Ergodicity of the 2D Navier-Stokes Equations with Degenerate Stochastic Forcing

Martin Hairer[1], **Jonathan C. Mattingly**[2]

[1] Department of Mathematics, The University of Warwick, Coventry CV4 7AL, United Kingdom. Email: hairer@maths.warwick.co.uk

[2] Department of Mathematics, Duke University, Durham NC 27708, USA. Email: jonm@math.duke.edu

Abstract

The stochastic 2D Navier-Stokes equations on the torus driven by degenerate noise are studied. We characterize the smallest closed invariant subspace for this model and show that the dynamics restricted to that subspace is ergodic. In particular, our results yield a purely geometric characterization of a class of noises for which the equation is ergodic in $L^2_0(\mathbb{T}^2)$. Unlike in previous works, this class is independent of the viscosity and the strength of the noise. The two main tools of our analysis are the *asymptotic strong Feller* property, introduced in this work, and an approximate integration by parts formula. The first, when combined with a weak type of irreducibility, is shown to ensure that the dynamics is ergodic. The second is used to show that the first holds under a Hörmander-type condition. This requires some interesting non-adapted stochastic analysis.

1 Introduction

In this article, we investigate the ergodic properties of the 2D Navier-Stokes equations. Recall that the Navier-Stokes equations describe the time evolution of an incompressible fluid and are given by

$$\partial_t u + (u \cdot \nabla) u = \nu \Delta u - \nabla p + \xi \;, \quad \operatorname{div} u = 0 \;, \tag{1.1}$$

where $u(x,t) \in \mathbf{R}^2$ denotes the value of the velocity field at time t and position x, $p(x,t)$ denotes the pressure, and $\xi(x,t)$ is an external force field acting on the fluid. We will consider the case when $x \in \mathbb{T}^2$, the two-dimensional torus. Our mathematical model for the driving force ξ is a Gaussian field which is white in time and colored in space. We are particularly interested in the case when only a few Fourier modes of ξ are non-zero, so that there is a well-defined "injection scale" L at which energy is pumped into the system. Remember that both the energy $\|u\|^2 = \int |u(x)|^2\,dx$ and the enstrophy $\|\nabla \wedge u\|^2$ are invariant under the nonlinearity of the 2D Navier-Stokes equations (*i.e.* they are preserved by the flow of (1.1) if $\nu = 0$ and $\xi = 0$).

From a careful study of the nonlinearity (see *e.g.* [Ros02] for a survey and [FJMR02] for some mathematical results in this field), one expects the enstrophy to cascade down to smaller and smaller scales, until it reaches a "dissipative scale" η at which the viscous term $\nu\Delta u$ dominates the nonlinearity $(u\cdot\nabla)u$ in (1.1). This picture is complemented by that of an inverse cascade of the energy towards larger and larger scales, until it is dissipated by finite-size effects as it reaches scales of order one. The physically interesting range of parameters for (1.1), where one expects to see both cascades and where the behavior of the solutions is dominated by the nonlinearity, thus corresponds to

$$1 \ll L^{-1} \ll \eta^{-1}\;. \tag{1.2}$$

The main assumptions usually made in the physics literature when discussing the behavior of (1.1) in the turbulent regime are ergodicity and statistical translational invariance of the stationary state. We give a simple geometric characterization of a class of forcings for which (1.1) is ergodic, including a forcing that acts only on 4 degrees of freedom (2 Fourier modes). This characterization is independent of the viscosity and is shown to be sharp in a certain sense. In particular, it covers the range of parameters (1.2). Since we show that the invariant measure for (1.1) is unique, its translational invariance follows immediately from the translational invariance of the equations.

From the mathematical point of view, the ergodic properties for infinite-dimensional systems are a field that has been intensely studied over the past two decades but is yet in its infancy compared to the corresponding theory for finite-dimensional systems. In particular, there is a gaping lack of results for truly hypoelliptic nonlinear systems, where the noise is transmitted to the relevant degrees of freedom only through the drift. The present article is an attempt to close this gap, at least for the particular case of the 2D Navier-Stokes equations. This particular case (and some closely related problems) has been an intense subject of study in recent years. However the results obtained so far require either a non-degenerate forcing on the "unstable" part of the equation [EMS01, KS00, BKL01, KS01, Mat02b, BKL02, Hai02, MY02], or the strong Feller property to hold. The latter was obtained only when the forcing acts on an infinite number of modes [FM95, Fer97, EH01, MS03]. The former used a change of measure via Girsanov's theorem and the pathwise contractive properties of the dynamics to prove ergodicity. In all of these works, the noise was sufficiently non-degenerate to allow in a way for an adapted analysis (see Section 4.5 below for the meaning of "adapted" in this context).

We give a fairly complete analysis of the conditions needed to ensure the ergodicity of the two dimensional Navier-Stokes equations. To do so, we employ information on the structure of the nonlinearity from [EM01] which was developed there to prove ergodicity of the finite dimensional Galerkin approximations under conditions on the forcing similar to this paper. However, our approach to the full

PDE is necessarily different and informed by the pathwise contractive properties and high/low mode splitting explained in the stochastic setting in [Mat98, Mat99] and the ideas of determining modes, inertial manifolds, and invariant subspaces in general from the deterministic PDE literature (cf. [FP67, CF88]). More directly, this paper builds on the use of the high/low splitting to prove ergodicity as first accomplished contemporaneously in [BKL01, EMS01, KS00] in the "essentially elliptic" setting (see section 4.5). In particular, this paper is the culmination of a sequence of papers by the authors and their collaborators [Mat98, Mat99, EH01, EMS01, Mat02b, Hai02, Mat03] using these and related ideas to prove ergodicity. Yet, this is the first to prove ergodicity of a stochastic PDE in a hypoelliptic setting under conditions which compare favorably to those under which similar theorems are proven for finite dimensional stochastic differential equations. One of the keys to accomplishing this is a recent result from [MP04] on the regularity of the Malliavin matrix in this setting.

One of the main technical contributions of the present work is to provide an infinitesimal replacement for Girsanov's theorem in the infinite dimensional non-adapted setting which the application of these ideas to the fully hypoelliptic setting seems to require. Another of the principal technical contributions is to observe that the strong Feller property is neither essential nor natural for the study of ergodicity in dissipative infinite-dimensional systems and to provide an alternative. We define instead a weaker *asymptotic strong Feller* property which is satisfied by the system under consideration and is sufficient to give ergodicity. In many dissipative systems, including the stochastic Navier-Stokes equations, only a finite number of modes are unstable. Conceivably, these systems are ergodic even if the noise is transmitted only to those unstable modes rather than to the whole system. The asymptotic strong Feller property captures this idea. It is sensitive to the regularization of the transition densities due to both probabilistic and dynamic mechanisms.

This paper is organized as follows. In Section 2 the precise mathematical formulation of the problem and the main results for the stochastic Navier-Stokes equations are given. In Section 3 we define the asymptotic strong Feller property and prove in Theorem 3.16 that, together with an irreducibility property it implies ergodicity of the system. We thus obtain the analog in our setting of the classical result often derived from theorems of Khasminskii and Doob which states that topological irreducibility, together with the strong Feller property, implies uniqueness of the invariant measure. The main technical results are given in Section 4, where we show how to apply the abstract results to our problem. Although this section is written with the stochastic Navier-Stokes equations in mind, most of the corresponding results hold for a much wider class of stochastic PDEs with polynomial nonlinearities.

Acknowledgements

We would like to thank G. Ben Arous, W. E, J. Hanke, X.-M. Li, E. Pardoux, M. Romito and Y. Sinai for motivating and useful discussions. We would also like to thank the anonymous referees for their careful reading of the text and their subsequent corrections and useful suggestions. The work of MH is partially supported by the Fonds National Suisse. The work of JCM was partially supported by the Institut Universitaire de France.

2 Setup and Main Results

Consider the two-dimensional, incompressible Navier-Stokes equations on the torus $\mathbb{T}^2 = [-\pi, \pi]^2$ driven by a degenerate noise. Since the velocity and vorticity formulations are equivalent in this setting, we choose to use the vorticity equation as this simplifies the exposition. For u a divergence-free velocity field, we define the vorticity w by $w = \nabla \wedge u = \partial_2 u_1 - \partial_1 u_2$. Note that u can be recovered from w and the condition $\nabla \cdot u = 0$. With these notations the vorticity formulation for the stochastic Navier-Stokes equations is as follows:

$$dw = \nu \Delta w \, dt + B(\mathcal{K}w, w) \, dt + Q \, dW(t) \;, \tag{2.1}$$

where Δ is the Laplacian with periodic boundary conditions and $B(u, w) = -(u \cdot \nabla)w$, the usual Navier-Stokes nonlinearity. The symbol $Q \, dW(t)$ denotes a Gaussian noise process which is white in time and whose spatial correlation structure will be described later. The operator $\mathcal{K}$ is defined in Fourier space by $(\mathcal{K}w)_k = -iw_k k^\perp / \|k\|^2$, where $(k_1, k_2)^\perp = (k_2, -k_1)$. By w_k, we mean the scalar product of w with $(2\pi)^{-1} \exp(ik \cdot x)$. It has the property that the divergence of $\mathcal{K}w$ vanishes and that $w = \nabla \wedge (\mathcal{K}w)$. Unless otherwise stated, we consider (2.1) as an equation in $\mathcal{H} = \mathrm{L}^2_0$, the space of real-valued square-integrable functions on the torus with vanishing mean. Before we go on to describe the noise process QW, it is instructive to write down the two-dimensional Navier-Stokes equations (without noise) in Fourier space:

$$\dot{w}_k = -\nu |k|^2 w_k - \frac{1}{4\pi} \sum_{j+\ell=k} \langle j^\perp, \ell \rangle \Big(\frac{1}{|\ell|^2} - \frac{1}{|j|^2} \Big) w_j w_\ell \;. \tag{2.2}$$

From (2.2), we see clearly that any closed subspace of $\mathcal{H}$ spanned by Fourier modes corresponding to a subgroup of $\mathbf{Z}^2$ is invariant under the dynamics. In other words, if the initial condition has a certain type of periodicity, it will be retained by the solution for all times.

In order to describe the noise $Q \, dW(t)$, we start by introducing a convenient way to index the Fourier basis of $\mathcal{H}$. We write $\mathbf{Z}^2 \setminus \{(0,0)\} = \mathbf{Z}^2_+ \cup \mathbf{Z}^2_-$, where

$$\begin{aligned} \mathbf{Z}^2_+ &= \{(k_1, k_2) \in \mathbf{Z}^2 \,|\, k_2 > 0\} \cup \{(k_1, 0) \in \mathbf{Z}^2 \,|\, k_1 > 0\} \;, \\ \mathbf{Z}^2_- &= \{(k_1, k_2) \in \mathbf{Z}^2 \,|\, -k \in \mathbf{Z}^2_+\} \;, \end{aligned}$$

(note that $\mathbf{Z}^2_+$ is essentially the upper half-plane) and set, for $k \in \mathbf{Z}^2 \setminus \{(0,0)\}$,

$$f_k(x) = \begin{cases} \sin(k \cdot x) & \text{if } k \in \mathbf{Z}^2_+, \\ \cos(k \cdot x) & \text{if } k \in \mathbf{Z}^2_-. \end{cases}$$

We also fix a set

$$\mathcal{Z}_0 = \{k_n \,|\, n = 1, \ldots, m\} \subset \mathbf{Z}^2 \setminus \{(0,0)\} , \tag{2.3}$$

which encodes the geometry of the driving noise. The set $\mathcal{Z}_0$ will correspond to the set of driven modes of the equation (2.1).

The process $W(t)$ is an m-dimensional Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. For definiteness, we choose Ω to be the Wiener space $\mathcal{C}_0([0,\infty), \mathbf{R}^m)$, W the canonical process, and $\mathbf{P}$ the Wiener measure. We denote expectations with respect to $\mathbf{P}$ by $\mathbf{E}$ and define $\mathcal{F}_t$ to be the σ-algebra generated by the increments of W up to time t. We also denote by $\{e_n\}$ the canonical basis of $\mathbf{R}^m$. The linear map $Q : \mathbf{R}^m \to \mathcal{H}$ is given by $Qe_n = q_n f_{k_n}$, where the q_n are some strictly positive numbers, and the wavenumbers k_n are given by the elements of $\mathcal{Z}_0$. With these definitions, QW is an $\mathcal{H}$-valued Wiener process. We also denote the average rate at which energy is injected into our system by $\mathcal{E}_0 = \operatorname{tr} QQ^* = \sum_n q_n^2$.

We assume that the set $\mathcal{Z}_0$ is symmetric, *i.e.* that if $k \in \mathcal{Z}_0$, then $-k \in \mathcal{Z}_0$. This is not a strong restriction and is made only to simplify the statements of our results. It also helps to avoid the possible confusion arising from the slightly non-standard definition of the basis f_k. This assumption always holds for example if the noise process QW is taken to be translation invariant. In fact, Theorem 2.1 below holds for non-symmetric sets $\mathcal{Z}_0$ if one replaces $\mathcal{Z}_0$ in the theorem's conditions by its symmetric part.

It is well-known [Fla94, MR] that (2.1) defines a stochastic flow on $\mathcal{H}$. By a stochastic flow, we mean a family of continuous maps $\Phi_t \colon \Omega \times \mathcal{H} \to \mathcal{H}$ such that $w_t = \Phi_t(W, w_0)$ is the solution to (2.1) with initial condition w_0 and noise W. Hence, its transition semigroup $\mathcal{P}_t$ given by $\mathcal{P}_t\varphi(w_0) = \mathbf{E}_{w_0}\varphi(w_t)$ is Feller. Here, φ denotes any bounded measurable function from $\mathcal{H}$ to $\mathbf{R}$ and we use the notation $\mathbf{E}_{w_0}$ for expectations with respect to solutions to (2.1) with initial condition w_0. Recall that an *invariant measure* for (2.1) is a probability measure $\mu_\star$ on $\mathcal{H}$ such that $\mathcal{P}_t^*\mu_\star = \mu_\star$, where $\mathcal{P}_t^*$ is the semigroup on measures dual to $\mathcal{P}_t$. While the existence of an invariant measure for (2.1) can be proved by "soft" techniques using the regularizing and dissipativity properties of the flow [Cru89, Fla94], showing its uniqueness is a challenging problem that requires a detailed analysis of the nonlinearity. The importance of showing the uniqueness of $\mu_\star$ is illustrated by the fact that it implies

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \varphi(w_t)\, dt = \int_{\mathcal{H}} \varphi(w)\, \mu_\star(dw) ,$$

for all bounded continuous functions φ and $\mu_\star$-almost every initial condition $w_0 \in \mathcal{H}$. It thus gives some mathematical ground to the *ergodic assumption* usually made in the physics literature when discussing the qualitative behavior of (2.1). The main results of this article are summarized by the following theorem:

Theorem 2.1 *Let $\mathcal{Z}_0$ satisfy the following two assumptions:*

A1 *There exist at least two elements in $\mathcal{Z}_0$ with different Euclidean norms.*

A2 *Integer linear combinations of elements of $\mathcal{Z}_0$ generate $\mathbf{Z}^2$.*

Then, (2.1) has a unique invariant measure in $\mathcal{H}$.

Remark 2.2 As pointed out by J. Hanke, condition **A2** above is equivalent to the easily verifiable condition that the greatest common divisor of the set $\{ \det(k, \ell) : k, \ell \in \mathcal{Z}_0\}$ is 1, where $\det(k, \ell)$ is the determinant of the 2×2 matrix with columns k and ℓ.

The proof of Theorem 2.1 is given by combining Corollary 4.2 with Proposition 4.4 below. A partial converse of this ergodicity result is given by the following theorem, which is an immediate consequence of Proposition 4.4.

Theorem 2.3 *There are two qualitatively different ways in which the hypotheses of Theorem 2.1 can fail. In each case there is a unique invariant measure supported on $\tilde{\mathcal{H}}$, the smallest closed linear subspace of $\mathcal{H}$ which is invariant under (2.1).*

- *In the first case the elements of $\mathcal{Z}_0$ are all collinear or of the same Euclidean length. Then $\tilde{\mathcal{H}}$ is the finite-dimensional space spanned by $\{f_k \,|\, k \in \mathcal{Z}_0\}$, and the dynamics restricted to $\tilde{\mathcal{H}}$ is that of an Ornstein-Uhlenbeck process.*
- *In the second case let $\mathcal{G}$ be the smallest subgroup of $\mathbf{Z}^2$ containing $\mathcal{Z}_0$. Then $\tilde{\mathcal{H}}$ is the space spanned by $\{f_k \,|\, k \in \mathcal{G} \setminus \{(0,0)\}\}$. Let k_1, k_2 be two generators for $\mathcal{G}$ and define $v_i = 2\pi k_i/|k_i|^2$, then $\tilde{\mathcal{H}}$ is the space of functions that are periodic with respect to the translations v_1 and v_2.*

Remark 2.4 That $\tilde{\mathcal{H}}$ constructed above is invariant is clear; that it is the smallest invariant subspace follows from the fact that the transition probabilities of (2.1) have a density with respect to the Lebesgue measure when projected onto any finite-dimensional subspace of $\tilde{\mathcal{H}}$, see [MP04].

