -closure of the space of smooth “cylinder functions”:

(2.5) $\mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d) = \{\psi | \psi(u, x, v) = \phi(\Pi_{\mathcal{K}}u, x, v), \mathcal{K} \subset \mathbb{K}, \phi \in C_0^\infty\},$

where $\Pi_{\mathcal{K}}$ denotes orthogonal projection onto $\mathbf{H}_{\mathcal{K}} \cong \mathbb{R}^{|\mathcal{K}|}$. These spaces are separable, closed subspaces on which finite-dimensional approximation can be made. It can be shown that, for example, C_V is strictly larger than $\mathring{C}_V$: indeed, $V_{\beta,\eta} \notin \mathring{C}_{V_{\beta,\eta}}$ (see Section 5.3 of [43]).

PROPOSITION 2.11. *For all $|p| \ll 1$, there exists $\psi_p \in \mathring{C}_V^1$, with the following properties:*

- (a) ψ_p is strictly positive and is bounded uniformly from below on bounded subsets of $\mathbf{H} \times P\mathbb{T}^d$.
- (b) $f_p(u, x, w) = |w|^{-p} \psi_p(u, x, w/|w|)$ is an eigenfunction of TP_t with eigenvalue $e^{-\Lambda(p)t}$.
- (c) As $p \rightarrow 0$ we have $\Lambda(p) = \lambda_1 p + o(p)$, where λ_1 the Lyapunov exponent (2.3).

REMARK 2.12. The value $-\Lambda(p)$ is sometimes referred to as the *moment Lyapunov exponent* [5], and arises naturally in the study of large deviations in the convergence of Lyapunov exponents; see, for example, [6, 7]. As we show, it satisfies the formula

$$-\Lambda(p) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbf{E} |D_x \phi_u^t v|^{-p}$$

for all $(u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$.

REMARK 2.13. As mentioned previously, the idea of using the two-point process to establish quenched correlation decay is due to [27], which studied quenched statistical properties of an incompressible SDE (in our terminology, a white-in-time velocity field). As in our setting, the phase space $\mathcal{D}^c$ for their two-point process is noncompact, and so a similar

²Here, $P\mathbb{T}^d \cong \mathbb{T}^d \times P^{d-1}$ is the projective bundle over $\mathbb{T}^d$, where $P^{d-1} = P(\mathbb{R}^d)$ is the projective space for $\mathbb{R}^d$.

drift condition is required: the one used in [27] cites the detailed study by Baxendale and Stroock [14] obtaining large deviations estimates sojourns to small neighborhoods of the diagonal $\mathcal{D}$. Our approach to drift conditions repelling from $\mathcal{D}$ borrows several ideas from [14], for example, the use of moment Lyapunov exponents and the linearized semigroup TP_t . We note, however, that the unbounded and infinite-dimensional phase space in our setting introduces numerous challenges to be overcome, for example, difficulties involving the use of infinitesimal generators in our setting.

2.4.3. Verifying the drift condition. Returning to the nonlinear two-point motion, in light of (2.4), it is natural to look for a Lyapunov function of the form

$$\mathcal{V}(u, x, y) = h_p(u, x, y) + V_{\beta+1, \eta}(u),$$

where β is sufficiently large, $\eta \in (0, \eta^*)$, and $h_p(u, x, y) := \chi(|w(x, y)|)f_p(u, x, w(x, y))$, where χ is a smooth cutoff satisfying $\chi|_{[0, 1/10]} \equiv 1$, $\chi|_{[1/5, \infty)} \equiv 0$.

For $\mathcal{V}$ as above, let us sketch how to satisfy Conditions 3 and 4. For Condition 3, by parabolic regularity and the strong Feller property in Condition 1, it suffices (Lemma 6.2) to show that for any $r > 0$, the sublevel set $\{\mathcal{V} \leq r\}$ satisfies

$$\{\mathcal{V} \leq r\} \subset \{\|u\|_{\mathbf{H}} \leq R_1 \text{ and } d(x, y) \geq R_2\},$$

where $R_i = R_i(r)$, $i = 1, 2$. We are guaranteed this as long as ψ_p is uniformly bounded from below on bounded sets as in Proposition 2.11(c).

For Condition 4, the validity of the linear approximation depends on the size of the velocity field. Since the latter can be arbitrarily large, the time scale on which this approximation can be used is nonuniform in u . For this reason, it is convenient to work at the time-infinitesimal level, that is, with infinitesimal generators. Temporarily neglecting technical issues regarding the domain, let $\mathcal{L}_{(2)} = \lim_{t \rightarrow 0} \frac{1}{t}(P_t^{(2)} - \text{Id})$ be the infinitesimal generator of the two-point motion. We will essentially show that

$$(2.6) \quad \mathcal{L}_{(2)}\mathcal{V} \leq -\Lambda(p)\mathcal{V} + C$$

for some $C > 0$. As one may expect, (2.6) cannot be rigorously justified exactly as such in infinite dimensions due to the fact that $V \notin \dot{C}_V$, but standard arguments are used to obtain an almost equivalent analogue (see Section 6.3).

Let us briefly sketch this argument, ignoring technical issues. When $\mathcal{L}_{(2)}$ hits h_p , we obtain

$$(2.7) \quad \mathcal{L}_{(2)}h_p \leq -\Lambda(p)h_p + C'V_{\beta+1, \eta};$$

the first term is good and comes from the fact that $TP_t f_p = e^{-\Lambda(p)t} f_p$ and that TP_t well approximates $P_t^{(2)}$ near the diagonal, while the second term is an error coming from the linearization approximation.

Unlike (2.6), we are able to show that h_p is in the domain of $\mathcal{L}_{(2)}$ and rigorously justify (2.7) (see Lemma 6.13). In order to make this perturbation argument, we crucially need that $\psi_p \in \dot{C}_V^1$, $V = V_{\beta, \eta}$.

For $\mathcal{V}$, we effectively deduce

$$\mathcal{L}_{(2)}\mathcal{V} \leq -\Lambda(p)h_p + C'V_{\beta+1, \eta} + \mathcal{L}V_{\beta+1, \eta},$$

where $\mathcal{L}$ denotes the (formal) infinitesimal generator of the (u_t) process (see Definition 3.4). From here, we will absorb the $C'V_{\beta+1, \eta}$ linearization error into $\mathcal{L}V_{\beta+1, \eta}$ by showing a “super drift condition”: formally, we can essentially view it as $\forall \kappa > 0, \exists C_\kappa > 0$ such that

$$\mathcal{L}V_{\beta+1, \eta} \leq -\kappa V_{\beta+1, \eta} + C_\kappa.$$

In other words, the drift condition for V as in Condition 4 can be taken with as strong an exponential decay rate as desired. This is much stronger than a standard drift condition; see (6.12) for details (also Remark 3.9). Taking $\kappa > 0$ sufficiently large absorbs the $C'V_{\beta+1,\eta}$ error term, resulting in (2.6).

The above sketch is far from complete and there are many details to fill in make it rigorous, including an appropriate C_0 semigroup framework for $P_t^{(2)}$ and issues relating to the domain of $\mathcal{L}_{(2)}$. Much of the difficulty in carrying out this argument rigorously stems from infinite dimensionality of $\mathbf{H} \times \mathcal{D}^c$; see Section 6 for a detailed discussion. Nevertheless, we ultimately verify the following.

PROPOSITION 2.14. *The Lyapunov function $\mathcal{V}$ satisfies Condition 4, that is, there exists $K > 0$ such that*

$$(2.8) \quad P_t^{(2)}\mathcal{V} \leq e^{-\Lambda(p)t}\mathcal{V} + K.$$

With Condition 4 in place, Theorem 2.2 now follows.

2.4.4. *Outline of the proof of Proposition 2.11.* In our analysis, it is convenient to transform the eigenproblem $T P_t f_p = e^{-\Lambda t} f_p$, with f_p as in (2.4), into the equivalent problem $\hat{P}_t^p \psi_p = e^{-\Lambda t} \psi_p$, where $\hat{P}_t^p$ denotes the following “twisted” Markov semigroup acting on observables $\psi : \mathbf{H} \times P\mathbb{T}^d \rightarrow \mathbb{R}$:

$$(2.9) \quad \hat{P}_t^p \psi(u, x, v) = \mathbf{E}_{(u,x,v)}(|D_x \phi^t v|^{-p} \psi(u_t, x_t, v_t)).$$

Above, we write (u_t, x_t, v_t) for the *projective process* on $\mathbf{H} \times P\mathbb{T}^d$, where $v_t \in P^{d-1}$ denotes the projective class of the vector $w_t^* = D_x \phi^t w$. We will refer to the Markov semigroup associated with the projective process as $\hat{P}_t = \hat{P}_t^0$. The projective process is natural here and also plays a major role in the proof of the positive Lyapunov exponent in [15].

Ultimately, we seek to define ψ_p to be an eigenfunction corresponding to the dominant eigenvalue of the semigroup $\hat{P}_t^p$ in some function space. Since we require $\psi_p \in \mathring{C}_V^1$, it is natural to consider spectral theory for the semigroup $\hat{P}_t^p$ on $\mathring{C}_V^1$. To work in this framework, however, entails significant technical problems. To start, it is already a challenge to prove that $\hat{P}_t$, let alone $\hat{P}_t^p$, is bounded on $C_V^1 \rightarrow C_V^1$ for any value of $t > 0$. Moreover, we are unable to show C_0 continuity for $\hat{P}_t^p$ in $\mathring{C}_V^1$. We note that this problem does not arise when working on the lower regularity space $\mathbf{W}$ as in [43].

We will show that $\hat{P}_{T_0}^p : C_V^1 \rightarrow C_V^1$ is bounded for a sufficiently large T_0 , permitting us to define ψ_p to be an eigenfunction of the discrete-time operator $\hat{P}_{T_0}^p$. Then we show that $\hat{P}_t^p$ defines a C_0 semigroup on $\mathring{C}_V$ and the relation $\hat{P}_t^p \psi_p = e^{-\Lambda(p)t} \psi_p$, $t \geq 0$ as in Proposition 2.11(b) is proved via semigroup theory.

Let us now make this more precise.

DEFINITION 2.15 (Spectral gap). Let A be a bounded linear operator on a Banach space $\mathfrak{B}$ with simple leading eigenvalue r . We say A has a *spectral gap* if there exists an $\epsilon > 0$ such that

$$\sigma(A) \setminus \{r\} \subseteq B_{|r|-\epsilon}(0).$$

For T_0 sufficiently large and all p sufficiently small, we will show that $\hat{P}_{T_0}^p$ is bounded $C_V^1 \rightarrow C_V^1$ and maps $\mathring{C}_V^1$ into itself (Lemmas 5.2 and 5.3). We then seek to prove that $\hat{P}_{T_0}^p$ has a spectral gap in $\mathring{C}_V^1$, and then construct ψ_p as an eigenfunction of $\hat{P}_{T_0}^p$ corresponding to its

leading eigenvalue (see (2.10) below). This is deduced from a spectral gap for the “untwisted” operator $\hat{P}_{T_0}$ in $\dot{C}_V^1$ and a spectral perturbation argument (see Lemma 5.1) using the fact that $\hat{P}_{T_0}^p \rightarrow \hat{P}_{T_0}$ in the C_V^1 norm as $p \rightarrow 0$ (Lemma 5.2).

In turn, a spectral gap for $\hat{P}_{T_0}$ in C_V^1 is deduced from a combination of a spectral gap for $\hat{P}_{T_0}$ in C_V and a Lasota–Yorke-type gradient estimate (Proposition 4.6) analogous to those used to prove asymptotic strong Feller in, for example, [42, 44]. The C_V spectral gap is deduced from Theorem 2.4 and essentially follows from ingredients already needed elsewhere (see Section 4.1). The gradient estimate is proved via an adaptation of the the Malliavin calculus scheme in [44] adapted to our more complicated nonlinearity, $\mathbf{H}$ regularity and the $\mathbf{H} \times P\mathbb{T}^d$ geometry. A brief sketch is provided in Section 4. In all, these arguments imply the following.

PROPOSITION 2.16. *Let $V = V_{\beta, \eta}$ be as in Lemma 2.9, where $\eta \in (0, \eta^*)$ is arbitrary and β is sufficiently large. Then there exist $T_0 > 0$, $p_0 > 0$ such that for all $p \in [-p_0, p_0]$, $\hat{P}_{T_0}^p$ has a spectral gap on $\dot{C}_V^1$ with real leading eigenvalue $e^{-T_0\Lambda(p)}$ for some $\Lambda(p) \in \mathbb{R}$.*

Moreover, the limit

$$(2.10) \quad \psi_p := \lim_{n \rightarrow \infty} e^{nT_0\Lambda(p)} \hat{P}_{nT_0}^p \mathbf{1}$$

exists in C_V^1 and defines an eigenfunction corresponding to the leading eigenvalue.

As observed earlier, to prove that ψ_p is an eigenfunction of $\hat{P}_t^p$ for all $t \geq 0$ as in Proposition 2.11(b), we cannot work on $\dot{C}_V^1$. Instead we will regard $\hat{P}_t^p$ as a semigroup on $\dot{C}_V$: here, $\hat{P}_t^p$ is bounded for all $t \geq 0$ and gives rise to a C_0 semigroup on $\dot{C}_V$ (Proposition 5.5). Carrying out an analogous spectral perturbation argument to Proposition 2.16 and using spectral theory for C_0 semigroups gives the following.

PROPOSITION 2.17. *Let V , p_0 , $\Lambda(p)$ and ψ_p be as in Proposition 2.16. For all $p \in [-p_0, p_0]$, $\hat{P}_t^p$ has a spectral gap on $\dot{C}_V$ for all $t > 0$ with leading eigenvalue $e^{-\Lambda(p)t}$, where $\Lambda(p)$ is as in Proposition 2.16. Moreover, ψ_p is an eigenfunction for $\hat{P}_t^p$ corresponding to this eigenvalue, and*

$$(2.11) \quad \hat{P}_t^p \psi_p = e^{-\Lambda(p)t} \psi_p.$$

At this point, we have the information required to prove Proposition 2.11. The formula (2.10) implies that ψ_p is nonnegative; positivity of $\hat{P}_{T_0}^p$ and irreducibility of (u_t, x_t, v_t) (Proposition 4.2) imply that $\psi_p > 0$ pointwise. The uniform positivity condition in Proposition 2.11 (a) follows from a compactness argument; see Lemma 5.7.

To prove the asymptotic $\Lambda(p) = \lambda_1 p + o(p)$ as in Proposition 2.11(c): first, we check that $p \mapsto \Lambda(p)$ is differentiable for p sufficiently small (Lemma 5.9) by combining (2.11) with Fréchet differentiability of $p \mapsto \hat{P}_t^p$ in the C_V operator norm. The formula $\Lambda'(0) = \lambda_1$ is then checked using a convexity argument and the characterization given in Remark 2.12; see Lemma 5.10 for details.

REMARK 2.18. Note that the twisted Markov semigroup in (2.9) can be written as

$$\hat{P}_t^p \psi(u, x, v) = \mathbf{E}_{(u, x, v)} \left(\exp \left(-p \int_0^t H(u_s, x_s, v_s) \, ds \right) \psi(u_t, x_t, v_t) \right),$$

where

$$H(u, x, v) := \langle v, Du(x)v \rangle,$$

so that $\hat{P}_t^p$ is the Feynman–Kac semigroup with potential H (see, e.g., [70]). This particular H and the resulting moment Lyapunov exponent $\Lambda(p)$ are useful in studying large deviations in the convergence of Lyapunov exponents for random dynamics; see, for example, [5–7].

REMARK 2.19. For the finite-dimensional models defined in Section 1.3, the construction of ψ_p and the corresponding Lyapunov function $\mathcal{V}$ can be simplified (though the above proof certainly suffices). Naturally, to treat System 5, everywhere there is (u_t) one must instead use (u_t, Z_t) . The conditions in Theorem 2.4 for geometric ergodicity for the projective process (u_t, x_t, v_t) (or (u_t, Z_t, x_t, v_t)) follow more-or-less immediately from hypoellipticity (see [15]) and the existence of Lyapunov functions for the (u_t) or (u_t, Z_t) processes (Remark 2.10). Moreover, $\hat{P}_t$ defines a C_0 semigroup on C_V . The spectral picture for the twisted Markov kernels $\hat{P}_t^p$ follows as discussed above (though simpler to check).

This provides, $0 < |p| \ll 1$, the ψ_p as in Proposition 2.11 (in the OU tower case, it also depends on Z). It is not necessary to study a spectral gap in C_V^1 , as ψ_p is C^∞ by Hörmander’s theorem and $D\psi_p \in C_V$ follows from a more refined use of hypoelliptic regularity using the scale of available V . The construction of $\mathcal{V}$ proceeds as above, as does the generator argument for checking the drift condition (2.8) (though these are much easier to justify). This completes the proof of Theorem 2.2 for these models.

REMARK 2.20. There are two fundamental places in our proof where we depend on the nondegenerate noise in Assumption 1. The first is in the Furstenberg criterion [Theorem 4.7, [15]] for the positive Lyapunov exponent (2.3). The Lasota–Yorke-type gradient bounds used in the asymptotically strong Feller frameworks of [42, 44] (and Section 4 below) do not seem to be the correct tool for obtaining this criterion. We are currently unsure what suitable hypoelliptic smoothing can be applied aside from strong Feller. The other place is in this work. The twisted Markov semigroup $\hat{P}_t^p$ does not seem to define a bounded semigroup on C_V^1 as currently required by the weak Harris theorem frameworks of [43, 45] due to the higher regularity needed for our methods. This forces a “strong” Harris theorem-type framework as in Theorem 2.4, which requires nondegenerate noise such as Assumption 1 (at least at high frequencies). *All of the other* uses of the strong Feller property/nonspatially smooth noise in this work and our previous [15] are not of fundamental importance and could be easily eliminated with some well-understood methods at the expense of additional technical complexity (e.g., those discussed in [38, 43] and the references therein). In these places, we use the nondegenerate noise only to reduce the length and complexity of the present manuscript.

2.5. *Notation.* We use the notation $f \lesssim g$ if there exists a constant $C > 0$ such that $f \leq Cg$ where C is independent of the parameters of interest. Sometimes we use the notation $f \approx_{a,b,c,\dots} g$ to emphasize the dependence of the implicit constant on the parameters, for example, $C = C(a, b, c, \dots)$. We denote $f \approx g$ if $f \lesssim g$ and $g \lesssim f$. Throughout, $\mathbb{R}^d$ is endowed with the standard Euclidean inner product $(\cdot, \cdot)$ and corresponding norm $|\cdot|$. We continue to write $|\cdot|$ for the corresponding matrix norm. When the domain of the L^p space is omitted, it is understood to be $\mathbb{T}^d$: $\|f\|_{L^p} = \|f\|_{L^p(\mathbb{T}^d)}$. We use the notation $\mathbf{E}X = \int_{\Omega} X(\omega) \mathbf{P}(d\omega)$ and $\|X\|_{L^p(\Omega)} = (\mathbf{E}|X|^p)^{1/p}$. When (z_t) is a Markov process, we write $\mathbf{E}_z, \mathbf{P}_z$ for the expectation and probability, respectively, conditioned on the event $z_0 = z$. We use the notation $\|f\|_{H^s} = \sum_{k \in \mathbb{Z}^d} |k|^{2s} |\hat{f}(k)|^2$ (denoting $\hat{f}(k) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{T}^d} e^{-ik \cdot x} f(x) dx$ the usual complex Fourier transform). We occasionally use Fourier multiplier notation $\widehat{m(\nabla)} f(\xi) := m(i\xi) \hat{f}(\xi)$. Additionally, we will often use r_0 to denote a number in $(\frac{d}{2} + 1, 3)$ such that the Sobolev embedding $H^{r_0} \hookrightarrow W^{1,\infty}$ holds.

We denote $P\mathbb{T}^d \cong \mathbb{T}^d \times P^{d-1}$ for projective bundle. We are often working with the Hilbert spaces $\mathbf{L}^2 \times T_v P\mathbb{T}^d$ and $\mathbf{H} \times T_v P\mathbb{T}^d$. For these spaces, we denote the inner product $\langle \cdot, \cdot \rangle_{\mathbf{W}}$ (resp., $\mathbf{H}$) and correspondingly for the norms as the finite-dimensional contribution to the inner product is unambiguous. For linear operators $A : \mathbf{L}^2 \times T_v P\mathbb{T}^d \rightarrow \mathbf{L}^2 \times T_v P\mathbb{T}^d$, we similarly denote the operator norm $\|A\|_{\mathbf{L}^2}$ and for linear operators $A : \mathbf{L}^2 \times T_v P\mathbb{T}^d \rightarrow \mathbb{R}$ we use the notation $\|A\|_{\mathbf{L}^{2*}}$ (analogously for $\mathbf{H}$). For $\mathcal{K} \subset \mathbb{K}$, define $\Pi_{\mathcal{K}} : \mathbf{L}^2 \times P\mathbb{T}^d \rightarrow \mathcal{K} \times P\mathbb{T}^d$ to be the orthogonal projection onto the subset of modes in $\mathcal{K}$. For $n \in \mathbb{N}$, Π_n denotes the orthogonal projection onto the modes with $k \in \mathbb{K}$, $|k| \leq n$.

3. Preliminaries, drift conditions and Jacobian estimates.

3.1. *Preliminaries.* We will write the Navier–Stokes system as an abstract evolution equation on $\mathbf{H}$ by

$$(3.1) \quad \partial_t u + B(u, u) + Au = Q\dot{W} = \sum_{m \in \mathbb{K}} q_m e_m \dot{W}^m,$$

where

$$B(u, v) = (\text{Id} - \nabla(-\Delta)^{-1}\nabla \cdot)(u \otimes v),$$

$$Au = \begin{cases} -\nu \Delta u & \text{if } d = 2, \\ -\nu' \Delta u + \nu \Delta^2 u & \text{if } d = 3. \end{cases}$$

The (u_t) process with initial data u is defined as the solution to (3.1) in the mild sense [25, 50]:

$$(3.2) \quad u_t = e^{-tA}u - \int_0^t e^{-(t-s)A}B(u_s, u_s)ds + \int_0^t e^{-(t-s)A}QdW(s),$$

where the above identity holds $\mathbf{P}$ almost surely for all $t > 0$. For (3.2), we have the following well-posedness theorem.

PROPOSITION 3.1 ([25, 50]). *For each of Systems 1–2, we have the following. For all initial $u \in \mathbf{H} \cap \mathbf{H}^\gamma$ with $\gamma < \alpha - \frac{d}{2}$ and all $T > 0$, $p \geq 1$, there exists a $\mathbf{P}$ almost surely unique solution (u_t) to (3.2). Moreover, (u_t) is $\mathcal{F}_t$ -adapted, and belongs to $L^p(\Omega; C([0, T]; \mathbf{H} \cap \mathbf{H}^\gamma)) \cap L^2(\Omega; L^2(0, T; \mathbf{H}^{\gamma+(d-1)}))$.*

Additionally, for all $p \geq 1$ and $0 \leq \gamma < \gamma' < \alpha - \frac{d}{2}$,

$$\mathbf{E} \sup_{t \in [0, T]} \|u_t\|_{\mathbf{H}^{\gamma'}}^p \lesssim_{T, p, \gamma} 1 + \|u_0\|_{\mathbf{H} \cap \mathbf{H}^\gamma}^p,$$

$$\mathbf{E} \int_0^T \|u_s\|_{\mathbf{H}^{\gamma+(d-1)}}^2 ds \lesssim_{T, \delta} 1 + \|u_0\|_{\mathbf{H}^\gamma}^2,$$

$$\mathbf{E} \sup_{t \in [0, T]} (t^{\frac{\gamma' - \gamma}{2(d-1)}} \|u_t\|_{\mathbf{H}^{\gamma'}})^p \lesssim_{p, T, \gamma, \gamma'} 1 + \|u_0\|_{\mathbf{H}^\gamma}^p.$$

Before proceeding, we need the following useful Sobolev interpolation inequalities.

LEMMA 3.2. *For all $r \in (0, 3)$, $\forall \sigma > 2$ and $\forall \epsilon > 0$, $\exists C = C(r, \epsilon, \sigma) > 0$ such that $\forall u \in \mathbf{H}^{\sigma+1}$,*

$$(3.3) \quad \|u\|_{\mathbf{H}^r} \leq \epsilon \|\Delta u\|_{L^2}^2 + \epsilon \frac{\|\nabla u\|_{\mathbf{H}^\sigma}^2}{1 + \|u\|_{\mathbf{H}^\sigma}^2} + C.$$

PROOF. Without loss of generality, we can assume $r > 2$ as otherwise the estimate is immediate. Similarly, we assume that $\|u\|_{\mathbf{H}^r} > \epsilon \|\Delta u\|_{L^2}^2$, as otherwise the estimate is automatic. By Sobolev interpolation,

$$\|u\|_{\mathbf{H}^r} \lesssim \|\Delta u\|_{L^2}^{1-\theta} \|\nabla u\|_{\mathbf{H}^\sigma}^\theta \lesssim \|u\|_{\mathbf{H}^r}^{\frac{1-\theta}{2}} \|\nabla u\|_{\mathbf{H}^\sigma}^\theta,$$

where $\theta = \frac{r-2}{\sigma-1}$. Therefore,

$$(3.4) \quad \|u\|_{\mathbf{H}^r} \lesssim \|\nabla u\|_{\mathbf{H}^\sigma}^{\frac{2\theta}{1+\theta}}.$$

A similar argument also implies that

$$(3.5) \quad \|u\|_{\mathbf{H}^\sigma} \lesssim \|u\|_{\mathbf{H}^r}^{\frac{1-\gamma}{2}} \|\nabla u\|_{\mathbf{H}^\sigma}^\gamma \lesssim \|\nabla u\|_{\mathbf{H}^\sigma}^{\frac{\gamma+\theta}{1+\theta}},$$

where $\gamma = \frac{\sigma-2}{\sigma-1}$. It follows from (3.4) and (3.5) that

$$\frac{\|\nabla u\|_{\mathbf{H}^\sigma}^2}{1 + \|u\|_{\mathbf{H}^\sigma}^2} \gtrsim \|\nabla u\|_{\mathbf{H}^\sigma}^{\frac{2(1-\gamma)}{1+\theta}} \gtrsim \|u\|_{\mathbf{H}^r}^{\frac{1-\gamma}{\theta}}.$$

Inequality (3.3) follows from the fact that for all $\epsilon > 0$ there exists a $C = C(\epsilon, \sigma, r)$ such that

$$\|u\|_{\mathbf{H}^r} \leq \epsilon \|u\|_{\mathbf{H}^\sigma}^{\frac{1-\gamma}{\theta}} + C$$

provided $\frac{1-\gamma}{\theta} = \frac{1}{r-2} > 1$, which holds if and only if $r < 3$. $\square$

LEMMA 3.3. *For all $\epsilon > 0$, $\sigma \geq 1$, there exists a $C_\epsilon > 0$ such that the following holds for all $u \in \mathbf{H}^{\sigma+1}$:*

$$\log(1 + \|u\|_{\mathbf{H}^\sigma}^2) \leq \epsilon \|\Delta u\|_{L^2}^2 + \epsilon \frac{\|\nabla u\|_{\mathbf{H}^\sigma}^2}{1 + \|u\|_{\mathbf{H}^\sigma}^2} + C_\epsilon.$$

PROOF. For all $\epsilon, \delta > 0$, $\exists C_{\delta, \epsilon} > 0$ such that the following holds $\forall x \geq 0$:

$$\log(1 + x^2) \leq C_{\delta, \epsilon} + \epsilon x^\delta.$$

Similar to Lemma 3.2, we may assume $\sigma > 1$ and

$$(3.6) \quad \|u\|_{\mathbf{H}^\sigma}^\delta > \epsilon \|\Delta u\|_{L^2}^2.$$

By Sobolev interpolation (for suitable $\theta \in (0, 1)$),

$$\|\Delta u\|_{L^2} \leq \|u\|_{\mathbf{H}^\sigma} \lesssim \|\Delta u\|_{L^2}^{1-\theta} \|\nabla u\|_{\mathbf{H}^\sigma}^\theta,$$

and hence by (3.6), there holds

$$\|u\|_{\mathbf{H}^\sigma}^{1-\frac{1}{2}\delta(1-\theta)} \lesssim \|\nabla u\|_{\mathbf{H}^\sigma}^\theta,$$

which implies

$$\frac{\|\nabla u\|_{\mathbf{H}^\sigma}^2}{1 + \|u\|_{\mathbf{H}^\sigma}^2} \gtrsim \|\nabla u\|_{\mathbf{H}^\sigma}^{2-\frac{4\theta}{2-\delta(1-\theta)}}.$$

The desired estimate follows taking δ sufficiently small. $\square$

3.2. *Fundamental estimates and super-Lyapunov property.* Let $C^2(\mathbf{H})$ be the space of twice Frechet differentiable functions on $\mathbf{H}$. We denote by $D_u G(u)$ the Frechet derivative at u , which is a linear functional on $\mathbf{H}$ defined for each $v \in \mathbf{H}$ by

$$D_u G(u)v = \lim_{h \rightarrow 0} h^{-1} (G(u + hv) - G(u)).$$

Likewise, the second Frechet derivative $D_u^2 G(u)$ can be identified is a bilinear form on $\mathbf{H}$ defined for each $v, w \in \mathbf{H}$ by

$$D_u^2 G(u)[v, w] = \lim_{h \rightarrow 0} h^{-1} (D_u G(u + hw)v - D_u G(u)v).$$

We will define the formal generator of the process (3.1) as follows.

