Nonlinear Differ. Equ. Appl. (2019) 26:43

1021-9722/19/060001-43
published online October 10, 2019
https://doi.org/10.1007/s00030-019-0589-z

Nonlinear Differential Equations and Applications NoDEA

Large Prandtl number asymptotics in randomly forced turbulent convection

Juraj Földes, Nathan E. Glatt-Holtz and Geordie Richards

Abstract. We establish the convergence of statistically invariant states for the stochastic Boussinesq equations in the infinite Prandtl number limit and in particular demonstrate the convergence of the Nusselt number (a measure of heat transport in the fluid). This is a singular parameter limit significant in mantle convection and for gasses under high pressure. The equations are subject to a both temperature gradient on the boundary and internal heating in the bulk driven by a stochastic, white in time, gaussian forcing. Here, the stochastic source terms have a strong physical motivation for example as a model of radiogenic heating. Our approach uses mixing properties of the formal limit system to reduce the convergence of invariant states to an analysis of the finite time asymptotics of solutions and parameter-uniform moment bounds. Here, it is notable that there is a phase space mismatch between the finite Prandtl system and the limit equation, and we implement methods to lift both finite and infinite time convergence results to an extended phase space which includes velocity fields. For the infinite Prandtl stochastic Boussinesq equations, we show that the associated invariant measure is unique and that the dual Markovian dynamics are contractive in an appropriate Kantorovich–Wasserstein metric. We then address the convergence of solutions on finite time intervals, which is still a singular perturbation. In the process we derive well-posed equations which accurately approximate the dynamics up to the initial time when the Prandtl number is large.

Mathematics Subject Classification. 76R05, 60H15, 35R60, 37L40.

Keywords. Convective turbulence, Stochastic Boussinesq equations, Large Prandtl number limit, Invariant measures, Ergodicity, Kantorovich–Wasserstein metrics, Singular perturbation analysis.

Contents

1. Introduction

Buoyancy driven convection plays a central role in a wide variety of physical processes: from Earth's climate system to the internal dynamics of stars. As such it is of fundamental importance to identify and predict robust statistical quantities in these complex flows and to connect such statistics with the basic equations governing their dynamics, for example the Boussinesq equations. In particular characterizing pattern formation, mean heat transport, and small scale dynamics as a function of physical parameters and boundary conditions remains a topic of intensive research theoretically, numerically, and experimentally; see e.g. [3,5,32,34] for a broad overview of recent developments.

It has long been understood that statistically invariant states of the nonlinear partial differential equations of fluid dynamics provide mathematical objects which are expected to contain various robust statistical quantities found in turbulent fluid flows. An ongoing challenge is therefore to address the existence, uniqueness, ergodicity, and dependence of these measures on parameters in a variety of specific contexts. While one may certainly pose such questions for deterministic equations [cf. [19]] the stochastic setting can be more tractable given the regularizing effect of noise on the associated probability distribution functions. Moreover, energy may be supplied to the system through both boundary or within the bulk of a fluid, the latter setting for instance models radioactive decay processes in the earth's mantle; see [4,24,33,39,41,51]. Both sources can therefore have an essentially stochastic character in situations of physical interest.

In this and a companion work, [21], we study statistically invariant states of the stochastically driven Boussinesq equations

$$\frac{1}{Pr}(\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u}) - \Delta\mathbf{u} = \nabla p + Ra\hat{\mathbf{k}}T, \quad \nabla\cdot\mathbf{u} = 0, \tag{1.1}$$

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

for the (non-dimensionalized) velocity field $\mathbf{u} = (u_1, u_2, u_3)$, pressure p, and temperature T of a buoyancy driven fluid. The system (1.1)–(1.2) evolves in a three dimensional domain $(x, y, z) = (\mathbf{x}, z) \in \mathcal{D} = [0, L]^2 \times [0, 1]$ and is supplemented with the boundary conditions

$$\begin{aligned} &\mathbf{u}_{|z=0} = \mathbf{u}_{|z=1} = 0, \quad T_{|z=0} = \tilde{Ra},\ T_{|z=1} = 0, \\ &\mathbf{u}, T \text{ are periodic in } \mathbf{x} = (x, y). \end{aligned} \tag{1.3}$$

The unitless physical parameters in the problem are the *Prandtl* number Pr and *Rayleigh* numbers Ra and $\tilde{Ra}$; see Remark 2.3 and [21] for further details concerning this choice of nondimensionalization. The unit vector $\hat{\mathbf{k}} = (0, 0, 1)$ points in the direction of the gravitational force. The driving noise in (1.2) is given by a collection of independent white noise processes $dW^k = dW^k(t)$ acting spatially through the functions $\sigma_k = \sigma_k(x, y, z)$ which form a complete orthogonal basis of eigenfunctions (ordered with respect to eigenvalues) of the Laplace operator on $\mathcal{D}$ supplemented with homogeneous Dirichlet boundary conditions for $z = 0, 1$ and periodic in $\mathbf{x} = (x, y)$. The stochastic terms in (1.2) have been normalized so that

$$\sum_{k=1}^{N} \|\sigma_k\|^2_{L^2(\mathcal{D})} = 1\,, \tag{1.4}$$

with the strength of the body forcing expressed in terms of the physical parameters Ra and $\tilde{Ra}$; see (2.7) below.

Our principal aim here is to establish convergence properties of statistically invariant states of (1.1)–(1.3) to invariant measures of the active scalar equation

$$-\Delta\mathbf{u} = \nabla p + Ra\hat{\mathbf{k}}T, \quad \nabla \cdot \mathbf{u} = 0, \tag{1.5}$$

$$dT + \mathbf{u} \cdot \nabla T dt = \Delta T dt + \sum_{k=1}^{N} \sigma_k dW^k \tag{1.6}$$

in the *Large Prandtl number limit*, that is, when Pr in (1.1) diverges to ∞. Here (1.5)–(1.6) is complemented with boundary conditions as in (1.3). Note that $\mathbf{u}$ and p are determined by T according to (1.5). We write this functional dependence as $\mathbf{u} = M(T)$ and denote $L(T) = (M(T), T)$.

The analysis of convection in the large Prandtl number limit is of basic interest in a variety of physical contexts, most notably in modeling certain portions of the earth's mantle and for convection in gasses under high pressure, where the Prandtl number can reach the order of 10^{24}, see [8,14,37]. It is worth emphasizing that the system (1.5)–(1.6) has very complex dynamics even without stochastic forcing when the Rayleigh number(s) are sufficiently large; see [3,5,6,8,14,32,37,38,44].

Overview of the main results

Let us now present a heuristic version of our main results; for the precise formulation see Theorem 2.2 below. Recall that for any function F and measure μ, the push-forward of μ under F is given by $F\mu := \mu \circ F^{-1}$.

Theorem 1.1. *Fix any $Ra, \tilde{Ra} > 0$ and consider (1.1)–(1.3) and (1.5)–(1.6) with N independently forced directions in the temperature equation. If $N = N(Ra, \tilde{Ra})$ is sufficiently large, then (1.5)–(1.6) possesses a unique ergodic* *invariant measure μ_∞. Let $\{\mu_{Pr}\}_{Pr\geq 1}$ be any sequence of statistically invariant states associated to (1.1)–(1.3) satisfying a uniform exponential moment bound (see (2.13) below noting that $\epsilon := \frac{1}{\mathrm{Pr}}$).*[1] *Then μ_{Pr} converges to $L_{\#}\mu_\infty$ in a suitable metric. In particular, for any sufficiently regular observable ϕ on the $(\mathbf{u}, T)$ phase space,*

$$\left| \int \phi(\mathbf{u}, T) d\mu_{Pr} - \int \phi(L(T)) d\mu_\infty \right| \leq C(Pr)^{-q}, \tag{1.7}$$

where $C = C(\phi, Ra, \tilde{Ra}), q = q(Ra, \tilde{Ra}) > 0$ are independent of Pr and q is independent of ϕ.

The proof of Theorem 1.1 contains several further results of independent interest. Firstly, we show that the Markovian dynamics of probability laws for the infinite Prandtl system, (1.5)– (1.6) is contractive in a suitable Wasserstein distance; see Theorem 2.2, (2.29) below. Secondly, we demonstrate that the finite time dynamics converge in the limit as $Pr \to \infty$.

Note that our results do not rely on the well-posedness of (1.1)–(1.2) or make any assertions concerning the convergence of (1.1)–(1.2) to the formal limit (1.5)–(1.6) for small times. On the other hand, as notable biproduct of our convergence analysis, we derive a well posed approximation of (1.1)–(1.2) up to the initial time $t = 0$ which is valid for large values of Pr. See Sect. 5.3 and Theorem 5.1 below for further details.

It is also worth emphasizing that our proof of Theorem 1.1 applies essentially verbatum to the two-dimensional version of (1.1)–(1.3), where the horizontal variable $\mathbf{x}$ is one-dimensional. Here *all* the statistically invariant states of the full system satisfy the uniform moment bound (2.13). Furthermore, in collaboration with Whitehead [21], the authors have established that with $N = \infty$ and $Pr = Pr(Ra, \tilde{Ra}) > 0$ sufficiently large, the 2D version of (1.1)–(1.3) possesses a unique ergodic invariant measure μ_{Pr}.

An empirical quantity of particular interest in convection is the *Nusselt number* Nu, a ratio of convective to conductive heat transfer, which is defined in terms of a statistical average (e.g. a time average) of the observable $\phi_{Nu} = \int_{\mathcal{D}} u_2 T\, dx$.[2] However, in the deterministic case, even in the turbulent regime of $Ra \gg 1$, Nu depends on initial condition, both at finite and infinite values of Pr and it is unclear that Nu is continuous at $Pr = \infty$. We show that the addition of a stochastic perturbation avoids these concerns.

