

ERGODIC AND MIXING PROPERTIES OF THE BOUSSINESQ EQUATIONS WITH A  
DEGENERATE RANDOM FORCING

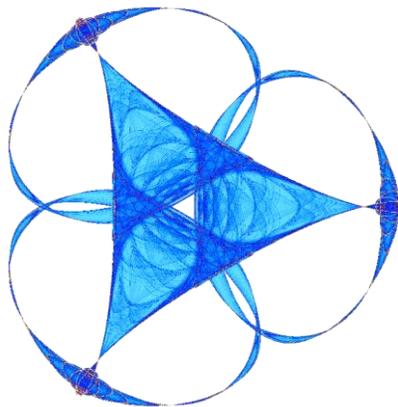
By

Juraj Földes♥, Nathan Glatt-Holtz♦, Geordie Richards♣, Enrique Thomann♠

♥Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN  
♦Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA  
♣Department of Mathematics, University of Rochester, Rochester, NY  
♠Department of Mathematics, Oregon State University, Corvallis, OR

**IMA Preprint Series #2421**

(November 2013)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS  
UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455-0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: <http://www.ima.umn.edu>

# Ergodic and Mixing Properties of the Boussinesq Equations with a Degenerate Random Forcing\*

Juraj Földes<sup>♥</sup>, Nathan Glatt-Holtz<sup>♦</sup>, Geordie Richards<sup>♣</sup>, Enrique Thomann<sup>♠</sup>

<sup>♥</sup>*Institute for Mathematics and its Applications, University of Minnesota  
Minneapolis, MN 55455*

<sup>♦</sup>*Department of Mathematics, Virginia Polytechnic Institute and State University  
Blacksburg, VA 24061*

<sup>♣</sup>*Department of Mathematics, University of Rochester  
Rochester, NY 14627*

<sup>♠</sup>*Department of Mathematics, Oregon State University  
Corvallis, OR 97331*

## Abstract

We establish the existence, uniqueness and attraction properties of an ergodic invariant measure for the Boussinesq Equations in the presence of a degenerate stochastic forcing acting only in the temperature equation and only at the largest spatial scales. The central challenge is to establish time asymptotic smoothing properties of the Markovian dynamics corresponding to this system. Towards this aim we encounter a Lie bracket structure in the associated vector fields with a complicated dependence on solutions. This leads us to develop a novel Hörmander-type condition for infinite-dimensional systems. Demonstrating the sufficiency of this condition requires new techniques for the spectral analysis of the Malliavin covariance matrix.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Mathematical Setting and Background</b>	<b>7</b>
<b>3</b>	<b>Gradient Estimates for the Markov Semigroup</b>	<b>13</b>
<b>4</b>	<b>Spectral Bounds for the Malliavin Covariance Matrix</b>	<b>21</b>
<b>5</b>	<b>Lie Bracket Computations</b>	<b>25</b>
<b>6</b>	<b>Noise Propagation in the Phase Space: Quantitative Estimates</b>	<b>30</b>
<b>7</b>	<b>Mixing and Other Convergence Properties</b>	<b>43</b>
<b>A</b>	<b>Appendix</b>	<b>47</b>

---

\*November 14, 2013

# 1 Introduction

In this work we analyze the stochastically forced Boussinesq equations for the velocity field  $\mathbf{u} = (u_1, u_2)$ , (density-normalized) pressure  $p$ , and temperature  $\theta$  of a viscous incompressible fluid. These equations take the form

$$d\mathbf{u} + (\mathbf{u} \cdot \nabla \mathbf{u})dt = (-\nabla p + \nu_1 \Delta \mathbf{u} + \mathbf{g}\theta)dt, \quad \nabla \cdot \mathbf{u} = 0, \quad (1.1)$$

$$d\theta + (\mathbf{u} \cdot \nabla \theta)dt = \nu_2 \Delta \theta dt + \sigma_\theta dW, \quad (1.2)$$

where the parameters  $\nu_1, \nu_2 > 0$  are respectively the kinematic viscosity and thermal diffusivity of the fluid and  $\mathbf{g} = (0, g)^T$  with  $g \neq 0$  is the product of the gravitational constant and the thermal expansion coefficient. The spatial variable  $x = (x_1, x_2)$  belongs to a two-dimensional torus  $\mathbb{T}^2$ . That is, we impose periodic boundary conditions in space. We consider a degenerate stochastic forcing  $\sigma_\theta dW$ , which acts only on a few Fourier modes and exclusively through the temperature equation.

We prove that there exists a unique statistically invariant state of the system (1.1)–(1.2). More precisely, we establish:

**Theorem 1.1.** *With white noise acting only on the two largest standard modes of the temperature equation (1.2),*

$$\sigma_\theta dW = \alpha_1 \cos x_1 dW^1 + \alpha_2 \sin x_1 dW^2 + \alpha_3 \cos x_2 dW^3 + \alpha_4 \sin x_2 dW^4,$$

*the Markov semigroup corresponding to (1.1)–(1.2) possesses a unique ergodic invariant measure. Moreover this measure is mixing, and it obeys a law of large numbers and a central limit theorem.*

The interaction between the nonlinear and stochastic terms in (1.1)–(1.2) is delicate, and leads us to develop a novel infinite-dimensional form of the Hörmander bracket condition. Our analysis generalizes techniques developed in the recent works [MP06, HM06, HM08, HM11], and we believe it has broader interest for systems of SPDEs.

## 1.1 Historical Background and Motivations

Going back to the early 1900’s Rayleigh [LR16] proposed the study of buoyancy driven fluid convection problems using the equations of Boussinesq [Bou97] in order to explain the experimental work of Bénard [Bó1]. Today this system of equations plays a fundamental role in a wide variety of physical settings including climate and weather, the study of plate tectonics, and the internal dynamical structure of stars, see e.g. [Ped82, Bus89, Get98, DG95, BPA00] and references therein for further background.

Physically speaking, the system (1.1)–(1.2) (with  $\sigma_\theta = 0$ ) arises as follows. Consider a fluid with velocity  $\mathbf{u}$  confined between two horizontal plates, where one fixes the temperature  $\theta$  of the fluid on the top  $\theta_t$ , and bottom  $\theta_b$ , with  $\theta_t \leq \theta_b$  (i.e. heating from below). It is typical to assume a linear relationship between density and temperature, and to impose the Boussinesq approximation, which posits that the only significant role played by density variations in the fluid arise through the gravitational terms, so that the fluid velocity  $\mathbf{u}$  and temperature  $\theta$  evolve according to (1.1)–(1.2). Due to the presence of viscosity, the fluid is not moving at the plates and one assumes no-slip boundary conditions  $\mathbf{u} = 0$ .<sup>1</sup>

The form of the Boussinesq equations we consider in this work, that is (1.1)–(1.2) supplemented with periodic boundary conditions, is sometimes referred to in the physics community as the ‘homogeneous Rayleigh-Bénard’ or ‘HRB’ system. It is derived as follows: One transforms the governing equations we have just described into an equivalent homogenous system by subtracting off a linear temperature profile. This introduces an additional excitation term in the temperature equation, and makes the temperature vanish at the plates. As a numerical simplification, one then replaces these boundary conditions with periodic ones

<sup>1</sup>One can consider consider the equations posed on an infinite channel, or assume periodic boundary conditions in the horizontal direction for both the velocity field and temperature.

(see [LT03, CLTT05]). This periodic setting is controversial in the physics community since it can produce unbounded (‘grow-up’) solutions, as has been observed both numerically and through explicit solutions, [CRG<sup>+</sup>06]. We will consider (1.1)–(1.2) in situations with *no temperature differential* ( $\theta_t = \theta_b$ , i.e. zero Rayleigh number), so that the additional excitation term is not present and such unbounded solutions do not exist. Extensions to more physically realistic boundary conditions for (1.1)–(1.2) will be addressed in forthcoming works.

In the mathematical community the deterministic Boussinesq equations with various boundary conditions on bounded and unbounded domains have attracted considerable attention. In one line of work the 2D system has been interpreted as an analogue of 3D axisymmetric flow where ‘vortex stretching’ terms appear in the ‘vorticity formulation’, see e.g. [CD80, CI99, CFdL04, HL05, Cha06, DP08, DP09, HK09, ACW10, LLT10, ACW10, DP11, CW12, CW13]. Other authors have sought to provide a rigorous mathematical framework for various physical and numerical observations in fluid convection problems, see e.g. [CD96, CD99, DC01, MW04, Wan04a, Wan04b, Wan05, Wan07, MW07, Wan08a, Wan08b, SW13].

Let us briefly motivate the stochastic forcing appearing in (1.1)–(1.2). Due to sensitivity with respect to initial data and parameters, individual solutions of the basic equations of fluid mechanics are unpredictable and seemingly chaotic. However, some of their statistical properties of solutions are robust. As early as the 19th century J.V. Boussinesq conjectured that turbulent flow cannot be solely described by deterministic methods, and indicated that a stochastic framework should be used, see [Sta88]. More recently the study of the Navier-Stokes equations with degenerate white noise forcing has been proposed a proxy for the large-scale ‘generic’ stirring which is assumed in the basic theories of turbulence; this setting is ubiquitous in the turbulence literature, see e.g. [Nov65, VKF79, Eyi96] and containing references. In this view, invariant measures of the stochastic equations of fluid dynamics would presumably contain the statistics posited by these theories.<sup>2</sup> The closely related question of unique ergodicity and mixing provides rigorous justification for the explicit and implicit measurement assumptions invoked by physicists and engineers when measuring statistical properties of turbulent systems.

The existence of invariant measures for forced-dissipative systems is often easy to prove with classical tools, namely by making use of the Krylov-Bogoliubov averaging procedure with energy (compactness) estimates, but the uniqueness of these measures is a deep and subtle issue. To establish this uniqueness one can follow the path laid out by the Doob-Khasminskii Theorem [Doo48, Km60, DPZ96], and more recently expanded upon in [HM06, HM08]. This strategy requires the proof of certain smoothing properties of the associated Markov semigroup, and to show that a common state can be reached by the dynamics regardless of initial conditions (irreducibility). Without stochastic forcing, solutions of our system converge to the trivial equilibrium, so that the proof of irreducibility is straightforward in our context. Thus the main challenge of this work is to establish sufficient smoothing properties for the Markov semigroup associated to (1.1)–(1.2).

## 1.2 Smoothing and Hypocoercivity in Infinite Dimensional Systems

In order to discuss the difficulties in establishing smoothing properties for the Markov semigroup it is useful to recall the canonical relationship between stochastic evolution equations and their corresponding Fokker-Planck (Kolmogorov) equations. Consider an abstract equation on a Hilbert space  $H$ ,

$$dU = F(U)dt + \sigma dW(t); \quad U(0) = U_0, \quad (1.3)$$

where  $\sigma dW(t) := \sum_{k=1}^N \sigma_k dW^k(t)$ , and  $\{W_k\}_{k=1}^N$  is a (finite or infinite) collection of independent 1D Brownian motions. We denote solutions  $U$  with the initial condition  $U_0$  at time  $t \geq 0$  by  $U(t, U_0)$ , and define

---

<sup>2</sup>In our context the 2D Batchelor-Krichanan theory [Kra67, Bat69] is probably the most relevant statistical theory. Note that its applicability would presumably require the imposition of a large scale damping term in the momentum equation as explained in [GHSV13] and see also e.g. [KM80, Tab02, FJMR02, Kup10]. Such a damping operator would not affect any of the conclusions drawn in the main results below. We also mention recent work on the statistics of turbulence in axisymmetric 3D flows [TDB13, NTC<sup>+</sup>10]. The governing equations for these systems bare a strong structural resemblance to (1.1)–(1.2) as has been pointed out in e.g. [HL05].

the Markov semigroup associated to (1.3) according to  $P_t\Phi(U_0) = \mathbb{E}\Phi(U(t, U_0))$ , where  $\Phi : H \rightarrow \mathbb{R}$  is any ‘observable’. Then  $\Psi(t) := P_t\Phi$  solves the Fokker-Planck equation corresponding to (1.3) given by

$$\partial_t\Psi = \frac{1}{2}\text{Tr}[(\sigma\sigma^*)D^2\Psi] + \langle F(U), D\Psi \rangle, \quad \Psi(0) := \Phi \quad (1.4)$$

where we view  $\sigma$  as an element in  $\mathcal{L}(\mathbb{R}^N, H)$ . The interested reader should consult [Cer01, DPZ02] for more on the general theory of second order PDEs posed on a Hilbert space.

There is a wide literature devoted to proving uniqueness and associated mixing properties of invariant measures for nonlinear stochastic PDEs when  $\sigma\sigma^*$  is non-degenerate or mildly degenerate. See e.g. [VKF79, Cru89, FM95, DPZ96, Fer97, Mat99, Mat02, E01, EMS01, BKL01, EH01, MY02, KS01, KS02, Mat02, Mat03, LW04, GM05, DPRRW06, AFS08, Deb11, KS12, CGHV13] and references therein. Roughly speaking, the fewer the number of driving stochastic terms in (1.3), the more degenerate the diffusion in (1.4), and the more difficult it becomes to establish smoothing properties for the Markov semigroup  $\{P_t\}_{t \geq 0}$ .<sup>3</sup> Moreover, while even the non-degenerate setting poses many interesting mathematical challenges, such stochastic forcing regimes are highly unsatisfactory from the point of view of turbulence where one typically assumes a clear separation between the forced and dissipative scales of motion.

Going back to the seminal work of Hörmander [Hö7] (and cf. [Kol34]), a theory of parabolic regularity for finite dimensional PDEs of the general form (1.4) with degenerate diffusion terms was developed. This theory of ‘hypoellipticity’ can be interpreted in terms of finite-dimensional stochastic ODEs for which these degenerate parabolic PDEs are the corresponding Kolmogorov equations. This connection suggested the potential for a more probabilistic approach, initiating the development of the so-called Malliavin calculus; see [Mal78] and subsequent authors [Shi80, Bis81a, Bis81b, Str81, IW89, KS84, KS85, Nor86, KS87]. In any case, the work of Hörmander and Malliavin has led to an extensive theory of unique ergodicity and mixing properties for finite-dimensional stochastic ODEs.

By comparison, for stochastic PDEs (which correspond to the situation when (1.4) is posed on an infinite dimensional space) this theory of hypoellipticity remains in its infancy. Recently, however, in a series of groundbreaking works [MP06, HM06, HM08, HM11], a theory of unique ergodicity for degenerately forced infinite-dimensional stochastic systems has emerged. These works produced two fundamental contributions: Firstly the authors demonstrated that to establish the uniqueness of the invariant measure it suffices to prove time asymptotic smoothing (asymptotic strong Feller property) instead of ‘instantaneous’ smoothing (strong Feller property). This is an abstract result from probability and it applies to very general settings including ours. Secondly the authors generalize the methods of Malliavin (and subsequent authors) in order to prove the asymptotic strong Feller property for certain infinite-dimensional stochastic systems. These works resulted in an infinite-dimensional analogue of the Hörmander bracket condition.

For the second point the application of the methods in [HM11] is more delicate and must be considered on a case-by-case basis; it requires a careful analysis of the interaction of the nonlinear and stochastic terms of the system. In our situation the bracket condition in [HM11] is not satisfied and needs to be replaced by a weaker notion. This required us to rework and generalize many basic elements of their approach.

### 1.3 The Hörmander bracket condition in Infinite Dimensions

To explain our contributions at a more technical level, let us recall what is meant by a ‘Hörmander bracket condition’ in the context of systems of the general form (1.3). Define

$$\mathcal{V}_0 := \text{span}\{\sigma_k : k = 1, \dots, d\},$$

and for  $m > 0$  take

$$\mathcal{V}_m := \text{span}\{[E, F], [E, \sigma_k], E : E \in \mathcal{V}_{m-1}, k = 1, \dots, d\}, \quad (1.5)$$

---

<sup>3</sup>On the other hand with more driving terms in (1.3), the well-posedness theory may become more difficult. For example this was the primary mathematical challenge in the pioneering work [FM95].

where  $F = F(U)$  is the drift term in (1.3), and for any Fréchet differentiable  $E_1, E_2 : H \rightarrow H$ ,

$$[E_1, E_2](U) := \nabla E_2(U)E_1(U) - \nabla E_1(U)E_2(U). \quad (1.6)$$

This operation  $[E_1, E_2]$  is referred to as the *Lie bracket* of the two ‘vector fields’  $E_1, E_2$ . In finite dimensions, that is when  $H = \mathbb{R}^N$ , the classical Hörmander condition is satisfied if  $\cup_{k \geq 0} \mathcal{V}_k = H$  for all  $U \in H$ .

This condition is not suitable for infinite-dimensional settings since the effect of randomness from the directly forced modes seems to weaken as it propagates out in infinite dimensional space. In [HM06, HM11] it was shown that, due to contractive properties of the flow on high modes, it is enough to require that the Lie brackets span all the unstable directions of the system. More precisely, the following non-degeneracy assumption is required: For every  $N > 0$ , there is an  $M > 0$  and a finite set  $\mathfrak{B} \subset \mathcal{V}_M$  such that the quadratic form

$$\langle \mathcal{Q}(U)\phi, \phi \rangle := \sum_{b(U) \in \mathfrak{B}} \langle \phi, b(U) \rangle^2 \quad (1.7)$$

satisfies, for each  $\alpha > 0$ ,

$$\langle \mathcal{Q}(U)\phi, \phi \rangle \geq \alpha C \|\phi\|^2, \quad \text{for every } \phi \in \mathcal{S}_{\alpha, N} := \{\phi \in H : \|P_N \phi\|^2 > \alpha \|\phi\|^2\}, \quad (1.8)$$

where  $U$  solves (1.1)-(1.2),  $C > 0$  is a constant independent of  $\alpha$ , and  $\{P_N\}$  is a sequence of projection operators onto successively larger spaces.<sup>4</sup>

It was suggested in [HM11] that for many equations with polynomial nonlinearities brackets of the type  $[\dots[[F(U), \sigma_{k_1}], \dots], \sigma_{k_m}]$  (where  $\sigma_{k_1}, \dots, \sigma_{k_m}$  are previously generated constant vector fields) suffice to build quadratic forms satisfying (1.8). Indeed, this approach has now been successfully employed for several important examples, including the 2D and 3D Navier-Stokes equations and the Ginzburg-Landau equations; see [EM01, Rom04, HM06, HM11]. In these works algebraic conditions on the set of stochastically forced modes have been derived which guarantee that any finite-dimensional space  $H_N$  can be generated from these types of brackets; one obtains a collection of  $U$ -independent elements  $\mathfrak{B}$  which form an orthonormal basis for  $H_N$ . This strategy has proven effective for certain scalar equations, but its limitations are evident in slightly more complicated situations. We believe that (1.1)–(1.2) provides an illuminating case study of these difficulties which has led us to generalize (1.8).

Observe that our model is distinguished by two key structural properties. Firstly, the buoyancy term  $\mathbf{g}\theta$  is the only means of spreading the effect of the stochastic forcing from the temperature equation, (1.1), to the momentum equations, (1.2). In particular note that this buoyancy term is linear, and therefore vanishes after two Lie bracket operations with constant vector fields. Secondly, the advective structure in (1.1)–(1.2) leads to a delicate ‘asymmetry’ in the nonlinear terms. For example this means that a more refined analysis is needed to address the spread of noise in the temperature equation (1.2) alone.<sup>5</sup> Concretely we find that, by combining these observations, one obtains  $[[F(U), \sigma_{k_1}], \sigma_{k_2}] = 0$  for any constant vector fields  $\sigma_{k_1}, \sigma_{k_2}$  concentrated in the temperature component of the phase space. In view of these discussions it is worth emphasizing that our chosen forcing - that is, noise acting only through the temperature equations - is the most interesting situation to consider from the point of view of ergodicity.

To get around these difficulties we make careful use of the interaction between the nonlinear, buoyancy and noise terms in (1.1)–(1.2) as follows: We see that noise activated through the temperature equation is spread to the momentum equations. It is then advected in the temperature equations and combines again with the noise terms to produce new directions in the temperature component of the phase space. These loose observations find concrete expressions in a series of Lie brackets of the form  $[[[[F, \sigma_{k_1}], F], \sigma_{k_2}]]$  for constant vector fields  $\sigma_{k_1}$  and  $\sigma_{k_2}$  which belong to the  $\theta$ -component. Remarkably, we found that this

<sup>4</sup>In fact the condition given in [HM11] is slightly more general than this. They also allow for the case when  $C$  might depend on  $U$  (subject to suitable moment bounds) but they require that  $C$  be almost everywhere positive which is *not* sufficient for our purposes.

<sup>5</sup>In simple language this ‘advective structure’ means that  $\mathbf{u} \cdot \nabla \theta$  is the only second order term in the temperature equation (1.2).

chain of admissible brackets leads to new constant vector fields exclusively in the  $\theta$ -component of the phase space. This surprising observation requires a series of detailed computations but is perhaps anticipated by the advective structure of the nonlinear terms.

It is in addressing the spread of noise in the momentum components of the phase space that the condition (1.8) breaks down. Forced directions in the temperature component are pushed to the momentum components through the buoyancy term. However, due to the presence of the nonlinearity, they are ‘mixed’ with terms which have an unavoidable and complicated directional (non-frequency-localized) dependence on  $U$ . Crucially, due to the advective structure in (1.1)–(1.2), these ‘error terms’ are concentrated only in the temperature component of the phase space. We are therefore able to push these error terms to small scales by using the ‘pure’  $\theta$  directions already generated (following the procedure described in the previous paragraph). More precisely, in the language of (1.5), (1.7), we are able to show that for every  $N, \tilde{N}$  we can find  $M > 0$  and sets  $\mathfrak{B} \subset \mathcal{V}_M$  consisting of elements of the form  $\Psi_k + J_k^{\tilde{N}}(U)$ , where the sequence  $\{\Psi_k\}$  (which are essentially vectors consisting of trigonometric functions) forms an orthonormal basis for the phase space and  $J_k^{\tilde{N}}$  are functions taking values in  $H$  with a complicated dependence on  $U$  but which are supported on ‘high’ frequencies, i.e. wavenumbers larger than  $\tilde{N}$ .

These structural observations for (1.1)–(1.2) lead us to formulate the following generalization of (1.8).

**Definition 1.2.** *Let  $H$  and  $\tilde{H}$  be Hilbert spaces with  $\tilde{H}$  compactly embedded in  $H$ . We say that (1.3) satisfies the generalized Hörmander condition if for every  $N > 0$  and every  $\epsilon > 0$  there exist  $M = M(\epsilon, N) > 0$ ,  $p \geq 1$ , and a finite set  $\mathfrak{B} \subset \mathcal{V}_M$  (where  $\mathcal{V}_M$  is defined according to (1.5)) such that*

$$\langle \mathcal{Q}(U)\phi, \phi \rangle \geq C \left( \alpha - \epsilon(1 + \|U\|_{\tilde{H}}^p) \right) \|\phi\|^2 \quad \text{for every } \phi \in \mathcal{S}_{\alpha, N}, \quad (1.9)$$

where  $C$  is independent of  $\alpha$  and  $\epsilon$ , and  $\mathcal{Q} = \mathcal{Q}_{N, \epsilon}$ ,  $\mathcal{S}_{\alpha, N}$  are defined in (1.7), (1.8) respectively.

Below we demonstrate that the condition (1.9) is sufficient to establish suitable time asymptotic smoothing properties for the Markov semigroup associated to (1.3)<sup>6</sup> This allows us to apply the abstract results from [HM06, HM08, KW12] to complete the proof of Theorem 1.1. The main technical challenge arising from the modified condition (1.9) is that it requires us to significantly rework the spectral analysis of the Malliavin matrix appearing in [MP06, HM06, HM11]. The technically oriented reader can skip immediately to Section 4 for further details.

## 1.4 Organization of the Manuscript

The manuscript is organized as follows: In Section 2 we restate our problem in an abstract functional setting and introduce some general definitions and notations. Then we reduce the question of uniqueness of the invariant measure to establishing a time asymptotic gradient estimate on the Markov semigroup. Next, in Section 3 we explain how, using the machinery of Malliavin calculus, this gradient bound reduces to a control problem for a linearization of (1.1)–(1.2). In turn we show that this control problem may be solved by establishing appropriate spectral bounds for the Malliavin covariance matrix  $\mathcal{M}$ . Section 4 is devoted to proving that our new form of the Hörmander condition, (1.9), implies these spectral bounds, modulo some technical estimates postponed for Section 6. Section 5 provides detailed Lie bracket computations leading to the modified condition (1.9). Finally in Section 7 we establish mixing properties, a law of large numbers and a central limit theorem for the invariant measure by making careful use of recent abstract results from [Shi06, HM08, KW12]. Appendices A.1 and A.2 collect respectively statistical moment bounds for (1.1)–(1.2) (and associated linearizations) and a brief review of some elements of the Malliavin calculus used in our analysis.

<sup>6</sup>For clarity and simplicity we present all the arguments in the context of the Boussinesq system but the implications for the Markov semigroup for systems satisfying Definition 1.2 could be shown to hold in a more general setting with essentially the same analysis.

## 2 Mathematical Setting and Background

In this section we formulate (1.1)–(1.2) as an abstract evolution equation on a Hilbert space and define its associated Markovian framework. Then we formulate our main result in Theorem 2.3 and give an outline of the proof, which sets the agenda for the work below.

In the rest of the paper, we consider (1.1) in the equivalent, vorticity formulation. Namely, if we denote  $\omega := \nabla^\perp \cdot \mathbf{u} = \partial_x u_2 - \partial_y u_1$ , then by a standard calculation we obtain

$$d\omega + (\mathbf{u} \cdot \nabla \omega - \nu_1 \Delta \omega) dt = g \partial_x \theta dt, \quad (2.1)$$

$$d\theta + (\mathbf{u} \cdot \nabla \theta - \nu_2 \Delta \theta) dt = \sigma_\theta dW. \quad (2.2)$$

The system (2.1)–(2.2) is posed on  $\mathbb{T}^2 \times (0, \infty)$ , where  $\mathbb{T}^2$  is the square torus  $\mathbb{T}^2 = [-\pi, \pi]^2 = \mathbb{R}^2 / (2\pi\mathbb{Z}^2)$ .<sup>7</sup> To close the system (2.1)–(2.2), we, as usual, calculate  $\mathbf{u}$  from  $\omega$  by Biot-Savart law, that is  $\mathbf{u} = K * \omega$ , where  $K$  is the Biot-Savart kernel, so that  $\nabla^\perp \cdot \mathbf{u} = \omega$  and  $\nabla \cdot \mathbf{u} = 0$ , see e.g. [Tem01, MB02] for further details. As mentioned in the introduction, the physically interesting case of non-periodic domains will be the subject of a future work.

We next introduce a functional setting for the system (2.1)–(2.2). The phase space is composed of mean zero, square integrable functions

$$H := \left\{ U := (\omega, \theta)^T \in (L^2(\mathbb{T}^2))^2 : \int_{\mathbb{T}^2} \omega dx = \int_{\mathbb{T}^2} \theta dx = 0 \right\} \quad (2.3)$$

equipped with the norm

$$\|U\|^2 := \frac{\lambda_1 \nu_1 \nu_2}{g^2} \|\omega\|_{L^2}^2 + \|\theta\|_{L^2}^2, \quad (2.4)$$

where  $\lambda_1 = 1$  (we omit  $\lambda_1$  below) is the principal eigenvalue of  $-\Delta$  on  $H$ . Observe that this norm is equivalent to the standard norm on the space  $(L^2(\mathbb{T}^2))^2$ .<sup>8</sup> Our choice of the norm is natural as both terms on the right hand side of (2.4) have been suitably non-dimensionalized. The associated inner product on  $H$  is denoted by  $\langle \cdot, \cdot \rangle$ . Note furthermore that the zero mean property embodied in the definition of  $H$ , (2.3) is to be maintained by the flow (2.1)–(2.2). The higher order Sobolev spaces are denoted

$$H^s := \left\{ U := (\omega, \theta)^T \in (W^{s,2}(\mathbb{T}^2))^2 : \int_{\mathbb{T}^2} \omega dx = \int_{\mathbb{T}^2} \theta dx = 0 \right\} \quad \text{for any } s \geq 0,$$

where  $W^{s,2}(\mathbb{T}^2)$  is classical Sobolev-Slobodeckii space, and  $H^s$  is equipped with the norm

$$\|U\|_{H^s}^2 := \frac{\nu_1 \nu_2}{g^2} \|\omega\|_{W^{s,2}}^2 + \|\theta\|_{W^{s,2}}^2.$$

For  $s > 0$ , we also denote  $H^{-s} := (H^s)^*$ , the dual space to  $H^s$ .

Since we need to estimates for linearizations of (2.1)–(2.2) around initial conditions and noise paths, we encounter elements in  $\mathcal{L}(H)$  and  $\mathcal{L}(H, \mathcal{L}(H))$  (where  $\mathcal{L}(X) = \mathcal{L}(X, X)$ ,  $\mathcal{L}(X, Y)$  is the space of linear operators from  $X$  to  $Y$ ). See, for example (3.4), below. We will sometimes abuse notation and denote

$$\|\mathcal{J}\| := \|\mathcal{J}\|_{\mathcal{L}(H)}, \quad \|\mathcal{J}^{(2)}\| := \|\mathcal{J}^{(2)}\|_{\mathcal{L}(H, \mathcal{L}(H))}, \quad (2.5)$$

for any  $\mathcal{J} \in \mathcal{L}(H)$ ,  $\mathcal{J}^{(2)} \in \mathcal{L}(H, \mathcal{L}(H))$ .

We also frequently work with finite dimensional subspaces of  $H$  along with the projection operators onto these spaces. Fix the trigonometric basis:

$$\sigma_k^0(x) := (0, \cos(k \cdot x))^T, \quad \sigma_k^1(x) := (0, \sin(k \cdot x))^T, \quad (2.6)$$

<sup>7</sup>Of course, by rescaling we obtain our results for any square torus  $\mathbb{R}^2 / (L\mathbb{Z}^2)$  and at the cost of more complicated notation and expressions below, one can also prove our results for non-square tori.

<sup>8</sup>Here and below we drop the dependence of various function spaces on  $\mathbb{T}^2$ .

and

$$\psi_k^0(x) := (\cos(k \cdot x), 0)^T, \quad \psi_k^1(x) := (\sin(k \cdot x), 0)^T. \quad (2.7)$$

We denote

$$\mathbb{Z}_+^2 := \{j = (j_1, j_2) \in \mathbb{Z}_0^2 : j_1 > 0 \text{ or } j_1 = 0, j_2 > 0\}$$

and for any  $N \geq 1$  define

$$H_N := \text{span}\{\sigma_k^l, \psi_k^l : |k| \leq N, l \in \{0, 1\}\}, \quad (2.8)$$

along with the associated projection operators

$$P_N : H \rightarrow H_N \text{ the orthogonal projection onto } H_N, \quad Q_N := I - P_N. \quad (2.9)$$

Note that  $Q_N$  maps  $H$  onto  $\text{span}\{\sigma_k^l, \psi_k^l : |k| > N, l \in \{0, 1\}\}$ .

In order to rewrite (2.1)–(2.2) in a functional form we introduce the following abstract operators associated to the various terms in the equation. For  $U := (\omega, \theta)$  and  $\tilde{U} = (\tilde{\omega}, \tilde{\theta})$ , let  $A : D(A) \subset H \rightarrow H$  be the linear symmetric positive definite operator defined by

$$AU := (-\nu_1 \Delta \omega, -\nu_2 \Delta \theta)^T,$$

for any  $U \in D(A)$ . Note that the scale of spaces  $H^s$ ,  $s \in \mathbb{R}$  coincides with the interpolation spaces between  $D(A)$  and  $H$  and  $A$  is a bounded operator from  $H^{s+2}$  to  $H^s$ .

For the inertial (non-linear) terms define  $B : H^1 \times H^1 \rightarrow H$  by

$$B(U, \tilde{U}) := ((K * \omega) \cdot \nabla \tilde{\omega}, (K * \omega) \cdot \nabla \tilde{\theta})^T, \quad (2.10)$$

for  $U, \tilde{U} \in H^1$ . It is well known that  $\|K * \omega\|_{H^s} \leq C \|\omega\|_{H^{s-1}}$ , and since  $H^2 \hookrightarrow L^\infty$ , we indeed obtain that  $B(U, \tilde{U}) \in H$ . Also set  $B(U) := B(U, U)$ . Finally, for the ‘buoyancy term’ define  $G : H^1 \rightarrow H$  by

$$GU = (g \partial_x \theta, 0)^T, \quad (2.11)$$

for  $U \in H^1$ .

Next we focus on the stochastic forcing terms appearing in (2.2). We introduce a finite set  $\mathcal{Z} \subset \mathbb{Z}_+^2$  which represents the forced directions in Fourier space. The driving noise process  $W := (W^{k,l})_{k \in \mathcal{Z}, l=0,1}$  is a  $d := 2 \cdot |\mathcal{Z}|$ -dimensional Brownian motion defined relative to a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  and we refer to the resulting tuple  $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}, W)$  as a *stochastic basis*.<sup>9</sup> Let  $\{e_k^l\}_{k \in \mathcal{Z}, l=0,1}$  be the standard basis of  $\mathbb{R}^{2|\mathcal{Z}|}$  and let  $\{\alpha_k^l\}_{k \in \mathcal{Z}, l=0,1}$  be a sequence of non-zero numbers. We define a linear map  $\sigma_\theta : \mathbb{R}^{2|\mathcal{Z}|} \rightarrow H$  such that

$$\sigma_\theta e_k^l := \alpha_k^l \sigma_k^l \quad \text{for any } k \in \mathcal{Z}, l \in \{0, 1\}. \quad (2.12)$$

where,  $\sigma_k^l$  are the basis elements defined in (2.6). Denote the Hilbert-Schmidt norm of  $\sigma_\theta$  by

$$\|\sigma_\theta\|^2 := \|\sigma_\theta^* \sigma_\theta\| = \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0,1\}}} (\alpha_k^l)^2.$$

<sup>9</sup>We may take the stochastic basis to be the standard ‘Wiener space’. Here we let  $\Omega = \{\omega \in C([0, \infty); \mathbb{R}^{2|\mathcal{Z}|}) : \omega(0) = 0\}$  with its Borelian  $\sigma$ -algebra and take  $\mathbb{P}$  to be the Wiener measure. On this space the stochastic process defined by the evaluation map  $W(t, \omega) = \omega(t)$  produces the statistics of Brownian motion. The filtration  $\mathcal{F}_t$  is then defined by the (completion) of the  $\sigma$ -algebra generated by  $W(s)$  for  $s \leq t$ . See e.g. [KS91] for further details.

We consider a stochastic forcing of the form<sup>10</sup>

$$\sigma_\theta dW := \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0,1\}}} \alpha_k^l \sigma_k^l dW^{k,l}. \quad (2.13)$$

The index  $\theta$  of  $\sigma_\theta$  indicates that  $\sigma_\theta$  attains nontrivial values only in the second component, in other words only the  $\theta$  component is directly forced.