DEFINITION 3.4. Suppose $G \in C^2(\mathbf{H})$ has the property that for each $u \in \mathbf{H}^{\sigma+1}$, and $v \in \mathbf{H}^{\sigma-1}$, the mapping $u \mapsto D_u G(u)v$ is continuous on $\mathbf{H}^{\sigma+1}$ and satisfies

$$|D_u G(u)v| \leq K(\|u\|_{\mathbf{H}})\|u\|_{\mathbf{H}^{\sigma+1}}\|v\|_{\mathbf{H}^{\sigma-1}}$$

for some continuous function $K(r)$, $r > 0$. We define the formal generator $\mathcal{L}$ of the process (3.1) to be the linear operator acting on G such that for each $u \in \mathbf{H}^{\sigma+1}$,

$$(3.7) \quad \mathcal{L}G(u) := -D_u G(u)(B(u, u) + Au) + \frac{1}{2} \sum_{m \in \mathbb{K}} |q_m|^2 D_u^2 G(u)[e_m, e_m].$$

REMARK 3.5. One cannot identify $\mathcal{L}$ as the infinitesimal generator of the Markov semi-group associated to the (3.1), but they coincide on a core of smooth cylinder functions (see [43] Lemma 5.11). Note that one can also apply $\mathcal{L}$ to functions that do not necessarily belong to the domain of the generator (see [43] Remark 5.7 for a discussion of this).

The next lemma provides the energy estimate that eventually implies the drift condition on $V_{\beta, \eta}$.

LEMMA 3.6. For all $V = V_{\beta, \eta}$ with $\beta \geq 2$ and $\eta > 0$, there exists a $C = C(\beta, \eta) > 0$ such that the following estimates hold for all $u \in \mathbf{H}^{\sigma+1}$:

$$(3.8) \quad \mathcal{L} \log V(u) \leq -v\eta \|\Delta u\|_{L^2}^2 - v\beta \frac{\|\nabla u\|_{\mathbf{H}}^2}{1 + \|u\|_{\mathbf{H}}^2} + C,$$

PROOF. First, consider (3.8). For simplicity, consider the case $d = 2$; the case $d = 3$ follows similarly with only trivial modification. By direct calculation, we find

$$D_u \log V(u)v = 2\eta \langle u, v \rangle_{\mathbf{W}} + \frac{2\beta \langle u, v \rangle_{\mathbf{H}}}{1 + \|u\|_{\mathbf{H}}^2},$$

and

$$D_u^2 \log V(u)[v, v] = 2\eta \|v\|_{\mathbf{W}}^2 + \frac{2\beta \|v\|_{\mathbf{H}}^2}{1 + \|u\|_{\mathbf{H}}^2} - \frac{4\beta |\langle u, v \rangle_{\mathbf{H}}|^2}{(1 + \|u\|_{\mathbf{H}}^2)^2},$$

where the $\langle \cdot, \cdot \rangle_{\mathbf{H}}$ inner product is to be interpreted as the natural pairing between $\mathbf{H}^{\sigma+1}$ and $\mathbf{H}^{\sigma-1}$. Consequently, we have

$$\begin{aligned} \mathcal{L} \log V \leq & \underbrace{-2\eta \langle u, B(u, u) + Au \rangle_{\mathbf{W}}}_{T_\omega} + \underbrace{-\frac{2\beta \langle u, B(u, u) + Au \rangle_{\mathbf{H}}}{1 + \|u\|_{\mathbf{H}}^2}}_{T_\sigma} \\ & + \underbrace{\sum_{k \in \mathbb{Z}_0^d} |q_k|^2 \left(\eta \|e_k\|_{\mathbf{W}}^2 + \frac{\beta \|e_k\|_{\mathbf{H}}^2}{1 + \|u\|_{\mathbf{H}}^2} \right)}_{T_Q}. \end{aligned}$$

By the divergence-free structure and standard $\mathbf{H}^\sigma$ estimates, $\exists C_\sigma > 0$ such that (see, e.g., [55]) the following holds for some $r \in (\frac{d}{2} + 1, 3)$,

$$\begin{aligned} T_\omega &= -2\nu\eta \|\Delta u\|_{L^2}^2, \\ T_\sigma &\leq -2\nu\beta \frac{\|\nabla u\|_{\mathbf{H}}^2}{1 + \|u\|_{\mathbf{H}}^2} + C_\sigma \|u\|_{\mathbf{H}^r}. \end{aligned}$$

The term involving $\|u\|_{\mathbf{H}^r}$ is then estimated by applying Lemma 3.2. Similarly, we have by $\|e_k\|_{\mathbf{W}} \leq |k|$, $\|e_k\|_{\mathbf{H}} = |k|^\sigma$ and $|q_k| \lesssim |k|^{-\alpha}$ and $\alpha - \sigma > d/2$ that

$$|T_Q| \leq \sum_{k \in \mathbb{Z}_0^d} |q_k|^2 (\eta \|e_k\|_{\mathbf{W}}^2 + \beta \|e_k\|_{\mathbf{H}}^2) \lesssim_{\beta, \eta} \sum_{k \in \mathbb{Z}_0^d} |k|^{-2(\alpha - \sigma)} \lesssim_{\beta, \eta} 1. \quad \square$$

Lemma 3.6 allows us to prove the following bound on (u_t) a solution to either of Systems 1 or 2. The following lemma is fundamental to our analysis and is used repeatedly in what follows. It controls not only an improvement in the moments of $V(u_t)$ but also controls the inclusion of exponential factors of time-integrations of $\|u_s\|_{\mathbf{H}^r}$ for $r \in (0, 3)$, which arise very naturally in our analysis.

LEMMA 3.7. *Let (u_t) solve either Systems 1 or 2. There exists a $\gamma_* > 0$, such that for all $0 \leq \gamma < \gamma_*$, $T \geq 0$, $r \in (0, 3)$, $\kappa \geq 0$ and $V(u) = V_{\beta, \eta}$ where $\beta \geq 0$ and $0 < e^{\gamma T} \eta < \eta_*$, there exists a constant $C = C(\gamma, T, r, \kappa, \beta, \eta) > 0$ such that the following estimate holds:*

$$(3.9) \quad \mathbf{E}_u \exp\left(\kappa \int_0^T \|u_s\|_{\mathbf{H}^r} ds\right) \sup_{0 \leq t \leq T} V^{e^{\gamma t}}(u_t) \leq C V(u).$$

REMARK 3.8. In fact, it suffices to take $\gamma_* = \frac{\nu}{8}$.

PROOF. By Itô's lemma applied to the functional $\log V(u)$ (see [50] Theorem 7.7.5), we know that

$$M_t := e^{\gamma t} \log V(u_t) - \log V(u) - \int_0^t e^{\gamma s} (\gamma \log V(u_s) + \mathcal{L} \log V(u_s)) ds$$

is a mean-zero, time-continuous, local martingale with quadratic variation satisfying

$$\langle M \rangle_t = \sum_{m \in \mathbb{K}} \int_0^t e^{2\gamma s} |q_k|^2 |D_u \log V(u_s) e_m|^2 ds \leq 8 \int_0^t e^{2\gamma s} (\beta^2 + \eta^2 \mathcal{Q} \|\Delta u_s\|_{L^2}^2) ds,$$

where $\mathcal{Q}$ is defined in Lemma 2.9. Recall the exponential martingale estimate

$$(3.10) \quad \mathbf{E}_u \exp\left(\sup_{t \geq 0} (M_t - \langle M \rangle_t)\right) \leq 2.$$

In light of this, note that

$$M_t - \langle M \rangle_t \geq e^{\gamma t} \log V(u_t) - \log V(u) - \int_0^t e^{\gamma s} (\gamma \log V(u_s) + \mathcal{L} \log V(u_s) + 8e^{\gamma s} (\beta^2 + \eta^2 \mathcal{Q} \|\Delta u_s\|_{L^2}^2)) ds.$$

Applying Lemma 3.3 to $\gamma \log V$ and Lemma 3.6 to $\mathcal{L} \log V$, collecting like terms, and bounding $\|u\|_{\mathbf{W}}$ by $\|\Delta u\|_{L^2}$ we find that for each $\epsilon > 0$,

$$(3.11) \quad \begin{aligned} & \gamma \log V(u_s) + \mathcal{L} \log V(u_s) + 8e^{\gamma s} (\beta^2 + \eta^2 \mathcal{Q} \|\Delta u_s\|_{L^2}^2) \\ & \leq -\eta(v - \gamma - 8e^{\gamma s} \eta \mathcal{Q}) \|\Delta u_s\|_{L^2}^2 - (v\beta - \epsilon) \frac{\|\nabla u_s\|_{\mathbf{H}^\sigma}^2}{1 + \|u_s\|_{\mathbf{H}^\sigma}^2} + Ce^{\gamma s}. \end{aligned}$$

Using that $e^{\gamma T} \eta \leq \eta_*$, we can choose γ and ϵ small enough so that

$$\gamma + 8\eta_* \mathcal{Q} \leq \frac{1}{2}v \quad \text{and} \quad \epsilon < \frac{1}{2}v\beta.$$

Therefore, we can apply the interpolation Lemma 3.2 to deduce that for each $\kappa > 0$ the right-hand side of inequality (3.11) is bounded by $-\kappa \|u_s\|_{\mathbf{H}^r} + Ce^{\gamma s}$ and, therefore,

$$M_t - \langle M \rangle_t \geq e^{\gamma t} \log V(u_t) - \log V(u) + \kappa \int_0^t e^{\gamma s} \|u_s\|_{\mathbf{H}^r} ds - C$$

Applying (3.10) and restricting the supremum to $0 \leq t \leq T$ gives the estimate. $\square$

REMARK 3.9. Note that Lemma 3.7 is strictly stronger than a drift condition. The improvement in the power of V is sometimes called a *super-Lyapunov property* and it provides an important strengthening of the notion of a drift condition. To see that (3.9) implies a drift condition, we write $P_1 \varphi(u) = \mathbf{E}_u \varphi(u_1)$ as the Markov semigroup for Navier–Stokes, set $\kappa = 0$ in (3.9), and then Jensen’s inequality implies that for some $C_L > 0$, there holds

$$P_1 V \leq (e^{C_L} V)^{e^{-\gamma}},$$

which immediately implies that for all $\delta > 0$, there exists $C_\delta > 0$ such that

$$P_1 V \leq \delta V + C_\delta.$$

Furthermore, the bound (3.9) can be iterated with repeated applications of Jensen’s inequality (cf. [Proposition 5.11, [44]]) to produce

$$(3.12) \quad P_n V \leq e^{C_L \frac{e^{-\gamma}}{1-e^{-\gamma n}}} V e^{-\gamma n}.$$

The inequality (3.12) gives a strong quantification of the tendency to return back to uniform sublevel sets of V irrespective of the initial distribution as $n \rightarrow \infty$ (i.e., the initial data is forgotten exponentially fast).

The following is a useful consequence of Lemma 3.7.

LEMMA 3.10. *For all $p \in [1, \infty)$, $\eta \in (0, \eta_*)$, there exists a (deterministic) $C_0 > 0$ and a random constant $D_0 : \mathbf{H} \times \Omega \rightarrow \mathbb{R}_{\geq 1}$, satisfying $\mathbf{E}(D_0(\cdot, u))^p \lesssim \exp(\eta \|u\|_{\mathbf{W}}^2)$, such that*

$$\exp\left(\int_0^t \|\nabla u_\tau\|_{L^\infty} d\tau\right) \leq D_0(\omega, u) e^{C_0 t}.$$

PROOF. We provide the proof when $p = 1$; other values of p require straightforward adjustments.

Set $V = V_{0,\eta}$, where $\eta \in (0, \eta_*)$ is arbitrary. To start, note that by Lemma 3.7 and Chebyshev's inequality, $\exists c = c(\eta) > 0$ such that

$$\mathbf{P}\left(\exp\left(2 \int_0^t \|\nabla u_s\|_{L^\infty} ds\right) > V(u)e^{4ct}\right) \leq \frac{\mathbf{E} \exp(2 \int_0^t \|\nabla u_s\|_{L^\infty} ds)}{V(u)e^{4ct}} \lesssim_\eta e^{-3ct}.$$

By Borel–Cantelli, there exists $N(\omega, u) \geq 1$ with $\mathbf{P}(N(\cdot, u) \geq n) \lesssim_\eta e^{-3cn}$ such that

$$\exp\left(\int_0^n \|\nabla u_s\|_{L^\infty} ds\right) \leq V(u)e^{4cn} \quad \text{for } n \geq N(\omega, u).$$

To bound when $n < N(\omega, u)$, we find

$$\exp\left(\int_0^n \|\nabla u_s\|_{L^\infty} ds\right) \leq D_0(\omega, u) := \exp\left(\int_0^{N(\omega, u)} \|\nabla u_s\|_{L^\infty} ds\right)$$

and note that by Cauchy–Schwarz,

$$\begin{aligned} \mathbf{E} D_0(\cdot, u) &\leq \sum_n (\mathbf{P}(N = n))^{1/2} \left[\mathbf{E} \exp\left(2 \int_0^n \|\nabla u_s\|_{L^\infty} ds\right) \right]^{1/2} \\ &\leq \sum_n (e^{-3cn})^{1/2} (V(u)e^{4cn})^{1/2} \leq V(u) \sum_n e^{-cn} \lesssim V(u). \end{aligned}$$

□

3.3. Estimates on the Jacobian of the projective process. In this section, we provide the necessary estimates on the Jacobian of the projective process. Recall the projective process $(\hat{z}_t) = (u_t, x_t, v_t)$ solves the abstract SDE in $\mathbf{H} \times P\mathbb{T}^d$,

$$\partial_t \hat{z}_t = F(\hat{z}_t) + Q \dot{W}_t,$$

where we view $Q \dot{W}$ as extended to $\mathbf{H} \times T_{v_t} P\mathbb{T}^d$ and for each $\hat{z} = (u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$,

$$F(\hat{z}) = \begin{pmatrix} -B(u, u) - Au \\ u(x) \\ (I - v \otimes v)(Du(x)v) \end{pmatrix}.$$

The Jacobian process $J_{s,t}$ denotes the Fréchet derivative of the solution $\hat{z}_t$ with respect to the value at time $s < t$. Hence, $J_{s,t}$ solves the operator-valued equation

$$(3.13) \quad \partial_t J_{s,t} = DF(\hat{z}_t) J_{s,t}, \quad J_{s,s} = \text{Id}.$$

We will prove that this is a bounded operator $J_{s,t} : \mathbf{W} \times T_{v_s} P\mathbb{T}^d \rightarrow \mathbf{W} \times T_{v_t} P\mathbb{T}^d$ (this is not obvious due to the evolution on $P\mathbb{T}^d$ requiring pointwise evaluations u and Du). Additionally, we let $K_{s,t} : \mathbf{W} \times T_{v_t} P\mathbb{T}^d \rightarrow \mathbf{W} \times T_{v_s} P\mathbb{T}^d$ denote the adjoint of $J_{s,t}$, in the sense that

$$\langle f, J_{s,t} \xi \rangle_{\mathbf{W}} = \langle K_{s,t} f, \xi \rangle_{\mathbf{W}}.$$

A straightforward calculation (see [44]) shows that $K_{s,t}$ solves the following backward-in-time equation:

$$\partial_s K_{s,t} = -DF(\hat{z}_s)^* K_{s,t}, \quad K_{t,t} = I,$$

where $DF(\hat{z}_s)^* : \mathbf{W} \times T_{v_s} P\mathbb{T}^d \rightarrow \mathbf{W} \times T_{v_s} P\mathbb{T}^d$ is the adjoint to $DF(\hat{z}_s)$.

In what follows, we will find it convenient to let $\tilde{z} = (\tilde{u}, \tilde{x}, \tilde{v}) \in \mathbf{W} \times T_{v_s} P\mathbb{T}^d$ an initial perturbation and denote

$$\tilde{z}_t := (\tilde{u}_t, \tilde{x}_t, \tilde{v}_t) = J_{s,t} \tilde{w} = (J_{s,t}^u \tilde{w}, J_{s,t}^x \tilde{w}, J_{s,t}^v \tilde{w}) \in \mathbf{W} \times T_{v_t} P\mathbb{T}^d,$$

which readily solves the linear evolution equation

$$\partial_t \tilde{z}_t = DF(\hat{z}_t) \tilde{z}_t, \quad \tilde{z}_s = \tilde{z}.$$

LEMMA 3.11. $\forall \sigma > \frac{d}{2} + 1, \forall r \in (\frac{d}{2} + 1, 3), \exists C, q' > 0$ such that the following holds pathwise:

$$(3.14a) \quad \|\tilde{u}_t\|_{\mathbf{W}} \leq \|\tilde{u}\|_{\mathbf{W}} \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right)$$

$$(3.14b) \quad \|J_{s,t}\|_{\mathbf{H}^\sigma \rightarrow \mathbf{H}^\sigma} \lesssim \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right) \left(1 + \langle t-s \rangle^3 \sup_{s < \tau < t} \|u_\tau\|_{\mathbf{H}^\sigma}^{q'}\right).$$

PROOF. We consider the case $d = 2$; $d = 3$ follows analogously. We start by estimating $\tilde{u}_t$, which solves

$$\partial_t \tilde{u}_t + B(u_t, \tilde{u}_t) + B(\tilde{u}_t, u_t) = A \tilde{u}_t.$$

Integration by parts, vector calculus and the divergence-free condition gives for any $p \in (2, \infty)$,

$$\begin{aligned} \frac{d}{dt} \|\tilde{u}_t\|_{\mathbf{W}}^2 + \nu \|\nabla \tilde{u}_t\|_{\mathbf{W}}^2 &= \langle \text{curl } \tilde{u}_t, \nabla \cdot (u_t \text{curl } \tilde{u}_t + \tilde{u}_t \text{curl } u_t) \rangle_{L^2} \\ &\lesssim \|\nabla \text{curl } u_t\|_{L^p} \|\tilde{u}_t\|_{\mathbf{W}} \|\tilde{u}_t\|_{L^{\frac{2p}{p-2}}}. \end{aligned}$$

Then Sobolev embedding implies that for $r = 2 + \frac{p-2}{2}$,

$$\frac{d}{dt} \|\tilde{u}_t\|_{\mathbf{W}}^2 \lesssim_p \|u_t\|_{H^r} \|\tilde{u}_t\|_{\mathbf{W}}^2,$$

which implies

$$\|\tilde{u}_t\|_{\mathbf{W}} \leq \|\tilde{u}\|_{\mathbf{W}} \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right).$$

This completes the proof of (3.14a) by choosing p sufficiently close to 2 (in the case $d = 2$). Note the case $d = 3$ is easier for (3.14a), as one can simply choose $p = \infty$ at the analogous step.

Next, we turn to the $\mathbf{H}^\sigma$ estimate on $\tilde{u}_t$ for the proof of (3.14b). By the divergence-free constraint,

$$(3.15) \quad \frac{d}{dt} \|\tilde{u}_t\|_{\mathbf{H}^\sigma}^2 + \nu \|\nabla \tilde{u}_t\|_{\mathbf{H}^\sigma}^2 = \underbrace{\langle \tilde{u}_t, \nabla \cdot (u_t \otimes \tilde{u}_t) \rangle_{\mathbf{H}^\sigma}}_{T1} + \underbrace{\langle \tilde{u}_t, \nabla \cdot (\tilde{u}_t \otimes u_t) \rangle_{\mathbf{H}^\sigma}}_{T2}.$$

Using the divergence-free property to introduce a commutator (recall the Fourier multiplier from Section 2.5), and the triangle inequality $\langle k \rangle^\sigma \lesssim_\sigma \langle k - \ell \rangle^\sigma + \langle \ell \rangle^\sigma$, there holds

$$\begin{aligned} (3.16) \quad T1 &= \langle \langle \nabla \rangle^\sigma \tilde{u}_t, \langle \nabla \rangle^\sigma \nabla \cdot (u_t \otimes \tilde{u}_t) - \nabla \cdot (u_t \otimes \langle \nabla \rangle^\sigma \tilde{u}_t) \rangle_{L^2} \\ &\leq \left(\sum_{|k-\ell| > |\ell|} + \sum_{|k-\ell| \leq |\ell|} \right) \langle k \rangle^\sigma |\hat{u}(k)| |\langle k \rangle^\sigma - \langle \ell \rangle^\sigma| |k| |\hat{u}_t(k-\ell)| |\hat{u}(\ell)| \\ &\lesssim \|\nabla \tilde{u}_t\|_{\mathbf{H}^\sigma} \|u_t\|_{\mathbf{H}^\sigma} \|\tilde{u}_t\|_{\mathbf{H}^{r-1}} + \|u_t\|_{\mathbf{H}^r} \|\tilde{u}_t\|_{\mathbf{H}^\sigma}^2, \end{aligned}$$

where in the penultimate line we used that in the second summation there holds $|\langle k \rangle^\sigma - \langle \ell \rangle^\sigma| \lesssim \langle \ell \rangle^{\sigma-1} |k - \ell|$ and in the last line we used Cauchy–Schwarz and Young’s convolution inequality. By a similar argument (but no commutator necessary), there holds

$$(3.17) \quad T2 \lesssim \|\nabla \tilde{u}_t\|_{\mathbf{H}^\sigma} \|u_t\|_{\mathbf{H}^\sigma} \|\tilde{u}_t\|_{\mathbf{H}^{r-1}} + \|u_t\|_{\mathbf{H}^r} \|\tilde{u}_t\|_{\mathbf{H}^\sigma}^2.$$

Note by interpolation, for some $\theta = \theta(\sigma) \in (0, 1)$, there holds

$$\|\tilde{u}_t\|_{\mathbf{H}^{r-1}} \leq \|\tilde{u}_t\|_{\mathbf{W}}^{1-\theta} \|\nabla \tilde{u}_t\|_{\mathbf{H}^\sigma}^\theta,$$

and, therefore, putting (3.16) and (3.17) together with (3.15) gives (for suitable $C > 0$),

$$\frac{d}{dt} \|\tilde{u}_t\|_{\mathbf{H}^\sigma}^2 \leq -\frac{\nu}{2} \|\nabla \tilde{u}_t\|_{\mathbf{H}^\sigma}^2 + C \|u_t\|_{\mathbf{H}^r} \|\tilde{u}_t\|_{\mathbf{H}^\sigma}^2 + C \|\tilde{u}_t\|_{\mathbf{W}}^2 \|u_t\|_{\mathbf{H}^\sigma}^{\frac{2}{1-\theta}}.$$

Integrating and (3.14a) completes the estimate on the velocity field necessary for (3.14b).

Next, consider the remaining contributions of the projective process necessary for (3.14b). Using that

$$\partial_t \tilde{x}_t = \nabla u_t(x_t) \tilde{x}_t + \tilde{u}_t(x_t)$$

and applying Grönwall's inequality followed by the estimate on the $\mathbf{H}^\sigma$ norm of the velocity field gives

$$\begin{aligned} |\tilde{x}_t| &\leq \exp\left(C \int_s^t \|\nabla u_{\tau'}\|_{L^\infty} d\tau'\right) |\tilde{x}| + \int_s^t \exp\left(C \int_\tau^t \|\nabla u_{\tau'}\|_{L^\infty} d\tau'\right) \|u_\tau\|_{L^\infty} d\tau \\ &\lesssim \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right) \left(1 + \langle t-s \rangle \sup_{s < \tau < t} \|u_\tau\|_{\mathbf{H}^\sigma}^{q'}\right) (|\tilde{x}| + \|\tilde{u}\|_{\mathbf{H}^\sigma}), \end{aligned}$$

which suffices for estimates $\tilde{x}_t$. Similarly, for the projective process we have

$$\begin{aligned} \partial_t \tilde{v}_t &= -\tilde{v}_t \otimes v_t \nabla u_t(x_t) v_t - v_t \otimes \tilde{v}_t \nabla u_t(x_t) v_t + (1 - v_t \otimes v_t) \nabla u_t(x_t) \tilde{v}_t \\ &\quad + (1 - v_t \otimes v_t) \tilde{x}_t \cdot \nabla^2 u_t(x_t) v_t + (1 - v_t \otimes v_t) \nabla \tilde{u}_t(x_t) v_t. \end{aligned}$$

Hence,

$$\frac{d}{dt} |\tilde{v}_t| \lesssim \|\nabla u\|_{L^\infty} |\tilde{v}_t| + |\tilde{x}_t| \|\nabla^2 u\|_{L^\infty} + \|\nabla \tilde{u}\|_{L^\infty},$$

and the required estimate on $\tilde{v}_t$ follows by Grönwall's inequality as in the case of $\tilde{x}_t$. $\square$

The pathwise estimates in Lemma 3.11 together with Lemma 3.7 implies the following.

LEMMA 3.12 (Jacobian bounds in expectation). *For all σ and all $\eta > 0$, there is a constant C_J such that the following holds for all $1 \leq p < \infty$:*

$$\sup_{s \leq t \leq 1} \mathbf{E} \|J_{s,t}\|_{\mathbf{H}^\sigma \rightarrow \mathbf{H}^\sigma}^p \leq V_{q',\eta}^p(u_s) \exp(pC_J).$$

Next, we deduce a parabolic smoothing estimate on the Jacobian. One small subtle point: pointwise evaluations of $\tilde{u}$ and $D\tilde{u}$ are appear in the ODEs for $\tilde{x}_t$ and $\tilde{v}_t$ (resp.), and hence a little care must be taken to control low regularity data for $\tilde{u}_t$ using the local-in-time parabolic smoothing.

LEMMA 3.13. *Let $\gamma \in [0, \alpha - \frac{d}{2})$ and $r \in (\frac{d}{2} + 1, 3)$. Then $\exists \varkappa'$ the following holds pathwise for $0 \leq s \leq t \leq 1$:*

$$(t-s)^{\frac{\gamma}{2(d-1)}} \|J_{s,t}\|_{\mathbf{W} \rightarrow \mathbf{H}^\gamma} \lesssim \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right) \left(1 + \sup_{\tau \in (s,t)} \|u_\tau\|_{\mathbf{H}^\sigma}^{\varkappa'}\right).$$

PROOF. We consider only the case $d = 2$; the case $d = 3$ is a straightforward variation. First, the desired estimate on $\tilde{u}_t$ follows from standard semilinear PDE methods (see, for instance, [44]) and is omitted for the sake of brevity.

