Ergodicity in Randomly Forced Rayleigh-Bénard Convection

J. Földes, N. Glatt-Holtz, G. Richards, J.P. Whitehead

emails: juraj.foldes@ulb.ac.be, negh@vt.edu, g.richards@rochester.edu, whitehead@mathematics.byu.edu

November 4, 2015

Abstract

We consider the Boussinesq approximation for Rayleigh-Bénard convection perturbed by an additive noise and with boundary conditions corresponding to heating from below. In two space dimensions, with sufficient stochastic forcing in the temperature component and large Prandtl number $Pr > 0$, we establish the existence of a unique ergodic invariant measure. In three space dimensions, we prove the existence of a statistically invariant state, and establish unique ergodicity for the infinite Prandtl Boussinesq system. Throughout this work we provide streamlined proofs of unique ergodicity which invoke an asymptotic coupling argument, a delicate usage of the maximum principle, and exponential martingale inequalities. Lastly, we show that the background method of Constantin-Doering [CD96] can be applied in our stochastic setting, and prove bounds on the Nusselt number relative to the unique invariant measure.

Contents

1 Introduction

Following experiments of Bénard [Bó1], Rayleigh [LR16] proposed the equations of Boussinesq [Bou97] as an effective model for the flow of a fluid driven by buoyancy forces due to heating from below and cooling from above, a phenomenon now referred to as Rayleigh-Bénard convection. These equations have since appeared in a wide variety of physical models, including descriptions of climate and weather processes and the internal dynamics of both planets and stars.

Individual solutions of the Boussinesq system can be unpredictable and seemingly chaotic, particularly in parameter ranges leading to turbulent regimes. However, some of the statistical properties of solutions are robust. It is therefore of fundamental significance to identify and predict statistical features of Rayleigh-Bénard convection, and to connect these features to rigorous theory at the level of the Boussinesq equations.

Indeed, the fine scale structure of flows, complex pattern formation, and the mean heat transport, for example, remain topics of intensive theoretical, numerical and experimental research. Here the role of analysis is particularly significant in parameter ranges beyond the capacity of direct numerical simulation or empirical observability. See [BPA00, Man06, AGL09, LX10] for a survey of recent developments in the physics literature.

A mathematically rigorous theory of statistical properties should include the analysis of invariant measures for the system, which contain important statistics of the flow. Natural questions include the existence, uniqueness, ergodicity and other attraction properties of invariant measures. Moreover one may seek to prove quantitative bounds on statistical quantities determined by flows in terms of these measures.

There is a significant literature devoted to proving rigorous quantitative bounds for the Boussinesq equations, primarily focused on estimating rates of convective heat transport. This direction of research was initiated by [Mal54, How72, Bus70, Bus78], advanced significantly with the invention of the "background flow method" [CD96, CD99, DC01] (see also [Hop40]), and refined in more recent works (e.g. [WD11, OS11]). It is noteworthy that practical methodologies for proving rigorous bounds on key statistical quantities have not been identified in a stochastic setting.

On the other hand, while some works have established existence and convergence properties of invariant measures for the Boussinesq system [Wan08], it is difficult, in general, to obtain uniqueness or erdogicity results for systems of deterministic partial differential equations (cf. [FMRT01]). These problems become more tractable by including a stochastic forcing, due to smoothing properties of the corresponding probability distribution functions induced by random perturbations in the equations. Moreover, as early as the 19th century, Boussinesq conjectured that turbulent flow cannot be described solely with deterministic methods, and indicated that a stochastic framework should be used [Sta88]. This setting is now ubiquitous in the turbulence literature, see e.g. [Nov65, VKF79, Eyi96] and containing references. In particular, note that some works have considered stochastic initial and boundary conditions for the Boussinesq system to predict qualitative features of the flow, including the onset of turbulence [VWK10, VCK12].

In the manuscript we consider a stochastic Boussinesq system

$$\frac{1}{Pr}(d\mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}dt) + \nabla pdt = \Delta\mathbf{u}dt + Ra\hat{\mathbf{e}}_d Tdt + \sum_{k=1}^{N_1}\tilde{\sigma}_k d\tilde{W}^k, \quad \nabla\cdot\mathbf{u} = 0, \tag{1.1}$$

$$dT + \mathbf{u}\cdot\nabla Tdt = \Delta Tdt + \sum_{k=1}^{N_2}\sigma_k dW^k, \tag{1.2}$$

for the (non-dimensionalized) velocity field $\mathbf{u} = (u_1,\ldots,u_d)$ (where $d = 2$ or $d = 3$), pressure p, and temperature T of a buoyancy driven fluid. The system (1.1)–(1.2) evolves in a domain $\mathbf{x} = (x_1,\ldots,x_d) \in \mathcal{D} = [0,L]^{d-1} \times [0,1]$ and is supplemented with boundary conditions to be specified below. Here $\hat{\mathbf{e}}_d = (0,\ldots,1)$ is a unit vector pointing in the vertical direction.

The driving noise is given by a collection of independent white noise processes $d\tilde{W}^k = d\tilde{W}^k(t)$ and $dW^k = dW^k(t)$ acting in spatial directions $\tilde{\sigma}_k = \tilde{\sigma}_k(x)$, $\sigma_k = \sigma_k(x)$ which form a complete orthogonal basis of eigenfunctions (ordered with respect to eigenvalues) of the Stokes and Laplace operators on $\mathcal{D}$, respectively, with appropriate boundary conditions. [1] The number of forced modes, N_1 and N_2, are both finite and will be further specified in theorem statements below. We can treat the case where the numbers N_1 or N_2 are infinite, provided we impose enough decay in the bases $\{\tilde{\sigma}_k\}$ and $\{\sigma_k\}$ such that the system (1.1)–(1.2) remains globally well-posed according to Propositions 2.1–2.2 (see Remark 2.2 below). This does not complicate our analysis in a significant way, and for simplicity of presentation, we assume $N_1, N_2 < \infty$ unless stated otherwise.

We will consider, in particular, the case $N_1 = 0$; that is, (1.1)–(1.2) with no stochastic forcing in the velocity component. This is partly motivated by investigations of the (determistic) Boussinesq system which

---

[1]Our analysis does not require stochastic perturbation in a diagonal basis of eigenfunctions (we assume this form for simplicity of presentation). More generally, we could consider any $\{\tilde{\sigma}_k\}$ and $\{\sigma_k\}$ which span a determining set of directions for the governing equations.

have considered convection driven by internal heating (see [Rob67, TZ67, LDB04, WD11, GS12, BN12]), to describe, for example, radioactive decay processes in the earth's mantle.

In previous work of the first three authors in collaboration with Thomann [FGHRT15], we established ergodic and mixing properties for (1.1)–(1.2) in the two-dimensional periodic domain (i.e. $\mathcal{D} = \mathbb{T}^2$) with a stochastic forcing acting on a small collection of low frequency modes in the temperature component only; that is, with $N_1 = 0$ and small $N_2 > 0$. This form of spatially degenerate random forcing (i.e. N_2 small) is motivated by the turbulence literature, where it is conjectured that nonlinear terms will propagate excitation to higher frequencies, and the system will converge to a unique statistical equilibrium (see [Nov65, VKF79, Eyi96]). The results of [FGHRT15] generalized recent progress of [HM06, HM08, HM11] on the stochastic Navier-Stokes equations and related systems. Indeed, note that with $N_1 = 0$ we are not forcing the velocity field in (1.1)–(1.2) directly, which is a more degenerate setting than was considered in [HM06, HM08, HM11].

From the physical point of view, using periodic boundary conditions in the vertical direction for (1.1)–(1.2) is not appropriate. Instead, one should fix the temperature on the upper and lower boundaries (corresponding to heating from below), and employ Dirichlet conditions in the velocity field, as follows

$$\mathbf{u}_{|x_d=0} = \mathbf{u}_{|x_d=1} = 0, \quad T_{|x_d=0} = \tilde{R}a, \ T_{|x_d=1} = 0, \quad \mathbf{u}, T \text{ are periodic in } \mathbf{x} = (x_1, \ldots, x_{d-1}). \qquad (1.3)$$

The positive unitless physical parameters in the problem (1.1)–(1.2) with boundary conditions (1.3) are the *Prandtl* number Pr and *Rayleigh* numbers Ra and $\tilde{R}a$; see Section A below for further details. Our first objective is to establish existence and uniqueness properties of invariant measures for (1.1)–(1.2) with boundary conditions given by (1.3).

Theorem 1.1. *Suppose the spatial dimension is $d = 2$. Then the system (1.1)–(1.2) with boundary conditions (1.3) possesses a unique ergodic invariant probability measure if at least one of the following holds:*

(i) $N_1 = N_1(Pr, Ra, \tilde{R}a) > 0$ and $N_2 = N_2(Pr, Ra, \tilde{R}a) > 0$ are both sufficiently large.

(ii) $Pr > 0$ and $N_2 = N_2(Pr, Ra, \tilde{R}a) > 0$ are both sufficiently large.

For the proof of Theorem 1.1, we use a streamlined argument which may be of broader interest in the theory of ergodicity for infinite-dimensional systems. A similar approach is used in concurrent work of two of the authors in collaboration with Mattingly [GHMR], where this technique has been implemented for a number of other nonlinear stochastic PDEs. The argument invokes an abstract framework developed for application to SDEs with delay [HMS11], allowing us to significantly reduce the length and technical detail of the proofs.

More precisely, by applying a theorem of [HMS11] (see Theorem 5.1), we can reduce the problem of uniqueness to the convergence of solutions of (1.1)–(1.2) to solutions of a shifted system (see (5.1)–(5.2) below). In order to apply the result of [HMS11], we invoke the Girsanov theorem to establish the equivalence of (the laws of) solutions to the shifted system to those of (1.1)–(1.2), and prove the desired convergence at time infinity by using Foias-Prodi type bounds, a stopping time argument, and a priori estimates on solutions. While the basic ingredients of this method are standard tools in the field (e.g. see [KS12]), we believe that its brevity and simplicity makes it useful. We emphasize that in the context of the Boussinesq system, the a priori estimates on solutions will be proven with a nontrivial comparison argument invoking the maximum principle and certain exponential martingale inequalities (see Section 3 for more details).

In three space dimensions, we prove the existence of statistically invariant states by means of a regularization along with suitable a priori estimates following the general strategy from [FG95] (see also [GHŠV15]). Here our approach is to consider a Galerkin truncation imposed only in the velocity equations (1.1). As such we are able to preserve the advection diffusion structure of (1.2), which plays a critical role in the analysis.

Theorem 1.2. *Suppose the spatial dimension is $d = 3$, then the system (1.1)–(1.2) with boundary conditions (1.3) possesses at least one statistically invariant state.*

The analysis of convection in the large Prandtl number limit is relevant in numerous contexts, such as modeling of the earth's mantle and for convection in high pressure gasses, where Pr can reach the order

of 10^{24} (see [CD99, DC01, OS11]). Taking $N_1 = 0$, and substituting $Pr = \infty$ into (1.1)–(1.2), we formally obtain the stochastic infinite Prandtl Boussinesq system

$$- \Delta \mathbf{u} = \nabla p + Ra\hat{\mathbf{e}}_d T, \quad \nabla \cdot \mathbf{u} = 0, \tag{1.4}$$

$$dT + \mathbf{u} \cdot \nabla T dt = \Delta T dt + \sum_{k=1}^{N_2} \sigma_k dW^k, \tag{1.5}$$

complemented with boundary conditions for T and $\mathbf{u}$ as in (1.3). Note that (1.4)–(1.5) is an active scalar equation for the temperature T, and the velocity field $\mathbf{u}$ is enslaved to T. We remark that the system (1.4)–(1.5) can have complex dynamics, even without stochastic forcing, provided the Rayleigh number Ra is sufficiently large; see [BH09, CD99, BPA00, DC01, Wan04, Par06, AGL09, LX10, OS11].

In a companion work of the first three authors [FGHR], we have recently established uniqueness and mixing properties of the invariant probability measure for the system (1.4)–(1.5). In this manuscript we present a more direct proof of uniqueness in order to highlight another application of the simplified method from [HMS11, GHMR].

Theorem 1.3. *Suppose the spatial dimension is $d = 2$ or $d = 3$. If $N_2 = N_2(Ra, \tilde{Ra}) > 0$ is sufficiently large, then the system (1.4)–(1.5) possesses a unique ergodic invariant probability measure.*

In [FGHR] we also studied asymptotics in the infinite Prandtl limit. Namely, we showed that as $Pr \to \infty$, statistically invariant states of (1.1)–(1.2) (which exist by Theorems 1.1–1.2) converge weakly (in the temperature component) to the unique invariant measure of (1.4)–(1.5). The proof was based on establishing that the Markovian dynamics of (1.4)–(1.5) are contractive with respect to an appropriate Kantorovich-Wasserstein metric. Using this contraction property, we reduced the question of weak convergence of invariant states as $Pr \to \infty$ to one of (fixed) finite time asymptotics, and Pr-uniform exponential moment bounds on invariant states of (1.1)–(1.2) and (1.4)–(1.5). In this paper we include complete proofs of the Pr-uniform bounds required in the analysis of [FGHR] (see Sections 3 and 4, and specifically Corollaries 4.2 and 4.5 below). The proofs of these estimates invoke a comparison argument based on the maximum principle and weighted bounds due to coupling in the systems (1.1)–(1.2) and (1.4)–(1.5).

For the final observation of this manuscript, we illustrate that the background method of Constantin and Doering [CD96, CD99, DC01] applies to the stochastic Boussinesq system (1.1)–(1.2). The Nusselt number Nu, defined as a long-time average, is the ratio of the convective to the conductive heat transport. The background method provides an upper bound on the Nusselt number as a function of the strength of the forcing mechanisms in the system (through both boundary and body forcing in the temperature equation), to illustrate an inherent restriction on the convection produced by heat sources in Boussinesq flows. In our setting, for Pr large, we will obtain rigorous bounds on the Nusselt number Nu for (1.1)–(1.2) in terms of Ra and $\tilde{Ra}$.

More precisely, for $d = 2$, $N_1 = 0$ and $N_2 = \infty$ (with sufficient decay in the basis functions $\{\sigma_k\}$ so that (1.1)–(1.2) is globally well-posed, see Remark 2.2), we define the Nusselt number (relative to μ) by [2]

$$Nu := \frac{1}{\tilde{Ra}|\mathcal{D}|} \int \int_{\mathcal{D}} (u_2 T - \partial_2 T) \, dx \, d\mu(\mathbf{u}, T) = 1 + \frac{1}{\tilde{Ra}|\mathcal{D}|} \int \int_{\mathcal{D}} u_2 T \, dx \, d\mu(\mathbf{u}, T), \tag{1.6}$$

where μ is the unique invariant measure of the system (1.1)–(1.2) (by Theorem 1.1), and ∂_2 denotes the partial derivative in the x_2 direction. Here the integral involving μ is taken over an appropriate phase space for (1.1)–(1.2), see Section 6 for details. Let us emphasize that we have defined the Nusselt number as a statistical average against the unique invariant measure μ, rather than using long time averages of the flow, which is a more standard interpretation. However, by invoking the ergodicity from Theorem 1.1, we can equate these two definitions, and apply the background method to obtain quantitative bounds on Nu. We will reproduce the simplest case of [CD96] which gives $Nu \lesssim (Ra\tilde{Ra})^{1/2}$ for $Ra, \tilde{Ra} \gg 1$ in our context, but it appears that one could adapt other arguments to produce sharper bounds (see e.g. [OS11]).

---

[2] In this context we define the conductive heat transport as $\tilde{Ra}$, the temperature difference between the plates.

Theorem 1.4. *Suppose $d = 2$, $N_1 = 0$, $N_2 = \infty$, and assume that $Pr > 0$ is sufficiently large such that, by Theorem 1.1, the system (1.1)–(1.2) possesses unique ergodic invariant measure μ. Then the Nusselt number Nu given by (1.6) satisfies:*

(i) For μ-almost every initial condition $(\mathbf{u}_0, T_0) \in H$,

$$Nu = \lim_{t \to \infty} \frac{1}{\tilde{Ra}|\mathcal{D}|} \mathbb{E} \left(\frac{1}{t} \int_0^t \int_{\mathcal{D}} (u_2 T - \partial_2 T)(x, s) dx ds \right),$$

where $(\mathbf{u}, T)$ is the solution of (1.1)–(1.2) with initial data $(\mathbf{u}_0, T_0)$.

(ii) $Nu \le C(Ra\tilde{Ra})^{1/2}$ for $Ra, \tilde{Ra} > 0$ large, where $C = C(|\mathcal{D}|) > 0$. [3]

The remainder of this manuscript is organized as follows. In Section 2 we provide the precise mathematical framework of the manuscript. In Section 3 we establish a priori estimates on solutions to (1.1)–(1.2), (1.4)–(1.5), and a larger class of stochastic drift-diffusion equations. We discuss the existence of invariant states in Section 4, and present the proof of Theorem 1.2. Section 5 is devoted to the proofs of our main unique ergodic theorems, Theorem 1.1 and Theorem 1.3. In Section 6 we show how the background method adapts to our stochastic setting, and provide the proof of Theorem 1.4. Lastly, we include an appendix (Appendix A) which provides some details of rescaling arguments for our model equations.

2 Mathematical Framework

In this manuscript we study the stochastic Boussinesq equations (1.1)–(1.2) and (1.4)–(1.5). As usual, these equations are rigorously understood in a time integrated sense. The unitless physical parameters in the problem are the *Prandtl number*, Pr, and the *Rayleigh numbers*, Ra and $\tilde{Ra}$. The system (1.1)–(1.2) is the result of a rescaling. In fact,

$$Pr = \frac{\nu}{\kappa}, \quad Ra = \frac{g\alpha\gamma h^{4-d/2}}{\nu\kappa^{3/2}}, \quad \tilde{Ra} = \frac{\sqrt{\kappa}h^{d/2-1}T_1}{\gamma}, \tag{2.1}$$

where h represents the height of the domain, T_1 the applied temperature difference, and γ is a stochastic heat flux representing the strength of the random forcing in the temperature component of the original variables. Also ν is the kinematic viscosity, κ is the coefficient of thermal diffusivity, g is the gravitational constant and α is the coefficient of thermal expansion. The orthogonal basis $\{\sigma_k\}$ appearing in the stochastic terms of (1.2) have been normalized, depending on the number of forced modes, such that for a given fixed $N_2 > 0$,

$$\|\sigma\|^2 := \sum_{k=1}^{N_2} \|\sigma_k\|_{L^2(\mathcal{D})}^2 = 1,$$

with the strength of the body forcing expressed through Ra and $\tilde{Ra}$. In particular we always assume forcing in the temperature equation, but we prescribe no normalization condition on forcing in the velocity equation. We provide more details on the formulation of our system in Appendix A below.