Corollary 1.1. *For fixed $Ra, \tilde{Ra} > 0$ and any $Pr = Pr(Ra, \tilde{Ra})$ sufficiently large, the system (1.1)–(1.3) posed in two space dimensions with $N = \infty$ and*

[1] Note that usual fundamental difficulties concerning the well-posedness of 3D Navier–Stokes apply to (1.1)–(1.2) and so, following [17], we consider only weak solutions whose laws do not change in time. The uniform exponential moment condition is analogous to a finite energy criterion for weak solutions of the 3D Navier–Stokes equations. In [21] we have established the existence of such states μ_{Pr}, see Proposition 2.1 below for a precise restatement. In particular we cannot rule out the existence of a collection $\{\mu_{Pr}\}_{Pr\in\mathbb{N}}$ which does not satisfy (2.13). Observe that none of these difficulties arise in the 2D case.

[2] Here u_2 represents the vertical velocity component for the 2D version of (1.1)–(1.3).

(1.4) *possesses a unique ergodic invariant measure* μ_{Pr}*, and the Nusselt number* Nu *given by*

$$Nu = (Nu)_{Pr} := 1 + \frac{1}{\tilde{Ra}|\mathcal{D}|} \int \int_{\mathcal{D}} u_2 T \, dx \, d\mu_{Pr}(\mathbf{u}, T) \tag{1.8}$$

satisfies

$$\lim_{Pr \to \infty} (Nu)_{Pr} = (Nu)_\infty \, .$$

Note that $(Nu)_\infty$ *is defined by* (1.8) *relative to the unique ergodic invariant measure* μ_∞ *of* (1.5)–(1.6)*.*

It should be emphasized that, because μ_{Pr} is ergodic for large Pr, our Nusselt number has an equivalent formulation in terms of time and noise averaged solutions. That is, for μ_{Pr}-almost every initial condition $(\mathbf{u}_0, T_0)$, the number Nu given by (1.8) satisfies

$$Nu = 1 + \lim_{t \to \infty} \frac{1}{\tilde{Ra}|\mathcal{D}|} \mathbb{E} \left(\frac{1}{t} \int_0^t \int_{\mathcal{D}} u_2 T(x, s) \, dx \, ds \right), \tag{1.9}$$

where $(\mathbf{u}, T)$ is the solution of (1.1)–(1.3) with initial condition $(\mathbf{u}_0, T_0)$. The same statement can be made with respect to the Nusselt number for the infinite Prandtl system (1.5)–(1.6). Note that, if the Nusselt number given by (1.9) is reformulated without the infinite time limit, then convergence as $Pr \to \infty$ follows as a consequence of convergence of solutions on finite time intervals (see Theorem 2.1 below). The reader is advised to consult [21, Theorem 1.4] for more details regarding the Nusselt number for (1.1)–(1.3), including bounds relative to the Rayleigh numbers.

Theorem 1.1 and Corollary 1.1 may be seen as complementary to a series of recent works [43–48] which address large Prandtl number asymptotics for the Boussinesq system in a deterministic framework. Here, we show that the addition of stochastic terms allows for stronger convergence results, but the proofs require a different framework. In particular, Corollary 1.1 resolves a conjecture of Wang [48] by confirming that stochastic forcing stabilizes the Nusselt number in the infinite Prandtl number limit.

Methods of analysis

The starting point of our analysis is to establish a strict contraction property for the Markov semigroup $\{P_t^0\}_{t \geq 0}$ associated to the formal limit system (1.5)–(1.6). We show that for some $t_* > 0$ sufficiently large and for any probability measures $\mu, \tilde{\mu}$ on the phase space associated with the T component of (1.5)–(1.6), one has

$$\rho(\mu P_{t_*}^0, \tilde{\mu} P_{t_*}^0) \leq \frac{1}{2} \rho(\mu, \tilde{\mu}), \tag{1.10}$$

where ρ is an appropriately chosen Kantorovich–Wasserstein metric. See (2.19) and Theorem 2.2, (i) for a precise formulation.

The bound (1.10) is crucial since it allows us to reduce the proof of the convergence of statistically invariant states in the infinite Prandtl limit to the convergence of solutions on finite time intervals. Indeed, suppose that μ_0 is the (unique) invariant measure for (1.5)–(1.6) and for $\varepsilon > 0$ let μ_ε be the T component of any stationary solution of (1.1)–(1.2) with $\varepsilon := 1/\Pr$. Utilizing the invariance of μ_0 and (1.10) we find

$$\begin{aligned}\rho(\mu_\varepsilon, \mu_0) = \rho(\mu_\varepsilon, \mu_0 P^0_{t_*}) &\leq \rho(\mu_\varepsilon, \mu_\varepsilon P^0_{t_*}) + \rho(\mu_\varepsilon P^0_{t_*}, \mu_0 P^0_{t_*}) \\ &\leq \rho(\mu_\varepsilon, \mu_\varepsilon P^0_{t_*}) + \frac{1}{2}\rho(\mu_\varepsilon, \mu_0)\,, \end{aligned} \tag{1.11}$$

and consequently

$$\rho(\mu_\varepsilon, \mu_0) \leq 2\rho(\mu_\varepsilon, \mu_\varepsilon P^0_{t_*})\,.$$

By properties of the Wasserstein metric, specifically (2.21), and using the stationarity of the solutions corresponding to μ_ε we therefore obtain the estimate

$$\rho(\mu_\varepsilon, \mu_0) \leq 2\mathbb{E}\rho(T^\varepsilon(t_*), T^{0,\varepsilon}(t_*)). \tag{1.12}$$

We have thus bounded the distance between invariant states by the mean distance between solutions at a fixed finite time t_*. Note that these two solutions satisfy identical initial conditions which are distributed as μ_ε.

Recently the strategy leading to (1.12) has proven effective for establishing the convergence of statistically invariant states for a variety of problems; see [7,22,25,27,31]. However, in order to implement this approach, one typically faces several major challenges. A first challenge is to prove the contraction estimate (1.10), where the semigroup $\{P^0_t\}_{t\geq 0}$ corresponds to (1.5)–(1.6). Moreover, in our setting, it is desirable to lift this contraction property to the extended phase space involving both the velocity $\mathbf{u}$ and temperature components T of our system. This is particularly relevant in view of the physical significance of the Nusselt number, a quantity involving both $\mathbf{u}$ and T as in (1.8). A second challenge is to show that $\mathbb{E}\rho(T^\varepsilon(t_*), T^{0,\varepsilon}(t_*)) \to 0$ as $\varepsilon \to 0$ in order to take advantage of (1.12). This task requires suitable $\epsilon = Pr^{-1}$ uniform moment bounds on the stationary statistics μ^ε and finite time convergence results for solutions in the limit as $Pr \to \infty$. As we describe presently the results established here require new ideas in comparison to the aforementioned related works. This is partially due to the presence of non-homogeneous boundary conditions for (1.5)–(1.6) and to the singular nature of the limit from (1.1)–(1.2) to (1.5)–(1.6).

Regarding the first challenge, guided by the classical Doob–Khasminskii Theorem [11,15,30] and as encompassed by the more recent developments in [25,27,29], one can establish a contraction of the type (1.10) when the Markov semigroup is smoothing, suitable moment bounds hold, and there is some form of irreducibility in the dynamics. The question of smoothing for the Markov semigroup can be translated to a control problem; see (B.10) below. In our setting, when the number of forced directions $N = N(Ra, \tilde{Ra})$ is sufficiently large, an appropriate control can be found through Foias–Prodi type considerations [18]. Since (1.5)–(1.6) may be seen as an advection diffusion system with $\mathbf{u}$ being two derivatives smoother than T, such a strategy largely repeats the approach used in previous works on the 2D stochastic Navier–Stokes equations [26,27,31,50]. On the other hand establishing suitable moment bounds is more delicate due to the non-homogenous boundary conditions imposed in (1.3) and requires a careful use of the maximum principle along with exponential martingale estimates. These bounds have been carried out in our companion work [21]. The main obstacle to proving (1.10) is to establish irreducibility, which does not follow from the approach set out in previous works, e.g. [10, 20, 22, 26, 49]. This is because the system (1.5)–(1.6) with its stochastic terms removed can have highly non-trivial dynamics, see [8, 14, 37, 38, 44]. We show that, despite this complication, the support of every invariant measure of (1.5)–(1.6) contains the basic conductive state. Indeed we establish with the use of another Foias–Prodi bound that a Girsanov shift of (1.5)–(1.6) converges to the conductive state with positive probability. We then employ moment estimates and stopping time arguments to translate this non-zero probability back to the original system (1.5)–(1.6) yielding the desired irreducibility.

In order to establish convergence of invariant states on the extended phase space, we adapt a methodology from recent joint work of the authors with Friedlander [22] which enhances (1.10) to a "lifted" contraction property with respect to a carefully chosen metric (see Lemma 3.1 below). By invoking this lifted contraction property, and appropriately modifying the argument in (1.11)–(1.12), the convergence of invariant states as $Pr \to \infty$ reduces to establishing the convergence of solutions of (1.1)–(1.2) to those of (1.5)–(1.6) at a fixed time $t_* > 0$, independent of ε, when the initial conditions have the same distribution in temperature only.

The second major challenge regards the convergence of solutions of (1.1)–(1.2) on finite time intervals as $Pr \to \infty$ for which we develop a suitable asymptotic analysis. This is a non-trivial task since the small parameter $1/Pr$ lies in front of the time derivative terms in (1.1). Moreover, the convergence analysis in [43–48] for a deterministic analogue of (1.1)–(1.2) requires significant modification. In particular these references crucially use higher temporal regularity properties which are missing in our stochastic setting. As a substitute we derive a stochastic evolution equation for the velocity component and use martingale properties of associated Itō integrals. Our analysis then takes advantage of uniform moment estimates from [21], some previously unobserved cancellations in certain error terms and delicate stopping time arguments.