Turn next to $\tilde{x}_t$ and $\tilde{v}_t$. Estimating the Lagrangian process as in Lemma 3.11 gives

$$|\tilde{x}_t| \lesssim \exp\left(\int_0^t \|\nabla u_\tau\|_{L^\infty} d\tau\right) |\tilde{x}| + \int_0^t \exp\left(\int_\tau^t \|\nabla u_s\|_{L^\infty} ds\right) \|\tilde{u}_\tau\|_{L^\infty} d\tau.$$

By Sobolev embedding and the smoothing estimates deduced on $\tilde{u}_t$, $\forall \delta > 0$,

$$\|\tilde{u}_\tau\|_{L^\infty} \lesssim \|\tilde{u}_\tau\|_{\mathbf{H}^{1+\delta}} \lesssim \frac{1}{\tau^{\delta/2}} \exp\left(C \int_0^\tau \|u_s\|_{\mathbf{H}^r} ds\right) \left(1 + \sup_{0 < s < \tau} \|u_s\|_{\mathbf{H}^\sigma}^{q'}\right) \|\tilde{u}\|_{\mathbf{W}},$$

which yields the desired estimate.

The situation for $\tilde{v}_t$ is similar but higher regularity is required. Observe that $\forall s, \delta > 0$,

$$\|\nabla \tilde{u}_\tau\|_{L^\infty} \lesssim_\delta \|\tilde{u}_\tau\|_{\mathbf{H}^{2+\delta}} \lesssim \tau^{-\frac{1+\delta}{2}} \exp\left(C \int_0^\tau \|u_s\|_{\mathbf{H}^r} ds\right) \left(1 + \sup_{0 < s < \tau} \|u_s\|_{\mathbf{H}^\sigma}^{q'}\right) \|\tilde{u}\|_{\mathbf{W}}.$$

Then the desired estimate then follows similar to the estimate on $|\tilde{x}_t|$. $\square$

REMARK 3.14. The above smoothing estimate implies the following: For some constant $C_\Pi \approx \log N$, there holds the following $\forall p$ with $p\eta < \eta^*$ and a β sufficiently large:

$$\mathbf{E} \|J_{0,1}(\text{Id} - \Pi_N)\|_{\mathbf{W} \rightarrow \mathbf{H}^\sigma}^p \leq V_{\beta,\eta}^p(u) \exp(-pC_\Pi).$$

REMARK 3.15. Lemma 3.13 is already nontrivial for $\gamma = 0$, that is, $J_{s,t}$ is a bounded linear operator.

The analogue of Lemma 3.11 holds also for $K_{s,t}$. The proof is actually easier and is omitted for brevity.

LEMMA 3.16. $\forall \sigma > \frac{d}{2} + 1, \forall r \in (\frac{d}{2} + 1, 3), \exists C, q' > 0$ such that the following hold pathwise:

$$\begin{aligned} \|K_{s,t}\|_{\mathbf{W} \rightarrow \mathbf{W}} &\lesssim \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right), \\ \|K_{s,t}\|_{\mathbf{H} \rightarrow \mathbf{H}} &\lesssim \exp\left(C \int_s^t \|u_\tau\|_{\mathbf{H}^r} d\tau\right) \left(1 + \langle t - s \rangle^3 \sup_{s < \tau < t} \|u_\tau\|_{\mathbf{H}^\sigma}^{q'}\right). \end{aligned}$$

4. Lasota–Yorke bound and spectral gap for the projective process.

4.1. Geometric ergodicity in C_V . In this section, we apply Theorem 2.4 to the projective process (u_t, x_t, v_t) . In [15], we already showed that there exists a unique stationary measure ν on $\mathbf{H} \times P\mathbb{T}^d$. Strong Feller in a scale of Sobolev spaces is proved in [15].

PROPOSITION 4.1 (From [15]). *For any $\sigma' \in (\alpha - 2(d-1), \alpha - \frac{d}{2})$ the projective Markov process (u_t, x_t, v_t) on $\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$ is strong Feller.*

By an easy variation of the irreducibility arguments in Section 6.2 together with those in [15], one deduces the following irreducibility property as well. The straightforward details are omitted for brevity.

PROPOSITION 4.2 (Essentially from [15]). *For any $\sigma' \in (\frac{d}{2} + 2, \alpha - \frac{d}{2})$, the projective process (u_t, x_t, y_t) , regarded as a process on $\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$, is topologically irreducible.*

The previous two properties imply equivalence of transition kernels by standard arguments (Lemma 3.2 in [39]; see also Theorem 4.1 in [36] for a special case).

LEMMA 4.3. *Propositions 4.1 and 4.2 imply that the family of transition kernels $\{\hat{P}_t((u, x, y), \cdot) : t > 0, (u, x, y) \in \mathbf{H} \times P\mathbb{T}^d\}$ are equivalent.*

Equivalence of transition measures provides the main tool for proving Condition 3.

LEMMA 4.4. *Let V be any of the Lyapunov functions for the projective process defined in Lemma 2.9. Let $K \subset \mathbf{H} \times P\mathbb{T}^d$ be any compact set with $\nu(K) > 0$. Fix arbitrary $t > 0$ and $r > 0$. Then we have*

$$\inf_{(u,x,v):V(u)\leq r}\hat{P}_t((u,x,v),K)>0.$$

PROOF. First, all transition kernels $\hat{P}_t((u,x,v),\cdot)$ are equivalent, and so we conclude that they are all equivalent with the unique stationary measure ν , which we regard as a measure on each $\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$. In particular, with K as in the hypothesis of Lemma 4.4, we have $\hat{P}_t((u,x,v),K)>0$ for all $(u,x,v)\in\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$, $\sigma'\in(\alpha-2(d-1),\alpha-\frac{d}{2})$.

For the sake of contradiction, assume Condition 3 fails. Then there is a sequence $\{z^n=(u^n,x^n,v^n):V(u^n)\leq r\}$ for which $\hat{P}_t(z^n,K)\rightarrow 0$ as $n\rightarrow\infty$. Fix $\sigma'\in(\alpha-2(d-1),\sigma)$. It follows from compact embedding and coercivity that $\{z^n\}$ admits a subsequence $\{z^{n'}\}$ converging in $\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$. For this sequence, we have

$$\lim_{n'\rightarrow\infty}\hat{P}_t(z^{n'},K)=0$$

by the strong Feller property on $\mathbf{H}^{\sigma'} \times P\mathbb{T}^d$ (Proposition 4.1). This contradicts Lemma 4.3 at σ' . $\square$

Putting everything together with Theorem 2.4 implies the geometric ergodicity of the projective process. Boundedness in C_V follows from Lemma 3.7. Recall ν is the stationary measure for the (u_t,x_t,v_t) process.

PROPOSITION 4.5. *For any V satisfying the conditions of Lemma 2.9, there exists a $\gamma>0$ (depending on V) such that for any ψ bounded measurable on $\mathbf{H} \times P\mathbb{T}^d$, there holds*

$$\left|\hat{P}_t\psi(z)-\int_{\mathbf{H}\times P\mathbb{T}^d}\psi\,\mathrm{d}\nu\right|\lesssim V(u)e^{-\gamma t}.$$

4.2. *Spectral gap in C_V^1 from a Lasota–Yorke estimate.* Recall in the outline that our approach to obtaining a spectral gap on $\hat{P}_t$ in $\mathring{C}_V^1$ is the following Lasota–Yorke gradient estimate. The version stated below was introduced by Hairer and Mattingly in [42] as a sufficient condition for the asymptotic strong Feller property as well as in [43] to prove spectral gaps in C_V^1 .

PROPOSITION 4.6 (Lasota–Yorke estimate). $\forall\beta'\geq 2$ sufficiently large and $\forall\eta'\in(0,\eta^*)$, $\exists C_1,\varkappa>0$ such that the following holds $\forall t>0$, and $\hat{z}=(u,x,v)\in\mathbf{H}\times P\mathbb{T}^d$:

$$\|D\hat{P}_t\psi(\hat{z})\|_{\mathbf{H}^*}\leq C_1V_{\beta',\eta'}(u)\Big(\sqrt{\hat{P}_t|\psi|^2(\hat{z})}+e^{-\varkappa t}\sqrt{\hat{P}_t\|D\psi\|_{\mathbf{H}^*}^2(\hat{z})}\Big).$$

Proposition 4.6 when combined with Proposition 4.5 is sufficient to deduce a spectral gap for $\hat{P}_{T_0}^P$ on $\mathring{C}_V^1$ for T_0 sufficiently large.

PROPOSITION 4.7. *For all $V=V_{\beta,\eta}$ with β sufficiently large and $\eta\in(0,\eta^*)$, we have that $\hat{P}_{T_0}^P$ has a spectral gap on C_V^1 for all T_0 sufficiently large.*

PROOF. For some $\delta > 0$ to be chosen below, define the equivalent norm on C_V^1 ,

$$\|\psi\|_{C_{V,\delta}^1} := \sup_{w \in \mathbf{H} \times P\mathbb{T}^d} \left(\frac{|\psi(\hat{z})|}{V(u)} + \frac{\delta \|D\psi(\hat{z})\|_{\mathbf{H}^*}}{V(u)} \right).$$

Choose $\beta' < \beta$ and $\eta' < \eta$. Then the Lasota–Yorke bound (Proposition 4.6) together with the super-Lyapunov property (3.12) implies

$$\begin{aligned} \|D\hat{P}_n\psi(\hat{z})\|_{\mathbf{H}^*} &\leq C_1 V_{\beta',\eta'} \sqrt{\hat{P}_n V^2(u)} (1 + e^{-\gamma n} \delta^{-1}) \|\psi\|_{C_{V,\delta}^1} \\ &\leq \frac{C}{\delta} V_{\beta',\eta'}(u) V^{e^{-\gamma n}}(u) (\delta + e^{-\gamma n}) \|\psi\|_{C_{V,\delta}^1}. \end{aligned}$$

Since $\beta > \beta'$ and $\eta > \eta'$, there exists a n_0 sufficiently large so that

$$\beta' + e^{-\gamma n} \beta < \beta \quad \text{and} \quad \eta' + e^{-\gamma n} \eta < \eta,$$

which implies $V_{\beta',\eta'}(u) V^{e^{-\gamma n_0}}(u) \leq V(u)$. Choosing $\delta > 0$ such that $C\delta < 1/8$ and $n_1 > n_0$ such that $Ce^{-\gamma n_1} < 1/8$, we have

$$\|D\hat{P}_{n_1}\psi(\hat{z})\|_{\mathbf{H}^*} \leq \frac{1}{4} \frac{V(u)}{\delta} \|\psi\|_{C_{V,\delta}^1}.$$

Combining this with the C_V spectral gap (Proposition 4.5) implies that $\hat{P}_{T_0}^P$ is a contraction on C_V^1 with respect to the $\|\cdot\|_{C_{V,\delta}^1}$ norm for all $T_0 > 0$ sufficiently large. This implies a spectral gap of $\hat{P}_{T_0}^P$ in C_V^1 . $\square$

4.3. Lasota–Yorke estimate: Proof of Proposition 4.6. Let us now turn to the proof of Proposition 4.6. The proof follows closely the discussion in [44] with some adjustments, and so we only provide a brief sketch. Aside from the fact that the necessary estimates on the $J_{s,t}$ and $K_{s,t}$ processes are more complex than in [44], the Malliavin matrix nondegeneracy requires some adjustments to deal with the degrees of freedom associated with $P\mathbb{T}^d$. Additionally, we need to use Lemma 3.13 and choose β' sufficiently large to control $\mathbf{H}$ regularity. Other than this, only minor modifications are needed.

4.3.1. Malliavin calculus and preliminaries. First, let us recall some basics of Malliavin calculus required to set up the framework of [44]. We will mostly be dealing with random variables $X \in \mathbf{W} \times P\mathbb{T}^d$. The Malliavin derivative $\mathcal{D}_h X$ of X in a Cameron–Martin direction $h = (h_t) \in L^2(\mathbb{R}_+, \mathbf{L}^2)$ is defined by

$$\mathcal{D}_h X := \frac{d}{d\epsilon} X(W + \epsilon H)|_{\epsilon=0}, \quad H = \int_0^\cdot h_s ds,$$

when the above derivative exists (in the Fréchet sense). If the derivative exists for each Cameron–Martin direction $h \in L^2(\mathbb{R}_+, \mathbf{L}^2)$, we say that X is *Malliavin differentiable*.

In practice, $\mathcal{D}_h X$ admits a representation of the form

$$\mathcal{D}_h X = \int_0^\infty \mathcal{D}_s X h_s ds,$$

where for a.e. $s \in \mathbb{R}_+$, $\mathcal{D}_s X$ is a random, bounded linear operator from $\mathbf{L}^2$ to $\mathbf{W} \times T_v P\mathbb{T}^d$ (see [59] for more details). It is standard that if X_t is adapted to the filtration $\mathcal{F}_t$ generated by W_t , then $\mathcal{D}_s X_t = 0$ if $s \geq t$. The Malliavin derivative $\mathcal{D}_h w_t$ is given by (recall $J_{s,t}$ is the Jacobian; see (3.13)),

$$\mathcal{D}_h w_t = \int_0^t J_{s,t} Q h_s ds := \mathcal{A}_t h.$$

The adjoint is given by $A_t^* \xi(s) = Q^* K_{s,t} \xi$ for $s \leq t$ and 0 for $s > t$ (recall $K_{s,t} = J_{s,t}^*$). We similarly denote $\mathcal{A}_{s,t} h = \int_s^t J_{\tau,t} Q h_\tau d\tau$. The Malliavin matrix $\mathcal{M}_{s,t}$ is a symmetric, positive semidefinite as a linear operator $\mathbf{W} \times T_{v_t} P\mathbb{T}^d \rightarrow \mathbf{W} \times T_{v_t} P\mathbb{T}^d$ defined by

$$\mathcal{M}_{s,t} := \mathcal{A}_{s,t} \mathcal{A}_{s,t}^* = \int_s^t J_{r,t} Q Q^* K_{r,t} dr.$$

For real-valued random variables, the Malliavin derivative can be realized as a Fréchet differential operator $\mathcal{D} : L^2(\Omega) \rightarrow L^2(\Omega; L^2(\mathbb{R}_+; \mathbf{L}^2))$. The adjoint operator $\mathcal{D}^* : L^2(\Omega; L^2(\mathbb{R}_+; \mathbf{L}^2)) \rightarrow L^2(\Omega)$ is referred to as the Skorohod integral, whose action on $h \in L^2(\Omega; L^2(\mathbb{R}_+; \mathbf{L}^2))$ we denote by

$$\int_0^\infty \langle h_t, \delta W_t \rangle_{\mathbf{L}^2} := \mathcal{D}^* h.$$

The Skorohod integral is an extension of the usual Itô integral; see [44, 59]. One moreover has the following:

$$\mathbf{E} \left(\int_0^\infty \langle h_t, \delta W_t \rangle_{\mathbf{L}^2} \right)^2 \leq \mathbf{E} \int_0^\infty \|h_t\|_{\mathbf{L}^2}^2 + \mathbf{E} \int_0^\infty \int_0^\infty \|\mathcal{D}_s h_t\|_{\mathbf{L}^2 \rightarrow \mathbf{L}^2}^2 ds dt.$$

A fundamental result in the theory of Malliavin calculus is the Malliavin integration by parts formula, stated below for the process $(\hat{z}_t)$ which takes values in $\mathbf{H} \times P\mathbb{T}^d$ (see, e.g., [24, 59]).

PROPOSITION 4.8. *Let ψ be a bounded differentiable function on $\mathbf{H} \times P\mathbb{T}^d$ with bounded derivatives and h_t be a process satisfying*

$$\mathbf{E} \int_0^T \|h_t\|_{\mathbf{L}^2}^2 dt + \mathbf{E} \int_0^T \int_0^T \|\mathcal{D}_s h_t\|_{\mathbf{L}^2 \rightarrow \mathbf{L}^2}^2 ds dt < \infty,$$

then the following relation holds:

$$\mathbf{E} \mathcal{D}_h \psi(\hat{z}_T) = \mathbf{E} \left(\psi(\hat{z}_T) \int_0^T \langle h_s, \delta W_s \rangle_{\mathbf{L}^2} \right).$$

4.3.2. Defining the control. By Proposition 4.8, for any choice of g satisfying the hypotheses, there holds

$$(4.1) \quad D \hat{P}_t \psi(\hat{z}) \xi = \mathbf{E} D \psi(\hat{z}_t) J_t \xi = \mathbf{E} D \psi(\hat{z}_t) \rho_t + \mathbf{E} \psi(\hat{z}_t) \int_0^t \langle h_s, \delta W_s \rangle_{\mathbf{L}^2},$$

where $\rho_t \in \mathbf{H} \times T_{v_t} P\mathbb{T}^d$ is the residual of the control,

$$\rho_t = J_t \xi - \mathcal{D}_h \hat{z}_t.$$

From (4.1), Cauchy–Schwarz implies

$$(4.2) \quad |D \hat{P}_t \psi(\hat{z}) \xi| \leq \sqrt{\mathbf{E} \left| \int_0^t \langle h_s, \delta W_s \rangle_{\mathbf{L}^2} \right|^2} \sqrt{\hat{P}_t |\psi|^2(z)} + \sqrt{\mathbf{E} \|\rho_t\|_{\mathbf{H}}^2} \sqrt{\hat{P}_t \|D \psi\|_{\mathbf{H}^*}^2(z)}.$$

The goal in [42, 44] is to choose h such that $\mathbf{E} \|\rho_t\|_{\mathbf{H}}^2$ decays exponentially fast while, at the same time, having a suitably controlled Skorohod integral. The same abstract formula for the control used in [44] will also work here (though obviously the control itself is different). For a random parameter $\beta_k > 0$ chosen below, we iteratively define the control as follows:

$$(4.3) \quad h_s = h_s^\xi = \begin{cases} Q^* K_{s,2n+1} (\beta_{2n} + \mathcal{M}_{2n,2n+1})^{-1} J_{2n,2n+1} \rho_{2n} & s \in [2n, 2n+1), \\ 0 & s \in [2n+1, 2n+2). \end{cases}$$

With the control defined as above, one obtains the following recurrence relation for the ρ_{2n} at even times:

$$\rho_{2n+2} = J_{2n+1,2n+2} \beta_{2n} (\beta_{2n} - \mathcal{M}_{2n,2n+1})^{-1} J_{2n,2n+1} \rho_{2n}$$

By (4.2), Proposition 4.6 follows from the following two lemmas. First, a decay estimate is shown in the following.

LEMMA 4.9. *For all $a \geq 2$ sufficiently large, $\forall \eta \in (0, \eta^*)$, and $\forall p \geq 2$ with $p\eta < \eta^*$, $\exists \kappa > 0$ such that the following holds,*

$$\mathbf{E} \|\rho_{2n}\|_{\mathbf{H}}^p \lesssim_{p,\eta,a} V_{a,\eta}^p(u) \exp(-p\kappa n) \|\xi\|_{\mathbf{H}}^p.$$

Second, a uniform estimate on the Skorohod integral of the control is shown in the following.

LEMMA 4.10. *Let h be defined as in (4.3). For all $a \geq 1$ sufficiently large, $\forall \eta \in (0, \eta^*)$, there holds*

$$\mathbf{E} \left| \int_0^{2n} \langle h_s, \delta W_s \rangle_{\mathbf{L}^2} \right|^2 \lesssim V_{a,\eta}(u) \|\xi\|_{\mathbf{H} \times T_v P\mathbb{T}^d}^2.$$

It remains to sketch the proofs of Lemmas 4.9 and 4.10. Let $\Pi : \mathbf{L}^2 \rightarrow \mathbf{L}^2$ be the projection to frequencies less than some high frequency N_0 chosen later and fixed through out the section. In what follows, it is convenient to define for $k \in \mathbb{K}$:

$$g_k := (q_k e_k, 0, 0).$$

In this notation, the Malliavin matrix takes the form for each $\xi \in \mathbf{W} \times T_{v_1} P\mathbb{T}^d$,

$$(4.4) \quad \langle \mathcal{M}_{s,t} \xi, \xi \rangle_{\mathbf{W}} = \sum_{k \in \mathbb{K}} \int_s^t \langle g_k, K_{r,t} \xi \rangle_{\mathbf{W}}^2 dr.$$

The precise statement of the Malliavin matrix nondegeneracy is as follows; we carry out the proof in Section 4.3.4.

PROPOSITION 4.11 (Malliavin matrix nondegeneracy). *For all $a \geq 1$ sufficiently large, and all $\eta \in (0, \eta^*)$, $p \geq 2$ with $\eta p \leq \eta^*$, $\forall \delta > 0$ there exists a $C^* = C^*(p, \delta, a, \eta)$ such that*

$$\mathbf{P} \left(\inf_{\xi \in S_\delta} \langle \xi, \mathcal{M}_{0,1} \xi \rangle_{\mathbf{W}} < \epsilon \right) \leq C^* V_{a,\eta}^p(u) \epsilon^p,$$

where

$$S_\delta = \{ \xi \in \mathbf{W} \times T_{v_1} P\mathbb{T}^d : \|\xi\|_{\mathbf{W}} = 1, \|\Pi \xi\|_{\mathbf{W}} > \delta \}.$$

Once Proposition 4.11 is obtained, we will set β_k as in [44] for suitably chosen parameters

$$\beta_k = \frac{\delta^3}{V_{a,\eta}(u_k) (C^*)^{1/p}}.$$

From [Corollary 5.15, [44]], Proposition 4.11 implies

$$\mathbf{E} \|\Pi \beta_{2n} (\beta_{2n} + \mathcal{M}_{2n,2n+1})^{-1}\|_{\mathbf{W} \rightarrow \mathbf{W}}^p \leq 2\delta^p.$$

Once one has suitable estimates on the $J_{s,t}$ and $K_{s,t}$ (and by extension, $\mathcal{A}$, $\mathcal{A}^*$, $\mathcal{M}$) as well as the Malliavin derivatives of $J_{s,t}$ and $K_{s,t}$ (in order to estimate the Skorohod integral in Lemma 4.10 via (4.3.1)) then Lemmas 4.9 and 4.10 follow essentially as in [44] (taking full advantage of Lemma 3.13 and Remark 3.14 to switch to and from the $\mathbf{W}$ and $\mathbf{H}$ spaces). Most of the required estimates are immediate from Section 3; we provide a brief sketch of the relevant Malliavin derivative estimates next.

4.3.3. *Malliavin derivatives of the Jacobian.* Estimating the Malliavin derivatives of $J_{s,t}$ and similar quantities requires estimates on the second variation, which we denote $J_{s,t}^{(2)}$. This operator is given by (see [44] or [15]):

$$\partial_t J_{s,t}^{(2)}(\xi, \zeta) = D^2 F(\hat{z}_t)(J_{s,t}\xi, J_{s,t}\zeta) + DF(\hat{z}_t)J_{s,t}^{(2)}(\xi, \zeta)s < t.$$

Observe that $D^2 F$ is of the form:

$$D^2 F(\hat{z}_t)(\xi, \zeta) = \begin{pmatrix} B(\xi, \zeta) + B(\zeta, \xi) \\ D\xi^u(x_t)\zeta^x + D\zeta^u(x_t)\xi^x + \xi^x \cdot D^2 u(x_t)\zeta^x \\ \mathfrak{V}_t(\xi, \eta) \end{pmatrix},$$

where $\mathfrak{V}$ is of the following general form: there exists a set of bounded operators $\mathcal{T}_j$ all multilinear in all but the last components,

$$(4.5) \quad \begin{aligned} \mathfrak{V}_t &= \mathcal{T}_1(\xi^v, \zeta^v, \nabla u_t(x_t), v_t) + \mathcal{T}_2(\xi^v, \zeta^x, \nabla^2 u_t(x_t), v_t) \\ &+ \mathcal{T}_3(\xi^v, \nabla \zeta^u(x_t), v_t) \\ &+ \mathcal{T}_4(\xi^x, \zeta^x, \nabla^3 u_t(x_t), v_t) + \mathcal{T}_5(\xi^x, \nabla^2 \zeta^u(x_t), v_t). \end{aligned}$$

The high numbers of derivatives appearing in $\mathfrak{V}$ makes obtaining estimates on $J_{s,t}^{(2)}$ nonobvious, indeed, it is bounded on high enough Sobolev spaces, but it is not clear that it defines a bounded operator on $(\mathbf{W} \times T_{v_s} P\mathbb{T}^d)^{\otimes 2} \rightarrow \mathbf{W} \times T_{v_t} P\mathbb{T}^d$. Nevertheless, the proof of Lemma 4.10 uses bounds on Malliavin derivatives of $J_{s,t}$ in such low regularities (see [44]). This requires taking advantage of the specific form of the Malliavin derivatives of $J_{s,t}$.

LEMMA 4.12 (Malliavin derivative bounds). *$\forall a \geq 1$ sufficiently large and all $p \geq 1$, $\eta \in (0, \eta^*)$ such that $p\eta \in (0, \eta^*)$, there exists a constant C_M such that the following holds (where C_L and γ are as in (3.12) above),*

$$(4.6) \quad \begin{aligned} \mathbf{E} \sup_{s,r \in [n,n+1]} \|\mathcal{D}_s J_{r,n+1}\|_{\mathbf{W} \rightarrow \mathbf{W}}^p &\lesssim e^{\frac{pC_L}{1-e^{-\gamma}} + pC_M} V_{a,\eta}^{pe^{-\gamma n}}(u), \\ \mathbf{E} \sup_{s \in [n,n+1]} \|\mathcal{D}_s \mathcal{A}_{n,n+1}\|_{\mathbf{W} \rightarrow \mathbf{W}}^p &\lesssim e^{\frac{pC_L}{1-e^{-\gamma}} + pC_M} V_{a,\eta}^{pe^{-\gamma n}}(u). \end{aligned}$$

REMARK 4.13. Estimates on $\mathcal{D}_s \mathcal{A}_{n,n+1}^*$ follow by duality.

PROOF. Consider the Malliavin derivative with respect to the k -th Brownian motion (denoted by $\mathcal{D}_s^k$), which is given by (see, e.g., [Lemma 5.13, [44]]),

$$\mathcal{D}_s^k J_{r,n+1}\xi = \begin{cases} J_{s,n+1}^{(2)}(J_{r,s}\xi, g_k) & r \leq s, \\ J_{r,n+1}^{(2)}(J_{s,r}g_k, \xi) & s \leq r. \end{cases}$$

Next, we observe that

$$J_{s,t}^{(2)}(\xi, \zeta) = \int_s^t J_{r',t}(D^2 F(\hat{z}_{r'})(J_{s,r'}\xi, J_{s,r'}\zeta)) dr'.$$

Therefore, for $r \leq s$

$$J_{s,n+1}^{(2)}(J_{r,s}\xi, g_k) = \int_s^{n+1} J_{r',n+1}(D^2 F(\hat{z}_{r'})(J_{r,r'}\xi, J_{s,r'}g_k)) dr'.$$

Denote for suitably large parameters C, q ,

$$\Gamma(r') = \exp\left(C \int_s^{r'} \|u_\tau\|_{H^{r_0}} d\tau\right) \sup_{s < \tau < r'} (1 + \|u_\tau\|_{\mathbf{H}}^q).$$

Due to the particular form of $J^{(2)}$ and in particular of $\mathfrak{V}$ in (4.5), that g_k is smooth and supported only in the velocity variables implies a significant simplification. For example, $|(J_{s,r'}g_k)^{x,v}| \lesssim |k|^{-\alpha+\frac{d+3}{2}}|s-r'|\Gamma(r')$. Hence, where this appears, we may use the gain in time to balance loss of regularity through the smoothing estimate in Lemma 3.13. For $r \leq s \leq r'$, (using Remark 3.15 as well),

$$\begin{aligned} & \|J_{r',n+1}(D^2F(\hat{z}_{r'})(J_{r,r'}\xi, J_{s,r'}g_k))\|_{\mathbf{W}} \\ & \lesssim \Gamma(r') \left(\frac{1}{((n+1)-r')^{\frac{1}{2(d-1)}}} + \frac{(s-r')}{(r-r')^{\frac{1}{2(d-1)}}} \right) \|\xi\|_{\mathbf{W}} \\ & \lesssim \Gamma(r') \left(\frac{1}{((n+1)-r')^{\frac{1}{2(d-1)}}} + 1 \right) \|\xi\|_{\mathbf{W}}. \end{aligned}$$

Applying Lemma 3.7 followed by Lemma 3.12 gives the desired result.