By Theorem 2.3 if the conditions of Theorem 2.1 are not satisfied then one of the modes with lowest wavenumber is in $\tilde{\mathcal{H}}^\perp$. In fact either $f_{(1,0)} \perp \tilde{\mathcal{H}}$ or $f_{(1,1)} \perp \tilde{\mathcal{H}}$. On the other hand for sufficiently small values of ν the low modes of (2.1) are expected to be linearly unstable [Fri95]. If this is the case, a solution to (2.1) starting in $\tilde{\mathcal{H}}^\perp$ will not converge to $\tilde{\mathcal{H}}$ and (2.1) is therefore expected to have several distinct invariant measures on $\mathcal{H}$. It is however known that the invariant measure is unique if the viscosity is sufficiently high, see [Mat99]. (At high viscosity, all modes are linearly stable. See [Mat03] for a more streamlined presentation.)

Example 2.5 The set $\mathcal{Z}_0 = \{(1,0), (-1,0), (1,1), (-1,-1)\}$ satisfies the assumptions of Theorem 2.1. Therefore, (2.1) with noise given by

$$\begin{aligned} QW(t,x) = W_1(t)\sin x_1 + W_2(t)\cos x_1 + W_3(t)\sin(x_1 + x_2) \\ + W_4(t)\cos(x_1 + x_2)\;, \end{aligned}$$

has a unique invariant measure in $\mathcal{H}$ for every value of the viscosity $\nu > 0$.

Example 2.6 Take $\mathcal{Z}_0 = \{(1,0),(-1,0),(0,1),(0,-1)\}$ whose elements are of length 1. Therefore, (2.1) with noise given by

$$QW(t,x) = W_1(t)\sin x_1 + W_2(t)\cos x_1 + W_3(t)\sin x_2 + W_4(t)\cos x_2 \;,$$

reduces to an Ornstein-Uhlenbeck process on the space spanned by $\sin x_1$, $\cos x_1$, $\sin x_2$, and $\cos x_2$.

Example 2.7 Take $\mathcal{Z}_0 = \{(2,0),(-2,0),(2,2),(-2,-2)\}$, which corresponds to case 2 of Theorem 2.3 with $\mathcal{G}$ generated by $(0,2)$ and $(2,0)$. In this case, $\tilde{\mathcal{H}}$ is the set of functions that are π-periodic in both arguments. Via the change of variables $x \mapsto x/2$, one can easily see from Theorem 2.1 that (2.1) then has a unique invariant measure on $\tilde{\mathcal{H}}$ (but not necessarily on $\mathcal{H}$).

3 An Abstract Ergodic Result

We start by proving an abstract ergodic result, which lays the foundations of the present work. Recall that a Markov transition semigroup $\mathcal{P}_t$ is said to be *strong Feller* at time t if $\mathcal{P}_t\varphi$ is continuous for every bounded measurable function φ. It is a well-known and much used fact that the strong Feller property, combined with some irreducibility of the transition probabilities implies the uniqueness of the invariant measure for $\mathcal{P}_t$ [DPZ96, Theorem 4.2.1]. If $\mathcal{P}_t$ is generated by a diffusion with smooth coefficients on $\mathbf{R}^n$ or a finite-dimensional manifold, Hörmander's theorem [Hör67, Hör85] provides us with an efficient (and sharp if the coefficients are analytic) criteria for the strong Feller property to hold. Unfortunately, no equivalent theorem exists if $\mathcal{P}_t$ is generated by a diffusion in an infinite-dimensional space, where the strong Feller property seems to be much "rarer". If the covariance of the noise is non-degenerate (*i.e.* the diffusion is elliptic in some sense), the strong Feller property can often be recovered by means of the Bismut-Elworthy-Li formula [EL94]. The only result to our knowledge that shows the strong Feller property for an infinite-dimensional diffusion where the covariance of the noise does not have a dense range is given in [EH01], but it still requires the forcing to act in a non-degenerate way on a subspace of finite codimension.

3.1 Preliminary definitions

Let $\mathcal{X}$ be a Polish (*i.e.* complete, separable, metrizable) space. Recall that a *pseudo-metric* for $\mathcal{X}$ is a continuous function $d : \mathcal{X}^2 \to \mathbf{R}_+$ such that $d(x,x) = 0$ and such that the triangle inequality is satisfied. We say that a pseudo-metric d_1 is larger than d_2 if $d_1(x,y) \geq d_2(x,y)$ for all $(x,y) \in \mathcal{X}^2$.

Definition 3.1 Let $\{d_n\}_{n=0}^\infty$ be an increasing sequence of (pseudo-)metrics on a Polish space $\mathcal{X}$. If $\lim_{n\to\infty} d_n(x,y) = 1$ for all $x \neq y$, then $\{d_n\}$ is a *totally separating system of (pseudo-)metrics* for $\mathcal{X}$.

Let us give a few representative examples.

Example 3.2 Let $\{a_n\}$ be an increasing sequence in **R** such that $\lim_{n\to\infty} a_n = \infty$. Then, $\{d_n\}$ is a *totally separating system of (pseudo-)metrics* for $\mathcal{X}$ in the following three cases.

1. Let d be an arbitrary continuous metric on $\mathcal{X}$ and set $d_n(x, y) = 1 \wedge a_n d(x, y)$.
2. Let $\mathcal{X} = \mathcal{C}_0(\mathbf{R})$ be the space of continuous functions on **R** vanishing at infinity and set $d_n(x, y) = 1 \wedge \sup_{s\in[-n,n]} a_n|x(s) - y(s)|$.
3. Let $\mathcal{X} = \ell^2$ and set $d_n(x, y) = 1 \wedge a_n \sum_{k=0}^n |x_k - y_k|^2$.

Given a pseudo-metric d, we define the following seminorm on the set of d-Lipschitz continuous functions from $\mathcal{X}$ to **R**:

$$\|\varphi\|_d = \sup_{\substack{x,y\in\mathcal{X}\\ x\neq y}} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)} . \tag{3.1}$$

This in turn defines a dual seminorm on the space of finite signed Borel measures on $\mathcal{X}$ with vanishing integral by

$$|\!|\!|\nu|\!|\!|_d = \sup_{\|\varphi\|_d=1} \int_{\mathcal{X}} \varphi(x)\, \nu(dx) . \tag{3.2}$$

Given μ_1 and μ_2, two positive finite Borel measures on $\mathcal{X}$ with equal mass, we also denote by $\mathscr{C}(\mu_1, \mu_2)$ the set of positive measures on $\mathcal{X}^2$ with marginals μ_1 and μ_2 and we define

$$\|\mu_1 - \mu_2\|_d = \inf_{\mu\in\mathscr{C}(\mu_1,\mu_2)} \int_{\mathcal{X}^2} d(x, y)\, \mu(dx, dy) . \tag{3.3}$$

The following lemma is an easy consequence of the Monge-Kantorovich duality, see *e.g.* [Kan42, Kan48, AN87], and shows that in most cases these two natural notions of distance can be used interchangeably.

Lemma 3.3 *Let d be a continuous pseudo-metric on a Polish space $\mathcal{X}$ and let μ_1 and μ_2 be two positive measures on $\mathcal{X}$ with equal mass. Then, one has $\|\mu_1 - \mu_2\|_d = |\!|\!|\mu_1 - \mu_2|\!|\!|_d$.*

Proof. This result is well-known if $(\mathcal{X}, d)$ is a separable metric space, see for example [Rac91] for a detailed discussion on many of its variants. If we define an equivalence relation on $\mathcal{X}$ by $x \sim y \Leftrightarrow d(x, y) = 0$ and set $\mathcal{X}_d = \mathcal{X}/\sim$, then d is well-defined on $\mathcal{X}_d$ and $(\mathcal{X}_d, d)$ is a separable metric space (although it may no longer be complete). Defining $\pi : \mathcal{X} \to \mathcal{X}_d$ by $\pi(x) = [x]$, the result follows from the Monge-Kantorovich duality in $\mathcal{X}_d$ and the fact that both sides of (3.3) do not change if the measures μ_i are replaced by $\pi^*\mu_i$. □

Recall that the total variation norm of a finite signed measure μ on $\mathcal{X}$ is given by $\|\mu\|_{\mathrm{TV}} = \frac{1}{2}(\mu^+(\mathcal{X})+\mu^-(\mathcal{X}))$, where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ. The next result is crucial to the approach taken in this paper.

Lemma 3.4 *Let $\{d_n\}$ be a bounded and increasing family of continuous pseudo-metrics on a Polish space $\mathcal{X}$ and define $d(x,y) = \lim_{n\to\infty} d_n(x,y)$. Then, one has $\lim_{n\to\infty} \|\mu_1 - \mu_2\|_{d_n} = \|\mu_1 - \mu_2\|_d$ for any two positive measures μ_1 and μ_2 with equal mass.*

Proof. The limit exists since the sequence is bounded and increasing by assumption, so let us denote this limit by L. It is clear from (3.3) that $\|\mu_1 - \mu_2\|_d \geq L$, so it remains to show the converse bound. Let μ_n be a measure in $\mathscr{C}(\mu_1, \mu_2)$ that realizes (3.3) for the distance d_n. (Such a measure is shown to exist in [Rac91].) The sequence $\{\mu_n\}$ is tight on $\mathcal{X}^2$ since its marginals are constant, so we can extract a weakly converging subsequence. Denote by μ_∞ the limiting measure. For $m \geq n$

$$\int_{\mathcal{X}^2} d_n(x,y)\, \mu_m(dx,dy) \leq \int_{\mathcal{X}^2} d_m(x,y)\, \mu_m(dx,dy) \leq L\;.$$

Since d_n is continuous, the weak convergence taking $m \to \infty$ implies that

$$\int_{\mathcal{X}^2} d_n(x,y)\, \mu_\infty(dx,dy) \leq L\;, \quad \forall\, n > 0\;.$$

It follows from the dominated convergence theorem that $\int_{\mathcal{X}^2} d(x,y)\, \mu_\infty(dx,dy) \leq L$, which concludes the proof. □

Corollary 3.5 *Let $\mathcal{X}$ be a Polish space and let $\{d_n\}$ be a totally separating system of pseudo-metrics for $\mathcal{X}$. Then, $\|\mu_1 - \mu_2\|_{\text{TV}} = \lim_{n\to\infty} \|\mu_1 - \mu_2\|_{d_n}$ for any two positive measures μ_1 and μ_2 with equal mass on $\mathcal{X}$.*

Proof. It suffices to notice that $\|\mu_1 - \mu_2\|_{\text{TV}} = \inf_{\mu\in\mathscr{C}(\mu_1,\mu_2)} \mu(\{(x,y) : x \neq y\}) = \|\mu_1 - \mu_2\|_d$ with $d(x,y) = 1$ whenever $x \neq y$ and then to apply Lemma 3.4. Observe that $d_n \to d$ by the definition of a totally separating system of pseudo-metrics and that Lemma 3.4 makes no assumptions on the continuity of the limiting pseudo-metric d. □

3.2 Asymptotic Strong Feller

Before we define the asymptotic strong Feller property, recall that:

Definition 3.6 A Markov transition semigroup on a Polish space $\mathcal{X}$ is said to be *strong Feller* at time t if $\mathcal{P}_t\varphi$ is continuous for every bounded measurable function $\varphi : \mathcal{X} \to \mathbf{R}$.

Note that if the transition probabilities $\mathcal{P}_t(x,\cdot\,)$ are continuous in x in the total variation topology, then $\mathcal{P}_t$ is strong Feller at time t.

Recall also that the support of a probability measure μ, denoted by supp(μ), is the intersection of all closed sets of measure 1. A useful characterization of the support of a measure is given by

Lemma 3.7 *A point $x \in \operatorname{supp}(\mu)$ if and only if $\mu(U) > 0$ for every open set U containing x.* □

It is well-known that if a Markov transition semigroup $\mathcal{P}_t$ is strong Feller and μ_1 and μ_2 are two distinct ergodic invariant measures for $\mathcal{P}_t$ (*i.e.* μ_1 and μ_2 are mutually singular), then $\operatorname{supp}\mu_1 \cap \operatorname{supp}\mu_2 = \phi$. (This can be seen *e.g.* by the same argument as in [DPZ96, Prop. 4.1.1].) In this section, we show that this property still holds if the strong Feller property is replaced by the following property, where we denote by $\mathcal{U}_x$ the collection of all open sets containing x.

Definition 3.8 A Markov transition semigroup $\mathcal{P}_t$ on a Polish space $\mathcal{X}$ is called *asymptotically strong Feller* at x if there exists a totally separating system of pseudo-metrics $\{d_n\}$ for $\mathcal{X}$ and a sequence $t_n > 0$ such that

$$\inf_{U\in\mathcal{U}_x} \limsup_{n\to\infty} \sup_{y\in U} \|\mathcal{P}_{t_n}(x,\cdot\,) - \mathcal{P}_{t_n}(y,\cdot\,)\|_{d_n} = 0\;, \tag{3.4}$$

It is called asymptotically strong Feller if this property holds at every $x \in \mathcal{X}$.

Remark 3.9 If $\mathcal{B}(x,\gamma)$ denotes the open ball of radius γ centered at x in some metric defining the topology of $\mathcal{X}$, then it is immediate that (3.4) is equivalent to

$$\lim_{\gamma\to 0} \limsup_{n\to\infty} \sup_{y\in\mathcal{B}(x,\gamma)} \|\mathcal{P}_{t_n}(x,\cdot\,) - \mathcal{P}_{t_n}(y,\cdot\,)\|_{d_n} = 0\;.$$

Remark 3.10 Notice that the definition of asymptotically strong Feller allows the possibility that $t_n = t$ for all n. In this case, the transition probabilities $\mathcal{P}_t(x,\cdot\,)$ are continuous in the total variation topology and thus $\mathcal{P}_s$ is strong Feller at times $s \geq t$. Conversely, it is *almost* true that all strong Feller processes are asymptotically strong Feller. The small discrepancy arises from the fact that strong Feller implies continuity of transition densities only in the strong topology while asymptotically strong Feller with t_n constant implies continuity in the topology of total variation convergence, which is a stronger topology. It might be useful to change the definition of asymptotically strong Feller to remove this discrepancy, however this change does not seem natural in the present context.

One way of seeing the connection to the strong Feller property is to recall that a standard criteria for $\mathcal{P}_t$ to be strong Feller is given by [DPZ96, Lem. 7.1.5]:

Proposition 3.11 *A semigroup $\mathcal{P}_t$ on a Hilbert space $\mathcal{H}$ is strong Feller if, for all $\varphi : \mathcal{H} \to \mathbf{R}$ with $\|\varphi\|_\infty \stackrel{\text{def}}{=} \sup_{x\in\mathcal{H}} |\varphi(x)|$ and $\|\nabla\varphi\|_\infty$ finite one has*

$$|\nabla\mathcal{P}_t\varphi(x)| \leq C(\|x\|)\|\varphi\|_\infty\;, \tag{3.5}$$

where $C : \mathbf{R}_+ \to \mathbf{R}$ is a fixed non-decreasing function. □

The following lemma provides a similar criteria for the asymptotic strong Feller property:

Proposition 3.12 *Let t_n and δ_n be two positive sequences with $\{t_n\}$ non-decreasing and $\{\delta_n\}$ converging to zero. A semigroup $\mathcal{P}_t$ on a Hilbert space $\mathcal{H}$ is asymptotically strong Feller if, for all $\varphi : \mathcal{H} \to \mathbf{R}$ with $\|\varphi\|_\infty$ and $\|\nabla\varphi\|_\infty$ finite one has*

$$|\nabla \mathcal{P}_{t_n}\varphi(x)| \leq C(\|x\|)(\|\varphi\|_\infty + \delta_n \|\nabla\varphi\|_\infty) \tag{3.6}$$

for all n, where $C : \mathbf{R}_+ \to \mathbf{R}$ is a fixed non-decreasing function.

Proof. For $\varepsilon > 0$, we define on $\mathcal{H}$ the distance $d_\varepsilon(w_1, w_2) = 1 \wedge \varepsilon^{-1}\|w_1 - w_2\|$, and we denote by $\|\cdot\|_\varepsilon$ the corresponding seminorms on functions and on measures given by (3.1) and (3.2). It is clear that if δ_n is a decreasing sequence converging to 0, $\{d_{\delta_n}\}$ is a totally separating system of metrics for $\mathcal{H}$.