Analogous to the results in [43–48] we derive an 'intermediate system', which we refer to as the 'corrector'. We show rigorously that this system approximates the finite Prandtl system in the velocity equation over bounded time intervals up the initial time; cf Theorem 5.1. While this corrector system is of independent interest we also provide a somewhat simpler and more direct analysis of the convergence of (1.1)–(1.2) to (1.5)–(1.6) which well approximates the infinite Prandtl system after an $O(1/Pr)$ time transient. Indeed, this more direct approach is sufficient for the upper bound in (1.12) since this bound only involves a fixed time $t_* > 0$.

Manuscript organization

The manuscript is organized as follows. In Sect. 2 we introduce the rigorous mathematical setting of the stochastic Boussinesq equations, (1.1)–(1.3), which serves as a foundation for the rest of the analysis. We also introduce the formalities of the Kantorovich–Wasserstein metric in Sect. 2.2, and provide a rigorous formulation of our main results in Sect. 2.3. Section 3 describes core of our strategy that reduces the question of convergence to finite time asymptotics and uniform moment bounds. Section 4 is devoted to establishing the contraction (1.10) for the infinite Prandtl system (1.5)–(1.6). In Sect. 5 we carry out the finite time convergence analysis. The section concludes with a derivation and analysis of the intermediate corrector system. In Sect. 6 we establish convergence of the Nusselt number. Finally two Appendices recall various elements essentially contained in previous works that we have used in our analysis. Appendix A is devoted to details for various moment estimates from [21] for a class of drift-diffusion equations which we use to bound (1.5)–(1.6). In Appendix B we outline gradient estimates on the Markov semigroup corresponding to (1.5)–(1.6) which are carried out in a similar fashion to e.g. [26].

2. Mathematical preliminaries and main results

We begin our analysis of the stochastic Boussinesq equations by recalling some details of their mathematical setting. The section concludes with a mathematically precise restatement of Theorem 1.1. Here and below we implicitly assume that C, c, C_0 etc. are constants depending on the domain $\mathcal{D}$ with any other dependency indicated explicitly.

For the forthcoming analysis it is convenient to consider an equivalent homogeneous, form of the stochastic Boussinesq equations. Introducing the 'small parameter' $\varepsilon = Pr^{-1} > 0$ and making the change of variable $\theta^\varepsilon = T - \tilde{Ra}(1-z)$ we rewrite (1.1)–(1.2) as[3]

$$\varepsilon(\partial_t \mathbf{u}^\varepsilon + \mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^\varepsilon) - \Delta \mathbf{u}^\varepsilon = \nabla \tilde{p}^\varepsilon + Ra\hat{\mathbf{k}}\theta^\varepsilon, \quad \nabla \cdot \mathbf{u}^\varepsilon = 0, \tag{2.1}$$

$$d\theta^\varepsilon + \mathbf{u}^\varepsilon \cdot \nabla \theta^\varepsilon dt = \tilde{Ra} \cdot u_3^\varepsilon dt + \Delta \theta^\varepsilon dt + \sum_{k=1}^{N} \sigma_k dW^k, \tag{2.2}$$

supplemented with the homogenous boundary conditions

$$\begin{gathered} \mathbf{u}^\varepsilon_{|z=0} = \mathbf{u}^\varepsilon_{|z=1} = 0, \quad \theta^\varepsilon_{|z=0} = \theta^\varepsilon_{|z=1} = 0, \\ \mathbf{u}^\varepsilon, \theta^\varepsilon \text{ are periodic in } \mathbf{x} = (x, y). \end{gathered} \tag{2.3}$$

Here, in reference to the $\tilde{Ra} \cdot u_3^\varepsilon$ term in (2.2) recall that $\mathbf{u}^\varepsilon = (u_1^\varepsilon, u_2^\varepsilon, u_3^\varepsilon)$. The corresponding infinite Prandtl system ($\varepsilon = 0$) is given by

$$-\Delta \mathbf{u}^0 = \nabla \tilde{p} + Ra\hat{\mathbf{k}}\theta^0, \quad \nabla \cdot \mathbf{u}^0 = 0, \tag{2.4}$$

$$d\theta^0 + \mathbf{u}^0 \cdot \nabla \theta^0 dt = \tilde{Ra} \cdot u_3^0 dt + \Delta \theta^0 dt + \sum_{k=1}^{N} \sigma_k dW^k, \tag{2.5}$$

again with initial conditions $\theta^0(0) = \theta_0^0$ and boundary conditions as in (2.3).

[3]Note that we have implicitly modified the pressure in (2.1) by $Ra\tilde{Ra}(z - \frac{1}{2}z^2)$ since $(1-z)\hat{\mathbf{k}} = \nabla(z - \frac{1}{2}z^2)$.

Remark 2.1. Notice that we do not prescribe an initial condition for $\mathbf{u}^0$ in (2.4)–(2.5) as this component does not satisfy an independent evolution equation. Indeed, (2.4)–(2.5) can be rewritten as

$$d\theta^0 + (M\theta^0) \cdot \nabla\theta^0 dt = \tilde{Ra}(M\theta^0)_3 dt + \Delta\theta^0 dt + \sum_{k=1}^{N} \sigma_k dW^k, \tag{2.6}$$

where the constitutive law M recovers $\mathbf{u}$ from θ according to (2.4) as in (2.9) below.

Remark 2.2. The systems (2.1)–(2.2) and (2.4)–(2.5) can be reformulated in terms of $T = \theta^\varepsilon + \tilde{Ra}(1 - z)$, which satisfies (1.1)–(1.2) or (1.5)–(1.6), respectively, and has boundary conditions (1.3). Our analysis makes use of both of these formulations.

Remark 2.3. As noted above, parameters in the problem are the Prandtl ($Pr = \varepsilon^{-1}$) and Rayleigh numbers (Ra, $\tilde{Ra}$), which are unit-less. In terms of basic physical quantities of interest we have that

$$\varepsilon^{-1} = Pr = \frac{\nu}{\kappa}, \quad Ra = \frac{g\alpha\gamma h^{5/2}}{\nu\kappa^{3/2}}, \quad \tilde{Ra} = \frac{\sqrt{\kappa h}(T_b - T_t)}{\gamma}, \tag{2.7}$$

where ν is the kinematic viscosity, κ the thermal diffusivity, g the gravitational constant, α the coefficient of thermal expansion, h the distance between the confining plates, $T_b - T_t$ the temperature differential, and $\gamma = \mathcal{H}/\rho c$ the intensity $\mathcal{H}$ of the volumetric heat flux normalized by the density ρ and specific heat c of the fluid. We refer the interested reader to [21], where the dimensionless form of the stochastically driven Boussinesq equations is derived.

2.1. Functional setting of the Boussinesq equations

We next define the phase space for the Boussinesq equations, which is very close to the classical framework for the Navier–Stokes equations; see e.g. [9,40] for further details.

We define $\mathbf{H} := H_1 \times H_2$ as the phase space for (2.1)–(2.3), where

$$H_1 := \{\mathbf{u} \in (L^2(\mathcal{D}))^3 : \nabla \cdot \mathbf{u} = 0, \mathbf{u} \cdot \mathbf{n}_{|z=0,1} = 0, \mathbf{u} \text{ is periodic in } \mathbf{x}\},$$
$$H_2 := \{\theta \in L^2(\mathcal{D}) : \theta \text{ is periodic in } \mathbf{x}\}$$

and we denote by $H = H_2$ the phase space for (2.4)–(2.5). The spaces $\mathbf{H}$ and H are endowed with the standard L^2-norm and we denote each of them by $\|\cdot\|$ as the appropriate meaning will be clear from the context.[4] All other norms are written as $\|\cdot\|_{\mathbb{X}}$ below for a given space $\mathbb{X}$. We define H^1 type spaces as

$$V_1 := \{\mathbf{u} \in (H^1(\mathcal{D}))^3 : \nabla \cdot \mathbf{u} = 0, \mathbf{u}_{|z=0,1} = 0, \mathbf{u} \text{ is periodic in } \mathbf{x}\},$$
$$V_2 := \{\theta \in H^1(\mathcal{D}) : \theta_{|z=0,1} = 0, \theta \text{ is periodic in } \mathbf{x}\}.$$

[4]Below will also consider the weighted metrics (2.22), (2.24) which generate an equivalent topology on $\mathbf{H}$ and H but are more suitable for the convergence of measures in the associated Wasserstein metric.

Let $\mathbf{V} = V_1 \times V_2$ and $V = V_2$. We will sometimes consider the $L^p(\mathcal{D})$ spaces of p-integrable functions for $p \in [1, \infty]$ and endow these spaces with their standard norms.

In what follows we frequently project or lift the dynamics to account for the phase space mismatch between (2.1)–(2.2) and (2.4)–(2.5). We define

$$\Pi : \mathbf{H} \to H_2 \text{ to be the projection onto the } \theta \text{ component of } \mathbf{H}. \tag{2.8}$$

Associated with the limit system (2.4)–(2.5) we have the constitutive law

$$M(\theta) = RaA^{-1}P\theta\hat{k}, \tag{2.9}$$

where A is the Stokes operator and P the Leray projector. In other words $\mathbf{u} = M(\theta)$ is the solution of

$$-\Delta\mathbf{u} = \nabla\tilde{p} + Ra\hat{\mathbf{k}}\theta, \quad \nabla \cdot \mathbf{u} = 0;$$

see Sect. 5.1 and in particular (5.1) below. We define the 'lifting map' $L : H \to \mathbf{H}$ from the temperature component to the extended phase space

$$L(\theta) = (M(\theta), \theta). \tag{2.10}$$

Finally, we denote $\Pr(\mathbb{X})$ as the space of Borel probability measures on a given complete metrizable space $\mathbb{X}$, typically $H, \mathbf{H}$ etc. For $\mu \in \Pr(\mathbf{H})$, we take $\Pi\mu(\cdot) = \mu(\Pi^{-1}(\cdot))$ to be the push-forward of μ by Π. Similarly $L\mu$ is the push-forward of μ by L when $\mu \in \Pr(H)$.