**Remark 2.1.**

- (i) For the rest of the manuscript we fix the physical constants  $\nu_1, \nu_2 > 0$ ,  $g \neq 0$  and non-zero noise coefficients  $(\alpha_k^l)_{k \in \mathcal{Z}, l \in \{0,1\}}$ . Below constants  $C, C_0, C^*$  etc. may change line by line and they implicitly depend on  $\nu_1, \nu_2 > 0$ ,  $g \neq 0$ , and  $\alpha_k^l \neq 0$ . All other parameter dependencies are indicated explicitly.
- (ii) Here and below  $d$  will always denote the number of driving Brownian motions  $2 \cdot |\mathcal{Z}|$ . Since the use of double indices in e.g. (2.12)–(2.13) can become notationally involved we sometimes simply denote the basis for  $\mathbb{R}^d$  as  $\{e_1 \dots e_d\}$ .

With these preliminaries in hand, the equations (2.1)–(2.2) may be written as an abstract stochastic evolution equation on  $H$

$$dU + (AU + B(U))dt = GUdt + \sigma_\theta dW, \quad U(0) = U_0, \quad (2.14)$$

where  $U_0 \in H$ . We say that  $U = U(t, U_0)$  is a solution of (2.14) if it is  $\mathcal{F}_t$ -adapted,

$$U \in C([0, \infty); H) \cap L_{loc}^2([0, \infty); H^1) \quad a.s., \quad (2.15)$$

and  $U$  satisfies (2.14) in the mild sense, that is,

$$U(t) = e^{-At}U_0 - \int_0^t e^{-A(t-s)}(B(U(s)) - GU(s))ds + \int_0^t e^{-A(t-s)}\sigma_\theta dW(s). \quad (2.16)$$

Note that  $B, G : H^1 \rightarrow H$  and the semigroup  $e^{-tA}$  maps  $H$  to  $H^2 \hookrightarrow H^1$ .

The following proposition summarizes the basic well-posedness, regularity, and smoothness with respect to data for (2.14).

**Proposition 2.2** (Existence/Uniqueness/Continuous Dependence on Data). *Fix  $\nu_1, \nu_2 > 0$ ,  $g \in \mathbb{R}$ , and a stochastic basis  $\mathcal{S}$ . Given any  $U_0 \in H$ , there exists a unique solution  $U : [0, \infty) \times \Omega \rightarrow H$  of (2.16) which is an  $\mathcal{F}_t$ -adapted process on  $H$  with the regularity (2.15).*

*For any  $t \geq 0$  and any realization of the noise  $W(\cdot, \omega)$ , the map  $U_0 \mapsto U(t, U_0)$  is Fréchet differentiable on  $H$ . On the other hand, for every fixed  $U_0 \in H$  and  $t \geq 0$ ,  $W \mapsto U(t, W)$  is Fréchet differentiable from  $C((0, t), \mathbb{R}^{2|\mathcal{Z}|})$  to  $H$ . Moreover,  $U$  is (spatially) smooth for all positive times, that is, for any  $t_0 > 0$  and any  $m \geq 0$ ,*

$$U \in C([t_0, \infty); H^m) \quad a.s.$$

Finally,  $U$  satisfy certain moment bounds as detailed below in Lemma A.1.

The well-posedness theory for (2.14) is standard and follows along the line of classical proofs for the stochastic 2D Navier-Stokes equations, particularly since we are considering the case of a spatially smooth, additive noise. See e.g. [KS12] for a detailed recent account close to our setting and e.g. [Roz90, DPZ92,

<sup>10</sup>Although we assume that  $|\mathcal{Z}|$  is finite, our results also hold true if we consider random stirring in all of the Fourier directions ( $\mathcal{Z} = \mathbb{Z}_+^2$ ) provided that we posit sufficient decay in the  $\alpha_k^l$ 's so that the resulting solutions are sufficiently spatially smooth. Note that it is for small values of  $|\mathcal{Z}|$  (and in particular when  $|\mathcal{Z}|$  is independent of  $\nu_1, \nu_2, g$ ) that makes proof of ergodicity for (2.1)–(2.2) difficult. As such we focus on our attention on the ‘smallest’ possible  $\mathcal{Z}$ .

PR07, HM11] and [CF88, Tem01, MB02] for more general background on the theory of infinite dimensional stochastic systems and mathematical fluids respectively. Some moment estimates are less standard, but can be found in a similar setting to ours in [HM06, Deb11, KS12]. For the purpose of completeness and in order to carefully track dependencies on parameters we include some details in Appendix A.1.

For simplicity we set

$$F(U) := -AU - B(U) + GU \quad (2.17)$$

and rewrite (2.14) in a more compact notation

$$dU = F(U)dt + \sigma dW, \quad U(0) = U_0,$$

which is particularly useful for the Hörmander type Lie bracket computations in Section 5.

## 2.1 The Markovian Framework and the Main Result

With the basic well-posedness of (2.14) in hand we next describe its associated Markovian semigroup. Let us first recall some further spaces. Denote by  $M_b(H)$  and  $C_b(H)$  respectively the spaces of bounded measurable and bounded continuous real valued functions on  $H$  equipped with the supremum norm. We also define

$$\mathcal{O}_\varsigma := \{\Phi \in C^1(H) : \|\Phi\|_\varsigma < \infty\}, \quad \text{where} \quad \|\Phi\|_\varsigma := \sup_{U_0 \in H} (\exp(-\varsigma\|U_0\|)(|\Phi(U_0)| + \|\nabla\Phi(U_0)\|)), \quad (2.18)$$

for any  $\varsigma > 0$ . Finally take  $Pr(H)$  to be the collection of Borelian probability measures on  $H$ .

Fix  $U = U(t, U_0) = U(t, U_0, W)$  and define the *Markovian transition function* associated to (2.14) by

$$P_t(U_0, E) = \mathbb{P}(U(t, U_0) \in E) \quad \text{for any } U_0 \in H, E \in \mathcal{B}(H), t \geq 0, \quad (2.19)$$

where  $\mathbb{P}$  is defined relative to the fixed stochastic basis (cf. Proposition 2.2) and  $\mathcal{B}(H)$  is the collection of Borel sets on  $H$ . We define the Markov semigroup  $\{P_t\}_{t \geq 0}$  with  $P_t : M_b(H) \rightarrow M_b(H)$  associated to (2.14) by

$$P_t\Phi(U_0) := \mathbb{E}\Phi(U(t, U_0)) = \int_H \Phi(U)P_t(U_0, dU) \quad \text{for any } \Phi \in M_b(H), t \geq 0. \quad (2.20)$$

By Proposition 2.2 and the dominated convergence theorem,  $\{P_t\}_{t \geq 0}$  is *Feller* meaning that  $P_t : C_b(H) \rightarrow C_b(H)$  for every  $t \geq 0$ . The dual operator  $P_t^*$  of  $P_t$ , which maps  $Pr(H)$  to itself, is given by

$$P_t^*\mu(A) := \int_H P_t(U_0, A)d\mu(U_0), \quad (2.21)$$

over  $\mu \in Pr(H)$ . Recall that  $\mu \in Pr(H)$  is an *invariant measure* if it is a fixed point of  $P_t^*$  for every  $t \geq 0$ , that is,  $P_t^*\mu = \mu$ . Such a measure  $\mu$  is *ergodic* if  $P_t$  is an ergodic map relative to  $\mu$  for every  $t \geq 0$ . In other words  $P_t\chi_A = \chi_A$ ,  $\mu$  a.e. implies  $\mu(A) \in \{0, 1\}$ .

We now formulate our main result, which asserts that statistically invariant states of (2.1)–(2.2) are unique and have strong attraction properties.

**Theorem 2.3.** *If  $Z = \{(1, 0), (0, 1)\}$ , then there exists a unique invariant measure  $\mu_*$  associated to (2.14) and for each  $t \geq 0$  the map  $P_t$  is ergodic relative to  $\mu_*$ . Moreover there exists a constant  $\varsigma^*$  such that  $\mu_*$  satisfies for each  $\varsigma \in (0, \varsigma^*)$*

(i) (*Mixing*) *There is  $\gamma = \gamma(\varsigma) > 0$  and  $C = C(\varsigma)$  such that*

$$\left| \mathbb{E}\Phi(U(t, U_0)) - \int_H \Phi(\bar{U})d\mu_*(\bar{U}) \right| \leq C \exp(-\gamma t + \varsigma\|U_0\|)\|\Phi\|_\eta. \quad (2.22)$$

for any  $\Phi \in \mathcal{O}_\varsigma$ ,  $U_0 \in H$  and any  $t \geq 0$ .

(ii) (Weak law of large numbers) For any  $\Phi \in \mathcal{O}_\zeta$  and any  $U_0 \in H$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \Phi(U(t, U_0)) dt = \int_H \Phi(\bar{U}) d\mu_*(\bar{U}) =: m_\Phi, \quad \text{in probability.} \quad (2.23)$$

(iii) (Central limit theorem) For every  $\Phi \in \mathcal{B}_\eta$  and every  $U_0 \in H$ ,

$$\lim_{T \rightarrow \infty} \mathbb{P} \left( \frac{1}{\sqrt{T}} \int_0^T (\Phi(U(t, U_0)) - m_\Phi) dt < \xi \right) = \mathcal{X}(\xi) \quad (2.24)$$

for any  $\xi \in \mathbb{R}$  where  $\mathcal{X}$  is the distribution function of a normal random variable with zero mean and variance equal to

$$\lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left( \int_0^T (\Phi(U(t, U_0)) - m_\Phi) dt \right)^2.$$

#### Remark 2.4.

- (i) We take  $\mathcal{Z} = \{(0, 1), (1, 0)\}$  in the statement of Theorem 2.3 for simplicity and clarity of exposition only. There are many other choices of finite  $\mathcal{Z}$  that imply our main results. For example, our approach applies trivially to any  $\mathcal{Z}$  with  $\{(0, 1), (1, 0)\} \subset \mathcal{Z}$ . More generally, using combinatorial arguments as in [HM06] and straightforward modifications of our proofs, one can show that our results hold if the integer linear combinations of elements in  $\mathcal{Z}$  generate  $\mathbb{Z}^2$ , for an easily verifiable criterion see [HM06, Remark 2.2].
- (ii) The most interesting case for the study of ergodicity for the Boussinesq system, is to consider stochastic forcing acting through the temperature equation only. Indeed, if the random stirring acts exclusively through the momentum equation (i.e. in (2.1)), then the temperature  $\theta$ , which is advected by the flow in (2.2), decays exponentially. Thus, in this case, conditions on the configuration of the forcing and the associated analysis is very close to [HM06]. On the other hand if the noise acts in both equations the proof is similar to ours and in many ways easier.<sup>11</sup> We remark that the case when the forcing is non-degenerate (nontrivial on all Fourier modes) in both the momentum and temperature equations, was addressed in [LW04] via coupling methods closely following the approach in [EMS01]. See also [Fer97].
- (iii) We emphasize that, when the random perturbation acts in the temperature equation only, this leads to a different geometric criteria for the noise structure compared to [HM06]. To see this difference at a heuristic level we write (2.1)–(2.2) in the Fourier representation:

$$\frac{d\omega_k}{dt} + \nu|k|^2\omega_k + \sum_{l+m=k} \langle l, m^\perp \rangle \left( \frac{1}{|l|^2} - \frac{1}{|m|^2} \right) \omega_l \omega_m = -ig \cdot k \theta_k, \quad (2.25)$$

$$d\theta_k + \left( \eta|k|^2\theta_k - \sum_{l+m=k} \frac{\langle l, m^\perp \rangle}{|m|^2} \theta_l \omega_m \right) dt = \mathbb{1}_{k \in \mathcal{Z}} dW^k. \quad (2.26)$$

Observe that at first only Fourier modes of  $\theta$  in  $\mathcal{Z}$  are excited. Then, through the buoyancy term on the right hand side of (2.25), the Fourier modes of  $\omega$  in  $\mathcal{Z}$  become excited. This is a purely formal argument as at the same time many modes in  $\theta$  become excited. If all elements of  $\mathcal{Z}$  have the same norm, such an excitation is not sufficient for the nonlinearity in (2.25), acting on its own, to propagate

<sup>11</sup>Indeed one observes that  $[[F(U), \sigma_k^l], \psi_{k'}^{l'}] \propto B(\psi_{k'}^{l'}, \sigma_k^l) \neq 0$ ,  $[[F(U), \psi_k^l], \psi_{k'}^{l'}] \propto B(\psi_{k'}^{l'}, \psi_k^l) \neq 0$  which allows us to generate new pure directions in both the temperature and vorticity components of the phase space; compare with Figure 5.1 in Section 5 and the discussion in Section 1.3.

the noise to higher Fourier modes.<sup>12</sup> However, excitation in the Fourier modes of  $\omega$  in  $\mathcal{Z}$  propagates to higher Fourier modes in  $\theta$  via the nonlinearity of (2.26); here the norm restriction is clearly absent. Thus, there is an additional mixing mechanism in the Boussinesq system compared to the Navier–Stokes equation.

(iv) The class of functions for which the mixing condition (2.22) holds is slightly restrictive. While it does allow for observables like individual Fourier coefficients of the solution or the total energy of solutions, a further analysis is required to extend to  $\Phi$ 's that involve pointwise spatial observations of the flow, for example ‘structure functions’. We leave these questions for future work.

## 2.2 Existence and Uniqueness of Invariant Measures and the Asymptotic Smoothing of the Markov Semigroup

Following [HM06] (and cf. [HM08, KW12]), we explain how the proof of Theorem 2.3 can be essentially reduced to establishing a time asymptotic gradient estimate, (2.29), on the Markov semigroup.

By the Krylov-Bogoliubov averaging method [KB37] it is immediate to prove the existence of an invariant measure in the present setting. Indeed, fix any  $U_0 \in H$ ,  $T > 0$  and define the probability measures  $\mu_T \in \text{Pr}(H)$  as

$$\mu_T(A) = \frac{1}{T} \int_0^T \mathbb{P}(U(t, U_0) \in A) dt \quad (A \in \mathcal{B}(H)).$$

From (A.9) and (A.10) it follows that

$$\min\{\nu_1, \nu_2\} \frac{1}{T} \mathbb{E} \int_0^T \|U\|_{H^1}^2 dt \leq \frac{\|U_0\|^2}{T} + \|\sigma_\theta\|^2. \quad (2.27)$$

Thus for any  $R > 0$  the set  $B_{H^1}(R) := \{U \in H : \|U\|_{H^1} \leq R\}$  is compact in  $H$  and by the Markov inequality

$$\begin{aligned} \mu_T(B_{H^1}(0, R)) &= \frac{1}{T} \int_0^T \mathbb{P}(\|U(t, U_0)\|_{H^1} \leq R) dt \geq 1 - \frac{1}{TR^2} \mathbb{E} \int_0^T \|U\|_{H^1}^2 dt \\ &\geq 1 - \frac{1}{R^2 \cdot \min\{\nu_1, \nu_2\}} \left( \frac{\|U_0\|^2}{T} + \|\sigma_\theta\|^2 \right), \end{aligned}$$

and therefore  $\{\mu_T\}_{T \geq 1}$  is tight, and hence weakly compact. Making use of the Feller property it then follows that any weak limit of this sequence is an invariant measure of (2.14).<sup>13</sup>

We now turn to the question of uniqueness which in contrast to existence is highly non-trivial. The classical theoretical foundation to our approach is the Doob-Khasminskii theorem, see [DPZ96]. While this approach has been fruitful for stochastic perturbations acting on all of the Fourier modes (see e.g. [DPZ96]); it requires an instantaneous (or at least finite time) smoothing of  $P_t$ , known as the *strong Feller property*.<sup>14</sup> This property is not expected to hold in the current hypo-elliptic setting.

In recent works [HM06, HM08, HM11] it has been shown that the strong Feller property can be replaced by a much weaker notion. The following theorem from [HM06] is the starting point for all of the work that follows below.

**Theorem 2.5** (Hairer–Mattingly [HM06]). *Let  $\{P_t\}_{t \geq 0}$  be a Feller Markov semigroup on a Hilbert space  $H$  and assume that the set of invariant measures  $\mathcal{I}$  of  $\{P_t\}_{t \geq 0}$  is compact. Suppose that*

<sup>12</sup>In fact, this structure in the nonlinearity is the reason for the additional condition that two modes of different length need to be stochastically forced in [HM06]

<sup>13</sup>Note that the Feller property and the bound (2.27) also show that the set of invariant measures  $\mathcal{I}$  for (2.14) is a compact, convex set. Since the extremal points of  $\mathcal{I}$  are ergodic invariant measures for (2.14), we therefore infer the existence of an ergodic invariant measure for (2.14). This also shows that if the invariant measure is unique, it is necessarily ergodic. See e.g. [DPZ96] for further details.

<sup>14</sup>More precisely,  $P_t$  is said to be strong Feller for some  $t > 0$  if  $P_t : M_b(H) \rightarrow C_b(H)$ .

(i) the semigroup  $\{P_t\}_{t \geq 0}$  is weakly irreducible namely there exists  $U_0 \in H$  such that  $\mu(B(\epsilon, U_0)) > 0$  for every  $\epsilon > 0$  and every  $\mu \in \text{Pr}(H)$  which is invariant under  $P_t^*$ . In other words there exists a point common to the support of every invariant measure.

(ii) There exists a non-decreasing sequence  $\{t_n\}_{n \geq 0}$  and a sequence  $\{\delta_n\}_{n \geq 0}$ , with  $\delta_n \rightarrow 0$  such that

$$\|\nabla P_{t_n} \Phi(U_0)\| \leq C (\|\Phi\|_\infty + \delta_n \|\nabla \Phi\|_\infty), \quad (2.28)$$

for every  $\Phi \in C_b^1(H)$  and where the constant  $C$  may depend on  $\|U_0\|$  (but not on  $\Phi$ ).<sup>15</sup>

Then the collection of invariant measures  $\mathcal{I}$  contains at most one element.

In our situation the proof of (i) is more or less standard and follows precisely as in [EM01, CGHV13]. The main difficulty is to establish the asymptotic smoothing property (ii) of Theorem 2.5. We prove the following stronger version of (2.28), which is also useful for other parts of Theorem 2.3 the proof of the mixing (2.22) and pathwise convergence properties (2.23), (2.24). We recall the convention in Remark 2.1.

**Proposition 2.6.** *For every  $\eta, \gamma_0 > 0$  and every  $U_0 \in H$ , the Markov semigroup  $\{P_t\}_{t \geq 0}$  defined by (2.20) satisfies the estimate*

$$\|\nabla P_t \Phi(U_0)\| \leq C \exp(\eta \|U_0\|^2) \left( \sqrt{P_t(|\Phi|^2)(U_0)} + e^{-\gamma_0 t} \sqrt{P_t(\|\nabla \Phi\|^2)(U_0)} \right), \quad (2.29)$$

for every  $t \geq 0$  and  $\Phi \in C_b^1(H)$ , where  $C = C(\eta, \gamma_0)$  is independent of  $t$  and  $\Phi$ .

With Proposition 2.6 the uniqueness of the invariant measure follows immediately from Theorem 2.5. With slightly more work we can also use (2.29) to establish the attraction properties (i)–(iii) in Theorem 2.3. Since this mainly requires the introduction of some further abstract machinery from [HM08, KW12] we postpone the rest of the proof of Theorem 2.3 to final part of Section 7.

### 3 Gradient Estimates for the Markov Semigroup

In this section we explain how the estimate on  $\nabla P_t \Phi$  in (2.29) can be translated to a control problem through the Malliavin integration by parts formula. This lead us to study the so called Malliavin covariance matrix  $\mathcal{M}$ , which links the existence of a desirable control to the properties of successive Hörmander-type Lie brackets of vector fields (on  $H$ ) associated to (2.14). Suitable spectral bounds for  $\mathcal{M}$  are given in Proposition 3.4 and we conclude this section by explaining how these bounds can be used in conjunction with a control built around  $\mathcal{M}$  to complete the proof of Proposition 2.6.

The involved proof of Proposition 3.4 is delayed for Sections 4, 5, 6 below. As we already noted in the introduction, although the statement of Proposition 3.4 looks similar to corresponding results in [MP06, HM06, BM07, HM11], the proof is significantly different due to the particular nonlinear structure of (2.1)–(2.2). As such Proposition 3.4 constitutes the main mathematical novelty of this work.

#### 3.1 Deriving the Control Problem

Let  $U = U(\cdot, U_0)$  be the solution of (2.14) and assume the convention from Remark 2.1 where we let  $d := 2 \cdot |\mathcal{Z}|$ . Then for any  $\Phi \in C_b^1(H)$ ,  $\xi \in H$  we have<sup>16</sup>

$$\nabla P_t \Phi(U_0) \cdot \xi = \mathbb{E}(\nabla \Phi(U(t, U_0)) \cdot \mathcal{J}_{0,t} \xi), \quad t \geq 0, \quad (3.1)$$

<sup>15</sup> Actually, the ‘gradient estimate’ (2.28) is a sufficient condition for a more general notion of infinite time smoothing referred to as the *asymptotically strong Feller property*. A precise topological definition using the Kantorovich-Wasserstein distance can be found in [HM06].

<sup>16</sup>For differentiability of  $U_0 \mapsto U(t, U_0)$  see Proposition 2.2 and [HM11, Section 3.3]

where for  $0 \leq s \leq t$ ,  $\mathcal{J}_{s,t}\xi$  denotes the unique solution of

$$\partial_t \rho + A\rho + \nabla B(U)\rho = G\rho, \quad \rho(s) = \xi, \quad (3.2)$$

and  $\nabla B(U)\rho := B(U, \rho) + B(\rho, U)$ .

The crucial step in establishing (2.29) is to ‘approximately remove’ the gradient from  $\Phi$  in (3.1). As such we seek to (approximately) identify  $\mathcal{J}_{0,t}\xi$  with a Malliavin derivative of some suitable random process and integrate by parts, in the Malliavin sense. In Appendix A.2 we recall some elements of this calculus which are used throughout this section. For an extended treatment of the Malliavin theory we refer to e.g. [Bel87, Mal97, Nua09, Nua06].

Recall, that in our situation the Malliavin derivative,  $\mathfrak{D} : L^2(\Omega, H) \rightarrow L^2(\Omega; L^2(0, t; \mathbb{R}^d) \otimes H)$  satisfies

$$\langle \mathfrak{D}U, v \rangle_{L^2([0, T], \mathbb{R}^d)} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left( U(T, U_0, W + \epsilon \int_0^T v ds) - U(T, U_0, W) \right).$$

We may infer that for  $v \in L^2(\Omega; L^2([0, T]; \mathbb{R}^d))$  one has (cf. [HM11])

$$\langle \mathfrak{D}U, v \rangle_{L^2([0, T], \mathbb{R}^d)} = \int_0^T \mathcal{J}_{s, T} \sigma_\theta v(s) ds, \quad (3.3)$$

and hence, by the Riesz representation theorem,

$$\mathfrak{D}_s^j U(T) = \mathcal{J}_{s, T} \sigma_\theta e_j \quad \text{for any } s \leq T, j = 1, \dots, d, \quad (3.4)$$

where the linearization  $\mathcal{J}_{r,t}\xi$  is the solution of (3.2),  $\sigma_\theta$  is given by (2.12), and  $\{e_j\}_{j=1, \dots, d}$  is the standard basis of  $\mathbb{R}^d$ .<sup>17</sup> Here and below, we adopt the standard notation  $\mathfrak{D}_s^j F := (\mathfrak{D}F)^j(s)$ , that is,  $\mathfrak{D}_s^j F$  is the  $j$ th component of  $\mathfrak{D}F$  evaluated at time  $s$ .

Motivated by (3.3), we define the random operator  $\mathcal{A}_{s,t} : L^2([s, t]; \mathbb{R}^d) \rightarrow H$  by

$$\mathcal{A}_{s,t} v := \int_s^t \mathcal{J}_{r,t} \sigma_\theta v(r) dr. \quad (3.5)$$

Notice that, by the Duhamel formula, for any  $0 \leq s < t$  the function  $\rho(t) := \mathcal{A}_{s,t} v$  satisfies

$$\partial_t \rho + A\rho + \nabla B(U)\rho = G\rho + \sigma_\theta v, \quad \rho(s) = 0.$$

With these preliminaries we now continue the computation started in (3.1). Using the Malliavin chain rule and integration by parts formula, as recalled in (A.31), (A.34), we infer that for any  $t \geq 0$  and any suitable (Skorokhod integrable)  $v \in L^2(\Omega \times [0, t], \mathbb{R}^d)$ <sup>18</sup>

$$\begin{aligned} \nabla P_t \Phi(U_0) \cdot \xi &= \mathbb{E}(\nabla \Phi(U) \cdot (\mathcal{A}_{0,t} v + \mathcal{J}_{0,t} \xi - \mathcal{A}_{0,t} v)) \\ &= \mathbb{E}(\nabla \Phi(U) \cdot \langle \mathfrak{D}U, v \rangle) + \mathbb{E}(\nabla \Phi(U) \cdot (\mathcal{J}_{0,t} \xi - \mathcal{A}_{0,t} v)) \\ &= \mathbb{E}(\langle \mathfrak{D} \Phi(U), v \rangle_{L^2([0, t], \mathbb{R}^d)}) + \mathbb{E}(\nabla \Phi(U) \cdot (\mathcal{J}_{0,t} \xi - \mathcal{A}_{0,t} v)) \\ &= \mathbb{E} \left( \Phi(U) \int_0^t v \cdot dW \right) + \mathbb{E}(\nabla \Phi(U) \cdot (\mathcal{J}_{0,t} \xi - \mathcal{A}_{0,t} v)), \end{aligned} \quad (3.6)$$

<sup>17</sup>Using the Malliavin chain rule and (A.30) we apply  $\mathfrak{D}$  to (2.14) and observe, at least formally that

$$\mathfrak{D}_s^j U(T) + \int_s^T (A \mathfrak{D}_s^j U(r) + \nabla B(U(r)) \mathfrak{D}_s^j U(r) - G \mathfrak{D}_s^j U(r)) dr = \sigma_\theta e_j,$$

for  $s < T$  and  $j = 1, \dots, d$ .

<sup>18</sup>Note that for non-adapted  $v$ ,  $\int_0^t v \cdot dW$  in (3.6) is understood as a stochastic integral in a generalized sense; see Appendix A.2 below or e.g. [Nua06] for further details.

where for brevity we have denoted  $U = U(t, U_0)$ . Observe that  $\rho(t) := \mathcal{J}_{0,t}\xi - \mathcal{A}_{0,t}v$  satisfies:

$$\partial_t \rho + A\rho + \nabla B(U)\rho = G\rho - \sigma_\theta v, \quad \rho(0) = \xi. \quad (3.7)$$

Directly from (3.6) we obtain (recalling the notation (2.20))

$$|\nabla P_t \Phi(U_0) \cdot \xi| \leq \left( \mathbb{E} \left| \int_0^t v \cdot dW \right|^2 \right)^{1/2} \sqrt{P_t |\Phi|^2(U_0)} + (\mathbb{E} \|\rho(t, \xi, v)\|^2)^{1/2} \sqrt{P_t \|\nabla \Phi\|^2(U_0)},$$

for any  $t \geq 0$  and  $\xi \in H$ . As such, (2.29) has been translated to the following control problem: For each  $\gamma_0, \eta > 0$  and each unit length element  $\xi \in H$ , find a (locally Skorokhod integrable)  $v = v(\xi) \in L^2(\Omega; L^2_{loc}([0, \infty), \mathbb{R}^d))$  such that

$$\sup_{\|\xi\|=1} \mathbb{E} \|\rho(t, \xi, v)\|^2 \leq C e^{-\gamma_0 t} \exp(\eta \|U_0\|^2) \quad (3.8)$$

and

$$\sup_{\|\xi\|=1, t \geq 0} \mathbb{E} \left| \int_0^t v \cdot dW \right|^2 \leq C \exp(\eta \|U_0\|^2), \quad (3.9)$$

where  $C = C(\gamma_0, \eta)$  is independent of  $t$ .

### 3.2 Choosing the Control

In the case when sufficiently many directions in Fourier space are forced or if  $\nu_1, \nu_2$  are sufficiently large, we can choose the control  $v$  in (3.7) from the use of determining modes. See [FP67] or more recently [FMRT01]. Specifically, we might choose  $\sigma_\theta v = \lambda P_N \rho$  in (3.7), where  $P_N$  is the projection onto the subspace  $H_N$  defined in (2.8) and  $\lambda > 0$  is a sufficiently large constant. The construction of such a control relies on the assumption that all of the modes with wave-numbers less  $N$  are directly forced, that is, they are in the range of  $\sigma_\theta$ . We remark, moreover, that  $N$  is a function of  $\nu_1, \nu_2 > 0$  and it diverges to infinity as  $\nu_1$  and  $\nu_2$  approach zero. The idea of using determining modes in the context of the ergodic theory of mathematical fluids equations (and other nonlinear SPDEs) has played a central role in a number of recent works. See e.g. [Mat02, Mat03, HM06, KS12, CGHV13].

If the range of  $\sigma_\theta$  is not sufficiently large, we are hindered by the fact that we cannot ‘directly control’ all the low (unstable) modes that are not dissipated by the diffusion. A different approach inspired by the finite dimensional case (cf. [Hai11]) would be to try to find an exact control. To achieve this we would seek for each  $\xi \in H$  a corresponding  $v$  such  $\mathcal{J}_{0,t}\xi = \mathcal{A}_{0,t}v$ . With the ansatz that  $v$  has the form  $v = \mathcal{A}_{0,t}^* \eta$  for some  $\eta \in H$  (where  $\mathcal{A}_{0,t}^*$  the adjoint of  $\mathcal{A}_{0,t}$  defined below in (3.11)) we could choose the control

$$v := \mathcal{A}_{0,t}^* (\mathcal{A}_{0,t} \mathcal{A}_{0,t}^*)^{-1} \mathcal{J}_{0,t} \xi. \quad (3.10)$$

The object  $\mathcal{M}_{0,t} := \mathcal{A}_{0,t} \mathcal{A}_{0,t}^*$ , referred to as the Malliavin covariance matrix, plays an important role in the theory of stochastic analysis. If one can establish the invertibility of  $\mathcal{M}_{0,t}$  one finds an exact control in (3.7) and with suitable bounds on  $v$  one shows that the Markov semigroup is smoothing in finite time (i.e. it is strongly Feller).

Sufficient conditions for the invertibility of the Malliavin matrix are well understood for finite dimensional problems. However, this invertibility is much harder to deduce in infinite dimensions and may not hold in general.<sup>19</sup> Instead, following the insights in [HM06], we now combine the strategy identified in (3.10) with the use of determining modes [FP67], and use a Tikhonov regularization of the Malliavin matrix to construct a control  $v$ , and corresponding  $\rho$ , which satisfy (3.8)–(3.9).

<sup>19</sup>Although  $\mathcal{M}_{0,t}$  is a linear, non-negative definite, self adjoint operator on a Hilbert space, which can be shown to be non-degenerate it is difficult to quantify the range of  $\mathcal{M}_{0,t}$ , or equivalently we are unable to characterize the domain of  $(\mathcal{M}_{0,t})^{-1}$ .

To make this more precise we first define several random operators. For any  $s < t$ , let  $\mathcal{A}_{s,t}^* : H \rightarrow L^2([s, t]; \mathbb{R}^d)$  be the adjoint of  $\mathcal{A}_{s,t}$  defined in (3.5). We observe that

$$(\mathcal{A}_{s,t}^* \xi)(r) = \sigma_\theta^* \mathcal{J}_{r,t}^* \xi =: \sigma_\theta^* \mathcal{K}_{r,t} \xi \quad \text{for any } \xi \in H, r \in [s, t], \quad (3.11)$$

where  $\sigma_\theta^* : H \rightarrow \mathbb{R}^d$  is the adjoint of  $\sigma_\theta$  defined in (2.12). Here, for  $s < t$ ,  $\mathcal{K}_{s,t} \xi = \mathcal{J}_{s,t}^* \xi$  is the solution of the ‘backward’ system (see [HM11])

$$\partial_s \rho^* = A \rho^* + (\nabla B(U(s)))^* \rho^* - G^* \rho^* = -(\nabla F(U))^* \rho^*, \quad \rho^*(t) = \xi. \quad (3.12)$$

We then define the Malliavin Matrix

$$\mathcal{M}_{s,t} := \mathcal{A}_{s,t} \mathcal{A}_{s,t}^* : H \rightarrow H. \quad (3.13)$$

We now build the control  $v$  and derive the associated  $\rho$  in (3.7) using the following iterative construction. Denote by  $v_{s,t}$  the control  $v$  restricted to the time interval  $[s, t]$  and let  $\rho_n := \rho(n)$ . Observe that  $\rho_0 = \xi \in H$  by definition. For each even non-negative integer  $n \in 2\mathbb{N}$ , having determined  $\rho_n$  and  $v_{0,n}$ , we set

$$v_{n,n+1}(r) = (\mathcal{A}_{n,n+1}^* (\mathcal{M}_{n,n+1} + I\beta)^{-1} \mathcal{J}_{n,n+1} \rho_n)(r), \quad v_{n+1,n+2}(r) = 0, \quad (3.14)$$

for  $r \in [n, n+2]$ . If we denote

$$\mathcal{R}_{n,n+1}^\beta := \beta (\mathcal{M}_{n,n+1} + I\beta)^{-1},$$

then using (3.7) we determine  $\rho_{n+2}$  according to

$$\begin{aligned} \rho_{n+2} &:= \mathcal{J}_{n+1,n+2} \rho_{n+1} \\ &= \mathcal{J}_{n+1,n+2} (\mathcal{A}_{n,n+1} \mathcal{A}_{n,n+1}^* (\mathcal{M}_{n,n+1} + I\beta)^{-1} \mathcal{J}_{n,n+1} + \mathcal{J}_{n,n+1}) \rho_n \\ &= \mathcal{J}_{n+1,n+2} \mathcal{R}_{n,n+1}^\beta \mathcal{J}_{n,n+1} \rho_n. \end{aligned} \quad (3.15)$$

### 3.3 Spectral Properties of $\mathcal{M}$ and Decay

Having defined the control  $v$ , and the associated error  $\rho$ , by (3.14) and (3.15) respectively, we now state and prove (modulo a spectral bound on  $\mathcal{M}_{n,n+1}$ , Proposition 3.4) the key decay estimate on  $\rho$ . This estimate is used in Section 3.4 to complete the proof of Proposition 2.6.