It follows immediately from (3.6) that for every Fréchet differentiable function φ from $\mathcal{H}$ to $\mathbf{R}$ with $\|\varphi\|_\varepsilon \leq 1$ one has

$$\int_{\mathcal{H}} \varphi(w)\,(\mathcal{P}_{t_n}(w_1, dw) - \mathcal{P}_{t_n}(w_2, dw)) \leq \|w_1 - w_2\| C(\|w_1\| \vee \|w_2\|)\Big(1 + \frac{\delta_n}{\varepsilon}\Big)\;. \tag{3.7}$$

Now take a Lipschitz continuous function φ with $\|\varphi\|_\varepsilon \leq 1$. By applying to φ the semigroup at time $1/m$ corresponding to a linear Strong Feller diffusion in $\mathcal{H}$, one obtains [Cer99, DPZ96] a sequence φ_m of Fréchet differentiable approximations φ_m with $\|\varphi_m\|_\varepsilon \leq 1$ and such that $\varphi_m \to \varphi$ pointwise. Therefore, by the dominated convergence theorem, (3.7) holds for Lipschitz continuous functions φ and so

$$\|\mathcal{P}_{t_n}(w_1, \cdot\,) - \mathcal{P}_{t_n}(w_2, \cdot\,)\|_\varepsilon \leq \|w_1 - w_2\| C(\|w_1\| \vee \|w_2\|)\Big(1 + \frac{\delta_n}{\varepsilon}\Big)\;,$$

Choosing $\varepsilon = a_n = \sqrt{\delta_n}$, we obtain

$$\|\mathcal{P}_{t_n}(w_1, \cdot\,) - \mathcal{P}_{t_n}(w_2, \cdot\,)\|_{a_n} \leq \|w_1 - w_2\| C(\|w_1\| \vee \|w_2\|)(1 + a_n)\;,$$

which in turn implies that $\mathcal{P}_t$ is asymptotically strong Feller since $a_n \to 0$. □

Example 3.13 Consider the SDE

$$dx = -x\,dt + dW(t)\;, \qquad dy = -y\,dt\;.$$

Then, the corresponding Markov semigroup $\mathcal{P}_t$ on $\mathbf{R}^2$ is not strong Feller, but it is asymptotically strong Feller. To see that $\mathcal{P}_t$ is not strong Feller, let $\varphi(x, y) = \mathrm{sgn}(y)$ and observe that $\mathcal{P}_t\varphi = \varphi$ for all $t \in [0, \infty)$. Since φ is bounded but not continuous, the system is not strong Feller. To see that the system is asymptotically strong Feller observe that for any differentiable $\varphi : \mathbf{R}^2 \to \mathbf{R}$ and any direction $\xi \in \mathbf{R}^2$ with $\|\xi\| = 1$, one has

$$\begin{aligned} |(\nabla \mathcal{P}_t\varphi)(x_0, y_0) \cdot \xi| &= |\mathbf{E}_{(x_0,y_0)}(\nabla\varphi)(x_t, y_t) \cdot (u_t, v_t)| \\ &\leq \|\nabla\varphi\|_\infty \mathbf{E}|(u_t, v_t)| \leq \|\nabla\varphi\|_\infty e^{-t}\;, \end{aligned}$$

where (u_t, v_t) is the linearized flow starting from ξ. In other words $(u_0, v_0) = \xi$, $du = -udt$, and $dv = -vdt$. This is a particularly simple example because the flow is globally contractive.

Example 3.14 Now consider the SDE

$$dx = (x - x^3)\,dt + dW(t)\;, \qquad dy = -y\,dt\;.$$

Again the function $\varphi(x, y) = \text{sgn}(y)$ is invariant under $\mathcal{P}_t$ implying that the system is not strong Feller. It is however not globally contractive as in the previous example. As in the previous example, let $\xi = (\xi_1, \xi_2) \in \mathbf{R}^2$ with and $\|\xi\| = 1$ and now let (u_t, v_t) denote the linearizion of this equation with $(u_0, v_0) = \xi$. Let $\mathcal{P}_t^x$ denote the Markov transition semigroup of the x_t process. It is a classical fact that for such a uniformly elliptic diffusion with a unique invariant measure one has $|\partial_x \mathcal{P}_t^x \varphi(x, y)| \leq C(|x|)\|\varphi\|_\infty$ for some non-decreasing function C and all $t \geq 1$. Hence differentiating with respect to both initial conditions produces

$$\begin{aligned}
|(\nabla \mathcal{P}_t \varphi)(x_0, y_0) \cdot \xi| &= |(\partial_x \mathcal{P}_t^x \varphi(x, y)\xi_1) + \mathbf{E}((\partial_y \varphi)(x_t, \tilde{y}_t)v_t)| \\
&\leq C(|x|)\|\varphi\|_\infty + \mathbf{E}|v_t|\|\nabla\varphi\|_\infty \\
&\leq (C(|x|) + 1)(\|\varphi\|_\infty + e^{-t}\|\nabla\varphi\|_\infty)
\end{aligned}$$

for $t \geq 1$ which implies that the system is asymptotically strong Feller.

Example 3.15 In infinite dimensions, even a seemingly non-degenerate diffusion can suffer from a similar problem. Consider the following infinite dimensional Ornstein-Uhlenbeck process $u(x, t) = \sum \hat{u}(k, t)\exp(ikx)$ written in terms of its complex Fourier coefficients. We take $x \in \mathbb{T} = [-\pi, \pi]$, $k \in \mathbf{Z}$ and

$$d\hat{u}(k, t) = -(1 + |k|^2)\hat{u}(k, t)\,dt + \exp(-|k|^3)\,d\beta_k(t)\;, \tag{3.8}$$

where the β_k are independent standard complex Brownian motions. The Markov transition densities $\mathcal{P}_t(x, \cdot)$ and $\mathcal{P}_t(y, \cdot)$ are singular for all finite times if $x - y$ is not sufficiently smooth. This implies that the diffusion (3.8) in $\mathcal{H} = \mathrm{L}^2([-\pi, \pi])$ is *not* strong Feller since by Lemma 7.2.1 of [DPZ96] the strong Feller property is equivalent to $\mathcal{P}_t(y, \cdot)$ being equivalent to $\mathcal{P}_t(x, \cdot)$ for all x and y. Another equivalent characterization of the strong Feller property is that the image$(S_t) \subset$ image(Q_t) where S_t is the linear semigroup generated by the deterministic part for the equation defined by $(S_t u)(k) = e^{-(1+|k|^2)t}u(k, 0)$ and $Q_t = \int_0^t S_r G S_r^* dr$ where G is the covariance operator of the noise defined by $(Gu)(k) = \exp(-2|k|^3)u(k)$. This captures the fact that the mean, controlled by S_t, is moving towards zero too slowly relative to the decay of the noise's covariance structure. However, one can easily check that the example is asymptotically strong Feller since the entire flow is pathwise contractive like in the first example.

The classical strong Feller property captures well the smoothing due to the random effects. When combined with irreducibility in the same topology, it implies that the transition densities starting from different points are mutually absolutely continuous. As the examples show, this is often not true in infinite dimensions. We see that the asymptotic strong Feller property better incorporates the smoothing due to the pathwise contraction of the dynamics. Comparing Proposition 3.11

with Proposition 3.12, one sees that the second term in Proposition 3.12 allows one to capture the progressive smoothing in time from the pathwise dynamics. This becomes even clearer when one examines the proofs of Proposition 4.3 and Proposition 4.11 later in the text. There one sees that the first term comes from shifting a derivative from the test function to the Wiener measure and the second is controlled using in an essential way the contraction due to the spatial Laplacian.

The usefulness of the asymptotic strong Feller property is seen in the following theorem and its accompanying corollary which are the main results of this section.

Theorem 3.16 *Let $\mathcal{P}_t$ be a Markov semigroup on a Polish space $\mathcal{X}$ and let μ and ν be two distinct ergodic invariant probability measures for $\mathcal{P}_t$. If $\mathcal{P}_t$ is asymptotically strong Feller at x, then $x \notin \operatorname{supp}\mu \cap \operatorname{supp}\nu$.*

Proof. Using Corollary 3.5, the proof of this result is a simple rewriting of the proof of the corresponding result for strong Feller semigroups.

For every measurable set A, every $t > 0$, and every pseudo-metric d on $\mathcal{X}$ with $d \le 1$, the triangle inequality for $\|\cdot\|_d$ implies

$$\|\mu - \nu\|_d \le 1 - \min\{\mu(A), \nu(A)\}\Big(1 - \max_{y,z\in A} \|\mathcal{P}_t(z,\cdot) - \mathcal{P}_t(y,\cdot)\|_d\Big)\;. \tag{3.9}$$

To see this, set $\alpha = \min\{\mu(A), \nu(A)\}$. If $\alpha = 0$ there is nothing to prove so assume $\alpha > 0$. Clearly there exist probability measures $\bar\nu$, $\bar\mu$, ν_A, and μ_A such that $\nu_A(A) = \mu_A(A) = 1$ and such that $\mu = (1-\alpha)\bar\mu + \alpha\mu_A$ and $\nu = (1-\alpha)\bar\nu + \alpha\nu_A$. Using the invariance of the measures μ and ν and the triangle inequality implies

$$\begin{aligned}\|\mu - \nu\|_d &= \|\mathcal{P}_t\mu - \mathcal{P}_t\nu\|_d \le (1-\alpha)\|\mathcal{P}_t\bar\mu - \mathcal{P}_t\bar\nu\|_d + \alpha\|\mathcal{P}_t\mu_A - \mathcal{P}_t\nu_A\|_d \\ &\le (1-\alpha) + \alpha\int_A\int_A \|\mathcal{P}_t(z,\cdot) - \mathcal{P}_t(y,\cdot)\|_d \mu_A(dz)\nu_A(dy) \\ &\le 1 - \alpha\Big(1 - \max_{y,z\in A}\|\mathcal{P}_t(z,\cdot) - \mathcal{P}_t(y,\cdot)\|_d\Big)\;.\end{aligned}$$

Continuing with the proof of the corollary, by the definition of the asymptotic strong Feller property, there exist constants $N > 0$, a sequence of totally separating pseudo-metrics $\{d_n\}$, and an open set U containing x such that $\|\mathcal{P}_{t_n}(z,\cdot) - \mathcal{P}_{t_n}(y,\cdot)\|_{d_n} \le 1/2$ for every $n > N$ and every $y, z \in U$. (Note that by the definition of totally separating pseudo-metrics $d_n \le 1$.)

Assume by contradiction that $x \in \operatorname{supp}\mu \cap \operatorname{supp}\nu$ and therefore that $\alpha = \min(\mu(U), \nu(U)) > 0$. Taking $A = U$, $d = d_n$, and $t = t_n$ in (3.9), we then get $\|\mu - \nu\|_{d_n} \le 1 - \frac{\alpha}{2}$ for every $n > N$, and therefore $\|\mu - \nu\|_{\mathrm{TV}} \le 1 - \frac{\alpha}{2}$ by Corollary 3.5, thus leading to a contradiction. □

As an immediate corollary, we have

Corollary 3.17 *If $\mathcal{P}_t$ is an asymptotically strong Feller Markov semigroup and there exists a point x such that $x \in \operatorname{supp}\mu$ for every invariant probability measure μ of $\mathcal{P}_t$, then there exists at most one invariant probability measure for $\mathcal{P}_t$.*

4 Applications to the Stochastic 2D Navier-Stokes Equations

To state the general ergodic result for the two-dimensional Navier-Stokes equations, we begin by looking at the algebraic structure of the Navier-Stokes nonlinearity written in Fourier space.

Remember that $\mathcal{Z}_0$ as given in (2.3) denotes the set of forced Fourier modes for (2.1). In view of Equation 2.2, it is natural to consider the set $\tilde{\mathcal{Z}}_\infty$, defined as the smallest subset of $\mathbf{Z}^2$ containing $\mathcal{Z}_0$ and satisfying that for every $\ell, j \in \tilde{\mathcal{Z}}_\infty$ such that $\langle \ell^\perp, j\rangle \neq 0$ and $|j| \neq |\ell|$, one has $j + \ell \in \tilde{\mathcal{Z}}_\infty$ (see [EM01]). Denote by $\tilde{\mathcal{H}}$ the closed subspace of $\mathcal{H}$ spanned by the Fourier basis vectors corresponding to elements of $\tilde{\mathcal{Z}}_\infty$. Then, $\tilde{\mathcal{H}}$ is invariant under the flow defined by (2.1).

Since we would like to make use of the existing results, we recall the sequence of subsets $\mathcal{Z}_n$ of $\mathbf{Z}^2$ defined recursively in [MP04] by

$$\mathcal{Z}_n = \left\{ \ell + j \,\middle|\, j \in \mathcal{Z}_0, \ell \in \mathcal{Z}_{n-1} \text{ with } \langle \ell^\perp, j\rangle \neq 0, |j| \neq |\ell| \right\} ,$$

as well as $\mathcal{Z}_\infty = \bigcup_{n=1}^\infty \mathcal{Z}_n$. The two sets $\mathcal{Z}_\infty$ and $\tilde{\mathcal{Z}}_\infty$ are the same even though from the definitions we only see $\mathcal{Z}_\infty \subset \tilde{\mathcal{Z}}_\infty$. The other inclusion follows from the characterization of $\mathcal{Z}_\infty$ given in Proposition 4.4 below.

The following theorem is the principal result of this article.

Theorem 4.1 *The transition semigroup on $\tilde{\mathcal{H}}$ generated by the solutions to (2.1) is asymptotically strong Feller.*

An almost immediate corollary of Theorem 4.1 is

Corollary 4.2 *There exists exactly one invariant probability measure for (2.1) restricted to $\tilde{\mathcal{H}}$.*

Proof of Corollary 4.2. The existence of an invariant probability measure μ for (2.1) is a standard result [Fla94, DPZ96, CK97]. By Corollary 3.17 it suffices to show that the support of every invariant measure contains the element 0. Applying Itô's formula to $\|w\|^2$ yields for every invariant measure μ the *a-priori* bound

$$\int_{\mathcal{H}} \|w\|^2 \,\mu(dw) \leq \frac{C\mathcal{E}_0}{\nu} \;.$$

(See [EMS01] Lemma B.1.) Therefore, denoting by $\mathcal{B}(\rho)$ the ball of radius ρ centered at 0, there exists $\tilde{C}$ such that $\mu(\mathcal{B}(\tilde{C})) > \frac{1}{2}$ for every invariant measure μ. On the other hand, [EM01, Lemma 3.1] shows that, for every $\gamma > 0$ there exists a time T_γ such that

$$\inf_{w \in \mathcal{B}(\tilde{C})} \mathcal{P}_{T_\gamma}(w, \mathcal{B}(\gamma)) > 0 \;.$$

(Note, though [EM01, Lemma 3.1] was about Galerkin approximations, inspection of the proof reveals that it holds equally for the full solution.) Therefore, $\mu(\mathcal{B}(\gamma)) > 0$ for every $\gamma > 0$ and every invariant measure μ, which implies that $0 \in \text{supp}(\mu)$ by Lemma 3.7. □

The crucial ingredient in the proof of Theorem 4.1 is the following result:

Proposition 4.3 *For every $\eta > 0$, there exist constants $C, \delta > 0$ such that for every Fréchet differentiable function φ from $\tilde{\mathcal{H}}$ to* **R** *one has the bound*

$$\|\nabla \mathcal{P}_n \varphi(w)\| \le C \exp(\eta \|w\|^2)(\|\varphi\|_\infty + \|\nabla\varphi\|_\infty e^{-\delta n}) , \tag{4.1}$$

for every $w \in \tilde{\mathcal{H}}$ and $n \in \mathbf{N}$.

The proof of Proposition 4.3 is the content of Section 4.6 below. Theorem 4.1 then follows from this proposition and from Proposition 3.12 with the choices $t_n = n$ and $\delta_n = e^{-\delta n}$. Before we turn to the proof of Proposition 4.3, we characterize $\mathcal{Z}_\infty$ and give an informal introduction to Malliavin calculus adapted to our framework, followed by a brief discussion on how it relates to the strong Feller property.