The estimate for $s \leq r$ follows similarly, completing the required estimate (4.6). The estimate on $\mathcal{D}_s\mathcal{A}_{n,n+1}$ then follows as in [44]. $\square$

4.3.4. Nondegeneracy of the Malliavin matrix. In this section, we sketch the proof of Proposition 4.11. The approach is similar to that of Hairer and Mattingly [44]. However, the proof in [44] does not exactly apply in our setting due to the fact that the nonlinearity in the (x, v) dynamics is not polynomial. However, our situation is much nicer due to nondegenerate noise, and Lemma 4.15 tells us that we only need one Lie bracket to span the phase space (as opposed to infinitely many brackets in [44]). As we saw in our previous paper [15], this is enough to avoid any quantitative noncancellation results like Norris' lemma or estimates on Wiener polynomials. Since our proof is relatively simple and different enough from [44] we include it below.

First, we record for the readers' convenience the following estimates proved in Section 3.1.

LEMMA 4.14. *For every $a \geq 2$, $\eta p \in (0, \eta^*)$, $p \geq 1$, we have*

$$\begin{aligned} \mathbf{E} \sup_{0 \leq t \leq 1} \|u_t\|_{\mathbf{H}}^p & \lesssim V_{a,\eta}^p(u), \\ \mathbf{E} \sup_{0 < s < 1} \|K_{s,1}\|_{\mathbf{W} \rightarrow \mathbf{W}}^p & \lesssim V_{a,\eta}^p(u). \end{aligned}$$

Naturally, one of the main ingredients is a uniform spanning condition on Lie brackets to verify hypoellipticity. The following is a straightforward consequence of the spanning lemma in our previous work [Lemma 5.3, [15]].

LEMMA 4.15. *For each initial $\hat{z} = (u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$, $\xi \in \mathbf{W} \times T_v P\mathbb{T}^d$ with $\|\xi\|_{\mathbf{W}} = 1$ and s such that $(2\alpha + d)/4 < s < \sigma - 2(d-1)$ we have*

$$\sup_{k \in \mathbb{K}} \{ |\langle g_k, \xi \rangle_{\mathbf{W}}|, |\langle DF(\hat{z})g_k, \xi \rangle_{\mathbf{W}}|^2 \} \gtrsim_s \frac{\|\xi\|_{\mathbf{H}^{-s}}^2}{(1 + \|u\|_{\mathbf{H}})^2}.$$

PROOF. Denote $\xi = (\xi_u, \xi_{x,v})$. First, for $s > (2\alpha + d)/4$ it is not hard to deduce

$$(4.7) \quad \sup_{k \in \mathbb{K}} |\langle g_k, \xi \rangle_{\mathbf{W}}| \gtrsim_s \|\xi^u\|_{\mathbf{H}^{-s}}^2.$$

Therefore, (4.7) follows since $4s - 2\alpha > d$ and $\|\xi^u\|_{\mathbf{W}} = 1$. Next, we note that

$$\langle DF(\hat{z})g_k, \xi \rangle_{\mathbf{W}} = q_k \langle B(e_k, u) + B(u, e_k), \xi_u \rangle_{\mathbf{W}} + \langle Ae_k, \xi_u \rangle_{\mathbf{W}} + \langle DF^{x,v}(\hat{z})g_k, \xi_{x,v} \rangle_{T_v P\mathbb{T}^d},$$

and that since $s < \sigma - 2$ and $\sup_{k \in \mathbb{K}} q_k \|e_k\|_{\mathbf{H}} < \infty$,

$$\begin{aligned} & \|\xi^u\|_{\mathbf{H}}^2 + \sup_{k \in \mathbb{K}} |\langle DF^{x,v}(\hat{z})g_k, \xi^{x,v} \rangle_{\mathbf{W}}|^2 \\ & \lesssim \sup_{k \in \mathbb{K}} |\langle DF(\hat{z})g_k, \xi \rangle_{\mathbf{W}}|^2 + (1 + \|u\|_{\mathbf{H}})^2 \|\xi^u\|_{\mathbf{H}}^2 \\ & \lesssim (1 + \|u\|_{\mathbf{H}})^2 \sup_{k \in \mathbb{K}} \{|\langle g_k, \xi \rangle_{\mathbf{W}}|, |\langle DF(\hat{z})g_k, \xi \rangle_{\mathbf{W}}|^2\}, \end{aligned}$$

where the last line followed from (4.7). The proof is complete upon appealing to [Lemma 5.3, [15]] and deducing

$$\sup_{k \in \mathbb{K}} |\langle DF^{x,v}(\hat{z})g_k, \xi^{x,v} \rangle_{T_v P \mathbb{T}^d}| \gtrsim \|\xi^{x,v}\|_{T_v P \mathbb{T}^d}.$$

□

Before continuing with the nondegeneracy calculation, we introduce the following useful shorthand notation from [Definition 6.9, [44]].

DEFINITION 4.16. Given a set $H = \{H^\epsilon\}_{\epsilon \leq 1}$, of measurable subsets $H^\epsilon \subset \Omega$, we will say “ H is a family of negligible events” if $\forall p \geq 1$, $\exists C_p$ such that $\mathbf{P}(H^\epsilon) \leq C_p \epsilon^p$. Given such a family and a statement Φ_ϵ , we say “ Φ_ϵ holds modulo H ” if $\forall \epsilon \leq 1$, the statement Φ_ϵ holds on $(H^\epsilon)^c$. Finally, we say the family H is “universal” if it does not depend on the problem at hand. We say a set H is “ Ψ -controlled” for a function Ψ if $C_p \lesssim_p \Psi^p(\hat{z})$ (the initial condition of the projective process).

Recall the Hölder seminorm of a function $f : [0, 1] \rightarrow \mathbb{R}$ for $\alpha \in (0, 1]$ is defined by

$$[f]_\alpha = \sup_{t \neq s} \frac{|f(t) - f(s)|}{|s - t|^\alpha}.$$

Recall the following standard interpolation lemma (see, e.g., [44]).

LEMMA 4.17. *Let f be a $C^{1,\alpha}$ function on $[0, 1]$, $\alpha \in (0, 1]$, then the following inequality holds:*

$$\|\partial_t f\|_{L^\infty} \leq 4 \|f\|_{L^\infty} \max\{1, \|f\|_{L^\infty}^{-\frac{1}{1+\alpha}} [\partial_t f]_\alpha^{\frac{1}{1+\alpha}}\}.$$

Recall the following formula for the Malliavin matrix:

$$\langle \xi, \mathcal{M}_{0,1} \xi \rangle_{\mathbf{W}} = \sum_{k \in \mathbb{K}} \int_0^1 \langle g_k, K_{s,1} \xi \rangle_{\mathbf{W}}^2 ds.$$

In light of this, we have the following implication.

LEMMA 4.18. *For all $a \geq 2$ sufficiently large and $\forall \eta \in (0, \eta^*)$ the following implication holds:*

$$\langle \xi, \mathcal{M}_{0,1} \xi \rangle_{\mathbf{W}} \leq \epsilon \quad \Rightarrow \quad \sup_k \sup_{0 \leq t \leq 1} |\langle g_k, K_{t,1} \xi \rangle_{\mathbf{W}}| \leq \epsilon^{1/8},$$

modulo a $V_{a,\eta}$ -controlled negligible set.

PROOF. Fix an arbitrary $\eta' \in (0, \eta^*)$ and $a' \geq 2$. Define $f_k(t) = \int_0^t \langle g_k, K_{s,1} \xi \rangle_{\mathbf{W}} ds$ and note that $f_k(t)$ is C^2 and satisfies

$$\partial_t f_k = \langle g_k, K_{t,1} \xi \rangle_{\mathbf{W}} \quad \text{and} \quad \partial_t^2 f_k = \langle DF(\hat{z}_t)g_k, K_{t,1} \xi \rangle_{\mathbf{W}}.$$

Additionally, when $\langle \xi, \mathcal{M}_1 \xi \rangle_{\mathbf{W}} \leq \epsilon$, $\|\xi\|_{\mathbf{W}} = 1$, we have $\sup_k \|f_k\|_{L^\infty} \leq \epsilon^{1/2}$ by (4.4) and Lemma 4.17.

Furthermore, by Lemma 4.14 there holds

$$\mathbf{E} \sup_k \|\partial_t^2 f_k\|_{L^\infty}^p \lesssim \left(\mathbf{E} \sup_{0 \leq t \leq 1} (1 + \|u_t\|_{\mathbf{H}^{2(d-1)}})^{2p} \right)^{1/2} \left(\mathbf{E} \sup_{0 \leq t \leq 1} \|K_{t,1}\|_{\mathbf{W} \rightarrow \mathbf{W}}^{2p} \right)^{1/2} \lesssim V_{\beta, \eta'}^p(u)$$

and, therefore, Chebyshev's inequality implies that modulo a $V_{4a', \eta'}$ -controlled negligible set one has the bound $\sup_k \|\partial_t^2 f_k\|_{L^\infty} \leq (1/16)\epsilon^{-1/4}$. Applying the interpolation Lemma 4.17 to f_k with $\alpha = 1$ modulo this family of negligible sets gives the implication. $\square$

LEMMA 4.19. *For all $a \geq 2$ sufficiently large and $\forall \eta \in (0, \eta^*)$ sufficiently small, the following implication holds:*

$$\sup_k \sup_{0 \leq t \leq 1} |\langle g_k, K_{t,1} \xi \rangle_{\mathbf{W}}| \leq \epsilon \quad \Rightarrow \quad \sup_k \sup_{0 \leq t \leq 1} |\langle DF(\hat{z}_t) g_k, K_{t,1} \xi \rangle_{\mathbf{W}}| \leq \epsilon^{1/8}$$

modulo a $V_{a, \eta}$ -controlled negligible set.

PROOF. Fix $\eta' \in (0, \eta^*)$ and $a' \geq 2$ and let $f_k = \int_0^t \langle g_k, K_{s,1} \xi \rangle_{\mathbf{W}} ds$. Our goal will be to apply the interpolation Lemma 4.17 to $\partial_t f_k$. However, since $DF(\hat{z}_t) g_k$ is only Hölder continuous, $\partial_t^2 f_k$ is not C^1 and we will need to obtain moment estimate on the $\alpha = 1/3$ Hölder seminorm of $\partial_t^2 f_k$. Indeed, applying the interpolation Lemma 4.17 to $\partial_t f_k$ with $\alpha = 1/3$ gives

$$(4.8) \quad \sup_k \|\partial_t^2 f_k\|_{L^\infty} \leq 4\epsilon^{1/4} \max \left\{ \epsilon^{3/4}, \sup_k [\partial_t^2 f_k]_{1/3}^{3/4} \right\}.$$

The required Hölder estimate on $\partial_t^2 f_k$ then follows from the high regularity of $u \in \mathbf{H}$, by standard time regularity estimates on Wiener processes and Lemma 4.14,

$$\mathbf{E} \sup_k [\partial_t^2 f_k]_{1/3}^p \lesssim V_{a', \eta'}^{2p}(u).$$

Therefore, modulo a $V_{12a', 12\eta'}$ -controlled negligible set one has the bound $\sup_k [\partial_t^2 f_k]_{1/3} < 2^{-8/3} \epsilon^{-1/6}$. Substituting this into (4.8) gives the desired result. $\square$

PROOF OF PROPOSITION 4.11. To prove this, we note that Lemmas 4.18 and 4.19 imply that for each $\delta \in (0, 1)$ and $\xi \in \mathcal{S}_\delta$ the following implication holds modulo a $V_{\beta, \eta}$ -controlled family of negligible events:

$$\langle \xi, \mathcal{M}_{0,1} \xi \rangle_{\mathbf{W}} < \epsilon \quad \Rightarrow \quad \begin{cases} \sup_k \sup_{0 \leq t \leq 1} |\langle g_k, K_{t,1} \xi \rangle_{\mathbf{W}}| \leq \epsilon^{1/8}, \\ \sup_k \sup_{0 \leq t \leq 1} |\langle DF(\hat{z}_t) g_k, K_{t,1} \xi \rangle_{\mathbf{W}}| \leq \epsilon^{1/64}. \end{cases}$$

By taking $t = 1$, this implies

$$\sup_k \{ |\langle g_k, \xi \rangle_{\mathbf{W}}|, |\langle DF(\hat{z}) g_k, \xi \rangle_{\mathbf{W}}|^2 \} \leq \epsilon^{1/32}.$$

However, appealing to Lemma 4.15 implies

$$\frac{\delta^2}{(1 + \|u\|_{\mathbf{H}})^2} \lesssim \frac{\|\xi\|_{\mathbf{H}^{-s}}^2}{(1 + \|u\|_{\mathbf{H}})^2} \lesssim \epsilon^{1/32}.$$

Therefore, choosing ϵ small enough like $\epsilon \lesssim_\delta (1 + \|u\|_{\mathbf{H}})^{-64}$, we deduce that

$$\mathbf{P}(\langle \xi, \mathcal{M}_{0,1} \xi \rangle_{\mathbf{W}} < \epsilon) \lesssim V_{a, \eta}^p(u) \epsilon^p.$$

We can extend this estimate to all $\epsilon > 0$, by noting the other case is $1 \lesssim (1 + \|u\|_{\mathbf{H}})^{64p} \epsilon^p$, and hence the proof is complete by choosing $a > 64$. $\square$

5. Spectral theory for the twisted Markov semigroup. As discussed in Section 2, our approach to proving a drift condition for the two-point process involves using spectral properties of the “twisted” Markov semigroup $\hat{P}_t^p$ defined for bounded measurable $\psi : \mathbf{H} \times P\mathbb{T}^d \rightarrow \mathbb{R}$, by

$$\hat{P}_t^p \psi(u, x, v) := \mathbf{E}_{(u,x,v)} |D\phi^t v|^{-p} \psi(u_t, x_t, v_t).$$

To simplify notation, we will denote $\hat{z} = (u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$ and $\hat{z}_t = (u_t, x_t, v_t)$. It is important to note that the semigroup $\hat{P}_t^p$ can be written as a Feynman–Kac semigroup

(5.1)
$$\hat{P}_t^p \psi(\hat{z}) = \mathbf{E}_{\hat{z}} \exp\left(-p \int_0^t H(\hat{z}_s) \, ds\right) \psi(\hat{z}_t),$$

where

$$H(u, x, v) := \langle v, \nabla u(x) v \rangle.$$

As discussed in the outline, our analysis requires that we study $\hat{P}_t^p$ with respect to both the C_V norm (as defined in (2.1)) as well as the stronger C_V^1 norm

$$\|\psi\|_{C_V^1} = \sup_{\hat{z} \in \mathbf{H} \times P\mathbb{T}^d} \left(\frac{|\psi(\hat{z})|}{V(u)} + \frac{\|D\psi(\hat{z})\|_{\mathbf{H}^*}}{V(u)} \right),$$

where $V(u) = V_{\beta,\eta}(u)$ for some choice of $\beta > 1$ and $\eta < \eta^*$.

The plan is as follows: We start in Section 5.1 in the C_V^1 framework by proving Proposition 2.16, establishing a spectral gap for $\hat{P}_{T_0}^p$ in C_V^1 for T_0 sufficiently large. In Section 5.2, we obtain C_0 continuity for $\hat{P}_t^p$ on the space $\mathring{C}_V$ as in Proposition 2.17. Finally, in Section 5.3 we pull this all together to prove the desired properties of the eigenfunction ψ_p in Proposition 2.11.

5.1. Proof of Proposition 2.16. As discussed in the outline, Proposition 2.16 follows from a spectral perturbation argument. Let us first record the following standard spectral perturbation lemma, which is a consequence of analytic functional calculus (see, e.g., [31]).

LEMMA 5.1. *Let L be a bounded linear operator on a Banach space $\mathfrak{B}$ with norm $\|\cdot\|$.*

(a) *For any $\epsilon > 0$, there exists $\delta > 0$ such that for any bounded linear L' on $\mathfrak{B}$ with $\|L - L'\| < \delta$, we have $\sigma(L') \subset B_\epsilon(\sigma(L))$.*

(b) *Let $S \subset \sigma(L)$ be a closed, isolated subset, that is, there exists an open set $\mathcal{U} \subset \mathbb{C}$ such that $S \cap \mathcal{U} = \sigma(L) \cap \mathcal{U}$. Let π_S denote the spectral projector corresponding to S . Then, for any $\epsilon > 0$, there exists $\delta > 0$ such that if $\|L - L'\| < \delta$ for some bounded linear L' , then $S' := \mathcal{U} \cap \sigma(L')$ is a closed, isolated subset of $\sigma(L')$ with the property that the spectral projector $\pi_{S'}$ for L' corresponding to S' satisfies $\|\pi_S - \pi_{S'}\| < \epsilon$.*

To apply Lemma 5.1, we show that $\hat{P}_t^p$ is a bounded perturbation of $\hat{P}_t$ for large enough t and small enough p .

LEMMA 5.2. *There exist $T_0 > 0$, $p_0 > 0$ such that $\forall \beta$ sufficiently large and all $\eta \in (0, \eta^*)$ there holds the following:*

- (a) $\hat{P}_{T_0}^p : C_V^1 \rightarrow C_V^1$ is a bounded linear operator for all $p \in [-p_0, p_0]$.
- (b) We have $\hat{P}_{T_0}^p \rightarrow \hat{P}_{T_0}$ in norm on C_V^1 as $p \rightarrow 0$.

PROOF. We first consider the proof of (a). Note that $|H(\hat{z})| \leq \|\nabla u\|_{L^\infty}$, and hence C_V boundedness follows from (5.1) and Lemma 3.7. Next, compute $D\hat{P}_t^p \psi \xi$ for some $\xi \in \mathbf{H} \times T_v P\mathbb{T}^d$:

$$D\hat{P}_t^p \psi \xi = \mathbf{E} \exp\left(-p \int_0^t H(z_s) ds\right) \left(D\psi(z_t) J_{0,t} \xi - p\psi(z_t) \int_0^t DH(z_s) J_{0,s} \xi ds \right).$$

By Lemma 3.11, there holds (for any $r \in (\frac{d}{2} + 1, 3)$),

$$\begin{aligned} |D\hat{P}_t^p \psi \xi| &\leq \|\psi\|_{C_V^1} \mathbf{E} \exp\left(C \int_0^t \|u_s\|_{H^r} ds\right) V_{\beta, \eta}(u_t) \left(1 + \sup_{0 < s < t} \|u_s\|_{\mathbf{H}}^{q'+1}\right) \|\xi\|_{\mathbf{H} \times T_v P\mathbb{T}^d} \\ &\leq \|\psi\|_{C_V^1} \mathbf{E} \exp\left(C \int_0^t \|u_s\|_{H^r} ds\right) \sup_{0 < s < t} V_{\beta+q'+1, \eta}(u_s) \|\xi\|_{\mathbf{H} \times T_v P\mathbb{T}^d}. \end{aligned}$$

Next, choosing T_0 sufficiently large such that

$$(\beta + q' + 1)e^{-\gamma T_0} < \beta,$$

Lemma 3.7 implies $\forall t \geq T_0, \exists C = C(t, \beta, \eta)$,

$$|D\hat{P}_t^p \psi \xi| \leq C(t, \beta, \eta) \|\psi\|_{C_V^1} V(u) \|\xi\|_{\mathbf{H} \times T_v P\mathbb{T}^d}.$$

This proves (a).

Next, consider (b). First, observe that for $x \geq 0$, $|e^{px} - 1| \leq |p|e^{(|p|+1)x}$, and hence

$$\begin{aligned} |\hat{P}_t^p \psi - \hat{P}_t \psi| &\leq \|\psi\|_{C_V} \mathbf{E} \left| \exp\left(-p \int_0^t H(z_s) ds\right) - 1 \right| V(z_t) \\ &\leq |p| \|\psi\|_{C_V} \mathbf{E} \exp\left((1 + |p|) \int_0^t \|\nabla u_s\|_{L^\infty} ds\right) V(z_t), \end{aligned}$$

and hence convergence in C_V holds by Lemma 3.7. For $\xi \in \mathbf{H} \times T_v P\mathbb{T}^d$,

$$\begin{aligned} |D\hat{P}_t^p \psi \xi - D\hat{P}_t \psi \xi| &\leq \mathbf{E} \left| \exp\left(-p \int_0^t H(z_s) ds\right) - 1 \right| |D\psi(z_t) J_{0,t} \xi| \\ &\quad + |p| \mathbf{E} \left| \int_0^t DH(z_s) J_{0,s} \xi ds \right| \exp\left(-p \int_0^t H(z_s) ds\right) |\psi(z_t)|. \end{aligned}$$

Convergence then holds by the same arguments used to prove boundedness for $t \geq T_0$ combined with that used to prove convergence in C_V . $\square$

As discussed in Section 2, we need to work with the spaces $\mathring{C}_V$ and $\mathring{C}_V^1$ which are, respectively, the C_V -closure and C_V^1 -closure of the space of smooth ‘cylinder functions’ $\mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$ (see (2.5)).

LEMMA 5.3. *For all β sufficiently large and all $\eta \in (0, \eta^*)$ there holds the following:*

- (a) *For all $t > 0$ and $p \in \mathbb{R}$, $\hat{P}_t^p : C_V \rightarrow C_V$ is bounded and $\hat{P}_t^p(\mathring{C}_V) \subset \mathring{C}_V$.*
- (b) *For T_0, p_0 as in Lemma 5.2, $\forall p \in [-p_0, p_0]$, we have that $\hat{P}_{T_0}^p(\mathring{C}_V^1) \subset \mathring{C}_V^1$.*

PROOF. Consider part (a) first. Boundedness in C_V was proved at the beginning of Lemma 5.2. To check $\hat{P}_t^p(\mathring{C}_V) \subset \mathring{C}_V$, note that since $\hat{P}_t^p$ is a continuous linear operator on $C_V \rightarrow C_V$ and $\mathring{C}_V$ is a closed subspace of C_V it suffices to prove that $\hat{P}_t^p$ maps a dense set of $\mathring{C}_V$ into $\mathring{C}_V$. To do so, we show that for all $\psi \in \mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$,

$$(5.2) \quad \lim_{n \rightarrow \infty} \|\hat{P}_t^p \psi - (\hat{P}_t \psi) \circ \Pi_n\|_{C_V} = 0,$$

where recall that $\Pi_n : \mathbf{H} \times P\mathbb{T}^d \rightarrow \mathbf{H}_n \times P\mathbb{T}^d$ is the projection onto Fourier modes satisfying $|k| \leq n$ (note that for any $\varphi \in C_V$, $\varphi \circ \Pi_n \in \mathring{C}_V$). With this in mind, denote $\hat{z} = (u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$ and $\hat{z}^n = \Pi_n \hat{z}$, and let Φ_t denote the random flow for the projective process on $\mathbf{H} \times P\mathbb{T}^d$ with $\hat{z}_t = \Phi_t(\hat{z})$, $\hat{z}_t^n = \Phi_t(\hat{z}^n)$. A direct calculation reveals, using that $\psi \in \mathring{C}_0^\infty$,

$$\begin{aligned} & |\hat{P}_t^P \psi(\hat{z}) - \hat{P}_t^P \psi(\hat{z}^n)| \\ & \leq \mathbf{E} \left| \exp\left(-p \int_0^t H(\hat{z}_s) ds\right) \psi(\hat{z}_t) - \exp\left(-p \int_0^t H(\hat{z}_s^n) ds\right) \psi(\hat{z}_t^n) \right| \\ & \lesssim_\psi \mathbf{E} \Gamma(t) \sup_{s \in [0, t]} d_{\mathbf{W} \times P\mathbb{T}^d}(\hat{z}_s, \hat{z}_s^n), \end{aligned}$$

where we have defined for $r \in (\frac{d}{2} + 1, 3)$,

$$\Gamma(t) := \exp\left(\int_0^t \|u_s\|_{H^r} + \|u_s^n\|_{H^r} ds\right) \sup_{s \in [0, t]} (1 + \|u_s\|_{\mathbf{H}} + \|u_s^n\|_{\mathbf{H}})$$

and $d_{\mathbf{W} \times P\mathbb{T}^d}(\hat{z}_1, \hat{z}_2)$ is the metric on $\mathbf{W} \times P\mathbb{T}^d$. In particular, the fact that $\psi \in \mathring{C}_0^\infty$ allowed to exchange $\mathbf{H} \times P\mathbb{T}^d$ for $\mathbf{W} \times P\mathbb{T}^d$. Following an analysis similar to the proof of the Jacobian estimate in Lemma 3.11, we find that $d_{H^r \times P\mathbb{T}^d}(\hat{z}_t, \hat{z}_t^n)$ satisfies the pathwise bound

$$(5.3) \quad d_{H^r \times P\mathbb{T}^d}(\hat{z}_t, \hat{z}_t^n) \lesssim \Gamma(t)^{c_1} d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n),$$

for some constant $c_1 > 1$.

A straightforward extension of Lemma 3.7 to the joint system $(\hat{z}_t, \hat{z}_t^n)$ gives for each $t > 0$,

$$(5.4) \quad \mathbf{E} \sup_{s \in [0, t]} \Gamma(t)^{1+c_1} \lesssim V_{\beta', \eta}(u),$$

for β' sufficiently large and for all $0 < \eta < \eta^*$. Therefore, for $V = V_{\beta'+1, \eta}$, we obtain

$$\|\hat{P}_t^P \psi - \hat{P}_t^P \psi \circ \Pi_n\|_{C_V} \lesssim \sup_{\hat{z} \in \mathbf{H} \times P\mathbb{T}^d} \frac{d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n)}{1 + \|u\|_{\mathbf{H}}^2} \lesssim \sup_{u \in \mathbf{H}} \frac{\|(\text{Id} - \Pi_n)u\|_{H^r}}{1 + \|u\|_{\mathbf{H}}^2}$$

Note that $\|(\text{Id} - \Pi_n)u\|_{H^r} \lesssim n^{r-\sigma} \|u\|_{\mathbf{H}}$, which completes the proof of (5.2).

Turning to part (b): by part (a), it suffices to prove the following for all $\psi \in \mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$:

$$\lim_{n \rightarrow \infty} \|\hat{P}_{T_0}^P \psi - (\hat{P}_{T_0}^P \psi) \circ \Pi_n\|_{C_V^1} = 0.$$

As above, let $\hat{z}^n = \Pi_n \hat{z}$ and $\hat{z}_t^n = \Phi_t(\hat{z}^n)$. It what follows we need to measure the difference between Jacobians of these different trajectories, namely $D_{\xi} \hat{z}_t^n =: J_t^n \xi \in \mathbf{H} \times T_{v_t^n} P\mathbb{T}^d$ and $D_{\xi} \hat{z}_t =: J_t \xi \in \mathbf{H} \times T_{v_t} P\mathbb{T}^d$. These Jacobians map onto different tangent spaces, but on the event $E_t^n = \{d_{P^{d-1}}(v_t, v_t^n) \leq 1/50\}$ we can place v_t, v_t^n in the same smooth chart and identify their tangent spaces with a copy of $\mathbb{R}^{d-1}$. Thus, on E_t^n we can always make sense of expressions like $\|J_t \xi - J_t^n \xi\|_{H^s}, \|J_t - J_t^n\|_{H^s}$.