We have the following general results concerning the existence and uniqueness of solutions of (2.1)–(2.3) and (2.4)–(2.5):

Proposition 2.1. (Existence, Uniqueness, and Continuous Dependence) *Fix any values $Ra, \tilde{Ra} > 0$.*

(i) *For every $\epsilon > 0$ and any given $\mu^0 \in \Pr(\mathbf{H})$ with $\int(\|\mathbf{u}\|^2 + \|\theta\|^2)d\mu^0(\mathbf{u}, \theta) < \infty$ there exists a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, W)$ upon which is defined an $\mathbf{H}$-valued stochastic process $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ with the regularity*

$$(\mathbf{u}^\varepsilon, \theta^\varepsilon) \in L^2(\Omega; L^2_{loc}([0, \infty); \mathbf{V}) \cap L^\infty_{loc}([0, \infty); \mathbf{H})),$$

which is weakly continuous in $\mathbf{H}$, adapted to $\{\mathcal{F}_t\}_{t\geq 0}$, satisfies (2.1)–(2.2) weakly and such that $(\mathbf{u}^\varepsilon(0), \theta^\varepsilon(0))$ is distributed as μ^0. We say that such a pair $(\mathcal{S}, (\mathbf{u}^\varepsilon, \theta^\varepsilon))$ is a weak-martingale solution of *(2.1)–(2.3). If, for some $p \geq 2$, and $\eta > 0$,*

$$\int_{\mathbf{H}} \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|^2_{L^p}))d\mu^0(\mathbf{u}, \theta) < \infty, \tag{2.11}$$

there exists $\eta_0 > 0$ and a weak martingale solution $(\mathcal{S}, (\mathbf{u}^\varepsilon, \theta^\varepsilon))$ such that

$$\mathbb{E}\exp\left(\eta_0\left(\sup_{s\in[0,t]}(\|\mathbf{u}^\varepsilon\|^2 + \|\theta^\varepsilon\|^2_{L^p}) + \int_0^t(\|\nabla\mathbf{u}^\varepsilon\|^2 + \|\nabla\theta^\varepsilon\|^2)ds\right)\right) \leq C < \infty \tag{2.12}$$

for each $t > 0$, where $C > 0$ is a constant independent of $\varepsilon \in (0, 1]$.

(ii) *Additionally, for any* $\varepsilon > 0$*, there exists a martingale solution* $(\mathcal{S}, (\mathbf{u}^\varepsilon_S, \theta^\varepsilon_S))$ *of* (2.1)–(2.2) *which is stationary in time, meaning that the law of the solution is independent of time. These stationary solutions* $(\mathcal{S}, (\mathbf{u}^\varepsilon_S, \theta^\varepsilon_S))$ *may be chosen in such a way that, for any* $p \geq 2$ *there is an* $\eta = \eta(p, Ra, \tilde{R}a) > 0$*, for which*

$$\sup_{1\geq\varepsilon>0} \int_{\mathbf{H}} \exp(\eta(\|\mathbf{u}\|^2 + \|\theta\|^2_{L^p}))d\mu_\varepsilon(\mathbf{u}, \theta) = C_0 < \infty, \tag{2.13}$$

where $\mu_\varepsilon(\cdot) = \mathbb{P}((\mathbf{u}^\varepsilon_S(t), \theta^\varepsilon_S(t)) \in \cdot)$ *for any fixed* $t \geq 0$*.*

(iii) *Now consider the case when* $\varepsilon = 0$*. Fix a stochastic basis* $\mathcal{S}$ *and any* $\mathcal{F}_0$*-measurable random variable* $\theta_0 \in L^2(\Omega, H)$*. Then there exists a unique process* θ^0 *with*

$$\theta^0 \in L^2(\Omega; L^2_{loc}([0, \infty); V) \cap C([0, \infty); H)), \tag{2.14}$$

which is $\mathcal{F}_t$*-adapted, weakly solves* (2.4)–(2.5)*, and satisfies the initial condition* $\theta^0(0) = \theta_0$*.*

(iv) *For a given stochastic basis* $\mathcal{S}$ *and each* $\theta_0 \in H$ *denote* $\theta^0(\cdot, \theta_0, W)$ *as the unique corresponding stochastic process satisfying* (2.4)–(2.5) *with* (2.14)*. We have that* $\theta_0 \mapsto \theta^0(t, \theta_0, W)$ *is Fréchet differentiable in* $\theta_0 \in H$ *for any* $t \geq 0$ *and any fixed realization* $W(\cdot) = W(\cdot, \omega)$*. On the other hand* $W \mapsto \theta^0(t, \theta_0, W)$ *is Fréchet differentiable in* W *from* $C_0([0, t], \mathbb{R}^N)$ *to* H *for each fixed* $\theta_0 \in H$ *and* $t > 0$*.*

These results are standard for a systems like (2.1)–(2.3) and (2.4)–(2.5); see e.g. [12,13,17,23]. The only novelty in view of existing methods is the uniform moment bound (2.13). The existence of such a collection of solutions is established using the maximum principle and exponential moment bounds in the companion work [21]; cf. Appendix A below.

The Markovian framework for (2.4)–(2.5) is defined as follows. The transition functions are given by

$$P^0_t(\theta_0, A) := \mathbb{P}(\theta^0(t, \theta_0) \in A), \quad t \geq 0,\ \theta_0 \in H, A \in \mathcal{B}(H), \tag{2.15}$$

where $\mathcal{B}(H)$ denotes the Borel sets of H, and the associated semigroup is given by

$$P^0_t \phi(\theta_0) := \mathbb{E}\phi(\theta^0(t, \theta_0)), \quad t \geq 0,\ \phi \in M_b(H), \tag{2.16}$$

where $M_b(H)$ is the set of bounded measurable functions on H. In view of the continuous dependence on initial conditions the semigroup $\{P^0_t\}_{t\geq0}$ is Feller, that is, it maps the set of continuous bounded functions on H, $C_b(H)$, to itself. This semigroup acts on Borelian probability measures μ according to

$$\mu P^0_t(A) = \int_H P^0_t(\theta, A)d\mu(\theta), \quad A \in \mathcal{B}(H). \tag{2.17}$$

A measure $\mu \in Pr(H)$ is said to be invariant with respect to $\{P^0_t\}_{t\geq0}$ if $\mu P^0_t = \mu$ for all $t \geq 0$. Recall that in three space dimensions the Markovian framework for the full system with $\varepsilon > 0$ cannot be implemented due to a lack of global well-posedness.

As an immediate consequence of bounds in Appendix A and the Krylov–Bogolyubov averaging technique we have

Lemma 2.1. *Under the assumptions of Proposition 2.1 and for any* $Ra, \tilde{Ra}$ *there exists an invariant measure* μ_0 *of the Markov semigroup* P_t^0*. Moreover for any such measure*

$$\int_H \exp(\eta\|\theta\|_{L^p}^2)d\mu_0(\theta) \le C_0 < \infty, \tag{2.18}$$

for any $p \ge 2$ *and any suitably small* $\eta = \eta(p, Ra, \tilde{Ra})$*.*

2.2. Wasserstein distance, weighted metrics and associated observables

We next recall the general setting of the Kantorovich–Wasserstein distance in which we establish our convergence results. We then introduce weighted metrics on H and $\mathbf{H}$ along with some associated classes of observable which are used to measure distances between measures in the analysis below.

Let $(\mathbb{X}, \rho)$ be a complete metric space and take $\mathrm{Pr}_1(\mathbb{X}, \rho)$ to be the set of Borel probability measures μ on $\mathbb{X}$ with $\int \rho_\eta(0,\theta)d\mu(\theta) < \infty$. On $\mathrm{Pr}_1(\mathbb{X}, \rho)$ we define the Kantorovich–Wasserstein metric, relative to ρ, equivalently as[5]

$$\begin{aligned} \rho(\mu, \tilde{\mu}) &:= \sup_{\|\phi\|_{Lip,\rho} \le 1} \left| \int_{\mathbb{X}} \phi(\theta)d\mu(\theta) - \int_{\mathbb{X}} \phi(\theta)d\tilde{\mu}(\theta) \right| \\ &= \inf_{\Gamma \in \mathcal{C}(\mu,\tilde{\mu})} \int_{\mathbb{X}\times\mathbb{X}} \rho(\theta, \tilde{\theta})d\Gamma(\theta, \tilde{\theta}), \end{aligned} \tag{2.19}$$

where

$$\|\phi\|_{Lip,\rho} := \sup_{\theta \ne \tilde{\theta}} \frac{|\phi(\theta) - \phi(\tilde{\theta})|}{\rho(\theta, \tilde{\theta})} \tag{2.20}$$

for $\phi : \mathbb{X} \to \mathbb{R}$, and $\mathcal{C}(\mu, \tilde{\mu})$ is the collection of Borel probability measures Γ in $Pr(\mathbb{X} \times \mathbb{X})$ with $\mu, \tilde{\mu}$ as its marginals. Hence, the last term in (2.19) is equivalent to

$$\rho(\mu, \tilde{\mu}) = \inf \mathbb{E}\rho(X, Y)\,, \tag{2.21}$$

where the infimum is taken over all $\mathbb{X}$-valued random variables X, Y distributed as $\mu, \tilde{\mu}$ respectively. See e.g. [16,42] for further background on these metrics.