4.1 The Structure of $\mathcal{Z}_\infty$

In this section, we give a complete characterization of the set $\mathcal{Z}_\infty$. We start by defining $\langle \mathcal{Z}_0 \rangle$ as the subset of $\mathbf{Z}^2 \setminus \{(0,0)\}$ generated by integer linear combinations of elements of $\mathcal{Z}_0$. With this notation, we have

Proposition 4.4 *If there exist $a_1, a_2 \in \mathcal{Z}_0$ such that $|a_1| \neq |a_2|$ and such that a_1 and a_2 are not collinear, then $\mathcal{Z}_\infty = \langle \mathcal{Z}_0 \rangle$. Otherwise, $\mathcal{Z}_\infty = \mathcal{Z}_0$. In either case, one always has that $\mathcal{Z}_\infty = \tilde{\mathcal{Z}}_\infty$.*

This also allows us to characterize the main case of interest:

Corollary 4.5 *One has $\mathcal{Z}_\infty = \mathbf{Z}^2 \setminus \{(0,0)\}$ if and only if the following holds:*

1. *Integer linear combinations of elements of $\mathcal{Z}_0$ generate $\mathbf{Z}^2$.*
2. *There exist at least two elements in $\mathcal{Z}_0$ with non-equal Euclidean norm.*

Proof of Proposition 4.4. It is clear from the definitions that if the elements of $\mathcal{Z}_0$ are all collinear or of the same Euclidean length, one has $\mathcal{Z}_\infty = \mathcal{Z}_0 = \tilde{\mathcal{Z}}_\infty$. In the rest of the proof, we assume that there exist two elements a_1 and a_2 of $\mathcal{Z}_0$ that are neither collinear nor of the same length and we show that one has $\mathcal{Z}_\infty = \langle \mathcal{Z}_0 \rangle$. Since it follows from the definitions that $\mathcal{Z}_\infty \subset \tilde{\mathcal{Z}}_\infty \subset \langle \mathcal{Z}_0 \rangle$, this shows that $\mathcal{Z}_\infty = \tilde{\mathcal{Z}}_\infty$.

Note that the set $\mathcal{Z}_\infty$ consists exactly of those points in $\mathbf{Z}^2$ that can be reached by a walk starting from the origin with steps drawn in $\mathcal{Z}_0$ and which does not contain any of the following “forbidden steps”:

Definition 4.6 A step with increment $\ell \in \mathcal{Z}_0$ starting from $j \in \mathbf{Z}^2$ is *forbidden* if either $|j| = |\ell|$ or j and ℓ are collinear.

Our first aim is to show that there exists $R > 0$ such that $\mathcal{Z}_\infty$ contains every element of $\langle \mathcal{Z}_0 \rangle$ with Euclidean norm larger than R. In order to achieve this, we start with a few very simple observations.

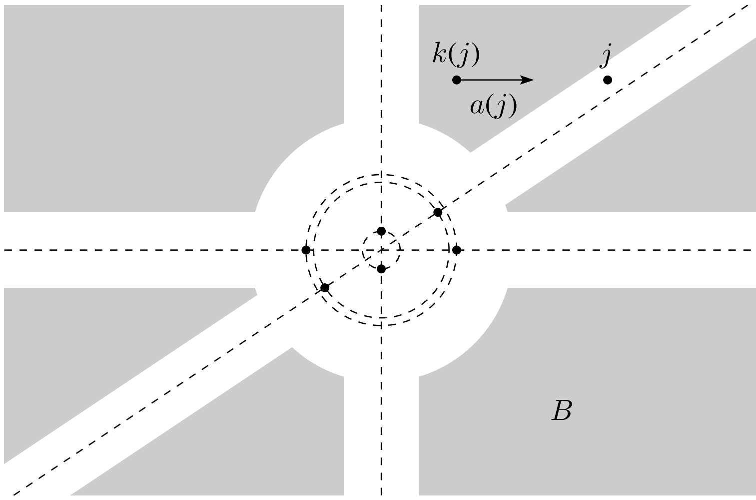

Figure 1: Construction from the proof of Proposition 4.4.

Lemma 4.7 *For every $R_0 > 0$, there exists $R_1 > 0$ such that every $j \in \langle \mathcal{Z}_0 \rangle$ with $|j| \le R_0$ can be reached from the origin by a path with steps in $\mathcal{Z}_0$ (some steps may be forbidden) which never exits the ball of radius R_1.* □

Lemma 4.8 *There exists $L > 0$ such that the set $\mathcal{Z}_\infty$ contains all elements of the form $n_1 a_1 + n_2 a_2$ with n_1 and n_2 in $\mathbf{Z} \setminus [-L, L]$.*

Proof. Assume without loss of generality that $|a_1| > |a_2|$ and that $\langle a_1, a_2 \rangle > 0$. Choose L such that $L\langle a_1, a_2 \rangle \ge |a_1|^2$. By the symmetry of $\mathcal{Z}_0$, we can replace (a_1, a_2) by $(-a_1, -a_2)$, so that we can assume without loss of generality that $n_2 > 0$. We then make first one step in the direction a_1 starting from the origin, followed by n_2 steps in the direction a_2. Note that the assumptions we made on a_1, a_2, and n_2 ensure that none of these steps is forbidden. From there, the condition $n_2 > L$ ensures that we can make as many steps as we want into either the direction a_1 or the direction $-a_1$ without any of them being forbidden. □

Denote by Z the set of elements of the form $n_1 a_1 + n_2 a_2$ considered in Lemma 4.8. It is clear that there exists $R_0 > 0$ such that every element in $\langle \mathcal{Z}_0 \rangle$ is at distance less than R_0 of an element of Z. Given this value R_0, we now fix R_1 as given from Lemma 4.7. Let us define the set

$$A = \mathbf{Z}^2 \cap (\{\alpha j \,|\, \alpha \in \mathbf{R}\,,\ j \in \mathcal{Z}_0\} \cup \{k \,|\, \exists j \in \mathcal{Z}_0 \text{ with } |j| = |k|\})\;,$$

which has the property that there is no forbidden step starting from $\mathbf{Z}^2 \setminus A$. Define furthermore

$$B = \{j \in \langle \mathcal{Z}_0 \rangle \,|\, \inf_{k \in A} |k - j| > R_1\}\;.$$

By Lemma 4.7 and the definition of B, every element of B can be reached by a path from Z containing no forbidden steps, therefore $B \subset \mathcal{Z}_\infty$. On the other hand, it is easy to see that there exists $R > 0$ such that for every element of $j \in \langle \mathcal{Z}_0 \rangle \setminus B$ with $|j| > R$, there exists an element $a(j) \in \mathcal{Z}_0$ and an element $k(j) \in B$ such that j can be reached from $k(j)$ with a finite number of steps in the direction $a(j)$. Furthermore, if R is chosen sufficiently large, none of these steps crosses A, and therefore none of them is forbidden. We have thus shown that there exists $R > 0$ such that $\mathcal{Z}_\infty$ contains $\{j \in \langle \mathcal{Z}_0 \rangle \,|\, |j|^2 \geq R\}$.

In order to help visualizing this construction, Figure 1 shows the typical shapes of the sets A (dashed lines) and B (gray area), as well as a possible choice of $a(j)$ and $k(j)$, given j. (The black dots on the intersections of the circles and the lines making up A depict the elements of $\mathcal{Z}_0$.)

We can (and will from now on) assume that R is an integer. The last step in the proof of Proposition 4.4 is

Lemma 4.9 *Assume that there exists an integer $R > 1$ such that $\mathcal{Z}_\infty$ contains $\{j \in \langle \mathcal{Z}_0 \rangle \,|\, |j|^2 \geq R\}$. Then $\mathcal{Z}_\infty$ also contains $\{j \in \langle \mathcal{Z}_0 \rangle \,|\, |j|^2 \geq R - 1\}$.*

Proof. Assume that the set $\{j \in \langle \mathcal{Z}_0 \rangle \,|\, |j|^2 = R - 1\}$ is non-empty and choose an element j from this set. Since $\mathcal{Z}_0$ contains at least two elements that are not collinear, we can choose $k \in \mathcal{Z}_0$ such that k is not collinear to j. Since $\mathcal{Z}_0$ is closed under the operation $k \mapsto -k$, we can assume that $\langle j, k \rangle \geq 0$. Consequently, one has $|j + k|^2 \geq R$, and so $j + k \in \mathcal{Z}_\infty$ by assumption. The same argument shows that $|j + k|^2 \geq |k|^2 + 1$, so the step $-k$ starting from $j + k$ is not forbidden and therefore $k \in \mathcal{Z}_\infty$. □

This shows that $\mathcal{Z}_\infty = \langle \mathcal{Z}_0 \rangle$ and therefore completes the proof of Proposition 4.4. □

4.2 Malliavin Calculus and the Navier-Stokes Equations

In this section, we give a brief introduction to some elements of Malliavin calculus applied to equation (2.1) to help orient the reader and fix notation. We refer to [MP04] for a longer introduction in the setting of equation (2.1) and to [Nua95, Bel87] for a more general introduction.

Recall from section 2, that $\Phi_t \colon C([0,t]; \mathbf{R}^m) \times \mathcal{H} \to \mathcal{H}$ was the map so that $w_t = \Phi_t(W, w_0)$ for initial condition w_0 and noise realization W. Given a $v \in \mathrm{L}^2_{loc}(\mathbf{R}_+, \mathbf{R}^m)$, the Malliavin derivative of the $\mathcal{H}$-valued random variable w_t in the direction v, denoted $\mathscr{D}^v w_t$, is defined by

$$\mathscr{D}^v w_t = \lim_{\varepsilon \to 0} \frac{\Phi_t(W + \varepsilon V, w_0) - \Phi_t(W, w_0)}{\varepsilon} \;,$$

where the limit holds almost surely with respect to the Wiener measure and where we set $V(t) = \int_0^t v(s)\, ds$. Note that we allow v to be random and possibly non-adapted to the filtration generated by the increments of W.

Defining the symmetrized nonlinearity $\tilde{B}(w, v) = B(\mathcal{K}w, v) + B(\mathcal{K}v, w)$, we use the notation $J_{s,t}$ with $s \leq t$ for the derivative flow between times s and t, *i.e.*

for every $\xi \in \mathcal{H}$, $J_{s,t}\xi$ is the solution of

$$\partial_t J_{s,t}\xi = \nu \Delta J_{s,t}\xi + \tilde{B}(w_t, J_{s,t}\xi) \quad t > s \;, \quad J_{s,s}\xi = \xi \;. \tag{4.2}$$

Note that we have the important cocycle property $J_{s,t} = J_{r,t}J_{s,r}$ for $r \in [s,t]$.

Observe that $\mathscr{D}^v w_t = A_{0,t}v$ where the random operator $A_{s,t} : \mathrm{L}^2([s,t], \mathbf{R}^m) \to \mathcal{H}$ is given by

$$A_{s,t}v = \int_s^t J_{r,t}Qv(r)\, dr \;.$$

To summarize, $J_{0,t}\xi$ is the effect on w_t of an infinitesimal perturbation of the initial condition in the direction ξ and $A_{0,t}v$ is the effect on w_t of an infinitesimal perturbation of the Wiener process in the direction of $V(s) = \int_0^s v(r)\, dr$.

Two fundamental facts we will use from Malliavin calculus are embodied in the following equalities. The first amounts to the chain rule, the second is integration by parts. For a smooth function $\varphi : \mathcal{H} \to \mathbf{R}$ and a (sufficiently regular) process v,

$$\mathbf{E}\langle (\nabla\varphi)(w_t), \mathscr{D}^v w_t\rangle = \mathbf{E}\Big(\mathscr{D}^v(\varphi(w_t))\Big) = \mathbf{E}\Big(\varphi(w_t)\int_0^t \langle v(s), dW_s\rangle\Big) \;. \tag{4.3}$$

The stochastic integral appearing in this expression is an Itô integral if the process v is adapted to the filtration $\mathcal{F}_t$ generated by the increments of W and a Skorokhod integral otherwise.

We also need the adjoint $A^*_{s,t} : \mathcal{H} \to \mathrm{L}^2([s,t], \mathbf{R}^m)$ defined by the duality relation $\langle A^*_{s,t}\xi, v\rangle = \langle \xi, A_{s,t}v\rangle$, where the first scalar product is in $\mathrm{L}^2([s,t], \mathbf{R}^m)$ and the second one is in $\mathcal{H}$. Note that one has $(A^*_{s,t}\xi)(r) = Q^* J^*_{r,t}\xi$, where $J^*_{r,t}$ is the adjoint in $\mathcal{H}$ of $J_{r,t}$.

One of the fundamental objects in the study of hypoelliptic diffusions is the Malliavin matrix $M_{s,t} \stackrel{\text{def}}{=} A_{s,t}A^*_{s,t}$. A glimpse of its importance can be seen from the following. For $\xi \in \mathcal{H}$, one sees that

$$\langle M_{0,t}\xi, \xi\rangle = \sum_{i=1}^m \int_0^t \langle J_{s,t}Qe_i, \xi\rangle^2\, ds \;.$$

Hence the quadratic form $\langle M_{0,t}\xi, \xi\rangle$ is zero for a direction ξ only if no variation whatsoever in the Wiener process at times $s \le t$ could cause a variation in w_t with a non-zero component in the direction ξ.

We also recall that the second derivative $K_{s,t}$ of the flow is the bilinear map solving

$$\begin{aligned}&\partial_t K_{s,t}(\xi,\xi') = \nu\Delta K_{s,t}(\xi,\xi') + \tilde{B}(w_t, K_{s,t}(\xi,\xi')) + \tilde{B}(J_{s,t}\xi', J_{s,t}\xi) \;,\\ &\quad K_{s,s}(\xi,\xi') = 0 \;.\end{aligned}$$

It follows from the variation of constants formula that $K_{s,t}(\xi,\xi')$ is given by

$$K_{s,t}(\xi,\xi') = \int_s^t J_{r,t}\tilde{B}(J_{s,r}\xi', J_{s,r}\xi)\, dr \;. \tag{4.4}$$

4.3 Motivating Discussion

It is instructive to proceed formally pretending that $M_{0,t}$ is invertible as an operator on $\mathcal{H}$. This is probably not true for the problem considered here and we will certainly not attempt to prove it in this article, but the proof presented in Section 4.6 is a modification of the argument in the invertible case and hence it is instructive to start there.

Setting $\xi_t = J_{0,t}\xi$, ξ_t can be interpreted as the perturbation of w_t caused by a perturbation ξ in the initial condition of w_t. Our goal is to find an infinitesimal variation in the Wiener path W over the interval $[0,t]$ which produces the same perturbation at time t as the shift in the initial condition. We want to choose the variation which will change the value of the density the least. In other words, we choose the path with the least action with respect to the metric induced by the inverse of Malliavin matrix. The least squares solution to this variational problem is easily seen to be, at least formally, $v = A_{0,t}^* M_{0,t}^{-1}\xi_t$ where $v \in \mathrm{L}^2([0,t],\mathbf{R}^m)$. Observe that $\mathscr{D}^v w_t = A_{0,t}v = J_{0,t}\xi$. Considering the derivative with respect to the initial condition w of the Markov semigroup $\mathcal{P}_t$ acting on a smooth function φ, we obtain

$$\begin{aligned}\langle \nabla \mathcal{P}_t\varphi(w), \xi\rangle &= \mathbf{E}_w((\nabla\varphi)(w_t)J_{0,t}\xi) = \mathbf{E}_w((\nabla\varphi)(w_t)\mathscr{D}^v w_t) \\ &= \mathbf{E}_w\left(\varphi(w_t)\int_0^t v(s)dW_s\right) \le \|\varphi\|_\infty \mathbf{E}_w\left|\int_0^t v(s)dW_s\right| ,\end{aligned} \tag{4.5}$$

were the penultimate estimate follows from the integration by parts formula (4.3). Since the last term in the chain of implications holds for functions which are simply bounded and measurable, the estimate extends by approximation to that class of φ. Furthermore since the constant $\mathbf{E}_w|\int_0^t v(s)dW_s|$ is independent of φ, if one can show it is finite and bounded independently of $\xi \in \tilde{\mathcal{H}}$ with $\|\xi\| = 1$, we have proved that $\|\nabla\mathcal{P}_t\varphi\|$ is bounded and thus that $\mathcal{P}_t$ is strong Feller in the topology of $\tilde{\mathcal{H}}$. Ergodicity then follows from this statement by means of Corollary 3.17. In particular, the estimate in (4.1) would hold.

In a slightly different language, since v is the infinitesimal shift in the Wiener path equivalent to the infinitesimal variation in the initial condition ξ, one can write down, via the Cameron-Martin theorem, the infinitesimal change in the Radon-Nikodym derivative of the "shifted" measure with respect to the original Wiener measure. This is not trivial since in order to compute the shift v, one uses information on $\{w_s\}_{s\in[0,t]}$, so it is in general not adapted to the Wiener process W_s. This non-adaptedness can be overcome as section 4.8 demonstrates. However the assumption in the above calculation that $M_{0,t}$ is invertible is more serious. We will overcome this by using the ideas and understanding which begin in [Mat98, Mat99, EMS01, KS00, BKL01].