We will often subtract a linear profile from T in order to replace (1.2) with a system satisfying homogeneous boundary conditions. Define

$$\theta := T - \tilde{Ra}(1 - x_d), \tag{2.2}$$

---

[3]Note that $Ra\tilde{Ra} = \frac{g\alpha h^3 T_1}{\nu\kappa}$, which is the usual Rayleigh number in deterministic convection problems.

we obtain

$$\frac{1}{Pr}(d\mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} dt) + \nabla p dt = \Delta \mathbf{u} dt + Ra\hat{\mathbf{e}}_d \theta dt + \sum_{k=1}^{N_1} \tilde{\sigma}_k d\tilde{W}^k, \quad \nabla \cdot \mathbf{u} = 0, \tag{2.3}$$

$$d\theta + \mathbf{u} \cdot \nabla \theta dt = \tilde{Ra} u_d dt + \Delta \theta dt + \sum_{k=1}^{N_2} \sigma_k dW^k, \tag{2.4}$$

with the boundary conditions modified on the physical boundary as

$$\mathbf{u}_{|x_d=0} = \mathbf{u}_{|x_d=1} = 0, \quad \theta_{|x_d=0} = \theta_{|x_d=1} = 0, \tag{2.5}$$

and initial conditions $\theta(t=0) = \theta_0 = T_0 - \tilde{Ra}(1-x_d)$. Notice that we have implicitly modified the pressure in (2.3) by $\tilde{Ra}(x_d - \frac{1}{2}x_d^2)$ since $(1-x_d)\hat{\mathbf{e}}_d = \nabla(x_d - \frac{1}{2}x_d^2)$.

We will also consider the infinite-Prandtl Boussinesq system

$$(-\Delta \mathbf{u}^0 + \nabla p)dt = Ra\hat{\mathbf{e}}_d \theta^0 dt, \quad \nabla \cdot \mathbf{u}^0 = 0, \tag{2.6}$$

$$d\theta^0 + \mathbf{u}^0 \cdot \nabla \theta^0 dt = \tilde{Ra} u_d^0 dt + \Delta \theta^0 dt + \sum_{k=1}^{N_2} \sigma_k dW^k \tag{2.7}$$

complemented with

$$\theta^0_{|x_d=0} = \theta^0_{|x_d=1} = 0,$$

and periodic boundary conditions in the horizontal directions.

For much of the analysis that follows, we will establish results (well-posedness, existence and uniqueness of invariant states) for the systems (2.3)–(2.4) and (2.6)–(2.7), but these results translate easily back to the original variables in (1.1)–(1.2) and (1.4)–(1.5).

2.1   Functional Setting

The equations (2.3)–(2.4) supplemented with (2.5) may be posed mathematically as follows. Define the phase space $H = H_1 \times H_2$ with

$$H_1 = \{\mathbf{u} \in (L^2(\mathcal{D}))^d : \nabla \cdot \mathbf{u} = 0, \mathbf{u} \cdot \mathbf{n}_{|x_d=0,1} = 0, \mathbf{u} \text{ is periodic in } (x_1, \ldots, x_{d-1})\},$$
$$H_2 = \{\theta \in L^2(\mathcal{D}) : \theta \text{ is periodic in } (x_1, \ldots, x_{d-1})\}.$$

Here $\mathbf{n} = (0, \ldots, 0, \pm 1)$ is the outward normal to $\mathcal{D}$. We next set $V = V_1 \times V_2$ with

$$V_1 = \{\mathbf{u} \in (H^1(\mathcal{D}))^d : \nabla \cdot \mathbf{u} = 0, \mathbf{u}_{|x_d=0,1} = 0, \mathbf{u} \text{ is periodic in } (x_1, \ldots, x_{d-1})\},$$
$$V_2 = \{\theta \in H^1(\mathcal{D}) : \theta_{|x_d=0,1} = 0, \theta \text{ is periodic in } (x_1, \ldots, x_{d-1})\}.$$

For further background on this general functional setting see e.g. [CF88, Tem01].

We have the following general well-posedness results concerning (2.3)–(2.5). The cases $d = 2$ and $d = 3$ are quite different reflecting the situation encountered for the Navier-Stokes equations (deterministic or stochastic) in $d = 2, 3$. We start with $d = 2$.

Proposition 2.1. *Suppose $d = 2$. Fix a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W)$ and any $U_0 = (\mathbf{u}_0, \theta_0) \in L^2(\Omega, H)$ which is $\mathcal{F}_0$-measurable relative to this basis. Then there exists a unique*

$$U = (\mathbf{u}, \theta) \in L^2(\Omega; L^2_{loc}([0, \infty); V) \cap C([0, \infty); H)) \tag{2.8}$$

which is predictable (in particular $\mathcal{F}_t$-adapted) satisfying (2.3)–(2.5) *weakly. Adopting the notation $U(t, U_0)$ for the solution of* (2.3)–(2.5) *corresponding to a given (deterministic) $U_0 \in H$ we have that U is continuous in U_0 for every fixed $t \geq 0$. As such,* (2.3)–(2.5) *generates a Markov semigroup according to*

$$P_t\phi(U_0) = \mathbb{E}\phi(U(t, U_0))$$

for any bounded measurable $\phi : H \to \mathbb{R}$. Furthermore, $\{P_t\}_{t\geq 0}$ is Feller, that is, P_t maps bounded continuous functions to bounded continuous functions for every $t \geq 0$.

In $d = 3$ the results are weaker, a reflection of our incomplete understanding of the 3d Navier-Stokes equation at present. In what follows we use $Pr(H)$ to denote the space of Borel probability measures on H, and let $\mathcal{B}(H)$ denote the space of Borel measurable subsets of H.

Proposition 2.2. *Suppose $d = 3$. Given any $\mathfrak{m} \in Pr(H)$ with $\int \|(\mathbf{u}, \theta)\|_{L^2}^2 d\mathfrak{m}(\mathbf{u}, \theta) < \infty$, there exists a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0}, W)$ and a stochastic process $U = (\mathbf{u}, \theta)$, relative to that basis, with*

$$U \in L^2(\Omega; L_{loc}^2([0, \infty); V) \cap L_{loc}^\infty([0, \infty); H)).$$

Also, U is $\mathcal{F}_t$ adapted, weakly continuous, satisfies (2.3)–(2.5)*, and the law of $U(0)$ is $\mathfrak{m}$. Moreover, such a process U exists satisfying the energy inequalities* (3.32) *and* (3.35)*, and if we assume that for some $p \geq 3$ we have $\int \|\theta\|_{L^p}^2 d\mathfrak{m}(\theta) < \infty$, then we can further suppose that $\theta \in L^2(\Omega; L_{loc}^\infty([0, \infty); L^p))$.*

Remark 2.1. *Notice that the stochastic elements in the problem are fixed in advance in $d = 2$ and are obtained as a part of the solution in $d = 3$. We say that the solutions given in Proposition 2.1 are 'pathwise' solutions whereas those in $d = 3$ are 'Martingale' solutions.*

The existence and uniqueness results given in Propositions 2.1, 2.2 can be established with the aid of a Faedo-Galerkin scheme. One establishes sufficient compactness from standard a priori estimates to pass to a limit on a new stochastic basis using the Skorokhod embedding theorem. Since a priori estimates yield (pathwise) uniqueness results for the 2 dimensional case, convergence in the given stochastic basis may be recovered from a variation of the Yamada-Watanabe theorem found in [GK96]. The details of these proofs are technically involved but follow very closely the analysis in numerous previous works, see e.g. [DPZ92, FG95, Ben95, GHZ09, DGHT11] for further details.

We also have the following well-posedness result for the infinite Prandtl system (2.6)–(2.7).

Proposition 2.3. *Suppose $d = 2$ or $d = 3$. Fix a stochastic basis $\mathcal{S}$ and any $\mathcal{F}_0$-measurable random variable $\theta_0^0 \in L^2(\Omega, H_2)$. Then there exists a unique process*

$$\theta^0 \in L^2(\Omega; L_{loc}^2([0, \infty); V_2) \cap C([0, \infty); H_2)),$$

which is $\mathcal{F}_t$-adapted, weakly solves (2.6)–(2.7) *and satisfies the initial condition $\theta^0(0) = \theta_0^0$. Adopting the notation $\theta^0(t, \theta_0^0)$ for the solution of* (2.6)–(2.7) *corresponding to a given (deterministic) $\theta_0^0 \in H_2$ we have that θ^0 is continuous in θ_0^0 for every fixed $t \geq 0$. As such,* (2.6)–(2.7) *generates a Feller Markov semigroup according to*

$$P_t\phi(\theta_0^0) = \mathbb{E}\phi(\theta^0(t, \theta_0^0))$$

for any bounded measurable $\phi : H_2 \to \mathbb{R}$.

Remark 2.2. *We remark that Propositions 2.1, 2.2, and 2.3 hold with $N_1 = \infty$ or $N_2 = \infty$ (note that $N_1 = 0$ for* (2.6)–(2.7)*) provided we impose sufficient decay in the bases $\{\tilde{\sigma}_k\}$ and $\{\sigma_k\}$ for H_1 and H_2, respectively. For example, if we impose that $\sum_{k=1}^\infty \|\tilde{\sigma}_k\|_{H^s}^2 < \infty$ and $\sum_{k=1}^\infty \|\sigma_k\|_{H^s}^2 < \infty$ for $s > 0$ sufficiently large, then these propositions hold as stated.*

2.2 Notation

We denote the L^2 norm by $\| \cdot \|$ in all that follows. To simplify notation we assume that the domain $\mathcal{D}$ is such that the Poincaré constant is equal to one:

$$\|\nabla f\| \geq \|f\|.$$

We use C to denote various constants which may depend on the size of the domain. If a constant depends on other quantities, we will state this explicitly.

3 A Priori Estimates

In this section we collect some a priori estimates on solutions to (2.3)–(2.4), (2.6)–(2.7) and related systems. We begin by establishing estimates on solutions for a class of stochastic drift-diffusion equations which appear throughout our analysis. The proofs will rely on a comparison argument, an application of the maximum principle, and certain exponential martingale inequalities.

3.1 Bounds for a class of Stochastic Drift-Diffusion Equations

Fix a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W)$ and consider the following stochastic drift diffusion equation

$$d\xi + \mathbf{v} \cdot \nabla \xi dt = (\tilde{R}a \cdot v_d + \Delta \xi)dt + \sum_{k=1}^{N} \sigma_k dW^k, \quad \xi(0) = \xi_0 \tag{3.1}$$

evolving on the domain $\mathcal{D} = [0, L]^{d-1} \times [0, 1]$, where $\mathbf{v}$ is a is weakly continuous, $\mathcal{F}_t$-adapted, and divergence free vector satisfying

$$\mathbf{v} = (v_1, \ldots, v_d) \in L^2(\Omega; L^2_{loc}([0, \infty); V_1) \cap L^\infty_{loc}([0, \infty); H_1)).$$

and both $\mathbf{v}$ and ξ satisfy the mixed Dirichlet-periodic boundary conditions

$$\mathbf{v}_{|x_d=0} = \mathbf{v}_{|x_d=1} = 0, \quad \xi_{|x_d=0} = \xi_{|x_d=1} = 0, \quad \mathbf{v}, \xi \text{ are periodic in } (x_1, \ldots, x_{d-1}). \tag{3.2}$$

Recall that by the change of variable $T = \xi + \tilde{R}a(1 - x_d)$ we may reformulate (3.1) as

$$dT + \mathbf{v} \cdot \nabla T dt = \Delta T dt + \sum_{k=1}^{N} \sigma_k dW^k, \quad T(0) = T_0 = \xi_0 + \tilde{R}a(1 - x_d), \tag{3.3}$$

where

$$\mathbf{v}_{|x_d=0} = \mathbf{v}_{|x_d=1} = 0, \quad T_{|x_d=0} = \tilde{R}a, \ T_{|x_d=1} = 0, \quad \mathbf{v}, T \text{ are periodic in } (x_1, \ldots, x_{d-1}). \tag{3.4}$$

Let $q \geq 2$ and consider any initial condition $\xi_0 \in H_2 \cap L^q(\mathcal{D})$ which is $\mathcal{F}_0$-measurable and assume

$$\mathbb{E} \exp(\eta^* \|\xi_0\|_{L^q}^2) < \infty,$$

for some $\eta^* > 0$. We say ξ is a solution of (3.1) if it is weakly continuous, $\mathcal{F}_t$-adapted, and satisfies

$$\xi \in L^2(\Omega; L^2_{loc}([0, \infty); V_2) \cap L^\infty_{loc}([0, \infty); L^q) \cap C([0, \infty); H_2)), \tag{3.5}$$

where for $d = 3$ we further assume $q \geq 3$ or $q \geq 2$ and $\mathbf{v} \in L^2(\Omega; L^2_{loc}([0, \infty); H^2(\mathcal{D}))$. Furthermore, the corresponding T satisfies (3.3) in weak sense (recall $\mathbf{v}$ is divergence free). Under these assumptions, there exists a unique solution of (3.1) (see e.g. [DPZ92]).

Proposition 3.1. *For any $2 \leq p \leq q$, there exist constants $C_0 = C_0(p)$ (with $C_0(2) = 1$), $C = C(\mathcal{D}, p)$ and $\gamma = \frac{C_0}{\|\sigma\|_{L^p}^2}$ such that for each $K > 0$, the solution ξ of* (3.1) *satisfies*

$$\mathbb{P}\left(\sup_{s \geq 0}\left(\frac{1}{2}\|\xi(s)\|_{L^p}^2 + \frac{C_0(p)}{2}\int_0^s \|\xi\|_{L^p}^2 ds' - 2\|\xi_0\|_{L^p}^2\right.\right.$$
$$\left.\left. - ((p-1)\|\sigma\|_{L^p}^2 + C\tilde{R}a^2)s\right) - C\tilde{R}a^2 \geq K \,\Big|\, \mathcal{F}_0\right) \leq e^{-\gamma K}. \quad (3.6)$$

Furthermore, there exists $\eta_0 = \eta_0(\tilde{R}a, p, \|\sigma\|_{L^p})$ such that for any $t \geq 0$, and any $\eta \leq \eta_0 \wedge (\eta^*/4)$,

$$\mathbb{E}\exp\left(\eta \sup_{s \in [0,t]} \|\xi(s)\|_{L^p}^2\right) \leq C'\mathbb{E}\exp\left(4\eta\|\xi_0\|_{L^p}^2\right), \quad (3.7)$$

for a constant $C' = C'(\tilde{R}a, p, \|\sigma\|_{L^p}, |\mathcal{D}|, t)$ independent of η and ξ_0. Also, for any fixed $t > 0$, letting $\psi(s) := \exp\left(C_0(s - t)/p\right)$, we have for each $K > 0$,

$$\mathbb{P}\left(\sup_{s \in [0,t]}\left(\frac{1}{2}\psi(s)\|\xi(s)\|_{L^p}^2 + \frac{C_0}{2p}\int_0^s \psi(s')\|\xi\|_{L^p}^2 ds' - 2\psi(0)\|\xi(0)\|_{L^p}^2\right.\right.$$
$$\left.\left. - \frac{p(p-1)}{C_0}\|\sigma\|_{L^p}^2 - C\tilde{R}a^2\right) \geq K \,\Big|\, \mathcal{F}_0\right) \leq e^{-\gamma K}, \quad (3.8)$$

and for $\eta \leq \eta_0 \wedge (\eta^*/4)$,

$$\mathbb{E}\exp\left(\eta\|\xi(t)\|_{L^p}^2\right) \leq C''\mathbb{E}\exp\left(4\eta(e^{-\kappa t}\|\xi_0\|^2)\right), \quad (3.9)$$

where $C'' = C''(\tilde{R}a, p, \|\sigma\|_{L^p}, |\mathcal{D}|)$ and $\kappa = \kappa(\tilde{R}a, p, |\mathcal{D}|) > 0$ are independent of t, η and ξ_0.

Proof. In order to establish the desired bounds we take advantage of a comparison principle to work with a homogeneous version of (3.3)–(3.4) in place of (3.1)–(3.2). Consider S, the solution to

$$dS + \mathbf{v} \cdot \nabla S dt = \Delta S dt + \sum \sigma_k dW^k, \quad S(0) = T_0,$$
$$S_{|x_d=0} = S_{|x_d=1} = 0, \quad S \text{ periodic in } (x_1, \ldots, x_{d-1}). \quad (3.10)$$

Note that, under the assumed conditions on the regularity of σ and $\mathbf{v}$, the systems (3.3) and (3.10) possess unique pathwise solutions. Then since $\tilde{T} = T - \tilde{R}a$ solves

$$d\tilde{T} + \mathbf{v} \cdot \nabla\tilde{T}dt = \Delta\tilde{T}dt + \sum \sigma_k dW^k, \quad \tilde{T}(0) = T_0 - \tilde{R}a,$$
$$\tilde{T}_{|x_d=0} = 0, \; \tilde{T}_{|x_d=1} = -\tilde{R}a, \quad \tilde{T} \text{ periodic in } (x_1, \ldots, x_{d-1}),$$

we have that $R = S - \tilde{T}$ is the unique solution of

$$\partial_t R + \mathbf{v} \cdot \nabla R = \Delta R, \quad R(0) = \tilde{R}a > 0, \quad (3.11)$$
$$R_{|x_d=0} = 0, \; R_{|x_d=1} = \tilde{R}a, \quad R \text{ periodic in } (x_1, \ldots, x_{d-1}). \quad (3.12)$$

Hence, by the maximum principle, we infer that $R \geq 0$ in $\mathcal{D} \times [0, \infty)$.