**Lemma 3.1.** *For any  $\varpi, \delta > 0$ , there exists  $\beta = \beta(\varpi, \delta) > 0$  which determines  $\rho$  in (3.15) so that for every even  $n \geq 0$ ,*

$$\mathbb{E}(\|\rho_{n+2}\|^8 | \mathcal{F}_n) \leq \delta \exp(\varpi \|U(n)\|^2) \|\rho_n\|^8. \quad (3.16)$$

Moreover, we have the block adapted structure

$$\rho(t), v(t) \text{ are } \mathcal{F}_{\vartheta(t)} \text{ measurable} \quad (3.17)$$

where  $\vartheta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$

$$\vartheta(t) := \begin{cases} [t] & \text{when } [t] \text{ is odd,} \\ t & \text{when } [t] \text{ is even,} \end{cases} \quad (3.18)$$

and  $[t]$  is the smallest integer greater than or equal to  $t$ .

**Remark 3.2.** *We choose the exponent 8 as it is sufficient for the estimates on  $v$ ; similar estimates are valid for any power greater or equal to two.*

To prove Lemma 3.1, we show that the control  $v$ , when it is active, is effective in pushing energy into small scales where it is dissipated by diffusion. To make this precise recall the definition of  $P_N$ ,  $Q_N$ , and  $H_N$  from (2.9) and (2.8). Fix  $N = N(\delta, \varpi)$  specified below and for  $n \in 2\mathbb{N}$  split  $\rho_{n+2} = \rho_{n+2}^L + \rho_{n+2}^H$ , defining

$$\rho_{n+2}^H := \mathcal{J}_{n+1, n+2} Q_N \mathcal{R}_{n, n+1}^\beta \mathcal{J}_{n, n+1} \rho_n, \quad (3.19)$$

$$\rho_{n+2}^L := \mathcal{J}_{n+1, n+2} P_N \mathcal{R}_{n, n+1}^\beta \mathcal{J}_{n, n+1} \rho_n. \quad (3.20)$$

While for large  $N$  estimates for  $\rho_{n+2}^H$  essentially make use of the parabolic character of (3.2), establishing suitable bounds on  $\rho_{n+2}^L$  requires a detailed understanding of the operator  $\mathcal{R}_{n, n+1}^\beta$ . We need to show, for sufficiently small  $\beta$ , that  $\mathcal{R}_{n, n+1}^\beta$  indeed pushes energy into small scales, that is,  $\|P_N \mathcal{R}_{n, n+1}^\beta\|$  is small. The following lemma from [HM11, Lemma 5.14], shows that this in turn follows from uniform positivity of  $\mathcal{M}_{n, n+1}$  on a cone around  $H_N$ .

**Lemma 3.3.** *Suppose that  $\mathcal{M}$  is a positive, self-adjoint linear operator on a separable Hilbert space  $H$ . Suppose that for some  $\alpha, \epsilon > 0$  and  $N \in \mathbb{N}$  we have that*

$$\inf_{\phi \in \mathcal{S}_{\alpha, N}} \frac{\langle \mathcal{M}\phi, \phi \rangle}{\|\phi\|^2} \geq \epsilon, \quad (3.21)$$

where

$$\mathcal{S}_{\alpha, N} := \{\phi \in H : \|P_N \phi\|^2 \geq \alpha \|\phi\|^2\}. \quad (3.22)$$

Then, for any  $\beta > 0$ ,

$$\|P_N \beta (\mathcal{M} + I\beta)^{-1}\| \leq \alpha \vee \sqrt{\beta/\epsilon}.$$

The next proposition shows that (3.21) holds true for  $\mathcal{M}_{n, n+1}$  on a large subset of the probability space.

**Proposition 3.4.** *Let  $\mathcal{M}_{n, n+1}$  be as in (3.13), relative to  $U$  solving (2.14). For any  $N \geq 1$ ,  $\alpha \in (0, 1]$ , and  $\eta > 0$  there exists a positive constant  $\epsilon^* = \epsilon^*(\alpha, \eta, N) > 0$ , such that, for any  $n \geq 0$ , and any  $0 < \epsilon \leq \epsilon^*$*

$$\mathbb{P} \left( \inf_{\phi \in \mathcal{S}_{\alpha, N}} \frac{\langle \mathcal{M}_{n, n+1} \phi, \phi \rangle}{\|\phi\|^2} < \epsilon | \mathcal{F}_n \right) \geq r(\epsilon) \exp(\eta \|U(n)\|^2), \quad (3.23)$$

where  $\mathcal{S}_{\alpha, N}$  is defined by (3.22) and  $r = r(\alpha, N, \eta) : (0, \epsilon^*] \rightarrow (0, \infty)$  is a non-negative, decreasing function with  $r(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . We emphasize that  $r$  is independent of  $n$ .

Proposition 3.4 is a direct consequence of the Markov property and Theorem 4.1 below. While similar results have appeared in previous works, the proof required us to develop a novel approach due to the particular nonlinear structure in (2.1)–(2.2).

We are now prepared to prove Lemma 3.1.

*Proof of Lemma 3.1.* We use the splitting  $\rho_{n+2} = \rho_{n+2}^H + \rho_{n+2}^L$  from (3.19), (3.20). The constant  $N$  appearing in the definition of this splitting is fixed in the estimate on  $\rho_{n+2}^H$  which we address first.

By the positive definiteness of  $\mathcal{M}_{n, n+1}$ , it follows that  $\|\mathcal{R}_{n, n+1}^\beta\| \leq 1$  almost surely, for any  $\beta > 0$ . Then, as  $\rho_n$  is  $\mathcal{F}_n$ -measurable, from (A.13), (A.17), and (A.2) we infer

$$\begin{aligned} \mathbb{E}(\|\rho_{n+2}^H\|^8 | \mathcal{F}_n) &\leq \|\rho_n\|^8 \cdot \mathbb{E}(\mathbb{E}(\|\mathcal{J}_{n+1, n+2} Q_N\|^8 | \mathcal{F}_{n+1}) \cdot \|\mathcal{J}_{n, n+1}\|^8 | \mathcal{F}_n) \\ &\leq \frac{\delta}{2^8} \exp(\varpi \|U(n)\|^2) \|\rho_n\|^8, \end{aligned} \quad (3.24)$$

for appropriate  $N = N(\varpi, \delta)$ . Fix such an  $N$  in (3.19), (3.20). Note that the bound (3.24) holds independently of the value of  $\beta$  appearing in (3.14).

Next, we estimate  $\rho_{n+2}^L$ . By (A.13), we infer

$$\begin{aligned} \mathbb{E}(\|\rho_{n+2}^L\|^8|\mathcal{F}_n) &\leq \|\rho_n\|^8 \cdot (\mathbb{E}(\|\mathcal{J}_{n+1,n+2}\|^{24}|\mathcal{F}_n))^{1/3} \left(\mathbb{E}(\|P_N \mathcal{R}_{n,n+1}^\beta\|^{24}|\mathcal{F}_n)\right)^{1/3} (\mathbb{E}(\|\mathcal{J}_{n,n+1}\|^{24}|\mathcal{F}_n))^{1/3} \\ &\leq C^* \|\rho_n\|^8 \exp(\varpi/2\|U(n)\|^2) \left(\mathbb{E}(\|P_N \mathcal{R}_{n,n+1}^\beta\|^{24}|\mathcal{F}_n)\right)^{1/3}, \end{aligned} \quad (3.25)$$

where  $C^* = C^*(\varpi)$ . For  $N$  fixed by (3.24) and for any  $\epsilon, \alpha > 0$ , and  $n \in 2\mathbb{N}$  consider the set

$$\Omega_{\epsilon, \alpha, N}^n := \left\{ \inf_{\phi \in \mathcal{S}_{\alpha, N}} \frac{\langle \mathcal{M}_{n,n+1} \phi, \phi \rangle}{\|\phi\|^2} \geq \epsilon \right\},$$

where  $\mathcal{S}_{\alpha, N}$  is defined in (3.22). Using Lemma 3.3, Proposition 3.4 with  $\eta$  replaced by  $3\varpi/2$ , and  $\|\mathcal{R}_{n,n+1}^\beta\| \leq 1$  we have

$$\begin{aligned} \mathbb{E} \left( \|P_N \mathcal{R}_{n,n+1}^\beta\|^{24} | \mathcal{F}_n \right) &= \mathbb{E} \left( \|P_N \mathcal{R}_{n,n+1}^\beta\|^{24} \mathbb{1}_{\Omega_{\epsilon, \alpha, N}^n} | \mathcal{F}_n \right) + \mathbb{E} \left( \|P_N \mathcal{R}_{n,n+1}^\beta\|^{24} \mathbb{1}_{(\Omega_{\epsilon, \alpha, N}^n)^c} | \mathcal{F}_n \right) \\ &\leq \left( \alpha \vee \sqrt{\beta/\epsilon} \right)^{24} + \mathbb{P} \left( (\Omega_{\epsilon, \alpha, N}^n)^c | \mathcal{F}_n \right) \\ &\leq \left( \left( \alpha \vee \sqrt{\beta/\epsilon} \right)^{24} + r(\epsilon) \right) \exp(3\varpi/2 \cdot \|U(n)\|^2), \end{aligned} \quad (3.26)$$

which holds for any  $\alpha \in (0, 1]$ ,  $\beta > 0$ , and any  $\epsilon < \epsilon^*(\alpha, \varpi, N) = \epsilon^*(\alpha, \varpi, \delta)$ , and where the function  $r$  is given by (3.23).

Next, we choose  $\alpha, \beta, \epsilon > 0$ , such that  $\epsilon < \epsilon^*(\alpha, \varpi, N)$ , and so that

$$\left( \left( \alpha \vee \sqrt{\beta/\epsilon} \right)^{24} + r(\epsilon) \right)^{1/3} \leq \frac{\delta}{2^8 C^*}, \quad (3.27)$$

where  $C^*$  is the constant from (3.25). First choose a sufficiently small  $\alpha = \alpha(\varpi, \delta) > 0$ . This choice of  $\alpha$  fixes  $\epsilon^* = \epsilon^*(\alpha, \varpi, \delta) = \epsilon^*(\varpi, \delta) > 0$ . We then choose a sufficiently small  $\epsilon < \epsilon^*$ ,  $\epsilon = \epsilon(\varpi, \delta)$  to control the  $r(\epsilon)$  term. Finally, based on this choice of  $\epsilon$ , we determine  $\beta = \beta(\varpi, \delta) > 0$ . Thus we infer that

$$\mathbb{E}(\|\rho_{n+2}^L\|^8|\mathcal{F}_n) \leq \frac{\delta}{2^8} \exp(\varpi\|U(n)\|^2) \|\rho_n\|^8.$$

Combining this bound with (3.24) and  $\|\rho_{n+2}\|^8 \leq 2^7(\|\rho_{n+2}^H\|^8 + \|\rho_{n+2}^L\|^8)$  establishes (3.16).

Pursuing the definitions of  $\rho$  and  $v$  through (3.14), (3.15) the block adapted structure in (3.17) clearly follows by induction and the definitions of the operators  $\mathcal{J}_{n,n+1}$ ,  $\mathcal{A}_{n,n+1}$ ,  $\mathcal{A}_{n,n+1}^*$ . The proof of Lemma 3.1 is thus complete.  $\square$

### 3.4 Proof of Proposition 2.6

This final section is devoted to the proof of Proposition 2.6 following the strategy identified in Sections 3.1–3.2: using the machinery of Malliavin calculus the desired estimate on the Markov semigroup (2.29) has been translated to the control problem (3.8), (3.9), where  $\rho$  is a solution of (3.7) with  $v$  defined by (3.14). Then, in Section 3.3, we derived a one time step decay estimate on  $\rho$  from Lemma 3.1 which we now use as follows.

*Proof of Proposition 2.6.* Fix any  $\eta, \gamma_0 > 0$  and any  $\rho(0) = \xi \in H$  with  $\|\xi\| = 1$ . We successively demonstrate (3.8), (3.9) for  $v$  and  $\rho$  determined by (3.7), (3.14) and (3.15).

First we prove (3.8), using Lemma 3.1. Define  $\Psi : \mathbb{R}^2 \rightarrow \mathbb{R}$  according to

$$\Psi(x, y) = \begin{cases} x/y & \text{for } y \neq 0, \\ 0 & \text{for } y = 0. \end{cases}$$

By (3.17),  $\rho_n$  is  $\mathcal{F}_n$  measurable, and therefore with (3.16) we infer

$$\mathbb{E}(\Psi(\|\rho_{n+2}\|, \|\rho_n\|)^8 \exp(-\varpi \|U(n)\|^2) | \mathcal{F}_n) = \mathbb{1}_{\|\rho_n\| > 0} \frac{\mathbb{E}(\|\rho_{n+2}\| | \mathcal{F}_n)}{\|\rho_n\|^8 \exp(\varpi \|U(n)\|^2)} \leq \delta, \quad (3.28)$$

for every  $\varpi, \delta > 0$ , and each  $n \in 2\mathbb{N}$ . For any  $k \geq 0$ , let

$$X_k := \Psi(\|\rho_{2k+2}\|, \|\rho_{2k}\|)^4 \exp(-\varpi/2 \cdot \|U(2k)\|^2), \quad Y_k := \exp(\varpi/2 \cdot \|U(2k)\|^2).$$

Repeated use of conditional expectation with (3.28) implies

$$\begin{aligned} \mathbb{E}\left(\prod_{k=0}^n X_k^2\right) &= \mathbb{E}\left(\mathbb{E}\left(\prod_{k=0}^n X_k^2 | \mathcal{F}_{2n}\right)\right) = \mathbb{E}\left(\prod_{k=0}^{n-1} X_k^2 \mathbb{E}(X_n^2 | \mathcal{F}_{2n})\right) \\ &\leq \delta \mathbb{E}\left(\prod_{k=0}^{n-1} X_k^2\right) \leq \dots \leq \delta^n. \end{aligned}$$

On the other hand, noting that for each  $n \in \mathbb{N}$ ,  $\mathbb{1}_{\|\rho_n\|=0} \rho(t) = 0$  for any  $t \geq n$  and that  $\|\rho_0\| = 1$  we have

$$\prod_{k=0}^n X_k Y_k = \|\rho_{2n+2}\|^4 \prod_{k=1}^n \mathbb{1}_{\|\rho_{2k}\| > 0} = \|\rho_{2n+2}\|^4.$$

Consequently, if  $\varpi \leq \eta^*$  (see Lemma A.1), Hölder's inequality and (A.4) yield

$$\begin{aligned} \mathbb{E}(\|\rho_{2n+2}\|^4) &\leq \left(\mathbb{E}\prod_{k=0}^n X_k^2\right)^{1/2} \left(\mathbb{E}\prod_{k=0}^n Y_k^2\right)^{1/2} \leq \delta^{n/2} \left(\mathbb{E}\exp\left(\varpi \sum_{k=0}^{2n} \|U(k)\|^2\right)\right)^{1/2} \\ &\leq \delta^{n/2} \exp\left(\frac{\varrho\varpi}{2} \|U_0\|^2\right) \exp(\varkappa n), \end{aligned}$$

where  $\varrho$  and  $\varkappa$  are the constants appearing in (A.4). Setting  $\delta := \exp(-4(\varkappa + \gamma_0))$ ,  $\varpi := \min\{\eta^*, 2\eta/\rho\}$ ,  $\beta = \beta(\eta, \gamma_0)$  as in Lemma 3.1, we obtain that for each integer  $n$

$$\mathbb{E}(\|\rho_{2n}\|^4) \leq C \exp(\eta \|U_0\|^2 - 2n\gamma_0). \quad (3.29)$$

In particular we infer (3.8) for  $t = 2n$ .

Next, observe that, for each  $n \in \mathbb{N}$ ,

$$\rho(t) = \begin{cases} \mathcal{J}_{2n,t} \rho_{2n} - \mathcal{A}_{2n,t} v_{2n,t} & \text{for } t \in [2n, 2n+1), \\ \mathcal{J}_{2n+1,t} \rho_{2n+1} & \text{for } t \in [2n+1, 2n+2]. \end{cases}$$

By (3.14) and (A.19)–(A.21), for any  $t \in [2n, 2n+2]$

$$\|v_{2n,t}\|_{L^2([2n,t]; \mathbb{R}^d)} \leq \|v_{2n,2n+1}\|_{L^2([2n,2n+1]; \mathbb{R}^d)} \leq \beta^{-1/2} \|\mathcal{J}_{2n,2n+1}\| \|\rho_{2n}\|, \quad (3.30)$$

and consequently for any  $t \in [2n, 2n+1)$

$$\begin{aligned} \|\rho(t)\| &\leq \|\mathcal{J}_{2n,t} \rho_{2n}\| + \|\mathcal{A}_{2n,t} v_{2n,t}\| \leq \|\mathcal{J}_{2n,t} \rho_{2n}\| + \|\mathcal{A}_{2n,t}\|_{\mathcal{L}(L^2([2n,t]; \mathbb{R}^d), H)} \|v_{2n,2n+1}\|_{L^2([2n,2n+1]; \mathbb{R}^d)} \\ &\leq C \beta^{-1/2} \left(1 + \sup_{s \in [2n,t]} \|\mathcal{J}_{s,t}\|^2\right) \|\rho_{2n}\|, \end{aligned}$$

and for any  $t \in [2n+1, 2n+2]$

$$\|\rho(t)\| \leq \sup_{s \in [2n+1,t]} \|\mathcal{J}_{s,t}\| \|\rho_{2n+1}\|.$$

Combining these observations, with (A.13) and (3.29) the desired estimate (3.8) now follows.<sup>20</sup>

We turn next to the proof of (3.9). Although  $v$  is not adapted, it follows from Lemma A.7 and (3.14) that  $v_{0,N} \in \mathbb{D}^{1,2}(L^2([0,N]; \mathbb{R}^d))$  for any  $N > 0$ , and we may thus use the generalized Itô isometry, (A.35). By (3.17),  $v(t)$  is  $\mathcal{F}_{\vartheta(t)}$  measurable, where  $\vartheta$  is defined in (3.18), and consequently by (A.30),  $\mathfrak{D}_s v(r) = 0$  if  $s > \vartheta(r)$ . Thus

$$\begin{aligned} \mathbb{E} \left( \int_0^{2N} v \cdot dW \right)^2 &= \mathbb{E} \int_0^{2N} |v(s)|_{\mathbb{R}^d}^2 ds + \mathbb{E} \int_0^{2N} \int_0^{2N} \chi_{s \leq \vartheta(r)} \chi_{r \leq \vartheta(s)} \text{Tr}(\mathfrak{D}_s v(r) \mathfrak{D}_r v(s)) ds dr. \\ &\leq \sum_{n=0}^{N-1} \left( \mathbb{E} \int_{2n}^{2n+1} |v_{2n,2n+1}(s)|_{\mathbb{R}^d}^2 ds + \mathbb{E} \int_{2n}^{2n+1} \int_{2n}^{2n+1} |\mathfrak{D}_s v_{2n,2n+1}(r)|_{\mathbb{R}^{d \times d}}^2 ds dr \right). \end{aligned} \quad (3.31)$$

For the first term in (3.31), we make use of (3.30)

$$\sum_{n=0}^{N-1} \mathbb{E} \int_{2n}^{2n+1} |v_{2n,2n+1}(s)|_{\mathbb{R}^d}^2 ds \leq \beta^{-1/2} \sum_{n=0}^{N-1} (\mathbb{E} \|\mathcal{J}_{2n,2n+1}\|^4 \cdot \mathbb{E} \|\rho_{2n}\|^4)^{1/2} \leq C \exp(\eta \|U_0\|^2) \quad (3.32)$$

for a constant  $C = C(\eta, \gamma_0)$  independent of  $N$ .

To bound the second term in (3.31), we compute an explicit expression for  $\mathfrak{D}_s v$ . By Lemma A.7, each of  $\mathcal{J}_{2n,2n+1}$ ,  $\mathcal{A}_{2n,2n+1}$ ,  $\mathcal{A}_{2n,2n+1}^*$ ,  $\mathcal{M}_{2n,2n+1} + \beta I$ , and  $(\mathcal{M}_{2n,2n+1} + \beta I)^{-1}$  are differentiable in the Malliavin sense and lie in the space  $\mathbb{D}^{1,p}$  for any  $p > 1$  (see (A.29)). It thus follows from (3.15) that  $\rho_{2n} \in \mathbb{D}^{1,p}$  for any  $p > 1$  and any  $n$ . Moreover, recalling that by (3.17),  $\rho_{2n}$  is  $\mathcal{F}_{2n}$  measurable, (A.30) implies that  $\mathfrak{D}_s \rho_{2n} = 0$  for any  $s \geq 2n$ . Then by the Malliavin product rule (see e.g. [PV, Lemma 3.6]) we compute

$$\begin{aligned} \mathfrak{D}_s^j v_{2n,2n+1} &= \mathcal{A}_{2n,2n+1}^* (\mathcal{M}_{2n,2n+1} + I\beta)^{-1} (\mathfrak{D}_s^j \mathcal{J}_{2n,2n+1}) \rho_{2n} + \mathcal{A}_{2n,2n+1}^* (\mathfrak{D}_s^j (\mathcal{M}_{2n,2n+1} + I\beta)^{-1}) \mathcal{J}_{2n,2n+1} \rho_{2n} \\ &\quad + (\mathfrak{D}_s^j \mathcal{A}_{2n,2n+1}^*) (\mathcal{M}_{2n,2n+1} + I\beta)^{-1} \mathcal{J}_{2n,2n+1} \rho_{2n} \end{aligned} \quad (3.33)$$

for any  $j \in \{1, \dots, d\}$  and  $s \in [2n, 2n+1]$ . Moreover, after differentiating the identity  $(\mathcal{M}_{2n,2n+1} + I\beta)^{-1} (\mathcal{M}_{2n,2n+1} + I\beta) = I$  and recalling that  $\mathcal{M}_{2n,2n+1} = \mathcal{A}_{2n,2n+1} \mathcal{A}_{2n,2n+1}^*$  we obtain

$$\begin{aligned} \mathfrak{D}_s^j (\mathcal{M}_{2n,2n+1} + I\beta)^{-1} &= \\ &(\mathcal{M}_{2n,2n+1} + I\beta)^{-1} ((\mathfrak{D}_s^j \mathcal{A}_{2n,2n+1}) \mathcal{A}_{2n,2n+1}^* + \mathcal{A}_{2n,2n+1} (\mathfrak{D}_s^j \mathcal{A}_{2n,2n+1}^*)) (\mathcal{M}_{2n,2n+1} + I\beta)^{-1}. \end{aligned} \quad (3.34)$$

By (3.33), (3.34), and the bounds (A.19)–(A.21), one has for each  $s \in [2n, 2n+1]$

$$\begin{aligned} \|\mathfrak{D}_s^j v_{2n,2n+1}\|_{L^2([2n,2n+1]; \mathbb{R}^d)} &\leq \beta^{-1/2} \|\mathfrak{D}_s^j \mathcal{J}_{2n,2n+1}\| \|\rho_{2n}\| + \beta^{-1} \|\mathfrak{D}_s^j \mathcal{A}_{2n,2n+1}\|_{\mathcal{L}(L^2([2n,2n+1], \mathbb{R}^d), H)} \|\mathcal{J}_{2n,2n+1}\| \|\rho_{2n}\| \\ &\quad + 2\beta^{-1} \|\mathfrak{D}_s^j \mathcal{A}_{2n,2n+1}^*\|_{\mathcal{L}(H, L^2([2n,2n+1], \mathbb{R}^d))} \|\mathcal{J}_{2n,2n+1}\| \|\rho_{2n}\|. \end{aligned} \quad (3.35)$$

Finally, we use (3.35), (A.23)–(A.25), (A.13), and (3.29) to conclude

$$\begin{aligned} \mathbb{E} \sum_{n=0}^{N-1} \int_{2n}^{2n+1} \int_{2n}^{2n+1} |\mathfrak{D}_s v_{n,n+2}(r)|_{\mathbb{R}^{d \times d}}^2 ds dr &= \mathbb{E} \sum_{n=0}^{N-1} \sum_{j=0}^d \int_{2n}^{2n+1} \|\mathfrak{D}_s^j v_{2n,2n+1}\|_{L^2([2n,2n+1]; \mathbb{R}^d)}^2 ds \\ &\leq C \beta^{-2} \exp(\eta/2 \|U_0\|^2) \sum_{n=0}^{\infty} [\mathbb{E} \|\rho_{2n}\|^4]^{1/2} \\ &\leq C \beta^{-2} \exp(\eta \|U_0\|^2), \end{aligned} \quad (3.36)$$

where  $C = C(\eta, \gamma_0)$ . Combining (3.36), (3.32) with (3.31) we now infer (3.9), completing the proof of Proposition 2.6.  $\square$

<sup>20</sup>Note that the constant we obtain for (3.8) grows as  $\beta^{-1/2}$ . This is inconsequential as we obtain a  $\beta^{-2}$  dependence in the constant for (3.9) below; see (3.36).

## 4 Spectral Bounds for the Malliavin Covariance Matrix

In this section we present the main technical result of this work, Theorem 4.1, which yields a probabilistic spectral bound on the Malliavin matrix  $\mathcal{M}_{0,T}$  (see (3.13)). Recall that in Section 2 we established the uniqueness of the invariant measure associated to (2.1)–(2.2) assuming a gradient estimate on the Markov semigroup, (2.29). Then, in Section 3, we established this gradient estimate (2.29) modulo Proposition 3.4, which is a corollary to Theorem 4.1. Hence, we have reduced the proof of the uniqueness in Theorem 2.3 to the proof of Theorem 4.1.

**Theorem 4.1.** *Let  $U_0 \in H$  and define  $\mathcal{M}_{0,T}$  according to (3.13), relative to  $U(\cdot) = U(\cdot, U_0)$  solving (2.14). Fix any  $\alpha \in (0, 1]$ ,  $N \geq 1$ , and  $\eta > 0$ . Then, there exists  $\epsilon^* := \epsilon^*(T, \alpha, \eta, N) > 0$  such that for any  $0 < \epsilon < \epsilon^*$ , there is a measurable set  $\Omega_\epsilon = \Omega_\epsilon(\alpha, N) \subset \Omega$  satisfying*

$$\mathbb{P}(\Omega_\epsilon^c) \leq r(\epsilon) \exp(\eta \|U_0\|^2),$$

where  $r = r(T, \alpha, \eta, N)$  is a nonnegative decreasing function such that  $r(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . On this set  $\Omega_\epsilon$

$$\inf_{\phi \in \mathcal{S}_{\alpha,N}} \frac{\langle \mathcal{M}_{0,T} \phi, \phi \rangle}{\|\phi\|^2} \geq \epsilon,$$

where  $\mathcal{S}_{\alpha,N} = \{\phi \in H : \|P_N \phi\|^2 > \alpha \|\phi\|^2\}$ . In particular,

$$\mathbb{P} \left( \inf_{\phi \in \mathcal{S}_{\alpha,N}} \frac{\langle \mathcal{M}_{0,T} \phi, \phi \rangle}{\|\phi\|^2} \geq \epsilon \right) \geq 1 - r(\epsilon) \exp(\eta \|U_0\|^2). \quad (4.1)$$

**Remark 4.2.**

- (i) An explicit form for  $r(\epsilon)$  is given in (4.12) below. While the decay rate in  $r(\epsilon)$  as  $\epsilon \rightarrow 0$  is much slower than in previous works, as observed in Section 3, it is sufficient for the proof of Proposition 2.6.
- (ii) Using the Markov property in the general form found in e.g. [DPZ92, Theorem 9.12], Theorem 4.1 immediately implies Proposition 3.4.
- (iii) Note that it is not enough to replace  $\mathcal{S}_{\alpha,N}$  by  $H_N \subset \mathcal{S}_{\alpha,N}$  in Theorem 4.1, as we made use of (4.1) for small  $\alpha > 0$  in the proof of Lemma 3.1; see e.g. (3.26), (3.27) above.

Broadly speaking, the proof of Theorem 4.1 involves an ‘iterative proof by contradiction’ following a strategy apparent even in e.g. [Nor86].<sup>21</sup> We show that, on sets of large probability, if the Malliavin matrix  $\mathcal{M}_{0,T}$  has a small eigenvalue, then a certain quadratic form  $\mathcal{Q}$  associated to (2.14) is small when evaluated at the corresponding eigenfunction. We then show that  $\mathcal{Q}$  has a suitable lower bound on  $\mathcal{S}_{\alpha,N}$ . This lower bound may be seen as an infinite-dimensional analogue of the Hörmander bracket condition [Hö67]. By combining these upper and lower bounds we conclude that, with large probability,  $\mathcal{S}_{\alpha,N}$  cannot contain eigenfunctions of  $\mathcal{M}_{0,T}$  corresponding to small eigenvalues.

We refer to proof of Theorem 4.1 as ‘iterative’ because the bounds on  $\mathcal{Q}$  are obtained by an inductive argument which yields a chain of quantitative bounds on certain functionals associated to Hörmander type Lie brackets. Although, we make significant use of a methodology recently developed in [MP06, BM07, HM11] to carry out this process in infinite dimensions, new and interesting difficulties emerge in our situation which reflect the interaction between the nonlinear structure of (2.1)–(2.2) and our choice of stochastic forcing (see Remark 2.4 above). Firstly, with the stochastic forcing in the temperature equation only, it is non-trivial to determine a sequence of suitable (Hörmander type) Lie brackets associated to (2.14); a completely different

<sup>21</sup>In [Nor86] as in [Mal78] the goal is to provide a probabilistic proof of Hörmander’s hypoellipticity theorem [Hö67]. These works link Hörmander’s bracket condition (associated to a hypo-elliptic evolution equation) to the invertibility of the Malliavin covariance matrix. In addition to [Nor86] we refer the interested reader to e.g. [Nua06, Hai11] for further details on the probabilistic approach to hypoellipticity.

methodology must be developed for (2.14) compared to the one used for the stochastic Navier-Stokes equation in [EM01, Rom04, HM06]. Secondly, the vector fields we obtain are  $U$  dependent. This situation forces us to use an infinite-dimensional analogue of the Hörmander bracket condition, which is weaker than the condition appearing in previous works.

The rest of this section is devoted to proof of Theorem 4.1 based on the lower and upper bounds on forms  $\mathcal{Q}$  associated to Hörmander type brackets involving (2.14). The lower and upper bounds are given below as Propositions 4.3 and 4.4, respectively. The detailed computations of the Hörmander brackets are postponed for Sections 5.

## 4.1 Quadratic Forms; Upper and Lower Bounds

Before precisely stating the lower and upper bounds, we briefly recall the origin of the quadratic forms found in these propositions. As explained in the introduction, the ‘admissible Hörmander brackets’ are elements in the sets

$$\mathcal{V}_m := \text{span} \left\{ [E, F], [E, \sigma_k^j], E : k \in \mathcal{Z}, j = \{0, 1\}, E \in \mathcal{V}_{m-1} \right\} \quad (4.2)$$

starting from  $\mathcal{V}_0 = \text{span}\{\sigma_k^j : k \in \mathcal{Z}, j = \{0, 1\}\}$ , where we recall that the Lie brackets are given by  $[E_1, E_2] := \nabla E_2 \cdot E_1 - \nabla E_1 \cdot E_2$ . We show in Section 5, that for each  $N < \tilde{N}$  there exists  $M = M(\tilde{N})$  such that the set

$$\mathfrak{B}_{N, \tilde{N}}(U) := \{\sigma_j^m, \psi_j^m + J_{j,m}^{\tilde{N}}(U) : m \in \{0, 1\}, j \in \mathbb{Z}_+^2, |j| \leq N\} \quad (4.3)$$

is contained in  $\mathcal{V}_M$ . Here, recall that  $\sigma_j^m$  and  $\psi_j^m$  are basis elements for  $H$  defined in (2.6), (2.7) above. The elements  $J_{j,m}^{\tilde{N}}(U)$  are  $U$  dependent ‘error’ terms, which reside in  $H^{\tilde{N}} := Q_{\tilde{N}}H$  (see (2.9)) and satisfy the bound (5.26); the explicit form for these terms is given in (5.23), (5.25) below.

The upshot is that for any (finite)  $M$  we are only able to identify  $U$ -dependent subsets of  $\mathcal{V}_M$ . Hence, we need to introduce a new form of the Hörmander described above in (1.9) (in more general terms) satisfied by

$$\langle \mathcal{Q}_{N, \tilde{N}}(U)\phi, \phi \rangle := \sum_{\tilde{b} \in \mathfrak{B}_{N, \tilde{N}}(U)} |\langle \phi, \tilde{b}(U) \rangle|^2 \quad (4.4)$$

for any  $\tilde{N} > N$ . We are ready to state the lower bound on  $\mathcal{Q}_{N, \tilde{N}}$ .

**Proposition 4.3.** *Fix any any integers  $N \leq \tilde{N}$  and define  $\mathfrak{B}_{N, \tilde{N}}$  by (4.3). Then, for any  $U \in H^2$  and any  $\alpha \in (0, 1]$  it holds that*

$$\langle \mathcal{Q}_{N, \tilde{N}}(U)\phi, \phi \rangle \geq \left( \frac{\alpha}{2} - C^* \frac{N^8}{\tilde{N}} (1 + \|U\|_{H^2}^2) \right) \|\phi\|^2 \quad (4.5)$$

for every  $\phi \in \mathcal{S}_{\alpha, N} = \{\phi \in H : \|P_N \phi\|^2 > \alpha \|\phi\|^2\}$ , where  $C^*$  is a universal constant (see Remark 2.1).

*Proof.* Since  $\{\sigma_j^m, \psi_j^m\}_{m \in \{0, 1\}, |j| \leq N}$  form an orthonormal basis of  $H_N = P_N H$ , we obtain for any  $\phi \in \mathcal{S}_{\alpha, N}$

$$\begin{aligned} \sum_{\tilde{b} \in \mathfrak{B}_{N, \tilde{N}}(U)} |\langle \phi, \tilde{b}(U) \rangle|^2 &= \sum_{\substack{|j| \leq N \\ m \in \{0, 1\}}} |\langle \phi, \sigma_j^m \rangle|^2 + |\langle \phi, \psi_j^m + J_{j,m}^{\tilde{N}}(U) \rangle|^2 \\ &= \|P_N \phi\|^2 + \sum_{\substack{|j| \leq N \\ m \in \{0, 1\}}} \left( 2\langle \phi, \psi_j^m \rangle \langle \phi, J_{j,m}^{\tilde{N}}(U) \rangle + \langle \phi, J_{j,m}^{\tilde{N}}(U) \rangle^2 \right) \\ &\geq \frac{1}{2} \|P_N \phi\|^2 - \sum_{\substack{|j| \leq N \\ m \in \{0, 1\}}} \langle \phi, J_{j,m}^{\tilde{N}}(U) \rangle^2 \geq \frac{\alpha}{2} \|\phi\|^2 - \|\phi\|^2 \sum_{\substack{|j| \leq N \\ m \in \{0, 1\}}} \|J_{j,m}^{\tilde{N}}(U)\|^2. \end{aligned}$$

Now observe that (5.26) with  $s = 1$  yields

$$\|J_{j,m}^{\tilde{N}}(U)\|^2 \leq C \frac{N^8}{\tilde{N}} (1 + \|U\|_{H^2}^2)$$

and the desired bound, (4.5) follows.  $\square$

Next we will state the ‘upper bound’ on  $\mathcal{Q}_{N,\tilde{N}}$ , whose proof is long and technical and is postponed to Section 4. It links the chain of Lie brackets presented in Section 5 (and summarized in Figure 5.1) to quantitative estimates on  $\mathcal{Q}_{N,\tilde{N}}$ .