The difficulty in inverting $M_{0,t}$ partly lies in our incomplete understanding of the natural space in which (2.1) lives. The knowledge needed to identify on what domain $M_{0,t}$ can be inverted seems equivalent to identifying the correct reference measure against which to write the transition densities. By "reference measure," we mean a replacement for the role of Lebesgue measure from finite dimensional diffusion theory. This is a very difficult proposition. An alternative was given in the papers [Mat98, Mat99, KS00, EMS01, BKL01, Mat02b, BKL02, Hai02, MY02]. The idea was to use the pathwise contractive properties of the flow at small scales due to the presence of the spatial Laplacian. Roughly speaking, the system has finitely many unstable directions and infinitely many stable directions. One can then use the noise to steer the unstable directions together and let the dynamics cause the stable directions to contract. This requires the small scales to be enslaved to the large scales in some sense. A stochastic version of such a determining modes statement (cf [FP67]) was developed in [Mat98]. Such an approach to prove ergodicity requires looking at the entire future to $+\infty$ (or equivalently the entire past) as the stable dynamics only brings solutions together asymptotically. In the first works in the continuous time setting [EMS01, Mat02b, BKL02], Girsanov's theorem was used to bring the unstable directions together completely, [Hai02] demonstrated the effectiveness of only steering all of the modes together asymptotically. Since all of these techniques used Girsanov's theorem, they required that all of the unstable directions be directly forced. This is a type of partial ellipticity assumption, which we will refer to as "effective ellipticity." The main achievement of this text is to remove this restriction. We also make another innovation which simplifies the argument considerably. We work infinitesimally, employing the linearization of the solution rather than looking at solutions starting from two different starting points.

4.4 Preliminary Calculations and Discussion

Throughout this and the following sections we fix once and for all the initial condition $w_0 \in \tilde{\mathcal{H}}$ for (2.1) and denote by w_t the stochastic process solving (2.1) with initial condition w_0. By $\mathbf{E}$ we mean the expectation starting from this initial condition unless otherwise indicated. Recall also the notation $\mathcal{E}_0 = \operatorname{tr} QQ^* = \sum |q_k|^2$. The following lemma provides us with the auxiliary estimates which will be used to control various terms during the proof of Proposition 4.3.

Lemma 4.10 *The solution of the 2D Navier-Stokes equations in the vorticity formulation (2.1) satisfies the following bounds:*

1. *There exist positive constants C and η_0, depending only on Q and ν, such that*

$$\mathbf{E}\exp\Big(\eta \sup_{t\geq s}\Big(\|w_t\|^2 + \nu\int_s^t \|w_r\|_1^2\,dr - \mathcal{E}_0(t-s)\Big)\Big) \leq C\exp(\eta e^{-\nu s}\|w_0\|^2)\;, \quad (4.6)$$

for every $s \geq 0$ and for every $\eta \leq \eta_0$. Here and in the sequel, we use the notation $\|w\|_1 = \|\nabla w\|$.

2. *There exist constants $\eta_1, a, \gamma > 0$, depending only on $\mathcal{E}_0$ and ν, such that*

$$\mathbf{E}\exp\Big(\eta\sum_{n=0}^{N}\|w_n\|^2-\gamma N\Big)\le\exp(a\eta\|w_0\|^2)\;, \tag{4.7}$$

holds for every $N > 0$, every $\eta \le \eta_1$, and every initial condition $w_0 \in \mathcal{H}$.

3. *For every $\eta > 0$, there exists a constant $C = C(\mathcal{E}_0, \nu, \eta) > 0$ such that the Jacobian $J_{0,t}$ satisfies almost surely*

$$\|J_{0,t}\|\le\exp\Big(\eta\int_0^t\|w_s\|_1^2\,ds+Ct\Big)\;, \tag{4.8}$$

for every $t > 0$.

4. *For every $\eta > 0$ and every $p > 0$, there exists $C = C(\mathcal{E}_0, \nu, \eta, p) > 0$ such that the Hessian satisfies*

$$\mathbf{E}\|K_{s,t}\|^p\le C\exp(\eta\|w_0\|^2)\;,$$

for every $s > 0$ and every $t \in (s, s + 1)$.

The proof of Lemma 4.10 is postponed to Appendix A.

We now show how to modify the discussion in Section 4.3 to make use of the pathwise contractivity on small scales to remove the need for the Malliavin covariance matrix to be invertible on all of $\tilde{\mathcal{H}}$.

The point is that since the Malliavin matrix is not invertible, we are not able to construct a $v \in L^2([0,T],\mathbf{R}^m)$ for a fixed value of T that produces the same infinitesimal shift in the solution as an (arbitrary but fixed) perturbation ξ in the initial condition. Instead, we will construct a $v \in L^2([0,\infty),\mathbf{R}^m)$ such that an infinitesimal shift of the noise in the direction v produces *asymptotically* the same effect as an infinitesimal perturbation in the direction ξ. In other words, one has $\|J_{0,t}\xi - A_{0,t}v_{0,t}\| \to 0$ as $t \to \infty$, where $v_{0,t}$ denotes the restriction of v to the interval $[0,t]$.

Set $\rho_t = J_{0,t}\xi - A_{0,t}v_{0,t}$, the residual error for the infinitesimal variation in the Wiener path W given by v. Then we have from (4.3) the *approximate* integration by parts formula:

$$\begin{aligned}\langle\nabla\mathcal{P}_t\varphi(w),\xi\rangle&=\mathbf{E}_w\Big(\langle\nabla(\varphi(w_t)),\xi\rangle\Big)=\mathbf{E}_w\Big((\nabla\varphi)(w_t)J_{0,t}\xi\Big)\\&=\mathbf{E}_w\Big((\nabla\varphi)(w_t)A_{0,t}v_{0,t}\Big)+\mathbf{E}_w((\nabla\varphi)(w_t)\rho_t)\\&=\mathbf{E}_w\Big(\mathscr{D}^{v_{0,t}}\varphi(w_t)\Big)+\mathbf{E}_w((\nabla\varphi)(w_t)\rho_t)\\&=\mathbf{E}_w\Big(\varphi(w_t)\int_0^t v(s)\,dW(s)\Big)+\mathbf{E}_w((\nabla\varphi)(w_t)\rho_t)\\&\le\|\varphi\|_\infty\mathbf{E}_w\Big|\int_0^t v(s)\,dW(s)\Big|+\|\nabla\varphi\|_\infty\mathbf{E}_w\|\rho_t\|\;.\end{aligned} \tag{4.9}$$

This formula should be compared with (4.5). Again if the process v is not adapted to the filtration generated by the increments of the Wiener process $W(s)$, the integral must be taken to be a Skorokhod integral otherwise Itô integration can be used. Note that the residual error satisfies the equation

$$\partial_t \rho_t = \nu \Delta \rho_t + \tilde{B}(w_t, \rho_t) - Qv(t) \;, \quad \rho_0 = \xi \;, \tag{4.10}$$

which can be interpreted as a control problem, where v is the control and $\|\rho_t\|$ is the quantity that one wants to drive to 0.

If we can find a v so that $\rho_t \to 0$ as $t \to \infty$ and $\mathbf{E}|\int_0^\infty v(s)\,dW(s)| < \infty$ then (4.9) and Proposition 3.12 would imply that w_t is asymptotically strong Feller. A natural way to accomplish this would be to take $v(t) = Q^{-1}\tilde{B}(w_t, \rho_t)$, so that $\partial_t \rho_t = \nu \Delta \rho_t$ and hence $\rho_t \to 0$ as $t \to \infty$. However for this to make sense it would require that $\tilde{B}(w_t, \rho_t)$ takes values in the range of Q. If the number of Brownian motions m is finite this is impossible. Even if $m = \infty$, this is still a delicate requirement which severely limits the range of applicability of the results obtained (see [FM95, Fer97, MS03]).

To overcome these difficulties, one needs to better incorporate the pathwise smoothing which the dynamics possesses at small scales. Though our ultimate goal is to prove Theorem 4.1, which covers (2.1) in a fundamentally hypoelliptic setting, we begin with what might be called the "essentially elliptic" setting. This allows us to outline the ideas in a simpler setting.

4.5 Essentially Elliptic Setting

To help to clarify the techniques used in the sections which follow and to demonstrate their applications, we sketch the proof of the following proposition which captures the main results of the earlier works on ergodicity, translated into the framework of the present paper.

Proposition 4.11 *Let $\mathcal{P}_t$ denote the semigroup generated by the solutions to (2.1) on $\mathcal{H}$. There exists an $N_* = N_*(\mathcal{E}_0, \nu)$ such that if $\mathcal{Z}_0$ contains $\{k \in \mathbf{Z}^2 \;,\; 0 < |k| \le N_*\}$, then for any $\eta > 0$ there exist positive constants c and γ so that*

$$|\nabla \mathcal{P}_t \varphi(w)| \le c \exp\left(\eta \|w\|^2\right)\Big(\|\varphi\|_\infty + e^{-\gamma t}\|\nabla\varphi\|_\infty\Big) \;.$$

This result translates the ideas in [EMS01, Mat02b, Hai02] to our present setting. (See also [Mat03] for more discussion.) The result does differ from the previous analysis in that it proceeds infinitesimally. However, both approaches lead to proving the system has a unique ergodic invariant measure.

The condition on the range of Q can be understood as a type of "effective ellipticity." We will see that the dynamics is contractive for directions orthogonal to the range of Q. Hence if the noise smooths in these directions, the dynamics will smooth in the other directions. What directions are contracting depends fundamentally on a scale set by the balance between $\mathcal{E}_0$ and ν (see [EMS01, Mat03]). Proposition 4.3 holds given a minimal non degeneracy condition independent of the viscosity ν, while Proposition 4.11 requires a non-degeneracy condition which depends on ν.

Proof of Proposition 4.11. Let π_h be the orthogonal projection onto the span of $\{f_k : |k| \geq N\}$ and $\pi_\ell = 1 - \pi_h$. We will fix N presently; however, we will proceed assuming $\mathcal{H}_\ell \stackrel{\text{def}}{=} \pi_\ell \mathcal{H} \subset \text{Range}(Q)$ and that $Q_\ell \stackrel{\text{def}}{=} \pi_\ell Q$ is invertible on $\mathcal{H}_\ell$. By (4.10) we therefore have full control on the evolution of $\pi_\ell \rho_t$ by choosing v appropriately. This allows for an "adapted" approach which does not require the control v to use information about the future increments of the noise process W.

Our approach is to first define a process ζ_t with the property that $\pi_\ell \zeta_t$ is 0 after a finite time and $\pi_h \zeta_t$ evolves according to the linearized evolution, and then choose v such that $\rho_t = \zeta_t$. Since $\pi_\ell \zeta_t = 0$ after some time and the linearized evolution contracts the high modes exponentially, we readily obtain the required bounds on moments of ρ_t. One can in fact pick any dynamics which are convenient for the modes which are directly forced. In the case when all of the modes are forced, the choice $\zeta_t = (1 - t/T) J_{0,t} \xi$ for $t \in [0, T]$ produces the well-known Bismut-Elworthy-Li formula [EL94]. However, this formula cannot be applied in the present setting as all of the modes are not necessarily forced.

For $\xi \in \mathcal{H}$ with $\|\xi\| = 1$, define ζ_t by

$$\partial_t \zeta_t = -\frac{1}{2} \frac{\pi_\ell \zeta_t}{\|\pi_\ell \zeta_t\|} + \nu \Delta \pi_h \zeta_t + \pi_h \tilde{B}(w_t, \zeta_t) \,, \quad \zeta_0 = \xi \,. \tag{4.11}$$

(With the convention that $0/0 = 0$.) Set $\zeta_t^h = \pi_h \zeta_t$ and $\zeta_t^\ell = \pi_\ell \zeta_t$. We define the infinitesimal perturbation v by

$$v(t) = Q_\ell^{-1} F_t \,, \quad F_t = \frac{1}{2} \frac{\zeta_t^\ell}{\|\zeta_t^\ell\|} + \nu \Delta \zeta_t^\ell + \pi_\ell \tilde{B}(w_t, \zeta_t) \,. \tag{4.12}$$

Because $F_t \in \mathcal{H}_\ell$, $Q_\ell^{-1} F_t$ is well defined. It is clear from (4.10) and (4.12) that ρ_t and ζ_t satisfy the same equation, so that indeed $\rho_t = \zeta_t$. Since ζ_t^ℓ satisfies $\partial_t \|\zeta_t^\ell\|^2 = -\|\zeta_t^\ell\|$, one has $\|\zeta_t^\ell\| \leq \|\zeta_0^\ell\| \leq \|\xi\| = 1$. Furthermore, for any initial condition w_0 and any ξ with $\|\xi\| = 1$, one has $\|\zeta_t^\ell\| = 0$ for $t \geq 2$. By calculations similar to those in Appendix A, there exists a constant C so that for any $\eta > 0$

$$\partial_t \|\zeta_t^h\|^2 \leq -\Big(\nu N^2 - \frac{C}{\nu \eta^2} - \eta \|w_t\|_1^2\Big) \|\zeta_t^h\|^2 + \frac{C}{\nu} \|w_t\|_1^2 \|\zeta_t^\ell\|^2 \,.$$

Hence,

$$\begin{aligned}
\|\zeta_t^h\|^2 &\leq \|\zeta_0^h\|^2 \exp\Big(-\Big[\nu N^2 - \frac{C}{\nu \eta^2}\Big] t + \eta \int_0^t \|w_s\|_1^2 ds\Big) \\
&\quad + C \exp\Big(-\Big[\nu N^2 - \frac{C}{\nu \eta^2}\Big] \Big[t - 2\Big] + \eta \int_0^t \|w_r\|_1^2 dr\Big) \int_0^2 \|w_s\|_1^2 ds \,.
\end{aligned}$$

By Lemma 4.10, for any $\eta > 0$ and $p \geq 1$ there exist positive constants C and γ so that for all N sufficiently large

$$\mathbf{E}\|\zeta_t^h\|^p \leq C(1 + \|\zeta_0^h\|^p)e^{\eta\|w_0\|^2}e^{-\gamma t} = 2Ce^{\eta\|w_0\|^2}e^{-\gamma t} \; . \tag{4.13}$$

It remains to get control over the size of the perturbation v. Since v is adapted to the Wiener path,

$$\Big(\mathbf{E}\Big|\int_0^t v(s)\,dW(s)\Big|\Big)^2 \leq \int_0^t \mathbf{E}\|v(s)\|^2\,ds \leq C\int_0^t \mathbf{E}\|F_s\|^2\,ds \; .$$

Now since $\|\pi_\ell \tilde{B}(u,w)\| \leq C\|u\|\|w\|$ (see [EMS01] Lemma A.4), $\|\zeta_t^\ell\| \leq 1$ and $\|\zeta_t^\ell\| = 0$ for $t \geq 2$, we see from (4.12) that there exists a $C = C(N)$ such that for all $s \geq 0$

$$\mathbf{E}\|F_s\|^2 \leq C\Big(1_{\{s\leq 2\}} + \mathbf{E}\|w_s\|^4\mathbf{E}\|\zeta_s\|^4\Big)^{1/2} \; .$$

By using (4.13) with $p = 4$, Lemma A.1 from the appendix to control $\mathbf{E}\|w_s\|^4$, and picking N sufficiently large, we obtain that for any $\eta > 0$ there is a constant C such that

$$\mathbf{E}\Big|\int_0^\infty v(s)\,dW(s)\Big| \leq C\exp(\eta\|w_0\|^2) \; . \tag{4.14}$$

Plugging (4.13) and (4.14) into (4.9), the result follows. □

4.6 Truly Hypoelliptic Setting: Proof of Proposition 4.3

We now turn to the truly hypoelliptic setting. Unlike in the previous section, we allow for unstable directions which are not directly forced by the noise. However, Proposition 4.4 shows that the randomness can reach all of the unstable modes of interest, *i.e.* those in $\tilde{\mathcal{H}}$. In order to show (4.1), we fix from now on $\xi \in \tilde{\mathcal{H}}$ with $\|\xi\| = 1$ and we obtain bounds on $\langle\nabla\mathcal{P}_n\varphi(w),\xi\rangle$ that are independent of ξ.

The basic structure of the argument is the same as in the preceding section on the essentially elliptic setting. We will construct an infinitesimal perturbation of the Wiener path over the time interval $[0,t]$ to approximately match the effect on the solution w_t of an infinitesimal perturbation of the initial condition in an arbitrary direction $\xi \in \tilde{\mathcal{H}}$.

However, since not all of the unstable directions are in the range of Q, we can no longer infinitesimally correct the effect of the perturbation in the low mode space as we did in (4.12). We rather proceed in a way similar to the start of section 4.3. However, since the Malliavin matrix is not invertible, we will regularize it and thus construct a v which compensates for the perturbation ξ only asymptotically as $t \to \infty$. Our construction produces a v which is *not adapted* to the Brownian filtration, which complicates a little bit the calculations analogous to (4.14). A more fundamental difficulty is that the Malliavin matrix is not invertible on any space which is easily identifiable or manageable, certainly not on L^2_0. Hence, the way of constructing v is not immediately obvious.