Since $R \geq 0$ in $\mathcal{D} \times [0, \infty)$, this gives $S \geq T - \tilde{R}a$, so that

$$(T - \tilde{R}a)_+ \leq S_+. \quad (3.13)$$

On the other hand, $\tilde{R} = S - T$ solves

$$\partial_t \tilde{R} + \mathbf{v} \cdot \nabla \tilde{R} = \Delta \tilde{R}, \quad \tilde{R}(0) = 0,$$
$$\tilde{R}_{|z=0} = -\tilde{R}a, \quad \tilde{R}_{|z=1} = 0, \quad \tilde{R} \text{ periodic in } \mathbf{x} = (x_1, x_2).$$

Hence, again due to the maximum principle, $\tilde{R} \leq 0$ in $\mathcal{D} \times [0, \infty)$, which gives $T \geq S$ and we have

$$T_- \leq S_-. \tag{3.14}$$

Combining (3.13) and (3.14) yields $|T| = T_+ + T_- \leq (T - \tilde{R}a)_+ + T_- + \tilde{R}a \leq |S| + \tilde{R}a$ so that

$$|\xi| \leq |T| + \tilde{R}a \leq |S| + 2\tilde{R}a. \tag{3.15}$$

With the bound (3.15) in mind we proceed to estimate the solution S of (3.10) as follows. By the L^p Itō lemma (see [Kry10]),

$$d\|S\|_{L^p}^p - p \int_{\mathcal{D}} \Delta S \cdot S|S|^{p-2} dx dt = \frac{p(p-1)}{2} \sum_{k=1}^{N} \int_{\mathcal{D}} \sigma_k^2 |S|^{p-2} dx dt + p \sum_{k=1}^{N} \int_{\mathcal{D}} \sigma_k S|S|^{p-2} dx dW_k, \tag{3.16}$$

where we used that $\mathbf{v}$ is divergence free to drop the advective terms $\mathbf{v} \cdot \nabla S$. Since we seek to estimate exponential moments for $\|S\|_{L^p}^2$ we next make a second application of Itō's lemma with $\phi_p(x) := (\delta + x)^{2/p}$ for any $\delta > 0$, $p \geq 2$. Direct computation yields

$$\phi_p'(x) = \frac{2}{p}(\delta + x)^{\frac{2-p}{p}}, \quad \phi_p''(x) = \frac{2(2-p)}{p^2}(\delta + x)^{\frac{2-2p}{p}}.$$

Noting that $\phi''(x)$ is non-positive for every $x \geq 0$, we can drop the Itō correction term and infer

$$\frac{1}{p} d\phi_p(\|S\|_{L^p}^p) - \phi_p'(\|S\|_{L^p}^p) \int_{\mathcal{D}} \Delta S \cdot S|S|^{p-2} dx dt \leq \frac{(p-1)}{2} \phi_p'(\|S\|_{L^p}^p) \sum_{k=1}^{N} \int_{\mathcal{D}} \sigma_k^2 |S|^{p-2} dx dt$$

$$+ \phi_p'(\|S\|_{L^p}^p) \sum_{k=1}^{N} \int_{\mathcal{D}} \sigma_k S|S|^{p-2} dx dW_k. \tag{3.17}$$

We can integrate by parts (due to the zero Dirichlet boundary condition in (3.10)) and use a version of the Poincaré inequality (see [KS12, Proposition 7.14.1]) to obtain

$$-\int_{\mathcal{D}} \Delta S S|S|^{p-2} dx = (p-1) \int_{\mathcal{D}} |S|^{p-2} |\nabla S|^2 dx \geq C_0 \int_{\mathcal{D}} |S|^p dx = C_0 \|S\|_{L^p}^p, \tag{3.18}$$

where $C_0 = C_0(p)$ with $C_0(2) = 1$. On the other hand

$$\phi_p'(\|S\|_{L^p}^p) \sum_{k=1}^{N} \int_{\mathcal{D}} \sigma_k^2 S^{p-2} dx \leq \phi_p'(\|S\|_{L^p}^p) \|\sigma\|_{L^p}^2 \|S\|_{L^p}^{p-2} \leq \frac{2}{p} \|\sigma\|_{L^p}^2. \tag{3.19}$$

Finally note that the final term in (3.17) is the differential of a (local) Martingale whose quadratic variation is given by

$$\int_0^t (\phi_p'(\|S\|_{L^p}^p))^2 \sum_{k=1}^{N} \left(\int_{\mathcal{D}} \sigma_k S|S|^{p-2} dx \right)^2 ds.$$

We bound this term using Hölder's inequality

$$\int_0^t (\phi_p'(\|S\|_{L^p}^p))^2 \sum_{k=1}^N \left(\int_{\mathcal{D}} \sigma_k S|S|^{p-2} dx\right)^2 ds \leq \int_0^t (\phi_p'(\|S\|_{L^p}^p))^2 \left(\|\sigma\|_{L^p}\|S\|_{L^p}^{p-1}\right)^2 ds$$

$$= \frac{2}{p}\|\sigma\|_{L^p}^2 \int_0^t \phi_p'(\|S\|_{L^p}^p)(\delta + \|S\|_{L^p}^p)^{\frac{2-p}{p}} \|S\|_{L^p}^{2(p-1)} ds$$

$$\leq \frac{2}{p}\|\sigma\|_{L^p}^2 \int_0^t \phi_p'(\|S\|_{L^p}^p)\|S\|_{L^p}^p ds. \tag{3.20}$$

We will now use the following estimate: for any continuous local martingale M_t with $M_0 = 0$, for any $\gamma, K > 0$,

$$\mathbb{P}\left(\sup_{t \geq 0}\left(M_t - \frac{\gamma}{2}\langle M\rangle_t\right) \geq K\right) \leq e^{-\gamma K}. \tag{3.21}$$

In particular, we can combine the bounds (3.17)–(3.20) with the exponential martingale estimate (3.21) for any $\gamma \leq \frac{C_0}{\|\sigma\|_{L^p}^2}$ to infer, by passing to the limit as $\delta \to 0$,

$$\mathbb{P}\left(\sup_{s \geq 0}\left(\|S(s)\|_{L^p}^2 + C_0(p)\int_0^s \|S\|_{L^p}^2 ds' - \|S_0\|_{L^p}^2 - (p-1)s\|\sigma\|_{L^p}^2\right) \geq K \,\Big|\, \mathcal{F}_0\right) \leq e^{-\gamma K}. \tag{3.22}$$

Note that from (3.15) and Minkowski inequality, we have

$$\|\xi\|_{L^p}^2 \leq (\|S\|_{L^p} + 2\tilde{R}a|\mathcal{D}|^{1/p})^2 \leq 2\|S\|_{L^p}^2 + C\tilde{R}a^2 \tag{3.23}$$

and

$$\|S_0\|_{L^p}^2 = \|T_0\|_{L^p}^2 = \|\xi_0 + \tilde{R}a(1-z)\|_{L^p}^2 \leq 2\|\xi_0\|_{L^p}^2 + C\tilde{R}a^2\,, \tag{3.24}$$

where $C = C(\mathcal{D})$. Then (3.6) follows by combining (3.22)–(3.24). Also the bound (3.7) is obtained from (3.6) by invoking the elementary formula

$$\mathbb{E}|X| = \int_0^\infty \mathbb{P}(|X| \geq y)\, dy. \tag{3.25}$$

In order to establish the remaining bounds, (3.8)–(3.9), we fix $t > 0$ and set $\psi(s) := \exp\left(C_1(s-t)\right)$ with $C_1 = \frac{C_0}{p}$. Using Itō formula for $\psi(s)\phi_p(\|S(s)\|_{L^p}^p)$ yields (cf. (3.17))

$$\frac{1}{p}d(\psi\phi_p(\|S\|_{L^p}^p)) - \psi\phi_p'(\|S\|_{L^p}^p)\int_{\mathcal{D}}\Delta S \cdot S|S|^{p-2}dxdt \leq \frac{(p-1)}{2}\psi\phi_p'(\|S\|_{L^p}^p)\sum_{k=1}^N \int_{\mathcal{D}}\sigma_k^2|S|^{p-2}dxdt$$

$$+ \frac{C_1}{p}\psi\phi_p(\|S\|_{L^p}^p)dt + \psi\phi_p'(\|S\|_{L^p}^p)\sum_{k=1}^N \int_{\mathcal{D}}\sigma_k S|S|^{p-2}dxdW_k\,.$$

Using $\psi \leq 1$, $\int_0^t \psi \leq 1$, and the exponential martingale inequality (3.21), we obtain as above

$$\mathbb{P}\left(\sup_{s \in [0,t]}\left(\psi(s)\phi_p(\|S(s)\|_{L^p}^p) + C_0\int_0^s \psi(s')\Big(\phi_p'(\|S\|_{L^p}^p)\|S\|_{L^p}^p - \frac{1}{p}\phi_p(\|S\|_{L^p}^p)\Big)ds'\right.\right.$$

$$\left.\left. - \psi(0)\phi_p(\|S(0)\|_{L^p}^p) - \frac{p(p-1)}{C_0}\|\sigma\|_{L^p}^2\right) \geq K \,\Big|\, \mathcal{F}_0\right) \leq e^{-\gamma K}. \tag{3.26}$$

By observing that $x\phi_p'(x) - \frac{1}{p}\phi(x) \to \frac{1}{p}x^{\frac{2}{p}}$ as $\delta \to 0$, and changing S to ξ, we obtain (3.8) from the dominated convergence theorem. Choosing $s = t$ and using $\psi(t) = 1$ and (3.25) we obtain (3.9), completing the proof. $\qquad\square$

3.2   Bounds for the Boussinesq System

In this section we establish bounds on solutions to (2.3)–(2.4). These proofs will invoke Proposition 3.1 and the exponential martingale inequality (3.21), and will rely on carefully chosen weighted norms to deal with interacting terms in the system.

Proposition 3.2. *Fix a stochastic basis* $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W)$, *and let* $U_0 = (\mathbf{u}_0, \theta_0) \in H$ *be* $\mathcal{F}_0$- *measurable initial conditions.*

(i) *When* $d = 2$, *let* $U = (\mathbf{u}, \theta)$ *denote the corresponding pathwise solution of* (2.3)–(2.4).

(ii) *When* $d = 3$, *assume* $\mathbb{E} \exp(\eta^* \|\theta_0\|_{L^3}^2) < \infty$ *for some* $\eta^* > 0$, *and let* $U = (\mathbf{u}, \theta)$ *denote a corresponding martingale solution of* (2.3)–(2.4) *satisfying the conclusions of Proposition 3.1 (with* $q = 3$).

Let

$$\gamma := \begin{cases} \min\left(\frac{1}{4Pr\|\tilde{\sigma}\|^2}, \frac{1}{4}\right), & \text{if } \|\tilde{\sigma}\| > 0, \\ \frac{1}{4} & \text{if } \|\tilde{\sigma}\| = 0. \end{cases} \tag{3.27}$$

Then there exist constants C_2, C_3, *depending on* $Ra, \tilde{Ra}, Pr$, *and* $\|\tilde{\sigma}\|$, *and a random constant* C_1 *which further depends on* $\|\mathbf{u}_0\|, \|\theta_0\|$, *such that for each* $K > 0$,

$$\mathbb{P}\left(\sup_{t \geq 0}\left(\|\mathbf{u}(t)\|^2 + \frac{Pr}{2}\|\theta(t)\|^2 + \frac{Pr}{2}\int_0^t\left(\|\nabla\mathbf{u}(s)\|^2 + \frac{1}{2}\|\nabla\theta(s)\|^2\right)ds\right.$$
$$\left.\left. - C_2 t - C_1\right) \geq C_3 K \,\right|\, \mathcal{F}_0\right) \leq 3e^{-\gamma K}. \tag{3.28}$$

Moreover, if $Pr \geq 2$, *then the constants* $\eta_1, C_4, \kappa > 0$ *given by*

$$\eta_1 := \frac{1}{64Ra^2 + 8}\begin{cases} \min\left(\frac{1}{2Pr\|\tilde{\sigma}\|^2}, 1\right), & \text{if } \|\tilde{\sigma}\| > 0, \\ 1 & \text{if } \|\tilde{\sigma}\| = 0, \end{cases} \tag{3.29}$$

$$C_4 := 4Pr\|\tilde{\sigma}\|^2 + C(1 + Ra^2)(1 + \tilde{Ra}^2), \quad \kappa := C(1 + Ra^2), \tag{3.30}$$

are such that for any $0 < \eta < \eta_1$ *and* $t^* > 0$, *we have*

$$\mathbb{E}\exp\left(\eta\|U(t^*)\|^2\right) \leq C\mathbb{E}\exp\left(\eta(\kappa e^{-t^*/2}\|U_0\|^2 + C_4)\right). \tag{3.31}$$

Remark 3.1. *We emphasize that when* $\tilde{\sigma} = 0$ *(i.e. no stochastic forcing in the velocity component) the constants* γ, η_1, C_4 *and* κ *are independent of* Pr.

Proof. Using (2.3) we compute with the Itō formula (note that this is an inequality for $d = 3$)

$$d\|\mathbf{u}\|^2 + 2Pr\|\nabla\mathbf{u}\|^2 dt \leq \left(2PrRa\langle\theta, u_d\rangle + (Pr)^2\|\tilde{\sigma}\|^2\right)dt + 2Pr\sum\langle\tilde{\sigma}_k, \mathbf{u}\rangle d\tilde{W}^k. \tag{3.32}$$

The last term in the previous line is a Martingale with quadratic variation $4(Pr)^2\int_0^t\sum\langle\tilde{\sigma}_k, \mathbf{u}\rangle^2 ds$. Then since

$$\sum\langle\tilde{\sigma}_k, \mathbf{u}\rangle^2 \leq \|\tilde{\sigma}\|^2\|\mathbf{u}\|^2,$$

by using the Poincaré inequality, one has

$$d\|\mathbf{u}\|^2 + \left(Pr\|\nabla\mathbf{u}\|^2 - Pr(Ra)^2\|\theta\|^2 - (Pr)^2\|\tilde{\sigma}\|^2 - 2\gamma(Pr)^2\|\tilde{\sigma}\|^2\|\mathbf{u}\|^2\right)dt$$
$$\leq -2\gamma(Pr)^2\sum\langle\tilde{\sigma}_k, \mathbf{u}\rangle^2 dt + 2Pr\sum\langle\tilde{\sigma}_k, \mathbf{u}\rangle d\tilde{W}^k.$$

Applying (3.21) with $\gamma \leq 1/(4Pr\|\tilde\sigma\|^2)$, we have for any $K > 0$, that $\mathbb{P}(E_{1,K}) \leq e^{-\gamma K}$ with

$$
E_{1,K} := \left\{ \sup_{t \geq 0} \left(\|\mathbf{u}(t)\|^2 + \int_0^t \left(\frac{Pr}{2} \|\nabla \mathbf{u}(s)\|^2 - Pr(Ra)^2 \|\theta(s)\|^2 \right) ds \right. \right.
$$
$$
\left. \left. - (Pr)^2 \|\tilde\sigma\|^2 t - \|\mathbf{u}_0\|^2 \right) \geq K \right\}. \tag{3.33}
$$

Note that when $\|\tilde\sigma\| = 0$, this holds for any $\gamma > 0$. Next, since θ satisfies a drift diffusion equation of the form (3.1), we obtain from (3.6) that for any $\gamma \leq 1/4$ one has $\mathbb{P}(E_{2,K}) \leq e^{-\gamma K}$, with

$$
E_{2,K} := \left\{ \sup_{t \geq 0} \left(\|\theta(t)\|^2 + \int_0^t \|\theta(s)\|^2 ds - (C\tilde{Ra}^2 + 2)t - (C\tilde{Ra}^2 + 4\|\theta_0\|^2) \right) \geq K \right\}, \tag{3.34}
$$

where we used $C_0(2) = 1$ and $\|\sigma\| = 1$. Note that from (2.4) one also has

$$
d\|\theta\|^2 + 2\|\nabla\theta\|^2 dt \leq (2\tilde{Ra}\langle\theta, u_d\rangle + 1)dt + 2\sum \langle\sigma_k, \theta\rangle dW^k, \tag{3.35}
$$

and, with $\gamma \leq 1/4$, another application of (3.21) gives $\mathbb{P}(E_{3,K}) \leq e^{-\gamma K}$, where

$$
E_{3,K} := \left\{ \sup_{t \geq 0} \left(\|\theta(t)\|^2 + \int_0^t \left(\frac{1}{2} \|\nabla\theta(s)\|^2 - \tilde{Ra}^2 \|u(s)\|^2 \right) ds - t - \|\theta_0\|^2 \right) \geq K \right\}. \tag{3.36}
$$

Define $E_K := \cup_{j=1}^3 E_{j,K}$. Then we have $\mathbb{P}(E_K) \leq 3e^{-\gamma K}$. Furthermore, on the complement of E_K, we evaluate $(1 + \tilde{Ra}^2)E_{1,K} + (1 + \tilde{Ra}^2)PrRa^2 E_{2,K} + \frac{Pr}{2} E_{3,K}$, meaning that we are adding the inequalities defining $E_{i,K}$, and using $\|\nabla u\| \leq \|u\|$ we obtain for any $t \geq 0$,

$$
\|\mathbf{u}(t)\|^2 + \frac{Pr}{2}\|\theta(t)\|^2 + \frac{Pr}{2} \int_0^t \left(\|\nabla \mathbf{u}(s)\|^2 + \frac{1}{2}\|\nabla\theta(s)\|^2 \right) ds
$$
$$
\leq \left((1 + \tilde{Ra}^2) \left(Pr^2\|\tilde\sigma\|^2 + Ra^2 Pr(C\tilde{Ra}^2 + 2) \right) + \frac{Pr}{2} \right) t
$$
$$
+ (1 + \tilde{Ra}^2) \left(\|\mathbf{u}_0\|^2 + Ra^2 Pr(C\tilde{Ra}^2 + 4\|\theta_0\|^2) \right) + \frac{Pr}{2}\|\theta_0\|^2 + \left((1 + \tilde{Ra}^2)(PrRa^2 + 1) + \frac{Pr}{2} \right) K.
$$

By defining C_1, C_2 and C_3 as

$$
C_1 := (1 + \tilde{Ra}^2) \left(\|\mathbf{u}_0\|^2 + Ra^2 Pr(C\tilde{Ra}^2 + 4\|\theta_0\|^2) \right) + \frac{Pr}{2}\|\theta_0\|^2, \tag{3.37}
$$
$$
C_2 := (1 + \tilde{Ra}^2) \left(Pr^2\|\tilde\sigma\|^2 + Ra^2 Pr(C\tilde{Ra}^2 + 2) \right) + \frac{Pr}{2}, \tag{3.38}
$$
$$
C_3 := (1 + \tilde{Ra}^2)(PrRa^2 + 1) + \frac{Pr}{2}, \tag{3.39}
$$

the proof of (3.28) is complete.