Specializing to our current setting, the following metrics on H and $\mathbf{H}$ prove useful for measuring the distance between the laws of solutions of (2.1)–(2.2) and (2.4)–(2.5). Following e.g. [27] we introduce, for $\eta > 0$, the weighted metric on H as

$$\rho_\eta(\theta, \tilde{\theta}) = \inf_{\substack{\gamma \in C^1([0,1];H) \\ \gamma(0)=\theta, \gamma(1)=\tilde{\theta}}} \int_0^1 \exp(\eta\|\gamma\|^2)\|\gamma'(s)\|ds, \tag{2.22}$$

for any $\theta, \tilde{\theta} \in H$. Notice that

$$\|\theta - \tilde{\theta}\| \le \rho_\eta(\theta, \tilde{\theta}) \le \exp(2\eta(\|\theta\|^2 + \|\tilde{\theta}\|^2))\|\theta - \tilde{\theta}\|\,, \tag{2.23}$$

[5]Here note slight abuse of notation wherein we denote both the underlying metric and its Wasserstein by ρ; the meaning of ρ will be clear from context in what follows.

for $\theta, \tilde{\theta} \in H$. For the extended phase space $\mathbf{H}$, similarly to our recent work [22], we take

$$\tilde{\rho}_\eta((\mathbf{u}, \theta), (\tilde{\mathbf{u}}, \tilde{\theta})) = \|\mathbf{u} - \tilde{\mathbf{u}}\|_{H^1} + \rho_\eta(\theta, \tilde{\theta}), \tag{2.24}$$

again defined for any $\eta > 0$.

For the statement of the main results we consider the following class of 'observables'

$$\mathcal{V}(\mathbf{H}) = \mathcal{V}_\eta(\mathbf{H}) := \left\{\phi \in C^1(\mathbf{H}) : [\phi]_\eta < \infty\right\},$$

where the semi-norm $[\cdot]_\eta$ is given by

$$[\phi]_\eta := \sup_{(u,\theta)\in\mathbf{H}} \left[\sup_{\zeta\in\mathbf{H}, \|\zeta\|=1} |\nabla_u \phi(u,\theta)\cdot\zeta| + \exp(-\eta\|\theta\|) \sup_{\xi\in H, \|\xi\|=1} |\nabla_\theta \phi(u,\theta)\cdot\xi| \right].$$

Note that, as in [27, Proposition 4.1],

$$\|\phi\|_{Lip, \tilde{\rho}_\eta} \leq C[\phi]_\eta, \tag{2.25}$$

for any $\phi \in C^1(\mathbf{H})$ with the constant C independent of ϕ.

2.3. Statement of the main results

We now precisely formulate the main results of this work on the convergence of solutions when $\Pr \to \infty$. We begin with the following finite time convergence result:

Theorem 2.1. *For each $\varepsilon \in (0,1)$, let $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ with its associated stochastic basis $\mathcal{S}$ be a martingale solution of (2.1)–(2.2) in the sense of Proposition 2.1. Relative to this $\mathcal{S}$, let θ^0 be a solution of (2.4)–(2.5). Suppose there exists $C_0, \eta > 0$ such that*[6]

$$\sup_{\varepsilon>0} \mathbb{E}\left[\exp(\eta(\|\mathbf{u}^\varepsilon(0)\|^2 + \|\theta^\varepsilon(0)\|^2_{L^3} + \|\theta^0(0)\|^2_{L^3}))\right] \leq C_0 < \infty, \tag{2.26}$$

and suppose that $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ maintains (2.13). Then, for each $t > 0$, there exists $\gamma_0 > 0$, $C > 0$ such that

$$\begin{aligned} &\mathbb{E}\left(\sup_{s\in[0,t]} \|\theta^\varepsilon(s) - \theta^0(s)\|^p + \int_0^t \|\mathbf{u}^\varepsilon(s) - M(\theta^0)(s)\|^2_{H^1} ds\right) \\ &\quad \leq C\left(\varepsilon^\gamma + \left(\mathbb{E}\|\theta^\varepsilon(0) - \theta^0(0)\|^2 + \varepsilon\mathbb{E}\|\mathbf{u}^\varepsilon(0) - M(\theta^0)(0)\|^2\right)^\gamma\right), \end{aligned} \tag{2.27}$$

for each $0 < \gamma \leq \gamma_0$ and any $p > \gamma$. Here, the constants $C = C(p, \eta, Ra, \tilde{Ra}, C_0, \|\sigma\|_{L^3}, t)$ and $\gamma_0 = \gamma_0(\eta, Ra, \tilde{Ra}, C_0, \|\sigma\|_{L^3}, t)$ are independent of $\varepsilon > 0$ and depends on the initial conditions only through C_0.

The proof of Theorem 2.1 is established in Sect. 5.2.

[6] Although we can relax the assumption on the initial velocity field to qth moment bounds for some $q \geq 4$, we have opted to impose an exponential moment condition for simplicity of presentation.

Remark 2.4. It is worth noting that since $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ are only martingale solutions the associated stochastic bases $\mathcal{S}$ are not unique and could in fact vary as a function of ε; that is, we cannot assume that these solutions are all defined relative to the *same* stochastic basis. Similar remarks apply to the bound (2.30) below. However, crucially, in both (2.27) (2.30) the constants do not depend on the choice of basis. Thus, since this subtlety does not cause any trouble in what follows, we shall henceforth suppress this technical point in order to avoid notational confusion.

We next state our results regarding the convergence of statistically invariant states to the unique invariant measure of the formal limit system (cf. Theorem 1.1).

Theorem 2.2. *Let $\{P_t^0\}_{t\geq 0}$ be the Markov semigroup associated to (2.4)–(2.5) defined in (2.16). There exists $N_0 > 0$ and $\eta_0 > 0$ depending only on Ra and $\tilde{Ra}$ such that if $N \geq N_0$, where N is the number of stochastically forced modes in (2.4)–(2.5), then the following bounds hold:*

(i) *For some $\gamma, C > 0$ depending only on Ra and $\tilde{Ra}$*

$$\rho_\eta(\mu P_t^0, \tilde{\mu} P_t^0) \leq C \exp(-\gamma t) \rho_\eta(\mu, \tilde{\mu}), \tag{2.28}$$

for any $\mu, \tilde{\mu} \in \mathrm{Pr}_1(H, \rho_\eta)$, $\eta \in (0, \eta_0)$ and every $t \geq 0$, where ρ_η is defined in (2.19). In particular, there exists a unique ergodic invariant measure $\mu_0 \in Pr_1(H, \rho_\eta)$ of (2.4)–(2.5).

(ii) *Suppose that $\{\mu_\varepsilon\}_{\varepsilon>0}$ is any collection of measures corresponding to stationary martingale solutions of (2.1)–(2.3) and satisfying the uniform bound (2.13) for any $\eta \in (0, \eta_0)$ and some $p \geq 3$. Let μ_0 be the unique invariant measure of (2.4)–(2.5). Then, there exists $\tilde{q} = \tilde{q}(Ra, \tilde{Ra})$, $\tilde{C} = \tilde{C}(Ra, \tilde{Ra})$, independent of $\varepsilon > 0$, such that*

$$\tilde{\rho}_\eta(\mu_\varepsilon, L\mu_0) \leq \tilde{C} \varepsilon^{\tilde{q}} \tag{2.29}$$

for every $\varepsilon > 0$. Consequently, for the stationary processes $(\mathbf{u}_S^\varepsilon, \theta_S^\varepsilon)$ and θ_S^0, distributed as μ_ε and μ_0, respectively,

$$|\mathbb{E}(\phi(\mathbf{u}_S^\varepsilon, \theta_S^\varepsilon) - \phi(L\theta_S^0))| \leq \tilde{C} [\phi]_\eta \varepsilon^{\tilde{q}} \tag{2.30}$$

for any $\phi \in \mathcal{V}(\mathbf{H})$.

The proof of (i) is carried out in Sect. 4 with some technical details relegated to Appendices A and B. In Sect. 3 we describe a general strategy which shows that, under the conditions of Theorem 2.2, (2.29) follows from (2.28) and (2.27).

We conclude this section by making several important remarks.

Remark 2.5. (i) Assertions of Theorem 2.2 also hold in two space dimensions and in addition one can show that (2.1)–(2.3) has a well defined Markov semigroup. Thus, any statistically invariant state corresponds to an invariant measure of the associated semigroup. This allows us to show in [21] that the ε-independent exponential moment bounds in (2.13) hold for all invariant measures.

(ii) In 3D, the existence of a sequence of statistically invariant states of (2.1)–(2.3) satisfying the uniform moment bound (2.13) is established in the companion work [21]. However we have not been able to show that *every* sequence of invariant states have such (uniform) exponential moments.
(iii) In Sect. 5 we also derive a 'corrector' system which well approximates the dynamics of the velocity field of the full system (2.1)–(2.2) up to the initial time for large values of Pr or equivalently small $\varepsilon > 0$. See Theorem 5.1 below for further details. Note however that this more refined version of (2.27) is not needed in order to achieve (2.29), (2.30).

3. Reduction to finite time dynamics

In this section we describe a general strategy for reducing the convergence of measures to finite time asymptotics and uniform moment bounds when the formal limit system satisfies a suitable mixing condition as in (2.28). To fix ideas we assume the conditions of Theorem 2.2 throughout this section. Also, we assume that both (2.28), (2.27) hold; we establish these bounds rigorously below in Sects. 4, 5 respectively. The reader should note that the presented method is flexible and can be applied in a variety of settings. See, for example, [22,25,27,31].

We adapt some ideas from our recent work [22, Section 5] to the present setting. For $\eta > 0$ take

$$\rho^*_\eta(\theta, \tilde{\theta}) = \tilde{\rho}_\eta(L(\theta), L(\tilde{\theta}))\,,$$

where $\tilde{\rho}_\eta$ is defined in (2.24) and recall that L is the lifting operator given in (2.10). It is not hard to show that the metrics ρ_η and ρ^*_η are equivalent (see [22] for details), and consequently the associated Wasserstein metrics on H are also equivalent. Then from (2.28) we obtain the following result, see [22, Corollary 5.4] and surrounding commentary for further details.