The main idea for the construction of v is to work with a regularized version $\widetilde{M}_{s,t} \stackrel{\text{def}}{=} M_{s,t} + \beta$ of the Malliavin matrix $M_{s,t}$, for some very small parameter β to be determined later. The resulting $\widetilde{M}^{-1}$ will be an inverse "up to a scale" depending on β. By this we mean that $\widetilde{M}^{-1}$ should not simply be thought of as an approximation of M^{-1}. It is an approximation with a very particular form. Theorem 4.12 which is taken from [MP04] shows that the eigenvectors with small eigenvalues are concentrated in the small scales with high probability. This means that $\widetilde{M}^{-1}$ is very close to M^{-1} on the large scales and very close to the identity times β^{-1} on the small scales. Hence $\widetilde{M}^{-1}$ will be effective in controlling the large scales but, as we will see, something else will have to be done for the small scales.

To be more precise, define for integer values of n the following objects:

$$\hat{J}_n = J_{n,n+\frac{1}{2}}\ ,\ \check{J}_n = J_{n+\frac{1}{2},n+1}\ ,\ A_n = A_{n,n+\frac{1}{2}}\ ,\ M_n = A_n A_n^*\ ,\ \widetilde{M}_n = \beta + M_n\ .$$

We will then work with a perturbation v which is given by 0 on all intervals of the type $[n+\frac{1}{2}, n+1]$, and by $v_n \in \mathrm{L}^2([n, n+\frac{1}{2}], \mathbf{R}^m)$ on the remaining intervals.

We define the infinitesimal variation v_n by

$$v_n = A_n^* \widetilde{M}_n^{-1} \hat{J}_n \rho_n\ , \tag{4.15}$$

where we denote as before by ρ_n the residual of the infinitesimal displacement at time n, due to the perturbation in the initial condition, which has not yet been compensated by v, *i.e.* $\rho_n = J_{0,n}\xi - A_{0,n}v_{0,n}$. From now on, we will make a slight abuse of notation and write v_n for the perturbation of the Wiener path on $[n, n+\frac{1}{2}]$ and its extension (by 0) to the interval $[n, n+1]$.

We claim that it follows from (4.15) that ρ_n is given recursively by

$$\rho_{n+1} = \check{J}_n \beta \widetilde{M}_n^{-1} \hat{J}_n \rho_n\ , \tag{4.16}$$

with $\rho_0 = \xi$. To see the claim observe that (4.16) implies $J_{n,n+1}\rho_n = \check{J}_n \hat{J}_n \rho_n = \check{J}_n A_n v_n + \rho_{n+1}$. Using this and the definitions of the operators involved, it is straightforward to see that indeed

$$\begin{aligned} A_{0,N} v_{0,N} &= \sum_{n=0}^{N-1} J_{(n+1),N} \check{J}_n A_n v_n = \sum_{n=0}^{N-1} (J_{n,N}\rho_n - J_{(n+1),N}\rho_{n+1}) \\ &= J_{0,N}\xi - \rho_N\ . \end{aligned}$$

Thus we see that at time N, the infinitesimal variation in the Wiener path $v_{0,N}$ corresponds to the infinitesimal perturbation in the initial condition ξ up to an error ρ_N.

It therefore remains to show that this choice of v has desirable properties. In particular we need to demonstrate properties similar to (4.13) and (4.14). The analogous statements are given by the next two propositions whose proofs will be the content of sections 4.7 and 4.8. Both of these propositions rely heavily on the following theorem obtained in [MP04, Theorem 6.2].

Theorem 4.12 *Denote by M the Malliavin matrix over the time interval $[0, \frac{1}{2}]$ and define $\tilde{\mathcal{H}}$ as above. For every α, η, p and every orthogonal projection π_ℓ on a finite number of Fourier modes, there exists $\tilde{C}$ such that*

$$\mathbf{P}(\langle M\varphi, \varphi\rangle < \varepsilon\|\varphi\|_1^2) \leq \tilde{C}\varepsilon^p \exp(\eta\|w_0\|^2)\;, \tag{4.17}$$

holds for every (random) vector $\varphi \in \tilde{\mathcal{H}}$ satisfying $\|\pi_\ell\varphi\| \geq \alpha\|\varphi\|_1$ almost surely, for every $\varepsilon \in (0,1)$, and for every $w_0 \in \tilde{\mathcal{H}}$.

The next proposition shows that we can construct a v which has the desired effect of driving the error ρ_t to zero as $t \to \infty$.

Proposition 4.13 *For any $\eta > 0$, there exist constants $\beta > 0$ and $C > 0$ such that*

$$\mathbf{E}\|\rho_N\|^{10} \leq \frac{C\exp(\eta\|w_0\|^2)}{2^N}\;, \tag{4.18}$$

holds for every $N > 0$. (Note that by increasing β further, the 2^N in the denominator could be replaced by K^N for an arbitrary $K \geq 2$ without altering the value of C.)

However for the above result to be useful, the "cost" of shifting the noise by v (*i.e.* the norm of v in the Cameron-Martin space) must be finite. Since the time horizon is infinite, this is not a trivial requirement. In the "essentially elliptic" setting, it was demonstrated in (4.14). In the "truly hypoelliptic" setting, we obtain

Proposition 4.14 *For any $\eta > 0$, there exists a constant C so that*

$$\mathbf{E}\Big|\int_0^N v_{0,s}\,dW(s)\Big|^2 \leq \frac{C}{\beta^2}e^{\eta\|w_0\|^2}\sum_{n=0}^{\infty}(\mathbf{E}\|\rho_n\|^{10})^{\frac{1}{5}} \tag{4.19}$$

(Note that the power 10 in this expression is arbitrary and can be brought as close to 2 as one wishes.)

Plugging these estimates into (4.9), we obtain Proposition 4.3. Note that even though Proposition 4.3 is sufficient for the present article, small modifications of (4.9) produce the following stronger bound.

Proposition 4.15 *For every $\eta > 0$ and every $\gamma > 0$, there exist constants $C_{\eta,\gamma}$ such that for every Fréchet differentiable function φ from $\tilde{\mathcal{H}}$ to $\mathbf{R}$ one has the bound*

$$\|\nabla\mathcal{P}_n\varphi(w)\| \leq \exp(\eta\|w\|^2)\Big(C_{\eta,\gamma}\sqrt{(\mathcal{P}_n|\varphi|^2)(w)} + \gamma^n\sqrt{(\mathcal{P}_n\|\nabla\varphi\|^2)(w)}\Big)\;,$$

for every $w \in \tilde{\mathcal{H}}$ and $n \in \mathbf{N}$.

Proof. Applying Cauchy-Schwarz to the terms on the right-hand side of the penultimate line of (4.9) one obtains

$$|\langle \nabla \mathcal{P}_n \varphi, \xi\rangle| \le \Big(\mathbf{E}\Big|\int_0^n v_{0,s}\, dW(s)\Big|^2 \mathcal{P}_n |\varphi|^2\Big)^{1/2} + \Big(\mathbf{E}\|\rho_n\|^{10} \mathcal{P}_n \|\nabla\varphi\|^2\Big)^{1/2} .$$

It now suffices to use the bounds from the above propositions and to note that the right-hand side is independent of the choice of ξ provided $\|\xi\| = 1$. □

4.7 Controlling the Error: Proof of Proposition 4.13

Before proving Proposition 4.13, we state the following lemma, which summarizes the effect of our control on the perturbation and shall be proved at the end of this section.

Lemma 4.16 *For every two constants $\gamma, \eta > 0$ and every $p \ge 1$, there exists a constant $\beta_0 > 0$ such that*

$$\mathbf{E}(\|\rho_{n+1}\|^p \,|\, \mathcal{F}_n) \le \gamma e^{\eta \|w_n\|^2} \|\rho_n\|^p$$

holds almost surely whenever $\beta \le \beta_0$.

Proof of Proposition 4.13. Define

$$C_n = \frac{\|\rho_{n+1}\|^{10}}{\|\rho_n\|^{10}} ,$$

with the convention that $C_n = 0$ if $\rho_n = 0$. Note that since $\|\rho_0\| = 1$, one has $\|\rho_N\|^{10} = \prod_{n=0}^{N-1} C_n$. We begin by establishing some properties of C_n and then use them to prove the proposition.

Note that $\|\beta \widetilde{M}_n^{-1}\| \le 1$ and so, by (4.8) and (4.16), for every $\eta > 0$ there exists a constant $C_\eta > 0$ such that

$$C_n \le \|\check{J}_n \beta \widetilde{M}_n^{-1} \hat{J}_n\|^{10} \le \|\check{J}_n\|^{10} \|\hat{J}_n\|^{10} \le \exp\Big(\eta \int_n^{n+1} \|w_s\|_1^2\, ds + C_\eta\Big) , \tag{4.20}$$

almost surely. Note that this bound is independent of β. Next, for given values of η and $R > 0$, we define

$$C_{n,R} = \begin{cases} e^{-\eta R} & \text{if } \|w_n\|^2 \ge 2R, \\ e^{\eta R} C_n & \text{otherwise.} \end{cases}$$

Obviously both C_n and $C_{n,R}$ are $\mathcal{F}_{n+1}$-measurable. Lemma 4.16 shows that for every $R > \eta^{-1}$, one can find a $\beta > 0$ such that

$$\mathbf{E}(C_{n,R}^2 \,|\, \mathcal{F}_n) \le \frac{1}{2} , \quad \text{almost surely} . \tag{4.21}$$

Note now that (4.20) and the definition of $C_{n,R}$ immediately imply that

$$C_n \le C_{n,R} \exp\Big(\eta \int_n^{n+1} \|w_s\|_1^2\, ds + \eta \|w_n\|^2 + C_\eta - \eta R\Big)\;, \tag{4.22}$$

almost surely. This in turn implies that

$$\begin{aligned}\prod_{n=0}^{N-1} C_n &\le \prod_{n=0}^{N-1} C_{n,R}^2 + \prod_{n=0}^{N-1} \exp\Big(2\eta \int_n^{n+1} \|w_s\|_1^2\, ds + 2\eta \|w_n\|^2 + 2C_\eta - 2\eta R\Big) \\ &\le \prod_{n=0}^{N-1} C_{n,R}^2 + \exp\Big(4\eta \sum_{n=0}^{N-1} \|w_n\|^2 + 2N(C_\eta - \eta R)\Big) \\ &\quad + \exp\Big(4\eta \int_0^N \|w_s\|_1^2\, ds + 2N(C_\eta - \eta R)\Big)\;.\end{aligned}$$

Now fix $\eta > 0$ (not too large). In light of (4.6) and (4.7), we can then choose R sufficiently large so that the two last terms satisfy the required bounds. Then, we choose β sufficiently small so that (4.21) holds and the estimate follows. □

To prove Lemma 4.16, we will use the following two lemmas. The first is simply a consequence of the dissipative nature of the equation. Because of the Laplacian, the small scale perturbations are strongly damped.

Lemma 4.17 *For every $p \ge 1$, every $T > 0$, and every two constants $\gamma, \eta > 0$, there exists an orthogonal projector π_ℓ onto a finite number of Fourier modes such that*

$$\mathbf{E}\|(1-\pi_\ell) J_{0,T}\|^p \le \gamma \exp(\eta \|w_0\|^2)\;, \tag{4.23}$$

$$\mathbf{E}\|J_{0,T}(1-\pi_\ell)\|^p \le \gamma \exp(\eta \|w_0\|^2)\;. \tag{4.24}$$

The proof of the above lemma is postponed to the appendix. The second lemma is central to the hypoelliptic results in this paper. It is the analog of (4.14) from the essentially elliptic setting and provides the key to controlling the “low modes” when they are not directly forced and Girsanov’s theorem cannot be used directly. This result makes use of the results in [MP04] which contains the heart of the analysis of the structure of the Malliavin matrix for equation (2.1) in the hypoelliptic setting.

Lemma 4.18 *Fix $\xi \in \tilde{\mathcal{H}}$ and define*

$$\zeta = \beta(\beta + M_0)^{-1} \hat{J}_0 \xi\;.$$

Then, for every two constants $\gamma, \eta > 0$ and every low-mode orthogonal projector π_ℓ, there exists a constant $\beta > 0$ such that

$$\mathbf{E}\|\pi_\ell \zeta\|^p \le \gamma e^{\eta \|w_0\|^2} \|\xi\|^p\;.$$

Remark 4.19 Since one has obviously that $\|\zeta\| \le \|\hat J_0\xi\|$, this lemma tells us that applying the operator $\beta(\beta + M_0)^{-1}$ (with a very small value of β) to a vector in $\tilde{\mathcal{H}}$ either reduces its norm drastically or transfers most of its "mass" into the high modes (where the cutoff between "high" and "low" modes is arbitrary but influences the possible choices of β). This explains why the control v is set to 0 for half of the time in Section 4.6: In order to ensure that the norm of ρ_n gets really reduced after one step, we choose the control in such a way that $\beta(\beta+M_n)^{-1}\hat J_n$ is composed by $\check J_n$, using the fact embodied in Lemma 4.17 that the Jacobian will contract the high modes before the low modes start to grow out of control.

Proof of Lemma 4.18. For $\alpha > 0$, let A_α denote the event $\|\pi_\ell\zeta\| > \alpha\|\zeta\|_1$. We also define the random vectors

$$\zeta_\alpha(\omega) = \zeta(\omega)\chi_{A_\alpha}(\omega)\;, \quad \bar\zeta_\alpha(\omega) = \zeta(\omega) - \zeta_\alpha(\omega)\;, \quad \omega \in \Omega\;,$$

where ω is the chance variable and χ_A is the characteristic function of a set A. It is clear that

$$\mathbf{E}\|\pi_\ell\bar\zeta_\alpha\|^p \le \alpha^p\,\mathbf{E}\|\zeta\|_1^p\;.$$

Using the bounds (4.8) and (4.6) on the Jacobian and the fact that M_0 is a bounded operator from $\mathcal{H}_1$ (the Sobolev space of functions with square integrable derivatives) into $\mathcal{H}_1$, we get

$$\mathbf{E}\|\pi_\ell\bar\zeta_\alpha\|^p \le \alpha^p\mathbf{E}\|\zeta\|_1^p \le \alpha^p\mathbf{E}\|\hat J_0\xi\|_1^p \le \frac{\gamma}{2}e^{\eta\|w_0\|^2}\|\xi\|^p\;, \tag{4.25}$$

(with η and γ as in the statement of the proposition) for sufficiently small α. From now on, we fix α such that (4.25) holds. One has the chain of inequalities

$$\begin{aligned}\langle\zeta_\alpha, M_0\zeta_\alpha\rangle &\le \langle\zeta, M_0\zeta\rangle \le \langle\zeta,(M_0+\beta)\zeta\rangle\\ &= \beta\langle\hat J_0\xi, \beta(M_0+\beta)^{-1}\hat J_0\xi\rangle \le \beta\|\hat J_0\xi\|^2\;.\end{aligned} \tag{4.26}$$

From Theorem 4.12, we furthermore see that, for every $p_0 > 0$, there exists a constant $\tilde C$ such that

$$\mathbf{P}(\langle M_0\zeta_\alpha,\zeta_\alpha\rangle < \varepsilon\|\zeta_\alpha\|_1^2) \le \tilde C\varepsilon^{p_0}\exp(\eta\|w_0\|^2)\;,$$

holds for every $w_0 \in \tilde{\mathcal{H}}$ and every $\varepsilon \in (0,1)$. Consequently, we have

$$\mathbf{P}\Big(\frac{\|\zeta_\alpha\|_1^2}{\|\hat J_0\xi\|^2} > \frac{1}{\varepsilon}\Big) \le \mathbf{P}(\langle M_0\zeta_\alpha,\zeta_\alpha\rangle < \varepsilon\beta\|\zeta_\alpha\|_1^2) \le \tilde C\beta^{p_0}\varepsilon^{p_0}\exp(\eta\|w_0\|^2)\;,$$

where we made use of (4.26) to get the first inequality. This implies that, for every $p, q \ge 1$, there exists a constant $\tilde C$ such that

$$\mathbf{E}\Big(\frac{\|\zeta_\alpha\|_1^p}{\|\hat J_0\xi\|^p}\Big) \le \tilde C\beta^q\exp(\eta\|w_0\|^2)\;. \tag{4.27}$$