**Proposition 4.4.** *Fix  $T > 0$ . There are positive constants  $q_i = q_i(T) > 0$ ,  $i = 0, \dots, 5$  such that the following holds. Fix any  $\eta > 0$ , any integer  $\tilde{N} > 0$  and define*

$$\mathcal{E}(\tilde{N}) := \min \left\{ q_0, \left( \frac{q_1}{\tilde{N}} \right)^{q_2^{\tilde{N}}} \right\}. \quad (4.6)$$

Then for every  $\epsilon \in (0, \mathcal{E}(\tilde{N}))$  there is a set  $\Omega_{\epsilon,\tilde{N}}^*$  and a constant  $C = C(\eta, T)$  such that

$$\mathbb{P}((\Omega_{\epsilon,\tilde{N}}^*)^c) \leq C \tilde{N}^{q_3} \exp(\eta \|U_0\|^2) \epsilon^{q_4^{\tilde{N}}} \quad (4.7)$$

and on  $\Omega_{\epsilon,\tilde{N}}^*$  one has (cf. (4.4))

$$\langle \mathcal{M}_{0,T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \langle \mathcal{Q}_{N,\tilde{N}}(U(T)) \phi, \phi \rangle \leq \epsilon^{q_5^{\tilde{N}}} \|\phi\|^2. \quad (4.8)$$

which is valid for  $N < \tilde{N}$  and any  $\phi \in H$ .

**Remark 4.5.** *We may suppose without loss of generality that*

$$q_2 > 1, \quad q_3 > 1, \quad \text{and} \quad q_4 < 1.$$

## 4.2 Proof of Theorem 4.1

Before turning to precise details, let us outline the proof of Theorem 4.1 using Propositions 4.3, 4.4. Observe that, on  $\Omega_{\epsilon,\tilde{N}}^*$  given in Proposition 4.4, we can combine (4.5), (4.8) to infer

$$\langle \mathcal{M}_{0,T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \epsilon^{q_5^{\tilde{N}}} \|\phi\|^2 \geq \left( \frac{\alpha}{2} - C^* \frac{N^8}{\tilde{N}} (1 + \|U(T)\|_{H^2}^2) \right) \|\phi\|^2, \quad (4.9)$$

for any  $\phi \in \mathcal{S}_{\alpha,N}$  and for any  $\epsilon < \mathcal{E}(\tilde{N})$ . If  $\|U(T)\|_{H^2}^2$  was bounded above by a deterministic constant, then we would prove Theorem 4.1 by taking  $\Omega_\epsilon = \Omega_{\epsilon,\tilde{N}}^*$  with  $\tilde{N} = \tilde{N}(\alpha, N)$  sufficiently large such that the conclusion of (4.9) produces a contradiction for  $\epsilon$  sufficiently small. Of course, since the  $U$  appearing in (4.9) is a solution to (2.14), such an upper bound is not to be expected. Instead, we make a modification of  $\Omega_{\epsilon,\tilde{N}}^*$  by intersecting with sets  $\Omega_{\epsilon,U,h}$  that quantify  $\|U(T)\|_{H^2}$  in terms of a function  $h = h(\epsilon)$ , which grows unboundedly as  $\epsilon \rightarrow 0$ . Returning to (4.9),  $\tilde{N}$  now depends on  $\epsilon$  through  $h$ . As such, care is needed in the choice of  $h$  to avoid a possibly circular argument.

*Proof of Theorem 4.1.* For each  $\epsilon \in (0, 1/e)$ , let

$$h(\epsilon) := \log(\log(\log(\epsilon^{-1}))), \quad \tilde{N}(\epsilon) := \left\lceil \frac{4C^* N^8 h(\epsilon)}{\alpha} \right\rceil, \quad (4.10)$$

where  $\lceil x \rceil$  denotes the smallest integer larger or equal to  $x$  and  $C^*$  is the constant from (4.5). Then

$$\frac{\alpha}{2} - C^* \frac{N^8}{\tilde{N}(\epsilon)} h(\epsilon) \geq \frac{\alpha}{4}. \quad (4.11)$$

Observe that, with these definitions, there exists  $\epsilon_1^* > 0$  such that  $\epsilon \in \mathcal{E}(\tilde{N}(\epsilon))$  whenever  $\epsilon < \epsilon_1^*$ , where  $\mathcal{E}$  is defined by (4.6).<sup>22</sup> For  $\epsilon < \epsilon_1^*$ ,  $\Omega_{\epsilon, \tilde{N}(\epsilon)}^*$  (see Proposition 4.4) is well defined and we set

$$\Omega_\epsilon := \hat{\Omega}_{\epsilon, U, h} \cap \Omega_{\epsilon, \tilde{N}(\epsilon)}^*,$$

where

$$\hat{\Omega}_{\epsilon, U, h} := \{1 + \|U(T)\|_{H^2}^2 \leq h(\epsilon)\}.$$

We now show that on the sets  $\Omega_\epsilon$  we obtain the desired conclusion (4.1). First observe that by (4.7), the Markov inequality, and (A.5), there is  $C = C(\eta, T)$  such that

$$\mathbb{P}(\Omega_\epsilon^c) \leq \mathbb{P}(\hat{\Omega}_{\epsilon, U, h}^c) + \mathbb{P}((\Omega_{\epsilon, \tilde{N}(\epsilon)}^*)^c) \leq C \left( \frac{1}{h(\epsilon)} + (\tilde{N}(\epsilon))^{q_3} \epsilon^{q_4^{\tilde{N}(\epsilon)}} \right) \exp(\eta \|U_0\|^2) =: r(\epsilon) \exp(\eta \|U_0\|^2) \quad (4.12)$$

whenever  $\epsilon < \epsilon_1^*$ . The later quantity  $r(\epsilon)$  decays to zero as  $\epsilon \rightarrow 0^+$ ; since  $h(\epsilon) \rightarrow \infty$  as  $\epsilon \rightarrow 0^+$

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0^+} \tilde{N}(\epsilon)^{q_3} \epsilon^{q_4^{\tilde{N}(\epsilon)}} &\leq \limsup_{\epsilon \rightarrow 0^+} h(\epsilon)^{2q_3} \epsilon^{\exp(-h(\epsilon)^2)} \\ &\leq \exp \left( \limsup_{\epsilon \rightarrow 0^+} (2q_3 \log h(\epsilon) + \exp(-h(\epsilon)^2) \log \epsilon) \right) \\ &= \exp \left( \limsup_{h \rightarrow \infty} (2q_3 \log(h) - \exp(-h^2) \exp(\exp(h))) \right) = 0. \end{aligned} \quad (4.13)$$

On the other hand, on  $\Omega_\epsilon$ , Propositions 4.3, 4.4, and (4.11) yield

$$\begin{aligned} \langle \mathcal{M}_{0, T} \phi, \phi \rangle &\leq \epsilon \|\phi\|^2 \\ \Rightarrow \epsilon^{q_5^{\tilde{N}(\epsilon)}} \|\phi\|^2 &\geq \left( \frac{\alpha}{2} - C^* \frac{N^8}{\tilde{N}} (1 + \|U(T)\|_{H^2}^2) \right) \|\phi\|^2 \geq \left( \frac{\alpha}{2} - C^* \frac{N^8}{\tilde{N}} h(\epsilon) \right) \|\phi\|^2 \geq \frac{\alpha}{4} \|\phi\|^2, \end{aligned}$$

for each  $\epsilon < \epsilon_1^*$  and any  $\phi \in \mathcal{S}_{\alpha, N}$ . We infer that if  $\epsilon < \epsilon_1^*$ , then on  $\Omega_\epsilon$ ,

$$\inf_{\phi \in \mathcal{S}_{\alpha, N}} \frac{\langle \mathcal{M}_{0, T} \phi, \phi \rangle}{\|\phi\|^2} > \epsilon \quad \text{whenever } \epsilon^{q_5^{\tilde{N}(\epsilon)}} < \frac{\alpha}{4}.$$

From the definition of  $\tilde{N}(\epsilon)$  in (4.10), we have  $\epsilon^{q_5^{\tilde{N}(\epsilon)}} \rightarrow 0$  as  $\epsilon \rightarrow 0^+$  (e.g. see (4.13) above), and the proof of Theorem 4.1 is complete.  $\square$

<sup>22</sup>Indeed, one can begin by supposing  $\epsilon_1^* < \min\{q_0, 1/e\}$  and observe that

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0^+} \epsilon \left( \frac{q_1}{\tilde{N}(\epsilon)} \right)^{-q_2^{\tilde{N}(\epsilon)}} &\leq \limsup_{\epsilon \rightarrow 0^+} \epsilon h(\epsilon)^{2q_2^{\tilde{N}(\epsilon)}} = \limsup_{\epsilon \rightarrow 0^+} \epsilon \exp \left( 2q_2^{\tilde{N}(\epsilon)} \log(h(\epsilon)) \right) \\ &\leq \limsup_{\epsilon \rightarrow 0^+} \epsilon \exp \left( \exp(h(\epsilon)^2) \right) = \limsup_{h \rightarrow \infty} \exp \left( -\exp(\exp(h)) \right) \exp \left( \exp(h^2) \right) = 0. \end{aligned}$$

## 5 Lie Bracket Computations

In this section we present, in functional setting, the chain of Lie brackets that approximately generate spanning sets for successively larger finite dimensional subspaces  $H_N$  of the phase space  $H$ . More precisely, we show that the approximate basis  $\mathfrak{B}_{N,\tilde{N}}(U)$  lie in admissible sets  $\mathcal{V}_M$  (see (4.2), (4.3)) for sufficiently large  $M = M(\tilde{N})$ , which allows us to define the quadratic forms  $\mathcal{Q}_{N,\tilde{N}}(U)$  leading to the upper and lower bounds, Propositions 4.3, 4.4 in Section 4.

As we described above, these computations are motivated by the celebrated Hörmander condition for the Kolmogorov-Fokker-Planck equations associated to (2.14), cf. (1.4). Our situation is notable in comparison to previous works in the infinite dimensional setting, [EM01, Rom04, HM06, HM11] as, to the best of our knowledge, we are the first to analyze a system, where the chain of vector fields and the associated quadratic forms  $\mathcal{Q}_{N,\tilde{N}} = \mathcal{Q}_{N,\tilde{N}}(U)$  depend on  $U$ , and are therefore random.

For contrast consider the 2D stochastic Navier-Stokes equation written in the vorticity formulation

$$d\omega + F_{NS}(\omega) = \sigma dW = \sum_{k,l} \alpha_k \psi_k^l dW^{k,l}, \quad \text{with } F_{NS}(\omega) := -\nu \Delta \omega + (K * \omega) \cdot \nabla \omega, \quad (5.1)$$

where as in (2.7) we set  $\psi_k^0 := \cos(k \cdot x)$  and  $\psi_k^1 := \sin(k \cdot x)$  for  $k \in \mathbb{Z}_+^2$ . Let  $\mathcal{V}_M^{NS}$ ,  $M \geq 0$  be the sequences of admissible vector fields corresponding to (5.1), defined analogously to (4.2). The following bracket structure for (5.1) was observed in [EM01]. Suppose that  $\psi_k^m, \psi_{k'}^{m'} \in \mathcal{V}_M^{NS}$  for some  $M \geq 0$ . Then, since  $(K * \omega) \cdot \nabla \omega$  is the only nonlinear (quadratic) term in (5.1), we obtain

$$[[F_{NS}(\omega), \psi_k^m], \psi_{k'}^{m'}] = (K * \psi_k^m) \cdot \nabla \psi_{k'}^{m'} + (K * \psi_{k'}^{m'}) \cdot \nabla \psi_k^m \in \mathcal{V}_{M+2}, \quad (5.2)$$

where the Lie brackets  $[\cdot, \cdot]$  are defined as in (1.6). Using elementary algebra, we obtain that  $\psi_{k+k'}^{m+m'} \in \mathcal{V}_{M+2}$  if  $k$  and  $k'$  are not parallel and  $|k| \neq |k'|$ . Under appropriate algebraic assumptions on the set of directly forced modes, one therefore obtains that for any  $N > 0$ ,  $H_N \subset \mathcal{V}_M$  for large enough  $M = M(N)$ . As we already noted in the introduction, this strategy of repeated brackets with constant vector field to generate exactly  $H_N$  has been used in all of the previously known examples; see [HM11].

Our situation is completely different. Since the random perturbation appears only in the temperature equation in (2.1)–(2.2), we immediately see, recalling the notation (2.6), that for any  $k, k' \in \mathbb{Z}_+^2$ ,  $m, m' \in \{0, 1\}$

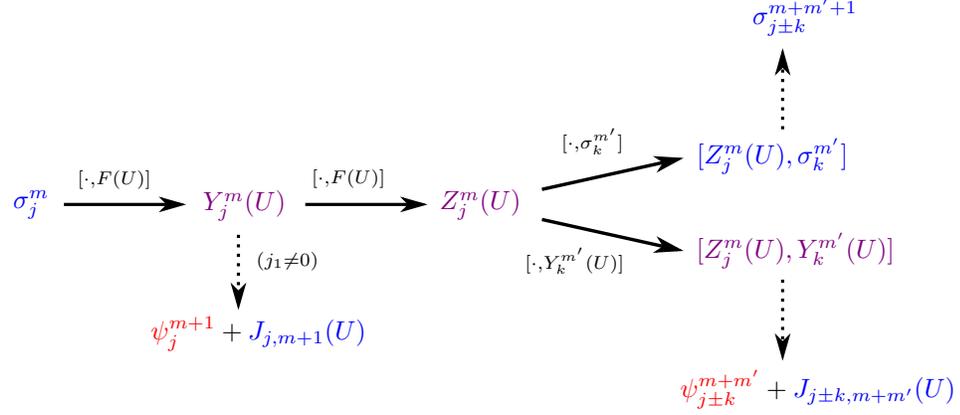
$$[[F(U), \sigma_k^m], \sigma_{k'}^{m'}] = B(\sigma_k^m, \sigma_{k'}^{m'}) + B(\sigma_{k'}^{m'}, \sigma_k^m) = 0, \quad (5.3)$$

and therefore no new modes are generated. The observation in (5.3) suggests that we need to make more carefully use of the interaction between the buoyancy term  $G$  and the advective structure in  $B$ .

The strategy which we devised to generate suitable directions is summarized in Figure 5.1 below. Strikingly, a bracket  $[[[F(U), \sigma_k^l], F(U)], \sigma_{k'}^{l'}] = c_1 B(\psi_k^{l+1}, \sigma_{k'}^{l'}) + c_2 B(\psi_{k'}^{l'+1}, \sigma_k^l)$  produces the desirable cancellation, where  $c_1, c_2$  are suitable constant.<sup>23</sup> While the computations leading to this cancellation are involved this ‘miracle’ is perhaps anticipated by the advective structure of  $B$ .

Having devised a strategy to generate  $\sigma$ -modes, we also have to generate the elements  $\psi_k^l$ ; that is, suitable directions in the  $\omega$  variable. In this case no additional cancellation is evident as we found for the  $\sigma$ -modes. Instead, we produce functions of the form  $\psi_l^m + J_{m,l}(U)$ , where  $J_{m,l}(U)$  is an ‘error term’ with a complicated dependence on  $U$ . Again as an artifact of the advective structure in  $B$ , these  $J_{m,l}(U)$  are concentrated in the  $\theta$  component only, and we can push these errors entirely into large wave-numbers by generating additional directions in  $\sigma$ . This in turn allows us to make use of the generalized Poincaré inequality to obtain bounds leading to our form of the Hörmander condition (1.9).

<sup>23</sup>As we already observed in Remark 2.4 it is precisely at this point that we are able to avoid the condition that the forcing in (1.1)–(1.2) contain wavenumbers of different magnitudes as is required for the 2D stochastic Navier-Stokes equation in [HM06].



**Figure 5.1:** The strategy for (approximately) generating  $H_N$  from Lie brackets. **Red** and **blue** text indicate that the corresponding functions reside exclusively in the  $\omega$  and  $\theta$  components, respectively. **Purple** text indicates that the element has non-trivial content in both components. Solid arrows mean that the new function is generated from a Lie bracket, with the type of bracket indicated above the arrow. Dotted arrows signify that the new element is generated as a linear combination of elements from the previous position.

## 5.1 Detailed Bracket Computations

Let us provide more details. In what follows  $U$  is arbitrary sufficiently smooth function.<sup>24</sup> In the forthcoming computations we make use of the following simple observations. Recalling that the Lie bracket between  $C^1$  vector fields  $E_1, E_2 : H \rightarrow H$  is given by  $[E_1, E_2](U) := \nabla E_2(U)E_1(U) - \nabla E_1(U)E_2(U)$  we have the antisymmetry and Jacobi identities

$$[E_1, E_2] = -[E_2, E_1], \quad [[E_1, E_2], E_3] + [[E_2, E_3], E_1] + [[E_3, E_1], E_2] = 0, \quad (5.4)$$

valid for any  $E_1, E_2, E_3$ . From (2.10) and (2.11) we have that for any  $\tilde{U} \in H^1$ ,

$$B(U, \tilde{U}) = 0 \quad \text{if } U = (0, \theta), \quad (5.5)$$

$$GU = 0 \quad \text{if } U = (\omega, 0), \quad (5.6)$$

$$G\sigma_j^m = (-1)^{m+1} g j_1 \psi_j^{m+1}. \quad (5.7)$$

Note also that in what follows the superscripts  $m$  appearing in the basis elements  $\sigma_k^m, \psi_k^m$  are understood modulo 2, for example by  $\sigma_k^{m+m'}$  we mean  $\sigma_k^{m+m' \pmod{2}}$ . For any  $j \in \mathbb{Z}^2$  we define  $j^\perp := (-j_2, j_1)$ .

We first show how the directions  $\sigma_k^m, k \in \mathbb{Z}_+^2, m \in \{0, 1\}$  can be obtained. Define  $Y_j^m(U) := [F(U), \sigma_j^m]$  and by (5.5), (5.7),

$$\begin{aligned} Y_j^m(U) &= A\sigma_j^m + B(\sigma_j^m, U) + B(U, \sigma_j^m) - G\sigma_j^m \\ &= \nu_2 |j|^2 \sigma_j^m + (-1)^m g j_1 \psi_j^{m+1} + B(U, \sigma_j^m). \end{aligned} \quad (5.8)$$

Now set

$$Z_j^m(U) := [F(U), Y_j^m(U)] = \nabla Y_j^m(U)F(U) - \nabla F(U)Y_j^m(U) \quad (5.9)$$

<sup>24</sup>At this stage in the argument,  $U$  is not necessarily a solution of (2.14). Of course we will use that  $U$  solves (2.14) later on in Section 6. This causes no problems since  $U$  is smooth; cf. Proposition 2.2.

and after some computations we derive the explicit formula

$$\begin{aligned} Z_j^m(U) &= B(F(U), \sigma_j^m) + \nu_2^2 |j|^4 \sigma_j^m + (-1)^m (\nu_1 + \nu_2) g j_1 |j|^2 \psi_j^{m+1} + A(B(U, \sigma_j^m)) + (-1)^m g j_1 B(\psi_j^{m+1}, U) \\ &\quad - B(U, -\nu_2 |j|^2 \sigma_j^m + (-1)^{m+1} g j_1 \psi_j^{m+1}) + B(U, B(U, \sigma_j^m)) - GB(U, \sigma_j^m). \end{aligned} \quad (5.10)$$

Remarkably, as alluded to above, the higher order bracket  $[Z_j^m(U), \sigma_k^{m'}]$  is independent of  $U$ . To see this and to derive a formula for  $[Z_j^m(U), \sigma_k^{m'}]$  one may proceed by explicit computations. Instead, we argue as follows: by (5.5), (5.8) we obtain  $[Y_j^m(U), \tilde{U}] = -B(\tilde{U}, \sigma_j^m) = 0$  for any  $\tilde{U} = (0, \tilde{U}_2)$ , and consequently by the Jacobi identity, (5.4)

$$[Z_j^m(U), \sigma_k^{m'}] = [[F(U), Y_j^m(U)], \sigma_k^{m'}] = -[Y_j^m(U), [F(U), \sigma_k^{m'}]] = -[Y_j^m(U), Y_k^{m'}(U)].$$

Moreover, since  $(-1)^{m+1} g j_1 \psi_j^{m+1}$  is the only term with non-zero first component in  $Y_j^m(U)$ , from (5.8) we obtain

$$[Y_j^m(U), Y_k^{m'}(U)] = B((-1)^m g j_1 \psi_j^{m+1}, \sigma_k^{m'}) - B((-1)^{m'} g k_1 \psi_k^{m'+1}, \sigma_j^m),$$

and therefore

$$[Z_j^m(U), \sigma_k^{m'}] = g \left( (-1)^{m+1} j_1 B(\psi_j^{m+1}, \sigma_k^{m'}) + (-1)^{m'} k_1 B(\psi_k^{m'+1}, \sigma_j^m) \right). \quad (5.11)$$

Next, note the following observation which is a consequence of simple trigonometric identities.

**Lemma 5.1.** *For any  $j, k \in \mathbb{Z}_+^2$  and  $m, m' \in \{0, 1\}$*

$$\begin{aligned} B(\psi_j^m, \sigma_k^{m'}) &= \frac{(-1)^{1+mm'}}{2} \frac{(j^\perp \cdot k)}{|j|^2} \left[ \sigma_{j+k}^{m+m'} + (-1)^{m'+1} \sigma_{j-k}^{m+m'} \right], \\ B(\psi_j^m, \psi_k^{m'}) &= \frac{(-1)^{1+mm'}}{2} \frac{(j^\perp \cdot k)}{|j|^2} \left[ \psi_{j+k}^{m+m'} + (-1)^{m'+1} \psi_{j-k}^{m+m'} \right]. \end{aligned}$$

Using Lemma 5.1 and (5.11) we have

$$[Z_j^m(U), \sigma_k^{m'}] = g (-1)^{(m+1)(m'+1)} \frac{(j^\perp \cdot k)}{2} \left[ (-1)^{m'} b(j, k) \sigma_{j-k}^{m+m'+1} - a(j, k) \sigma_{j+k}^{m+m'+1} \right], \quad (5.12)$$

where

$$a(j, k) := \frac{j_1}{|j|^2} + \frac{k_1}{|k|^2} \quad \text{and} \quad b(j, k) := \frac{j_1}{|j|^2} - \frac{k_1}{|k|^2}. \quad (5.13)$$

From these relations, the following proposition follows easily.

**Proposition 5.2.** *Let  $j, k \in \mathbb{Z}_+^2$ ,  $a(j, k), b(j, k)$  be as in (5.13). Then, with  $Z_j^m$  given by (5.9),*

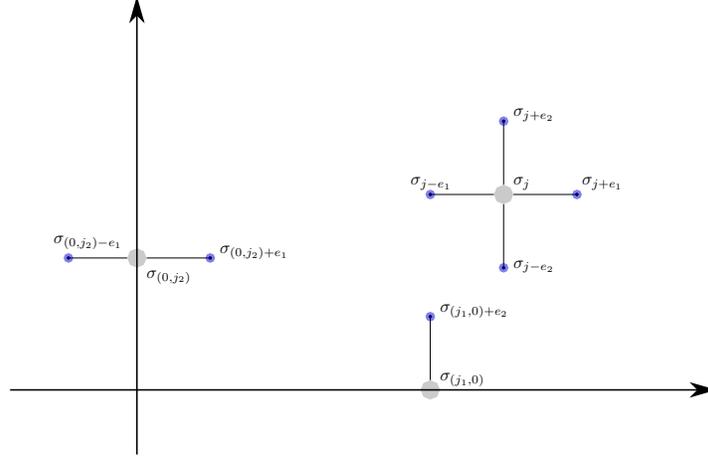
$$g(j^\perp \cdot k) a(j, k) \sigma_{j+k}^0 = -[Z_j^0(U), \sigma_k^1] - [Z_j^1(U), \sigma_k^0], \quad (5.14)$$

$$g(j^\perp \cdot k) a(j, k) \sigma_{j+k}^1 = [Z_j^0(U), \sigma_k^0] - [Z_j^1(U), \sigma_k^1], \quad (5.15)$$

$$g(j^\perp \cdot k) b(j, k) \sigma_{j-k}^0 = [Z_j^1(U), \sigma_k^0] - [Z_j^0(U), \sigma_k^1], \quad (5.16)$$

$$g(j^\perp \cdot k) b(j, k) \sigma_{j-k}^1 = -[Z_j^1(U), \sigma_k^1] - [Z_j^0(U), \sigma_k^0]. \quad (5.17)$$

**Remark 5.3.** *The diagram in Figure 5.2 and an induction argument detailed in Section 6.3 and illustrated in Figure 6.2 show that starting with the forced directions  $\sigma_{(1,0)}^m, \sigma_{(0,1)}^m$  for each  $m \in \{0, 1\}$ , it is possible to reach  $\sigma_k^m$  for any  $k \in \mathbb{Z}_+^2$  and  $m \in \{0, 1\}$ . If we replaced the vectors  $(1, 0), (0, 1)$  in the definition of  $\mathcal{Z}$  by other elements in  $\mathbb{Z}_+^2$ , Figure 5.2 would change, namely the segments parallel to axes would be changed to segments parallel to the new directions in  $\mathcal{Z}$ . In this case more a complicated algebraic condition as in [HM06] is needed to demonstrate that  $\mathcal{Z}$  generates a spanning set for  $\mathbb{Z}_+^2$ .*



**Figure 5.2:** New basis elements  $\sigma_j^m$  in the  $\theta$  components of the phase space (see (2.6)) that can be generated via the chain of brackets leading to (5.12) with Proposition 5.2. Recall that we index these basis elements by  $\mathbb{Z}_+^2$  so that the points in figure lie on lattice points. Note furthermore that elements  $\sigma_l$  correspond to both  $\sigma_l^0, \sigma_l^1$ . Grey points represent existing directions and purple points represent new directions generated via one iteration of the chain of brackets illustrated in the upper part of Figure 5.1. Perusing (5.14)–(5.17) and the definition of  $a$  and  $b$  in (5.13) we see why new directions are restricted along the  $x$  and  $y$  coordinate axes respectively.

Directions  $\psi_k^m$  are different and they include an error term with a component in the  $\theta$ -direction. First note that (5.8) can be rewritten as

$$(-1)^m g j_1 \psi_j^{m+1} = Y_j^m(U) - (-1)^m g j_1 J_{j,m+1}^*(U), \quad (5.18)$$

where  $J_{j,m+1}^*(U) := \frac{(-1)^m}{g j_1} (\nu_2 |j|^2 \sigma_j^m + B(U, \sigma_j^m))$ . Note that, by (2.10), we see that  $J_{j,m+1}^*$  is concentrated only in its  $\theta$  component. Since we can generate  $Y_j^m(U)$ , by (5.18), we can also generate  $\psi_j^{m+1}$  (with an error term), whenever  $j_1 \neq 0$ . This constitutes the first downward branch in the lower portion of Figure 5.1.

To reach the basis function  $\psi_j^m$  along the  $j_2$  axis ( $j_1 = 0$ ) we can mimic the approach in [EM01] by considering brackets of the form  $[[F(U), \psi_j^m], \psi_k^{m'}]$ ; cf. (5.2). Since we did not generate  $\psi_j^m, \psi_k^{m'}$ , we instead use elements  $Y_j^m(U)$ , that is,  $\psi_j^m$  with error terms and calculate

$$[[F(U), Y_j^m(U)], Y_k^{m'}(U)] = [Z_j^m(U), Y_k^{m'}(U)] = [[Z_j^m(U), F(U)], \sigma_k^{m'}] - [[Z_j^m(U), \sigma_k^{m'}], F(U)]. \quad (5.19)$$

Notice, that the second identity follows from (5.4) and shows that  $[[F(U), Y_j^m(U)], Y_k^{m'}(U)]$  can be obtained from admissible bracket operations. On the other hand, with (5.9)

$$\begin{aligned} [[F(U), Y_j^m(U)], Y_k^{m'}(U)] &= \nabla^2 F(U) \{Y_j^m(U), Y_k^{m'}(U)\} + \nabla F(U) \{\nabla Y_j^m(U) \{Y_k^{m'}(U)\}\} \\ &\quad - \nabla(\nabla Y_j^m(U) \{F(U)\}) \{Y_k^{m'}(U)\} + \nabla Y_k^{m'}(U) \{[F(U), Y_j^m(U)]\}, \end{aligned}$$

where  $\nabla^k G(U) \{X_1, \dots, X_k\}$  denotes the  $k$ th derivative of  $G$  in the directions  $X_1, \dots, X_k$ . With (5.5), (5.18), and (2.17) we obtain

$$\begin{aligned} \nabla^2 F(U) \{Y_j^m(U), Y_k^{m'}(U)\} &= -B(Y_j^m(U), Y_k^{m'}(U)) - B(Y_k^{m'}(U), Y_j^m(U)) \\ &= (-1)^{m+m'+1} g^2 j_1 k_1 \left( B(\psi_j^{m+1}, \psi_k^{m'+1}) + B(\psi_k^{m'+1}, \psi_j^{m+1}) \right. \\ &\quad \left. + B(\psi_j^m, J_{k,m'}^*(U)) + B(\psi_k^{m'}, J_{j,m}^*(U)) \right). \end{aligned}$$

Also, since  $\nabla Y_j^m(U)\{Y_k^{m'}(U)\} = -B(Y_k^{m'}(U), \sigma_j^m) = (-1)^{m'} g k_1 B(\psi_k^{m'}, \sigma_j^m)$  has a zero  $\omega$  component, we have

$$\nabla F(U)\{\nabla Y_j^m(U)\{Y_k^{m'}(U)\}\} = (-1)^{m'} g k_1 G B(\psi_k^{m'}, \sigma_j^m) + D_{j,k}^{m,m'}(U),$$

where  $U \mapsto D_{j,k}^{m,m'}(U)$  is affine and has a zero  $\omega$  component. Finally, since  $\nabla Y(U)\{X\}$  has a zero  $\omega$  component for any  $X$  (of course the same is true for a derivative of  $\nabla Y(U)\{X\}$ ), we obtain

$$\begin{aligned} [Z_j^m(U), Y_k^{m'}(U)] &= (-1)^{m+m'+1} g^2 j_1 k_1 \left( B(\psi_j^{m+1}, \psi_k^{m'+1}) + B(\psi_k^{m'+1}, \psi_j^{m+1}) \right) \\ &\quad + (-1)^{m'} g k_1 G B(\psi_k^{m'+1}, \sigma_j^m) + H_{j,k}^{m,m'}(U), \end{aligned} \quad (5.20)$$

where  $U \mapsto H_{j,k}^{m,m'}(U)$  is affine and it is concentrated entirely in the  $\theta$  component.

Using (5.20) we are able to reach basis function that are not accessible by the brackets leading to (5.18). Note that this second case ( $j_1 = 0$ ) is represented graphically the last lower branch of Figure 5.1.

**Proposition 5.4.** *Let  $\ell_2 > 0$ ,  $\ell = (0, \ell_2)$ ,  $\ell' = (1, \ell_2)$ , and  $\bar{e}_1 = (1, 0)$ . Then*

$$g^2 \frac{|\ell|^3}{|\ell'|^2} \psi_\ell^0 = -[Z_{\ell'}^0(U), Y_{\bar{e}_1}^0(U)] - [Z_{\ell'}^1(U), Y_{\bar{e}_1}^1(U)] + H_{\ell', \bar{e}_1}^{0,0}(U) + H_{\ell', \bar{e}_1}^{1,1}(U),$$

and

$$g^2 \frac{|\ell|^3}{|\ell'|^2} \psi_\ell^1 = [Z_{\ell'}^1(U), Y_{\bar{e}_1}^0(U)] - [Z_{\ell'}^0(U), Y_{\bar{e}_1}^1(U)] + H_{\ell', \bar{e}_1}^{0,1}(U) - H_{\ell', \bar{e}_1}^{1,0}(U).$$

*Proof.* From Lemma 5.1 and the fact  $G\sigma_\ell^m = 0$  (see (5.6)) one has

$$G B(\psi_{e_1}^{m'+1}, \sigma_{\ell'}^m) = (-1)^{mm'+m'+1} g \ell_2 \psi_{e_1+\ell'}^{m+m'}. \quad (5.21)$$

Combining (5.20), (5.21), and Lemma 5.1, we have

$$[Z_{\ell'}^m(U), Y_{e_1}^{m'}(U)] = (-1)^{mm'+1} \frac{g^2 \ell_2}{2} \left[ \frac{2 + \ell_2^2}{1 + \ell_2^2} \psi_{\ell'+e_1}^{m+m'} + \frac{\ell_2^2}{1 + \ell_2^2} (-1)^{m'} \psi_{\ell'-e_1}^{m+m'} \right] + H_{\ell', e_1}^{m,m'}(U). \quad (5.22)$$

The proposition follows after eliminating the term involving  $\psi_{\ell'+e_1}^{m+m'}$  by considering first the cases  $(m, m') = (0, 0)$ ,  $(m, m') = (1, 1)$  to determine  $\psi_\ell^0$  and  $(m, m') = (0, 1)$ ,  $(m, m') = (1, 0)$  to determine  $\psi_\ell^1$ .  $\square$

Combining (5.18) and Proposition 5.4 we now define the error term in the  $\psi$ -directions

$$J_{j,m}(U) = \begin{cases} (-1)^m \frac{\mu |j|^2}{g j_1} \sigma_j^{m+1} + (-1)^m \frac{1}{g j_1} B(U, \sigma_j^{m+1}) & \text{if } j_1 \neq 0, \\ \frac{1 + |j|^2}{g^2 |j|^3} (-H_{j+e_1, e_1}^{0,0}(U) - H_{j+e_1, e_1}^{1,1}(U)) & \text{if } j_1 = 0, m = 0, \\ \frac{1 + |j|^2}{g^2 |j|^3} (-H_{j+e_1, e_1}^{0,1}(U) + H_{j+e_1, e_1}^{1,0}(U)) & \text{if } j_1 = 0, m = 1. \end{cases} \quad (5.23)$$

By combining (5.8), Proposition 5.4, and (5.23), we have for each  $j \in \mathbb{Z}_+^2$ ,  $m \in \{0, 1\}$ ,

$$\psi_j^m + J_{j,m}(U) = \begin{cases} \frac{(-1)^m}{g j_1} Y_j^{m+1}(U) & \text{if } j_1 \neq 0, \\ \frac{1 + |j|^2}{g^2 |j|^3} (-[Z_{j+\bar{e}_1}^0(U), Y_{\bar{e}_1}^0(U)] - [Z_{j+\bar{e}_1}^1(U), Y_{\bar{e}_1}^1(U)]) & \text{if } j_1 = 0, m = 0, \\ \frac{1 + |j|^2}{g^2 |j|^3} ([Z_{j+\bar{e}_1}^1(U), Y_{\bar{e}_1}^0(U)] - [Z_{j+\bar{e}_1}^0(U), Y_{\bar{e}_1}^1(U)]) & \text{if } j_1 = 0, m = 1. \end{cases} \quad (5.24)$$

## 5.2 Estimates and Related Properties for the Error Terms

We next summarize some basic properties of  $J_{j,m}(U)$  in the following lemmata.