To estimate the distance between J_t, J_t^n , we use the following estimate on the second variation (itself a consequence of the estimates in Section 3.3): for some $C, q'' > 0$,

$$\|J_{0,t}^{(2)}[\xi, \zeta]\|_{H^r} \lesssim \exp\left(C \int_0^t \|u\|_{H^r} ds\right) \left(1 + \sup_{0 < s < t} \|u\|_{H^r}^{q''}\right) \|\xi\|_{H^r} \|\zeta\|_{H^r}.$$

On the event E_t^n , this implies the estimate

$$(5.5) \quad \|J_t - J_t^n\|_{H^r} \lesssim \Gamma(t)^{c_2} d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n),$$

for some $c_2 > 1$. Then, similar to part (a), for some $r \in (\frac{d}{2} + 1, 3)$,

$$\begin{aligned} & |D\hat{P}_t^p \psi(\hat{z})\xi - D\hat{P}_t^p \psi(\hat{z}^n)\xi| \\ & \leq \left| \mathbf{E} \left(\exp \left(-p \int_0^t H(\hat{z}_s) ds \right) D\psi(\hat{z}_t) J_{0,t} \xi - \exp \left(-p \int_0^t H(\hat{z}_s^n) ds \right) D\psi(\hat{z}_t^n) J_{0,t}^n \xi \right) \right| \\ & \quad + \left| \mathbf{E} \left(p\psi(\hat{z}_t) \exp \left(-p \int_0^t H(\hat{z}_s) ds \right) \int_0^t DH(\hat{z}_s) J_{0,s} \xi ds \right. \right. \\ & \quad \left. \left. - p\psi(\hat{z}_t^n) \exp \left(-p \int_0^t H(\hat{z}_s^n) ds \right) \int_0^t DH(\hat{z}_s^n) J_{0,s}^n \xi ds \right) \right| \end{aligned}$$

We split the above into expectations on E_t^n , $(E_t^n)^c$. The integrand on E_t^n can be bounded using (5.3) and (5.5), resulting in a bound $\lesssim \Gamma(t)^{c_3} d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n)$ for some $c_3 > 0$. On $(E_t^n)^c$, we can bound the integrand as in Lemma 3.11, while the pathwise estimate (5.3) and the bound (5.4) yields

$$|D\hat{P}_t^p \psi(\hat{z})\xi - D\hat{P}_t^p \psi(\hat{z}^n)\xi| \lesssim \mathbf{E}_{\hat{z}, \hat{z}^n} \Gamma^{c_4}(t) d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n) + V_{\beta, \eta}(u) d_{H^r \times P\mathbb{T}^d}(\hat{z}, \hat{z}^n)$$

for some sufficiently large $\beta > 0$. From here, the proof proceeds as in part (a) (possibly after increasing β further). $\square$

We are now ready to prove Proposition 2.16.

PROOF OF PROPOSITION 2.16. By Lemma 5.2, the operator $\hat{P}_{T_0}$ has a dominant, simple eigenvalue at 1 (as an operator $C_V^1 \rightarrow C_V^1$). Let $r_0 \in (0, 1)$ such that $\sigma(\hat{P}_{T_0}) \setminus \{1\} \subset B_{r_0}(0)$, where $r_0 \in (0, 1)$. Fix $\epsilon \ll 1 - r_0$. By Lemma 5.1 and the fact that $\hat{P}_{T_0}^p \rightarrow \hat{P}_{T_0}$ in the operator norm on $\dot{C}_V^1$, it follows that for all $|p|$ sufficiently small, we have that $\sigma(\hat{P}_{T_0}^p) \subset B_{r_0+\epsilon}(0) \cup B_\epsilon(1)$. Taking $|p|$ sufficiently small, Lemma 5.1(b) implies that the spectral projector π_p for $\hat{P}_{T_0}^p$ corresponding to $\sigma(\hat{P}_{T_0}^p) \cap B_\epsilon(1)$ is close in operator norm to the rank-one spectral projector π_0 for $\hat{P}_{T_0}$ corresponding to $\{1\}$. Since rank-one projection operators are an open set in the space of bounded linear operators on $\dot{C}_V^1$, it follows that π_p is rank-one when $|p| \ll 1$.

We conclude that for each such p , there is a unique, simple eigenvalue $\lambda_p \in B_\epsilon(1)$. To show that this eigenvalue is real, note that the complexification of the operator $\hat{P}_{T_0}^p$ sends real parts of functions to real parts and imaginary parts to imaginary parts. Thus, if $\Im(\lambda_p) \neq 0$ then the complex conjugate $\overline{\lambda_p}$ would also be an eigenvalue. This contradicts the fact that the spectral projector π_p is rank-one. We conclude that λ_p is real, positive and coincides with the spectral radius $\rho(\hat{P}_{T_0}^p)$. The value $\Lambda(p)$ is defined so that $\lambda_p = e^{-T_0 \Lambda(p)}$.

Finally, convergence of the limit formula (2.10) follows from the standard Gelfand formula and that $\pi_p \mathbf{1} \neq 0$ (by Lemma 5.1 and $\pi_0 \mathbf{1} = \mathbf{1}$). $\square$

5.2. Proof of Proposition 2.17. We next prove Proposition 2.17, namely that for all $t > 0$, $\hat{P}_t^p$ has a spectral gap on $\dot{C}_V$ and that ψ_p is an eigenfunction for the dominant eigenvalue for all $t \geq 0$. As in the proof of Proposition 2.16 we will make a spectral perturbation argument. We will need the following lemma, which is a simple variation of Lemma 5.2 and is proved in the same way.

LEMMA 5.4. *There exists $p_0 > 0$ such that the following holds for all $p \in [p_0, p_0]$, all β sufficiently large and all $\eta \in (0, \eta^*)$:*

- (a) *For all $t > 0$, we have that $\hat{P}_t^p$ is a bounded linear operator on C_V .*
- (b) *For each $t > 0$ fixed, we have $\lim_{p \rightarrow 0} \|\hat{P}_t^p - \hat{P}_t\|_{C_V} = 0$.*

Next, let us examine the problem of finding a suitable function-space framework for which $t \mapsto \hat{P}_t^p$ is a C_0 semigroup. As it turns out, to check C_0 continuity in C_V it does not suffice to check boundedness in C_V , since it does not admit the countable dense subset of smooth functions we need to make a strong continuity argument. Instead, we will follow the approach carried out in [43] and prove C_0 continuity on $\mathring{C}_V$, which provides a natural, separable, closed subspace of C_V .

PROPOSITION 5.5. *For all $p \in [-p_0, p_0]$ all $\beta \geq 1$ sufficiently large and all $0 < \eta < \eta^*$, $\hat{P}_t^p$ extends to a C_0 -semigroup on $\mathring{C}_V$ for $V = V_{\beta, \eta}$. That is, $\hat{P}_t^p(\mathring{C}_V) \subset \mathring{C}_V$ for all $t > 0$, and $t \mapsto \hat{P}_t^p$ is a C_0 semigroup on $\mathring{C}_V$.*

PROOF. Following [43] Theorem 5.10, it is sufficient to show that $\hat{P}_t^p$ maps $\mathring{C}_V$ into itself and that $t \mapsto \hat{P}_t^p$ is strongly continuous in $\mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$ in the C_V topology. The first step, $\hat{P}_t^p(\mathring{C}_V) \subset \mathring{C}_V$ for all $t > 0$, is proved in Lemma 5.3, hence it suffices to check the strong continuity.

To do this, we fix $\psi \in \mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$. Let $\mathcal{K} \subset \mathbb{K}$ be the finite set such that we can write $\psi = \varphi \circ \Pi_{\mathcal{K}}$ for some function $\varphi \in C_0^\infty$. By Itô's formula, it follows that for each $r \in (1 + d/2, 3)$ there holds for $t \in (0, 1]$,

$$\begin{aligned} |\hat{P}_t^p \psi(\hat{z}) - \psi(\hat{z})| &\lesssim_{\psi} t \mathbf{E} \exp \left(\int_0^t \|u_s\|_{H^r} ds \right) \sup_{s \in [0, t]} (1 + \|\Pi_{\mathcal{K}} u_s\|_{\mathbf{H}} + \|F(\Pi_{\mathcal{K}} \hat{z}_s)\|_{\mathbf{H} \times P\mathbb{T}^d}) \\ &\lesssim_{\psi, n} t V_{1, \eta}(u), \end{aligned}$$

where we used the fact that $\|F(\Pi_{\mathcal{K}} \hat{z}_s)\|_{\mathbf{H} \times P\mathbb{T}^d} \lesssim_{\mathcal{K}} \|u_s\|_{\mathbf{H}}^2$ and used Lemma 3.7. Putting this together, we conclude that for $V = V_{\beta, \eta}$, $\beta \geq 3/2$, we have

$$\|\hat{P}_t^p \psi - \psi\|_{C_V} \lesssim_{\psi, n} t \rightarrow 0 \quad \text{as } t \rightarrow 0.$$

By density of $\mathring{C}_0^\infty(\mathbf{H} \times P\mathbb{T}^d)$ in $\mathring{C}_V$, we conclude strong continuity of $\hat{P}_t^p$ in $\mathring{C}_V$. $\square$

We are now ready to prove Proposition 2.17.

PROOF OF PROPOSITION 2.17. Let $\Lambda(p)$ be as in Proposition 2.16 and let $s(p) < e^{-T_0 \Lambda(p)}$ be such that $\sigma(P_{T_0}^p) \setminus \{e^{-T_0 \Lambda(p)}\} \subset B_{s(p)}(0)$ as an operator on C_V^1 .

Lemmas 5.1 and 5.4, together with Proposition 4.5, imply that for all $\epsilon > 0$, for all p sufficiently small, there are $0 < \tilde{s}(p) < e^{T_0 \tilde{\Lambda}(p)}$ such that $e^{T_0 \tilde{\Lambda}(p)} \in B_\epsilon(1)$ and $\sigma(P_{T_0}^p) \setminus \{e^{-T_0 \tilde{\Lambda}(p)}\} \subset B_{\tilde{s}(p)}(0)$ as an operator on C_V . Since $\psi_p \in \mathring{C}_V^1 \subset \mathring{C}_V$ is already an eigenfunction for $\hat{P}_{T_0}^p$, we conclude that in fact $\tilde{\Lambda}(p) = \Lambda(p)$ for all p sufficiently small.

To complete the proof of Proposition 2.17, we establish the spectral picture for $\hat{P}_t^p$ for all $t > 0$ using semigroup theory. To start, the spectral mapping theorem for the point spectrum ([4] A-III Theorem 6.3) implies that $-\Lambda(p)$ is an eigenvalue of the infinitesimal generator $\mathcal{A}^p$ of $\hat{P}_t^p$. Corollary 6.4 in Chapter A-III of [4] implies that for all $\eta \in \mathbb{C}$,

$$(5.6) \quad \ker(\eta \text{Id} - \mathcal{A}^p) = \bigcap_{s \geq 0} \ker(e^{s\eta} \text{Id} - \hat{P}_s^p).$$

Applying (5.6) to $\eta = -\Lambda(p)$, we have that ψ_p is (up to rescaling) the unique eigenvector for $\mathcal{A}^p$ for $-\Lambda(p)$; by another application of (5.6), we conclude that $\hat{P}_t^p \psi_p = e^{-\Lambda(p)t} \psi_p$ for all $t > 0$. $\square$

REMARK 5.6. Recall that $\mathbf{H}$ is defined to be the space of divergence free, mean-zero velocity fields in H^σ , where σ is drawn from the interval $(\alpha - 2(d-1), \alpha - \frac{d}{2})$ (see Section 1.1 for notation; cf. Propositions 4.1, 4.2). In particular, the preceding arguments can be repeated with σ replaced by any other $\sigma' < \sigma$ in this interval: with $\hat{P}_{T_0}^p$, $|p| \ll 1$ regarded as an operator on the corresponding C_V^1 -space, we obtain a corresponding dominant eigenfunction ψ'_p , defined on $H^{\sigma'}$ velocity fields and continuous in the $H^{\sigma'}$ topology. By parabolic regularization, ψ'_p coincides with ψ_p when restricted to the domain $\mathbf{H} \times P\mathbb{T}^d$.

5.3. *Properties of ψ_p : Proof of Proposition 2.11.* It remains to establish the desired properties of ψ_p and $\Lambda(p)$, namely:

- (A) Uniform positivity of ψ_p on bounded subsets of $\mathbf{H} \times P\mathbb{T}^d$; and
- (B) The asymptotic $\Lambda(p) = p\lambda_1 + o(p)$, $p \rightarrow 0$, where $\lambda_1 > 0$ is the top Lyapunov exponent of the Lagrangian flow ϕ^t .

First, we prove strict positivity.

LEMMA 5.7. *For any $R > 0$, we have that*

$$(5.7) \quad \inf_{\|u\|_{\mathbf{H}} \leq R, (x,v) \in P\mathbb{T}^d} \psi_p(u, x, v) > 0.$$

PROOF. The nonnegative cone $C_{V,+}^1 = \{\psi \in C_V^1 : \psi \geq 0\}$ is closed in C_V^1 ; since $\hat{P}_{nT_0}^p C_{V,+}^1 \subset C_{V,+}^1$ and $e^{nT_0\Lambda(p)} \hat{P}_{nT_0}^p \mathbf{1} \rightarrow \psi_p$ as $n \rightarrow \infty$ (Proposition 2.16), we conclude $\psi_p \geq 0$.

To verify that $\psi_p > 0$ pointwise, we use the fact that the Markov semigroup $\hat{P}_t$ is topologically irreducible (Proposition 4.2). To wit, assume for the sake of contradiction that $\psi_p(z) = 0$ for some point $\hat{z} = (u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$. Since ψ_p is continuous and not identically equal to zero, it holds that $U_p := \{\psi_p > 0\} \subset \mathbf{H} \times P\mathbb{T}^d$ is nonempty and open. Therefore,

$$\psi_p(\hat{z}) \geq \mathbf{P}_{\hat{z}}(\hat{z}_t \in U_p) \mathbf{E}_{\hat{z}}(|D_x \phi_u^t v|^{-p} \psi_p(\hat{z}_t) | \hat{z}_t \in U_p) > 0.$$

Note that the same arguments apply to ψ'_p corresponding to $\sigma' < \sigma$ as in Remark 5.6, hence $\psi'_p > 0$ pointwise as well.

To conclude (5.7), assume for the sake of contradiction that there is a bounded sequence $\{\hat{z}^n = (u^n, x^n, v^n)\} \subset \mathbf{H} \times P\mathbb{T}^d$ for which $\psi_p(\hat{z}^n) \rightarrow 0$. With $\sigma' < \sigma$ fixed as in Remark 5.6, let $\hat{z}^{n'}$ be a subsequence converging in $H^{\sigma'}$ to some $\hat{z}^* \in \mathbf{H}^{\sigma'} \times P\mathbb{T}^d$. Since ψ'_p is $H^{\sigma'}$ -continuous and coincides with ψ_p on H^σ , it follows that $\psi'_p(\hat{z}^*) = 0$, contradicting the pointwise positivity established earlier. $\square$

Item (B) is a version of classical results in the ergodic theory of stochastic differential equations in finite dimensions. In that literature (see, e.g., the survey [7]), the value $-\Lambda(p)$ is referred to as the *moment Lyapunov exponent*. This terminology is justified by the following:

LEMMA 5.8. *For all $(u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$, we have*

$$\Lambda(p) = - \lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbf{E} |D_x \phi_u^t v|^{-p}.$$

Moreover, the above limit is uniform on bounded subsets of $\mathbf{H} \times P\mathbb{T}^d$.

PROOF. The argument of the logarithm on the right-hand side is equal to $\hat{P}_t^p \mathbf{1}$ evaluated at (u, x, v) . Uniform convergence on bounded subsets of $\mathbf{H} \times P\mathbb{T}^d$ now follows from the C_V limit $\psi_p = \lim_{t \rightarrow \infty} e^{t\Lambda(p)} \hat{P}_t^p \mathbf{1}$ (a consequence of the C_V spectral gap) and (5.7). $\square$

Next, we verify (B) by relating the value $\Lambda(p)$ with the Lyapunov exponent λ_1 of the Lagrangian flow ϕ^t . The first step is to show that $\Lambda(p)$ is in fact differentiable.

LEMMA 5.9. *The function $p \mapsto \Lambda(p)$ is differentiable in a neighborhood of $p = 0$.*

PROOF. Let us first show that $p \mapsto \hat{P}_1^p$ is Fréchet differentiable as an operator-valued function in C_V . Formally, we expect the derivative $\frac{d}{dp} \hat{P}_1^p$ to be given by

$$\frac{d}{dp} \hat{P}_1^p \psi(u, x, v) = -\mathbf{E}_{(u,x,v)}(\log |D_x \phi_u^1 v| \cdot |D_x \phi_u^1 v|^{-p} \psi(u_1, x_1, v_1)).$$

A direct application of Lemma 3.7 implies that the right-hand side defines a bounded linear operator on C_V . Fréchet differentiability of $p \mapsto \hat{P}_1^p$ follows if

$$\lim_{h \rightarrow 0} \left\| \mathbf{E} \left(\frac{1}{h} (|D_x \phi^1 v|^{-p-h} - |D_x \phi^1 v|^{-p}) + \log |D_x \phi^1 v| \cdot |D_x \phi^1 v|^{-p} \right) V(u_1) \right\|_{C_V} = 0.$$

By the mean value theorem, this is justified as long as

$$\|\mathbf{E} V(u_t) |D_x \phi^1 v|^{-p} \log^2 |D_x \phi^1 v|\|_{C_V} < \infty,$$

which follows from Lemma 3.7.

Having established that $p \mapsto \hat{P}_1^p$ is Fréchet differentiable, it now follows from the spectral gap for $\hat{P}_1^p$ in C_V and the standard contour integral formula for spectral projectors that $p \mapsto \pi_p$ is likewise Fréchet differentiable, hence $p \mapsto \psi_p = \pi_p(\mathbf{1})$ is also Fréchet differentiable. Fixing $\hat{z}^* \in \mathbf{H} \times P\mathbb{T}^d$, we may now express

$$e^{-\Lambda(p)} = \frac{1}{\psi_p(\hat{z}^*)} \hat{P}_1^p \psi_p(\hat{z}^*).$$

Since the right-hand side is a ratio of differentiable functions and the denominator is nonvanishing, we conclude $p \mapsto \Lambda(p)$ is differentiable. $\square$

LEMMA 5.10. *Let $p_0 > 0$ be as in Proposition 2.17.*

- (a) *The mapping $p \mapsto \Lambda(p)$ is convex on $[-p_0, p_0]$.*
- (b) *We have*

$$\Lambda'(0) = \lambda_1.$$

In particular, since $\lambda_1 > 0$, we have that $\Lambda(p) > 0$ for all $p > 0$ sufficiently small.

With our preparations in place, the proof of Lemma 5.10 follows from straightforward versions of standard arguments; see, for example, [5, 7]. We sketch the proof below for the sake of completeness.

PROOF. For convexity, let $p, q \in (-p_0, p_0)$ and $\lambda \in [0, 1]$ and fix an arbitrary $(u, x, v) \in \mathbf{H} \times P\mathbb{T}^d$. By Hölder's inequality,

$$\mathbf{E} |D_x \phi_u^t v|^{-\lambda p - (1-\lambda)q} \leq (\mathbf{E} |D_x \phi_u^t v|^{-p})^\lambda \cdot (\mathbf{E} |D_x \phi_u^t v|^{-q})^{1-\lambda}$$

Convexity follows on taking the log of both sides, dividing by t , taking $t \rightarrow \infty$ and applying Lemma 5.8.

Next, for all $p \in (-p_0, p_0)$, we have by Jensen's inequality that

$$\mathbf{E}|D_x \phi_u^t v|^{-p} = \mathbf{E}e^{-p \log |D_x \phi_u^t v|} \geq \exp(-p \mathbf{E} \log |D_x \phi_u^t v|).$$

Taking the log of both sides, dividing by t and taking $t \rightarrow \infty$ results in the inequality

$$\Lambda(p) \leq p\lambda_1.$$

In particular, $\Lambda(p)/p \geq \lambda_1$ for $p \in (-p_0, 0)$ and $\Lambda(p)/p \leq \lambda_1$ for $p \in (0, p_0)$. By convexity, the left- and right-hand derivatives $\Lambda'(0+)$ and $\Lambda'(0-)$ exist; the inequalities above imply that $\Lambda'(0-) \geq \lambda_1$ and $\Lambda'(0+) \leq \lambda_1$. By differentiability, these values coincide. This completes the proof. $\square$

6. Geometric ergodicity for the two point motion. We now turn to our study of the two point Lagrangian motion (u_t, x_t, y_t) . Recall that that given two initial points $(x, y) \in \mathcal{D}^c = \{(x, y) \in \mathbb{T}^d \times \mathbb{T}^d : x \neq y\}$, (x_t, y_t) are defined by

$$x_t = \phi^t(x), \quad y_t = \phi^t(y).$$

This induces a Feller Markov semigroup $P_t^{(2)}$ defined on bounded measurable $\varphi : \mathbf{H} \times \mathcal{D}^c \rightarrow \mathbb{R}$ by

$$P_t^{(2)} \varphi(u, x, y) := \mathbf{E}_{(u, x, y)} \varphi(u_t, x_t, y_t).$$

We eventually apply Theorem 2.4 to $P_t^{(2)}$ to prove Theorem 2.2. As discussed in Section 2, Conditions 1 and 2 follow from Propositions 2.6 and 2.7 (strong Feller and topological irreducibility), which we prove below. This shows that $\mu \times \text{Leb} \times \text{Leb}$ is the unique stationary measure for $P_t^{(2)}$ (see [25]). Similarly, Propositions 2.6 and 2.7 imply equivalence of transition kernels.

LEMMA 6.1. *The family of transition kernels $\{P_t^{(2)}((u, x, y), \cdot) : t > 0, (u, x, y) \in \mathbf{H} \times \mathcal{D}^c\}$ are equivalent measures.*

As we saw in the proof of Lemma 4.4, equivalence of the transition kernels implies the following.

LEMMA 6.2. *Let $K \subset \mathbf{H} \times \mathcal{D}^c$ be any compact set with $(\mu \times \text{Leb} \times \text{Leb})(K) > 0$. Then for arbitrary $t > 0$ and $R_1, R_2 > 0$, we have*

$$\inf_{\|u\| \leq R_1, d(x, y) \geq R_2} P_t^{(2)}((u, x, y), K) > 0.$$

Therefore, Conditions 3 and 4 follow from Proposition 2.14, which we prove below, and the fact that the Lyapunov function $\mathcal{V}$ is *coercive* in the sense that for each $r > 0$, we can always find $R_1, R_2 > 0$ such that

$$\begin{aligned} & \{(u, x, v) \in \mathbf{H} \times \mathcal{D}^c : \mathcal{V}(u, x, y) \leq r\} \\ & \subset \{(u, x, y) \in \mathbf{H} \times \mathcal{D}^c : \|u\| \leq R_1, \text{ and } d(x, y) \geq R_2\}. \end{aligned}$$

The remainder of this section will be focused on proving Propositions 2.6 and 2.7 and Theorem 2.14.

6.1. Strong Feller property: Proof of Proposition 2.6.

6.1.1. Uniform parabolic Hörmander conditions. Due to Assumption 1, we are only concerned with spreading the noise to the degrees of freedom on $\mathcal{D}^c$. Accordingly, denote the vector field

$$X_m(x,y)=\begin{pmatrix}e_m(x)\\e_m(y)\end{pmatrix}\in T_{(x,y)}\mathcal{D}^c\simeq\mathbb{R}^{2d}.$$

In order to prove Proposition 2.6, we use the following uniform spanning. It is helpful to observe that vector fields in the set $\mathcal{A}_2$ in the statement below can be identified as incompressible velocity fields; the uniform spanning condition is simply a quantitative statement about the infinitesimal controllability, that is, that the particles can be nudged in any direction by velocity fields accessible via the noise.

LEMMA 6.3. *There holds the following uniform spanning condition: for all $x,y\in\mathbb{T}^d$ with $x\neq y$, if we denote the set of unit vectors in the span of the available vector fields:*

$$\mathcal{A}_2=\Big\{\sum_{\substack{m=(k,i)\in\mathbb{K} \\ |k|\leq 2}}c_mX_m:\max_m|c_m|\leq 1\Big\},$$

then

(6.1)
$$\inf_{h\in\mathbb{R}^{2d}:|h|=1}\max_{\Gamma\in\mathcal{A}_2}|\langle\Gamma(x,y),h\rangle_{\mathbb{R}^{2d}}|\gtrsim d(x,y).$$

REMARK 6.4. The proof shows that Assumption 2 is not quite sharp, especially for Galerkin–Navier–Stokes, where hypoellipticity in (u_t) will fill all available degrees of freedom. However, for simplicity of exposition it is easier simply to use the less complicated condition.

PROOF. For definiteness, parameterize $\mathbb{T}^d$ as $(-\pi,\pi]^d$. Note that by Assumption 2 and trigonometric identities, $\mathcal{A}_2$ is translation invariant in the sense that if (6.1) holds at some point (x,y) then (6.1) holds also at $(x',y')=(x+\beta\mod 2\pi\mathbb{Z}^d,y+\beta\mod 2\pi\mathbb{Z}^d)$ for any vector $\beta\in\mathbb{R}^d$.

At any point $(x,y)\in\mathcal{D}^c$, we divide the tangent space into $\mathbb{R}_x^d\oplus\mathbb{R}_y^d$ where the first $\mathbb{R}_x^d$ is associated with infinitesimal motions of x and the second $\mathbb{R}_y^d$ associated with the infinitesimal motions of y . We are able to restrict ourselves to vectors $X\in\mathcal{A}_2$, which vanish on either $\mathbb{R}_x^d$ or $\mathbb{R}_y^d$ as uniform spanning follows by linear combinations (after slightly adjusting the constant). By the above symmetry considerations, it suffices to consider the case $x=0,y\neq 0$ and show that we can uniformly span $\mathbb{R}_y^d$ with vector fields in $\mathcal{A}_2$ that also vanish at zero.

We first consider the case $d=2$. In what follows, we denote $y=(y^{(1)},y^{(2)})$. Let $\delta\in(0,\frac{1}{10})$ be fixed and arbitrary. Define the set of points where shear flows which vanish at $x=0$ cannot span $\mathbb{R}_y^d$.

$$\mathfrak{D}:=\{y\in\mathbb{T}^2:y=(0,a),\text{ or }y=(a,0),\text{ for some }a\in(-\pi,\pi]\}.$$

There are essentially three cases.

Case 1 (2D): y is at least $\delta d(x,y)$ away from $\mathfrak{D}$:

By trigonometric interpolation, $\exists h(\zeta)$ a linear combination of $\cos 2\zeta, \sin 2\zeta, \cos \zeta, \sin \zeta$ with coefficients having absolute value less than one such that satisfies $h(0)=0$

and $h(y^{(2)}) \gtrsim \delta|y|^{(2)}$. Hence, the vector field, $Y(x, y)|_{\mathbb{R}_x^2} = (h(x^{(2)}), 0)$, $Y(x, y)|_{\mathbb{R}_y^2} = (h(y^{(2)}), 0)$,

$$(6.2) \quad \left| \left\langle Y(x, y)|_{\mathbb{R}_y^2}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle_{\mathbb{R}_y^2} \right| \gtrsim_\delta d(x, y).$$

The analogous transverse shear flows span the vertical direction in $\mathbb{R}_y^2$.

Case 2 (2D): y is less than $\delta d(x, y)$ from $\mathfrak{D}$ but more than $\delta\pi$ away from the points $(0, \pi)$, $(\pi, 0)$

Suppose without loss of generality that y is close to the horizontal line $y^{(2)} = 0$. Shear flows span the vertical direction of $\mathbb{R}_y^2$ as in Case 1. To span the horizontal direction, we use the flow

$$(6.3) \quad Y(x, y)|_{\mathbb{R}_{y'}^2} = \begin{pmatrix} -\sin y^{(1)} \cos y^{(2)} \\ \cos y^{(1)} \sin y^{(2)} \end{pmatrix},$$

which gives (6.2).