It remains to prove (3.31). Fix $t^* > 0$ and define $\phi(t) := \exp(\frac{Pr}{2}(t - t^*))$. It follows from (3.32) that

$$
d(\phi\|\mathbf{u}\|^2) + 2\phi Pr\|\nabla\mathbf{u}\|^2 dt = \phi \left(2PrRa\langle\theta, u_2\rangle + \frac{Pr}{2}\|\mathbf{u}\|^2 + (Pr)^2\|\tilde\sigma\|^2 \right) dt + 2Pr\phi \sum \langle\tilde\sigma_k, \mathbf{u}\rangle d\tilde{W}^k.
$$

Rearranging and applying standard estimates as above we find

$$
d(\phi\|\mathbf{u}\|^2) + \phi \left(\frac{Pr}{2}\|\nabla\mathbf{u}\|^2 - 2PrRa^2\|\theta\|^2 - 2\delta Pr^2\phi\|\tilde\sigma\|^2\|\mathbf{u}\|^2 - (Pr)^2\|\tilde\sigma\|^2 \right) dt
$$
$$
\leq 2Pr\phi \sum \langle\tilde\sigma_k, \mathbf{u}\rangle d\tilde{W}^k - 2\delta Pr^2\phi^2 \sum \langle\tilde\sigma_k, \mathbf{u}\rangle^2 dt.
$$

Set $\delta := \frac{1}{4Pr\|\tilde{\sigma}\|^2}$ if $\|\tilde{\sigma}\| > 0$, or $\delta := 1$ if $\|\tilde{\sigma}\| = 0$. Using that $\phi \leq 1$, and $\int_0^{t^*} \phi \, dt < \frac{2}{Pr}$ we obtain that $\mathbb{P}(\tilde{E}_{1,K}) \leq e^{-\delta K}$, where

$$\tilde{E}_{1,K} := \left\{ \|\mathbf{u}(t^*)\|^2 - 2PrRa^2 \int_0^{t^*} \phi(t)\|\theta\|^2 dt - 2Pr\|\tilde{\sigma}\|^2 - e^{-Prt^*/2}\|\mathbf{u}_0\|^2 \geq K \right\}.$$

Since θ satisfies a drift-diffusion equation of the form (3.1), we obtain from (3.8) that $\mathbb{P}(\tilde{E}_{2,K}) \leq e^{-K/2}$, where $\psi(t) = \exp((t - t^*)/2)$ and

$$\tilde{E}_{2,K} := \left\{ \sup_{t\in[0,t^*]} \left(\frac{\psi(t)}{2}\|\theta(t)\|^2 + \frac{1}{2}\int_0^t \psi\|\theta\|^2 ds - 2e^{-t^*/2}\|\theta_0\|^2 - 1 - C\tilde{R}a^2 \right) \geq K \right\}$$

$$\supseteq \left\{ \frac{1}{2}\|\theta(t^*)\|^2 + \int_0^{t^*} \frac{\psi(s)}{2}\|\theta\|^2 ds - 2e^{-t^*/2}\|\theta_0\|^2 - 1 - C\tilde{R}a^2 \geq K \right\}.$$

We now define $\tilde{E}_K = \tilde{E}_{1,K} \cup \tilde{E}_{2,K}$ and infer that, as above $\mathbb{P}(\tilde{E}_K) \leq 2e^{-\gamma K}$ where $\gamma = \delta \wedge \frac{1}{2}$ is independent of Pr if $\tilde{\sigma} = 0$. Now on $\tilde{E}_K^c$ we estimate

$$\|\mathbf{u}(t^*)\|^2 + \frac{1}{2}\|\theta(t^*)\|^2$$

$$\leq 2PrRa^2 \int_0^{t^*} \phi(t)\|\theta\|^2 dt + 2Pr\|\tilde{\sigma}\|^2 + e^{-Prt^*/2}\|\mathbf{u}_0\|^2 + 2e^{-t^*/2}\|\theta_0\|^2 + 2 + C\tilde{R}a^2 + 2K.$$

Observing that $\phi^{1/2} \leq \psi$ for $Pr \geq 2$ we obtain on $\tilde{E}_K^c$,

$$\int_0^{t^*} \phi\|\theta\|^2 dt \leq \int_0^{t^*} \phi^{1/2}\psi\|\theta\|^2 dt \leq \left(4e^{-t^*/2}\|\theta_0\|^2 + 2 + C\tilde{R}a^2 + 2K \right) \int_0^{t^*} \phi^{1/2} dt$$

$$\leq \frac{1}{Pr}\left(C\left(e^{-t^*/2}\|\theta_0\|^2 + 1 + \tilde{R}a^2 \right) + 8K \right).$$

Noting the definitions of C_4, κ above, we conclude that, for any $t^* > 0$,

$$\mathbb{P}\left(\|U(t^*)\|^2 - C_4 - \kappa e^{-t^*/2}\|U_0\|^2 > (32Ra^2 + 4)K \right) \leq 2e^{-\gamma K}.$$

By using the formula (3.25) we can infer (3.31) for any $\eta \leq \gamma/(32Ra^2 + 4) = \eta_1$, and the proof is complete. $\quad\square$

3.3   Bounds for the Infinite Prandtl System

Next we establish estimates analogous to those in Proposition 3.2 for the infinite Prandlt system (2.6)–(2.7).

Proposition 3.3. *There exist constants $\tilde{C}_1, \tilde{C}_2, \tilde{C}_3 > 0$ depending on $Ra, \tilde{R}a$ and $\|\theta_0^0\|$ such that for each $\gamma \leq 1/4$, and any $K > 0$, the solution θ^0 to (2.6)–(2.7) with initial data θ_0^0 satisfies*

$$\mathbb{P}\left(\sup_{t\geq 0} \left(\|\theta^0(t)\|^2 + \frac{1}{2}\int_0^t \|\nabla\theta^0(s)\|^2 ds - \tilde{C}_1 - \tilde{C}_2 t \right) \geq \tilde{C}_3 K \right) \leq 2e^{-\gamma K}. \tag{3.40}$$

Furthermore, there exists $C, \tilde{C} > 0$, $\eta_0 = \eta_0(Ra, \tilde{R}a) > 0$ and $\tilde{C}_4 = \tilde{C}_4(Ra, \tilde{R}a) > 0$ such that for any $t^* > 0$ and $0 < \eta < \eta_0$,

$$\mathbb{E}\left(\exp\left(\eta\|\theta^0\|^2 + \frac{\eta e^{-t^*/4}}{2}\int_0^{t^*} \|\nabla\theta^0\|^2 ds \right) \right) \leq C \exp\left(\tilde{C}\eta(1 + Ra\tilde{R}a)e^{-t^*/4}\|\theta_0^0\|^2 + \tilde{C}_4 \right). \tag{3.41}$$

Proof. The proof of (3.40) is a modification of the proof of Proposition 3.2. First observe that from (2.6), the Poincaré inequality yields

$$\|\nabla \mathbf{u}^0\|^2 = Ra\langle u_d^0, \theta^0 \rangle \le \frac{1}{2}\|\nabla \mathbf{u}^0\|^2 + \frac{Ra^2}{2}\|\theta^0\|^2, \tag{3.42}$$

and therefore

$$\|\mathbf{u}^0\| \le \|\nabla \mathbf{u}^0\| \le Ra\|\theta^0\|. \tag{3.43}$$

Next, we consider the sets $E_{2,K}$ and $E_{3,K}$ as defined in (3.34) and (3.36), respectively. Define $A_K = \cup_{j=2,3}E_{j,K}$, and note that by the reasoning from the proof of Proposition 3.2 above, for $\gamma \le 1/4$ we have $\mathbb{P}(A_K) \le 2e^{-\gamma K}$. Furthermore, on the complement of A_K, for any $t \ge 0$, we can use (3.43) to find

$$\begin{aligned}
\|\theta^0(t)\|^2 + \frac{1}{2}\int_0^t \|\nabla\theta^0(s)\|^2 ds &\le \|\theta_0^0\|^2 + \tilde{Ra}^2 \int_0^t \|\mathbf{u}^0(s)\|^2 ds + t + K \\
&\le \|\theta_0^0\|^2 + \tilde{Ra}^2 Ra^2 \int_0^t \|\theta^0(s)\|^2 ds + t + K \\
&\le C(\tilde{Ra}^2 Ra^2 + 1)\left(\|\theta_0^0\|^2 + \tilde{Ra}^2 + \left(\tilde{Ra}^2 + 1\right)t + K\right).
\end{aligned}$$

By defining $\tilde{C}_1, \tilde{C}_2$ and $\tilde{C}_3$ in the obvious way, we obtain (3.40).

It remains to establish (3.41). Define $\psi(t) = \exp((t - t^*)/4)$, and note that from (2.4) (and $\|\sigma\| = 1$) one has

$$\begin{aligned}
d(\psi\|\theta^0\|^2) + \psi\left(\|\nabla\theta^0\|^2 - 2\tilde{Ra}\langle\theta^0, u_d^0\rangle - 2\psi\gamma\|\theta^0\|^2 - 1\right)dt \\
\le 2\psi\sum\langle\tilde{\sigma}_k, \theta^0\rangle d\tilde{W}^k - 2\psi^2\gamma\sum\langle\tilde{\sigma}_k, \theta^0\rangle^2 dt.
\end{aligned} \tag{3.44}$$

Observe that by combining (3.42) and (3.43) we have

$$\langle\theta^0, u_d^0\rangle = \frac{1}{Ra}\|\nabla\mathbf{u}^0\|^2 \le Ra\|\theta^0\|^2. \tag{3.45}$$

Then, for $\gamma = 1/4$, the exponential martingale inequality (3.21) and (3.45) give that $\mathbb{P}(F_{1,K}) \le e^{-K/4}$, where

$$F_{1,K} := \left\{\|\theta^0(t^*)\|^2 + \int_0^{t^*}\psi(s)\left(\frac{1}{2}\|\nabla\theta^0(s)\|^2 - 2Ra\tilde{Ra}\|\theta^0(s)\|^2\right)ds - 4 - e^{-\frac{t^*}{4}}\|\theta_0^0\|^2 \ge K\right\}. \tag{3.46}$$

Also define $F_{2,K}$ as $\tilde{E}_{2,K}$ in the proof of Proposition 3.2 with θ replaced by θ^0, and with the modified definition of $\psi(t)$ given in this proof. As above we obtain $\mathbb{P}(F_{2,K}) \le e^{-K/4}$, and find that on $(F_{1,K} \cup F_{2,K})^c$ we have

$$\begin{aligned}
\|\theta^0(t^*)\|^2 + \frac{e^{-\frac{t^*}{4}}}{2}\int_0^{t^*}\|\nabla\theta^0(s)\|^2 ds &\le Ra\tilde{Ra}\int_0^{t^*}\psi(s)\|\theta^0(s)\|^2 ds + 1 + e^{-\frac{t^*}{4}}\|\theta_0^0\|^2 + K \\
&\le C(Ra\tilde{Ra} + 1)(e^{-\frac{t^*}{4}}\|\theta_0^0\|^2 + 1 + \tilde{Ra}^2 + K).
\end{aligned}$$

That is, with $\tilde{C}_4 = C(Ra\tilde{Ra} + 1)(1 + \tilde{Ra}^2)$, we conclude that

$$\mathbb{P}\left(\|\theta^0(t^*)\|^2 + \frac{e^{-\frac{t^*}{4}}}{2}\int_0^{t^*}\|\nabla\theta^0(s)\|^2 ds - \tilde{C}_4 - C(Ra\tilde{Ra} + 1)e^{-\frac{t^*}{4}}\|\theta_0^0\|^2 \ge C(Ra\tilde{Ra} + 1)K\right) \le 2e^{-K/4}.$$

By using the formula (3.25) the inequality (3.41) follows, completing the proof. $\quad\square$

4   Existence of Invariant States

In this section we apply the Krylov-Bogoliubov averaging procedure [KB37] to establish the existence of ergodic invariant measures (2.3)–(2.4) when $d = 2$, and for the infinite Prandtl system (2.6)–(2.7) when $d = 2$ or $d = 3$. We also prove the existence of a statistically invariant state for (2.3)–(2.4) when $d = 3$ by adapting the method of [FG95]. That is, we provide the proof of Theorem 1.2 at the end of this section.

4.1   Invariant Measures

Proposition 4.1. *When $d = 2$ there exists an ergodic invariant measure $\mu \in Pr(H)$ for the Markov semigroup P_t corresponding to (2.3)–(2.4). In other words, $\mu P_t = \mu$ for every $t \geq 0$ and if $P_t \chi_A = \chi_A$, μ-almost surely then $\mu(A) \in \{0, 1\}$. Moreover each invariant measure μ of (2.3)–(2.4) satisfies the exponential moment bound*

$$\int_H \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|_{L^p}^2))d\mu(U) \leq C \exp(\eta C_5) < \infty \tag{4.1}$$

for all $0 < \eta < \eta_1$, $p \geq 2$, where $\eta_1 = \eta_1(p, Pr, Ra, \tilde{Ra}, \|\tilde{\sigma}\|)$ and $C_5 = C_5(p, Pr, Ra, \tilde{Ra}, \|\tilde{\sigma}\|)$. These constants are independent of Pr if $\tilde{\sigma} = 0$.

Before we present the proof of Proposition 4.1, let us remark on an important consequence which appeared in our companion work [FGHR].

Corollary 4.2. *Suppose $\tilde{\sigma} = 0$ and $d = 2$. Let $\{\mu_{Pr}\}_{Pr \geq 2}$ denote any sequence of invariant measures for (2.3)–(2.4) corresponding to increasing values of $Pr \geq 2$. For any $p \geq 2$, there exists $\eta = \eta(p, Ra, \tilde{Ra}, \|\tilde{\sigma}\|)$, independent of Pr, such that*

$$\sup_{Pr \geq 2} \int_H \exp\left(\eta\left(\|\mathbf{u}\|^2 + \|\theta\|_{L^p}^2\right)\right) d\mu_{Pr}(\mathbf{u}, \theta) < \infty. \tag{4.2}$$

Proof of Proposition 4.1. Suppose $d = 2$, fix any $(\mathbf{u}_0, \theta_0) \in H$, $T > 0$ and let $\mathbf{u}(t) = \mathbf{u}(t, \mathbf{u}_0)$, $\theta(t) = \theta(t, \theta_0)$ denote the solution of (2.3)–(2.4). Define the probability measures $\mu_T \in Pr(H)$ as

$$\mu_T(A) = \frac{1}{T} \int_0^T \mathbb{P}((\mathbf{u}, \theta)(t) \in A)dt, \qquad (A \in \mathcal{B}(H)). \tag{4.3}$$

From (3.28) we have

$$\sup_{T \geq 1} \frac{1}{T} \mathbb{E} \int_0^T \left(\|\nabla \mathbf{u}\|^2 + \|\nabla \theta\|^2\right) ds \leq C < \infty, \tag{4.4}$$

and hence with the Chebyshev inequality we infer that the collection $\{\mu_T\}_{T \geq 1}$ is tight and thus weakly compact. It follows immediately that any sub-sequence converges to an invariant measure for (2.3)–(2.4); see e.g. [DPZ92].

We next establish (4.1) as follows. Consider any invariant measure $\mu \in Pr(H)$ of (2.3)–(2.4). Fix any $R > 0$ and define $\Phi_R : H \to \mathbb{R}$ as $\phi_R(U) = \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|_{L^p}^2)) \wedge R$, where $U = (\mathbf{u}, \theta)$, η is given as in (4.1) and $R > 0$. Since ϕ_R is continuous and bounded on H, and μ is invariant, it follows that

$$\int_H \phi_R(U)d\mu(U) = \int_H P_t \phi_R(U)d\mu(U) = \int_H \int_H \phi_R(\tilde{U})P_t(U, d\tilde{U})d\mu(U), \tag{4.5}$$

for any $t \geq 0$. On the other hand, by applying (3.31) and (3.9) we infer that, for any $U \in H$,

$$\int_H \phi_R(\tilde{U})P_t(U_0, d\tilde{U}) = \mathbb{E}\phi_R(U(t, U_0)) \leq C \exp\left(\eta\left(e^{-\epsilon_1 t}\kappa_1(\|u_0\|^2 + \|\theta_0\|_{L^p}^2) + C_5\right)\right), \tag{4.6}$$

where $\epsilon_1, \kappa_1, C_5 > 0$ depend on $Ra, \tilde{Ra}, p$ and $|\mathcal{D}|$ and Pr. Note that these constants are independent of Pr if $\tilde{\sigma} = 0$.

Combining these observations we infer that, for any $t > 0$, $\rho > 0$, $R > 0$,

$$
\int_H \phi_R(U)d\mu(U) = \int_{B(\rho)} \int_H \phi_R(\tilde{U})P_t(U, d\tilde{U})d\mu(U) + \int_{B(\rho)^c} \int_H \phi_R(\tilde{U})P_t(U, d\tilde{U})d\mu(U)
$$
$$
\leq C \exp\left(\eta \left(e^{-\epsilon_1 t}\kappa_1\rho^2 + C_5\right)\right) + R\mu(B(\rho)^c). \tag{4.7}
$$

The desired result now follows by first taking $t \to \infty$ then $\rho \to \infty$ and finally $R \to \infty$ and using the monotone convergence theorem.

Working from (4.1) we can establish further uniform moment bounds in H^1. Returning to (3.32), (3.35) we infer

$$
2\mathbb{E}(\|\nabla\mathbf{u}\|^2 + \|\nabla\theta\|^2) = \mathbb{E}\left(2(Ra + \tilde{Ra})\langle\theta, u_2\rangle + Pr\|\tilde{\sigma}\|^2 + 1\right)
$$
$$
\leq \mathbb{E}\left((Ra + \tilde{Ra})(\|\theta\|^2 + \|u_2\|^2) + Pr\|\tilde{\sigma}\|^2 + 1\right) \tag{4.8}
$$

for any stationary solution of (2.3)–(2.4). Taking $\mathcal{I}$ to be the collection of invariant measures of $\{P_t\}_{t\geq 0}$, and combining (4.8) with (4.1) we infer

$$
\sup_{\mu\in\mathcal{I}} \int_H \|\nabla U\|^2 d\mu(U) < \infty. \tag{4.9}
$$

So far we have only obtained the existence of an invariant measure, in particular $\mathcal{I}$ is non-empty. It is easy to see that $\mathcal{I}$ is convex and closed. In view of (4.9) we have furthermore that the collection $\mathcal{I}$ is tight and hence compact. We may thus infer the existence of an extremal point in $\mathcal{I}$ from Krein-Milman theorem. Since, cf. [DPZ92], the ergodic invariant measure consist of the extremal points of $\mathcal{I}$, we infer that $\mathcal{I}$ must contain an ergodic invariant measure, completing the proof of the first item. □

Proposition 4.3. *The Markov semigroup corresponding to* (2.6)–(2.7) *possesses an ergodic invariant measure* $\mu \in Pr(H_2)$, *and each invariant measure satisfies the exponential moment bound*

$$
\int_H \exp(\eta\|\theta^0\|_{L^p}^2)d\mu(\theta^0) \leq C \exp(\eta\tilde{C}_4) < \infty, \tag{4.10}
$$

for all $0 < \eta < \eta_2$, $p \geq 2$, *where* $\eta_2 = \eta_2(p, Ra, \tilde{Ra}) > 0$, $\tilde{C}_5 = \tilde{C}_5(p, Ra, \tilde{Ra})$.