**Lemma 5.5.** *Fix any  $j \in \mathbb{Z}_+^2$  with  $|j| \leq N$ ,  $m \in \{0, 1\}$ , and any  $U \in H^1$ . Then,  $(J_{j,m}(U))_\omega = 0$  (the  $\omega$  component of  $J_{j,m}(U)$  is zero) and*

$$\|J_{j,m}(U)\| \leq CN^3(1 + \|U\|_{H^1}),$$

where the constant  $C$  is independent of  $N$ . Moreover  $U \mapsto J_{j,m}(U) - J_{j,m}(0)$  is linear, that is,  $U \mapsto J_{j,m}(U)$  is affine.

*Proof.* The proof is a straightforward consequence of definitions. For example by careful inspection we have an estimate

$$\|H_{j+e_1, e_1}^{m, m'}(U)\| \leq C|j|^4(1 + \|U\|_{H^1}),$$

of any  $m, m' \in \{0, 1\}$ . □

Lemma 5.5 does not provide us with sufficient estimate for  $J_{j,m}$  as it grows both in  $N$  and  $U$ . However, we crucially use the fact  $(J_{j,m})_\omega = 0$  as follows. We can generate sufficiently many, and consequently subtract from  $J_{j,m}$ , pure modes  $\sigma_k^m$ . Hence, we generate all modes  $\sigma_k^m$  with  $|k| \leq \tilde{N}$  but for approximation of  $H_N$  we use only those with  $|k| \leq N \ll \tilde{N}$ , the rest we use for controlling the size of the error  $J_{j,m}$  (for details see the proof of Lemma 6.12).

To this end we derive estimates for projections of  $J_{j,m}(U)$  into high Fourier modes. Recall that  $Q_N$  is the orthogonal projection on complement of  $H_N$  and denote

$$J_{j,m}^{\tilde{N}}(U) := Q_{\tilde{N}} J_{j,m}(U). \quad (5.25)$$

**Lemma 5.6.** *For every integers  $N, \tilde{N}$  with  $\tilde{N} \geq N > 0$ , and every integer  $s \geq 1$  and  $U \in H^{s+1}$*

$$\|J_{j,m}^{\tilde{N}}(U)\| \leq C \frac{N^{s+3}}{\tilde{N}^{s/2}} (1 + \|U\|_{H^{s+1}}) \quad (|j| \leq N, m \in \{0, 1\}), \quad (5.26)$$

where  $C = C(s)$  is independent of  $N, \tilde{N}$  and  $U$ .

*Proof.* Since the  $\tilde{N}^{\text{th}}$  eigenvalue  $\lambda_{\tilde{N}} \sim \tilde{N}$ , cf. [CF88], one has by the generalized Poincaré inequality that

$$\|J_{j,m}^{\tilde{N}}(U)\| \leq C \frac{1}{\lambda_{\tilde{N}}^{s/2}} \|J_{j,m}(U)\|_{H^s} \leq \frac{1}{\tilde{N}^{s/2}} \|J_{j,m}^{\tilde{N}}(U)\|_{H^s}.$$

By careful inspection of (5.23), noting that  $J_{j,m}$  is affine in  $U$ , we obtain

$$\|J_{j,m}^{\tilde{N}}(U)\|_{H^s} \leq C(1 + \|U\|_{H^{s+1}}),$$

where  $C = C(N, s)$ . The exact dependence of the right hand side on  $N$  can be inferred from the fact that each derivative of  $J_{j,m}(U)$  can produce at most one factor of  $|j| \leq N$ . □

## 6 Noise Propagation in the Phase Space: Quantitative Estimates

This section is devoted to the proof of Proposition 4.4. To establish this “upper bound” on the quadratic forms  $\mathcal{Q}_{N, \tilde{N}}$  defined in Section 4, recall that in Section 5 we showed that  $\mathfrak{B}_{N, \tilde{N}}(U) \subset \mathcal{V}_M$  (see (4.2), (4.3))

for sufficiently large  $M = M(\tilde{N})$ . We thus to translate each of the Lie bracket computations in Section 5 into quantitative bounds. Roughly speaking, we would like to show that

$$\langle \mathcal{M}_{0,T}\phi, \phi \rangle \text{ is 'small' implies that } \langle \phi, \sigma_k^l \rangle \text{ are all 'small' for all } k \in \mathcal{Z}, l \in \{0, 1\} \quad (6.1)$$

and that, starting from any admissible vector field  $E \in \mathcal{V}_M$ , cf. (4.2)

$$\langle \phi, E \rangle \text{ is 'small' implies that } \langle \phi, [E, \sigma_k^l] \rangle, \langle \phi, [E, F] \rangle \text{ are 'small' for all } k \in \mathcal{Z}, l \in \{0, 1\}. \quad (6.2)$$

To achieve (6.1), (6.2) we broadly follow an approach recently developed in [MP06, BM07, HM11].<sup>25</sup> Notice that

$$\langle \mathcal{M}_{s,t}\phi, \phi \rangle = \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0,1\}}} (\alpha_k^l)^2 \int_s^t \langle \sigma_k^l, \mathcal{K}_{r,t}\phi \rangle^2 dr \quad (6.3)$$

and define  $g_\phi(t) := \langle \mathcal{K}_{t,T}\phi, E(U(t)) \rangle$  over test functions  $\phi$  and admissible vector fields  $E$ . To address the first case in (6.2) we make use of a change of variable  $\tilde{U} := U - \sigma W$ . Expanding  $E(U)$  in this new variable we obtain a Wiener polynomial with coefficient similar to  $[E, \sigma_k^l]$  and we infer the ‘smallness’ from time regularity results for Wiener polynomial derived in [HM11] and recalled here as Theorem 6.4. For the second case in (6.2) we differentiate  $g_\phi$ , and find that  $g'_\phi = \langle \mathcal{K}_{t,T}\phi, [E, F] \rangle$ , at least up to a change of variable. We then make use of the fact that we can bound the maximum of  $g'_\phi$  in terms of, for example,  $g_\phi$  and  $C^\alpha$  norms of  $g'_\phi$  to deduce the desired implication.

Observe that our quadratic forms  $\mathcal{Q}_{N,\tilde{N}}$  depend on  $U(T)$  and thus have a strong probabilistic dependence. Indeed the existence of these ‘error’ terms in (4.4) means that we have to carefully track the growth of constants as a function of the number of Lie brackets we take. We also need to explain, at a quantitative level, how we are able to push error terms to entirely to high wavenumbers. Neither of these concerns can be addressed from an ‘obvious inspection’ of the methods in [HM11]. In addition to these mathematical concerns, we have developed several Lemmata 6.1, 6.2 which we believe streamline the presentation of some of the arguments in comparison to previous works.

The rest of the section is organized as follows: We begin with some generalities introducing or recalling some general lemmata that will be used repeatedly in the course of arguments leading to the rigorous form of (6.1)–(6.2). In Subsection 6.2 we present the series of Lemmas 6.5–6.9 each of which corresponds to one (or more) of the Lie brackets computed in Section 5. As we proceed we refer to Figure 6.1 to help guide the reader through some admittedly involved computations. In Subsection 6.3 we piece together the proved implications in an inductive argument to complete the proof of Proposition 4.4 In all that follows we maintain the convention from Remark 2.1, that is, all constants are implicitly dependent on given parameters  $\nu_1, \nu_2, g, \alpha, \dots$  of the problem. Note also that we carry out our arguments on a general time interval  $[T/2, T]$  for some  $T > 0$ , which makes all constants  $T$  dependent; we apply the conclusion only for  $T = 1$  above for the proof of Theorem 2.3.

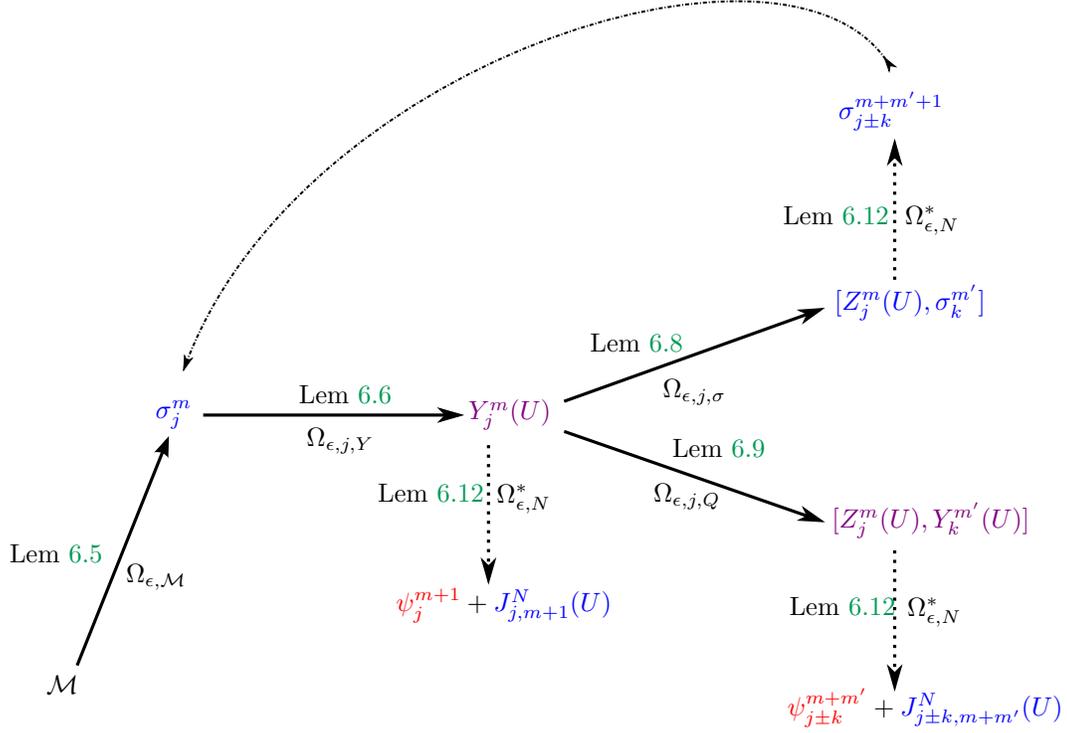
## 6.1 Preliminaries

We begin by introducing some further notational conventions and some general Lemmata 6.1, 6.2, 6.4 which are be used frequently in the course of the analysis.

For any  $a < b$ ,  $\beta \in \mathbb{R}$  and  $\alpha \in (0, 1]$  define the semi-norms

$$\|U\|_{C^\alpha([a,b], H^\beta)} := \sup_{\substack{t_1 \neq t_2 \\ t_1, t_2 \in [a,b]}} \frac{\|U(t_1) - U(t_2)\|_{H^\beta}}{|t_1 - t_2|^\alpha}.$$

<sup>25</sup>As in these works, the more classical methods using the Norris lemma do not apply, since it requires the inversion of the operators  $\mathcal{J}_{0,t}$ . See [Nor86] and also e.g. [Nua06, Hai11] for further details.



**Figure 6.1:** An illustration of the structure of the lemmata in Section 6 that leads to the proof of Proposition 4.4. The arrows indicate that if one term is ‘small’ then the other one is ‘small’ on a set of large measure (displayed below the arrow), where the meaning of ‘smallness’ is made precise in each lemma. Above the arrow we indicate in which lemma is this implication proved. The arrow at the top of the diagram shows that the process is iterative.

If  $a = T/2$  and  $b = T$  we will write  $\|\cdot\|_{C^\alpha H^\beta}$  instead of  $\|\cdot\|_{C^\alpha([T/2, T], H^\beta)}$  and denote  $\|U\|_{C^0 H^\beta} := \sup_{t \in [T/2, T]} \|U\|_{H^\beta}$ . Similar notations will be employed for the Hölder spaces  $C^\alpha([a, b])$ ,  $C^{1, \alpha}([a, b])$  etc. Recalling the notation in (1.6) we will define the ‘generalized Lie bracket’

$$[E_1(U), E_2(\tilde{U})] := \nabla E_2(\tilde{U})E_1(U) - \nabla E_1(U)E_2(\tilde{U}),$$

for all suitably regular  $E_1, E_2 : H \rightarrow H$  and  $U, \tilde{U} \in H$ . Below we often consider  $\bar{U} = U - \sigma_\theta W$  which satisfies the shifted equation (cf. (2.14))

$$\partial_t \bar{U} = F(U) = F(\bar{U} + \sigma_\theta W), \quad \bar{U}(0) = U_0. \quad (6.4)$$

Note that, in contrast to  $U$ ,  $\bar{U}$  is  $C^{1, \alpha}$  in time for any  $\alpha < 1/2$ .

We next prove two auxiliary lemmata which encapsulates the process of obtaining  $[E, F]$  type brackets from time differentiation.

**Lemma 6.1.** *Suppose  $E : H \rightarrow H$  is Fréchet differentiable,  $\phi \in H$ ,  $U$  solves (2.14), and  $\bar{U}$  is defined by (6.4). Then for any  $p \geq 1$  and any  $\eta > 0$ , we have that*

$$\mathbb{E} \sup_{t \in [T/2, T]} |\partial_t \langle \mathcal{K}_{t, T} \phi, E(\bar{U}) \rangle|^p \leq C \|\phi\|^p \exp(\eta \|U_0\|^2) \left( \mathbb{E} \sup_{t \in [T/2, T]} \|[E(\bar{U}), F(U)]\|^{2p} \right)^{1/2}, \quad (6.5)$$

where  $C = C(\eta, p, T)$ . Moreover, for any  $\alpha \in (0, 1]$ ,

$$\begin{aligned} & \mathbb{E} \left( \|\partial_t \langle \mathcal{K}_{t,T} \phi, E(\bar{U}) \rangle\|_{C^\alpha}^p \right) \\ & \leq C \|\phi\|^p \exp(\eta \|U_0\|^2) \cdot \left[ \left( \mathbb{E} \sup_{t \in [T/2, T]} \|[E(\bar{U}), F(U)]\|_{H^2}^{2p} \right)^{1/2} + \left( \mathbb{E} \|[E(\bar{U}), F(U)]\|_{C^\alpha H}^{2p} \right)^{1/2} \right], \end{aligned} \quad (6.6)$$

with  $C = C(\eta, p, T)$ .

*Proof.* Since  $\mathcal{K}_{t,T} \phi$  solves (3.12) and  $\bar{U}$  satisfies (6.4) we have

$$\begin{aligned} \partial_t \langle \mathcal{K}_{t,T} \phi, E(\bar{U}) \rangle &= \langle \partial_t \mathcal{K}_{t,T} \phi, E(\bar{U}) \rangle + \langle \mathcal{K}_{t,T} \phi, \nabla E(\bar{U}) \cdot \partial_t \bar{U} \rangle \\ &= -\langle \mathcal{K}_{t,T} \phi, \nabla F(U) E(\bar{U}) \rangle + \langle \mathcal{K}_{t,T} \phi, \nabla E(\bar{U}) F(U) \rangle \\ &= \langle \mathcal{K}_{t,T} \phi, [F(U), E(\bar{U})] \rangle. \end{aligned} \quad (6.7)$$

Now, (6.5) immediately follows from Hölder inequality and (A.14). To prove (6.6), we use that for any  $\alpha \in (0, 1)$ ,  $s, s' \in \mathbb{R}$ , and any suitably regular  $A, B$  one has

$$\begin{aligned} \|\langle A, B \rangle\|_{C^\alpha} &:= \sup_{\substack{t \neq s \\ s, t \in [T/2, T]}} \left| \frac{\langle A(t), B(t) \rangle - \langle A(s), B(s) \rangle}{|s - t|^\alpha} \right| = \sup_{\substack{t \neq s \\ s, t \in [T/2, T]}} \left| \frac{\langle A(t) - A(s), B(t) \rangle + \langle A(s), B(t) - B(s) \rangle}{|s - t|^\alpha} \right| \\ &\leq \|A\|_{L^\infty H^{-s}} \|B\|_{C^\alpha H^s} + \|A\|_{C^\alpha H^{-s'}} \|B\|_{L^\infty H^{s'}}. \end{aligned} \quad (6.8)$$

Combining (6.8) with (6.7) and using Hölder's inequality,

$$\begin{aligned} \mathbb{E} \left( \|\partial_t \langle \mathcal{K}_{t,T} \phi, E(\bar{U}) \rangle\|_{C^\alpha}^p \right) &\leq C \left( \mathbb{E} \left( \sup_{t \in [T/2, T]} \|\mathcal{K}_{t,T} \phi\|^{2p} \right) \right)^{1/2} \left( \mathbb{E} \left( \|[E(\bar{U}), F(U)]\|_{C^\alpha H}^{2p} \right) \right)^{1/2} \\ &\quad + C \left( \mathbb{E} \left( \|\mathcal{K}_{t,T} \phi\|_{C^\alpha H^{-2}}^{2p} \right) \right)^{1/2} \left( \mathbb{E} \left( \|[E(\bar{U}), F(U)]\|_{H^2}^{2p} \right) \right)^{1/2} \end{aligned}$$

and (6.6) follows from (A.14) and (A.16).  $\square$

**Lemma 6.2.** Fix  $T > 0$ ,  $\alpha \in (0, 1]$  and an index set  $\mathcal{I}$ . Consider a collection of random functions  $g_\phi$  taking values in  $C^{1,\alpha}([T/2, T])$  and indexed by  $\phi \in \mathcal{I}$ . Define, for each  $\epsilon > 0$ ,

$$\Lambda_{\epsilon, \alpha} := \bigcup_{\phi \in \mathcal{I}} \Lambda_{\epsilon, \alpha}^\phi, \quad \text{where} \quad \Lambda_{\epsilon, \alpha}^\phi := \left\{ \sup_{t \in [T/2, T]} |g_\phi(t)| \leq \epsilon \text{ and } \sup_{t \in [T/2, T]} |g'_\phi(t)| > \epsilon^{\frac{\alpha}{2(1+\alpha)}} \right\}. \quad (6.9)$$

Then, there is  $\epsilon_0 = \epsilon_0(\alpha, T)$  such that for each  $\epsilon \in (0, \epsilon_0)$

$$\mathbb{P}(\Lambda_{\epsilon, \alpha}) \leq C \epsilon \mathbb{E} \left( \sup_{\phi \in \mathcal{I}} \|g_\phi\|_{C^{1,\alpha}([T/2, T])}^{2/\alpha} \right). \quad (6.10)$$

*Proof.* As observed in [HM11, Lemma 6.14] we have the elementary bound

$$\|f'\|_{L^\infty} \leq 4 \|f\|_{L^\infty} \max \left\{ \frac{2}{T}, \|f\|_{L^\infty}^{-1/(1+\alpha)} \|f'\|_{C^\alpha}^{1/(1+\alpha)} \right\}, \quad (6.11)$$

which is valid for any  $f \in C^{1,\alpha}([T/2, T])$ . Fix any  $\phi \in \mathcal{I}$ . On the set  $\Lambda_{\epsilon, \alpha}^\phi$ , if  $g_\phi$  attains the maximum in (6.11) in the first term, then

$$\epsilon^{\frac{\alpha}{2(1+\alpha)}} < \|g'_\phi\|_{L^\infty} \leq 4 \|g_\phi\|_{L^\infty} \frac{2}{T} \leq \epsilon \frac{8}{T}.$$

Clearly this cannot happen if  $\epsilon < \epsilon_0(\alpha, T) := (T/8)^{2(1+\alpha)/(2+\alpha)}$ . Thus for any  $\epsilon < \epsilon_0$ , on  $\Lambda_{\epsilon, \alpha}^\phi$ , one has

$$\epsilon^{\frac{\alpha}{2(1+\alpha)}} < 4\epsilon^{\alpha/(1+\alpha)} \|g'_\phi\|_{C^\alpha}^{1/(1+\alpha)}, \quad \text{i.e.} \quad \|g'_\phi\|_{C^\alpha}^{2/\alpha} \geq 4^{-2/\alpha(1+\alpha)} \epsilon^{-1} := C(\alpha)\epsilon^{-1}.$$

Since this lower bound is independent of  $\phi \in \mathcal{I}$  we infer

$$\Lambda_{\epsilon, \alpha} \subset \left\{ \sup_{\phi \in \mathcal{I}} \|g'_\phi\|_{C^\alpha}^{2/\alpha} \geq C(\alpha)\epsilon^{-1} \right\}.$$

With this observation and the Markov inequality we infer (6.10), completing the proof.  $\square$

**Remark 6.3.** *Observe that*

$$\Lambda_{\epsilon, \alpha}^c = \bigcap_{\phi \in \mathcal{I}} \left\{ \sup_{t \in [T/2, T]} |g_\phi(t)| > \epsilon \text{ or } \sup_{t \in [T/2, T]} |g'_\phi(t)| \leq \epsilon^{\frac{\alpha}{2(1+\alpha)}} \right\}.$$

Thus, on  $\Lambda_{\epsilon, \alpha}^c$ ,

$$\sup_{t \in [T/2, T]} |g_\phi(t)| < \epsilon \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |g'_\phi(t)| \leq \epsilon^{\frac{\alpha}{2(1+\alpha)}} \quad (6.12)$$

for every  $\phi \in \mathcal{I}$ .

Finally, we recall in our notations, a crucial quantitative bound on Wiener polynomials established in [HM11]. In particular this restatement avoids the language of ‘almost implication’ introduced in [HM11, Hai11].

Given any multi-index  $\alpha := (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$  recall the standard notation  $W^\alpha := W_1^{\alpha_1} \dots W_d^{\alpha_d}$ .

**Theorem 6.4** (Hairer-Mattingly, [HM11]). *Fix  $M, T > 0$ . Consider the collection  $\mathfrak{P}_M$  of  $M^{\text{th}}$  degree of ‘Wiener polynomials’ of the form*

$$F = A_0 + \sum_{|\alpha| \leq M} A_\alpha W^\alpha,$$

where for each multi-index  $\alpha$ , with  $|\alpha| \leq M$ ,  $A_\alpha : \Omega \times [0, T] \rightarrow \mathbb{R}$  is an arbitrary stochastic process. Then, for all  $\epsilon \in (0, 1)$  and  $\beta > 0$ , there exists a measurable set  $\Omega_{\epsilon, M, \beta}$  with

$$\mathbb{P}(\Omega_{\epsilon, M, \beta}^c) \leq C\epsilon,$$

such that on  $\Omega_{\epsilon, M, \beta}$  and for every  $F \in \mathfrak{P}_M$

$$\sup_{t \in [0, T]} |F(t)| < \epsilon^\beta \quad \Rightarrow \quad \begin{cases} \text{either} & \sup_{|\alpha| \leq M} \sup_{t \in [0, T]} |A_\alpha(t)| \leq \epsilon^{\beta 3^{-M}}, \\ \text{or} & \sup_{|\alpha| \leq M} \sup_{s \neq t \in [0, T]} \frac{|A_\alpha(t) - A_\alpha(s)|}{|t - s|} \geq \epsilon^{-\beta 3^{-(M+1)}}. \end{cases}$$

## 6.2 Implications Starting from Small Eigenvalues

We now start proving the implications depicted in Figure 6.1. Note that throughout what follows we fix a small constant  $\epsilon_0 = \epsilon_0(T)$  which gives the range of  $\epsilon$  values for which Lemmas 6.5–6.9 hold. The first lemma explains how a lower bounds bound on the eigenvalues of  $\mathcal{M}_{0, T}$  initiates the iteration.

**Lemma 6.5.** *For every  $0 < \epsilon < \epsilon_0(T)$  and every  $\eta > 0$  there exist a set  $\Omega_{\epsilon, \mathcal{M}}$  and  $C = C(\eta, T)$  with*

$$\mathbb{P}(\Omega_{\epsilon, \mathcal{M}}^c) \leq C \exp(\eta \|U_0\|^2) \epsilon$$

such that on the set  $\Omega_{\epsilon, \mathcal{M}}$

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, \sigma_k^l \rangle| \leq \epsilon^{1/8} \|\phi\|,$$

for each  $k \in \mathcal{Z}$ ,  $l \in \{0, 1\}$ , and every  $\phi \in H$ . We recall that  $\mathcal{Z} \subset \mathbb{Z}_+^2$  is the set of directly forced modes as in (2.13) and the elements  $\sigma_k^l$  are given by (2.6).

*Proof.* For any  $\phi \in H$  with  $\|\phi\| = 1$ , define

$$g_\phi(t) := \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0, 1\}}} (\alpha_k^l)^2 \int_0^t \langle \sigma_k^l, \mathcal{K}_{r, T} \phi \rangle^2 dr \leq \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0, 1\}}} (\alpha_k^l)^2 \int_0^T \langle \sigma_k^l, \mathcal{K}_{r, T} \phi \rangle^2 dr = \langle \mathcal{M}_{0, T} \phi, \phi \rangle,$$

see (6.3). Note that

$$g'_\phi(t) = \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0, 1\}}} (\alpha_k^l)^2 \langle \sigma_k^l, \mathcal{K}_{t, T} \phi \rangle^2, \quad g''_\phi(t) = 2 \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0, 1\}}} (\alpha_k^l)^2 \langle \sigma_k^l, \mathcal{K}_{t, T} \phi \rangle \langle \sigma_k^l, \partial_t \mathcal{K}_{t, T} \phi \rangle.$$

Let  $\Omega_{\epsilon, \mathcal{M}} := \Lambda_{\epsilon, 1}^c$ , where  $\Lambda_{\epsilon, \alpha}$  is as in (6.9) with  $\mathcal{I} := \{\phi \in H : \|\phi\| = 1\}$ . Then by Lemma 6.2 with  $\alpha = 1$ , (A.14), and (A.16) one has

$$\mathbb{P}(\Omega_{\epsilon, \mathcal{M}}^c) \leq C \epsilon \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0, 1\}}} (\alpha_k^l)^4 \mathbb{E} \left( \sup_{\substack{t \in [T/2, T] \\ \|\phi\|=1}} |\langle \sigma_k^l, \mathcal{K}_{t, T} \phi \rangle \langle \sigma_k^l, \partial_t \mathcal{K}_{t, T} \phi \rangle|^2 \right) \leq C \exp(\eta \|U_0\|^2) \epsilon$$

for any  $\epsilon < \epsilon_0 = \epsilon_0(T)$ , where  $C = C(\eta, T)$ . Finally, on  $\Omega_{\epsilon, \mathcal{M}}$  we have, cf. (6.12), that

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\alpha_{k, l}| |\langle \mathcal{K}_{t, T} \phi, \sigma_j^l \rangle| \leq \epsilon^{1/2} \|\phi\|,$$

for each  $k \in \mathcal{Z}$ ,  $l \in \{0, 1\}$  and any  $\phi \in H$ . Since  $\alpha_k^l \neq 0$ , the assertion of the lemma follows for  $\epsilon \leq \epsilon_0(T)$ .  $\square$

We next turn to implications of the form  $\sigma \rightarrow [F, \sigma] = Y$ ; see Figure 6.1.

**Lemma 6.6.** *Fix any  $j \in \mathbb{Z}_+^2$ . For each  $0 < \epsilon < \epsilon_0(T)$  and  $\eta > 0$  there exist a set  $\Omega_{\epsilon, j, Y}$  and  $C = C(\eta, T)$  with*

$$\mathbb{P}(\Omega_{\epsilon, j, Y}^c) \leq C |j|^8 \exp(\eta \|U_0\|^2) \epsilon,$$

such that on the set  $\Omega_{\epsilon, j, Y}$ , for each  $m \in \{0, 1\}$ , it holds that

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle| \leq \epsilon \|\phi\| \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Y_j^m(U) \rangle| \leq \epsilon^{1/4} \|\phi\|. \quad (6.13)$$

*Proof.* By expanding  $U = \bar{U} + \sigma W$ , and using (5.5), we observe that

$$Y_j^m(U) = Y_j^m(\bar{U}). \quad (6.14)$$

Then for fixed  $m \in \{0, 1\}$  and any  $\phi \in \mathcal{I} := \{\phi \in H : \|\phi\| = 1\}$  define  $g_\phi(t) := \langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle$  and observe by (6.7) and (6.14) that  $g'_\phi(t) = \langle \mathcal{K}_{t, T} \phi, [F(U), \sigma_j^m] \rangle = \langle \mathcal{K}_{t, T} \phi, Y_j^m(U) \rangle = \langle \mathcal{K}_{t, T} \phi, Y_j^m(\bar{U}) \rangle$ . Let  $\Omega_{\epsilon, j, Y} := \Lambda_{\epsilon, 1}^c$

with  $\Lambda_{\epsilon,1}$  as in (6.9). Once again, with (6.12), we see that (6.13) holds on  $\Omega_{\epsilon,j,Y}$ . On the other hand, by (6.10), (6.5), (6.14), (5.9), and (A.5), we have

$$\begin{aligned} \mathbb{P}(\Omega_{\epsilon,j,Y}^c) &\leq C\epsilon \mathbb{E} \left( \sup_{\phi \in \mathcal{I}} \sup_{t \in [T/2, T]} |\partial_t \langle \mathcal{K}_{t,T} \phi, Y_j^m(\bar{U}) \rangle|^2 \right) \leq C\epsilon \exp\left(\frac{\eta}{2} \|U_0\|^2\right) \mathbb{E} \sup_{t \in [T/2, T]} \|Z_j^m(U)\|^4 \\ &\leq C\epsilon |j|^8 \exp\left(\frac{\eta}{2} \|U_0\|^2\right) \mathbb{E} \left( 1 + \sup_{t \in [T/2, T]} \|U\|_{H^2}^8 \right) \leq C\epsilon |j|^8 \exp(\eta \|U_0\|^2) \end{aligned}$$

for any  $\epsilon < \epsilon^*(T)$ , where  $C = C(\eta, T)$ . For the third inequality above we have also used the estimate

$$\sup_{t \in [T/2, T]} \|Z_j^m(U)\|_{H^s} \leq C|j|^{4+s} \left( 1 + \sup_{t \in [T/2, T]} \|U\|_{H^{s+2}}^2 \right), \quad (6.15)$$

which follows from (5.10) by counting derivatives and applying the Hölder and Poincaré inequalities.  $\square$

**Remark 6.7.** *The constants in the exponents of  $|j|$  and  $\epsilon$  in the forthcoming Lemmas 6.8, 6.9 rapidly become; however, there is nothing special about these numbers. We simply need to track that in the bounds  $|j|$  and  $\epsilon$  grow like  $|j|^\tau$  and  $\epsilon^\kappa$  respectively for some  $\kappa, \tau > 0$ .*

We next establish implications corresponding the chain of brackets  $Y \rightarrow Z \rightarrow [Z, \sigma]$ . We refer again to the Figure 6.1.

**Lemma 6.8.** *Fix  $j \in \mathbb{Z}_+^2$ . For each  $0 < \epsilon < \epsilon_0(T)$ , and  $\eta > 0$  there exist a set  $\Omega_{\epsilon,j,\sigma}$  and  $C = C(\eta, T)$  with*

$$\mathbb{P}(\Omega_{\epsilon,j,\sigma}^c) \leq C|j|^{90 \times 6} \exp(\eta \|U_0\|^2) \epsilon, \quad (6.16)$$

such that on the set  $\Omega_{\epsilon,j,\sigma}$ , for each  $m \in \{0, 1\}$ , it holds that

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, Y_j^m(U) \rangle| \leq \epsilon \|\phi\| \quad \Rightarrow \quad \begin{cases} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, Z_j^m(\bar{U}) \rangle| \leq \epsilon^{1/30} \|\phi\|, \\ \sup_{\substack{k \in \mathcal{Z} \\ l \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, [Z_j^m(U), \sigma_k^l] \rangle| \leq \epsilon^{1/60} \|\phi\|. \end{cases}$$

*Proof.* In the course of the proof we suppress, for the sake of brevity, the subscript  $\sigma$  in the definition of various sets leading to  $\Omega_{\epsilon,j,\sigma}$ . For fixed  $m \in \{0, 1\}$  and  $\phi \in H$  let  $g_\phi(t) := \langle \mathcal{K}_{t,T} \phi, Y_j^m(U) \rangle = \langle \mathcal{K}_{t,T} \phi, Y_j^m(\bar{U}) \rangle$  (cf. (6.14)) so that  $g'_\phi(t) = \langle \mathcal{K}_{t,T} \phi, [Y_j^m(\bar{U}), F(U)] \rangle = -\langle \mathcal{K}_{t,T} \phi, Z_j^m(U) \rangle$  (see (6.7), (5.9)). Let  $\Omega_{\epsilon,j}^1 = \Lambda_{\epsilon,1/4}^c$ , where  $\Lambda_{\epsilon,\alpha}^c$  is as in (6.9) over with  $\mathcal{I} := \{\phi \in H : \|\phi\| = 1\}$ . Then, on  $\Omega_{\epsilon,j}^1$  one has, in view of (6.12),

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, Y_j^m(U) \rangle| \leq \epsilon \|\phi\| \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, Z_j^m(U) \rangle| \leq \epsilon^{1/10} \|\phi\|. \quad (6.17)$$

By Lemma 6.2 with  $\alpha = 1/4$  and (6.6), (A.5), (A.6) we have

$$\begin{aligned} \mathbb{P}((\Omega_{\epsilon,j}^1)^c) &\leq C\epsilon \mathbb{E} \left( \sup_{\phi \in \mathcal{I}} \|g'\|_{C^{1/4}}^8 \right) \\ &\leq C\epsilon \exp\left(\frac{\eta}{2} \|U_0\|^2\right) \left[ \left( \mathbb{E} \sup_{t \in [T/2, T]} \|Z_j^m(U)\|_{H^2}^{16} \right)^{1/2} + \left( \mathbb{E} \|Z_j^m(U)\|_{C^{1/4}H}^{16} \right)^{1/2} \right] \\ &\leq C\epsilon |j|^{48} \exp\left(\frac{\eta}{2} \|U_0\|^2\right) \left[ \left( \mathbb{E} \left( 1 + \sup_{t \in [T/2, T]} \|U\|_{H^4}^{32} \right) \right)^{1/2} \right. \\ &\quad \left. + \left( \mathbb{E} \left( \|U\|_{C^{1/4}H^2}^{16} \left( 1 + \sup_{t \in [T/2, T]} \|U\|_{H^2}^{16} \right) \right) \right)^{1/2} \right] \\ &\leq C\epsilon |j|^{48} \exp(\eta \|U_0\|^2), \end{aligned} \quad (6.18)$$

where  $C = C(\eta, T)$  and we used the bilinearity of  $Z$  with estimates like those leading to (6.15). Next, by expanding  $U = \bar{U} + \sigma W$  we find

$$Z_j^m(U) = Z_j^m(\bar{U}) - \sum_{\substack{k \in \mathcal{Z} \\ l \in \{0,1\}}} \alpha_k^l [Z_j^m(U), \sigma_k^l] W^{k,l}. \quad (6.19)$$

In view of (5.11), all of the second order terms in (6.19) of the form  $[[Z(U), \sigma_k^l], \sigma_{k'}^{l'}] W^{k,l} W^{k',l'}$  are zero and  $[Z_j^m(U), \sigma_k^l] = [Z_j^m(\bar{U}), \sigma_k^l]$ .