Case 3 (2D): y is less than $\delta\pi$ away from one of $(0, \pi)$, $(\pi, 0)$: This case is the most difficult. Suppose that y is close to $(\pi, 0)$; the case $(0, \pi)$ is treated analogously. As in Case 2, shear flows span the vertical direction. However, there is a new degeneracy at $(\pi, 0)$ in the horizontal direction, as the cellular flow in (6.3) vanishes. To rectify this, we choose the flow

$$Y(x, y)|_{\mathbb{R}_{y'}^d} = \frac{1}{2} \begin{pmatrix} \cos 2y^{(1)} \cos 2y^{(2)} \\ \sin 2y^{(1)} \sin 2y^{(2)} \end{pmatrix} - \frac{1}{2} \begin{pmatrix} \cos y^{(1)} \cos y^{(2)} \\ \sin y^{(1)} \sin y^{(2)} \end{pmatrix},$$

which satisfies (6.2).

Next, we discuss the extension to $d = 3$.

In this case, we redefine $\mathfrak{D}$ in the analogous way:

$$\mathfrak{D} := \{y \in \mathbb{T}^2 : y = (a, 0, 0), \text{ or } y = (0, a, 0), \text{ or } y = (0, 0, a), \text{ for some } a \in (-\pi, \pi]\}.$$

Case 1 (3D): y is at least $\delta d(x, y)$ away from $\mathfrak{D}$

This case is analogous to the Case 1 above simply by using shear flows in each of the coordinate directions separately.

Case 2 (3D): y is less than $\delta d(x, y)$ away from $\mathfrak{D}$

Suppose without loss of generality that y is close to the horizontal plane. The vertical direction (normal to the plane) is spanned by shear flows. To span the horizontal plane, we use flows that are independent of the normal direction and the problem reduces to the 2D case treated above.

This completes the uniform spanning. $\square$

6.1.2. Sketch of proof of Proposition 2.6. We can essentially apply the same proof as we did in [15] (which draws heavily from [33] combined with some simplifications in the Malliavin calculus and a more sophisticated choice of control to deal with the more complicated nonlinearity).

The strategy here is to regularize the process for large values of $\|u\|_{\mathbf{H}}$. This will be done through the use of an auxiliary Wiener process $Z_t \in \mathbb{R}^{2d}$, which will couple to the dynamics through the vector field L on $\mathbb{R}^{2d}$ defined for each $Z \in \mathbb{R}^{2d}$ by

$$L(Z) := \sum_{j=1}^{2d} \hat{e}_j \frac{Z^j}{\sqrt{1 + |Z^j|^2}},$$

where $\{\hat{e}_j\}_{j=1}^{2d}$ are the canonical basis elements for $\mathbb{R}^{2d}$. The cut-off process $\bar{z}_t^\rho \in \mathbf{H} \times \mathbb{T}^{2d} \times \mathbb{R}^{2d}$ is then defined by augmenting by Z_t so that $\bar{z}_t^\rho = (u_t^\rho, x_t^\rho, y_t^\rho, Z_t)$ satisfies the cut-off equation

$$\partial_t \bar{z}^\rho = F^\rho(\bar{z}_t^\rho) - A \bar{z}_t^\rho + \tilde{Q} \dot{\bar{W}}_t, \quad \bar{z}_0^\rho = \bar{z} \in \mathbf{H} \times \mathcal{D}^c \times \mathbb{R}^{2d},$$

where $\tilde{Q} \dot{\bar{W}}_t = (Q W_t, 0, 0, W_t^Z)$ for W_t^Z a Wiener process on $\mathbb{R}^{2d}$ independent from W_t , and (for F, A , and L suitably extended to vector fields on $\mathbf{H} \times \mathcal{D}^c \times \mathbb{R}^{2d}$), for each $\bar{z} = (z, Z) \in \mathbf{H} \times \mathcal{D}^c \times \mathbb{R}^{2d}$,

$$F^\rho(\bar{z}) := (1 - \chi_{2\rho}(\|u\|_{\mathbf{H}}))F(z) + \chi_\rho(\|u\|_{\mathbf{H}})L(Z),$$

where $\chi_\rho(r) = \chi(r/\rho)$ with $\chi \in C^\infty(\mathbb{R}_+)$ monotone increasing, nonnegative, with $\chi(r) = 0$ for $r \leq 1$ and $\chi(r) = 1$ for $r > 2$.

Let $P_t^{(2),\rho}$ denote the Markov semigroup associated with the cut-off process $(\bar{z}_t^\rho)$. The main step in [15] (and [33]) is to prove the following gradient bound. This is done via Malliavin calculus with a low frequency approximation and short-time perturbation argument to obtain the control. The main ingredient specific to the two-point motion is the uniform parabolic Hörmander condition proved above in Lemma 6.3.

PROPOSITION 6.5. *There exists constants $a, b > 0$ such that for each $\varphi \in C^2(\mathbf{H} \times \mathbb{T}^{2d} \times \mathbb{R}^{2d})$ and each $\bar{z} = (u, x, y, Z) \in \mathbf{H} \times \mathcal{D}^c \times \mathbb{R}^{2d}$, the derivative $DP_t^{(2),\rho} \varphi(\bar{z})$ exists and satisfies for each $\xi \in \mathbf{H} \times \mathbb{R}^{4d}$,*

$$|DP_t^{(2),\rho} \varphi(\bar{z}) \xi| \lesssim_\rho t^{-a} d(x, y)^{-b} (1 + \|u\|_{\mathbf{H}} + |Z|)^b \|\varphi\|_{L^\infty} \|\xi\|_{\mathbf{H} \times \mathbb{R}^{4d}}.$$

Using Proposition 6.5, one can prove the strong Feller property for the noncut-off process z_t using the following metric on $\mathbf{H} \times \mathcal{D}^c$:

$$d_b(z^1, z^2) := \inf_{\gamma: z^1 \rightarrow z^2} \int_0^1 d(x_s, y_s)^{-b} (1 + \|u_s\|_{\mathbf{H}})^b \|\dot{\gamma}_s\|_{\mathbf{H} \times \mathbb{R}^{2d}} \, ds,$$

where the infimum is taken over all differentiable curves $[0, 1] \ni t \mapsto \gamma_t = (u_t, x_t, y_t)$ in $\mathbf{H} \times \mathcal{D}^c$ connecting z^1 and z^2 . It is not hard to see that the metric $d_b(\cdot, \cdot)$ generates the $\mathbf{H} \times \mathcal{D}^c$ topology since the extremal trajectories avoid the diagonal.

SKETCH OF PROOF OF PROPOSITION 2.6. Fix $t > 0$ and $\epsilon > 0$ and let $z^1, z^2 \in \mathbf{H} \times \mathcal{D}^c$ and take the initial $Z = 0$. From the moment estimates on (u_t) in Proposition 3.1, we can choose the cut-off large enough (depending on $\|\varphi\|_{L^\infty}$ and $\|u_1\|_{\mathbf{H}}$ and $\|u_2\|_{\mathbf{H}}$) such that (see [15] for more detail),

$$|P_t^{(2)} \varphi(z^1) - P_t^{(2)} \varphi(z^2)| \leq |P_t^{(2),\rho} \varphi(z^1, 0) - P_t^{(2),\rho} \varphi(z^2, 0)| + 2\epsilon.$$

By Proposition 6.5 and minimizing along all curves connecting $(z^1, 0)$ and $(z^2, 0)$, one derives

$$|P_t^{(2)} \varphi(z_1) - P_t^{(2)} \varphi(z_2)| \lesssim_\rho t^{-a} d_b(z_1, z_2) + 2\epsilon.$$

The proof is completed by taking $d_b(z_1, z_2)$ sufficiently small. $\square$

6.2. *Irreducibility: Proof of Proposition 2.7.* In this section, we prove Proposition 2.7, that is, we show that the transition kernel $P_t^{(2)}(z, \cdot)$ is locally positive on $\mathbf{H} \times \mathcal{D}^c$ for $t > 0$ and $z \in \mathbf{H} \times \mathcal{D}^c$. Proposition 2.7 is an immediate consequence of the following lemma.

LEMMA 6.6. *Let $z, z' \in \mathbf{H} \times \mathcal{D}^c$ be arbitrary. Then $\forall \epsilon > 0$ and $\forall t > 0$,*

$$P_t^{(2)}(z, B_\epsilon(z')) > 0,$$

where we denote $B_\epsilon(z')$ the ϵ -ball in $\mathbf{H} \times \mathcal{D}^c$.

As usual, Lemma 6.6 is proved via an approximate control argument. Consider the following for g_t a deterministic control:

$$\begin{aligned} \partial_t u_t + B(u_t, u_t) + Au_t &= Qg_t, \\ \partial_t x_t &= u_t(x_t), \\ \partial_t y_t &= u_t(y_t). \end{aligned}$$

We prove that for any z, z' as in the statement of Lemma 6.6, we construct a $g_t \in \mathbf{L}^2$ such that at time t ,

$$\|u' - u_t\|_{H^\sigma} + d(x', x_t) + d(y', y_t) < \frac{\epsilon}{2}.$$

Moreover, $Qg_t \in C^\infty$ and the size of $\|Qg_t\|_{L_t^\infty H^\sigma}$ will depend only $d(x, y)$, $d(x', y')$, u , u' (where we denote $z = (u, x, y)$, $z' = (u', x', y')$) and can be chosen uniformly over compact sets in $\mathbf{H} \times \mathcal{D}^c$. Local positivity of the Wiener measure together with a stability argument then implies Lemma 6.6; see, for example, [Lemma 7.3 [15]] for how to carry out such details.

Constructing g_t is a three step procedure: use a “scaling” (see, e.g., the discussion in [38] and the references therein) to force u_t to (approximately) zero in an arbitrarily short time-window. Then we use arguments involving well-chosen sequences of shear flows to exactly control the two particles to the desired locations (here the proof is vaguely reminiscent of that of Lemma 6.3). Then we again use a “scaling” to force the (u_t) to (approximately) u' while simultaneously not disturbing the particles by more than $O(\epsilon)$.

LEMMA 6.7. *Let $u \in H^\sigma$ be arbitrary. Then $\forall \epsilon > 0$, $\exists \delta < \epsilon$ and a control $g : [0, \delta] \rightarrow \mathbf{L}^2$ such that $\|u_\delta\|_{H^\sigma} \leq \frac{\epsilon}{4}$ and $\sup_{0 \leq t \leq \delta} \|u_t\|_{H^\sigma} \leq 3\|u\|_{H^\sigma}$.*

REMARK 6.8. This lemma is simplified by the use of fully nondegenerate noise; however, using the methods in [38] and the references therein, one can obtain essentially the same lemma from any noise that satisfies the Hörmander bracket conditions for (u_t) discussed in [32, 65].

PROOF. The lemma follows by choosing g_t as the following for suitably chosen δ , N :

$$g_t = -Q^{-1}\delta^{-1}\Pi_N u;$$

see, for example, discussions in [38] for more information. $\square$

The next lemma constructs a control to move x to x' and y to y' , assuming that the velocity is initially zero.

LEMMA 6.9. *Let $a \in (0, \frac{1}{2})$ and suppose $u_a = 0$, $(x_a, y_a) = (x, y)$. For all $x, x', y, y' \in \mathcal{D}^c$, $\exists C_g$, (depending only on $d(x, y)$ and $d(x', y')$) and a control $g =: g^{ctr, a}$ satisfying $\sup_{t \in (a, 1-a)} \|g_t^{ctr, a}\| \leq C_g$ such that $u_{1-a} = 0$ and $(x_{1-a}, y_{1-a}) = (x', y')$.*

PROOF. First, observe that, by first moving x and then moving y , (and the vice versa) it suffices to fix one particle and move the other. Parameterize $\mathbb{T}^d$ as $(-\pi, \pi]^d$ and suppose without loss of generality that $x = 0$, $y \neq 0$. We will choose the velocity field to fix $x_t = 0$ and satisfy $y_{\frac{1}{2}} = y'$.

We carry out the proof in $d = 2$ for notational simplicity; the $d = 3$ case follows similarly. Let $\delta > 0$ be fixed arbitrary. There are essentially two cases.

Case 1: y, y' do not lie within δ of the same coordinate axis.

Case 1a: Neither y nor y' lie within a δ of any coordinate axis.

Denote $y = (y^{(1)}, y^{(2)})$ and $y' = (y'^{(1)}, y'^{(2)})$. By Assumption 2, trigonometric interpolation and that all shears are stationary solutions of the 2D Euler equations, for any smooth f_t with $f_a = 0$, $\exists g_t$ such that

$$u_t(\zeta) = f_t \begin{pmatrix} h(\zeta^{(2)}) \\ 0 \end{pmatrix},$$

with h a linear combination of $\cos y, \sin y, \cos 2y, \sin 2y$ satisfying $h(0) = 0$ and $h(y^{(2)}) = 1$. It is clear that we can choose f_t (and hence g_t) such that $y_{\frac{a}{2} + \frac{1}{4}} = (y'^{(1)}, y^{(2)})$. Over the time interval $t \in (\frac{a}{2} + \frac{1}{4}, \frac{1}{2})$ we then similarly move the second component of y using a shear flow of the form $(0, b(\zeta^{(2)}))$.

Case 1b: One or both y of y' lie within a δ of a coordinate axis.

Unlike the previous case, the order in which we apply the shear flows (i.e., in the horizontal or in the vertical direction) matters.

Suppose that y' lies on the vertical axis. Then either y lies on the horizontal axis or lies away from either axis. We apply the same procedure of two repeated shear flows as the previous case; however, we *first* adjust $y^{(2)}$ to $y'^{(2)}$ and *then* adjust $y^{(1)}$ to $y'^{(1)}$. If y' lies on the horizontal axis, we proceed similarly, but this time first adjusting $y^{(1)}$ to $y'^{(1)}$ and then adjusting $y^{(2)}$ to $y'^{(2)}$.

Case 2: y, y' lie within a δ of the same coordinate axis.

By symmetry, without loss of generality we can assume that the coordinate axis is the horizontal. Since $y \neq 0$, by trigonometric interpolation, there exists a shear flow $(0, h(y))$ where h is a trigonometric polynomial supported only in the first two harmonics satisfying

$$h(0) = 0, \quad h(y^{(1)}) = 1.$$

Hence, in any time window we can move y back into Case 1b, at which point we proceed as above. $\square$

The next lemma is simply the reverse of Lemma 6.7.

LEMMA 6.10. *Let $u' \in \mathbf{H}$ be arbitrary. Then $\forall \epsilon > 0$, $\exists \delta \ll 1$ and a control $g : [1 - \delta, 1] \rightarrow \mathbf{L}^2$ such that if $\|u_{1-\delta}\|_{\mathbf{H}} \leq \frac{\epsilon}{4}$, then there holds $\|u_1 - u'\|_{\mathbf{H}} < \frac{\epsilon}{4}$, $\sup_{1-\delta \leq t \leq 1} \|u_t\|_{\mathbf{H}} \leq 3\|u'\|_{\mathbf{H}}$ and $d(x_{1-\delta}, x_1) + d(y_{1-\delta}, y_1) \lesssim \delta \|u'\|_{\mathbf{H}}$.*

Finally, we briefly sketch how to assemble the control and prove the necessary stability for the problem.

PROOF OF LEMMA 6.6. We use the control (where $g^{ctr, \delta}$ is chosen to send x_δ, y_δ exactly to x', y' at time $1 - \delta$),

$$g_t = \begin{cases} -\delta^{-1} Q^{-1} u_{\leq N} & t \in (0, \delta), \\ g_t^{ctr, \delta} & t \in (\delta, 1 - \delta), \\ \delta^{-1} Q^{-1} u'_{\leq N} & t \in (1 - \delta, 1). \end{cases}$$

For suitable choices of δ , N , the lemma now follows from Lemmas 6.7, 6.10 and 6.9 and local positivity of the Wiener measure as in, for example, [Lemma 7.3, [15]]. $\square$

6.3. Verification of the drift condition: Proof of Proposition 2.14. We now commence the analysis of $\mathcal{V}$ defined in Section 2.4 using the eigenfunction ψ_p constructed in Section 5. We will assume for the rest of this section that $p \in (0, p_0)$ is fixed, ensuring the existence ψ_p by Proposition 2.11. Recall that ψ_p belongs to $\dot{C}_V^1$, where $V = V_{\beta, \eta}$ for all β large enough and $\eta \in (0, \eta^*)$. In what follows, we will increase this lower bound on β finitely many time without explicitly keeping track of the value.

For $(x, y) \in \mathcal{D}^c$, we define $w(x, y) = y - x \bmod 2\pi\mathbb{Z}^d \in \mathbb{R}^d$ (i.e., (x, y) is the shortest displacement vector from x to y). Then, for each $(u, x, w) \in \mathbf{H} \times \mathbb{T}^d \times \mathbb{R}^d$, $\mathcal{V}$ is of the form

$$\mathcal{V}(u, x, y) = \hat{h}_p(u, x, w(x, y)) + V(u),$$

where

$$\hat{h}_p(u, x, w) := |w|^{-p} \psi_p\left(u, x, \frac{w}{|w|}\right) \chi(|w|),$$

and χ is a smooth, strictly positive cutoff supported in $B(0, \frac{1}{50})$. The cut-off function ensures continuity, since $w(x, y)$ is continuous on $\mathcal{D}_{1/10}$. Consequently, this cut-off allows us to pull h_p back to a continuous function $\hat{h}_p$ on $\mathbf{H} \times \mathcal{D}^c$ by

$$h_p(u, x, y) := \hat{h}_p(u, x, w(x, y)).$$

Our strategy for verifying the drift condition (Condition 4) is to show that $\hat{h}_p$ is an approximate eigenfunction for the two point Markov semigroup $P_t^{(2)}$. To do this, it is convenient to work with the infinitesimal generator $\mathcal{L}_{(2)}$ of $P_t^{(2)}$. Therefore, we will need to show that it is a legitimate C_0 semigroup on an appropriate Banach space. Moreover, we will need to deal with observables that are unbounded both for large u and as (x, y) approach the diagonal $\mathcal{D}$. To do this, we introduce the following weight:

$$\widehat{V}(u, x, y) := \widehat{V}_{p, \beta, \eta}(u, x, y) = d(x, y)^{-p} V_{\beta, \eta}(u),$$

where $p \in (0, p_0)$ and $V_{\beta, \eta}$ is defined by (2.2), with $\eta \in (0, \eta^*)$, $\beta \geq 1$ and $d(x, y)$ denotes the natural distance metric on $\mathbb{T}^d$. Treating p, η, β as fixed for now we then define the following weighted supremum norm:

$$\|\varphi\|_{C_{\widehat{V}}} := \sup_{z \in \mathbf{H} \times \mathcal{D}^c} \frac{|\varphi(z)|}{\widehat{V}(z)},$$

and denote $C_{\widehat{V}}$ to be space of continuous functions on $\mathbf{H} \times \mathcal{D}^c$ whose $\|\cdot\|_{C_{\widehat{V}}}$ norm is finite. Note that since $\mathcal{V} \lesssim \widehat{V}$ (for appropriate p, β, η), we have that $\mathcal{V} \in C_{\widehat{V}}$.

First, we show that $P_t^{(2)}$ is a bounded linear operator on $C_{\widehat{V}}$ (although it lacks strong continuity in this space).

LEMMA 6.11. *For all $p \in (0, p_0)$, $\beta \geq 1$ and $\eta \in (0, \eta^*)$, $P_t^{(2)}$ extends to a bounded linear operator on $C_{\widehat{V}}$. Specifically, there exists a constant C such that for each $\varphi \in C_{\widehat{V}}$,*

$$\|P_t^{(2)} \varphi\|_{C_{\widehat{V}}} \leq e^{Ct} \|\varphi\|_{C_{\widehat{V}}}.$$

PROOF. To prove this, first note that

$$|P_t^{(2)} \varphi| \leq \|\varphi\|_{C_{\widehat{V}}} \mathbf{E}(d(x_t, y_t)^{-p} V(u_t)).$$

Our first step will be to deduce a lower bound for $d(x_t, y_t)$. To this end, we note that when $d(x_t, y_t) < 1/10$ we find a local chart and represent x_t and y_t as vectors in $\mathbb{R}^d$ so that $d(x_t, y_t) = |x_t - y_t|$ and, therefore, deduce the differential inequality

$$(6.4) \quad \frac{d}{dt} \log(|x_t - y_t|) \geq -\|u_t\|_{H^{r_0}}.$$

In order to integrate the above inequality, we must be careful that we integrate over time intervals where the local chart we used remains valid. With this in mind, suppose that $t > 0$ is such that $d(x_t, y_t) < d(x, y)/100$ and define

$$\tau_t = \sup\{s : 0 \leq s < t, d(x_s, y_s) \geq d(x, y)/10\},$$

to be the last time before t that $d(x_s, y_s)$ was outside the chart. Note that τ_t is well defined, since $d(x_0, y_0) \geq d(x, y)/10$ and is strictly less than t by continuity of (x_t, y_t) . Consequently, integrating (6.4) from τ_t to t yields

$$|x_t - y_t| \geq |x_{\tau_t} - y_{\tau_t}| \exp\left(-\int_{\tau_t}^t \|u_s\|_{H^{r_0}} ds\right) \geq \frac{d(x, y)}{10} \exp\left(-\int_0^t \|u_s\|_{H^{r_0}} ds\right).$$

Of course, when $d(x_t, y_t) > d(x, y)/100$, a lower bound is automatic and, therefore, we obtain

$$d(x_t, y_t) \gtrsim d(x, y) \exp\left(-\int_0^t \|u_s\|_{H^{r_0}} ds\right).$$

It then follows from Lemma 3.7 that we can bound

$$\begin{aligned} \mathbf{E}_z(d(x_t, y_t)^{-p} V(u_t)) &\lesssim d(x, y)^{-p} \mathbf{E}_u \exp\left(p \int_0^t \|u_s\|_{H^{r_0}} ds\right) V(u_t) \\ &\lesssim e^{Ct} d(x, y)^{-p} V(u). \end{aligned} \quad \square$$

As we saw for the twisted Markov semigroup, boundedness in a Banach space is not enough to ensure that $P_t^{(2)}$ gives rise to a C_0 -semigroup on that space. Indeed, we must define the space $\mathring{C}_{\widehat{V}}$ obtained as the closure of the space of smooth cylinder functions

$$\mathring{C}_0^\infty(\mathbf{H} \times \mathcal{D}^c) = \{\varphi | \varphi(u, x, y) = \phi(\Pi_K u, x, y), |\mathcal{K}| < \infty, \phi \in C_0^\infty(\mathbb{R}^{|\mathcal{K}|})\}$$

with respect to the norm $\|\cdot\|_{\mathring{C}_{\widehat{V}}}$. An analogous argument to the proof of Proposition 5.5 for the twisted Markov semigroup $\hat{P}_t^p$ (in fact a strictly simpler one since it does not involve derivatives) gives the C_0 semigroup property of $P_t^{(2)}$. We omit the proof for brevity.

PROPOSITION 6.12. *Let $\widehat{V} = \widehat{V}_{p, \beta, \eta}$, where $\eta \in (0, \eta^*)$, $p \in (0, p_0)$, and $\beta > 1$ is taken large enough. Then the Markov semigroup $P_t^{(2)}$ extends to a C_0 semigroup on $\mathring{C}_{\widehat{V}}$.*

Consequently, Proposition 6.12 implies that there is a well-defined generator $\mathcal{L}_{(2)}$ for $P_t^{(2)}$ on $\mathring{C}_{\widehat{V}}$ with dense domain $\text{Dom}(\mathcal{L}_{(2)}) \subseteq \mathring{C}_{\widehat{V}}$. The key estimate of this section is the following approximate drift condition for $\mathcal{L}_{(2)}$.

LEMMA 6.13. *For all $p \in (0, p_0)$, $\eta \in (0, \eta^*)$ and $\beta \geq 1$ taken large enough, h_p belongs to $\text{Dom}(\mathcal{L}_{(2)})$ on $\mathring{C}_{\widehat{V}_{p, \beta, \eta}}$ and there exists a constant C' such that*

$$\mathcal{L}_{(2)} h_p \leq -\Lambda(p) h_p + C' V_{\beta+1, \eta},$$

A priori, it is not clear that h_p actually belongs to the domain of the generator $\mathcal{L}_{(2)}$ since it involves ψ_p , which belongs to the domain of the generator of $\hat{P}_t^p$ and, therefore, more readily belongs to the domain of the generator of the semigroup associated to the linearized motion

$$T P_t \hat{h}(u, x, w) := \mathbf{E}_{(u, x, w)} \hat{h}(u_t, x_t, D\phi^t w).$$

Therefore, in order to prove Lemma 6.13, we will need to approximate $P_t^{(2)} \hat{h}_p$ by

$$T^* P_t \hat{h}_p(u, x, y) := T P_t \hat{h}_p(u, x, w(x, y)).$$

To this end, we write

$$(6.5) \quad \frac{P_t^{(2)} h_p - h_p}{t} = \frac{T^* P_t \hat{h}_p - h_p}{t} + \frac{P_t^{(2)} h_p - T^* P_t \hat{h}_p}{t},$$

and show that each limit on the right-hand side of (6.5) converges separately. It is important to remark that this approximation is only effective $t \rightarrow 0$, indeed $T^* P_t \hat{h}_p$ is not even continuous on $\mathcal{D}^c$ due to the discontinuous nature of $w(x, y)$.

For the first term in (6.5), we have the following.