Proof of Proposition 4.3. By invoking Proposition 3.1, this requires a simple modification of the proof of Proposition 4.1, and we omit the details. □

4.2 Existence of Statistically Invariant States

In three space dimensions we establish the existence of statistically invariant states. Namely, we rephrase and prove Theorem 1.2 as follows.

Proposition 4.4. *When* $d = 3$ *there exists a measure* $\mu \in Pr(H)$ *and a corresponding (Martingale) solution of* (2.3)–(2.4) *which is stationary in time, namely* $\mu_t(\cdot) = \mathbb{P}(U(t) \in \cdot)$ *is identically equal to* μ. *Moreover* μ *satisfies an exponential moment bound as in* (4.1).

As above, we mention a useful corollary of Proposition 4.4 which appeared in the manuscript [FGHR].

Corollary 4.5. *Suppose* $\tilde{\sigma} = 0$ *and* $d = 3$. *For any* $p \geq 2$, *the system* (2.3)–(2.4) *possesses a sequence of statistically invariant states* $\{\mu_{Pr}\}_{Pr\geq 2}$ *such that* (4.2) *is satisfied.*

One can prove Corollary 4.5 by combining (4.1) with Pr-uniform bounds on the $\|\theta\|_{L^p}^2$ term in (4.2). These latter bounds follow by using (3.9) and mimicking the proof of (4.1) in the case $d = 3$ (see below).

Proof of Proposition 4.4. We establish the existence of a statistically invariant state using a modified Galerkin truncation scheme. Let P_N be the projection onto the first N eigenfunctions of the Stokes operator, which are of course divergence free. Define

$$d\mathbf{u}^N + P_N(\mathbf{u}^N \cdot \nabla \mathbf{u}^N) = Pr P_N \Delta \mathbf{u}^N dt + Pr Ra P_N \hat{\mathbf{e}}_d \theta^N dt + Pr \sum_{k=1}^{N_1} P_N \tilde{\sigma}_k d\tilde{W}^k, \quad \mathbf{u}^N(0) = P_N \mathbf{u}_0 \quad (4.11)$$

$$d\theta^N + \mathbf{u}^N \cdot \nabla \theta^N dt = u_d^N dt + \Delta \theta^N dt + \sum_{k=1}^{N_2} \sigma_k dW^k, \quad \theta^N(0) = \theta_0. \quad (4.12)$$

Observe that for each N there is a unique pathwise solution $U^N(\cdot, U_0)$ which depends continuously on its initial condition $U_0 = (\mathbf{u}_0, \theta_0) \in H$. Take $\{P^N\}_{t\geq 0}$ to be the associated Markov semigroup. As in Proposition 3.2 we have

$$\mathbb{P}\left(\sup_{t\geq 0}\left(\|\mathbf{u}^N(t)\|^2 + \frac{Pr}{2}\|\theta^N(t)\|^2 + \frac{Pr}{2}\int_0^t \left(\|\nabla \mathbf{u}^N\|^2 + \frac{1}{2}\|\nabla \theta^N\|^2\right)ds - C_2 t - C_1\right) \geq C_3 K\right) \leq 3e^{-\gamma K}, \quad (4.13)$$

where the constants C_1, C_2, C_3 have the same meaning as in (3.37)–(3.39). To see this observe that we have only used the cancellation property in the nonlinear term to analyze the momentum equation. We have deliberately avoided truncating the equation for θ^N so that θ^N still satisfies drift-diffusion equation (3.1).

Making use of (4.13) we may now obtain the existence of an invariant measure μ^N for each N by implementing the Krylov-Bogoliubov procedure and arguing precisely as in the $d = 2$ case. We observe moreover that

$$\int_H \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|_{L^p}^2))d\mu^N(U) \leq \exp(\eta C_1), \quad (4.14)$$

where, crucially, the constant C_1 in this upper bound is independent of N. Here again the arguments in the proof of (4.1) pass through virtually line by line to the present case.

To a sequence of invariant measures $\{\mu_N\}_{N\geq 1}$ we may now associate a a sequence of stationary solutions $\bar{U}^N$ of (4.11)–(4.12). An involved limiting procedure very similar to e.g. [FG95] (see also [KS12, GHŠV15]) can be used to pass to a limit in this class of stationary solutions. We briefly sketch some details.

By repeating the computations leading to (4.13) with the stationary solutions $\bar{U}^N$ we obtain that

$$\bar{U}^N \text{ is uniformly bounded in } L^p(\Omega; L^2([0,t^*];V) \cap L^\infty([0,t^*];H)) \quad (4.15)$$

for any $p \geq 1$ and any $t^* > 0$. On the other hand writing (4.11)–(4.12) as an abstract evolution equation on H we may write $\bar{U}^N(t) = \bar{U}^N(0) + \int_0^t F_N(\bar{U}^N)ds + \sigma_N W(t)$ where F_N is a suitable abstract operator; see e.g. [CF88, Tem01] for precise details. Define now $X = V \cap (H^3(\mathcal{D}))^4$ let X^* be its dual, relative to H. For any $\alpha \in (0, 1/2)$ and $p > 1$ we have that

$$\mathbb{E}\|\bar{U}^N\|_{W^{\alpha,p}([0,t^*];X^*)}^p \leq C\mathbb{E}\left\|\bar{U}^N(0) + \int_0^\cdot F_N(\bar{U}^N)ds\right\|_{W^{1,p}([0,t*];X^*)}^p + C\mathbb{E}\|\sigma_N W(t)\|_{W^{\alpha,p}([0,t^*];H)}^p$$

here $W^{\alpha,p}([0,T];X^*)$ is the Sobolev-Slobodeckij space. Since $\|F_N(U)\|_{X^*} \leq C(\|U\|^2 + 1)$ for any $U \in H$, where the constant C does not depend on N we infer that

$$\mathbb{E}\left\|\bar{U}^N(0) + \int_0^\cdot F_N(\bar{U}^N)ds\right\|_{W^{1,p}([0,t^*];X^*)}^p \leq C\mathbb{E}\left(\sup_{s\in[0,t^*]}\|\bar{U}^N\|^{2p} + 1\right).$$

Moreover, given the standard time regularity properties of Brownian motions we obtain the uniform bound

$$\sup_{N \geq 1} \mathbb{E}\|\sigma_N W(t)\|^p_{W^{\alpha,p}([0,t^*];H)} < \infty\,.$$

Combining these bounds with (4.15) we now obtain that

$$\bar{U}^N \text{ is uniformly bounded in } L^p(\Omega; W^{\alpha,p}([0,T]; X^*)) \tag{4.16}$$

for any $\alpha < 1/2$ and $p > 1$.

Take $Y = V \cap (H^4(\mathcal{D}))^4$ and Y^* its dual relative to H. We now show that the collection of measures $\{\mu_N\}_{N \geq 1}$ given the laws of $\bar{U}^N$ and defined on the path-space $C([0,t^*]; Y^*)$ are compact. To this end recall that we have the compact embeddings

$$L^2([0,t^*]; V) \cap W^{1/4,2}([0,t^*]; X^*) \subset\subset L^2([0,t^*]; H), \quad W^{1/4,8}([0,t^*]; X^*) \subset\subset C([0,t^*]; Y^*) \tag{4.17}$$

which are variations on the Aubin-Lions and Arzelà-Ascoli compactness theorems respectively; see [FG95] and also [DGHT11] for further details. Combining the (4.17) with (4.15), (4.16) we obtain that, after passing to a subsequence μ_N converges weakly in $L^2([0,t^*]; H) \cap C([0,t^*]; Y^*)$.

By changing the stochastic basis we can, by Skorokhod's embedding theorem, assume that $\bar{U}^N$ converges almost surely in $L^2([0,t^*]; H)$, and in $C([0,t^*]; Y^*)$, to an element $\bar{U}$, while maintaining the uniform bound (4.15). Additionally, since $\bar{U}^N$ is stationary, in view of the convergence in $C([0,t^*]; Y^*)$, it follows that $\bar{U}$ is stationary too. In order to show that $\bar{U}$ is a Martingale solution of the system (2.3)–(2.4), one needs to obtain an appropriate stochastic basis in conjunction with this limiting procedure. Here several approaches have been developed and can be used, see [DPZ92, FG95, Ben95, DGHT11, HS12].

Finally we remark that the measure μ we obtain as the weak limit of $\{\mu^N\}$ satisfies (4.1). Indeed by (4.14), $\int_H \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|^2_{L^p}) \wedge R d\mu^N(U) \leq \exp(nC_1)$ for every $R > 0$, where the constant C_1 is independent of R. using first weak convergence and then the monotone convergence theorem we can now pass to a limit first in N and then in R. With this the proof of Proposition 4.4, and thus of Theorem 1.2 is complete. $\square$

5 Unique Ergodicity Results

We proceed to prove the main results on unique ergodicity in this work, Theorem 1.1 and Theorem 1.3. We will begin by presenting an abstract framework for proving unique ergodicity by asymptotic coupling following [HMS11]. This framework will serve to simplify our analysis. Recall that the strategy for proving uniqueness via this framework was expanded upon for multiple examples in forthcoming work of a subset of the coauthors in collaboration with Mattingly, [GHMR].

5.1 Abstract Results from Ergodic Theory

To prove our ergodic theorems we will make use of the following abstract result from [HMS11]. Let P be a Markov transition function on a Polish space (H, ρ). Denote the collection of all Borel probability measures on H by $Pr(H)$. Take P_∞ to be the associated transition function on the path space of infinite one-sided sequences $H_\infty = H^{\mathbb{N}}$, and for $\mu \in Pr(H)$, let $\mu P_\infty \in Pr(H_\infty)$ be the measure defined by $\int_{H_\infty} P_\infty(x, \cdot) d\mu(x)$. Given $\mu_1, \mu_2 \in \Pr(H)$, consider

$$\tilde{\mathcal{C}}_\infty(\mu_1, \mu_2) := \left\{\Gamma \in Pr(H_\infty \times H_\infty) : \Pi_i^\# \Gamma \ll \mu_i P_\infty \text{ for each } i \in \{1, 2\}\right\},$$

where Π_i is the projection onto the ith coordinate and $f^\# \mu$ is a push forward of the measure μ, that is, $f^\# \mu(A) = \mu(f^{-1}(A))$. We will also denote $D := \{(x, y) \in H_\infty \times H_\infty : \lim_{n \to \infty} \rho(x_n, y_n) = 0\}$.

Theorem 5.1 ([HMS11])**.** *Suppose there exists a Borel measurable set $\mathcal{A} \subset H$ such that*

(i) for any P invariant Borel probability measure μ, $\mu(\mathcal{A}) > 0$,

(ii) there exists a measurable map $\Gamma : \mathcal{A} \times \mathcal{A} \to Pr(H_\infty \times H_\infty)$ *such that, for all* $x, y \in \mathcal{A}$, $\Gamma(x, y) \in \tilde{\mathcal{C}}_\infty(\delta_x, \delta_y)$ *and* $\Gamma(x, y)(D) > 0$,

then there is at most one invariant probability measure for P.

5.2   Proof of Theorem 1.1

Proof of Theorem 1.1. Let P_∞ be the transition function on H_∞, with the product topology, corresponding to equations (2.3)–(2.4) evaluated at integer times. For a given $U = (\mathbf{u}, \theta)$ satisfying (2.3)–(2.4) and $\tilde{U}_0 = (\tilde{\mathbf{u}}_0, \tilde{\theta}_0)$, define $\tilde{U} = (\tilde{\mathbf{u}}, \tilde{\theta})$ as the solution of

$$\frac{1}{Pr}(d\tilde{\mathbf{u}} + \tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} dt) + \nabla \tilde{p} dt = \Delta \tilde{\mathbf{u}} dt + Ra \hat{\mathbf{e}}_d \tilde{\theta} dt + \sum_{k=1}^{N_1} \tilde{\sigma}_k d\tilde{W}^k \tag{5.1}$$
$$- \lambda_1 P_{N_1}(\tilde{\mathbf{u}} - \mathbf{u}) \mathbb{1}_{t < \tau_R} dt, \quad \nabla \cdot \tilde{\mathbf{u}} = 0,$$

$$d\tilde{\theta} + \tilde{\mathbf{u}} \cdot \nabla \tilde{\theta} dt = \tilde{Ra} \tilde{u}_d dt + \Delta \tilde{\theta} dt + \sum_{k=1}^{N_2} \sigma_k dW^k - \lambda_2 P_{N_2}(\tilde{\theta} - \theta) \mathbb{1}_{t < \tau_R} dt, \tag{5.2}$$

with initial data $\tilde{U}_0$. Here $\mathbb{1}_{t < \tau_R}$ is the characteristic function of $\{t < \tau_R\}$, where

$$\tau_R = \inf \left\{ t > 0 : \int_0^t \left(\|\sigma^{-1} P_{N_1}(\tilde{\mathbf{u}} - \mathbf{u})\|^2 + \|\tilde{\sigma}^{-1} P_{N_2}(\tilde{\theta} - \theta)\|^2 \right) ds \geq R \right\}. \tag{5.3}$$

Let $\tilde{P}_\infty$ be the transition function on H_∞ associated to (5.1)–(5.2), and as above, for fixed $\mu \in Pr(H)$, let $\mu P_\infty \in Pr(H_\infty)$ denote the probability measure $\mu \tilde{P}_\infty(\cdot) = \int_{H_\infty} \tilde{P}_\infty(x, \cdot) d\mu(x)$. This proof will proceed by applying Theorem 5.1 with $\mathcal{A} = H$ and $\Gamma = \Gamma_{U_0, \tilde{U}_0} := \delta_{U_0} P_\infty \times \delta_{\tilde{U}_0} \tilde{P}_\infty$ for each $U_0, \tilde{U}_0 \in H$. In order to do this, we first show that, with fixed initial conditions, the laws of solutions to (2.3)–(2.4) and (5.1)–(5.2) are equivalent.

Recall that $\mathbf{W} = (\tilde{W}, W) : \Omega \times (0, \infty) \to \mathbb{R}^{N_1 + N_2}$ is a standard Wiener process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We define the "Girsanov shift"

$$a(t) := \begin{pmatrix} \tilde{\sigma}^{-1} \lambda_1 P_{N_1}(\tilde{\mathbf{u}} - \mathbf{u}) \\ \sigma^{-1} \lambda_2 P_{N_2}(\tilde{\theta} - \theta) \end{pmatrix} \mathbb{1}_{t < \tau_R},$$

and let

$$\mathbf{W}_0(t) = \mathbf{W}(t) - \int_0^t a(s) ds. \tag{5.4}$$

We also consider

$$D(t) := \exp \left(\int_0^t a(s) \cdot d\mathbf{W} - \frac{1}{2} \int_0^t \|a(s)\|^2 ds \right). \tag{5.5}$$

We can easily check that $D(t)$ is a continuous martingale, and since $a(t)$ satisfies the Novikov condition (see Proposition 1.15 of [RY99]),

$$\mathbb{E} \exp \left(\frac{1}{2} \int_0^\infty \|a(s)\|^2 ds \right) = \mathbb{E} \exp \left(\frac{1}{2} \int_0^{\tau_R} \|a(s)\|^2 ds \right) \leq e^{\frac{1}{2}R} < \infty,$$

it follows that $\{D(t)\}_{t \geq 0}$ is uniformly integrable. By the Girsanov theorem, (see Propositions 1.1-1.4 of [RY99]) there is a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$ such that the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathcal{F}_t$ is $D(t)$, and under $\mathbb{Q}$, $\mathbf{W}_0$ is a standard Wiener process. Moreover, it holds that $\mathbb{P} \sim \mathbb{Q}$ on $\mathcal{F}_\infty = \sigma(\cup_{t \geq 0} \mathcal{F}_t)$.

Let $U = U(t, U_0, \mathbf{W}(t))$ and $\tilde{U} = \tilde{U}(t, U_0, \mathbf{W}(t))$ denote the solutions to (2.3)–(2.4) and (5.1)–(5.2), respectively, at time $t \geq 0$, with initial condition U_0, and with stochastic forcing $\mathbf{W}$. Then by the uniqueness of solutions (see e.g. Theorem 2.4.6 of [KS12]) we have that, almost surely,

$$\tilde{U}(\cdot, U_0, \mathbf{W}(\cdot)) = U(\cdot, U_0, \mathbf{W}_0(\cdot)). \tag{5.6}$$

Next recall that, for each fixed initial condition U_0, the solutions of (2.3)–(2.4) and (5.1)–(5.2) induce measurable maps, which we denote by Φ_{U_0} and $\tilde{\Phi}_{U_0}$, respectively, from $(\Omega, \mathcal{F}, \mathbb{P})$ into $C([0, \infty); H)$. It follows that the law of solutions to (2.3)–(2.4) is given by $(\Phi_{U_0})_\# \mathbb{P}$, and similarly, the law associated to (5.1)–(5.2) is $(\tilde{\Phi}_{U_0})_\# \mathbb{P}$. By (5.6), for any $\mathcal{F}_\infty$-measurable $G \subset C([0, \infty); H)$,

$$(\tilde{\Phi}_{U_0})_\# \mathbb{P}(G) = \mathbb{P}\left(\tilde{U}(\cdot, U_0, \mathbf{W}(\cdot)) \in G\right) = \mathbb{P}\left(U(\cdot, U_0, \mathbf{W}_0(\cdot)) \in G\right), \tag{5.7}$$

whereas, since $\mathbf{W}_0$ is a standard Wiener process under $\mathbb{Q}$,

$$(\Phi_{U_0})_\# \mathbb{P}(G) = \mathbb{P}\left(U(\cdot, U_0, \mathbf{W}(\cdot)) \in G\right) = \mathbb{Q}\left(U(\cdot, U_0, \mathbf{W}_0(\cdot)) \in G\right). \tag{5.8}$$

By combining (5.7) and (5.8) with $\mathbb{P} \sim \mathbb{Q}$ on $\mathcal{F}_\infty$, we conclude that $(\tilde{\Phi}_{U_0})_\# \mathbb{P} \sim (\Phi_{U_0})_\# \mathbb{P}$. That is, with the same initial conditions, the laws of solutions to (2.3)–(2.4) and (5.1)–(5.2) are equivalent on $C([0, \infty); H)$.