To estimate each of the terms in (6.19), we introduce for  $s \in \{0, 1\}$ ,  $\phi \in H$

$$\mathcal{N}_s(\phi) := \max_{k \in \mathcal{Z}, l \in \{0,1\}} \left\{ \left\| \langle \mathcal{K}_{t,T} \phi, Z_j^m(\bar{U}) \rangle \right\|_{C^s}, |\alpha_{k,l}| \left\| \langle \mathcal{K}_{t,T} \phi, [Z_j^m(U), \sigma_k^l] \rangle \right\|_{C^s} \right\}.$$

By Theorem 6.4, there exists a set  $\Omega_\epsilon^\sharp$  such that  $\mathbb{P}((\Omega_\epsilon^\sharp)^c) < C\epsilon$ , and on  $\Omega_\epsilon^\sharp$

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, Z_j^m(U) \rangle| \leq \epsilon^{1/10} \implies \begin{cases} \text{either } \mathcal{N}_0(\phi) \leq \epsilon^{1/30}, \\ \text{or } \mathcal{N}_1(\phi) \geq \epsilon^{-1/90}. \end{cases} \quad (6.20)$$

Recalling that  $\mathcal{I} = \{\phi \in H : \|\phi\| = 1\}$ , let

$$\Omega_{\epsilon,j}^2 := \bigcap_{\phi \in \mathcal{I}} \{\mathcal{N}_0(\phi) < \epsilon^{1/30}\} \cap \Omega_{\epsilon,j}^\sharp.$$

By (6.17) on the set  $\Omega_{\epsilon,j,\sigma}^1 := \Omega_{\epsilon,j}^1 \cap \Omega_{\epsilon,j}^2$  we obtain the desired conclusion for each  $\epsilon < \epsilon_0(T)$ . Thus it remains to estimate the size of  $\Omega_{\epsilon,j,\sigma}^c$ . By (6.20), (6.18), and the Markov inequality we have

$$\begin{aligned} \mathbb{P}(\Omega_{\epsilon,j,\sigma}^c) &\leq \mathbb{P}((\Omega_{\epsilon,j}^1)^c) + \mathbb{P}((\Omega_{\epsilon,j}^\sharp)^c) + \mathbb{P}\left(\sup_{\phi \in \mathcal{I}} \mathcal{N}_1(\phi) \geq \epsilon^{-1/90}\right) \\ &\leq C|j|^{48} \exp(\eta \|U_0\|^2) \epsilon + C\epsilon \mathbb{E}\left(\sup_{\phi \in \mathcal{I}} (\mathcal{N}_1(\phi))^{90}\right). \end{aligned} \quad (6.21)$$

However, by (6.5) and (A.5) along with further estimates along the lines leading to (6.15) we have

$$\begin{aligned} \mathbb{E} \left\| \langle \mathcal{K}_{t,T} \phi, Z_j^m(\bar{U}) \rangle \right\|_{C^1([T/2, T]; \mathbb{R})}^{90} &\leq C \exp(\eta/2 \|U_0\|^2) \left( \mathbb{E} \sup_{t \in [T/2, T]} \|[Z_j^m(\bar{U}), F(U)]\|^{180} \right)^{1/2} \\ &\leq C \exp(\eta/2 \|U_0\|^2) |j|^{90 \times 6} (\mathbb{E}(1 + \|U\|_{H^4}^{3 \times 180}))^{1/2} \\ &\leq C \exp(\eta \|U_0\|^2) |j|^{90 \times 6}, \end{aligned} \quad (6.22)$$

where  $C = C(\eta, T)$ . Finally, due to (5.11) and similar applications of (6.5) and (A.5) the estimate

$$\mathbb{E} \left\| \langle \mathcal{K}_{t,T} \phi, [Z_j^m(U), \sigma_k^l] \rangle \right\|_{C^1([T/2, T]; \mathbb{R})}^{90} \leq C \exp(\eta \|U_0\|^2) |j|^{90 \times 2} \quad (6.23)$$

follows. By combining (6.21)-(6.23) we obtain (6.16), and the proof is complete.  $\square$

The final lemma of this section corresponds to brackets of the form  $Y \rightarrow Z \rightarrow [Z, Y]$ . For fixed  $j \in \mathbb{Z}_+^2$ , define  $\mathcal{Z}_j$  as the union of  $j$  with the set of points in  $\mathbb{Z}_+^2$  adjacent to  $j$ , that is,

$$\mathcal{Z}_j := \{k \in \mathbb{Z}_+^2 : k = j \pm m \text{ for some } m \in \{0\} \cup \mathcal{Z}\}.$$

**Lemma 6.9.** Fix  $j \in \mathbb{Z}_+^2$ . For each  $0 < \epsilon < \epsilon_0(T)|j|^{-2}$  and  $\eta > 0$  there exists  $C = C(\eta, T)$  and a measurable set  $\Omega_{\epsilon, j, Q}$  with

$$\mathbb{P}((\Omega_{\epsilon, j, Q})^c) \leq C|j|^{14 \times 5400} \exp(\eta \|U_0\|^2) \epsilon, \quad (6.24)$$

such that on the set  $\Omega_{\epsilon, j, Q}$  it holds that, for every  $\phi \in H$ ,

$$\begin{aligned} \sum_{\substack{i \in \mathcal{Z}_j \\ m \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Y_i^m(U) \rangle| &\leq \epsilon \|\phi\| \\ \Rightarrow \sum_{\substack{k \in \mathcal{Z} \\ m, l \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [Z_j^m(U), Y_k^l(U)] \rangle| &\leq \epsilon^{1/3600} \|\phi\|. \end{aligned} \quad (6.25)$$

*Proof.* By (5.19) it suffices to find  $\Omega_{\epsilon, j, Q}$  satisfying (6.24) such that, on  $\Omega_{\epsilon, j, Q}$ , assuming (6.25), it follows that

$$\sum_{\substack{k \in \mathcal{Z} \\ m, l \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [[Z_j^m(U), \sigma_k^l], F(U)] \rangle| \leq \epsilon^{1/2}, \quad (6.26)$$

$$\sum_{\substack{k \in \mathcal{Z} \\ m, l \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [[Z_j^m(U), F(U)], \sigma_k^l] \rangle| \leq \epsilon^{1/1800}. \quad (6.27)$$

To obtain (6.26), we have by (5.12) and the definition of  $Y_j^m(U)$ , that if (6.25) holds (note  $j \pm k \in \mathcal{Z}_j$  for  $k \in \mathcal{Z}$ ), then

$$\begin{aligned} |\langle \mathcal{K}_{t, T} \phi, [[Z_j^m(U), \sigma_k^l], F(U)] \rangle| &= |g(j^\perp \cdot k)| |\langle \mathcal{K}_{t, T} \phi, [a(j, k) \sigma_{j+k}^{m+l+1} + (-1)^{l+1} b(j, k) \sigma_{j-k}^{m+l+1}, F(U)] \rangle| \\ &\leq C|j| (|\langle \mathcal{K}_{t, T} \phi, Y_{j+k}^{m+l+1}(U) \rangle| + |\langle \mathcal{K}_{t, T} \phi, Y_{j-k}^{m+l+1}(U) \rangle|) \\ &\leq C|j| \epsilon \|\phi\| \end{aligned} \quad (6.28)$$

and (6.26) follows for any  $\epsilon < (C|j|)^{-2}$ .

It remains to prove that (6.25) implies (6.27) on an appropriate set. By Lemma 6.8, there exists a set  $\Omega_{\epsilon, j}^1$  satisfying (6.16) such that on  $\Omega_{\epsilon, j}^1$ , for each  $m \in \{0, 1\}$ , and each  $\phi \in \mathcal{I} = \{\phi \in H : \|\phi\| = 1\}$

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Y_j^m(U) \rangle| \leq \epsilon \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Z_j^m(\bar{U}) \rangle| \leq \epsilon^{1/30}. \quad (6.29)$$

For fixed  $m \in \{0, 1\}$ , and each  $\phi \in \mathcal{I}$ , let  $g_\phi(t) := \langle \mathcal{K}_{t, T} \phi, Z_j^m(\bar{U}) \rangle$  so that  $g'_\phi(t) = \langle \mathcal{K}_{t, T} \phi, [Z_j^m(\bar{U}), F(U)] \rangle$  (see (6.7)). Let  $\Omega_{\epsilon, j}^2 := \Lambda_{\epsilon^{1/30}, 1/4}^c$ , where  $\Lambda_{\epsilon, \alpha}$  is defined in (6.9). Thus on  $\Omega_{\epsilon, j}^3 := \Omega_{\epsilon, j}^1 \cap \Omega_{\epsilon, j}^2$  we have (invoking (6.29)) for each  $m \in \{0, 1\}$ , and  $\phi \in \mathcal{I}$

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Y_j^m(U) \rangle| \leq \epsilon \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [Z_j^m(\bar{U}), F(U)] \rangle| \leq \epsilon^{1/300}. \quad (6.30)$$

Similarly as in the proof of Lemma 6.8,

$$\mathbb{P}((\Omega_{\epsilon, j}^3)^c) \leq C\epsilon \mathbb{E} \left( \sup_{\phi \in \mathcal{I}} \|g'_\phi\|_{C^{1/4}([T/2, T])}^{8 \times 30} \right) \leq C\epsilon |j|^{240 \times 6} \exp(\eta \|U_0\|^2),$$

where the last inequality is analogous to estimates in (6.22). Next, we establish (6.30) with  $Z_j^m(\bar{U})$  replaced by  $Z_j^m(U)$ , for which we use the expansion (6.19). Specifically, as in (6.28)

$$\sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [[Z_j^m(\bar{U}), \sigma_k^l], F(U)] \rangle| \|W^{k, l}(t)\| \leq C|j| \epsilon \|\phi\| \sup_{t \in [T/2, T]} |W^{k, l}(t)|. \quad (6.31)$$

Since  $\mathbb{E}\|W^{k,l}\|_{L^\infty} < \infty$ , Markov inequality yields  $\mathbb{P}((\Omega_{\epsilon,k,l}^4)^c) \leq C\epsilon^{1/2}$ , with  $\Omega_{\epsilon,k,l}^4 := \{\sup_{t \in [T/2, T]} |W^{k,l}(t)| \leq \epsilon^{-1/2}\}$ . By combining (6.19), (6.30), and (6.31) on  $\Omega_{\epsilon,j}^5 := \Omega_{\epsilon,j}^3 \cap \Omega_{\epsilon,j}^4$  for any  $\epsilon < \epsilon_0(T)|j|^{-2}$  it holds that

$$\begin{aligned} & \sum_{\substack{i \in \mathcal{Z}_j \\ m \in \{0,1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T}\phi, Y_i^m(U) \rangle| \leq \epsilon \\ & \Rightarrow \sum_{m \in \{0,1\}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T}\phi, [Z_j^m(U), F(U)] \rangle| \leq \epsilon^{1/600}. \end{aligned}$$

Similarly as in Lemma 6.9, we expand  $[Z_j^m(U), F(U)]$  with respect to  $U = \bar{U} + \sigma W$  and again we use Theorem 7.1 of [HM11] to establish

$$\begin{aligned} & \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T}\phi, [Z_j^m(U), F(U)] \rangle| \leq \epsilon^{1/600} \|\phi\| \\ & \Rightarrow \sup_{l \in \{0,1\}, k \in \mathcal{Z}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T}\phi, [[Z_j^m(\bar{U}), F(\bar{U})], \sigma_k^l] \rangle| \leq \epsilon^{1/1800} \|\phi\| \end{aligned}$$

on a set  $\Omega_\epsilon^6$  satisfying

$$\mathbb{P}((\Omega_\epsilon^6)^c) \leq C|j|^{14 \times 5400} \exp(\eta \|U_0\|^2) \epsilon. \quad (6.32)$$

The proof is finished if we set  $\Omega_{\epsilon,j,Q} := \Omega_{\epsilon,j}^5 \cap \Omega_{\epsilon,j}^6$ , and note by a (5.3) and (5.12) that

$$[[Z_j^m(\bar{U}), F(\bar{U})], \sigma_k^l] = [[Z_j^m(U), F(U)], \sigma_k^l].$$

□

### 6.3 Spanning Sets for $H^N$ from Brackets and Associated Tails

With all elements in Figure 6.1 now established, we explain how the lemmata are pieced together to conclude the proof of Proposition 4.4. To simplify the forthcoming calculations we denote by  $\kappa$  the power of  $\epsilon$ , and  $\tau$  the power of  $|j|$ , appearing in the statement of Lemma 6.9, that is,  $\kappa = 1/3600$  and  $\tau = 14 \times 5400$ . Then assertions of Lemmata 6.5 – 6.9 are of the form: for each  $\phi \in H$  with  $\|\phi\| \leq 1$  and for any sufficiently small  $\epsilon$  one has

$$A_j(\phi) \leq \epsilon \quad \Rightarrow \quad B_j(\phi) \leq \epsilon^\kappa$$

on a set  $\Omega_\epsilon$  with  $\mathbb{P}(\Omega_\epsilon^c) \leq C\epsilon|j|^\tau \exp(\eta \|U_0\|^2)$ , where  $C = C(T)$  and  $A_j, B_j$  are appropriate functionals.

Denote

$$\mathcal{I}_N := \{j \in \mathbb{Z}_+^2 : |j_1| + |j_2| \leq N + 1\} \setminus \{(0, N + 1), (0, N), (N + 1, 0), (N, 0)\}$$

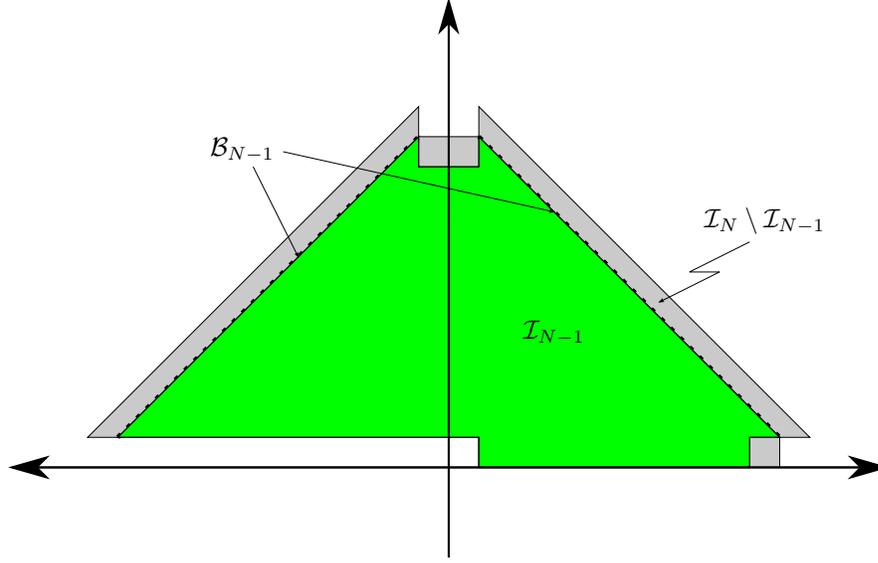
see Figure 6.2. Note that the choice to ‘delete’ the corners of the triangular set  $\mathcal{I}_N$  is to assure that points in  $\mathcal{I}_N \setminus \mathcal{I}_{N-1}$  can be reached from points in  $\mathcal{I}_{N-1}$  using only moves depicted in Figure 5.2.

**Lemma 6.10.** *Let  $\kappa$  be as above, and for every  $N \geq 0$  denote*

$$p_N := \kappa^{2N+2}.$$

*Then there exists*

$$\epsilon_0 = \epsilon_0(N, T) := C(T) \min \left\{ 1, \left( \frac{g}{2N^2} \right)^{2/(\kappa p_{N-1})} \right\},$$



**Figure 6.2:** An illustration of the sets  $I_N$  and the associated induction procedure in Lemma (6.10).

such that for each  $\epsilon \in (0, \epsilon_0)$ , there is a set  $\Omega_{\epsilon, N}$  and  $C = C(T)$  with

$$\mathbb{P}(\Omega_{\epsilon, N}^c) \leq CN^{\tau+2} \exp(\eta \|U_0\|^2) \epsilon^{p_N},$$

such that on  $\Omega_{\epsilon, N}$  for any  $\phi \in H$ ,  $j \in \mathcal{I}_N$ ,  $m \in \{0, 1\}$ , one has

$$\begin{aligned} \langle \mathcal{M}_{0, T} \phi, \phi \rangle &\leq \epsilon \|\phi\|^2 \\ \Rightarrow \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle| &\leq \epsilon^{p_N \kappa} \|\phi\| \quad \text{and} \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, Y_j^m(U) \rangle| \leq \epsilon^{p_N \kappa} \|\phi\|. \end{aligned} \quad (6.33)$$

**Remark 6.11.** Lemma 6.10 establishes the smallness of  $|\langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle|$  for all  $j \in \mathcal{I}_N$ . The smallness of  $|\langle \mathcal{K}_{t, T} \phi, \psi_j^m \rangle|$  requires more work and it is discussed below in Lemma 6.12.

*Proof.* We first remark that the function  $n \mapsto (\frac{q}{n^2})^{2/(\kappa p_{n-1})}$  decreases, and therefore  $\epsilon_0 \leq (\frac{q}{n^2})^{2/(\kappa p_{n-1})}$  for each  $n \leq N$ .

We proceed by induction in  $N \geq 1$ . For the first step,  $N = 1$ , we show that the result holds on the set  $\mathcal{I}_1 = \{(1, 1), (-1, 1)\}$ . To this end we first establish (6.33) for  $j \in \mathcal{Z} = \{(0, 1), (1, 0)\}$ , which are the directly forced modes in (2.14). Indeed, Lemma 6.5 and Lemma 6.6 imply that (6.33) holds for each  $j \in \mathcal{Z}$ ,  $m \in \{0, 1\}$ , with  $p_N$  replaced by  $\kappa$ , on the set

$$\Omega_{\epsilon, 1}^1 = \Omega_{\epsilon, \mathcal{M}} \cap \Omega_{\epsilon^\kappa, (0, 1), Y} \cap \Omega_{\epsilon^\kappa, (1, 0), Y},$$

with  $\mathbb{P}((\Omega_{\epsilon, 1}^1)^c) \leq C \exp(\eta \|U_0\|^2) \epsilon^\kappa$ .

We now establish (6.33) for  $j \in \mathcal{I}_1$ . By (6.33) with  $j' \in \mathcal{Z}$ , and Lemma 6.8, one has

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{j', k \in \mathcal{Z}} \sup_{m, l \in \{0, 1\}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [Z_{j'}^m(U), \sigma_k^l] \rangle| \leq \epsilon^{\kappa^3} \|\phi\|,$$

on a set

$$\Omega_{\epsilon, 1}^2 := \Omega_{\epsilon, 1}^1 \cap \Omega_{\epsilon^{\kappa^2}, (0, 1), \sigma} \cap \Omega_{\epsilon^{\kappa^2}, (1, 0), \sigma}$$

with

$$\mathbb{P}((\Omega_{\epsilon, 1}^2)^c) \leq C \exp(\eta \|U_0\|^2) \epsilon^{\kappa^2}.$$

Using Proposition 5.2 with  $j' = (0, 1)$ ,  $k = (1, 0)$ , and all combinations of  $m, l \in \{0, 1\}$ , we obtain (since  $(j'^\perp \cdot k), a(j', k), b(j', k) \neq 0$ ) that on the set  $\Omega_{\epsilon, 1}^2$ , for each  $j \in \mathcal{I}_1$ ,  $m \in \{0, 1\}$ ,

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle| \leq \frac{1}{2g} \epsilon^{\kappa^3} \|\phi\| \leq \epsilon^{\frac{1}{2}\kappa^3} \|\phi\| \leq \epsilon^{\kappa^4} \|\phi\|.$$

Notice we have used the inequality  $\epsilon \leq (2g)^{2/\kappa^3}$ . The first part of (6.33) follows with  $p_1 = \kappa^3$ . The second part of (6.33) with  $p_1 = \kappa^4$  follows from the first part and Lemma 6.6 on

$$\Omega_{\epsilon, 1} := \Omega_{\epsilon, 1}^2 \cap \Omega_{\epsilon^{\kappa^4}, (-1, 1), Y} \cap \Omega_{\epsilon^{\kappa^4}, (1, 1), Y},$$

so that  $\mathbb{P}(\Omega_{\epsilon, 1}^c) \leq C \exp(\eta \|U_0\|^2) \epsilon^{\kappa^4}$ . This completes the proof of the base case  $N = 1$ .

Next we establish the inductive step. That is, assuming (6.33) holds for each  $j \in \mathcal{I}_{N-1}$  (with  $N-1 \geq 1$ ) on a set  $\Omega_{\epsilon, N-1}$ , we will show that (6.33) holds true for  $j \in \mathcal{I}_N$  on a set  $\Omega_{\epsilon, N}$ . We introduce the set

$$\mathcal{B}_{N-1} := \{j' = (j'_1, j'_2) \in \mathcal{I}_{N-1} : |j'_1|' + |j'_2|' = N\},$$

which is the ‘boundary of  $\mathcal{I}_{N-1}$  excluding the  $x$  and  $y$  axes’ as illustrated by the broken line segments in Figure 6.2. Denote

$$\Omega_{\epsilon, N}^1 := \Omega_{\epsilon, N-1} \cap \bigcap_{j' \in \mathcal{B}_{N-1}} \Omega_{\epsilon^{p_{N-1}}, j', \sigma},$$

then the inductive hypothesis and Lemma 6.8 imply

$$\mathbb{P}((\Omega_{\epsilon, N}^1)^c) \leq C \exp(\eta \|U_0\|^2) ((N-1)^{\tau+2} \epsilon^{p_{N-1}} + |\mathcal{B}_{N-1}| N^\tau \epsilon^{p_{N-1}}) \leq C N^{\tau+2} \exp(\eta \|U_0\|^2) \epsilon^{p_{N-1}}. \quad (6.34)$$

Then on the set  $\Omega_{\epsilon, N}^1$  we have

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{\substack{l, m \in \{0, 1\} \\ j' \in \mathcal{B}_{N-1}, k \in \mathcal{Z}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, [Z_{j'}^m(U), \sigma_k^l] \rangle| \leq \epsilon^{p_{N-1}\kappa} \|\phi\|. \quad (6.35)$$

To complete the inductive step it is enough to establish (6.33) for any fixed  $j \in \mathcal{I}_N \setminus \mathcal{I}_{N-1}$  and  $m \in \{0, 1\}$ . We observe that for each  $j \in \mathcal{I}_N \setminus \mathcal{I}_{N-1}$ , there exists  $j' \in \mathcal{B}_{N-1}$  such that  $k := j - j' \in \mathcal{Z} \cup (-\mathcal{Z})$ . In other words, any point in  $\mathcal{I}_N \setminus \mathcal{I}_{N-1}$  can be reached from  $\mathcal{B}_{N-1}$  via ‘allowable directions’ as shown in Figure 5.2. Since  $k$  is parallel to one of the axes and  $j'$  is not, we have  $k^\perp \cdot j' \neq 0$  and  $a(j', k) \neq 0$ ,  $b(j', k) \neq 0$ , where  $a, b$  are defined by (5.13). Using (6.35) and Proposition 5.2, we infer that on the set  $\Omega_{\epsilon, N}^1$ , for each fixed  $j \in \mathcal{I}_N \setminus \mathcal{I}_{N-1}$ ,  $m \in \{0, 1\}$  (for our choice of  $\epsilon \leq (\frac{g}{2N^2})^{2/(\kappa p_{N-1})}$ )

$$\langle \mathcal{M}_{0, T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t, T} \phi, \sigma_j^m \rangle| \leq \frac{2N^2}{g} \epsilon^{p_{N-1}\kappa} \|\phi\| \leq \epsilon^{p_{N-1}\kappa/2} \|\phi\|, \quad (6.36)$$

where we used that  $|(k \perp j')| \geq 1$ ,  $|a(k, j')| \geq N^{-2}$ , and  $|b(k, j')| \geq N^{-2}$ .<sup>26</sup>

To complete the induction it remains to establish the second part of (6.33). Define

$$\Omega_{\epsilon, N} = \Omega_{\epsilon, N}^1 \cap \bigcap_{j' \in \mathcal{B}_{N-1}} \Omega_{\epsilon^{p_{N-1}\kappa/2}, j', Y}.$$

<sup>26</sup> To see the estimate for  $b$  (estimates for  $a$  are analogous) we observe that

$$|b((1, 0), j')| = 1 - \frac{j_1^2}{j_1(j_1^2 + j_2^2)} \geq 1 - \frac{1}{|j_1|}, \quad |b((0, 1), j')| = \frac{|j_1|}{(j_1^2 + j_2^2)} \geq \frac{1}{1 + (N-1)^2} \geq \frac{1}{N^2},$$

and the desired bound follows when  $|j_1| \geq 2$ . If  $|j_1| = 1$ , then  $b((1, 0), (\pm 1, N-1)) = 1 - \frac{1}{1+(N-1)^2} \geq \frac{1}{2}$ .

We then obtain (analogously to (6.34)) that

$$\mathbb{P}(\Omega_{\epsilon,N}^c) \leq C \exp(\eta \|U_0\|^2) (N^{\tau+2} \epsilon^{p_{N-1\kappa}} + N^{\tau+1} \epsilon^{p_{N-1\kappa}/2}) \leq CN^{\tau+2} \exp(\eta \|U_0\|^2) \epsilon^{p_N}.$$

On  $\Omega_{\epsilon,N}$ , (6.36) and Lemma 6.6 yield that (6.33) holds with the desired  $p_N$ .  $\square$

**Lemma 6.12.** *Fix  $N \geq 2$  and let  $p_N$  and  $\epsilon_0$  be as in Lemma 6.10. There is*

$$\epsilon_1 = \epsilon_1(N, T) := C(T) \min \left\{ 1, \left( \frac{C(T)}{1 + N^4} \right)^{2/(p_{2N\kappa})} \right\}.$$

Then for every  $\epsilon \in (0, \epsilon_1)$  and  $\eta > 0$  there exists a set  $\Omega_{\epsilon,N}^*$  with

$$\mathbb{P}((\Omega_{\epsilon,N}^*)^c) \leq CN^{\tau+3} \exp(\eta \|U_0\|^2) \epsilon^{p_{2N\kappa}},$$

where  $C = C(\eta, T)$ , such that on the set  $\Omega_{\epsilon,N}^*$  for all  $\phi \in H$ ,  $j$  with  $|j| \leq N$ , and  $m \in \{0, 1\}$ , one has

$$\langle \mathcal{M}_{0,T} \phi, \phi \rangle \leq \epsilon \|\phi\|^2 \quad \Rightarrow \quad \begin{cases} |\langle \phi, \sigma_j^m \rangle| \leq \epsilon^{p_{2N}} \|\phi\|, \\ |\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) \rangle| \leq \epsilon^{p_{2N\kappa}/2} \|\phi\|. \end{cases} \quad (6.37)$$

$$(6.38)$$

For the definition of  $J_{j,m}^N$  see (5.25) and (5.23).

*Proof.* Below, without further notice, we use that  $\epsilon_0(2N, T) \geq \epsilon_1(N, T)$  for appropriate  $C(T)$ , where  $\epsilon_0$  is as in Lemma 6.10. Observe that  $|j| \leq N$  implies  $j \in \mathcal{I}_{2N}$ , because

$$N^2 \geq j_1^2 + j_2^2 \geq \frac{1}{2}(|j_1| + |j_2|)^2.$$

Then by Lemma 6.10, for each  $j$  such that  $|j| \leq N$ , on the set  $\Omega_{\epsilon,2N}$ , (6.33) with  $t = T$  implies

$$|\langle \phi, \sigma_j^m \rangle| \leq \epsilon^{p_{2N}} \|\phi\| \quad (6.39)$$

and (6.37) follows.

To establish (6.38) we first fix  $|j| \leq N$  with  $j_1 \neq 0$  and  $m \in \{0, 1\}$ . Then, by Lemma 6.10, (5.24), and (5.25), on the set  $\Omega_{\epsilon,2N}$ ,

$$\begin{aligned} \epsilon^{p_{2N\kappa}} \|\phi\| &\geq |\langle \phi, Y_j^{m+1}(U(T)) \rangle| = g|j_1| |\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) + P_N J_{j,m}(U(T)) \rangle| \\ &\geq g|j_1| |\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) \rangle| - g|j_1| |\langle \phi, P_N J_{j,m}(U(T)) \rangle|. \end{aligned} \quad (6.40)$$

Next, fix  $|j| \leq N$  with  $j_1 = 0$  and set  $j' := \vec{e}_1 + j$ . It is easy to check that  $j'$ ,  $j' \pm \vec{e}_2$ ,  $j' \pm \vec{e}_1$  belong to  $\mathcal{I}_{2N}$  whenever  $N \geq 2$ , so that, by the second part of (6.33), (6.25) is satisfied (with  $j$  replaced by  $j'$ ) on the set  $\Omega_{\epsilon,2N}$ . Then by Lemma 6.9 (the smallness conditions on  $\epsilon$  required by Lemma 6.9 are satisfied if  $\epsilon < \epsilon_1$  for appropriate  $C(T)$ ),

$$\sum_{\substack{k \in \mathcal{Z} \\ m, l \in \{0, 1\}}} \sup_{t \in [T/2, T]} |\langle \mathcal{K}_{t,T} \phi, [Z_{j'}^m(U), Y_k^l(U)] \rangle| \leq \epsilon^{p_{2N\kappa}} \|\phi\|, \quad (6.41)$$

on the set

$$\Omega_{\epsilon,N}^{*,1} := \Omega_{\epsilon,2N} \cap \bigcap_{|j'| \leq N, j_1=1} \Omega_{\epsilon^{p_{2N}}, j', Q},$$

with  $\mathbb{P}((\Omega_{\epsilon,N}^{*,1})^c) \leq C(2N)^{\tau+3} \epsilon^{p_{2N}} \exp(\eta \|U_0\|^2)$ . Then by (5.24) and (6.41) (with  $t = T$ ), on  $\Omega_{\epsilon,N}^{*,1}$  one has

$$2\epsilon^{p_{2N}} \|\phi\| \geq \frac{g^2 |j|^3}{1 + |j|^2} (|\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) \rangle| - |\langle \phi, P_N J_{j,m}(U(T)) \rangle|). \quad (6.42)$$

Combining both cases, (6.40) and (6.42), one has for any  $|j| \leq N$ , on the set  $\Omega_{\epsilon, N}^{*,1}$ , that

$$|\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) \rangle| \leq C(\epsilon^{p_{2N}} \|\phi\| + |\langle \phi, P_N J_{j,m}(U(T)) \rangle|). \quad (6.43)$$

Since by Lemma 5.5 the first component of  $J_{j,m}(U(T))$  vanishes, there exists  $(\beta_k^l)_{|k| \leq N, l \in \{0,1\}}$  such that

$$P_N J_{j,m}(U(T)) = \sum_{\substack{|k| \leq N \\ l \in \{0,1\}}} \beta_k^l \sigma_k^l.$$

Consequently, by (6.39) and Lemma 5.5, on the set  $\Omega_{\epsilon, 2N}$ ,

$$\begin{aligned} |\langle \phi, P_N J_{j,m}(U(T)) \rangle| &= \left| \sum_{\substack{|k| \leq N \\ l \in \{0,1\}}} \beta_k^l \langle \phi, \sigma_k^l \rangle \right| \leq C \sup_{\substack{|k| \leq N \\ l \in \{0,1\}}} |\langle \phi, \sigma_k^l \rangle| \sum_{\substack{|k| \leq N \\ l \in \{0,1\}}} |\beta_k^l| \\ &\leq CN \|J_{j,m}(U(T))\| \epsilon^{p_{2N}} \|\phi\| \leq CN^4 (1 + \|U(T)\|_{H^1}) \epsilon^{p_{2N}} \|\phi\|. \end{aligned} \quad (6.44)$$

If we set  $\hat{\Omega}_{\epsilon, N} := \{1 + \|U(T)\|_{H^1} \leq \epsilon^{-p_{2N}/2}\}$ , then (A.5) and the Markov inequality imply

$$\mathbb{P}(\hat{\Omega}_{\epsilon, N}^c) \leq C \epsilon^{p_{2N}/2} \exp(\eta \|U_0\|^2),$$

where  $C = C(T)$ . Combining (6.43) and (6.44), on  $\Omega_{\epsilon, N}^* := \hat{\Omega}_{\epsilon, N} \cap \Omega_{\epsilon, N}^{*,1}$ , it holds that

$$|\langle \phi, \psi_j^m + J_{j,m}^N(U(T)) \rangle| \leq C \|\phi\| (\epsilon^{p_{2N\kappa}} + N^4 \epsilon^{p_{2N}/2}) \leq C \|\phi\| (1 + N^3) \epsilon^{p_{2N\kappa}} \leq \epsilon^{p_{2N\kappa}/2} \|\phi\|,$$

provided  $\epsilon \leq \epsilon_1$ . Finally, notice that

$$\mathbb{P}((\Omega_{\epsilon, N}^*)^c) \leq \mathbb{P}(\hat{\Omega}_{\epsilon, N}^c) + \mathbb{P}(\Omega_{\epsilon, N}^{*,1c}) \leq CN^{\tau+3} \exp(\eta \|U_0\|^2) \epsilon^{p_{2N\kappa}}.$$

□

*Proof of Proposition 4.4.* By Lemma 6.12 on a set  $\Omega_{\epsilon, \tilde{N}}$  one has

$$\sum_{\tilde{b}(U) \in \mathfrak{B}_{N, \tilde{N}}(U)} |\langle \phi, \tilde{b}(U(T)) \rangle|^2 \leq \sum_{\substack{|j| \leq \tilde{N} \\ m \in \{0,1\}}} |\langle \phi, \sigma_j^m \rangle|^2 + |\langle \phi, \psi_j^m + J_{j,m}^{\tilde{N}}(U(T)) \rangle|^2 \leq \tilde{N}^2 \epsilon^{p_{2\tilde{N}\kappa}} \|\phi\|^2 \leq \epsilon^{p_{2\tilde{N}\kappa}/2} \|\phi\|^2,$$

whenever  $\epsilon \in (0, \epsilon_2)$ .