Lemma 3.1. *Under the same conditions as Theorem 2.2 (i), we have*

$$\tilde{\rho}_\eta(L(\mu P^0_t), L(\tilde{\mu} P^0_t)) \leq C e^{-\gamma t} \tilde{\rho}_\eta(L(\mu), L(\tilde{\mu})) \tag{3.1}$$

for any $\mu, \tilde{\mu} \in Pr_1(H, \rho_\eta)$, and every $t \geq 0$.

Using (3.1) choose $t^* > 0$ to guarantee that

$$\tilde{\rho}_\eta(L(\mu P^0_{t^*}), L(\tilde{\mu} P^0_{t^*})) \leq \frac{1}{2} \tilde{\rho}_\eta(L(\mu), L(\tilde{\mu})). \tag{3.2}$$

By the invariance of μ_0

$$\begin{aligned}\tilde{\rho}_\eta(\tilde{\mu}, L\mu_0) &\leq \tilde{\rho}_\eta(\tilde{\mu}, L((\Pi\tilde{\mu}) P^0_{t+t^*})) + \tilde{\rho}_\eta(L((\Pi\tilde{\mu}) P^0_{t+t^*}), L(\mu_0 P^0_{t+t^*}))\\ &\leq \tilde{\rho}_\eta(\tilde{\mu}, L((\Pi\tilde{\mu}) P^0_{t+t^*})) + \frac{1}{2}\left[\tilde{\rho}_\eta(L((\Pi\tilde{\mu}) P^0_t), \tilde{\mu}) + \tilde{\rho}_\eta(\tilde{\mu}, L\mu_0)\right]\end{aligned}$$

for any $t \geq 0$ and any other measure $\tilde{\mu} \in \Pr(\mathbf{H})$. Here recall that Π is the projection operator defined in (2.8). Rearranging, taking a time average we obtain

$$\tilde{\rho}_\eta(\tilde{\mu}, L\mu_0) \leq \frac{2}{t^*}\int_0^{t^*} \left[\tilde{\rho}_\eta(\tilde{\mu}, L((\Pi\tilde{\mu})P^0_{t+t^*})) + \tilde{\rho}_\eta(\tilde{\mu}, L((\Pi\tilde{\mu})P^0_t))\right] dt$$
$$= \frac{2}{t^*}\int_0^{2t^*} \tilde{\rho}_\eta(\tilde{\mu}, L((\Pi\tilde{\mu})P^0_t)) dt\,. \tag{3.3}$$

With (3.3) now in hand, we consider a sequence of stationary martingale solutions $\{(\mathbf{u}^\varepsilon_S, \theta^\varepsilon_S)\}_{\varepsilon>0}$ and take $\{\mu_\varepsilon\}_{\varepsilon>0} \subset \Pr(\mathbf{H})$ to be the corresponding collection of stationary measures. We suppose that $\{\mu_\varepsilon\}_{\varepsilon>0}$ satisfies the uniform moment condition (2.13) as in Proposition 2.1, (ii). We also denote $\theta^{0,\varepsilon}_S$ (and $M(\theta^{0,\varepsilon}_S)$) the solution of (2.4), (2.5) with the initial condition $\theta^\varepsilon(0)$ so that, for every $t > 0$, the law of $\theta^{0,\varepsilon}_S(t)$ is $(\Pi\mu_\varepsilon)P^0_t$. Consequently, (2.21), (2.24) yield

$$\tilde{\rho}_\eta(\mu_\varepsilon, L((\Pi\mu_\varepsilon)P^0_t)) \leq \mathbb{E}\|\mathbf{u}^\varepsilon_S(t) - M(\theta^{0,\varepsilon}_S(t))\|_{H^1} + \mathbb{E}\rho_\eta(\theta^\varepsilon_S(t), \theta^{0,\varepsilon}_S(t)), \tag{3.4}$$

where we recall M is defined as in (2.9). By (2.23) one has, for any $q > 0$,

$$\mathbb{E}\rho_\eta(\theta^\varepsilon_S(t), \theta^{0,\varepsilon}_S(t)) \leq \mathbb{E}\left(\exp(2\eta(\|\theta^\varepsilon_S(t)\|^2 + \|\theta^{0,\varepsilon}_S(t)\|^2))\|\theta^\varepsilon_S(t) - \theta^{0,\varepsilon}_S(t))\|\right)$$
$$\leq C\mathbb{E}\left(\exp(3\eta(\|\theta^\varepsilon_S(t)\|^2 + \|\theta^{0,\varepsilon}_S(t)\|^2))\|\theta^\varepsilon_S(t) - \theta^{0,\varepsilon}_S(t)\|^{q/2}\right)$$
$$\leq C\Big(\mathbb{E}\exp(12\eta\|\theta^\varepsilon_S(t)\|^2)\cdot\mathbb{E}\exp(12\eta\|\theta^{0,\varepsilon}_S(t)\|^2))\Big)^{1/4}$$
$$\cdot\left(\mathbb{E}\|\theta^\varepsilon_S(t) - \theta^{0,\varepsilon}_S(t)\|^q\right)^{1/2}.$$

Using (A.4) with $p = 2$ we obtain

$$\mathbb{E}\rho_\eta(\theta^\varepsilon_S(t), \theta^{0,\varepsilon}_S(t))$$
$$\leq C\big(\mathbb{E}\exp(96\eta\|\theta^\varepsilon_S(0)\|^2)\big)^{1/2}\left(\mathbb{E}\|\theta^\varepsilon_S(t) - \theta^{0,\varepsilon}_S(t)\|^q\right)^{1/2}. \tag{3.5}$$

Finally combining (3.3) with (3.4), (3.5) we obtain

$$\tilde{\rho}_\eta(\mu_\varepsilon, L\mu_0) \leq C\mathbb{E}\int_0^{2t*} \|\mathbf{u}^\varepsilon_S(t) - M(\theta^{0,\varepsilon}_S(t))\|_{H^1} dt$$
$$+ C\left(\mathbb{E}\exp(96\eta\|\theta^\varepsilon_S(0)\|^2)\right)^{1/2}\sup_{t\in[0,2t^*]}\left(\mathbb{E}\|\theta^\varepsilon_S(t) - \theta^{0,\varepsilon}_S(t)\|^q\right)^{1/2}, \tag{3.6}$$

which holds for any $t^* > 0$ such that (3.2) holds.

With (3.6) established we conclude this section by detailing the proof of Theorem 2.2, (ii) up to the supporting results proven in Sects. 4 and 5.

Proof of Theorem 2.2, (ii). The inequality (2.28) implies (3.1) which in turn implies the bound (3.6). Applying (2.27) with $(\mathbf{u}^\varepsilon, \theta^\varepsilon) = (\mathbf{u}^\varepsilon_S, \theta^\varepsilon_S)$ and $\theta^0 = \theta^{0,\varepsilon}_S$, noting $\theta^{0,\varepsilon}_S(0) = \theta^\varepsilon(0)$, and recalling the assumed bound (2.13) we infer (2.29). To prove (2.30), let C be as in (2.25). Since the Lipschitz norm, with metric $\tilde{\rho}_\eta$ of $\psi := \phi/C[\phi]_\eta$ is at most one, then, by (2.19)

$$|\mathbb{E}(\psi(\mathbf{u}_S^\varepsilon,\theta_S^\varepsilon)-\psi(L\theta_S^0))| = \left|\int_{\mathbf{H}}\psi(\mathbf{u},\theta)\mu_\varepsilon(\mathbf{u},\theta)-\int_H\psi(\mathbf{u},\theta)L(\mu_0)(\mathbf{u},\theta)\right| \\ \leq \tilde{\rho}_\eta(\mu_\varepsilon, L\mu_0)\,, \tag{3.7}$$

and the result follows from (2.29). The proof is complete. □

4. Contraction in the Wasserstein distance for the infinite Prandtl system

In this section we establish some properties of the infinite Prandtl system (2.4)–(2.5), which provide a sufficient condition for proving Theorem 2.2 (i) as a consequence of a general result in [27, Theorem 3.4]. These properties are summarized as follows:

Proposition 4.1. *There exist $\eta_0 > 0$ and N_0, depending only on $Ra, \tilde{Ra}$, such that for any $0 < \eta < \eta_0$, whenever the number of forced modes N exceeds N_0, we have*

(a) **Lyapunov structure:** *For all $t^* > 0$, there exists $C_1 = C_1(t^*, \eta)$ such that for each $\theta_0^0 \in H$ and every $t \in [0, t^*]$,*

$$\mathbb{E}\left(\exp(\eta\|\theta^0(t,\theta_0^0)\|^2)(1+\|\mathcal{J}_{0,t}\|)\right) \\ \leq C_1\exp(\eta(1+4Ra\tilde{Ra})e^{-t/2}\|\theta_0^0\|^2)\,, \tag{4.1}$$

where the operator $\mathcal{J}_{0,t}$ is the Fréchet derivative of $\theta^0(t,\theta_0)$ with respect to initial condition θ_0^0; see (B.1) and (B.8) below.

(b) **Gradient Bound for Markov semigroup:** *for any $\phi \in C_b^1(H)$, and every $t \geq 0$, $\theta \in H$*

$$\|\nabla P_t^0\phi(\theta)\| \leq C\exp(\eta\|\theta\|^2)\left(\sqrt{P_t^0(|\phi(\theta)|^2)}+\delta(t)\sqrt{P_t^0(\|\nabla\phi(\theta)\|^2)}\right), \tag{4.2}$$

where $\delta(t) \to 0$ as $t \to \infty$. Here again $\delta : [0,\infty) \to [0,\infty)$ and $C > 0$ depend only on $Ra, \tilde{Ra}$, and η.