Taking $\mathcal{A} = H$ and setting $\Gamma = \Gamma_{U_0, \tilde{U}_0} := \delta_{U_0} P_\infty \times \delta_{\tilde{U}_0} \tilde{P}_\infty$ for each $U_0, \tilde{U}_0 \in H$, we proceed to check that the assumptions of Theorem 5.1 are satisfied. Clearly H is a Polish space of nonzero measure for every invariant measure. From what was proven above, we have $\Pi_2^\# \Gamma_{U_0, \tilde{U}_0} = \delta_{\tilde{U}_0} \tilde{P}_\infty \ll \delta_{\tilde{U}_0} P_\infty$, and it follows that $\Gamma_{U_0, \tilde{U}_0} \in \tilde{\mathcal{C}}_\infty(\delta_{U_0}, \delta_{\tilde{U}_0})$. It remains to show that $\Gamma_{U_0, \tilde{U}_0}(D) > 0$.

Set $V = (\mathbf{v}, \phi) := \tilde{U} - U$. We obtain that

$$\frac{1}{Pr}(\partial_t \mathbf{v} + \tilde{\mathbf{u}} \cdot \nabla \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{u}) + \nabla p = \Delta \mathbf{v} + Ra\hat{\mathbf{e}}_d \phi - \lambda_1 P_{N_1} \mathbf{v} \mathbb{1}_{t < \tau_R}, \tag{5.9}$$

$$\partial_t \phi + \tilde{\mathbf{u}} \cdot \nabla \phi + \mathbf{v} \cdot \nabla \theta = \tilde{Ra} v_d + \Delta \phi - \lambda_2 P_{N_2} \phi \mathbb{1}_{t < \tau_R}. \tag{5.10}$$

Standard energy estimates yield

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{v}\|^2 + Pr\|\nabla \mathbf{v}\|^2 + \lambda_1 Pr \mathbb{1}_{t < \tau_R}\|P_{N_1} \mathbf{v}\|^2 = \int_{\mathcal{D}} (PrRa\phi v_d - \mathbf{v} \cdot \nabla \mathbf{u} \cdot \mathbf{v}) dx$$

and

$$\frac{1}{2}\frac{d}{dt}\|\phi\|^2 + \|\nabla \phi\|^2 + \lambda_2 \mathbb{1}_{t < \tau_R}\|P_{N_2} \phi\|^2 = \int_{\mathcal{D}} (\tilde{Ra} v_d \phi - \phi \mathbf{v} \cdot \nabla \theta) dx.$$

That is,

$$\frac{1}{2}\frac{d}{dt}(\|\mathbf{v}\|^2 + \|\phi\|^2) + Pr\|\nabla \mathbf{v}\|^2 + \|\nabla \phi\|^2 + \lambda_1 Pr \mathbb{1}_{t < \tau_R}\|P_{N_1} \mathbf{v}\|^2 + \lambda_2 \mathbb{1}_{t < \tau_R}\|P_{N_2} \phi\|^2$$

$$\leq \int_{\mathcal{D}} \left((\tilde{Ra} + PrRa)|\phi v_d| + |\mathbf{v} \cdot \nabla \mathbf{u} \cdot \mathbf{v}| + |\phi \mathbf{v} \cdot \nabla \theta|\right) dx. \tag{5.11}$$

We estimate the last two terms on the right-hand side of (5.11) as follows

$$\int_{\mathcal{D}} (|\mathbf{v} \cdot \nabla \mathbf{u} \cdot \mathbf{v}| + |\phi \mathbf{v} \cdot \nabla \theta|) dx \leq \|\mathbf{v}\|_{L^4}^2 \|\nabla \mathbf{u}\| + \|\phi\|_{L^4}\|\mathbf{v}\|_{L^4}\|\nabla \theta\|$$

$$\leq C \left(\|\mathbf{v}\|\|\nabla \mathbf{v}\|\|\nabla \mathbf{u}\| + \|\phi\|^{1/2}\|\nabla \phi\|^{1/2}\|\mathbf{v}\|^{1/2}\|\nabla \mathbf{v}\|^{1/2}\|\nabla \theta\|\right)$$

$$\leq \frac{Pr}{2}\|\nabla \mathbf{v}\|^2 + \frac{1}{2}\|\nabla \phi\|^2 + \frac{C}{Pr}\|\mathbf{v}\|^2\|\nabla \mathbf{u}\|^2 + \frac{C}{\sqrt{Pr}}\left(\|\phi\|^2 + \|\mathbf{v}\|^2\right)\|\nabla \theta\|^2.$$

From this estimate and (5.11) we obtain

$$\frac{1}{2}\frac{d}{dt}(\|\mathbf{v}\|^2 + \|\phi\|^2) + \frac{1}{2}(Pr\|\nabla\mathbf{v}\|^2 + \|\nabla\phi\|^2) + Pr\lambda_1 \mathbb{1}_{t<\tau_R}\|P_{N_1}\mathbf{v}\|^2 + \lambda_2 \mathbb{1}_{t<\tau_R}\|P_{N_2}\phi\|^2$$
$$\leq (\tilde{Ra} + PrRa)\int_{\mathcal{D}} |\phi v_d|dx + \frac{C}{\sqrt{Pr}}(\|\mathbf{v}\|^2\|\nabla\mathbf{u}\|^2 + (\|\phi\|^2 + \|\mathbf{v}\|^2)\|\nabla\theta\|^2). \qquad (5.12)$$

We proceed to establish part (i) of Theorem 1.1. For this we consider, for each $K > 0$, the events

$$\tilde{E}_K := \left\{ \sup_{t\geq 0}\left(\|\mathbf{u}(t)\|^2 + \frac{Pr}{2}\|\theta(t)\|^2 + \frac{Pr}{2}\int_0^t \left(\|\nabla\mathbf{u}(s)\|^2 + \frac{1}{2}\|\nabla\theta(s)\|^2\right)ds - C_1 - C_2 t\right) \leq C_3 K\right\},$$

where C_1, C_2, C_3 are the constants defined in (3.37)–(3.39). By Proposition 3.2 we can choose $K = K(Pr, Ra) > 0$ such that $\mathbb{P}(\tilde{E}_K) > 0$. We first show that for already fixed $K > 0$ there exists $R = R(K, \|\mathbf{u}_0\|, \|\theta_0\|, \|\tilde{\sigma}\|, Pr, Ra, \tilde{Ra})$, such that $\tau_R = \infty$ on $\tilde{E}_K$. Indeed, assume τ_R is finite for some element of $\tilde{E}_K$. Then for $t < \tau_R$, we obtain

$$\lambda_1 \mathbb{1}_{t<\tau_R}\|P_{N_1}\mathbf{v}\|^2 = \lambda_1\|P_{N_1}\mathbf{v}\|^2 \qquad (5.13)$$

Moreover, for any $\lambda_1 > 0$ (specified below), we take integer $N_1 \geq \sqrt{2\lambda_1/C}$, where C is a constant depending only on the domain, such that

$$\lambda_1\|\mathbf{v}\|^2 \leq \lambda_1\|P_{N_1}\mathbf{v}\|^2 + \frac{C\lambda_1}{N_1^2}\|\nabla Q_{N_1}\mathbf{v}\|^2 \leq \lambda_1\|P_{N_1}\mathbf{v}\|^2 + \frac{1}{2}\|\nabla\mathbf{v}\|^2, \qquad (5.14)$$

where we used $\|\nabla Q_N v\|^2 \geq \kappa_N\|Q_N v\|^2$, with κ_N being the Nth eigenvalue of the Stokes operator with the Dirichlet boundary conditions, and $\kappa_N \approx N^2$.

Similarly, for any $\lambda_2 > 0$, we take $N_2 \geq \sqrt{2\lambda_2/C}$ and denote $\lambda := (Pr\lambda_1) \wedge \lambda_2$. Then, we have

$$\frac{1}{2}\frac{d}{dt}(\|\mathbf{v}\|^2 + \|\phi\|^2) + \lambda(\|\mathbf{v}\|^2 + \|\phi\|^2)$$
$$\leq (\tilde{Ra} + PrRa)\int_{\mathcal{D}}|\phi v_d|dx + \frac{C}{\sqrt{Pr}}(\|\mathbf{v}\|^2\|\nabla\mathbf{u}\|^2 + (\|\phi\|^2 + \|\mathbf{v}\|^2)\|\nabla\theta\|^2). \qquad (5.15)$$

Using

$$\int_{\mathcal{D}}|\phi v_d|dx \leq \|\phi\|^2 + \|\mathbf{v}\|^2,$$

we obtain, under the condition $\lambda \geq 2(\tilde{Ra} + PrRa)$,

$$\frac{d}{dt}(\|\mathbf{v}\|^2 + \|\phi\|^2) + \lambda\left(\|\mathbf{v}\|^2 + \|\phi\|^2\right) \leq \frac{C}{\sqrt{Pr}}\left(\|\phi\|^2 + \|\mathbf{v}\|^2\right)\left(\|\nabla\mathbf{u}\|^2 + \|\nabla\theta\|^2\right). \qquad (5.16)$$

Then by Grönwall's inequality,

$$\|\mathbf{v}(t)\|^2 + \|\phi(t)\|^2 \leq (\|\mathbf{v}_0\|^2 + \|\phi_0\|^2)\exp\left(\frac{C}{\sqrt{Pr}}\int_0^t (\|\nabla\mathbf{u}(s)\|^2 + \frac{1}{2}\|\nabla\theta(s)\|^2)ds - \lambda t\right)$$

and thus on $\tilde{E}_K$,

$$\|\mathbf{v}(t)\|^2 + \|\phi(t)\|^2 \leq (\|\mathbf{v}_0\|^2 + \|\phi_0\|^2)\exp\left(\frac{C}{Pr^{3/2}}(C_3 K + C_1 + C_2 t) - \lambda t\right). \qquad (5.17)$$

It follows that, for $t < \tau_R$, by fixing $\lambda > 0$ such that $\lambda > \frac{CC_2}{Pr^{3/2}}$, we have

$$\int_0^t \left(\|\lambda_1 \tilde{\sigma}^{-1} P_{N_1} \mathbf{v}(s)\|^2 + \|\lambda_2 \sigma^{-1} P_{N_2} \phi(s)\|^2 \right) ds$$

$$\leq \tilde{C} \lambda (\|\mathbf{v}_0\|^2 + \|\phi_0\|^2) \exp\left(\frac{C}{Pr^{3/2}} (C_3 K + C_1) \right) \int_0^t \exp\left(\left(\frac{CC_2}{Pr^{3/2}} - \lambda \right) s \right) ds$$

$$\leq \tilde{C} (\|\mathbf{v}_0\|^2 + \|\phi_0\|^2) \exp\left(\frac{C}{Pr^{3/2}} (C_3 K + C_1) \right),$$

where $\tilde{C} = \tilde{C}(Pr, \|\tilde{\sigma}\|)$ is, in particular, independent of R. Choose $R > 0$ such that

$$R > 2\tilde{C} (\|\mathbf{v}_0\|^2 + \|\phi_0\|^2) \exp\left(\frac{C}{Pr^{3/2}} (C_3 K + C_1) \right). \tag{5.18}$$

Then, for $0 < t < \tau_R$, we have

$$\int_0^t \left(\|\lambda_1 \tilde{\sigma}^{-1} P_{N_1} \mathbf{v}(s)\|^2 + \|\lambda_2 \sigma^{-1} P_{N_2} \phi(s)\|^2 \right) ds < \frac{R}{2},$$

which contradicts the assumption that τ_R is finite.

It follows that for these choices of K, R, and λ, one has that on $\tilde{E}_K$, the estimate

$$\|\mathbf{v}(t)\|^2 + \|\phi(t)\|^2 \leq C \exp\left(\left(\frac{CC_2}{Pr^{3/2}} - \lambda \right) t \right)$$

holds for all $t > 0$, and therefore

$$\tilde{E}_K \subset \left\{ \{(U(n, U_0), \tilde{U}(n, \tilde{U}_0))\}_{n \in \mathbb{N}} \in D \right\}, \tag{5.19}$$

where $D = \{(v, w) \in H_\infty \times H_\infty : \lim_{n \to \infty} \|v_n - w_n\| = 0\}$. Moreover, from (5.19), Proposition 3.2, and our choice of K, it follows that

$$\Gamma_{U_0, \tilde{U}_0}(D) = P\left(\{(U(n, U_0), \tilde{U}(n, \tilde{U}_0))\}_{n \in \mathbb{N}} \in D \right) \geq P(\tilde{E}_K) > 0. \tag{5.20}$$

By Theorem 5.1 the proof of part (i) of Theorem 1.1 is complete.

We proceed to establish part (ii) of Theorem 1.1, for which we will take $\lambda_1 = 0$.

Once again, with an appropriate choice of R, we will show that $\tau_R = \infty$ on $\tilde{E}_K$. Indeed, assume τ_R is finite for some element of $\tilde{E}_K$. Then for $t < \tau_R$, $N_2 \geq \sqrt{2\lambda_2/C} > 0$, we have from (5.12), using the inverse Poicaré inequality as before, that

$$\frac{1}{2} \frac{d}{dt} (\|\mathbf{v}\|^2 + \|\phi\|^2) + \frac{Pr}{2} \|\mathbf{v}\|^2 + \lambda_2 \|\phi\|^2$$

$$\leq (\tilde{R}a + PrRa) \int_{\mathcal{D}} |\phi v_d| dx + \frac{C}{\sqrt{Pr}} (\|\mathbf{v}\|^2 \|\nabla \mathbf{u}\|^2 + (\|\phi\|^2 + \|\mathbf{v}\|^2) \|\nabla \theta\|^2)$$

$$\leq \frac{Pr}{4} \|\mathbf{v}\|^2 + \frac{C(\tilde{R}a + PrRa)^2}{Pr} \|\phi\|^2 + \frac{C}{\sqrt{Pr}} (\|\mathbf{v}\|^2 \|\nabla \mathbf{u}\|^2 + (\|\phi\|^2 + \|\mathbf{v}\|^2) \|\nabla \theta\|^2).$$

That is, with $\lambda_2 \geq 2C(\tilde{R}a + PrRa)^2/Pr$, taking $\lambda = (\frac{1}{2} Pr) \wedge \lambda_2$, we have

$$\frac{d}{dt} (\|\mathbf{v}\|^2 + \|\phi\|^2) + \lambda (\|\mathbf{v}\|^2 + \|\phi\|^2) \leq \frac{C}{\sqrt{Pr}} (\|\phi\|^2 + \|\mathbf{v}\|^2) (\|\nabla \mathbf{u}\|^2 + \|\nabla \theta\|^2), \tag{5.21}$$

as in (5.16) above. The rest of the proof follows the proof of part (i) above. Note that, from (3.38) (specifically from the dependence of C_2 on Pr), we can choose Pr large enough such that $\lambda = (\frac{1}{2} Pr) \wedge \lambda_2 > \frac{CC_2}{Pr^{3/2}}$, as required in the proof.

$$\square$$

5.3   Proof of Theorem 1.3

Proof of Theorem 1.3. We write $H = H_2$ as the initial data phase space for (2.6)–(2.7) during this proof. We assume that the spatial dimension is $d = 3$, as the proof when $d = 2$ is nearly identical. Let P_∞ be the transition function on H_∞ corresponding to equations (2.6)–(2.7) evaluated at integer times. For given $U^0 = (\mathbf{u}^0, \theta^0)$ satisfying (2.6)–(2.7) and initial condition $\tilde{\theta}_0 \in H$, define $\tilde{U} = (\tilde{\mathbf{u}}, \tilde{\theta})$ as the solution of

$$-\Delta\tilde{\mathbf{u}} + \nabla\tilde{p} = Ra\hat{\mathbf{e}}_d\tilde{\theta}, \quad \nabla \cdot \tilde{\mathbf{u}} = 0, \tag{5.22}$$

$$d\tilde{\theta} + \tilde{\mathbf{u}} \cdot \nabla\tilde{\theta}dt = \tilde{Ra}\tilde{u}_d dt + \Delta\tilde{\theta}dt + \sum_{k=1}^{N_2} \sigma_k dW^k - \lambda_2 P_{N_2}(\tilde{\theta} - \theta^0)\mathbb{1}_{t<\tilde{\tau}_R}dt \tag{5.23}$$

where

$$\tilde{\tau}_R = \inf\left\{t > 0 : \int_0^t \|\sigma^{-1}\lambda_2 P_{N_2}(\tilde{\theta} - \theta^0)\|^2 ds \geq R\right\}. \tag{5.24}$$

As in the proof of Theorem 1.1, let $\tilde{P}_\infty$ be the transition function on H_∞ corresponding to equations (5.22) and (5.23). Again, by the Girsanov theorem, for any initial data $\tilde{\theta}_0 \in H$, one has $\Gamma_{\theta_0^0,\tilde{\theta}_0} \in \tilde{C}_\infty(\delta_{\theta_0^0}, \delta_{\tilde{\theta}_0})$ with $\Gamma_{\theta_0^0,\tilde{\theta}_0} = \delta_{\theta_0^0}P_\infty \times \delta_{\tilde{\theta}_0}\tilde{P}_\infty$. Following the discussion in the proof of Theorem 1.1 and using Theorem 5.1, it remains to prove that $\Gamma_{\theta_0^0,\tilde{\theta}_0}(D) > 0$.