□

## 7 Mixing and Other Convergence Properties

In this final section we show how the abstract results developed in [HM08, KW12] (and cf. [Shi06, HMS11, KS12]) can be applied in our setting to establish mixing and pathwise attraction properties for the unique invariant measure associated to (2.1)–(2.2) and to thus complete the proof of the main result Theorem 2.3.

We begin by introducing some notations. For any  $r \in (0, 1]$  and any  $\varsigma > 0$  define

$$\rho_r(U_1, U_2) := \inf_{\substack{\gamma \in C^1([0,1], H), \\ \gamma(0) = U_1, \gamma(1) = U_2}} \int_0^1 \exp(\varsigma r \|\gamma(\varrho)\|^2) \|\gamma'(\varrho)\| d\varrho. \quad (7.1)$$

The fixed value of  $\varsigma$  is determined in the course of the proof of Theorem 2.3, (i). As shown in [HM08],  $\rho_r$  is metric on  $H$  for any  $r > 0$  and for any  $U_1, U_2 \in H$  one has

$$\|U_1 - U_2\| \leq \rho_r(U_1, U_2) \leq \exp(\varsigma r \max\{\|U_1\|, \|U_2\|\}) (\|U_1 - U_2\|). \quad (7.2)$$

For brevity of notation we set  $\rho := \rho_1$  and for any  $\Phi : H \rightarrow \mathbb{R}$  define

$$\|\Phi\|_{Lip} := \sup_{U_1 \neq U_2} \frac{|\Phi(U_1) - \Phi(U_2)|}{\rho(U_1, U_2)}.$$

On the other hand, for any  $\mu_1, \mu_2 \in Pr(H)$ , the set of Borelian probability measures on  $H$ , denote

$$\mathcal{C}(\mu_1, \mu_2) := \{\Gamma \in Pr(H \times H) : \Gamma(A \times H) = \mu_1(A), \Gamma(H \times A) = \mu_2(A), \text{ for any } A \in \mathcal{B}(H)\}, \quad (7.3)$$

where  $\mathcal{B}(H)$  is a collection of all Borel subsets of  $H$ . Elements  $\Gamma \in \mathcal{C}(\mu_1, \mu_2)$  are typically referred to as a *coupling* of  $\mu_1, \mu_2$ . The distance  $\rho$  defined by (7.1) can be used to induce a Wasserstein-Kantorovich distance (see e.g. [GS02, HM08, KS12] for further details) on the set

$$Pr_1(H) := \left\{ \mu \in Pr(H) : \int_H \rho(0, u) d\mu(u) < \infty \right\}.$$

This distance is defined by two equivalent formulas

$$\rho(\mu_1, \mu_2) := \sup_{\|\Phi\|_{Lip} \leq 1} \left| \int_H \Phi(U) d\mu_1(U) - \int_H \Phi(U) d\mu_2(U) \right| = \inf_{\Gamma \in \mathcal{C}(\mu_1, \mu_2)} \int_H \rho(U_1, U_2) d\Gamma(U_1, U_2). \quad (7.4)$$

While the metric defined in (7.4) is a useful for proving contraction properties of  $\{P_t^*\}_{t \geq 0}$ , it is less transparent for applications involving observables on  $H$ . As such we consider the class of observables  $\mathcal{O}_\varsigma$  defined in (2.18). By [HM08, Proposition 4.1],  $\|\Phi\|_{Lip} \leq C\|\Phi\|_\varsigma$ , and therefore  $\|\Phi\|_{Lip} < \infty$  for  $\Phi \in \mathcal{O}_\varsigma$ .

Next, we recall in our setting abstract results from [HM08, Theorem 3.4, Theorem 4.5] and from [KW12] (and cf. [Shi06, KS12]).

**Theorem 7.1** (Hairer-Mattingly, [HM08]). *Suppose  $U = U(t, U_0)$  is a stochastic (semi)flow on a Hilbert space  $H$  with a  $C^1$  dependence on  $U_0 \in H$ . Define the Markov semigroups  $\{P_t\}_{t \geq 0}$ ,  $\{P_t^*\}_{t \geq 0}$  associated to  $U(t, U_0)$  as in (2.20), (2.21) and assume that there exists  $\varsigma > 0$  such that*

(a) *there exists  $C > 0$  and a decreasing function  $\xi : [0, 1] \rightarrow [0, 1]$  with  $\xi(1) < 1$  such that<sup>27</sup>*

$$\mathbb{E}(\exp(r\varsigma\|U(t, U_0)\|^2)(1 + \|\nabla_{U_0} U(t, U_0)\|)) \leq C \exp(r\varsigma\xi(t)\|U_0\|^2)$$

*for every  $U_0 \in H$ ,  $r \in [1/4, 3]$  and  $t \in [0, 1]$ .*

(b) *a gradient estimate on the Markov semigroup, (2.29), holds for  $\eta = \varsigma/2$ .*

(c) *given any  $\mathfrak{J} > 0$ ,  $r \in (0, 1)$ , and  $\epsilon > 0$ , there exists  $T^* = T^*(\mathfrak{J}, r, \epsilon, \varsigma)$  such that for any  $T > T^*$ ,*

$$\inf_{\|U_1\|, \|U_2\| \leq \mathfrak{J}} \sup_{\Gamma \in \mathcal{C}(P_T^* \delta_{U_1}, P_T^* \delta_{U_2})} \Gamma\{(U', U'') \in H \times H : \rho_r(U', U'') < \epsilon\} > 0. \quad (7.5)$$

*Here,  $\delta_U$  is the Dirac measure concentrated at  $U$  and  $\mathcal{C}(\delta_{U_1}, \delta_{U_2})$  is defined in (7.3).*

*Then there exist positive constants  $C, \gamma > 0$  such that*

$$\rho(P_t^* \mu_1, P_t^* \mu_2) \leq C \exp(-\gamma t) \rho(\mu_1, \mu_2), \quad (7.6)$$

*for every  $\mu_1, \mu_2 \in Pr_1(H)$  and every  $t > 0$ . Moreover, there exists a unique invariant measure  $\mu_*$  (that is  $P_t^* \mu_* = \mu_*$  for every  $t \geq 0$ ) and*

$$\left| \mathbb{E} \Phi(U(t, U_0)) - \int_H \Phi(\bar{U}) d\mu_*(\bar{U}) \right| \leq C \exp(-\gamma t + \varsigma \|U_0\|^2) \|\Phi - \int \Phi d\mu_*\|_\eta \quad (7.7)$$

*which holds for every  $U_0 \in H$  and every  $\Phi \in \mathcal{O}_\varsigma$  (cf. (2.18)).*

<sup>27</sup>The statement of this result in [HM08] is slightly more general and involves the use of Lyapunov functions  $V$ . Here we simply set  $V(x) = \exp(\varsigma|x|^2)$ .

**Theorem 7.2** (Komorowski-Walczuk [KW12]). *Let  $\{P_t\}_{t \geq 0}$  be a Feller Markov semigroup on a metric space  $(H, \rho)$  with the continuity property:  $\lim_{t \rightarrow 0} P_t \Phi(U_0) = \Phi(U_0)$  for all  $\Phi \in C_b(H)$ ,  $U_0 \in H$ . Let  $P_t(U_0, A)$  be the associated transition functions (cf. (2.19)). Suppose that for some  $C, \gamma > 0$  the contraction property (7.6) holds for every  $\mu_1, \mu_2 \in Pr_1(H)$ . Assume moreover that, for every  $R > 0$ <sup>28</sup>*

$$\sup_{t \geq 0} \sup_{U_0 \in B_R} \int_H [\rho(0, U)]^3 P_t(U_0, dU) < \infty, \quad (7.8)$$

where  $B_R := \{U_0 \in H : \rho(0, U_0) < R\}$ . Then, there exists a unique invariant probability measure  $\mu_* \in Pr_1(H)$  such that, for any  $\Phi \in C^1(H)$  and any  $U_0 \in H$

$$\frac{1}{T} \int_0^T \Phi(U(t, U_0)) dt \rightarrow \int_H \Phi(U) d\mu_*(U) \quad \text{in probability.}$$

Moreover, the limit  $\sigma^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left( \int_0^T (\Phi(U(t, U_0)) dt - \int_H \Phi(U) d\mu_*(U)) dt \right)^2$  exists and

$$\lim_{T \rightarrow \infty} \mathbb{P} \left( \frac{1}{\sqrt{T}} \int_0^T \left( \Phi(U(t, U_0)) - \int_H \Phi(U) d\mu_*(U) \right) dt < \xi \right) = \mathcal{X}_\sigma(\xi),$$

where  $\mathcal{X}_\sigma$  is the distribution function of a normal random variable with mean zero and variance  $\sigma^2$ .

Using Theorem 7.1, 7.2 we now establish the attraction properties (i)–(iii) to complete the proof of Theorem 2.3.

*Proof of Theorem 2.3.* We begin by establishing the conditions for (a)–(c) of Theorem 7.1. To prove (a), note that for any  $\varsigma, r > 0$ , (A.12) with  $\eta = r\varsigma \frac{\kappa}{4} e^{-\kappa/4}$  implies we have for any  $t \in [0, 1]$

$$\|\nabla_{U_0} U(t, U_0)\| = \|\mathcal{J}_{0,t}\| \leq C \exp \left( r\varsigma \frac{\kappa}{4} e^{-\kappa/4} \int_0^t \|U(s)\|_{H^1}^2 ds \right) \leq C \exp \left( r\varsigma \frac{\kappa}{4} e^{-\kappa t/4} \int_0^t \|U(s)\|_{H^1}^2 ds \right),$$

where  $C = C(r, \varsigma)$  and  $\kappa = \min\{\nu_1, \nu_2\}$ . As such we have established (a) with  $\xi(t) := \exp(-t\kappa/2)$  follows from (A.2) and for any  $\varsigma < \eta^*/3$ . Set  $\varsigma := \eta^*/6$ . Since Proposition 2.6 holds for any  $\eta > 0$  we infer the second condition (b) for the given  $\varsigma$ .

To establish (c), (7.5), observe that, for any  $\mathfrak{J} > 0$  and  $\epsilon > 0$  there exists  $T_* = T_*(\mathfrak{J}, \epsilon) \geq 0$  such that

$$\inf_{\|U_0\| \leq \mathfrak{J}} P_T(U_0, \{U \in H; \|U\| \leq \epsilon\}) > 0, \quad (7.9)$$

for every  $T > T_*$ . A detailed proof of (7.9) which applies in our setting can be found in [EM01, CGHV13], here we just briefly sketch the essential ideas. If there is no forcing, that is, if there is no Brownian motion, then by the dissipativity of (2.1)–(2.2), there is a  $T^* = T^*(\|U_0\|, \epsilon)$  such that  $\|U(T, U_0)\| < \epsilon/2$ . However, for any open ball  $B(\delta) := \{x \in \mathbb{R}^{2|\mathcal{Z}|} : |x| < \delta\}$  there is a non-zero probability that the Brownian motion remains in  $B(\delta)$  over the whole interval  $[0, T]$ . Then by the continuous dependence of solutions with respect to external forcing we conclude (7.9) for sufficiently small  $\delta > 0$ . More precisely we can use the change of variable  $\tilde{U} = U - \sigma W$  and standard estimates to show that  $\|U(T, U_0)\| < \epsilon$ .

We now establish (7.5) from (7.9) as follows. For  $U_1, U_2 \in H$  and  $T > 0$  define  $\tilde{\Gamma}_{U_1, U_2} \in Pr(H \times H)$  by

$$\tilde{\Gamma}_{U_1, U_2}(A_1 \times A_2) := P_T(U_1, A_1) P_T(U_2, A_2) \text{ for any } A_1, A_2 \in \mathcal{B}(H).$$

<sup>28</sup>The condition (7.8) given here is slightly stronger than the conditions (which appear as H2, H3) given for the results appearing in [KW12].

From (7.3) it follows that  $\tilde{\Gamma}_{U_1, U_2} \in \mathcal{C}(P_T^* \delta_{U_1}, P_T^* \delta_{U_2})$ , that is,  $\tilde{\Gamma}_{U_1, U_2}$  couples  $P_T^* \delta_{U_1}$  and  $P_T^* \delta_{U_2}$ . As such, using (7.2) and (7.9), we infer that, for any  $\mathfrak{J}, \epsilon > 0$ ,  $r \in (0, 1)$ , and any  $T > T^*(\mathfrak{J}, \min\{\epsilon \exp(-3\varsigma)/2, 1\})$

$$\begin{aligned} & \inf_{\|U_1\|, \|U_2\| \leq \mathfrak{J}} \sup_{\Gamma \in \mathcal{C}(P_T^* \delta_{U_1}, P_T^* \delta_{U_2})} \Gamma\{(U', U'') \in H \times H : \rho_r(U', U'') < \epsilon\} \\ & \geq \inf_{\|U_1\|, \|U_2\| \leq \mathfrak{J}} \tilde{\Gamma}_{U_1, U_2} \{(U', U'') \in B_1 \times B_1 : \|U'\| + \|U''\| < \epsilon \exp(-\varsigma r)\} \\ & \geq \left( \inf_{\|U_1\| \leq \mathfrak{J}} P_T(U_1, \{U' \in H : \|U'\| < \min(\epsilon/2 \cdot \exp(-3\eta), 1)\}) \right)^2 > 0, \end{aligned}$$

where  $B_1 := \{U \in H : \|U\| < 1\}$  and (7.5) follows.

Having established the conditions (a)–(c) in Theorem 7.1 we infer (7.6) and (7.7) for (2.1)–(2.2). To prove (2.22) it suffices to show that  $|\int \Phi d\mu_*| \leq C \|\Phi\|_\varsigma$ , where  $\mu_*$  is the unique invariant measure of (2.1)–(2.2). However,

$$\begin{aligned} \left| \int \Phi(U) d\mu_*(U) \right| & \leq \int \exp(\varsigma \|U\|^2) \exp(-\varsigma \|U\|^2) |\Phi(U)| d\mu_*(U) \\ & \leq \|\Phi\|_\varsigma \int_H \exp(\varsigma \|U\|^2) d\mu_*(U), \end{aligned}$$

and therefore it suffices to show

$$\int_H \exp(\varsigma \|U\|^2) d\mu_*(U) \leq C < \infty, \quad (7.10)$$

where  $\varsigma = \eta^*/6$ . For any  $R > 0$ , define

$$\Phi_R(U) = \begin{cases} \exp(\varsigma \|U\|^2) & \text{for } \|U\| < R, \\ \exp(\varsigma R^2) & \text{for } \|U\| \geq R, \end{cases}$$

and note that  $\Phi_R \in C_b(H)$ . Now using that  $P_t^* \mu_* = \mu_*$  and (A.2) we have for any  $T > 0$ ,  $\mathfrak{J} > 0$

$$\begin{aligned} \left| \int \Phi_R(U) d\mu_*(U) \right| & \leq \left| \int_{B_2(H)} P_T \Phi_R(U) d\mu_*(U) \right| + \left| \int_{B_2(H)^c} P_T \Phi_R(U) d\mu_*(U) \right| \\ & \leq \left| \int_{B_2(H)} \mathbb{E} \exp(\varsigma \|U(t, U_0)\|^2) d\mu_*(U_0) \right| + \exp(\varsigma R^2) \mu_*(H \setminus B_2(H)) \\ & \leq C \left| \int_{B_2(H)} \exp\left(\varsigma e^{-\frac{T}{2} \min\{\nu_1, \nu_2\}} \|U_0\|^2\right) d\mu_*(U_0) \right| + \exp(\varsigma R^2) \mu_*(H \setminus B_2(H)) \\ & \leq C \exp\left(\varsigma e^{-\frac{T}{2} \min\{\nu_1, \nu_2\}} \mathfrak{J}^2\right) + \exp(\varsigma R^2) \mu_*(H \setminus B_2(H)). \end{aligned}$$

Now since  $T \geq 0$  is arbitrary we infer that for  $\mathfrak{J}, R > 0$  that

$$\left| \int \Phi_R(U) d\mu_*(U) \right| \leq C + \exp(\varsigma R^2) \mu_*(H \setminus B_2(H)),$$

where  $C$  is independent of  $\mathfrak{J}$ . Passing  $\mathfrak{J} \rightarrow \infty$  and then  $R \rightarrow \infty$  and using the monotone convergence theorem we conclude (7.10) and (2.22) follows.

The remaining convergence properties (2.23), (2.24) follow once we show that the conditions for Theorem 7.2 are met. The Feller property and stochastic continuity of  $P_t$  follow immediately from the well-posedness properties of (2.1)–(2.2) as recalled above in Proposition 2.2. It remains to verify the bound in (7.8).

By (7.2) and (A.2) we have for already fixed  $\varsigma = \eta^*/6$  and for any  $U_0 \in H$  and any  $t \geq 0$  that

$$\begin{aligned} \int [\rho(0, U)]^3 P_t(U_0, dU) &\leq \int \exp(3\varsigma \|U\|^2) \|U\|^3 P_t(U_0, dU) \leq C \int \exp(\eta^* \|U\|^2) P_t(U_0, dU) \\ &= C \mathbb{E}(\exp(\eta^* \|U(t, U_0)\|^2)) \leq C \exp(\eta^* \|U_0\|^2), \end{aligned}$$

where the constant  $C$  is independent of  $U_0$  and  $t \geq 0$ . Thus, in view of (7.2), we infer the bound (7.8) and hence the convergences (2.23), (2.24). This completes the proof of Theorem 2.3.  $\square$

## A Appendix

Section A.1 collects various moment estimates for (2.1)–(2.2). In Section A.2 we briefly review of some aspects of the Malliavin calculus relevant to our analysis above.

### A.1 Moment Estimates

In this section we provide details for the moments bound used throughout the manuscript. As above the dependence on physical parameter in constants is suppressed in what follows, see Remark 2.1. Denote

$$\zeta^* := \frac{\nu_1 \nu_2}{g^2}.$$

Then, cf. (2.4),

$$\|U\|^2 = \zeta^* \|\omega\|_{L^2}^2 + \|\theta\|_{L^2}^2, \quad \|U\|_{H^1}^2 = \zeta^* \|\nabla \omega\|_{L^2}^2 + \|\nabla \theta\|_{L^2}^2, \quad (\text{A.1})$$

Also recall, that our domain is  $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z}^2)$ , and therefore the Poincarè inequality takes the form  $\|U\|_H \leq \|U\|_{H^1}$ .

Most of the forthcoming bounds have previously been obtained in the context of the stochastic Navier-Stokes equations and some other nonlinear SPDEs with a dissipative (parabolic) structure. In order to modify them for the Boussinesq system, we need to compensate for the ‘buoyancy’ term  $g\partial_x\theta$  when carrying out energy estimates. This is accomplished by differently weighting the temperature and momentum equations.<sup>29</sup> We illustrate this strategy in the proof of (A.2); proofs of other estimates use the same approach in combination with a straightforward modification of existing methods for the (stochastic) Navier-Stokes equation (see e.g. [HM06, Deb11, KS12]) and they are omitted.

In the first lemma we state a priori bounds on  $U$ . These estimates reflect parabolic type regularization properties of (2.1)–(2.2), and are particularly useful for obtaining spectral bounds on the Malliavin matrix carried out in Section 4–6.

**Lemma A.1.** *Fix any  $U_0 \in H$  and let  $U(\cdot) = U(\cdot, U_0)$  be the unique solution of (2.1)–(2.2) with  $U(0) = U_0$ . Denote  $\kappa := \min\{\nu_1, \nu_2\}$ . There exists  $\eta^* > 0$  such that:*

(i) *For any  $T > 0$  and  $\eta \in (0, \eta^*]$ ,*

$$\mathbb{E} \exp \left( \eta \|U(T)\|^2 + \eta \frac{\kappa}{4} e^{-\kappa T/4} \int_0^T \|U(t)\|_{H^1}^2 dt \right) \leq C \exp \left( \eta e^{-\kappa T/2} \|U_0\|^2 \right), \quad (\text{A.2})$$

$$\mathbb{E} \exp \left( \sup_{\tau \in [0, T]} \eta \|U(\tau)\|^2 + \eta \frac{\kappa}{2} \int_0^T \|U(t)\|_{H^1}^2 dt \right) \leq C \exp \left( \eta \|U_0\|^2 + \eta \|\sigma_\theta\|^2 T \right), \quad (\text{A.3})$$

for a constant  $C$  independent of  $T$ .

<sup>29</sup>As noted in the introduction, we are considering the so called ‘HRB approximation’ in which the Boussinesq equation is considered with periodic boundary conditions after subtracting off the temperature differential profile, see [CRG+06]. In our setup the temperature differential is zero, and therefore the dissipativity properties we derived here do not contradict the situation illuminated in [CRG+06]; we can exclude the possibility of ‘grow-up’ solutions.

(ii) For any  $N > 0$  and  $\eta \in (0, \eta^*]$ ,

$$\mathbb{E} \exp \left( \eta \sum_{k=0}^N \|U(k)\|^2 \right) \leq \exp(\varrho \eta \|U_0\|^2 + \varkappa N), \quad (\text{A.4})$$

where  $\varrho, \varkappa > 0$  are positive constants independent of  $N$  and  $U_0$ .

(iii) For any  $s \geq 0$ ,  $p \geq 2$ , and  $\eta \in (0, \eta^*]$  there exists  $C = C(\eta, s, T, p)$  such that

$$\mathbb{E} \left( \sup_{t \in [T/2, T]} \|U(t)\|_{H^s}^p \right) \leq C \exp(\eta \|U_0\|^2). \quad (\text{A.5})$$

(iv) For any  $p \geq 2$ ,  $s \geq 0$ ,  $\eta > 0$ , and  $T > 0$  there is  $C = C(s, T, p, \eta)$  such that

$$\mathbb{E} \left( \|U\|_{C^{\frac{1}{4}}([T/2, T], H^s)}^p \right) \leq C \exp(\eta \|U_0\|^2). \quad (\text{A.6})$$

*Proof.* For any  $f \in H^1$  denote

$$\|f\|_{D(A^{1/2})}^2 := \zeta^* \nu \|\nabla f_1\|_{L^2}^2 + \mu \|\nabla f_2\|_{L^2}^2.$$

From (2.1)–(2.2) we have

$$d\|\omega\|_{L^2}^2 + 2\nu_1 \|\nabla \omega\|_{L^2}^2 dt = 2g \langle \partial_x \theta, \omega \rangle dt, \quad (\text{A.7})$$

$$d\|\theta\|_{L^2}^2 + 2\nu_2 \|\nabla \theta\|_{L^2}^2 dt = \|\sigma_\theta\|^2 dt + 2\langle \sigma_\theta, \theta \rangle dW. \quad (\text{A.8})$$

Now we weight differently the equations (note the difference to the Navier-Stokes equation). Multiplying (A.7) by  $\zeta^*$  and adding to (A.8) we obtain (recall (A.1)) for any  $\eta > 0$

$$d(\eta \|U\|^2) = \eta(2g\zeta^* \langle \partial_x \theta, \omega \rangle + \|\sigma_\theta\|^2 - 2\|U\|_{D(A^{1/2})}^2) dt + 2\eta \langle \sigma_\theta, \theta \rangle dW. \quad (\text{A.9})$$

Since by Poincarè inequality

$$\begin{aligned} 2\zeta^* g |\langle \partial_x \theta, \omega \rangle| &\leq \nu_1 \zeta^* \|\omega\|_{L^2}^2 + \nu_2 \|\nabla \theta\|_{L^2}^2 \leq \nu_1 \zeta^* \|\nabla \omega\|_{L^2}^2 + \nu_2 \|\nabla \theta\|_{L^2}^2 = \|U\|_{D(A^{1/2})}^2, \\ \kappa \|U\|^2 &\leq \|U\|_{D(A^{1/2})}^2, \end{aligned} \quad (\text{A.10})$$

we have for  $Z(t) := \frac{\eta}{\kappa} \|U(t)\|_{D(A^{1/2})}^2$  and  $V(t) := \eta \|U(t)\|^2$  that  $V \leq Z$  and

$$\begin{aligned} \eta(2g\zeta^* \langle \partial_x \theta, \omega \rangle + \|\sigma_\theta\|^2 - 2\|U\|_{D(A^{1/2})}^2) &\leq \eta \|\sigma_\theta\|^2 - \kappa Z, \\ 4\eta^2 |\langle \sigma_\theta, \theta \rangle|^2 &\leq 4\eta^2 \|\sigma_\theta\|^2 \|\theta\|_{L^2}^2 \leq 4\eta^2 \|\sigma_\theta\|^2 \|U\|^2 \leq 4\eta \|\sigma_\theta\|^2 Z. \end{aligned}$$

Thus, by [HM08, Lemma 5.1], for any  $\eta \in (0, \kappa/(4\|\sigma_\theta\|^2))$  one has

$$\mathbb{E} \exp \left( \eta \|U(T)\|^2 + \frac{\kappa e^{-\kappa T/4}}{4} \int_0^T Z(s) ds \right) \leq C \exp \left( \eta e^{-\kappa T/2} \|U_0\|^2 \right).$$

Now, (A.2) follows from  $\kappa Z(s) = \eta \|U(s)\|_{D(A^{1/2})}^2 \geq \eta \kappa \|U(s)\|_{H^1}^2$ .

The estimates (A.3), (A.4) follow similarly as in [HM06, proof of Lemma 4.10] and (A.5) follows as in [KS12, Proposition 2.4.12]. Finally, (A.6) follows from (A.5) and the fact that  $\|W_k^t\|_{C_{[T/2, T]}^{1/4}}$  has finite  $p$ -th moment for any  $p \geq 1$ , see [BBAK92].  $\square$

**Remark A.2.** In the following estimates  $\eta$  appears only on the right hand side, and therefore they remain valid if  $\eta$  is increased, thus we do not assume any upper bound on  $\eta$ .

Next lemmata collect estimates on linearizations of (2.1)–(2.2). Recall the definitions of the operators  $\mathcal{J}_{0,t}$  and its adjoint  $\mathcal{K}_{0,t}$  given in (3.2) and (3.12) respectively. Moreover, for any  $t \geq s \geq 0$  let  $\mathcal{J}_{s,t}^{(2)} : H \rightarrow \mathcal{L}(H, \mathcal{L}(H))$  be the second derivative of  $U$  with respect to an initial condition  $U_0$ . Observe that for fixed  $U_0$  and any  $\xi, \xi' \in H$  the function  $\rho(t) := \mathcal{J}_{s,t}^{(2)}(\xi, \xi')$  is the solution of

$$\partial_t \rho + A\rho + \nabla B(U)\rho + \nabla B(\mathcal{J}_{s,t}\xi)\mathcal{J}_{s,t}\xi' = G\rho, \quad \rho(s) = 0. \quad (\text{A.11})$$

**Lemma A.3.** For each  $\eta > 0$  and  $0 < s < t$ , we have the pathwise estimate

$$\|\mathcal{J}_{s,t}\| \leq \exp\left(\eta \int_s^t \|U(s)\|_{H^1}^2 ds + C(t-s)\right), \quad (\text{A.12})$$

where  $C = C(\eta)$  is independent of  $s, t$ . Moreover, for each  $\tau \leq T$ ,  $p \geq 1$ , and  $\eta > 0$  there is  $C = C(\eta, T - \tau, p)$  such that

$$\mathbb{E} \sup_{s < t \in [\tau, T]} \|\mathcal{J}_{s,t}\|^p \leq C \exp(\eta \|U_0\|^2), \quad (\text{A.13})$$

$$\mathbb{E} \sup_{s < t \in [\tau, T]} \|\mathcal{K}_{s,t}\|^p \leq C \exp(\eta \|U_0\|^2), \quad (\text{A.14})$$

$$\mathbb{E} \sup_{s < t \in [\tau, T]} \|\mathcal{J}_{s,t}^{(2)}\|^p \leq C \exp(\eta \|U_0\|^2). \quad (\text{A.15})$$

*Proof.* Proof of (A.12) follows along the same lines as [HM06, Lemma 4.10.3]. By taking expectation, (A.13) follows from (A.12), (A.2), and (A.3). Finally, (A.14) follows from (A.13) by duality and (A.15) is similar to [HM06, Lemma 4.10.4].  $\square$

The next lemma provides us with estimates to initial time in a weak norm, which allows us to avoid some technical arguments in Section 6.2 (cf. [HM11]).

**Lemma A.4.** For any  $p \geq 2$ ,  $T > 0$ , and  $\eta > 0$  there is  $C = C(p, T, \eta)$  such that

$$\mathbb{E} \sup_{t \in [T/2, T]} \|\partial_t \mathcal{K}_{t,T} \xi\|_{H^{-2}}^p \leq C \exp(\eta \|U_0\|^2) \|\xi\|^p. \quad (\text{A.16})$$

*Proof.* Recall that  $\rho^* = \mathcal{K}_{t,T} \xi$  solves (3.12) and notice that  $\|B(U', U'')\| \leq \|U'\|_{H^1} \|U''\|_{H^1}$  for any  $U', U'' \in H^1$ . Since  $\|A\rho^*\|_{H^{-2}} \leq \|\rho^*\|$  and

$$\begin{aligned} \|(\nabla B(U))^* \rho^* - (\nabla G(U))^* \rho^*\|_{H^{-2}} &\leq \sup_{\|\psi\|_{H^2} \leq 1} (|\langle (\nabla B(U))^* \rho^*, \psi \rangle| + |\langle (\nabla G(U))^* \rho^*, \psi \rangle|) \\ &\leq \sup_{\|\psi\|_{H^2} \leq 1} (|\langle \rho^*, B(U, \psi) \rangle| + |\langle \rho^*, B(\psi, U) \rangle| + |\langle \rho^*, \nabla G(U)\psi \rangle|) \\ &\leq \|\rho^*\| \sup_{\|\psi\|_{H^2} \leq 1} (2\|U\|_{H^1} \|\psi\|_{H^1} + |g| \|\psi\|_{H^1}) \\ &\leq C \|\rho^*\| (\|U\|_{H^1} + 1), \end{aligned}$$

then (A.16) follows from (A.14) and (A.5).  $\square$

The next lemma is a version of the Foias-Prodi estimate, [FP67], used in this work. Specifically the estimate (A.17) is employed in the decay estimate (3.16).

**Lemma A.5.** *For every  $p \geq 1$ ,  $T > 0$ ,  $\delta, \gamma > 0$  there exists  $N_* = N_*(p, T, \delta, \gamma)$ , such that for any  $N \geq N_*$  one has (recall that  $Q_N$  was defined in (2.9))*

$$\begin{aligned}\mathbb{E}\|Q_N \mathcal{J}_{0,T}\|^p &\leq \gamma \exp(\delta \|U_0\|^2), \\ \mathbb{E}\|\mathcal{J}_{0,T} Q_N\|^p &\leq \gamma \exp(\delta \|U_0\|^2).\end{aligned}\tag{A.17}$$

*Proof.* The proof is analogous to [HM06, Lemma 4.17] and it essentially follows from the fact that  $A$  has compact resolvent (see [HM11, proof of Theorem 8.1]).  $\square$

We next present estimates on the operators  $\mathcal{A}_{s,t}$  and the inverse of the regularized Malliavin matrix that are primarily used in Section 3.4.

**Lemma A.6.** *For  $0 < s < t$  define  $\mathcal{A}_{s,t}$ ,  $\mathcal{A}_{s,t}^*$ , and  $\mathcal{M}_{s,t}$  by to (3.5), (3.11), and (3.13) respectively. Then*

$$\|\mathcal{A}_{s,t}\|_{\mathcal{L}(L^2([s,t];\mathbb{R}^d), H)} \leq C \left( \int_s^t \|\mathcal{J}_{r,t}\|^2 dr \right)^{1/2}\tag{A.18}$$

for a constant  $C$  independent of  $s, t$ . Moreover,

$$\|\mathcal{A}_{s,t}^* (\mathcal{M}_{s,t} + I\beta)^{-1/2}\|_{\mathcal{L}(H, L^2([s,t];\mathbb{R}^d))} \leq 1,\tag{A.19}$$

$$\|(\mathcal{M}_{s,t} + I\beta)^{-1/2} \mathcal{A}_{s,t}\|_{\mathcal{L}(L^2([s,t];\mathbb{R}^d), H)} \leq 1,\tag{A.20}$$

$$\|(\mathcal{M}_{s,t} + I\beta)^{-1/2}\|_{\mathcal{L}(H, H)} \leq \beta^{-1/2}.\tag{A.21}$$

Here,  $\mathcal{L}(X, Y)$  denotes the operator norm of the linear map between the given Hilbert spaces  $X$  and  $Y$ .

*Proof.* The first bound (A.18) follows from the definition of  $\mathcal{A}_{s,t}$  and Hölder's inequality. Next, since  $\mathcal{M}_{s,t}$  is self-adjoint

$$\begin{aligned}\|(\mathcal{M}_{s,t} + I\beta)^{1/2} U\|_H^2 &= \langle (\mathcal{M}_{s,t} + I\beta)^{1/2} U, (\mathcal{M}_{s,t} + I\beta)^{1/2} U \rangle = \langle (\mathcal{M}_{s,t} + I\beta) U, U \rangle \\ &= \beta \|U\|_H^2 + \langle \mathcal{A}_{s,t} \mathcal{A}_{s,t}^* U, U \rangle = \beta \|U\|_H^2 + \|\mathcal{A}_{s,t}^* U\|_{L^2([s,t];\mathbb{R}^d)}^2\end{aligned}$$

for any  $U \in H$ . Setting  $U = (\mathcal{M}_{s,t} + I\beta)^{-1/2} V$  ( $(\mathcal{M}_{s,t} + I\beta)^{-1/2}$  is invertible) we immediately obtain (A.19) and (A.21). Finally, (A.20) follows from (A.19) by duality.  $\square$

For the ‘cost of control’ bounds (3.9) on  $v$  in Section 3.4 we also made use of bounds on the Malliavin derivative of the random operators  $\mathcal{J}_{s,t}$ ,  $\mathcal{A}_{s,t}$ , and  $\mathcal{A}_{s,t}^*$  for  $0 < s < t$ . Observe that for  $\tau \leq t$  (see [HM11])

$$\mathfrak{D}_\tau^j \mathcal{J}_{s,t} \xi = \begin{cases} \mathcal{J}_{\tau,t}^{(2)}(\sigma_\theta e_j, \mathcal{J}_{s,\tau} \xi) & \text{if } s \leq \tau, \\ \mathcal{J}_{s,t}^{(2)}(\mathcal{J}_{\tau,s} \sigma_\theta e_j, \xi) & \text{if } s > \tau. \end{cases}\tag{A.22}$$

We refer the reader to Appendix A.2 for further details on the Malliavin derivative operator  $\mathfrak{D}$  and the associated spaces  $\mathbb{D}^{1,p}$  on which it acts.