(c) **Irreducibility condition:** *for any $M > 0$, $\epsilon > 0$ there is a $t_* = t_*(M,\epsilon,\eta)$ such that for each $t \geq t_*$*

$$\inf_{\|\theta_0\|,\|\tilde{\theta}_0\|\leq M}\ \sup_{\Gamma\in\mathcal{C}(\delta_{\theta_0}P_t^0,\delta_{\tilde{\theta}_0}P_t^0)} \Gamma\{(\theta,\tilde{\theta})\in H\times H : \rho_\eta(\theta,\tilde{\theta}) < \epsilon\} > 0, \tag{4.3}$$

where, as above in (2.19), $\mathcal{C}(\delta_{\theta_0}P_t^0, \delta_{\tilde{\theta}_0}P_t^0)$ denotes the collection of all couplings of the measures $\delta_{\theta_0}P_t^0$ and $\delta_{\tilde{\theta}_0}P_t^0$.

Proving the first item, (a), essentially reduces to establishing a moment bound which follows from estimates found in [21], and which we recall below in Appendix A (see Proposition A.2). The second condition, (4.2), can be translated to a control problem through the use of Malliavin calculus which in our setting amounts to proving a relatively straightforward Foias-Prodi type estimate. Once again (b) can be established by methods essentially contained in previous works and we relegate further details to Appendix B. As already mentioned above, the principal novel challenge here is to prove the irreducibility property (c) which we turn to next.

4.1. Irreducibility

In previous related works the proof of irreducibility essentially relies on the fact that the governing equations without stochastic forcing have a trivial attractor which is stable under small force perturbations; see e.g. [10,20,26,49]. In our present situation, (2.4)–(2.5), the dynamics without body forces can be highly non-trivial.[7] Our approach to proving (4.3) is based on a control argument and Foias-Prodi type estimates. With these estimates in place, by invoking the Girsanov theorem and stopping time arguments, (4.3) is established following previous proofs of support theorems for SPDEs with additive noise, see e.g. [11, Theorem 7.4.1]. For clarity and precision of exposition, we provide a self-contained proof of irreducibility.

As a preliminary step we show that (4.3) follows from a simpler bound.

Lemma 4.1. *For a given $N \geq 0$ consider (2.4)–(2.5) with N independently forced directions. If for every $M, \epsilon > 0$ there is a $t_* = t_*(M, \epsilon) > 0$ such that*

$$\inf_{\|\theta_0\| \leq M} \mathbb{P}(\|\theta^0(t, \theta_0)\| < \epsilon) > 0, \text{ for each } t \geq t_*, \tag{4.4}$$

then (4.3) holds for such an N and any $\eta > 0$.

Proof. For any $\theta_0, \tilde{\theta}_0 \in H$ consider the element $\tilde{\Gamma} \in \mathcal{C}(\delta_{\theta_0} P_t^0, \delta_{\tilde{\theta}_0} P_t^0)$ defined on cylindrical sets as

$$\tilde{\Gamma}(A \times B) = P_t(\theta_0, A) \times P_t(\tilde{\theta}_0, B), \quad A, B \in \mathcal{B}(H).$$

For each $t > 0$ and any $M, \eta, \gamma > 0$ one has

$$\begin{aligned}
&\inf_{\|\theta_0\|, \|\tilde{\theta}_0\| \leq M} \sup_{\Gamma \in \mathcal{C}(\delta_{\theta_0} P_t, \delta_{\tilde{\theta}_0} P_t)} \Gamma\{(\theta, \tilde{\theta}) \in H \times H : \rho_\eta(\theta, \tilde{\theta}) < \gamma\} \\
&\geq \inf_{\|\theta_0\|, \|\tilde{\theta}_0\| \leq M} \tilde{\Gamma} \left\{(\theta, \tilde{\theta}) \in B_1 \times B_1 : \|\theta\| + \|\tilde{\theta}\| < \gamma \exp(-4\eta)\right\} \\
&\geq \left(\inf_{\|\theta_0\| \leq M} P_t \left(\theta_0, \left\{\theta \in H : \|\theta\| < \min\{\gamma/2 \cdot \exp(-4\eta), 1\}\right\}\right)\right)^2 \\
&= \left(\inf_{\|\theta_0\| \leq M} \mathbb{P}(\|\theta(t, \theta_0)\| < \min\{\gamma/2 \cdot \exp(-4\eta), 1\})\right)^2,
\end{aligned}$$

where we have used (2.23) in the first inequality. Applying (4.4) with $\epsilon = \min\{\gamma/2 \cdot \exp(-4\eta), 1\})$ and the given $M > 0$ yields the desired result. □

In order to establish (4.3) the rest of the section is therefore devoted to

Proposition 4.2. *There exists $N_0 = N_0(Ra, \tilde{Ra})$ sufficiently large (cf. (4.9)) such that, for any $N \geq N_0$ and every $M, \epsilon > 0$, there is a $t_* = t_*(M, \epsilon) > 0$ such that (4.4) is satisfied.*

[7] Note that the geometric control methods developed in [1,2], and in [23] for the Boussinesq system, would be difficult to apply, as these methods seemingly require a detailed understanding of the wave-number interactions in (2.4)–(2.5).

Proof of Proposition 4.2. We first establish the analogue of (4.4) for the modified system

$$-\Delta \bar{\mathbf{u}} = \nabla \bar{p} + Ra\hat{\mathbf{k}}\bar{\theta}\,, \quad \nabla\cdot\bar{\mathbf{u}} = 0, \tag{4.5}$$

$$d\bar{\theta} + \bar{\mathbf{u}}\cdot\nabla\bar{\theta}dt = (\tilde{Ra}\cdot\bar{u}_3 + \Delta\bar{\theta} - \lambda_N P_N\bar{\theta})dt + \sum_{k=1}^{N}\sigma_k dW^k,$$

$$\bar{\theta}(0) = \theta_0\,, \tag{4.6}$$

when N is sufficiently large.[8] As in (2.1)–(2.2) we supplement (4.5)–(4.6) with the homogeneous boundary conditions (2.3). Denote $\psi := \bar{\theta} - \sum_{k=1}^{N}\sigma_k W^k = \bar{\theta} - \sigma W$ which satisfies

$$\partial_t\psi + \bar{\mathbf{u}}\cdot\nabla\psi = \tilde{Ra}\cdot\bar{u}_3 + \Delta\psi - \lambda_N P_N\psi + (\Delta\sigma W - \lambda_N P_N\sigma W - \bar{\mathbf{u}}\cdot\nabla(\sigma W)),$$
$$\psi(0) = \theta_0.$$

Taking an inner product with ψ, using that $\bar{\mathbf{u}}$ is divergence free, the inverse Poincaré inequality [see (B.14) below] and the bound

$$\|\nabla\bar{\mathbf{u}}\| \leq Ra\|\bar{\theta}\| \leq Ra(\|\psi\| + \|\sigma W\|) \tag{4.7}$$

which follows from (4.5) we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\psi\|^2 + \lambda_N\|\psi\|^2 &\leq (\tilde{Ra}\|\bar{\mathbf{u}}\| + \|\Delta\sigma W\| + \lambda_N\|\sigma W\| + \|\bar{\mathbf{u}}\|\|\nabla\sigma W\|_{L^\infty})\|\psi\|\\ &\leq C(\tilde{Ra}Ra(\|\psi\| + \|\sigma W\|) + \|\Delta\sigma W\| + \lambda_N\|\sigma W\|\\ &\quad + Ra(\|\psi\| + \|\sigma W\|)\|\nabla\sigma W\|_{L^\infty})\|\psi\|.\end{aligned}$$

For any $t > 0$, let $\xi_t = \sup\{s \in [0,t] : \|\psi(s)\| = 0\}$ with the convention that the supremum of the empty set is zero. Thus, for any $t > 0$, on the interval $[\xi_t, t]$ it follows that

$$\begin{aligned}&\frac{d}{dt}\|\psi\| + (\lambda_N - CRa(\tilde{Ra} + \|\nabla\sigma W\|_{L^\infty}))\|\psi\|\\ &\quad \leq C\left((Ra\tilde{Ra} + \lambda_N + Ra\|\nabla\sigma W\|_{L^\infty})\|\sigma W\| + \|\Delta\sigma W\|\right).\end{aligned} \tag{4.8}$$

Next, we use the fact that, with positive probability, each of $\|\sigma W\|, \|\nabla\sigma W\|$, $\|\Delta\sigma W\|$ stays close to zero over finite time intervals. For $\gamma > 0$, $t > 0$, $N > 0$ consider the sets

$$\begin{aligned}\mathcal{X}_{\gamma,t,N} := \Big\{ &\sup_{s\in[0,t]}\|\nabla\sigma W\|_{L^\infty} \leq 1, \sup_{s\in[0,t]}\|\Delta\sigma W\| \leq \frac{\gamma}{2C},\\ &\sup_{s\in[0,t]}\|\sigma W\| \leq \gamma\left(\frac{1}{2C(Ra\tilde{Ra} + \lambda_N + Ra)} \wedge 1\right)\Big\}.\end{aligned}$$

Since σ is spatially smooth we infer from standard properties of Brownian motion that $\mathbb{P}(\mathcal{X}_{\gamma,t,N}) > 0$ for any $\gamma > 0$, $t > 0$, and $N > 0$. On the other

[8]Here recall that P_N denotes the projection onto the first N modes of $-\Delta$ [with boundary conditions as in (2.3)] and λ_N is the corresponding largest eigenvalue in this collection.

hand, on $\mathcal{X}_{\gamma,t,N}$ the differential inequality

$$\frac{d}{dt}\|\psi\| + (\lambda_N - CRa(\tilde{R}a + 1))\|\psi\| \leq \gamma$$

holds over the interval $[\xi_t, t]$.