Since $\|\pi_\ell \zeta_\alpha\| \le \|\zeta_\alpha\|_1$ and

$$\mathbf{E}\|\zeta_\alpha\|_1^p \le \sqrt{\mathbf{E}\Big(\frac{\|\zeta_\alpha\|_1^{2p}}{\|\hat J_0 \xi\|^{2p}}\Big)\mathbf{E}\|\hat J_0 \xi\|^{2p}}\;,$$

it follows from (4.27) and the bound (4.8) on the Jacobian that, by choosing β sufficiently small, one gets

$$\mathbf{E}\|\pi_\ell \zeta_\alpha\|^p \le \frac{\gamma}{2} e^{\eta\|w_0\|^2}\|\xi\|^p\;. \tag{4.28}$$

Note that $\mathbf{E}\|\pi_\ell \zeta\|^p = \mathbf{E}\|\pi_\ell \zeta_\alpha\|^p + \mathbf{E}\|\pi_\ell \bar\zeta_\alpha\|^p$ since only one of the previous two terms is nonzero for any given realization ω. The claim thus follows from (4.25) and (4.28). □

Using Lemma 4.17 and Lemma 4.18, we now give the

Proof of Lemma 4.16. Define $\zeta_n = \beta \widetilde{M}_n^{-1}\hat J_n \rho_n$, so that $\rho_{n+1} = \check J_n \zeta_n$. It follows from the definition of $\widetilde{M}_n$ and the bounds (4.8) and (4.6) on the Jacobian that there exists a constant C such that

$$\mathbf{E}(\|\zeta_n\|^p \,|\, \mathcal{F}_n) \le C e^{\frac{\eta}{2}\|w_n\|^2}\|\rho_n\|^p\;,$$

uniformly in $\beta > 0$. Applying (4.24) to this bound yields the existence of a projector π_ℓ on a finite number of Fourier modes such that

$$\mathbf{E}(\|\check J_n (1-\pi_\ell)\zeta_n\|^p \,|\, \mathcal{F}_n) \le \gamma e^{\eta\|w_n\|^2}\|\rho_n\|^p\;.$$

Furthermore, Lemma 4.18 shows that, for an arbitrarily small value $\tilde\gamma$, one can choose β sufficiently small so that

$$\mathbf{E}(\|\pi_\ell \zeta_n\|^p \,|\, \mathcal{F}_n) \le \tilde\gamma e^{\frac{\eta}{2}\|w_n\|^2}\|\rho_n\|^p\;.$$

Applying again the *a priori* estimates (4.8) and (4.6) on the Jacobian, we see that one can choose $\tilde\gamma$ (and thus β) sufficiently small so that

$$\mathbf{E}(\|\check J_n \pi_\ell \zeta_n\|^p \,|\, \mathcal{F}_n) \le \gamma e^{\eta\|w_n\|^2}\|\rho_n\|^p\;,$$

and the result follows. □

4.8 Cost of the Control : Proof of Proposition 4.14

Since the process $v_{0,s}$ is not adapted to the Wiener process $W(s)$, the integral must be taken to be a Skorokhod integral. We denote by $\mathscr{D}_s F$ the Malliavin derivative of a random variable F at time s (see [Nua95] for definitions). Suppressing the dependence on the initial condition w, we obtain from the definition of the Skorokhod integral and from the corresponding Itô isometry (see *e.g.* [Nua95, p. 39])

$$\mathbf{E}\Big|\int_0^N v(s)\,dW(s)\Big|^2 \le \mathbf{E}\|v_{0,N}\|^2 + \sum_{n=0}^N \int_n^{n+\frac12}\int_n^{n+\frac12} \mathbf{E}|\!|\!|\mathscr{D}_s v_n(t)|\!|\!|^2\,ds\,dt\;.$$

(Remember that $v_n(t) = 0$ on $[n+\frac{1}{2}, n+1]$.) In this expression, the norm $\|\!|\cdot|\!\|$ denotes the Hilbert-Schmidt norm on $m \times m$ matrices, so one has

$$\int_n^{n+\frac{1}{2}} \int_n^{n+\frac{1}{2}} \mathbf{E}\|\!|\mathscr{D}_s v_n(t)|\!\|^2 \, ds\, dt = \sum_{i=1}^m \int_n^{n+\frac{1}{2}} \mathbf{E}\|\mathscr{D}_s^i v_n\|^2 \, ds \;,$$

where the norm $\|\cdot\|$ is in $\mathrm{L}^2([n, n+\frac{1}{2}], \mathbf{R}^m)$ and $\mathscr{D}_s^i$ denotes the Malliavin derivative with respect to the ith component of the noise at time s.

In order to obtain an explicit expression for $\mathscr{D}_s^i v_n$, we start by computing separately the Malliavin derivatives of the various expressions that enter into its construction. Recall from [Nua95] that $\mathscr{D}_s^i w_t = J_{s,t} Q e_i$ for $s < t$. It follows from this and the expression (4.2) for the Jacobian that the Malliavin derivative of $J_{s,t}\xi$ is given by

$$\partial_t \mathscr{D}_r^i J_{s,t}\xi = \nu\Delta \mathscr{D}_r^i J_{s,t}\xi + \tilde{B}(w_t, \mathscr{D}_r^i J_{s,t}\xi) + \tilde{B}(J_{r,t}Qe_i, J_{s,t}\xi) \;.$$

From the variation of constants formula and the expression (4.4) for the process K, we get

$$\mathscr{D}_r^i J_{s,t}\xi = \begin{cases} K_{r,t}(Qe_i, J_{s,r}\xi) & \text{if } r \geq s, \\ K_{s,t}(J_{r,s}Qe_i, \xi) & \text{if } r \leq s. \end{cases} \tag{4.29}$$

In the remainder of this section, we will use the convention that if $A : \mathcal{H}_1 \to \mathcal{H}_2$ is a random linear map between two Hilbert spaces, we denote by $\mathscr{D}_s^i A : \mathcal{H}_1 \to \mathcal{H}_2$ the random linear map defined by

$$(\mathscr{D}_s^i A)h = \langle \mathscr{D}_s(Ah), e_i \rangle \;.$$

With this convention, (4.29) yields immediately

$$\mathscr{D}_r^i \hat{J}_n w = K_{r,n+\frac{1}{2}}(J_{n,r}w, Qe_i) \quad \text{for } r \in [n, n+\tfrac{1}{2}]. \tag{4.30}$$

Similarly, we see from (4.29) and the definition of A_n that the map $\mathscr{D}_r^i A_n$ given by

$$\begin{aligned} \mathscr{D}_r^i A_n h = &\int_n^r K_{r,n+\frac{1}{2}}(J_{s,r}Qh(s), Qe_i))\, ds \\ &+ \int_r^{n+\frac{1}{2}} K_{r,n+\frac{1}{2}}(Qh(s), J_{r,s}Qe_i))\, ds \;. \end{aligned} \tag{4.31}$$

We denote its adjoint by $\mathscr{D}_r^i A_n^*$. Since $\widetilde{M}_n = \beta + A_n A_n^*$, we get from the chain rule

$$\mathscr{D}_s^i \widetilde{M}_n^{-1} = -\widetilde{M}_n^{-1}\Big((\mathscr{D}_s^i A_n)A_n^* + A_n(\mathscr{D}_s^i A_n^*)\Big)\widetilde{M}_n^{-1} \;.$$

Since ρ_n is $\mathcal{F}_n$-measurable, one has $\mathscr{D}_r^i \rho_n = 0$ for $r \geq n$. Therefore, combining the above expressions with the Leibniz rule applied to the definition (4.15) of v_n yields

$$\mathscr{D}_s^i v_n = (\mathscr{D}_s^i A_n^*)\widetilde{M}_n^{-1}\hat{J}_n \rho_n + A_n^* \widetilde{M}_n^{-1}(\mathscr{D}_s^i \hat{J}_n)\rho_n$$

$$- A_n^* \widetilde{M}_n^{-1} \Big((\mathscr{D}_s^i A_n) A_n^* + A_n (\mathscr{D}_s^i A_n^*) \Big) \widetilde{M}_n^{-1} \hat{J}_n \rho_n \;.$$

Since $\widetilde{M}_n = \beta + A_n A_n^*$, one has the almost sure bounds

$$\|A_n^* \widetilde{M}_n^{-1/2}\| \le 1 \;, \quad \|\widetilde{M}_n^{-1/2} A_n\| \le 1 \;, \quad \|\widetilde{M}_n^{-1/2}\| \le \beta^{-1/2} \;.$$

This immediately yields

$$\|\mathscr{D}_s^i v_n\| \le 3\beta^{-1} \|\mathscr{D}_s^i A_n\| \|\hat{J}_n\| \|\rho_n\| + \beta^{-1/2} \|\mathscr{D}_s^i \hat{J}_n\| \|\rho_n\| \;.$$

Combining this with (4.31), (4.30), and Lemma 4.10, we obtain, for every $\eta > 0$, the existence of a constant C such that

$$\mathbf{E}\|\mathscr{D}_s^i v_n\|^2 \le C e^{\eta \|w\|^2} \beta^{-2} (\mathbf{E}\|\rho_n\|^{10})^{\frac{1}{5}} \;.$$

Applying Lemma 4.10 to the definition of v_n we easily get a similar bound for $\mathbf{E}\|v_n\|^2$, which then implies the quoted result.

5 Discussion and Conclusion

Even though the results obtained in this work are relatively complete, they still leave a few questions open.

Do the transition probabilities for (2.1) converge towards the invariant measure and at which rate? In other words, do the solutions to (2.1) have the mixing property? We expect this to be the case and plan to answer this question in a subsequent publication.

What happens if $\mathcal{H} \neq \tilde{\mathcal{H}}$ and one starts the system with an initial condition $w_0 \in \mathcal{H} \setminus \tilde{\mathcal{H}}$? If the viscosity is sufficiently large, we know that the component of w_t orthogonal to $\tilde{\mathcal{H}}$ will decrease exponentially with time. This is however not expected to be the case when ν is small. In this case, we expect to have (at least) one invariant measure associated to every (closed) subspace V invariant under the flow.

Appendix A A Priori Estimates for the Navier-Stokes Equations

Note: The letter C denotes generic constants whose value can change from one line to the next even within the same equation. The possible dependence of C on the parameters of (2.1) should be clear from the context.

We define for $\alpha \in \mathbf{R}$ and for w a smooth function on $[0, 2\pi]^2$ with mean 0 the norm $\|w\|_\alpha$ by

$$\|w\|_\alpha^2 = \sum_{k \in \mathbf{Z}^2 \setminus \{0,0\}} |k|^{2\alpha} w_k^2 \;,$$

where of course w_k denotes the Fourier mode with wavenumber k. Define furthermore $(\mathcal{K}w)_k = -i w_k k^\perp / \|k\|^2$, $B(u, v) = (u \cdot \nabla) v$ and $\mathcal{S} = \{s = (s_1, s_2, s_3) \in$

$\mathbf{R}^3_+ : \sum s_i \geq 1, s \neq (1,0,0), (0,1,0), (0,0,1)\}$. Then the following relations are useful (cf. [CF88]):

$$\langle B(u,v), w\rangle = -\langle B(u,w), v\rangle \qquad \text{if } \nabla\cdot u = 0 \tag{A.1}$$

$$|\langle B(u,v), w\rangle| \leq C\|u\|_{s_1}\|v\|_{1+s_2}\|w\|_{s_3} \qquad (s_1,s_2,s_3)\in\mathcal{S} \tag{A.2}$$

$$\|\mathcal{K}u\|_\alpha = \|u\|_{\alpha-1} \tag{A.3}$$

$$\|u\|^2_\beta \leq \varepsilon\|u\|^2_\alpha + \varepsilon^{-2\frac{\gamma-\beta}{\beta-\alpha}}\|u\|^2_\gamma \qquad \text{if } 0\leq\alpha<\beta<\gamma \text{ and } \varepsilon>0. \tag{A.4}$$

Before we turn to the proof of Lemma 4.10, we give the following essential bound on the solutions of (2.1).

Lemma A.1 *There exist constants $\eta_0 > 0$ and $C > 0$, such that for every $t > 0$ and every $\eta \in (0, \eta_0]$, the bound*

$$\mathbf{E}\exp(\eta\|w_t\|^2) \leq C\exp(\eta e^{-\nu t}\|w_0\|^2) \tag{A.5}$$

holds.

Proof. From (A.1) and Itôs formula, we obtain

$$\|w_t\|^2 - \|w_0\|^2 + 2\nu\int_0^t \|w_r\|_1^2\,dr = \int_0^t \langle w_r, Q\,dW(r)\rangle + \mathcal{E}_0 t\;,$$

where we set $\mathcal{E}_0 = \operatorname{tr} QQ^*$. Using the fact that $\|w_r\|_1^2 \geq \|w_r\|^2$, we get

$$\|w_t\|^2 \leq e^{-\nu t}\|w_0\|^2 + \frac{\mathcal{E}_0}{\nu} + \int_0^t e^{-\nu(t-r)}\langle w_r, Q\,dW(r)\rangle - \nu\int_0^t e^{-\nu(t-r)}\|w_r\|^2\,dr\;.$$

There exists a constant $\alpha > 0$ such that $\nu\|w_r\|^2 > \frac{\alpha}{2}\|Q^* w_r\|^2$, so that [Mat02a, Lemma A.1] implies

$$\mathbf{P}\Big(\|w_t\|^2 - e^{-\nu t}\|w_0\|^2 - \frac{\mathcal{E}_0}{\nu} > \frac{K}{\alpha}\Big) \leq e^{-K}\;.$$

Note now that if a random variable X satisfies $\mathbf{P}(X \geq C) \leq 1/C^2$ for all $C \geq 0$, then $\mathbf{E}X \leq 2$. The bound (A.5) thus follows immediately with for example $\eta_0 = \alpha/2$ and $C = 2\exp(\frac{\alpha\mathcal{E}_0}{2\nu})$. □

We now turn to the proof of Lemma 4.10.

Proof of Lemma 4.10. **Point 1.** From (A.1) and Itô's formula, for any $\eta > 0$ we obtain

$$\begin{aligned}\eta\|w_t\|^2 + \eta\nu\int_s^t \|w_r\|_1^2\,dr - \eta\mathcal{E}_0(t-s)\\ = \eta\|w_s\|^2 + \eta\int_s^t\langle w_r, Q\,dW(r)\rangle - \eta\nu\int_s^t\|w_r\|_1^2\,dr\end{aligned}$$

where we set $\mathcal{E}_0 = \operatorname{tr} QQ^*$. Denote by $M(s,t)$ the first two terms on the right-hand side of the last expression and set $N(s,t) = M(s,t) - \eta\nu \int_s^t \|w_r\|_1^2\,dr$. Now observe that with α as in the proof of Lemma A.1 above, one has $N(s,t) \le M(s,t) - \frac{\alpha}{2\eta}\langle M\rangle(s,t)$ where $\langle M\rangle(s,t)$ is the quadratic variation of the continuous L^2-martingale M. Hence by the standard exponential martingale estimate $\mathbf{P}(\sup_{t\ge s} N(s,t) \ge K \,|\, \mathcal{F}_s) \le \exp(\eta\|w_s\|^2 - \frac{\alpha K}{\eta})$ for all $s \ge 0$. Here we use the notation $\mathcal{F}_s$ to denote the filtration generated by the noise up to the time s. Thus, for all $\eta \in (0, \alpha/2]$ and $s \ge 0$,

$$\mathbf{E}\exp\Big(\eta \sup_{t\ge s}\Big(\|w_t\|^2 + \nu\int_s^t \|w_r\|_1^2\,dr - \mathcal{E}_0(t-s)\Big)\,\Big|\,\mathcal{F}_s\Big) \le 2\exp(\eta\|w_s\|^2)\;.$$

Choosing η_0 as above and using Lemma A.1 to bound the expected value of the right-hand side completes the proof.

Point 2. Taking conditional expectations with respect to $\mathcal{F}_{N-1}$ on the left hand side of (4.7) and applying Lemma A.1, one has

$$\mathbf{E}\exp\Big(\eta\sum_{n=0}^{N}\|w_n\|^2\Big) \le C\mathbf{E}\exp\Big(\eta e^{-\nu}\|w_{N-1}\|^2 + \eta\sum_{n=0}^{N-1}\|w_n\|^2\Big)\;.$$

Applying this procedure repeatedly, one obtains

$$\mathbf{E}\exp\Big(\eta\sum_{n=0}^{N}\|w_n\|^2\Big) \le C^N \exp(a\eta\|w_0\|^2)\;,$$

where $a = \sum_{n=0}^{\infty} e^{-\nu n}$. This computation is valid, provided $a\eta$ is smaller than η_0, so the result follows by taking $\eta_1 = \eta_0/a$.