**Lemma A.7.** *For any  $0 \leq s < t$  the random operators  $\mathcal{J}_{s,t}$ ,  $\mathcal{A}_{s,t}$ ,  $\mathcal{A}_{s,t}^*$  are differentiable in the Malliavin sense. Moreover, for any  $\eta > 0$  and  $p \geq 1$  we have the bounds*

$$\mathbb{E}\|\mathfrak{D}_\tau^j \mathcal{J}_{s,t}\|^p \leq C \exp(\eta \|U_0\|^2),\tag{A.23}$$

$$\mathbb{E}\|\mathfrak{D}_\tau^j \mathcal{A}_{s,t}\|_{\mathcal{L}(L^2([s,t];\mathbb{R}^d), H)}^p \leq C \exp(\eta \|U_0\|^2),\tag{A.24}$$

$$\mathbb{E}\|\mathfrak{D}_\tau^j \mathcal{A}_{s,t}^*\|_{\mathcal{L}(H, L^2([s,t];\mathbb{R}^d))}^p \leq C \exp(\eta \|U_0\|^2),\tag{A.25}$$

where  $C = C(\eta, p)$ .

*Proof.* The proof of (A.23)–(A.25) is based on the observation (A.22) and the bound (A.15). Further details can be found in [HM06].  $\square$

## A.2 Some Elements of the Malliavin Calculus

In this section, we recall in our context and notations some elements of the Malliavin calculus used in above. For further general background on this vast subject see, for example, [Bel87, Mal97, Nua06, Nua09].

Fix a stochastic basis  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P}, W)$ , where  $W = (W_1, \dots, W_d)$  is a  $d$ -dimensional standard Brownian motion,  $\{\mathcal{F}_t\}_{t \geq 0}$  is a filtration to which this process  $W$  is adapted. In application to (2.1)–(2.2),  $d = 2|\mathcal{Z}|$  represents the number of independent noise processes driving the system. Fix any  $T \in (0, \infty)$ .

We first recall the definition of the *Malliavin derivative*  $\mathfrak{D}$  which is defined on a subset of  $L^p(\Omega)$  for  $p > 1$  (see [Nua09, Nua06] or [Mal97]). We begin by explaining how this operator  $\mathfrak{D}$  acts on ‘smooth random variables’. For any given  $n \geq 1$ , consider a Schwartz function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , that is,  $f$  that satisfies

$$\sup_{z \in \mathbb{R}^n} |z^\alpha D^\beta f(z)| < \infty$$

for any multi-indices  $\alpha, \beta$ . For such functions define  $F \in L^p(\Omega)$ ,  $p > 1$  by

$$F = f \left( \int_0^T g_1 \cdot dW, \dots, \int_0^T g_n \cdot dW \right),$$

where  $g_1, \dots, g_n$  are deterministic elements in  $L^2([0, T], \mathbb{R}^d)$ . For such  $F$  the Malliavin derivative is defined as

$$\mathfrak{D}F := \sum_{k=1}^n \frac{\partial f}{\partial x_k} \left( \int_0^T g_1 \cdot dW, \dots, \int_0^T g_n \cdot dW \right) g_k. \quad (\text{A.26})$$

Notice that  $\mathfrak{D}F \in L^p(\Omega; L^2([0, T], \mathbb{R}^d))$ . To extend  $\mathfrak{D}$  to a broader class of elements we adopt the norm

$$\|F\|_{\mathbb{D}^{1,p}}^p := \mathbb{E}|F|^p + \mathbb{E}\|\mathfrak{D}F\|_{L^2([0,T],\mathbb{R}^d)}^p,$$

and denote  $\text{Dom}(\mathfrak{D}) = \mathbb{D}^{1,p}$  be the closure of the above defined functions  $F$  under this norm  $\|\cdot\|_{\mathbb{D}^{1,p}}$ .

We can repeat the above construction for random variables taking values in a separable Hilbert space  $\mathcal{H}$ . In this case start by considering ‘elementary’ functions of the form

$$F := \sum_{i \in \mathcal{I}} f_i \left( \int_0^T g_1^i \cdot dW, \dots, \int_0^T g_{n_i}^i \cdot dW \right) h_i =: \sum_{i \in \mathcal{I}} F_i h_i, \quad (\text{A.27})$$

where  $i \in \mathcal{I}$  is a finite index set,  $n_i \in \mathbb{N}$ ,  $f_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$  are Schwartz functions,  $h_i$  elements in  $\mathcal{H}$  and, as above,  $g_1^i, \dots, g_{n_i}^i$ ,  $i \in \mathcal{I}$  are deterministic elements in  $L^2([0, T], \mathbb{R}^d)$ . Define

$$\mathfrak{D}F := \sum_{i \in \mathcal{I}} \mathfrak{D}(F_i) h_i. \quad (\text{A.28})$$

Then  $\mathfrak{D}$  is a closeable operator from  $L^p(\Omega, \mathcal{H})$  to  $L^p(\Omega; L^2([0, T], \mathbb{R}^d) \otimes \mathcal{H})$ . With a slight abuse of notation we denote

$$\|F\|_{\mathbb{D}^{1,p}}^p := \mathbb{E}\|F\|_{\mathcal{H}}^p + \mathbb{E}\|\mathfrak{D}F\|_{L^2([0,T],\mathbb{R}^d) \otimes \mathcal{H}}^p. \quad (\text{A.29})$$

As above, we take  $\mathbb{D}^{1,p} = \mathbb{D}^{1,p}(\mathcal{H})$  to be the closure of the functions  $F$  of the form (A.28) under the norm  $\|\cdot\|_{\mathbb{D}^{1,p}}$ .

For  $F \in \mathbb{D}^{1,2}$ , we adopt the notations

$$\mathfrak{D}_s F := (\mathfrak{D}F)(s), \quad s \in [0, T], \quad \mathfrak{D}^j F := (\mathfrak{D}F)^j, \quad j = 1, \dots, d,$$

i.e.  $(\mathfrak{D}F)^j$  is the  $j$ th component of  $\mathfrak{D}F$  as an element in  $\mathbb{R}^d$  (or  $\mathbb{R}^d \otimes \mathcal{H}$ ), for fixed  $\omega, t \in \Omega \times [0, T]$ . Furthermore, in view of (A.26), we have that if  $F \in \mathbb{D}^{1,2}$  is  $\mathcal{F}_s$  measurable for some  $s \in [0, T]$ , then

$$\mathfrak{D}_t F = 0, \quad \text{for every } t > s. \quad (\text{A.30})$$

With these basic definitions in place, we now introduce two important elements of Malliavin's theory, the chain rule and the integration by part formula. The Malliavin chain rule states that for any  $\phi \in C_b^1(\mathbb{R}^m)$  (continuously differentiable functions with bounded derivatives) and  $F = (F_1, \dots, F_m)$  with  $F_i \in \mathbb{D}^{1,2}$  one has that  $\phi(F) \in \mathbb{D}^{1,2}$  and

$$\mathfrak{D}^j \phi(F) = \nabla \phi(F) \cdot \mathfrak{D}^j F = \sum_{i=1}^m \partial_{x_i} \phi(F) \mathfrak{D}^j F_i \quad (j \in \{1, \dots, d\}), \quad (\text{A.31})$$

see e.g. [Nua06, Proposition 1.2.3]. Note also that this chain rule extends to the Hilbert space setting; if  $\phi \in C_b^1(\mathcal{H})$  and  $F \in \mathbb{D}^{1,2}(\mathcal{H})$  then  $\phi(F) \in \mathbb{D}^{1,2}$  and  $\mathfrak{D}^j \phi(F) = \nabla \phi(F) \cdot \mathfrak{D}^j F$ .<sup>30</sup>

Next, we introduce the Malliavin integration by part formula which can be understood in terms of the adjoint operator to  $\mathfrak{D}$ . For  $\mathfrak{D} : \mathbb{D}^{1,2} \subset L^2(\Omega) \rightarrow L^2(\Omega; L^2[0, T], \mathbb{R}^d)$  define its adjoint  $\mathfrak{D}^* : \text{Dom}(\mathfrak{D}^*) \subset L^2(\Omega; L^2[0, T], \mathbb{R}^d) \rightarrow L^2(\Omega)$  by

$$\mathbb{E} \langle \mathfrak{D}F, v \rangle_{L^2([0, T], \mathbb{R}^d)} = \mathbb{E}(F \mathfrak{D}^* v), \quad (\text{A.32})$$

for any  $F \in \mathbb{D}^{1,2}$  and any  $v \in \text{Dom}(\mathfrak{D}^*)$ . If  $F$  has the form (A.27) we define

$$\mathbb{E} \langle \mathfrak{D}F, v \rangle_{L^2([0, T], \mathbb{R}^d)} := \sum_{i \in \mathcal{I}} \mathbb{E} \langle \mathfrak{D}F_i, v \rangle_{L^2([0, T], \mathbb{R}^d)} h_i = \sum_{i \in \mathcal{I}} \mathbb{E}(F_i \mathfrak{D}^* v) h_i = \mathbb{E} \left( \sum_{i \in \mathcal{I}} F_i h_i \mathfrak{D}^* v \right) = \mathbb{E}(F \mathfrak{D}^* v),$$

and therefore, after passing to the limit, we see that (A.32) holds true for  $\mathcal{H}$  valued elements  $F \in \mathbb{D}^{1,2}(\mathcal{H})$  and  $v \in \text{Dom}(\mathfrak{D}^*) \subset L^2(\Omega; L^2([0, T], \mathbb{R}^d))$ . Note in particular that even in this infinite dimensional setting the duality in (A.32) remains in  $L^2([0, T], \mathbb{R}^d)$ .

The map  $\mathfrak{D}^*$  is called the *Skorokhod integral* (see [Nua06]) and is often written as

$$\mathfrak{D}^* v =: \int_0^T v \cdot dW, \quad (\text{A.33})$$

so that (A.32) reads as

$$\mathbb{E} \langle \mathfrak{D}F, v \rangle_{L^2([0, T], \mathbb{R}^d)} = \mathbb{E} \left( F \int_0^T v \cdot dW \right). \quad (\text{A.34})$$

The reason behind this notation is that if  $v \in L^2(\Omega; L^2[0, T], \mathbb{R}^d)$  and is adapted to  $\mathcal{F}_t$ , then  $v \in \text{Dom}(\mathfrak{D}^*)$  and  $\int_0^T v \cdot dW$  is the classical Doebelin-Itô integral. In general, interpreting  $v$  as an  $\mathcal{H} = L^2([0, T]; \mathbb{R}^d)$  valued random variable, we have that  $\mathbb{D}^{1,2}(L^2([0, T]; \mathbb{R}^d)) \subset \text{Dom}(\mathfrak{D}^*)$ .

In order to make quantitative estimates for (A.33) we finally recall a generalized form of the classical Itô isometry. If  $v \in \mathbb{D}^{1,2}(L^2([0, T]; \mathbb{R}^d))$ , then  $\mathfrak{D}v \in L^2(\Omega; L^2([0, T]; \mathbb{R}^d) \otimes L^2([0, T]; \mathbb{R}^d)) = L^2(\Omega; L^2([0, T]^2; \mathbb{R}^{d \times d}))$  and the generalized Itô isometry takes the form:

$$\mathbb{E} \left( \int_0^T v \cdot dW \right)^2 = \mathbb{E} \|v\|_{L^2([0, T], \mathbb{R}^d)}^2 + \mathbb{E} \int_0^T \int_0^T \text{Tr}(\mathfrak{D}_s v(r) \mathfrak{D}_r v(s)) ds dr, \quad (\text{A.35})$$

see e.g. [Nua09, Chapter 1, (1.54)]. In view of (A.30), the classical Itô isometry is recovered from (A.35) when  $v$  is  $\mathcal{F}_t$ -adapted. More generally such observations concerning the  $\mathcal{F}_t$  measurability of  $v$  in conjunction with (A.30), (A.35) are used in a crucial fashion to obtain the bound (3.9).

<sup>30</sup>For some of the estimates in Section 3.4 (cf. (3.9)) we used a more general form of the product rule that can be found in e.g. [PV].

## Acknowledgements

The authors gratefully acknowledge the support of the Institute for Mathematics and its Applications (IMA) where this work was conceived. JF, NGH, GR were postdoctoral fellows and ET was a New Directions Professor during the academic year 2012-2013. We have also benefited from the hospitality of the Department of Mathematics Virginia Tech and from the Newton Institute for Mathematical Sciences, University of Cambridge where the final stage of the writing was completed. NGH's work has been partially supported under the grant NSF-DMS-1313272.

We would like to thank T. Beale, C. Doering, D. Faranda, J. Mattingly, V. Šverák, and V. Vicol for numerous inspiring discussions along with many helpful references. We would also like to express our appreciation to L. Capogna and H. Bessaih for helping us to initiate the reading group at the IMA that eventually lead to this work.

## References

- [ACW10] D. Adhikari, C. Cao, and J. Wu, *The 2D Boussinesq equations with vertical viscosity and vertical diffusivity*, J. Differential Equations **249** (2010), no. 5, 1078–1088. MR 2652164 (2011d:35388)
- [AFS08] S. Albeverio, F. Flandoli, and Y. G. Sinai, *SPDE in hydrodynamic: recent progress and prospects*, Lecture Notes in Mathematics, vol. 1942, Springer-Verlag, Berlin, 2008, Lectures given at the C.I.M.E. Summer School held in Cetraro, August 29–September 3, 2005, Edited by Giuseppe Da Prato and Michael Röckner. MR 2459087 (2009g:76120)
- [B01] M. Bénard, Ann. d. Chimie et de Physique **xxiii** (1901), 62.
- [Bat69] G.K. Batchelor, *Computation of the energy spectrum in homogeneous two-dimensional turbulence*, Physics of fluids **12** (1969), no. Suppl. II, 233–239.
- [BBAK92] P. Baldi, G. Ben Arous, and G. Kerkycharian, *Large deviations and the Strassen theorem in Hölder norm*, Stoch. Proc. Appl. **42** (1992), no. 1.
- [Bel87] D. R. Bell, *The Malliavin calculus*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 34, Longman Scientific & Technical, Harlow, 1987. MR 902583 (88m:60155)
- [Bis81a] J.-M. Bismut, *Martingales, the Malliavin calculus and Hörmander's theorem*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 85–109. MR 620987 (82h:60114)
- [Bis81b] ———, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Z. Wahrsch. Verw. Gebiete **56** (1981), no. 4, 469–505. MR 621660 (82k:60134)
- [BKL01] J. Bricmont, A. Kupiainen, and R. Lefevere, *Ergodicity of the 2D Navier-Stokes equations with random forcing*, Comm. Math. Phys. **224** (2001), no. 1, 65–81, Dedicated to Joel L. Lebowitz. MR 1868991 (2003c:76032)
- [BM07] Y. Bakhtin and J. C. Mattingly, *Malliavin calculus for infinite-dimensional systems with additive noise*, Journal of Functional Analysis **249** (2007), no. 2, 307 – 353.
- [Bou97] J.V. Boussinesq, *Théorie de l'écoulement tourbillonnant et tumultueux des liquides dans les lits rectilignes a grande section*, Des comptes rendus des séances des sciences (1897).
- [BPA00] E. Bodenschatz, W. Pesch, and G. Ahlers, *Recent developments in Rayleigh-Bénard convection*, Annual review of fluid mechanics, Vol. 32, Annu. Rev. Fluid Mech., vol. 32, Annual Reviews, Palo Alto, CA, 2000, pp. 709–778. MR 1744317 (2000k:76059)
- [Bus89] F.H. Busse, *Fundamentals of thermal convection*, Mantle Convections, Plate Tectonics and Global Dynamics (W. Peltier, ed.), 1989.
- [CD80] J. R. Cannon and E. DiBenedetto, *The initial value problem for the Boussinesq equations with data in  $L^p$* , Approximation methods for Navier-Stokes problems (Proc. Sympos., Univ. Paderborn, Paderborn, 1979), Lecture Notes in Math., vol. 771, Springer, Berlin, 1980, pp. 129–144. MR 565993 (81f:35101)
- [CD96] P. Constantin and C. R. Doering, *Heat transfer in convective turbulence*, Nonlinearity **9** (1996), no. 4, 1049–1060. MR 1399486 (97g:76042)
- [CD99] ———, *Infinite Prandtl number convection*, J. Statist. Phys. **94** (1999), no. 1-2, 159–172. MR 1679670 (2000a:76152)
- [Cer01] S. Cerrai, *Second order pde's in finite and infinite dimension: a probabilistic approach*, vol. 1762, Springer, 2001.
- [CF88] P. Constantin and C. Foias, *Navier-Stokes equations*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1988. MR 972259 (90b:35190)

- [CFdIL04] D. Córdoba, C. Fefferman, and R. de la Llave, *On squirt singularities in hydrodynamics*, SIAM J. Math. Anal. **36** (2004), no. 1, 204–213. MR 2083858 (2005i:76032)
- [CGHV13] P. Constantin, N. Glatt-Holtz, and V. Vicol, *Unique ergodicity for fractionally dissipated, stochastically forced 2d euler equations*, Comm. Math. Phys. (2013), (to appear).
- [Cha06] D. Chae, *Global regularity for the 2D Boussinesq equations with partial viscosity terms*, Adv. Math. **203** (2006), no. 2, 497–513. MR 2227730 (2007e:35223)
- [CI99] D. Chae and O. Y. Imanuvilov, *Generic solvability of the axisymmetric 3-D Euler equations and the 2-D Boussinesq equations*, J. Differential Equations **156** (1999), no. 1, 1–17. MR 1700862 (2000d:76009)
- [CLTT05] E. Calzavarini, D. Lohse, F. Toschi, and R. Tripiccione, *Rayleigh and prandtl number scaling in the bulk of rayleigh-bénard turbulence*, Physics of Fluids **17** (2005), no. 5, 055107–055107–7.
- [CRG<sup>+</sup>06] E. Calzavarini, Doering C. R., J. D. Gibbon, D. Lohse, A. Tanabe, and Toschi F., *Exponentially growing solutions in homogeneous rayleigh-bénard convection*, Physical Review E **73** (2006), no. 3.
- [Cru89] A. B. Cruzeiro, *Solutions et mesures invariantes pour des équations d'évolution stochastiques du type Navier-Stokes*, Exposition. Math. **7** (1989), no. 1, 73–82. MR MR982157 (90c:35161)
- [CW12] D. Chae and J. Wu, *The 2D Boussinesq equations with logarithmically supercritical velocities*, Adv. Math. **230** (2012), no. 4-6, 1618–1645. MR 2927350
- [CW13] C. Cao and J. Wu, *Global Regularity for the Two-Dimensional Anisotropic Boussinesq Equations with Vertical Dissipation*, Arch. Ration. Mech. Anal. **208** (2013), no. 3, 985–1004. MR 3048599
- [DC01] C. R. Doering and P. Constantin, *On upper bounds for infinite Prandtl number convection with or without rotation*, J. Math. Phys. **42** (2001), no. 2, 784–795. MR 1809253 (2002d:76107)
- [Deb11] A. Debussche, *Ergodicity results for the stochastic navier-stokes equations: an introduction.*, Preprint (2011), (to appear).
- [DG95] C. R. Doering and J. D. Gibbon, *Applied analysis of the Navier-Stokes equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1995. MR MR1325465 (96a:76024)
- [Doo48] J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. **63** (1948), 393–421. MR 0025097 (9,598c)
- [DP08] R. Danchin and M. Paicu, *Existence and uniqueness results for the Boussinesq system with data in Lorentz spaces*, Phys. D **237** (2008), no. 10-12, 1444–1460. MR 2454598 (2009h:76184)
- [DP09] ———, *Global well-posedness issues for the inviscid Boussinesq system with Yudovich's type data*, Comm. Math. Phys. **290** (2009), no. 1, 1–14. MR 2520505 (2010f:35298)
- [DP11] ———, *Global existence results for the anisotropic Boussinesq system in dimension two*, Math. Models Methods Appl. Sci. **21** (2011), no. 3, 421–457. MR 2782720 (2012g:35254)
- [DPRRW06] G. Da Prato, M. Rockner, B. L. Rozovskii, and F. Y. Wang, *Strong solutions of stochastic generalized porous media equations: existence, uniqueness, and ergodicity*, Communications in Partial Differential Equations **31** (2006), no. 2, 277–291.
- [DPZ92] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992. MR MR1207136 (95g:60073)
- [DPZ96] ———, *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996. MR MR1417491 (97k:60165)
- [DPZ02] ———, *Second order partial differential equations in hilbert spaces*, no. 293, Cambridge University Press, 2002.
- [E01] W. E, *Stochastic hydrodynamics*, Current developments in mathematics, 2000, Int. Press, Somerville, MA, 2001, pp. 109–147. MR 1882534 (2003a:76040)
- [EH01] J. P. Eckmann and M. Hairer, *Uniqueness of the invariant measure for a stochastic pde driven by degenerate noise*, Communications in Mathematical Physics **219** (2001), no. 3, 523–565.
- [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** (2001), no. 11, 1386–1402. MR 1846802 (2002g:76075)
- [EMS01] W. E, J. C. Mattingly, and Y. G.. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation*, Comm. Math. Phys. **224** (2001), no. 1, 83–106, Dedicated to Joel L. Lebowitz. MR 1868992 (2002m:76024)
- [Eyi96] G.L. Eyink, *Exact results on stationary turbulence in 2d: consequences of vorticity conservation*, Physica D: Nonlinear Phenomena **91** (1996), no. 1, 97–142.
- [Fer97] B. Ferrario, *The Bénard problem with random perturbations: dissipativity and invariant measures*, NoDEA Nonlinear Differential Equations Appl. **4** (1997), no. 1, 101–121. MR 1433314 (98c:35174)

- [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** (2002), no. 3-4, 591–645. MR 1914189 (2004k:76067)
- [FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2-D Navier-Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), no. 1, 119–141. MR MR1346374 (96g:35223)
- [FMRT01] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes equations and turbulence*, Encyclopedia of Mathematics and its Applications, vol. 83, Cambridge University Press, Cambridge, 2001. MR 1855030 (2003a:76001)
- [FP67] C. Foias 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. MR 0223716 (36 #6764)
- [Get98] A. V. Getling, *Rayleigh-Bénard convection*, Advanced Series in Nonlinear Dynamics, vol. 11, World Scientific Publishing Co. Inc., River Edge, NJ, 1998, Structures and dynamics. MR 1629996 (99g:76138)
- [GHSV13] N. Glatt-Holtz, V. Sverak, and V. Vicol, *On inviscid limits for the stochastic navier-stokes equations and related models*.
- [GM05] B. Goldys and B. Maslowski, *Exponential ergodicity for stochastic burgers and 2d navier-stokes equations*, Journal of Functional Analysis **226** (2005), no. 1, 230–255.
- [GS02] A. L. Gibbs and F. E. Su, *On choosing and bounding probability metrics*, International Statistical Review **70** (2002), no. 3, 419–435.
- [H67] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR 0222474 (36 #5526)
- [Hai11] M. Hairer, *On Malliavin’s proof of Hörmander’s theorem*, Bull. Sci. Math. **135** (2011), no. 6-7, 650–666. MR 2838095 (2012m:60119)
- [HK09] T. Hmidi and S. Keraani, *On the global well-posedness of the Boussinesq system with zero viscosity*, Indiana Univ. Math. J. **58** (2009), no. 4, 1591–1618. MR 2542974 (2011b:35208)
- [HL05] T. Y. Hou and C. Li, *Global well-posedness of the viscous Boussinesq equations*, Discrete Contin. Dyn. Syst. **12** (2005), no. 1, 1–12. MR 2121245 (2005j:76026)
- [HM06] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, Ann. of Math. (2) **164** (2006), no. 3, 993–1032. MR 2259251 (2008a:37095)
- [HM08] ———, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091. MR 2478676 (2010i:35295)
- [HM11] ———, *A theory of hypoellipticity and unique ergodicity for semilinear stochastic pdes*, Electron. J. Probab. **16** (2011), no. 23, 658–738.
- [HMS11] M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259. MR 2773030
- [IW89] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam, 1989. MR 1011252 (90m:60069)
- [KB37] N. Kryloff and N. Bogoliouboff, *La théorie générale de la mesure dans son application à l’étude des systèmes dynamiques de la mécanique non linéaire*, Ann. of Math. (2) **38** (1937), no. 1, 65–113. MR 1503326
- [Km60] R. Z. Khas’minskii, *Ergodic properties of recurrent diffusion processes and stabilization of the solution to the cauchy problem for parabolic equations*, Theory of Probability & Its Applications **5** (1960), no. 2, 179–196.
- [KM80] R. H. Kraichnan and D. Montgomery, *Two-dimensional turbulence*, Rep. Prog. Phys. **43** (1980), 574–619.
- [Kol34] A. Kolmogoroff, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2) **35** (1934), no. 1, 116–117. MR 1503147
- [Kra67] R. H. Kraichnan, *Inertial ranges in two-dimensional turbulence*, Phys. Fluids **10** (1967), no. 7, 1417–1423.
- [KS84] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306. MR 780762 (86k:60100a)
- [KS85] ———, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76. MR 783181 (86k:60100b)
- [KS87] ———, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442. MR 914028 (89c:60093)
- [KS91] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR MR1121940 (92h:60127)
- [KS01] S. Kuksin and A. Shirikyan, *A coupling approach to randomly forced nonlinear PDE’s. I*, Comm. Math. Phys. **221** (2001), no. 2, 351–366. MR 1845328 (2002e:35253)

- [KS02] ———, *Coupling approach to white-forced nonlinear PDEs*, J. Math. Pures Appl. (9) **81** (2002), no. 6, 567–602. MR 1912412 (2003g:37146)
- [KS12] ———, *Mathematics of two-dimensional turbulence*, Cambridge Tracts in Mathematics, no. 194, Cambridge University Press, 2012.
- [Kup10] A. Kupiainen, *Ergodicity of two dimensional turbulence*, arXiv:1005.0587v1 [math-ph] (2010).
- [KW12] T. Komorowski and A. Walczuk, *Central limit theorem for Markov processes with spectral gap in the Wasserstein metric*, Stochastic Process. Appl. **122** (2012), no. 5, 2155–2184. MR 2921976
- [LLT10] A. Larios, E. Lunasin, and E. S. Titi, *Global well-posedness for the 2d boussinesq system without heat diffusion and with either anisotropic viscosity or inviscid voigt-regularization*.
- [LR16] O.M.F.R.S. Lord Rayleigh, *On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side*, Philosophical Magazine Series 6 **32** (1916), no. Issue 192, 529–546.
- [LT03] D. Lohse and F. Toschi, *Ultimate state of thermal convection*, Physical Review Letters **90** (2003), no. 3.
- [LW04] J. Lee and M. Y. Wu, *Ergodicity for the dissipative Boussinesq equations with random forcing*, J. Statist. Phys. **117** (2004), no. 5-6, 929–973. MR 2107902 (2005i:35219)
- [Mal78] P. Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York), Wiley, 1978, pp. 195–263. MR 536013 (81f:60083)
- [Mal97] ———, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 313, Springer-Verlag, Berlin, 1997. MR 1450093 (99b:60073)
- [Mat99] J. C. Mattingly, *Ergodicity of 2D Navier-Stokes equations with random forcing and large viscosity*, Comm. Math. Phys. **206** (1999), no. 2, 273–288. MR MR1722141 (2000k:76040)
- [Mat02] ———, *Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics*, Comm. Math. Phys. **230** (2002), no. 3, 421–462. MR 1937652 (2004a:76039)
- [Mat03] ———, *On recent progress for the stochastic Navier Stokes equations*, Journées “Équations aux Dérivées Partielles”, Univ. Nantes, Nantes, 2003, pp. Exp. No. XI, 52. MR 2050597 (2005b:60155)
- [MB02] A.J. Majda and A.L. Bertozzi, *Vorticity and incompressible flow*, Cambridge Texts in Applied Mathematics, vol. 27, Cambridge University Press, Cambridge, 2002. MR 1867882 (2003a:76002)
- [MP06] J. C. Mattingly and É. Pardoux, *Malliavin calculus for the stochastic 2D Navier-Stokes equation*, Comm. Pure Appl. Math. **59** (2006), no. 12, 1742–1790. MR MR2257860 (2007j:60082)
- [MW04] T. Ma and S. H. Wang, *Dynamic bifurcation and stability in the Rayleigh-Bénard convection*, Commun. Math. Sci. **2** (2004), no. 2, 159–183. MR 2119936 (2005j:76038)
- [MW07] T. Ma and S. Wang, *Rayleigh-Bénard convection: dynamics and structure in the physical space*, Commun. Math. Sci. **5** (2007), no. 3, 553–574. MR 2352331 (2008g:76047)
- [MY02] N. Masmoudi and L. S. Young, *Ergodic theory of infinite dimensional systems with applications to dissipative parabolic pdes*, Communications in mathematical physics **227** (2002), no. 3, 461–481.
- [Nor86] J. Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130. MR 942019 (89f:60058)
- [Nov65] E. A. Novikov, *Functionals and the random-force method in turbulence theory*, Soviet Physics JETP **20** (1965), 1290–1294. MR 0191516 (32 #8921)
- [NTC<sup>+</sup>10] A. Naso, S. Thalabard, G. Collette, P. H. Chavanis, and B. Dubrulle, *Statistical mechanics of beltrami flows in axisymmetric geometry: equilibria and bifurcations*, Journal of Statistical Mechanics: Theory and Experiment **2010** (2010), no. 06, P06019.
- [Nua06] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR 2200233 (2006j:60004)
- [Nua09] ———, *Malliavin calculus and its applications*, CBMS Regional Conference Series in Mathematics, vol. 110, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2009. MR 2498953 (2010b:60164)
- [Ped82] J. Pedlosky, *Geophysical fluid dynamics*, Springer Verlag, 1982.
- [PR07] C. Prévôt and M. Röckner, *A concise course on stochastic partial differential equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007. MR MR2329435
- [PV] M. Pronk and M. Veraar, *Tools for malliavin calculus in umd banach spaces*, Potential Analysis, (to appear).
- [Rom04] M. Romito, *Ergodicity of the finite dimensional approximation of the 3D Navier-Stokes equations forced by a degenerate noise*, J. Statist. Phys. **114** (2004), no. 1-2, 155–177. MR 2032128 (2005a:76128)

- [Roz90] B. L. Rozovskii, *Stochastic evolution systems*, Mathematics and its Applications (Soviet Series), vol. 35, Kluwer Academic Publishers Group, Dordrecht, 1990, Linear theory and applications to nonlinear filtering, Translated from the Russian by A. Yarkho. MR 1135324 (92k:60136)
- [Shi80] I. Shigekawa, *Derivatives of Wiener functionals and absolute continuity of induced measures*, J. Math. Kyoto Univ. **20** (1980), no. 2, 263–289. MR 582167 (83g:60051)
- [Shi06] A. Shirikyan, *Law of large numbers and central limit theorem for randomly forced pde's*, Probability theory and related fields **134** (2006), no. 2, 215–247.
- [Sta88] M. M. Stanišić, *The mathematical theory of turbulence*, second ed., Universitext, Springer-Verlag, New York, 1988. MR 918774 (88i:76020)
- [Str81] D. W. Stroock, *The Malliavin calculus and its applications*, Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980), Lecture Notes in Math., vol. 851, Springer, Berlin, 1981, pp. 394–432. MR 620997 (82k:60092)
- [SW13] T. Sengul and S. H. Wang, *Pattern formation in Rayleigh-Bénard convection*, Commun. Math. Sci. **11** (2013), no. 1, 315–343. MR 2975373
- [Tab02] P. Tabeling, *Two-dimensional turbulence: a physicist approach*, Physics Reports **362** (2002), no. 1, 1–62.
- [TDB13] S. Thalabard, B. Dubrulle, and F. Bouchet, *Statistical mechanics of the 3d axisymmetric euler equations in a taylor-couette geometry*, arXiv preprint arXiv:1306.1081 (2013).
- [Tem01] R. Temam, *Navier-Stokes equations: Theory and numerical analysis*, AMS Chelsea Publishing, Providence, RI, 2001, Reprint of the 1984 edition. MR MR1846644 (2002j:76001)
- [VKF79] M. I. Vishik, A. I. Komech, and A. V. Fursikov, *Some mathematical problems of statistical hydromechanics*, Uspekhi Mat. Nauk **34** (1979), no. 5(209), 135–210, 256. MR 562801 (83e:35098)
- [Wan04a] X. Wang, *Infinite Prandtl number limit of Rayleigh-Bénard convection*, Comm. Pure Appl. Math. **57** (2004), no. 10, 1265–1282. MR 2069723 (2005d:76015)
- [Wan04b] ———, *Large Prandtl number behavior of the Boussinesq system of Rayleigh-Bénard convection*, Appl. Math. Lett. **17** (2004), no. 7, 821–825. MR 2072841 (2005g:76052)
- [Wan05] ———, *A note on long time behavior of solutions to the Boussinesq system at large Prandtl number*, Nonlinear partial differential equations and related analysis, Contemp. Math., vol. 371, Amer. Math. Soc., Providence, RI, 2005, pp. 315–323. MR 2143874 (2006a:76105)
- [Wan07] ———, *Asymptotic behavior of the global attractors to the Boussinesq system for Rayleigh-Bénard convection at large Prandtl number*, Comm. Pure Appl. Math. **60** (2007), no. 9, 1293–1318. MR 2337505 (2009a:35196)
- [Wan08a] ———, *Bound on vertical heat transport at large Prandtl number*, Phys. D **237** (2008), no. 6, 854–858. MR 2452172 (2009i:76064)
- [Wan08b] ———, *Stationary statistical properties of Rayleigh-Bénard convection at large Prandtl number*, Comm. Pure Appl. Math. **61** (2008), no. 6, 789–815. MR 2400606 (2010b:76098)

Juraj Földes

Institute for Mathematics and its Applications (IMA)

University of Minnesota

Web: [ima.umn.edu/~foldes/](http://ima.umn.edu/~foldes/)

Email: [foldes@ima.umn.edu](mailto:foldes@ima.umn.edu)

Nathan Glatt-Holtz

Department of Mathematics

Virginia Polytechnic Institute and State University

Web: [www.math.vt.edu/people/negh/](http://www.math.vt.edu/people/negh/)

Email: [negh@vt.edu](mailto:negh@vt.edu)

Geordie Richards

Department of Mathematics

University of Rochester

Web: [www.math.rochester.edu/grichar5/](http://www.math.rochester.edu/grichar5/)

Email: [grichar5@z.rochester.edu](mailto:grichar5@z.rochester.edu)

Enrique Thomann

Department of Mathematics

Oregon State University

Web: [www.math.oregonstate.edu/people/view/thomann](http://www.math.oregonstate.edu/people/view/thomann)

Email: [thomann@math.orst.edu](mailto:thomann@math.orst.edu)