For $V = \tilde{U} - U^0 = (\mathbf{v}, \phi)$, we obtain

$$-\Delta\mathbf{v} + \nabla p = Ra\hat{\mathbf{e}}_d\phi, \tag{5.25}$$

$$\partial_t\phi + \tilde{\mathbf{u}} \cdot \nabla\phi + \mathbf{v} \cdot \nabla\theta^0 = v_d + \Delta\phi - \lambda_2 P_{N_2}\phi\mathbb{1}_{t<\tilde{\tau}_R}. \tag{5.26}$$

Energy estimates yield

$$\|\nabla\mathbf{v}\|^2 \leq Ra^2\|\phi\|^2,$$

$$\frac{1}{2}\frac{d}{dt}\|\phi\|^2 + \|\nabla\phi\|^2 + \lambda_2\|P_{N_2}\phi\|^2\mathbb{1}_{t<\tilde{\tau}_R} \leq \int_{\mathcal{D}}(\tilde{Ra}|v_d\phi| + |\phi\mathbf{v} \cdot \nabla\theta^0|)dx, \tag{5.27}$$

and we find

$$\int_{\mathcal{D}} |\phi\mathbf{v} \cdot \nabla\theta^0|dx \leq C\|\mathbf{v}\|_{L^6}\|\nabla\theta^0\|\|\phi\|_{L^3} \leq C\|\nabla\mathbf{v}\|\|\nabla\theta^0\|\|\phi\|^{1/2}\|\nabla\phi\|^{1/2}$$

$$\leq CRa\|\phi\|^{3/2}\|\nabla\theta^0\|\|\nabla\phi\|^{1/2} \leq \frac{1}{2}\|\nabla\phi\|^2 + CRa^{4/3}\|\phi\|^2\|\nabla\theta^0\|^{4/3}$$

$$\leq \frac{1}{2}\|\nabla\phi\|^2 + CRa^4\|\phi\|^2 + \frac{1}{2}\|\nabla\theta^0\|^2\|\phi\|^2.$$

Combining with (5.27) we obtain

$$\frac{d}{dt}\|\phi\|^2 + \|\nabla\phi\|^2 + 2\lambda_2\|P_{N_2}\phi\|^2\mathbb{1}_{t<\tilde{\tau}_R} \leq (CRa^4 + 2Ra\tilde{Ra})\|\phi\|^2 + \|\nabla\theta^0\|^2\|\phi\|^2. \tag{5.28}$$

Let

$$\tilde{A}_K = \left\{\sup_{t\geq 0}\left(\|\theta^0(t)\|^2 + \int_0^t \|\nabla\theta^0(s)\|^2 ds - \tilde{C}_1 - \tilde{C}_2 t\right) \leq \tilde{C}_3 K\right\},$$

where $\tilde{C}_1, \tilde{C}_2, \tilde{C}_3$ are the constants determined in Proposition 3.3. With an appropriate choice of R, we will show that $\tilde{\tau}_R = \infty$ on $\tilde{A}_K$. Indeed, assume $\tilde{\tau}_R$ is finite for some element of $\tilde{A}_K$. Then for $t < \tilde{\tau}_R$, by choosing $N_2 \geq \sqrt{\lambda_2/C}$, where C is such that $2\lambda_2\|\phi\|^2 \leq 2\lambda_2\|P_{N_2}\phi\|^2 + \frac{1}{2}\|\nabla\phi\|^2$ (cf. (5.14)), we obtain from (5.28) that

$$\frac{d}{dt}\|\phi\|^2 + \frac{1}{2}\|\nabla\phi\|^2 + 2\lambda_2\|\phi\|^2 \leq (CRa^4 + 2Ra\tilde{Ra})\|\phi\|^2 + \|\nabla\theta\|^2\|\phi\|^2. \tag{5.29}$$

Now by selecting $\lambda_2 > (CRa^4 + 2Ra\tilde{Ra}) \vee \tilde{C}_2$ we find

$$\frac{d}{dt}\|\phi\|^2 + \lambda_2\|\phi\|^2 \leq \|\nabla\theta\|^2\|\phi\|^2, \tag{5.30}$$

and by Grönwall's inequality,

$$\|\phi(t)\|^2 \leq \|\phi_0\|^2 \exp\left(\int_0^t \|\nabla\theta(s)\|^2 ds - \lambda_2 t\right).$$

Therefore, on $\tilde{A}_K$,

$$\|\phi(t)\|^2 \leq \|\phi_0\|^2 \exp\left(\tilde{C}_3 K + \tilde{C}_1\right)\exp\left((\tilde{C}_2 - \lambda_2)t\right). \tag{5.31}$$

It follows that, for all $t < \tilde{\tau}_R$,

$$\int_0^t \|\sigma^{-1}\lambda_2 P_{N_2}\phi\|^2 ds \leq \lambda_2\tilde{C}\|\phi_0\|^2 \exp\left(\tilde{C}_3 K + \tilde{C}_1\right)\int_0^t \exp\left((\tilde{C}_2 - \lambda_2)s\right) ds$$

$$\leq \tilde{C}_0\|\phi_0\|^2 \exp\left(\tilde{C}_3 K + \tilde{C}_1\right),$$

where $\tilde{C}_0 = \tilde{C}_0(\tilde{C}_2)$. Choose $R > 0$ such that $R^2 > 2\tilde{C}_0\|\phi_0\|^2 \exp\left(\tilde{C}_3 K + C_1\right)$. Then, since $\lambda_2 \geq \tilde{C}_2$ for all $0 < t < \tilde{\tau}_R$, we have

$$\int_0^t \|\sigma^{-1}\lambda_2 P_{N_2}\phi\|^2 ds < \frac{R}{2}, \tag{5.32}$$

which contradicts the assumption that $\tilde{\tau}_R$ is finite. Therefore, for these choices of K, R and λ_2, we have that on $\tilde{A}_K$, the estimate (5.31) holds for all $t > 0$, and the remainder of the proof that $\Gamma_{\theta_0^0, \tilde{\theta}_0}(D) > 0$ follows as in (5.19)–(5.20) above. $\qquad\square$

6 Bounds on the Nusselt Number

In this section we include the proof of Theorem 1.4, which states that the Nusselt number Nu relative to the unique invariant measure of (1.1)–(1.2) is observable, and provides quantitative bounds in terms of Ra, $\tilde{Ra}$. Throughout this section, we assume $d = 2$ and for a function $f = f(x_1, x_2, t)$ we use the notation

$$\langle f \rangle = \langle f \rangle(x_1, x_2) = \lim_{t \to \infty} \frac{1}{t}\int_0^t f(t, x_1, x_2)dt \tag{6.1}$$

to denote the infinite temporal average. We will also use the symbol

$$\overline{f} = \overline{f}(t, x_2) = \frac{1}{|\mathcal{D}|}\int_0^L f(t, x_1, x_2)dx_1$$

to refer to the horizontal average of f, and $\fint_{\mathcal{D}} dx$ for averages over $\mathcal{D}$.

Proof of Theorem 1.4. We assume $d = 2$, $N_1 = 0$, $N_2 = \infty$, and Pr is sufficiently large such that, by Theorem 1.1, the system (1.1)–(1.2) possesses a unique ergodic invariant measure μ. Recall that the Nusselt number Nu, relative to μ, is given by (1.6). This notation is imprecise, and we should now interpret (1.6) according to (2.2) as

$$Nu = 1 + \int_H u_2(\theta + \tilde{Ra}(1 - x_2))d\mu(\mathbf{u}, \theta). \tag{6.2}$$

Next consider $f : H \to \mathbb{R}$ given by

$$f(\tilde{\mathbf{u}}, \tilde{\theta}) = \int_{\mathcal{D}} \tilde{u}_2(\tilde{\theta} + \tilde{R}a(1 - x_2))dx.$$

Notice that from (4.1) we have $f \in L^1(H; \mu)$. Indeed,

$$\int_H \left(\int_{\mathcal{D}} \tilde{u}_2(\tilde{\theta} + \tilde{R}a(1 - x_2))dx \right) d\mu(\tilde{\mathbf{u}}, \tilde{\theta}) \le \frac{1}{2} \int_H \left(\|\tilde{\mathbf{u}}\|^2 + \|\tilde{\theta}\|^2 + \tilde{C} \right) d\mu(\tilde{\mathbf{u}}, \tilde{\theta}) < \infty,$$

where $\tilde{C} = \tilde{C}(\tilde{R}a)$. By the Birkhoff ergodic theorem it follows that, for μ almost every initial condition $(\mathbf{u}_0, \theta_0) \in H$, recalling (1.6),

$$Nu = 1 + \frac{1}{\tilde{R}a} \int_H \fint_{\mathcal{D}} \tilde{u}_2(\tilde{\theta} + \tilde{R}a(1 - x_2)) \, dx \, d\mu(\tilde{\mathbf{u}}, \tilde{T}) = 1 + \frac{1}{\tilde{R}a} \left\langle \mathbb{E} \left(\fint_{\mathcal{D}} u_2(\theta + \tilde{R}a(1 - x_2))dx \right) \right\rangle$$

$$= 1 + \frac{1}{\tilde{R}a} \left\langle \mathbb{E} \left(\fint_{\mathcal{D}} u_2 T dx \right) \right\rangle \tag{6.3}$$

where $(\mathbf{u}, T)$ is the solution to (1.1)–(1.2) with initial data $(\mathbf{u}_0, T_0)$ (note that θ_0 and T_0 are related by (2.2)). It remains to establish part (ii) of Theorem 1.4.

We follow the method of [CD96]. Fix initial conditions $(\mathbf{u}_0, T_0) \in \operatorname{supp}(\mu)$ such that (6.3) is satisfied. Notice that from (1.2) and integration by parts, recalling that $\|\sigma\| = 1$, we have

$$d\|T\|^2 = \left(2 \int_{\mathcal{D}} \Delta T T dx + 1 \right) dt + 2\langle T, \sigma \rangle dW$$

$$= \left(-2\|\nabla T\|^2 - 2|\mathcal{D}|\tilde{R}a\overline{\partial_2 T}|_{x_2=0} + 1 \right) dt + 2\langle T, \sigma \rangle dW. \tag{6.4}$$

Also, by Proposition 3.2 (specifically (3.31)),

$$\frac{1}{t} \mathbb{E}(\|T(t)\|^2 - \|T_0\|^2) \le \frac{C}{t} \to 0, \tag{6.5}$$

as $t \to \infty$. We then take the expectation of the integrated form of (6.4), and conclude by (6.5) that

$$\frac{1}{|\mathcal{D}|} \langle \mathbb{E}\|\nabla T\|^2 \rangle = -\tilde{R}a \langle \mathbb{E} \left(\overline{\partial_2 T}|_{x_2=0} \right) \rangle + \frac{1}{2|\mathcal{D}|}. \tag{6.6}$$

Next observe that (1.2) also gives

$$d\overline{T} = \left(\overline{\partial_2^2 T} - \overline{\partial_2(u_2 T)} \right) dt + \overline{\sigma} dW,$$

and therefore

$$(1 - x_2)d\overline{T} = (1 - x_2) \left(\overline{\partial_2^2 T} - \overline{\partial_2(u_2 T)} \right) dt + (1 - x_2)\overline{\sigma} dW. \tag{6.7}$$

We now integrate (6.7) in x_2, and integrate by parts, to obtain

$$d \left(\int_0^1 (1 - x_2)\overline{T} dx_2 \right) = \left(\int_0^1 (1 - x_2)\overline{\partial_2^2 T} dx_2 + \fint_{\mathcal{D}} (1 - x_2)\partial_2(u_2 T)dx \right) dt + \fint_{\mathcal{D}} (1 - x_2)\sigma dW$$

$$= \left(-\overline{\partial_2 T}|_{x_2=0} + \int_0^1 \overline{\partial_2 T} dx_2 - \fint_{\mathcal{D}} u_2 T dx \right) dt + \fint_{\mathcal{D}} (1 - x_2)\sigma dW$$

$$= \left(-\overline{\partial_2 T}|_{x_2=0} - \tilde{R}a - \fint_{\mathcal{D}} u_2 T dx \right) dt + \fint_{\mathcal{D}} (1 - x_2)\sigma dW. \tag{6.8}$$

Note that we have used the boundary conditions (1.3) in the last line. By taking the expectation and infinite temporal average, we have by Proposition 3.2, another integration by parts, and (6.3),

$$-\left\langle \mathbb{E}\left(\overline{\partial_2 T}|_{x_2=0}\right)\right\rangle - \tilde{Ra} = \left\langle \mathbb{E}\fint_{\mathcal{D}} u_2 T dx\right\rangle = \tilde{Ra}(Nu-1).$$

Therefore

$$\tilde{Ra}Nu = -\left\langle \mathbb{E}\left(\overline{\partial_2 T}|_{x_2=0}\right)\right\rangle,$$

and by (6.6),

$$\frac{1}{|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla T\|^2\right\rangle = -\tilde{Ra}\left\langle \mathbb{E}\left(\overline{\partial_2 T}|_{x_2=0}\right)\right\rangle + \frac{1}{2|\mathcal{D}|} = \tilde{Ra}^2 Nu + \frac{1}{2|\mathcal{D}|},$$

or equivalently,

$$Nu = \frac{1}{\tilde{Ra}^2|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla T\|^2\right\rangle - \frac{1}{2\tilde{Ra}^2|\mathcal{D}|}. \tag{6.9}$$

Next observe that (1.1) combined with Proposition 3.2 and (6.3) gives

$$\frac{1}{|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla \mathbf{u}\|^2\right\rangle = Ra\left\langle \mathbb{E}\fint_{\mathcal{D}} u_2 T\, dx\right\rangle = Ra\tilde{Ra}(Nu-1),$$

so we also have

$$Nu = \frac{1}{Ra\tilde{Ra}|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla \mathbf{u}\|^2\right\rangle + 1. \tag{6.10}$$

We multiply (6.10) by $-1/2$ and add it to (6.9) to obtain

$$Nu = \frac{2}{\tilde{Ra}^2|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla T\|^2\right\rangle - \frac{1}{Ra\tilde{Ra}|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla \mathbf{u}\|^2\right\rangle - \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1. \tag{6.11}$$

Consider a deterministic background profile $\tau = \tau(x_2)$ satisfying $\tau(0) = \tilde{Ra}$, $\tau(1) = 0$ and consider $T(t,x) = \tau(x_2) + \theta(t,x)$. Then the fluctuations θ satisfy

$$d\theta + (\mathbf{u}\cdot\nabla\theta - \Delta\theta - \tau'' + u_2\tau')dt = \sigma dW,$$

with Dirichlet boundary conditions at $x_2 = 0, 1$. This gives

$$d\|\theta\|^2 = \left(2\int_{\mathcal{D}}(\Delta\theta\theta + \tau''\theta - u_2\tau'\theta)dx + 1\right)dt + 2\langle T, \sigma\rangle dW$$

$$= \left(-2\|\nabla\theta\|^2 - 2\int_{\mathcal{D}}(\tau'\partial_2\theta + u_2\tau'\theta)dx + 1\right)dt + 2\langle T, \sigma\rangle dW.$$

By taking the expectation and infinite temporal average, we have again by Proposition 3.2,

$$-\frac{1}{|\mathcal{D}|}\left\langle \mathbb{E}\|\nabla\theta\|^2\right\rangle - \left\langle \mathbb{E}\fint_{\mathcal{D}}(\tau'\partial_2\theta + u_2\tau'\theta)dx\right\rangle + \frac{1}{2|\mathcal{D}|} = 0. \tag{6.12}$$

Next we expand

$$\|\nabla T\|^2 = \|\nabla\theta\|^2 + 2\int_{\mathcal{D}}\tau'\partial_2\theta dx + |\mathcal{D}|\int_0^1 (\tau')^2 dx_2. \tag{6.13}$$

We take the expectation and temporal average of (6.13), and add (6.12) twice to obtain

$$\frac{1}{|\mathcal{D}|}\left\langle\mathbb{E}\|\nabla T\|^2\right\rangle = -\frac{1}{|\mathcal{D}|}\left\langle\mathbb{E}\|\nabla\theta\|^2\right\rangle + \int_0^1 (\tau')^2 dx_2 - 2\left\langle\mathbb{E}\fint_\mathcal{D} u_2\tau'\theta dx\right\rangle + \frac{1}{|\mathcal{D}|}. \tag{6.14}$$

Inserting (6.14) into (6.11) we find

$$Nu = \frac{2}{\tilde{Ra}^2}\int_0^1 (\tau')^2 dx_2 + \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1 - \frac{2}{\tilde{Ra}^2|\mathcal{D}|}Q_\tau(\theta,\mathbf{u}), \tag{6.15}$$

where

$$Q_\tau(\theta,\mathbf{u}) = \left\langle\mathbb{E}\left(\|\nabla\theta\|^2 + \frac{\tilde{Ra}}{2Ra}\|\nabla\mathbf{u}\|^2 + 2\int_\mathcal{D} u_2\tau'\theta dx\right)\right\rangle. \tag{6.16}$$