LEMMA 6.14. *For all $p \in (0, p_0)$, $\eta \in (0, \eta^*)$ and $\beta \geq 1$ large enough, the following limit holds in $C_{\hat{V}_{p, \beta, \eta}}$:*

$$\lim_{t \rightarrow 0} \frac{T^* P_t \hat{h}_p - h_p}{t} = -\Lambda(p) h_p + \mathcal{E}_p,$$

where $\mathcal{E}_p(u, x, y) = \hat{\mathcal{E}}_p(u, x, w(x, y))$ and (recall the definition of H in Remark 2.18)

$$\hat{\mathcal{E}}_p(u, x, w) = H(u, x, w/|w|) |w|^{1-p} \psi_p(u, x, w/|w|) \chi'(|w|).$$

PROOF. To begin, note that $|w|^{-p} \psi_p$ is an eigenfunction of $T P_t$ with eigenvalue $e^{-\Lambda(p)t}$. Denote the linearized process $w_t^* = D_x \phi^t w$ and note that w_t^* is a solution to $\partial_t w_t^* = D u_t(x_t) w_t^*$. Using this, we find

$$T P_t \hat{h}_p = e^{-\Lambda(p)t} h_p + \mathbf{E} |w_t|^{-p} \psi_p(u_t, x_t, w_t^*) (\chi(|w_t^*|) - \chi(|w|)).$$

Noting that

$$\chi(|w_t^*|) - \chi(|w|) = \int_0^t |w_s^*| H(u_s, x_s, v_s) \chi'(|w_s^*|) ds,$$

where $v_s = w_s^*/|w_s^*|$ is the projective process with initial data $v = w/|w|$, allows us to write

$$(6.6) \quad \frac{T P_t \hat{h}_p - \hat{h}_p}{t} = \frac{e^{-\Lambda(p)t} - 1}{t} \hat{h}_p + \hat{\mathcal{E}}_p + \hat{R}_t,$$

where

$$\hat{R}_t := \mathbf{E} \left(|w_t^*|^{-p} \psi_p(u_t, x_t, v_s) \frac{1}{t} \int_0^t |w_s^*| H(u_s, x_s, v_s) \chi'(|w_s^*|) ds - \mathcal{E}_p \right).$$

The fact that $\psi_p \in \dot{C}_V$, where $V = V_{\beta_0, \eta}$, for some $\beta_0 \geq 1$ means that we can find $\psi_p^{(n)} \in \dot{C}_0^\infty$ where $\psi_p^{(n)}$ only depends on finitely many Fourier modes $\Pi_n u$ of u , such that $\psi_p^{(n)} \rightarrow \psi_p$ in C_V . Using the fact that

$$|w_t^*| = |w| \exp \left(\int_0^t H(u_s, x_s, v_s) ds \right)$$

and that $\Pi_n u_t$ is finite dimensional, a direct calculation yields for each n

$$(6.7) \quad |\hat{R}_t| \lesssim |w|^{1-p} \exp\left(C_p \int_0^t \|u_s\|_{H^{r_0}}\right) \sup_{s \in (0,t)} V_{\beta_0+1,\eta}(u_s) (C_n \rho_t + \|\psi_p - \psi_p^{(n)}\|_{C_V}),$$

where C_n is a constant depending (badly) on n and $D\psi_p^{(n)}$ and

$$\rho_t := \sup_{s \in (0,t)} (\|u_s - u\|_{H^{r_0}} + d_{\mathbb{T}^d}(x_s, x) + d_{p^{d-1}}(v_s, v)).$$

A straightforward consequence the evolution equation for the projective process (u_t, x_t, v_t) directly implies

$$\rho_t \leq t \sup_{s \in (0,t)} (1 + \|u_s\|_{\mathbf{H}}^2) + \sup_{s \in (0,t)} \|QW_s\|_{H^{r_0}}.$$

By the BDG inequality, we can bound H^{r_0} norm of the Wiener process using

$$\mathbf{E} \sup_{s \in (0,t)} \|QW_s\|_{H^{r_0}}^2 \lesssim_Q t.$$

Therefore, taking expectation of (6.7) applying Cauchy–Schwarz and the exponential estimates from Lemma (3.7), we conclude that for some $\beta_1 \geq 1$ large enough,

$$\mathbf{E}|\hat{R}_t| \lesssim_Q |w|^{1-p} V_{\beta_1,\eta}(u) (C_n t^{1/2} + \|\psi_p - \psi_p^{(n)}\|_{C_V}).$$

Denote $R_t(u, x, y) = \hat{R}_t(u, x, w(x, y))$ and note that while R_t is not necessarily continuous in (x, y) due to the discontinuity in $w(x, y)$ away from the diagonal, the quantity $|w(x, y)| = d_{\mathbb{T}^d}(x, y)$ is continuous on $\mathcal{D}^c$ and therefore $\mathbf{E}|R_t|$ is bounded above by a continuous function. The corresponding estimate on R_t implies after first sending $t \rightarrow 0$ and then $n \rightarrow \infty$ that

$$\lim_{t \rightarrow 0} \|\mathbf{E}|R_t|\|_{C_{\hat{V}_{p,\beta_1,\eta}}} = 0.$$

This, coupled with the fact that $\frac{e^{-\Lambda(p)t}-1}{t} \rightarrow -\Lambda(p)$ as $t \rightarrow 0$ and $h_p \in C_{\hat{V}_{p,\beta_1,\eta}}$, is sufficient to conclude the proof in light of equation (6.6). $\square$

The second term in (6.5) involves controlling the error involved in approximating $P_t^{(2)} h_p$ by $T^* P_t \hat{h}_p$. As discussed in Section 2.4, this is one of the main difficulties in proving a valid drift condition and is the only reason we need $\hat{C}_V^1$ estimates on ψ_p , that is, so that we can differentiate ψ_p with respect to the projective coordinate and bound it by $V(u)$.

LEMMA 6.15. *For all $p \in (0, p_0)$, $\eta \in (0, \eta^*)$, and $\beta \geq 1$ large enough, the following limit holds in $C_{\hat{V}_{p,\beta,\eta}}$:*

$$\lim_{t \rightarrow 0} \frac{P_t^{(2)} h_p - T^* P_t \hat{h}_p}{t} = (\nabla_y h_p - \nabla_x) \cdot \Sigma,$$

where $\Sigma(u, x, y) = u(y) - u(x) - Du(x)w(x, y)$.

PROOF. Define $w_t = w(x_t, y_t)$ and $w_t^* = D\phi^t w$, where $w = w(x, y)$, and note that

$$\begin{aligned} \frac{P_t^{(2)} h_p - T^* P_t \hat{h}_p}{t} &= \frac{1}{t} \mathbf{E}(\hat{h}_p(u_t, x_t, w_t) - \hat{h}_p(u_t, x_t, w_t^*)) \\ &= \mathbf{E} \int_0^1 \nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta \cdot \frac{w_t - w_t^*}{t}, \end{aligned}$$

where $w_t^\theta = \theta w_t + (1 - \theta)w_t^*$. To continue, for each $t \geq 0$, we define the events

$$A_t := \left\{ t \sup_{s \in (0, t)} \|\nabla u_s\|_\infty \leq \frac{1}{100} \right\},$$

$$B_t := \left\{ t \sup_{s \in (0, t)} (\|\nabla u_s\|_\infty (|w_s| + |w_s^*|)) \leq \frac{|w_t^*|}{2} \right\}.$$

Note that both the events A_t , B_t implicitly depend on the initial data (u, x, y) and we have that

$$\lim_{t \rightarrow \infty} \mathbf{P}(A_t \cap B_t) = 1.$$

The event $A_t \subset \Omega$ is chosen so that on it w_t has not moved far from its starting point w ,

$$|w_t - w| \leq t \sup_{s \in (0, t)} \|\nabla u_s\|_{L^\infty} \leq \frac{1}{100},$$

and, therefore, on A_t we can write

$$w_t - w_t^* = \int_0^t u_s(y_s) - u_s(x_s) - Du_s(x_s)w_s^* ds.$$

It follows that

$$\frac{w_t - w_t^*}{t} \mathbb{1}_{A_t} = \Sigma \mathbb{1}_{A_t} + R_t,$$

where

$$R_t := \frac{\mathbb{1}_{A_t}}{t} \int_0^t (u_s(y_s) - u(y)) - (u_s(x_s) - u(x)) - (Du_s(x_s)w_s^* - Du(x)w) d\tau.$$

Directly computing all the differences it is straightforward to see that $|R_t|$ can be bounded by

$$(6.8) \quad |R_t| \lesssim \sup_{s \in (0, t)} (|u(y_s) - u(x_s) - u(y) + u(x)| + \|u_s - u\|_{H^{r_0}} |w| + \|u_s\|_{H^{r_0}} |w_s^* - w|)$$

Using the evolution equation for $u_t(x_t)$, and denoting $F^u(u) = -B(u, u) - Au$ we find

$$\begin{aligned} & u_t(y_t) - u_t(x_t) \\ &= u(y) - u(x) + \int_0^t F^u(u_s)(y_s) - F^u(u_s)(x_s) ds \\ & \quad + \int_0^t u_s(y_s) \cdot \nabla u_s(y_s) - u_s(x_s) \cdot \nabla u_s(x_s) ds + \int_0^t (Q(y_s) - Q(x_s)) dW_s. \end{aligned}$$

Using this, similar to the proof of Lemma 6.14, an application of the BDG inequality gives

$$(6.9) \quad \sqrt{\mathbf{E}(R_t)^2} \lesssim |w|t^{1/2} \mathbf{E} \exp\left(C \int_0^t \|u\|_{H^{r_0}}\right) \sup_{s \in [0, t]} (1 + \|u_s\|_{\mathbf{H}}^2).$$

The event $B_t \subset \Omega$ is chosen such that on B_t , w_t^* satisfies the following geometric constraint:

$$|w_t - w_t^*| \leq |w_t - w| + |w_t^* - w| \leq t \sup_{s \in (0, t)} (\|\nabla u_s\|_\infty (|w_s| + |w_s^*|)) \leq \frac{1}{2} |w_t^*|$$

and, therefore, on B_t we have a lower bound for $|w_t^\theta|$:

$$|w_t^\theta| \geq |w_t^*| - |w_t - w_t^*| \geq \frac{1}{2} |w_t^*|.$$

Using the fact that $|\nabla_w \hat{h}_p| \leq |w|^{-p-1} \|\psi_p\|_{C_V^1} V_{\beta_0, \eta}(u)$ for some $\beta_0 \geq 1$,

$$\begin{aligned}
 (6.10) \quad & \mathbb{1}_{B_t} \left| \int_0^1 \nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta \right| \\
 & \lesssim \|\psi_p\|_{C_V^1} V_{\beta_0, \eta}(u_t) \mathbb{1}_{B_t} \int_0^1 |w_t^\theta|^{-p-1} d\theta \\
 & \lesssim \|\psi_p\|_{C_V^1} V_{\beta_0, \eta}(u_t) |w_t^*|^{-p-1} \\
 & \lesssim \|\psi_p\|_{C_V^1} \exp\left(C_p \int_0^t \|u_s\|_{H^{r_0}} ds\right) |w|^{-p-1} V_{\beta_0, \eta}(u_t).
 \end{aligned}$$

Combining (6.8) with equation (6.10), using Cauchy–Schwarz, estimate (6.9) and Lemma 3.7 gives

$$(6.11) \quad \mathbf{E} \mathbb{1}_{A_t \cap B_t} \left| \int_0^1 \nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta \cdot R_t \right| \lesssim t^{1/2} \|\psi_p\|_{C_V^1} |w|^{-p} V_{\beta_1, \eta}(u)$$

for some $\beta_1 > \beta_0$ and all $\eta \in (0, \eta^*)$. Again similar to the proof of Lemma 6.14, by density of $\mathcal{C}_0^\infty(\mathbf{H} \times \mathbb{T}^d \times P\mathbb{T}^d)$ in $\mathring{C}_V^1$ we can take an approximating sequence of $\psi_p^{(n)}$ converging to ψ_p in C_V^1 such that $\psi_p^{(n)}$ only depends on finitely many Fourier modes $\Pi_n u$. Using the fact that $\Pi_n u_t$ is finite dimensional, we conclude that

$$\begin{aligned}
 & \mathbb{1}_{A_t \cap B_t} \left| \int_0^1 (\nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta - (\nabla_y - \nabla_x) h_p) \cdot \frac{w_t - w_t^*}{t} \right| \\
 & \leq |w|^{-p} \exp\left(C \int_0^t \|u_s\|_{H^{r_0}} ds\right) \sup_{s \in (0, t)} V_{\beta_1, \eta}(u_s) (C_n \tilde{\rho}_t + \|D_v \psi_p - D_v \psi_p^{(n)}\|_{C_V}),
 \end{aligned}$$

where

$$\tilde{\rho}_t = \sup_{s \in (0, t)} (\|u_s - u\|_{H^{r_0}} + d_{\mathbb{T}^d}(x_s, x) + \mathbb{1}_{A_s} |w_s - w| + \mathbb{1}_{A_s} |w_s^* - w|).$$

Using the evolution equation for (u_t, x_t, w_t) and the BDG inequality to deal with $\tilde{\rho}_t$, an analogous argument to the one in the proof of Lemma 6.14 implies that

$$\lim_{t \rightarrow 0} \mathbf{E} \mathbb{1}_{A_t \cap B_t} \left| \int_0^1 (\nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta - (\nabla_y - \nabla_x) h_p) \cdot \frac{w_t - w_t^*}{t} \right| = 0,$$

where the limit holds in $C_{\hat{V}_{p, \beta_2, \eta}}$ for some β_2 large enough. Combining this with (6.11) and the fact that $R_t = \mathbb{1}_{A_t} \frac{w_t - w_t^*}{t} - \mathbb{1}_{A_t} \Sigma$ yields

$$\lim_{t \rightarrow 0} \mathbf{E} \mathbb{1}_{A_t \cap B_t} \int_0^1 \nabla_w \hat{h}_p(u_t, x_t, w_t^\theta) d\theta \cdot \frac{w_t - w_t^*}{t} = (\nabla_y - \nabla_x) h_p \cdot \Sigma,$$

where the limit holds in $C_{\hat{V}_{p, \beta_3, \eta}}$ for β_3 large enough.

On the complement $A_t^c \cup B_t^c$, we use the fact that for each $\delta > 0$,

$$\begin{aligned}
 \mathbb{1}_{A_t^c \cup B_t^c} & \lesssim t^{1+\delta} \left(\sup_{s \in (0, t)} \|\nabla u_s\|_{L^\infty}^{1+\delta} + |w_t^*|^{-1-\delta} \sup_{s \in (0, t)} (\|\nabla u_s\|_{L^\infty}^{1+\delta} (|w_s| + |w_s^*|))^{1+\delta} \right) \\
 & \lesssim t^{1+\delta} \exp\left(2(1+\delta) \int_0^t \|u_s\|_{H^{r_0}} ds\right) \sup_{s \in (0, t)} \|u_s\|_{H^{r_0}}^{1+\delta},
 \end{aligned}$$

and, therefore,

$$\begin{aligned} & \frac{1}{t} \mathbb{1}_{A_t^c \cup B_t^c} (\hat{h}_p(u_t, x_t, w_t) - \hat{h}_p(u_t, x_t, w_t^*)) \\ & \leq \frac{1}{t} \mathbb{1}_{A_t^c \cup B_t^c} (|w_t|^{-p} + |w_t^*|^{-p}) V_{\beta_0, \eta}(u_t) \\ & \lesssim t^\delta |w|^{-p} \exp\left(C_{p, \delta} \int_0^t \|u_s\|_{H^{r_0}} ds\right) \sup_{s \in (0, t)} V_{\beta_0+1, \eta}(u_s) \end{aligned}$$

which implies by the exponential estimates of Lemma 3.7 that the following limit holds in $C_{\hat{V}_{p, \beta_0+1, \eta}}$:

$$\lim_{t \rightarrow 0} \frac{1}{t} \mathbb{E} \mathbb{1}_{A_t^c \cup B_t^c} (\hat{h}_p(u_t, x_t, w_t) - \hat{h}_p(u_t, x_t, w_t^*)) = 0.$$

Putting all the limits together completes the proof. $\square$

Lemma 6.13 is now a simple consequence of the previous two lemmas.

PROOF OF LEMMA 6.13. Applying Lemmas 6.14, and 6.15 to the splitting (6.5), we can deduce

$$\mathcal{L}_{(2)} h_p = -\Lambda(p) h_p + \mathcal{E}_p + (\nabla_y - \nabla_x) h_p \cdot \Sigma.$$

Note that

$$\mathcal{E}_p \leq |w|^{1-p} \|\nabla u\|_{L^\infty} \|\psi_p\|_{C_V} V_{\beta, \eta}(u)$$

Similarly, since $|(\nabla_y - \nabla_x) h_p| \lesssim |d(x, y)|^{-p-1} \|\psi_p\|_{C_V^1}$ and by Taylor's theorem $|\Sigma| \leq |d(x, y)|^2 \|\nabla^2 u\|_{L^\infty}$, we deduce that since $p < 1$,

$$|\mathcal{E}_p + (\nabla_y - \nabla_x) h_p \cdot \Sigma| \lesssim |d(x, y)|^{1-p} \|u\|_{W^{2, \infty}} \|\psi_p\|_{C_V^1} V_{\beta, \eta}(u) \lesssim V_{\beta+1, \eta}(u). \quad \square$$

We are now ready to complete the proof of Proposition 2.14.

PROOF OF PROPOSITION 2.14. Let $\mathcal{L}$ denote the formal generator of the Navier–Stokes equations defined by equation (3.7) of Section 3. First, observe that for any $\beta > 0$, $\eta \in (0, \eta^*)$ we have for $V = V_{\beta, \eta}$ and all $u \in \mathbf{H}^{\sigma+d-1}$ that

$$\mathcal{L}V(u) = \left(\mathcal{L} \log V(u) + \sum_{m \in \mathbb{K}} |q_m|^2 |D_u \log V(u) e_m|^2 \right) V(u).$$

and, therefore, using the fact that

$$\sum |q_k|^2 |D_u \log V(u) e_m|^2 \leq 8\beta^2 + 8\eta^2 \mathcal{Q} \|\Delta u\|_{L^2}^2,$$

and applying inequality (3.8) of Lemma 3.6 we deduce that

$$\mathcal{L}V(u) \leq \left(-\eta(v - 8\eta\mathcal{Q}) \|\Delta u\|_{L^2}^2 - v\beta \frac{\|\nabla u\|_{H^\sigma}^2}{1 + \|u\|_{H^\sigma}^2} + C \right) V(u).$$

Note that $\eta \leq \eta^*$ ensures that $\frac{\eta}{2}(v - 16\eta\mathcal{Q})$ is positive. Applying Lemma 3.3 implies that $\forall \delta > 0$, $\exists C_\delta > 0$ such that

$$\mathcal{L}V \leq -\delta \log(1 + \|u\|_{H^\sigma}^2) V.$$

We treat the right-hand side above by dividing into regions where $\|u\|_{H^\sigma} \leq 1$ and $\|u\|_{H^\sigma} > 1$, in the former case everything is bounded by a constant, in the latter case we can bound the logarithmic factor below by $\log(2)$. This gives for all $0 < \kappa < \log(2)\delta$, and some constant $C > 0$,

$$-\delta \log(1 + \|u\|_{H^\sigma}^2) V \leq -\kappa V + C.$$

As δ was arbitrary, it follows that for each $\kappa > 0$, we have the bound (for a suitable $C_\kappa > 0$),

$$(6.12) \quad \mathcal{L}V(u) \leq -\kappa V(u) + C_\kappa.$$

Note that inequality (6.12) is an infinitesimal version of a drift condition for Navier–Stokes. However, V does not belong to the domain of $\mathcal{L}_{(2)}$ in $\mathring{C}_{\hat{V}}$, and so we must proceed with more care to deduce a corresponding drift condition on the semigroup P_t . Define $\tau_n = \inf\{t > 0 : \|u_t\|_{\mathbf{H}} > n\}$ (note $\tau_n \rightarrow \infty$ as $n \rightarrow \infty$ with probability one). Applying Itô's formula (Theorem 7.7.5 [50]) to $e^{\Lambda(p)t} V(u_t)$ implies that

$$\begin{aligned} \mathbf{E}_u e^{\Lambda(p)t \wedge \tau_n} V(u_{t \wedge \tau_n}) - V(u) &= \mathbf{E}_u \int_0^{t \wedge \tau_n} e^{\Lambda(p)s} (\Lambda(p)V(u_s) + \mathcal{L}V(u_s)) ds, \\ &\leq \mathbf{E}_u \int_0^{t \wedge \tau_n} e^{\Lambda(p)s} ((\Lambda(p) - \kappa)V(u_s) + C) ds. \end{aligned}$$

As we know that $\mathbf{E} \sup_{s \in (0,t)} V(u_s) < \infty$ (e.g., from Lemma 3.7), we can use dominated convergence to pass the $n \rightarrow \infty$ limit on both sides of the above inequality to deduce

$$(6.13) \quad e^{\Lambda(p)t} P_t V - V \leq \int_0^t e^{\Lambda(p)s} P_s ((\Lambda(p) - \kappa)V(u_s) + C) ds,$$

where P_t denotes the Markov semigroup for the Navier–Stokes equations.

Recall that $\mathcal{V}$ takes the form

$$\mathcal{V} = h_p + V_{\beta+1,\eta},$$

where $h_p \in C_{\hat{V}_{p,\beta,\eta}}$. Naturally, using the C_0 semigroup property of $P_t^{(2)}$ on functions in $\mathring{C}_{\hat{V}_{p,\beta,\eta}}$ and Lemma 6.13 we also find that

$$\begin{aligned} (6.14) \quad e^{\Lambda(p)t} P_t^{(2)} h_p - h_p &= \int_0^t e^{\Lambda(p)s} P_s^{(2)} (\Lambda(p)h_p + \mathcal{L}_{(2)}h_p) ds \\ &\leq \int_0^t e^{\Lambda(p)s} C' P_s V_{\beta+1,\eta} ds. \end{aligned}$$

Using the fact that

$$P_t^{(2)} \mathcal{V} = P_t^{(2)} h_p + P_t V_{\beta+1,\eta},$$

we complete the proof by adding (6.14) and (6.13) and taking κ large enough so that $\kappa - \Lambda(p) \geq C'$ to conclude that there is a constant K such that

$$e^{\Lambda(p)t} P_t^{(2)} \mathcal{V} - \mathcal{V} \leq C \int_0^t e^{\Lambda(p)s} ds \leq K e^{\Lambda(p)t}.$$

□

7. Correlation decay: Proof of Theorem 1.4.

PROOF OF THEOREM 2.2. Geometric ergodicity for the two-point process (u_t, x_t, y_t) as in Theorem 2.2 follows from the general framework given in Theorem 2.4 and Conditions 1, 2, 3 and 4 listed there. A proof of the strong Feller property in Condition 1 was sketched in Section 6.1, while topological irreducibility as in Condition 2 was obtained in Section 6.2.

Our desired Lyapunov function $\mathcal{V}$ for the two-point process was identified in the previous section the proof of Proposition 2.14, validating Condition 4, while Lemma 4.4 affirms Condition 3 holds for this choice of $\mathcal{V}$. $\square$

It remains to complete the proof of Theorem 1.4, which occupies the remainder of the paper.

Our goal is to prove the following. Let $s > 0$ and $p \geq 1$ be fixed. Then there exists a deterministic constant $\hat{\gamma} = \hat{\gamma}(s, p)$ and a measurable function $D : \Omega \times \mathbf{H} \rightarrow [1, \infty)$, with the following property: for $\mathbf{P} \times \mu$ -almost all $(\omega, u) \in \Omega \times \mathbf{H}$ and for all mean-zero $f, g \in H^s$, we have

$$(7.1) \quad \left| \int f(x) g(\phi_{\omega, u}^t(x)) dx \right| \leq D(\omega, u) e^{-\hat{\gamma} t} \|f\|_{H^s} \|g\|_{H^s}.$$

The proof proceeds in several steps: (1) establish correlation bounds in discrete time; and (2) extending these correlation bounds to cover all real times $t \geq 0$. At the end, we will (3) estimate p th moments for the “random constant” D in terms of u .

NOTATION. Let $\mathcal{V}$ be the Lyapunov function appearing in Theorem 2.2 and note (see Sections 2.4 and 6) that $\mathcal{V}(u, x, y) \leq (V(u))^2 \hat{W}(x, y)$, where $V = V_{\beta, \eta}$ for some $\beta, \eta > 0$ (see (2.2) in Section 2.4 for notation) and $\hat{W} \in L^1(\text{Leb})$.

Hereafter, $s > 0$, $p \geq 1$ are fixed. Without loss of generality, we will assume $s \in (0, 1)$. Fix as well $\hat{\gamma} \in (0, \frac{\hat{\alpha}}{2})$, where $\hat{\alpha} > 0$ is the mixing rate as in Proposition 2.14 for the two-point process; further constraints will be placed on $\hat{\gamma}$ as we proceed. Throughout, $\hat{\gamma}' \in (\hat{\gamma}, \frac{\hat{\alpha}}{2})$ will be a dummy parameter, also chosen appropriately. Given generic ω, u , we write $\phi^t = \phi_{\omega, u}^t$ for short. Let f, g be fixed, smooth, mean-zero observables. Finally, for $k = (k_1, \dots, k_d) \in \mathbb{Z}^d$, we write $|k| = |k|_\infty = \max\{|k_i|, 1 \leq i \leq d\}$.

7.1. Correlation bounds in discrete time. Let $\{\hat{e}_k\}_{k \in \mathbb{Z}_0^d}$ be the real Fourier basis of mean-zero functions on $\mathbb{T}^d$ (see, for instance, the notation in [42]).

Fix $u \in \mathbf{H}$, regarded as a fixed initial condition for the velocity field process (u_t) . By a variant of the Borel–Cantelli argument given in Section 2.1, we have that for each $(k, k') \in \mathbb{Z}_0^d \times \mathbb{Z}_0^d$, the random variable

$$N_{k, k'}(u) = \max \left\{ n \geq 0 : \left| \int \hat{e}_k(x) \hat{e}_{k'}(\phi^n(x)) dx \right| > e^{-\hat{\gamma}' n} V(u) \right\}$$

is finite with probability 1, where the tail estimate

$$(7.2) \quad \mathbf{P}\{N_{k, k'}(u) > n\} \lesssim e^{-(\hat{\alpha} - 2\hat{\gamma}')n}$$

holds uniformly in u, k, k' . In particular, it holds that $|\int \hat{e}_k(x) \hat{e}_{k'}(\phi^n(x)) dx| \leq e^{\hat{\gamma}'(N_{k, k'}(u) - n)} V(u)$ for all $n \geq 0$. Hereafter, let us suppress the “ u ” (which remains fixed) in $N_{k, k'}$.

Expand

$$f = \sum_{k \in \mathbb{Z}_0^d} f_k e_k, \quad g = \sum_{k \in \mathbb{Z}_0^d} g_k e_k,$$

so that

$$(7.3) \quad \left| \int f(x) g(\phi^n(x)) dx \right| \leq V(u) e^{-\hat{\gamma}' n} \sum_{k, k' \in \mathbb{Z}_0^d} |f_k| |g_{k'}| e^{\hat{\gamma}' N_{k, k'}}$$

To form a comparison of the right-hand side of (7.3) with a Sobolev norm, observe that due to the uniformity of the estimate (7.2) in k, k' , we have

$$(7.4) \quad \mathbf{P}\{e^{\hat{\gamma}' N_{k,k'}} > |k||k'|\} \lesssim (|k||k'|)^{-\frac{\hat{\alpha}-2\hat{\gamma}'}{\hat{\gamma}'}} ,$$

again uniformly in k, k' . The right-hand side is summable over $\mathbb{Z}_0^d \times \mathbb{Z}_0^d$ when $\hat{\gamma}$ is sufficiently small. Applying Borel–Cantelli to the countable collection of events $\mathfrak{S}_{k,k'} = \{e^{\hat{\gamma}' N_{k,k'}} > |k||k'|\}$, we conclude that the measurable function³

$$K = \max\{|k| \vee |k'| : k, k' \in \mathbb{Z}_0^d \text{ and } e^{\hat{\gamma}' N_{k,k'}} > |k||k'|\}$$

satisfies the tail estimate $\mathbf{P}\{K > \ell\} \lesssim \ell^{2d - \frac{\hat{\alpha}-2\hat{\gamma}'}{\hat{\gamma}'}}$ uniformly in u . Observe that by definition, we have $e^{\hat{\gamma}' N_{k,k'}} \leq |k||k'|$ for all $(k, k') \in \mathbb{Z}_0^d \times \mathbb{Z}_0^d$ with $|k| \vee |k'| > K$.

Define

$$\hat{D} = \max_{|\hat{k}|, |\hat{k}'| \leq K} e^{\hat{\gamma}' N_{\hat{k}, \hat{k}'}} .$$

Plugging the simple bound $e^{\hat{\gamma}' N_{k,k'}} \leq \hat{D}(1 + |k||k'|)$ into (7.3), we conclude that for all $n \geq 0$,

$$\begin{aligned} \left| \int_{\mathbb{T}^d} f(x) g(\phi^n(x)) dx \right| &\leq \hat{D} V(u) \left(\sum_k |k| |f_k| \right) \left(\sum_k |k| |g_k| \right) e^{-\hat{\gamma}' n} \\ &\lesssim \hat{D} V(u) \|f\|_{H^{\frac{d}{2}+2}} \|g\|_{H^{\frac{d}{2}+2}} e^{-\hat{\gamma}' n} \end{aligned}$$

To control the right-hand side in terms of H^s norms, we can compensate for reduced regularity of f, g by reducing the exponential decay rate. For this, we will use the following standard approximation Lemma (cf. Lemma 4.2 in [22], where similar ideas are used to control correlation decay in terms of an H^1 norm).