Point 3. We define $\xi_t = J_{0,t}\xi_0$ for some $\xi_0 \in \mathcal{H}$. The evolution of ξ_t is then given by (4.2). We thus have for the $\mathcal{H}$-norm of ξ the equation

$$\partial_t\|\xi_t\|^2 = -2\nu\|\nabla\xi_t\|^2 + 2\langle B(\mathcal{K}\xi_t, w_t), \xi_t\rangle\;.$$

Equation (A.1) yields the existence of a constant C such that $2|\langle B(\mathcal{K}h, w), \zeta\rangle| \le C\|w\|_1\|h\|\|\zeta\|_{1/2}$ for example. By interpolation, we get

$$2|\langle B(\mathcal{K}h, w), \zeta\rangle| \le \nu\|\zeta\|_1^2 + \frac{C}{\eta^2\nu}\|\zeta\|^2 + \frac{\eta}{2}\|w\|_1^2\|h\|^2\;, \tag{A.6}$$

and therefore

$$\partial_t\|\xi_t\|^2 \le -\nu\|\nabla\xi_t\|^2 + \frac{C}{\eta^2\nu}\|\xi_t\|^2 + \frac{\eta}{2}\|w_t\|_1^2\|\xi_t\|^2\;, \tag{A.7}$$

for every $\eta > 0$. This yields (4.8).

Point 4. This bound follows in a rather straightforward way from (A.8) which is in the next proof. Standard Sobolev estimates and interpolation inequalities give for the symmetrized nonlinearity $\tilde{B}$ the bound

$$\|\tilde{B}(u,w)\| \le C(\|u\|_{1/2}\|w\|_1 + \|u\|_1\|w\|_{1/2})$$

$$\leq C(\|u\|^{1/2}\|u\|_1^{1/2}\|w\|_1 + \|w\|^{1/2}\|w\|_1^{1/2}\|u\|_1) \;.$$

Combining this with the definition (4.4) of $K_{s,t}$ and bound (A.8) yields for $s,t \in [0,1]$

$$\begin{aligned}\|K_{s,t}\| &\leq C\int_s^t \|J_{r,t}\|\|J_{s,r}\|_1^{3/2}\|J_{s,r}\|^{1/2}\,dr \\ &\leq C\exp\Big(\eta\int_0^1 \|w_r\|_1^2\,dr\Big)\;,\end{aligned}$$

where we used the integrability of $|r-s|^{-3/4}$ in the second step. This concludes the proof of Lemma 4.10. □

Proof of Lemma 4.17. In order to get (4.23), we show that with the above notations, one can get bounds on $\|\xi_t\|_1$ as well. To achieve this we define, for a constant $\varepsilon > 0$ to be fixed later, $\zeta_t = \|\xi_t\|^2 + t\varepsilon\|\xi_t\|_1^2$. Using (A.7) to bound the derivative of the first term and combining (A.2) with (A.3) for the other terms, we then get in a straightforward way

$$\begin{aligned}\partial_t\zeta_t \leq (\varepsilon-\nu)\|\nabla\xi_t\|^2 + \frac{C}{\eta^2\nu}\|\xi_t\|^2 + \frac{\eta}{2}\|w_t\|_1^2\|\xi_t\|^2 \\ - 2t\varepsilon\nu\|\xi_t\|_2^2 + 2t\varepsilon C\|w_t\|_1\|\xi_t\|_2\|\xi_t\|_{1/2}\;,\end{aligned}$$

By (A.4), we get

$$\begin{aligned}2C\|w\|_1\|\xi\|_2\|\xi\|_{1/2} &\leq 2\sqrt{\eta\nu}\|w\|_1\|\xi\|_2\|\xi\|_1 + \frac{C}{\eta\nu}\|w\|_1\|\xi\|_2\|\xi\| \\ &\leq \nu\|\xi\|_2^2 + \eta\|w\|_1^2\|\xi\|_1^2 + \frac{C}{\eta^2\nu^3}\|w\|_1^2\|\xi\|^2\;.\end{aligned}$$

this immediately yields

$$\begin{aligned}\partial_t\zeta_t \leq (\varepsilon-\nu)\|\xi_t\|_1^2 + \frac{C}{\eta^2\nu}\|\xi_t\|^2 + \Big(\frac{\eta}{2} + \frac{t\varepsilon C}{\eta^2\nu^3}\Big)\|w_t\|_1^2\|\xi_t\|^2 \\ - t\varepsilon\nu\|\xi_t\|_2^2 + t\varepsilon\eta\|w_t\|_1^2\|\xi_t\|_1^2\;.\end{aligned}$$

If we take ε sufficiently small (of the order $\eta^3\nu^3$), we get

$$\partial_t\zeta_t \leq \Big(\frac{C}{\eta^2\nu} + \eta\|w_t\|_1^2\Big)\zeta_t\;,\quad \text{for } t\in[0,1],$$

and therefore

$$\|J_t\xi_0\|_1^2 \leq \frac{\tilde C}{t}\exp\Big(\eta\int_0^1 \|w_s\|_1^2\,ds\Big)\|\xi_0\|^2\;, \tag{A.8}$$

for some (possibly rather large) constant $\tilde C$. If we now define π_N as the orthogonal projection on the set of Fourier modes with $|k| \geq N$, we have

$$\|\pi_N\xi_t\| \leq \frac{1}{N}\|\xi_t\|_1\;.$$

The bound (4.23) immediately follows by taking $\pi_\ell = 1 - \pi_N$ for N sufficiently large.

We now turn to the proof of the bound (4.24). We define π_ℓ as above (but reserve the right to choose the precise value of N later) and set $\xi_t^\ell = \pi_\ell \xi_t$ and $\xi_t^h = (1-\pi_\ell)\xi_t$. With these notations, (4.24) amounts to obtaining bounds on $\|\xi_t\|$ with $\xi_0^\ell = 0$. Using the identity (A.1) we have

$$\begin{aligned}\partial_t \|\xi_t^\ell\|^2 &= -2\nu \|\xi_t^\ell\|_1^2 + 2\langle B(\mathcal{K}\xi_t^\ell, w_t), \xi_t^\ell\rangle \\ &\quad - 2\langle B(\mathcal{K}\xi_t^h, \xi_t^\ell), w_t\rangle - 2\langle B(\mathcal{K}w_t, \xi_t^\ell), \xi_t\rangle\;, \\ \partial_t \|\xi_t^h\|^2 &= -2\nu \|\xi_t^h\|_1^2 - 2\langle B(\mathcal{K}\xi_t, \xi_t^h), w_t\rangle - 2\langle B(\mathcal{K}w_t, \xi_t^h), \xi_t\rangle\;.\end{aligned}$$

Applying (A.2) to the right-hand side allows to get the bound

$$\begin{aligned}\partial_t \|\xi_t^\ell\|^2 &\le -2\nu \|\xi_t^\ell\|_1^2 + 2\langle B(\mathcal{K}\xi_t^\ell, w_t), \xi_t^\ell\rangle + C\|w_t\|_{1/2}\|\xi_t^\ell\|_1 \|\xi_t^h\|\;, \\ \partial_t \|\xi_t^h\|^2 &\le -2\nu \|\xi_t^h\|_1^2 + C\|w_t\|_{1/2}\|\xi_t^h\|_1\|\xi_t\|\;.\end{aligned}$$

We then bound the first line using (A.6) and the second line using $\|\xi_t^h\|_1^2 \ge N^2\|\xi_t^h\|^2$. We thus obtain

$$\begin{aligned}\partial_t \|\xi_t^\ell\|^2 &\le \Big(\frac{C}{\eta^2} + \eta\|w_t\|_1^2\Big)\|\xi_t^\ell\|^2 + C\|w_t\|_{1/2}^2 \|\xi_t^h\|^2\;, \\ \partial_t \|\xi_t^h\|^2 &\le -\nu N^2 \|\xi_t^h\|^2 + C\|w_t\|_{1/2}^2\|\xi_t\|^2\;,\end{aligned} \tag{A.9}$$

for an arbitrary value of η and for a constant C depending on ν but independent of N and η. Using the *a priori* bound from point 3 above for the Jacobian ξ and the interpolation inequality $\|w_s\|_{1/2}^2 \le \|w_s\|\|w_s\|_1$ immediately produces the bound

$$\begin{aligned}\|\xi_t^h\|^2 &\le e^{-\nu N^2 t}\|\xi_0^h\|^2 + C\int_0^t e^{-\nu N^2(t-s)}\|w_s\|_{1/2}^2 \|\xi_s\|^2\,ds \\ &\le e^{-\nu N^2 t}\|\xi_0^h\|^2 + \frac{C(T)\|\xi_0^h\|^2}{N} e^{\eta\int_0^t \|w_s\|_1^2\,ds}\sqrt{\int_0^t \|w_s\|_1^2\,ds}\,\sup_{s\in[0,t]}\|w_s\| \\ &\le \|\xi_0^h\|^2\Big(e^{-\nu N^2 t} + \frac{C(T)}{N}\exp\Big(\eta'\int_0^t \|w_s\|_1^2\,ds\Big)\sup_{s\in[0,t]}\|w_s\|\Big)\;,\end{aligned}$$

for an arbitrary $\eta' > \eta$. Combining this with the bound of point 1 above shows that, for every η, every γ, every p, and every T, there exists a constant N_0 such that

$$\mathbf{E}_w \|\xi_t^h\|^p \le 2e^{-\nu N^2 pt}\|\xi_0^h\|^p + \gamma e^{\eta\|w\|^2}\|\xi_0^h\|^p\;,$$

for all $t \in [0,T]$ and all $N \ge N_0$. Since $\xi_0^\ell = 0$ by assumption, it follows from (A.9) that

$$\|\xi_t^\ell\|^2 \le C\int_0^t \exp\Big(\frac{C(t-s)}{\eta^2} + \eta\int_s^t \|w_r\|_1^2\,dr\Big)\|w_s\|_{1/2}^2\|\xi_s^h\|^2\,ds$$

$$\le \Big(\int_0^t e^{\frac{8C(t-s)}{\eta^2}+8\eta\int_s^t \|w_r\|_1^2\,dr}\,ds\Big)^{1/8}\Big(\int_0^t \|\xi_s^h\|^8\,ds\Big)^{1/4}$$
$$\times \sqrt{\int_0^t \|w_s\|_1^2\,ds}\,\sup_{s\in[0,t]}\|w_s\|\;.$$

The required bound (4.24) now follows easily by taking expectations and using the previous bounds. □

References

[AN87] E. J. ANDERSON and P. NASH. *Linear programming in infinite-dimensional spaces*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Ltd., Chichester, 1987. Theory and applications, A Wiley-Interscience Publication.

[Bel87] D. R. BELL. *The Malliavin calculus*. Longman Scientific & Technical, Harlow, 1987.

[BKL01] J. BRICMONT, A. KUPIAINEN, and R. LEFEVERE. Ergodicity of the 2D Navier-Stokes equations with random forcing. *Comm. Math. Phys.* **224**, no. 1, (2001), 65–81. Dedicated to Joel L. Lebowitz.

[BKL02] J. BRICMONT, A. KUPIAINEN, and R. LEFEVERE. Exponential mixing of the 2D stochastic Navier-Stokes dynamics. *Comm. Math. Phys.* **230**, no. 1, (2002), 87–132.

[Cer99] S. CERRAI. Ergodicity for stochastic reaction-diffusion systems with polynomial coefficients. *Stochastics Stochastics Rep.* **67**, no. 1-2, (1999), 17–51.

[CF88] P. CONSTANTIN and C. FOIAŞ. *Navier-Stokes Equations*. University of Chicago Press, Chicago, 1988.

[CK97] P.-L. CHOW and R. Z. KHASMINSKII. Stationary solutions of nonlinear stochastic evolution equations. *Stochastic Anal. Appl.* **15**, no. 5, (1997), 671–699.

[Cru89] A. B. CRUZEIRO. Solutions et mesures invariantes pour des équations d'évolution stochastiques du type Navier-Stokes. *Exposition. Math.* **7**, no. 1, (1989), 73–82.

[DPZ96] G. DA PRATO and J. ZABCZYK. *Ergodicity for Infinite Dimensional Systems*. Cambridge, 1996.

[EH01] J.-P. ECKMANN and M. HAIRER. Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Commun. Math. Phys.* **219**, no. 3, (2001), 523–565.

[EL94] K. D. ELWORTHY and X.-M. LI. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125**, no. 1, (1994), 252–286.

[EM01] W. E and J. C. MATTINGLY. Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation. *Comm. Pure Appl. Math.* **54**, no. 11, (2001), 1386–1402.

[EMS01] W. E, J. C. MATTINGLY, and Y. G. SINAI. Gibbsian dynamics and ergodicity for the stochastic forced Navier-Stokes equation. *Comm. Math. Phys.* **224**, no. 1.

[Fer97] B. FERRARIO. Ergodic results for stochastic Navier-Stokes equation. *Stochastics and Stochastics Reports* **60**, no. 3–4, (1997), 271–288.

[FJMR02] C. FOIAS, M. S. JOLLY, O. P. MANLEY, and R. ROSA. Statistical estimates for the Navier-Stokes equations and the Kraichnan theory of 2-D fully developed turbulence. *J. Statist. Phys.* **108**, no. 3-4, (2002), 591–645.

[Fla94] F. FLANDOLI. Dissipativity and invariant measures for stochastic Navier-Stokes equations. *NoDEA* **1**, (1994), 403–426.

[FM95] F. FLANDOLI and B. MASLOWSKI. Ergodicity of the 2-D Navier-Stokes equation under random perturbations. *Commun. Math. Phys.* **172**, no. 1, (1995), 119–141.

[FP67] C. FOIAŞ and G. PRODI. Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2. *Rend. Sem. Mat. Univ. Padova* **39**, (1967), 1–34.

[Fri95] U. FRISCH. *Turbulence: the legacy of A. N. Kolmogorov*. Cambridge, 1995.

[Hai02] M. HAIRER. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probab. Theory Related Fields* **124**, no. 3, (2002), 345–380.

[Hör67] L. HÖRMANDER. Hypoelliptic second order differential equations. *Acta Math.* **119**, (1967), 147–171.

[Hör85] L. HÖRMANDER. *The Analysis of Linear Partial Differential Operators I–IV*. Springer, New York, 1985.

[Kan42] L. V. KANTOROVICH. On the translocation of masses. *Dokl. Akad. Nauk SSSR* **37**, (1942), 194–201.

[Kan48] L. V. KANTOROVICH. On a problem of Monge. *Uspekhi Mat. Nauk* **3**, no. 2, (1948), 225–226.

[KS00] S. KUKSIN and A. SHIRIKYAN. Stochastic dissipative PDEs and Gibbs measures. *Comm. Math. Phys.* **213**, no. 2, (2000), 291–330.

[KS01] S. B. KUKSIN and A. SHIRIKYAN. A coupling approach to randomly forced nonlinear PDE's. I. *Commun. Math. Phys.* **221**, (2001), 351–366.

[Mat98] J. C. MATTINGLY. *The Stochastically forced Navier-Stokes equations: energy estimates and phase space contraction*. Ph.D. thesis, Princeton University, 1998.

[Mat99] J. C. MATTINGLY. Ergodicity of 2D Navier-Stokes equations with random forcing and large viscosity. *Comm. Math. Phys.* **206**, no. 2, (1999), 273–288.

[Mat02a] J. C. MATTINGLY. The dissipative scale of the stochastics Navier-Stokes equation: regularization and analyticity. *J. Statist. Phys.* **108**, no. 5-6, (2002), 1157–1179.

[Mat02b] J. C. MATTINGLY. Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. *Commun. Math. Phys.* **230**, no. 3, (2002), 421–462.

[Mat03] J. C. MATTINGLY. On recent progress for the stochastic Navier-Stokes equations. In *Journées Équations aux dérivées partielles*, Forges-les-Eaux. 2003.

[MP04] J. C. MATTINGLY and E. PARDOUX. Malliavin calculus and the randomly forced Navier-Stokes equation, 2004. To be published in Comm. Pure Appl. Math.

[MR] R. MIKULEVICIUS and B. L. ROZOVSKII. Stochastic Navier-Stokes equations for turbulent flows. To appear in SIAM J. Math. Analysis.

[MS03] J. C. MATTINGLY and T. M. SUIDAN. The small scales of the stochastic Navier-Stokes equations under rough forcing, 2003. JSP ??????

[MY02] N. MASMOUDI and L.-S. YOUNG. Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs. *Comm. Math. Phys.* **227**, no. 3, (2002), 461–481.

[Nua95] D. NUALART. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, New York, 1995.

[Rac91] S. T. RACHEV. *Probability metrics and the stability of stochastic models*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1991.

[Ros02] R. M. S. ROSA. Some results on the Navier-Stokes equations in connection with the statistical theory of stationary turbulence. *Appl. Math.* **47**, no. 6, (2002), 485–516. Mathematical theory in fluid mechanics (Paseky, 2001).