Consider a background temperature given by

$$\tau = \tau(x_2) = \tilde{Ra} - \frac{\tilde{Ra}}{\delta}\int_0^{x_2}\psi(x/\delta)dx,$$

where $\psi = \psi(x)$ is a smooth function on $\mathbb{R}$ such that $0 \le \psi \le C$, $\text{supp}(\psi) \subset [0,1]$ and $\int_\mathbb{R}\psi(x)dx = 1$. It follows that, for $\delta < 1$, $\tau(0) = \tilde{Ra}$ and $\tau(1) = 0$. Next, observe that for any $x_2 \in [0,1]$, we can exploit the Dirichlet boundary condition at $x_2 = 0$ and write

$$|\overline{u_2\theta}(x_2)| = \left|\int_0^{x_2}\overline{u_2\theta}'(z)dz\right| = \frac{1}{|\mathcal{D}|}\left|\int_0^L\int_0^{x_2}(\theta\partial_2 u_2 + u_2\partial_2\theta)(x_1,z)dzdx_1\right|$$

$$= \frac{1}{|\mathcal{D}|}\left|\int_0^L\int_0^{x_2}\left[\left(\int_0^z \partial_2\theta(x_1,y)dy\right)\partial_2 u_2(x_1,z) + \left(\int_0^z \partial_2 u_2(x_1,y)dy\right)\partial_2\theta(x_1,z)\right]dzdx_1\right|$$

$$\le |x_2|\|\nabla u\|\|\nabla\theta\|,$$

where we have applied Cauchy-Schwarz in each variable to obtain the last line. It follows that

$$\left|\int_\mathcal{D} u_2\tau'\theta dx\right| = \left|\int_0^1\overline{u_2\theta}(x_2)\tau'(x_2)dx_2\right| \le \tilde{Ra}\|\nabla u\|\|\nabla\theta\|\int_0^1 |x_2/\delta|\psi(x_2/\delta)dx_2$$

$$= \tilde{Ra}\|\nabla u\|\|\nabla\theta\|\delta\int_0^1 |z|\psi(z)dz \le \sqrt{\tilde{Ra}/2Ra}\|\nabla u\|\|\nabla\theta\|,$$

by taking $\delta \sim 1/\sqrt{2\tilde{Ra}Ra}$. With this choice for τ, by (6.16) we have

$$Q_\tau(\theta,\mathbf{u}) \ge \left\langle\mathbb{E}\left[\left(\|\nabla\theta\| - \sqrt{\tilde{Ra}/2Ra}\|\nabla\mathbf{u}\|\right)^2\right]\right\rangle \ge 0,$$

for all $(\theta,\mathbf{u}) \in H$, and therefore

$$Nu \le \frac{2}{\tilde{Ra}^2}\int_0^1 (\tau')^2 dx_2 + \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1 = \frac{2}{\delta^2}\int_0^1 (\psi(x_2/\delta))^2 dx_2 + \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1$$

$$= \frac{2}{\delta}\int_0^1 (\psi(z))^2 dz + \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1$$

$$\le C\sqrt{\tilde{Ra}Ra} + \frac{1}{\tilde{Ra}^2|\mathcal{D}|} - 1. \tag{6.17}$$

$$\square$$

Remark 6.1. *Notice that, due to the stochastic forcing, our bound on the Nusselt number in* (6.17) *includes an "Itô correction" term of the form* $\frac{1}{\tilde{R}a^2|\mathcal{D}|}$. *Further note that, from* (2.1), *by keeping all other parameters fixed and taking the stochastic heat flux γ large, $Ra\tilde{R}a$ is constant and $\tilde{R}a$ goes to zero, so that this extra term dominates the right-hand side of* (6.17), *and we have weaker estimates on the Nusselt number Nu. We do not claim that this reflects an enhanced convective heat transport for* (1.1)–(1.2) *as γ increases, but rather that our mathematical methods produce an extra term in this case, which may or may not be misleading.*

A   Appendix: Non-dimensionalization of the Equations

In this appendix we will describe how to obtain the non-dimensional system (1.1)–(1.2) from the standard form of these equations through a rescaling. We also derive the system considered in [FGHR]. The Boussinesq equations in their original form can be written (with stochastic forcing included) as

$$d\mathbf{v}^* + (\mathbf{v}^* \cdot \nabla \mathbf{v}^* + \nabla p^*)\, dt^* = (g\alpha\hat{\mathbf{e}}_d T^* + \nu\Delta\mathbf{v}^*)dt^* + \tilde{\gamma}\sum_{k=1}^{N_1}\tilde{\sigma}_k^* d\tilde{W}_*^k, \quad \nabla\cdot\mathbf{v}^* = 0, \tag{A.1}$$

$$dT^* + (\mathbf{v}^*\cdot\nabla T^*)\, dt^* = \kappa\Delta T^* dt^* + \gamma\sum_{k=1}^{N_2}\sigma_k^* dW_*^k. \tag{A.2}$$

We denote the spatial variable by $\mathbf{x}^* = (x_1^*,\ldots,x_d^*)$, where $d = 2$ or $d = 3$, assume that $\mathbf{v}^* = (v_1,\ldots,v_d)$ and T^* are periodic in $(x_1^*,\ldots,x_{d-1}^*)$, and that

$$\mathbf{v}^*|_{x_d^*=0} = \mathbf{v}^*|_{x_d^*=h} = 0, \quad T^*|_{x_d^*=0} = T_1, \quad T^*|_{x_d^*=h} = 0,$$

where $T_1 > 0$. In these equations, g represents the gravitational acceleration, ν the kinematic viscosity, κ the thermal conductivity, α the thermal expansion coefficient, and $\tilde{\gamma},\gamma$ the volumetric flux coefficients, which will be specified in more detail depending on context below. Here $\tilde{\sigma}_k^*$ and σ_k^* denote bases of eigenfunctions of the (time independent) Stokes and Laplace operators, respectively, with appropriate boundary conditions, and $\tilde{W}_*^k, W_*^k$ are sequences of one-dimensional, mutually independent, standard Brownian motions.

We abuse notation for computational purposes, and rewrite (A.1)-(A.2) using standard derivative notation

$$\frac{\partial\mathbf{v}^*}{\partial t^*} + \mathbf{v}^*\cdot\nabla\mathbf{v}^* + \nabla p^* = g\alpha\hat{\mathbf{e}}_d T^* + \nu\Delta\mathbf{v}^* + \tilde{\gamma}\sum_{k=1}^{N_1}\tilde{\sigma}_k^*\frac{\partial\tilde{W}_*^k}{\partial t^*}, \tag{A.3}$$

$$\frac{\partial T^*}{\partial t^*} + \mathbf{v}^*\cdot\nabla T^* = \kappa\Delta T^* + \gamma\sum_{k}\sigma_k^*\frac{\partial W_*^k}{\partial t^*}. \tag{A.4}$$

To obtain the non-dimensional version of (A.3)-(A.4), we change variables as follows

$$\mathbf{x}^* = h\mathbf{x},$$
$$t^* = \beta t,$$
$$\mathbf{v}^*(\mathbf{x}^*, t^*) = \lambda\mathbf{u}(\mathbf{x}, t),$$
$$T^*(\mathbf{x}^*, t^*) = \tilde{T}T(\mathbf{x}, t),$$
$$p^*(\mathbf{x}^*, t^*) = \frac{\nu\lambda}{h}p(\mathbf{x}, t),$$

where $\beta = h^2/\kappa$ and $\lambda = \kappa/h$. We compute

$$\frac{\lambda}{\beta}\frac{\partial\mathbf{u}}{\partial t} + \frac{\lambda^2}{h}\mathbf{u}\cdot\nabla\mathbf{u} + \frac{\nu\lambda}{h^2}\nabla p = g\alpha\tilde{T}\hat{\mathbf{e}}_d T + \frac{\nu\lambda}{h^2}\Delta\mathbf{u} + \frac{\tilde{\gamma}}{h^{d/2}\sqrt{\beta}}\sum_{k=1}^{N_1}\tilde{\sigma}_k\frac{\partial\tilde{W}^k}{\partial t},$$

$$\frac{\tilde{T}}{\beta}\frac{\partial T}{\partial t} + \frac{\lambda\tilde{T}}{h}\mathbf{u}\cdot\nabla T = \frac{\tilde{T}\kappa}{h^2}\Delta T + \frac{\gamma}{h^{d/2}\sqrt{\beta}}\sum_{k=1}^{N_2}\sigma_k\frac{\partial W^k}{\partial t},$$

where we have used that $\frac{1}{\beta^{1/2}}W_*^k(\beta t) =: W^k(t)$ is a standard Brownian motion, and $\|\sigma_k^*\| = h^{-d/2}\|\sigma_k\|$.

Notice that $\frac{\tilde{T}}{\beta} = \frac{\lambda\tilde{T}}{h} = \frac{\tilde{T}\kappa}{h^2}$ and $\frac{\lambda}{\beta} = \frac{\lambda^2}{h}$, and if we recall $Pr = \nu/\kappa$, we obtain

$$\frac{1}{Pr}\left(\frac{\partial\mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) + \nabla p = \frac{g\alpha\tilde{T}h^3}{\kappa\nu}\hat{\mathbf{e}}_d T + \Delta\mathbf{u} + \frac{\tilde{\gamma}}{\nu\sqrt{\kappa}h^{\frac{d}{2}-2}}\sum_{k=1}^{N_1}\tilde{\sigma}_k\frac{\partial\tilde{W}^k}{\partial t}, \tag{A.5}$$

$$\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T = \Delta T + \frac{\gamma}{\tilde{T}\sqrt{\kappa}h^{d/2-1}}\sum_{k=1}^{N_2}\sigma_k\frac{\partial W^k}{\partial t}. \tag{A.6}$$

with the boundary conditions

$$\mathbf{u}|_{x_d=0} = \mathbf{u}|_{x_d=h} = 0, \quad T|_{x_d=0} = \frac{T_1}{\tilde{T}}, \quad T|_{x_d=h} = 0.$$

The non-dimensionalization proceeds as follows: we select $\beta = h^2/\kappa$ as the unit of time, h the unit of space, and further select $\tilde{T} = \frac{\gamma\sqrt{\beta}}{h^{d/2}} = \frac{\gamma}{\sqrt{\kappa}h^{d/2-1}}$ as the unit of temperature. We also absorb the constant $\frac{\tilde{\gamma}}{\nu\sqrt{\kappa}h^{\frac{d}{2}-2}}$ into the definition of the basis $\{\tilde{\sigma}_k\}$, leading to the system (1.1)–(1.2). The parameters in the problem are then the unit-less "Rayleigh numbers" $Ra := \frac{g\alpha\tilde{T}h^3}{\kappa\nu} = \frac{g\alpha\gamma h^{4-d/2}}{\nu\kappa^{3/2}}$ and $\tilde{Ra} := \frac{T_1}{\tilde{T}} = \frac{\sqrt{\kappa}h^{d/2-1}T_1}{\gamma}$. Here we have defined the units of γ and $\tilde{\gamma}$ as required to maintain consistency and reach a non-dimensional form for (1.1)–(1.2).

For example, in the work [FGHR] we consider $d = 3$, $\tilde{\gamma} = 0$, and interpret $\gamma = \frac{H}{\rho c}$ as the volumetric "stochastic" heat flux H (units of power/$\sqrt{\text{volume*time}}$) normalized by the density ρ and the specific heat c of the fluid. Then (A.5)–(A.6) becomes

$$\frac{1}{Pr}\left(\frac{\partial\mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) + \nabla p = Ra\hat{\mathbf{e}}_d T + \Delta\mathbf{u}, \tag{A.7}$$

$$\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T = \Delta T + \sum_{k=1}^{N_2}\sigma_k\frac{\partial W^k}{\partial t}, \tag{A.8}$$

with the boundary conditions

$$\mathbf{u}|_{x_d=0} = \mathbf{u}|_{x_d=1} = 0, \quad T|_{x_d=0} = \tilde{Ra}, \quad T|_{x_d=1} = 0.$$

Acknowledgments

References

[AGL09]   G. Ahlers, S. Grossmann, and D. Lohse. Heat transfer and large scale dynamics in turbulent rayleigh-bénard convection. *Reviews of Modern Physics*, 81(2):503, 2009.

[Bó1]   M. Bénard. Les tourbillons cellulaires dans une nappe liquide propageant de la chaleur par convection en régime permanent. *Ann. d. Chimie et de Physique*, xxiii:62–144, 1901.

[Ben95]   A. Bensoussan. Stochastic Navier-Stokes equations. *Acta Appl. Math.*, 38(3):267–304, 1995.

[BH09]   M. Breuer and U. Hansen. Turbulent convection in the zero reynolds number limit. *Euro. phys. Lett.*, 86(24004), 2009.

[BN12]   A. Barletta and D. A. Nield. On the rayleigh–bénard–poiseuille problem with internal heat generation. *International Journal of Thermal Sciences*, 57:1–16, 2012.

[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. In *Annual review of fluid mechanics, Vol. 32*, volume 32 of *Annu. Rev. Fluid Mech.*, pages 709–778. Annual Reviews, Palo Alto, CA, 2000.

[Bus70]   F.H. Busse. Bounds for turbulent shear flow. *Journal of Fluid Mechanics*, 41(01):219–240, 1970.

[Bus78]   F. H. Busse. The optimum theory of turbulence. *Advances in applied mechanics*, 18:77–121, 1978.

[CD96]   P. Constantin and C. R. Doering. Heat transfer in convective turbulence. *Nonlinearity*, 9(4):1049–1060, 1996.

[CD99]   P. Constantin and C. R. Doering. Infinite Prandtl number convection. *J. Statist. Phys.*, 94(1-2):159–172, 1999.

[CF88]   P. Constantin and C. Foias. *Navier-Stokes equations*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1988.

[DC01]   C. R. Doering and P. Constantin. On upper bounds for infinite Prandtl number convection with or without rotation. *J. Math. Phys.*, 42(2):784–795, 2001.

[DGHT11]   A. Debussche, N. Glatt-Holtz, and R. Temam. Local martingale and pathwise solutions for an abstract fluids model. *Physica D*, 240(14-15):1123–1144, 2011.

[DPZ92]   G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1992.

[Eyi96]   G.L. Eyink. Exact results on stationary turbulence in 2d: consequences of vorticity conservation. *Physica D: Nonlinear Phenomena*, 91(1):97–142, 1996.

[FG95]   F. Flandoli and D. Gatarek. Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, 102(3):367–391, 1995.

[FGHR]   J. Földes, N. Glatt-Holtz, and G. Richards. Large prandtl number asymptotics in randomly forced turbulent convection. submitted.

[FGHRT15]   J. Földes, N. Glatt-Holtz, G. Richards, and E. Thomann. Ergodic and mixing properties of the boussinesq equations with a degenerate random forcing. *Journal of Functional Analysis*, 269(8):2427 – 2504, 2015.

[FMRT01]   C. Foias, O. Manley, R. Rosa, and R. Temam. *Navier-Stokes equations and turbulence*, volume 83 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2001.

[GHMR]   N. Glatt-Holtz, J. Mattingly, and G. Richards. A simple framework for the analysis of ergodicity in infinite dimensional stochastic systems. (to appear).

[GHŠV15]   N. Glatt-Holtz, V. Šverák, and V. Vicol. On inviscid limits for the stochastic navierstokes equations and related models. *Archive for Rational Mechanics and Analysis*, 217(2):619–649, 2015.

[GHZ09]   N. Glatt-Holtz and M. Ziane. Strong pathwise solutions of the stochastic Navier-Stokes system. *Advances in Differential Equations*, 14(5-6):567–600, 2009.

[GK96]   I. Gyöngy and N. Krylov. Existence of strong solutions for itô's stochastic equations via approximations. *Probability theory and related fields*, 105(2):143–158, 1996.

[GS12]   D. Goluskin and E. A. Spiegel. Convection driven by internal heating. *Physics Letters A*, 377(1):83–92, 2012.

[HM06]   M. Hairer and J. C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)*, 164(3):993–1032, 2006.

[HM08]   M. Hairer and J. C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[HM11]   M. Hairer and J. C. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic pdes. *Electron. J. Probab.*, 16(23):658–738, 2011.

[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(1-2):223–259, 2011.

[Hop40] E. Hopf. Ein allgemeiner endlichkeitssatz der hydrodynamik. *Mathematische Annalen*, 117(1):764–775, 1940.

[How72] L. N. Howard. Bounds on flow quantities. *Annual Review of Fluid Mechanics*, 4(1):473–494, 1972.

[HS12]   M. Hofmanová and J. Seidler. On weak solutions of stochastic differential equations. *Stochastic Analysis and Applications*, 30(1):100–121, 2012.

[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(1):65–113, 1937.

[Kry10] N. V. Krylov. Itô's formula for the L_p-norm of stochastic W_p^1-valued processes. *Probab. Theory Related Fields*, 147(3-4):583–605, 2010.

[KS12]   S. Kuksin and A. Shirikyan. *Mathematics of Two-Dimensional Turbulence*. Number 194 in Cambridge Tracts in Mathematics. Cambridge University Press, 2012.

[LDB04] L. Lu, C. R. Doering, and F. H. Busse. Bounds on convection driven by internal heating. *Journal of Mathematical Physics*, 45(7):2967–2986, 2004.

[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(Issue 192):529–546, 1916.

[LX10]   D. Lohse and K. Q. Xia. Small-scale properties of turbulent rayleigh-bénard convection. *Annual Review of Fluid Mechanics*, 42:335–364, 2010.

[Mal54] W. V. R. Malkus. The heat transport and spectrum of thermal turbulence. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 225(1161):196–212, 1954.

[Man06] P. Manneville. Rayleigh-bénard convection: Thirty years of experimental, theoretical, and modeling work. In *Dynamics of Spatio-Temporal Cellular Structures*, pages 41–65. Springer, 2006.

[Nov65] E. A. Novikov. Functionals and the random-force method in turbulence theory. *Soviet Physics JETP*, 20:1290–1294, 1965.

[OS11]   F. Otto and C. Seis. Rayleigh–bénard convection: improved bounds on the nusselt number. *Journal of mathematical physics*, 52(8):083702, 2011.

[Par06] J. Park. Dynamic bifurcation theory of rayleigh-benard convection with infinite prandtl number. *Discrete Contin. Dyn. Syst.*, 6(3):591, 2006.

[Rob67] P. H. Roberts. Convection in horizontal layers with internal heat generation. theory. *Journal of Fluid Mechanics*, 30(01):33–49, 1967.

[RY99]   D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[Sta88] M. M. Stanišić. *The mathematical theory of turbulence*. Universitext. Springer-Verlag, New York, second edition, 1988.

[Tem01] R. Temam. *Navier-Stokes equations: Theory and numerical analysis*. AMS Chelsea Publishing, Providence, RI, 2001. Reprint of the 1984 edition.

[TZ67]   D. J. Tritton and M. N. Zarraga. Convection in horizontal layers with internal heat generation. experiments. *Journal of Fluid Mechanics*, 30(01):21–31, 1967.

[VCK12] D. Venturi, M. Choi, and G.E. Karniadakis. Supercritical quasi-conduction states in stochastic rayleighbénard convection. *Int. J. Heat Mass Transfer*, 55(13–14):3732–3743, 2012.

[VKF79] M. I. Vishik, A. I. Komech, and A. V. Fursikov. Some mathematical problems of statistical hydromechanics. *Uspekhi Mat. Nauk*, 34(5(209)):135–210, 256, 1979.

[VWK10] D. Venturi, X. Wan, and G.E. Karniadakis. Stochastic bifurcation analysis of rayleighbénard convection. *J. Fluid Mech.*, 650:391–413, 2010.

[Wan04] X. Wang. Infinite Prandtl number limit of Rayleigh-Bénard convection. *Comm. Pure Appl. Math.*, 57(10):1265–1282, 2004.

[Wan08] X. Wang. Stationary statistical properties of Rayleigh-Bénard convection at large Prandtl number. *Comm. Pure Appl. Math.*, 61(6):789–815, 2008.

[WD11]   J. P. Whitehead and C. R. Doering. Internal heating driven convection at infinite prandtl number. *Journal of Mathematical Physics*, 52(9):093101, 2011.