LEMMA 7.1. *Let $0 < s < s'$ and let $h \in H^s$ have zero mean. Then, for any $\epsilon > 0$ there exists a mean-zero $h_\epsilon \in H^{s'}$ such that (i) $\|h_\epsilon\|_{L^2} \lesssim \|h\|_{L^2}$, (ii) $\|h_\epsilon - h\|_{L^2} \lesssim \epsilon \|h\|_{H^s}$ and (iii) $\|h_\epsilon\|_{H^{s'}} \lesssim_{s,s'} \epsilon^{-\frac{s'-s}{s}} \|h\|_{H^s}$.*

We apply the above with $s' = \frac{d}{2} + 2$ and for the value of s specified at the beginning of the section. Fixing $\epsilon > 0$, we estimate

$$\begin{aligned} \left| \int f(x) g(\phi^n(x)) dx \right| &\leq \left| \int f_\epsilon(x) g_\epsilon(\phi^n(x)) dx \right| + \|g\|_{L^2} \|f - f_\epsilon\|_{L^2} + \|f_\epsilon\|_{L^2} \|g - g_\epsilon\|_{L^2} \\ &\lesssim \hat{D} V(e^{-\hat{\gamma}' n} \epsilon^{-\frac{d+4-2s}{2s}} + \epsilon) \|f\|_{H^s} \|g\|_{H^s} . \end{aligned}$$

Optimizing in ϵ on the right-hand side yields $\epsilon = e^{-\frac{2s\hat{\gamma}'}{d+4}n}$, resulting in the estimate

$$(7.5) \quad \left| \int f(x) g(\phi^n(x)) dx \right| \leq \hat{D} V e^{-\hat{\gamma}'' n} \|f\|_{H^s} \|g\|_{H^s} ,$$

where $\hat{\gamma}'' := \frac{2s\hat{\gamma}'}{d+4}$ (having absorbed an s -dependent constant into $\hat{D}$).

³Given real numbers $a, b > 0$, we write $a \vee b = \max\{a, b\}$.

7.2. *Correlation bounds in continuous time.* To estimate (7.1) at continuous times $t \in \mathbb{R}_{\geq 0}$, let $t = n + \hat{t}$, $\hat{t} \in [0, 1)$, $n \in \mathbb{Z}_{\geq 0}$. Applying (7.5), we have

$$\left| \int f(x) g(\phi_{\omega, u}^t(x)) dx \right| \leq e^{\hat{\gamma}''} \hat{D}(\omega, u) V(u) \|f\|_{H^s} \|g \circ \phi_{\theta^n \omega, u_n}^{\hat{t}}\|_{H^s} e^{-\hat{\gamma}'' t}.$$

We will estimate the factor $\|g \circ \phi_{\theta^n \omega, u_n}^{\hat{t}}\|_{H^s}$ by

$$(7.6) \quad \sup_{\hat{t} \in [0, 1)} \|g \circ \phi_{\omega', u'}^{\hat{t}}\|_{H^s} \leq \Gamma(\omega', u') \|g\|_{H^s},$$

where $\Gamma : \Omega \times \mathbf{H} \rightarrow [1, \infty)$ is defined by (for some suitable constant $C > 0$),

$$\Gamma(\omega', u') = \exp\left(C \int_0^1 \|\nabla u'_t\|_{L^\infty} dt\right).$$

To verify (7.6), we use the characterization $\|f\|_{H^s} \approx \|f\|_{L^2} + (\int \int \frac{|f(x) - f(y)|^2}{|x - y|^{2s+d}} dx dy)^{1/2}$ for the H^s norm, $s \in (0, 1)$, to estimate

$$\|g \circ \phi_{\omega', u'}^{\hat{t}}\|_{H^s} \lesssim \|g\|_{L^2} + \left(\int \int \frac{|g(x_{\hat{t}}) - g(y_{\hat{t}})|^2}{|x - y|^{2s+d}} dx dy \right)^{1/2} \lesssim \|g\|_{H^s} \|D\phi_{\omega', u'}^{\hat{t}}\|_{L^\infty}^{s+d/2}.$$

The bound (7.6) now follows from Grönwall's inequality.

Fix $\epsilon \ll \hat{\gamma}''$ and define

$$\bar{\Gamma}(\omega, u) = \max_{n \leq M} \Gamma(\theta^n \omega, u_n),$$

where $M := \max\{n \geq 0 : \Gamma(\theta^n \omega, u_n) \geq e^{\epsilon n}\}$, so that $\Gamma(\theta^n \omega, u_n) \leq e^{\epsilon n} \bar{\Gamma}(\omega, u)$ for all n (that M is finite, hence $\bar{\Gamma}$ is defined and will be checked shortly). Overall, we have

$$\left| \int f(x) g(\phi_{\omega, u}^t(x)) dx \right| \leq \underbrace{e^{\hat{\gamma}''} \hat{D} V \bar{\Gamma}}_{=: D(\omega, u)} \|f\|_{H^s} \|g\|_{H^s} e^{-(\hat{\gamma}'' - \epsilon)t}.$$

Finally, we set $\epsilon = \frac{1}{100} \hat{\gamma}''$ and $\hat{\gamma} = \hat{\gamma}'' - \epsilon$. This completes the proof of the continuous-time correlation estimate as in item (2) for any sufficiently small value of $\hat{\gamma}$.

7.3. *Estimating moments of $D(\omega, u)$.* With u fixed, let us now estimate moments $\mathbf{E} D(\omega, u)^p$ by breaking up $D \approx \hat{D} V \bar{\Gamma}$ as above.

Let us first estimate the $\mathbf{P}$ -moments of $\bar{\Gamma}$. To start, by Lemma 3.10 applied at time $t = 1$, $\exists D_0 : \Omega \times \mathbf{H} \rightarrow \mathbb{R}_{\geq 1}$ such that $\Gamma(\omega', u') \lesssim D_0(\omega', u')$, where for any $p \geq 1$, we can arrange so that $\mathbf{E}(D_0(\cdot, u'))^p \lesssim V(u')$. In particular,

$$\mathbf{P}(\Gamma(\theta^n \omega, u_n) > e^{\epsilon n}) \lesssim V(u) e^{-\epsilon n}$$

by Chebyshev's inequality, hence by Borel–Cantelli the random time $M = M(\omega, u)$ is almost-surely finite, with the tail estimate $\mathbf{P}(M > n) \lesssim V(u) e^{-\epsilon n}$. Now,

$$\begin{aligned} \mathbf{E}(\bar{\Gamma}(\cdot, u))^p &= \mathbf{E} \max_{n \leq M} (\Gamma(\theta^n \cdot, u_n))^p \\ &\leq \mathbf{E} \sum_{n=0}^{\infty} \chi_{n \leq M} (\Gamma(\theta^n \cdot, u_n))^p \\ &\leq \sum_{n=0}^{\infty} (\mathbf{P}(M \geq n))^{1/2} (\mathbf{E}(\Gamma(\theta^n \cdot, u_n))^{2p})^{1/2}. \end{aligned}$$

We already have the estimate $\mathbf{P}(M \geq n) \lesssim V(u)e^{-\epsilon n}$, while for the other term we have

$$\mathbf{E}(\Gamma(\theta^n \cdot, u_n))^{2p} \lesssim \mathbf{E}V(u_n) \lesssim V(u)$$

using the tower property of conditional expectation and the drift condition for V . In total, $\mathbf{E}(\bar{\Gamma}(\cdot, u))^p \lesssim_p V(u)$.

It remains to estimate p th moments of $\hat{D}$. Note that although $\hat{D}$ depends nontrivially on the initial velocity u , all of our tails estimates for $\hat{D}$ are uniform in u . In particular, all of the following estimates are uniform in u :

$$\begin{aligned} \mathbf{E}\hat{D}^p &= \sum_{K_0=1}^{\infty} \mathbf{E}\mathbf{1}_{K=K_0} \max_{|k|, |k'| \leq K_0} e^{\hat{\gamma}' p N_{k,k'}} \\ &\leq \sum_{K_0=1}^{\infty} (\mathbf{P}\{K = K_0\})^{1/2} \cdot \left\| \max_{|k|, |k'| \leq K_0} e^{\hat{\gamma}' p N_{k,k'}} \right\|_{L^2(\Omega)} \\ &\lesssim \sum_{K_0=1}^{\infty} K_0^{\frac{d}{2} - \frac{\hat{\alpha} - 2\hat{\gamma}'}{2\hat{\gamma}'}} \sum_{|k|, |k'| \leq K_0} \|e^{\hat{\gamma}' p N_{k,k'}}\|_{L^2(\Omega)}. \end{aligned}$$

The $\|e^{\hat{\gamma}' p N_{k,k'}}\|_{L^2(\Omega)}$ terms on the right-hand side are all uniformly bounded (independent of k, k' and u) by (7.4) if $\hat{\gamma}'$ is sufficiently small. Since there are $\approx K_0^{2d}$ such terms, we obtain

$$\mathbf{E}\hat{D}^p \lesssim \sum_{K_0=1}^{\infty} K_0^{\frac{5d}{2} - \frac{\hat{\alpha} - 2\hat{\gamma}'}{2\hat{\gamma}'}} ,$$

which provides a finite moment estimate when $\hat{\gamma}'$ is sufficiently small.

Funding. J. Bedrossian was supported by NSF CAREER grant DMS-1552826 and NSF RNMS #1107444 (Ki-Net).

This material was based upon work supported by the National Science Foundation under Award No. DMS-1604805. A. Blumenthal would like Dmitry Dolgopyat for useful insights and helpful discussions.

This material was based upon work supported by the National Science Foundation under Award No. DMS-1803481.

REFERENCES

- [1] AIMINO, R., NICOL, M. and VAIENTI, S. (2015). Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields* **162** 233–274. [MR3350045](https://doi.org/10.1007/s00440-014-0571-y) <https://doi.org/10.1007/s00440-014-0571-y>
- [2] ALBERTI, G., CRIPPA, G. and MAZZUCATO, A. L. (2019). Exponential self-similar mixing by incompressible flows. *J. Amer. Math. Soc.* **32** 445–490. [MR3904158](https://doi.org/10.1090/jams/913) <https://doi.org/10.1090/jams/913>
- [3] ANTONSEN, T. M. JR, FAN, Z., OTT, E. and GARCIA-LOPEZ, E. (1996). The role of chaotic orbits in the determination of power spectra of passive scalars. *Phys. Fluids* **8** 3094–3104.
- [4] ARENDT, W., GRABOSCH, A., GREINER, G., GROH, U., LOTZ, H. P., MOUSTAKAS, U., NAGEL, R., NEUBRANDER, F. and SCHLOTTERBECK, U. (1986). *One-Parameter Semigroups of Positive Operators. Lecture Notes in Math.* **1184**. Springer, Berlin. [MR0839450](https://doi.org/10.1007/BFb0074922) <https://doi.org/10.1007/BFb0074922>
- [5] ARNOLD, L. (1984). A formula connecting sample and moment stability of linear stochastic systems. *SIAM J. Appl. Math.* **44** 793–802. [MR0750951](https://doi.org/10.1137/0144057) <https://doi.org/10.1137/0144057>
- [6] ARNOLD, L. and KLIEMANN, W. (1987). Large deviations of linear stochastic differential equations. In *Stochastic Differential Systems* 115–151. Springer, Berlin.
- [7] ARNOLD, L., KLIEMANN, W. and OELJEKLAUS, E. (1986). Lyapunov exponents of linear stochastic systems. In *Lyapunov Exponents (Bremen, 1984). Lecture Notes in Math.* **1186** 85–125. Springer, Berlin. [MR0850072](https://doi.org/10.1007/BFb0076835) <https://doi.org/10.1007/BFb0076835>

- [8] AYYER, A., LIVERANI, C. and STENLUND, M. (2009). Quenched CLT for random toral automorphism. *Discrete Contin. Dyn. Syst.* **24** 331–348. MR2486579 <https://doi.org/10.3934/dcds.2009.24.331>
- [9] BABIANO, A., BOFFETTA, G., PROVENZALE, A. and VULPIANI, A. (1994). Chaotic advection in point vortex models and two-dimensional turbulence. *Phys. Fluids* **6** 2465–2474. MR1279363 <https://doi.org/10.1063/1.868194>
- [10] BABIANO, A. and PROVENZALE, A. (2007). Coherent vortices and tracer cascades in two-dimensional turbulence. *J. Fluid Mech.* **574** 429–448. MR2292791 <https://doi.org/10.1017/S0022112006004265>
- [11] BAHOUN, W., BOSE, C. and RUZIBOEV, M. (2019). Quenched decay of correlations for slowly mixing systems. *Trans. Amer. Math. Soc.* **372** 6547–6587. MR4024530 <https://doi.org/10.1090/tran/7811>
- [12] BALADI, V., BENEDICKS, M. and MAUME-DESCHAMPS, V. (2002). Almost sure rates of mixing for i.i.d. unimodal maps. *Ann. Sci. Éc. Norm. Supér. (4)* **35** 77–126. MR1886006 [https://doi.org/10.1016/S0012-9593\(01\)01083-7](https://doi.org/10.1016/S0012-9593(01)01083-7)
- [13] BATCHELOR, G. K. (1959). Small-scale variation of convected quantities like temperature in turbulent fluid. I. General discussion and the case of small conductivity. *J. Fluid Mech.* **5** 113–133. MR0100457 <https://doi.org/10.1017/S002211205900009X>
- [14] BAXENDALE, P. H. and STROOCK, D. W. (1988). Large deviations and stochastic flows of diffeomorphisms. *Probab. Theory Related Fields* **80** 169–215. MR0968817 <https://doi.org/10.1007/BF00356102>
- [15] BEDROSSIAN, J., BLUMENTHAL, A. and PUNSHON-SMITH, S. (2018). Lagrangian chaos and scalar advection in stochastic fluid mechanics. *J. Eur. Math. Soc.* To appear. Available at [arXiv:1809.06484](https://arxiv.org/abs/1809.06484).
- [16] BEDROSSIAN, J., BLUMENTHAL, A. and PUNSHON-SMITH, S. (2019). The batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics. *Comm. Pure Appl. Math.* To appear. Available at [arXiv:1911.11014](https://arxiv.org/abs/1911.11014).
- [17] BEDROSSIAN, J., BLUMENTHAL, A. and PUNSHON-SMITH, S. (2021). Almost-sure enhanced dissipation and uniform-in-diffusivity exponential mixing for advection-diffusion by stochastic Navier–Stokes. *Probab. Theory Related Fields* **179** 777–834. MR4242626 <https://doi.org/10.1007/s00440-020-01010-8>
- [18] BRESSAN, A. (2003). A lemma and a conjecture on the cost of rearrangements. *Rend. Semin. Mat. Univ. Padova* **110** 97–102. MR2033002
- [19] BRICMONT, J., KUPIAINEN, A. and LEFEVERE, R. (2002). Exponential mixing of the 2D stochastic Navier–Stokes dynamics. *Comm. Math. Phys.* **230** 87–132. MR1930573 <https://doi.org/10.1007/s00220-002-0708-1>
- [20] CARVERHILL, A. (1987). Furstenberg’s theorem for nonlinear stochastic systems. *Probab. Theory Related Fields* **74** 529–534. MR0876254 <https://doi.org/10.1007/BF00363514>
- [21] CHIRIKOV, B. V. (1979). A universal instability of many-dimensional oscillator systems. *Phys. Rep.* **52** 264–379. MR0536429 [https://doi.org/10.1016/0370-1573\(79\)90023-1](https://doi.org/10.1016/0370-1573(79)90023-1)
- [22] COTI ZELATI, M., DELGADINO, M. G. and ELGINDI, T. M. (2020). On the relation between enhanced dissipation timescales and mixing rates. *Comm. Pure Appl. Math.* **73** 1205–1244. MR4156602 <https://doi.org/10.1002/cpa.21831>
- [23] CRIPPA, G., LUCA, R. and SCHULZE, C. (2019). Polynomial mixing under a certain stationary Euler flow. *Phys. D* **394** 44–55. MR3958293 <https://doi.org/10.1016/j.physd.2019.01.009>
- [24] DA PRATO, G. (2014). *Introduction to Stochastic Analysis and Malliavin Calculus*, 3rd ed. *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]* **13**. Edizioni della Normale, Pisa. MR3186829 <https://doi.org/10.1007/978-88-7642-499-1>
- [25] DA PRATO, G. and ZABCZYK, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge Univ. Press, Cambridge. MR1417491 <https://doi.org/10.1017/CBO9780511662829>
- [26] DOLGOPYAT, D. (1998). On decay of correlations in Anosov flows. *Ann. of Math. (2)* **147** 357–390. MR1626749 <https://doi.org/10.2307/121012>
- [27] DOLGOPYAT, D., KALOSHIN, V. and KORALOV, L. (2004). Sample path properties of the stochastic flows. *Ann. Probab.* **32** 1–27. MR2040774 <https://doi.org/10.1214/aop/1078415827>
- [28] DONZIS, D. A., SREENIVASAN, K. R. and YEUNG, P. K. (2010). The Batchelor spectrum for mixing of passive scalars in isotropic turbulence. *Flow Turbul. Combust.* **85** 549–566.
- [29] DRAGIČEVIĆ, D., FROYLAND, G., GONZÁLEZ-TOKMAN, C. and VAIENTI, S. (2018). A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys.* **360** 1121–1187. MR3803820 <https://doi.org/10.1007/s00220-017-3083-7>
- [30] DUARTE, P. (1994). Plenty of elliptic islands for the standard family of area preserving maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **11** 359–409. MR1287238 [https://doi.org/10.1016/S0294-1449\(16\)30180-9](https://doi.org/10.1016/S0294-1449(16)30180-9)
- [31] DUNFORD, N. and SCHWARTZ, J. T. (1957). *Linear Operators. Part 1: General Theory*. Interscience, New York.

- [32] E, W. and MATTINGLY, J. C. (2001). Ergodicity for the Navier–Stokes equation with degenerate random forcing: Finite-dimensional approximation. *Comm. Pure Appl. Math.* **54** 1386–1402. [MR1846802](#) <https://doi.org/10.1002/cpa.10007>
- [33] ECKMANN, J.-P. and HAIRER, M. (2001). Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Comm. Math. Phys.* **219** 523–565. [MR1838749](#) <https://doi.org/10.1007/s002200100424>
- [34] ELGINDI, T. M. and ZLATOŠ, A. (2019). Universal mixers in all dimensions. *Adv. Math.* **356** 106807. [MR4008523](#) <https://doi.org/10.1016/j.aim.2019.106807>
- [35] FENG, Y. and IYER, G. (2019). Dissipation enhancement by mixing. *Nonlinearity* **32** 1810–1851. [MR3942601](#) <https://doi.org/10.1088/1361-6544/ab0e56>
- [36] FLANDOLI, F. and MASLOWSKI, B. (1995). Ergodicity of the 2-D Navier–Stokes equation under random perturbations. *Comm. Math. Phys.* **172** 119–141. [MR1346374](#)
- [37] FURSTENBERG, H. (1963). Noncommuting random products. *Trans. Amer. Math. Soc.* **108** 377–428. [MR0163345](#) <https://doi.org/10.2307/1993589>
- [38] GLATT-HOLTZ, N. E., HERZOG, D. P. and MATTINGLY, J. C. (2018). Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations. *Ann. PDE* **4** 16. [MR3862850](#) <https://doi.org/10.1007/s40818-018-0052-1>
- [39] GOLDBY, B. and MASLOWSKI, B. (2005). Exponential ergodicity for stochastic Burgers and 2D Navier–Stokes equations. *J. Funct. Anal.* **226** 230–255. [MR2158741](#) <https://doi.org/10.1016/j.jfa.2004.12.009>
- [40] GORODETSKI, A. (2012). On stochastic sea of the standard map. *Comm. Math. Phys.* **309** 155–192. [MR2864790](#) <https://doi.org/10.1007/s00220-011-1365-z>
- [41] HAIRER, M. (2011). On Malliavin’s proof of Hörmander’s theorem. *Bull. Sci. Math.* **135** 650–666. [MR2838095](#) <https://doi.org/10.1016/j.bulsci.2011.07.007>
- [42] HAIRER, M. and MATTINGLY, J. C. (2006). Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)* **164** 993–1032. [MR2259251](#) <https://doi.org/10.4007/annals.2006.164.993>
- [43] HAIRER, M. and MATTINGLY, J. C. (2008). Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. *Ann. Probab.* **36** 2050–2091. [MR2478676](#) <https://doi.org/10.1214/08-AOP392>
- [44] HAIRER, M. and MATTINGLY, J. C. (2011). A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electron. J. Probab.* **16** 658–738. [MR2786645](#) <https://doi.org/10.1214/EJP.v16-875>
- [45] HAIRER, M. and MATTINGLY, J. C. (2011). Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI. Progress in Probability* **63** 109–117. Birkhäuser/Springer Basel AG, Basel. [MR2857021](#) https://doi.org/10.1007/978-3-0348-0021-1_7
- [46] IYER, G., KISELEV, A. and XU, X. (2014). Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. *Nonlinearity* **27** 973–985. [MR3207161](#) <https://doi.org/10.1088/0951-7715/27/5/973>
- [47] KUKSIN, S., NERSESYAN, V. and SHIRIKYAN, A. (2020). Exponential mixing for a class of dissipative PDEs with bounded degenerate noise. *Geom. Funct. Anal.* **30** 126–187. [MR4081057](#) <https://doi.org/10.1007/s00039-020-00525-5>
- [48] KUKSIN, S., PIATNITSKI, A. and SHIRIKYAN, A. (2002). A coupling approach to randomly forced nonlinear PDEs. II. *Comm. Math. Phys.* **230** 81–85. [MR1927233](#) <https://doi.org/10.1007/s00220-002-0707-2>
- [49] KUKSIN, S. and SHIRIKYAN, A. (2002). Coupling approach to white-forced nonlinear PDEs. *J. Math. Pures Appl. (9)* **81** 567–602. [MR1912412](#) [https://doi.org/10.1016/S0021-7824\(02\)01259-X](https://doi.org/10.1016/S0021-7824(02)01259-X)
- [50] KUKSIN, S. and SHIRIKYAN, A. (2012). *Mathematics of Two-Dimensional Turbulence*. *Cambridge Tracts in Mathematics* **194**. Cambridge Univ. Press, Cambridge. [MR3443633](#) <https://doi.org/10.1017/CBO9781139137119>
- [51] LEDRAPPIER, F. (1986). Positivity of the exponent for stationary sequences of matrices. In *Lyapunov Exponents (Bremen, 1984)*. *Lecture Notes in Math.* **1186** 56–73. Springer, Berlin. [MR0850070](#) <https://doi.org/10.1007/BFb0076833>
- [52] LIN, Z., THIFFEAULT, J.-L. and DOERING, C. R. (2011). Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.* **675** 465–476. [MR2801050](#) <https://doi.org/10.1017/S0022112011000292>
- [53] LIVERANI, C. (2004). On contact Anosov flows. *Ann. of Math. (2)* **159** 1275–1312. [MR2113022](#) <https://doi.org/10.4007/annals.2004.159.1275>
- [54] LUNASIN, E., LIN, Z., NOVIKOV, A., MAZZUCATO, A. and DOERING, C. R. (2012). Optimal mixing and optimal stirring for fixed energy, fixed power, or fixed palenstrophy flows. *J. Math. Phys.* **53** 115611. [MR3026556](#) <https://doi.org/10.1063/1.4752098>
- [55] MAJDA, A. J. and BERTOZZI, A. L. (2002). *Vorticity and Incompressible Flow*. *Cambridge Texts in Applied Mathematics* **27**. Cambridge Univ. Press, Cambridge. [MR1867882](#)

- [56] MATTINGLY, J. C. (2002). Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics. *Comm. Math. Phys.* **230** 421–462. [MR1937652](#) <https://doi.org/10.1007/s00220-002-0688-1>
- [57] MEYN, S. P. and TWEEDIE, R. L. (2012). *Markov Chains and Stochastic Stability*. Springer, New York.
- [58] MILES, C. J. and DOERING, C. R. (2018). Diffusion-limited mixing by incompressible flows. *Nonlinearity* **31** 2346–2350. [MR3816677](#) <https://doi.org/10.1088/1361-6544/aab1c8>
- [59] NUALART, D. (1995). *The Malliavin Calculus and Related Topics. Probability and Its Applications (New York)*. Springer, New York. [MR1344217](#) <https://doi.org/10.1007/978-1-4757-2437-0>
- [60] PESIN, Y. and CLIMENHAGA, V. (2010). Open problems in the theory of non-uniform hyperbolicity. *Discrete Contin. Dyn. Syst.* **27** 589–607. [MR2600681](#) <https://doi.org/10.3934/dcds.2010.27.589>
- [61] PETERSEN, K. (1989). *Ergodic Theory. Cambridge Studies in Advanced Mathematics 2*. Cambridge Univ. Press, Cambridge. [MR1073173](#)
- [62] PIERREHUMBERT, R. T. (1994). Tracer microstructure in the large-eddy dominated regime. *Chaos Solitons Fractals* **4** 1091–1110.
- [63] PROVENZALE, A. (1999). Transport by coherent barotropic vortices. In *Annual Review of Fluid Mechanics, Vol. 31. Annu. Rev. Fluid Mech.* **31** 55–93. Annual Reviews, Palo Alto, CA. [MR1670944](#) <https://doi.org/10.1146/annurev.fluid.31.1.55>
- [64] PROVENZALE, A., BABIANO, A., BRACCO, A., PASQUERO, C. and WEISS, J. B. (2008). Coherent vortices and tracer transport. In *Transport and Mixing in Geophysical Flows, Springer* 101–118.
- [65] ROMITO, M. (2004). Ergodicity of the finite dimensional approximation of the 3D Navier–Stokes equations forced by a degenerate noise. *J. Stat. Phys.* **114** 155–177. [MR2032128](#) <https://doi.org/10.1023/B:JOSS.0000003108.92097.5c>
- [66] RUELLE, D. (1983). Flots qui ne mélangent pas exponentiellement. *C. R. Acad. Sci. Paris Sér. I Math.* **296** 191–193. [MR0692974](#) https://doi.org/10.1142/9789812833709_0024
- [67] SARIG, O. (2002). Subexponential decay of correlations. *Invent. Math.* **150** 629–653. [MR1946554](#) <https://doi.org/10.1007/s00222-002-0248-5>
- [68] SEIS, C. (2013). Maximal mixing by incompressible fluid flows. *Nonlinearity* **26** 3279–3289. [MR3141856](#) <https://doi.org/10.1088/0951-7715/26/12/3279>
- [69] SHRAIMAN, B. I. and SIGGIA, E. D. (2000). Scalar turbulence. *Nature* **405** 639.
- [70] VARADHAN, S. R. S. (1984). *Large Deviations and Applications. CBMS-NSF Regional Conference Series in Applied Mathematics 46*. SIAM, Philadelphia, PA. [MR0758258](#) <https://doi.org/10.1137/1.9781611970241.bm>
- [71] WARHAFT, Z. (2000). Passive scalars in turbulent flows. In *Annual Review of Fluid Mechanics, Vol. 32. Annu. Rev. Fluid Mech.* **32** 203–240. Annual Reviews, Palo Alto, CA. [MR1744308](#) <https://doi.org/10.1146/annurev.fluid.32.1.203>
- [72] WUNSCH, C. and FERRARI, R. (2004). Vertical mixing, energy, and the general circulation of the oceans. In *Annual Review of Fluid Mechanics. Vol. 36. Annu. Rev. Fluid Mech.* **36** 281–314. Annual Reviews, Palo Alto, CA. [MR2062315](#) <https://doi.org/10.1146/annurev.fluid.36.050802.122121>
- [73] YAO, Y. and ZLATOŠ, A. (2017). Mixing and un-mixing by incompressible flows. *J. Eur. Math. Soc. (JEMS)* **19** 1911–1948. [MR3656475](#) <https://doi.org/10.4171/JEMS/709>
- [74] YOUNG, L.-S. (2013). Mathematical theory of Lyapunov exponents. *J. Phys. A* **46** 254001. [MR3067028](#) <https://doi.org/10.1088/1751-8113/46/25/254001>