Hence, fixing N_0 sufficiently large, we have for any $N \geq N_0$,

$$\lambda_N \geq \max\{2CRa(\tilde{R}a + 1), 1\}, \tag{4.9}$$

and we infer that on $\mathcal{X}_{\gamma,t,N}$,

$$\|\bar{\theta}(t, \theta_0)\| \leq \|\psi(t)\| + \|\sigma W\| \leq 2\gamma + e^{-\lambda_N t/2}\|\theta_0\|,$$

where note that $\|\psi(t)\| = 0$ on the set where $\xi_t > 0$. Therefore, for a given $M > 0$, $\epsilon > 0$, by choosing $\gamma = \epsilon/4$ and $t_* = t_*(M, \epsilon)$ such that $e^{-\lambda_N t_*}M \leq \frac{\epsilon}{2}$, we have for any $t \geq t_*$

$$\inf_{\|\theta_0\| \leq M} \mathbb{P}(\|\bar{\theta}(t, \theta_0)\| < \epsilon) \geq \mathbb{P}(\mathcal{X}_{\epsilon/4,t,N}) > 0\,. \tag{4.10}$$

In order to now infer (4.4) from (4.10) we apply the Girsanov theorem and make further bounds to a slightly modified version of (4.5)–(4.6). For $K > 0$ and $\theta_0 \in H$ define $\tilde{\theta}_K = \tilde{\theta}_K(\cdot, \theta_0)$ as the solution of (4.5)–(4.6) with the term $-\lambda_N P_N \bar{\theta}$ replaced with $-\lambda_N P_N \tilde{\theta}_K \chi_K(\|P_N \tilde{\theta}_K\|)$. Here χ_K is a smooth, non-negative cut-off function with $\chi_K \equiv 1$ for $|x| \leq K$ and $\chi_K \equiv 0$ for $|x| \geq K+1$. Consider the stopping times

$$\tau_K(\theta_0) = \inf_{s \geq 0}\left\{\|P_N \tilde{\theta}_K(s, \theta_0)\| \geq K\right\},$$

for any $K > 0$ and any $\theta_0 \in H$. It is not hard to see that for any $K > 0$ and any $\theta_0 \in H$

$$\mathbb{P}\left(\bar{\theta}(t \wedge \tau_K(\theta_0), \theta_0) = \tilde{\theta}_K(t \wedge \tau_K(\theta_0), \theta_0), \text{ for every } t \geq 0\right) = 1. \tag{4.11}$$

On the other hand, for any $\theta_0 \in H$ and $K > 0$, the law of $\tilde{\theta}_K(\cdot, \theta_0)$ is absolutely continuous with respect to the law of the processes $\theta^0(\cdot, \theta_0)$ solving (2.4)–(2.5). Indeed, for $\theta_0 \in H$ and $K > 0$ define

$$\mathcal{M}_{\theta_0,K}(t) = \exp\left(-\int_0^t \alpha_{\theta_0,K}\,dW - \frac{1}{2}\int_0^t |\alpha_{\theta_0,K}|^2 ds\right), \tag{4.12}$$

where

$$\alpha_{\theta_0,K}(s) = -\lambda_N \sigma^{-1} P_N \tilde{\theta}_K(s, \theta_0)\chi_K(\|P_N \tilde{\theta}_K(s, \theta_0)\|)$$

and take

$$d\mathbb{Q}_{\theta_0,K,t} := \mathcal{M}_{\theta_0,K}(t)\,d\mathbb{P}\,.$$

Notice that, since $|\sigma^{-1} P_N \tilde{\theta}_K(s, \theta_0)\chi_K(\|P_N \tilde{\theta}_K\|)| \leq \|\sigma^{-1}\| \cdot (K+1)$, the Novikov condition is satisfied and for any $\epsilon > 0$, $t \geq 0$, $K > 0$, and $\theta \in H$, the Girsanov theorem yields

$$\mathbb{P}(\|\theta(t, \theta_0)\| < \epsilon) = \mathbb{Q}_{\theta_0,K,t}(\|\tilde{\theta}_K(t, \theta_0)\| < \epsilon) = \mathbb{E}\left(\mathcal{M}_{\theta_0,K}(t)\mathbb{1}_{\|\tilde{\theta}_K(t,\theta_0)\| < \epsilon}\right).$$

Hence, for any $\epsilon > 0$, $\theta_0 \in H$, and for any $\beth, K, t > 0$, the Markov inequality implies

$$\mathbb{P}(\|\theta(t,\theta_0)\| < \epsilon) \geq \beth\mathbb{P}\left(\|\tilde{\theta}_K(t,\theta_0)\| < \epsilon, \mathcal{M}_{\theta_0,K}(t) \geq \beth\right)$$
$$\geq \beth\mathbb{P}\left(\|\bar{\theta}(t,\theta_0)\| < \epsilon, \mathcal{M}_{\theta_0,K}(t) \geq \beth, \tau_K(\theta_0) > t\right),$$

where we used (4.11) for the final inequality. On the other hand

$$\mathbb{P}\left(\|\bar{\theta}(t,\theta_0)\| < \epsilon\right) \leq \mathbb{P}\left(\|\bar{\theta}(t,\theta_0)\| < \epsilon, \mathcal{M}_{\theta_0,K}(t) \geq \beth\right) + \mathbb{P}\left(\mathcal{M}_{\theta_0,K}(t) < \beth\right)$$
$$\leq \mathbb{P}\left(\|\bar{\theta}(t,\theta_0)\| < \epsilon, \mathcal{M}_{\theta_0,K}(t) \geq \beth, \tau_K(\theta_0) > t\right)$$
$$+ \mathbb{P}\left(\mathcal{M}_{\theta_0,K}(t) < \beth\right) + \mathbb{P}(\tau_K(\theta_0) < t).$$

These two bounds yield

$$\begin{aligned}\frac{1}{\beth}&\inf_{\|\theta_0\|\leq M}\mathbb{P}(\|\theta(t,\theta_0)\| < \epsilon)\\ &\geq \inf_{\|\theta_0\|\leq M}\mathbb{P}\left(\|\bar{\theta}(t,\theta_0)\| < \epsilon\right)\\ &\quad - \sup_{\|\theta_0\|\leq M}\left(\mathbb{P}\left(\mathcal{M}_{\theta_0,K}(t) < \beth\right) + \mathbb{P}(\tau_K(\theta_0) < t)\right),\end{aligned} \tag{4.13}$$

for any $M, \epsilon, t > 0$ and for any $K, \beth > 0$.

Since the first term on the the right-hand side of (4.13) is independent of $K > 0$ and $\beth > 0$, we finish the argument by showing that for every fixed $M, K, t > 0$

$$\sup_{\|\theta_0\|\leq M}\mathbb{P}\left(\mathcal{M}_{\theta_0,K}(t) < \beth\right) \to 0, \quad \text{as } \beth \to 0, \tag{4.14}$$

and for every given $M, t > 0$

$$\sup_{\|\theta_0\|\leq M}\mathbb{P}(\tau_K(\theta_0) < t) \to 0, \quad \text{as } K \to \infty. \tag{4.15}$$

For the first bound (4.14), we have from (4.12) and Itō isometry

$$\begin{aligned}\mathbb{P}\left(\mathcal{M}_{\theta_0,K}(t) < \beth\right) &= \mathbb{P}\left(\int_0^t \alpha_{\theta_0,K}dW + \frac{1}{2}\int_0^t |\alpha_{\theta_0,K}|^2 ds > \log(\beth^{-1})\right)\\ &\leq \frac{1}{\log(\beth^{-1})}\mathbb{E}\left(\left|\int_0^t \alpha_{\theta_0,K}dW\right| + \frac{1}{2}\int_0^t |\alpha_{\theta_0,K}|^2 ds\right)\\ &\leq \frac{2}{\log(\beth^{-1})}\mathbb{E}\left(1 + \lambda_N^2\|\sigma^{-1}P_N\|^2\int_0^t \|P_N\tilde{\theta}(t,\theta_0)\|^2\chi_K(\|P_N\tilde{\theta}_K(t,\theta_0)\|)ds\right),\\ &\leq \frac{2\left(1 + \lambda_N^2\|\sigma^{-1}P_N\|^2(K+1)^2 t\right)}{\log(\beth^{-1})},\end{aligned}$$

valid for any $\beth \in (0,1)$, $K > 0$, and any $\theta_0 \in H$. For the second bound, (4.15) observe that, in view of (4.11),

$$\mathbb{P}(\tau_K(\theta_0) < t) \leq \mathbb{P}\left(\sup_{s\in[0,t]} \|P_N\bar{\theta}(s,\theta_0)\| \geq K\right) \leq \frac{1}{K^2}\mathbb{E}\left(\sup_{s\in[0,t]} \|\bar{\theta}(s,\theta_0)\|^2\right). \tag{4.16}$$

From the Itō formula, and (1.4), it follows that

$$d\|\bar{\theta}\|^2 + 2\lambda_N\|P_N\bar{\theta}\|^2dt + 2\|\nabla\bar{\theta}\|^2dt = \left(2\tilde{Ra}\langle\tilde{u}_3,\bar{\theta}\rangle + 1\right)dt + \langle\sigma,\bar{\theta}\rangle dW.$$

Integrating in time and using (4.5), inverse Poincaré inequality [see (B.14)], and (4.7) we infer for any $s \geq 0$

$$\|\bar{\theta}(s)\|^2+2\lambda_N\int_0^s \|\bar{\theta}\|^2dr \leq \|\theta_0\|^2 + 2Ra\tilde{Ra}\int_0^s \|\bar{\theta}\|^2dr + s + 2\sup_{r\in[0,s]}\left|\int_0^r\langle\sigma,\bar{\theta}\rangle dW\right|.$$

Using the assumption (4.9) and the Birkholder–Davis–Gundy inequality we infer

$$\mathbb{E}\left(\sup_{s\in[0,t]} \|\bar{\theta}(s,\theta_0)\|^2\right) \leq \|\theta_0\|^2 + 17t. \tag{4.17}$$

Combining (4.16) and (4.17) thus yields the second bound (4.15).

Using (4.10), (4.13), (4.14), and (4.15), we conclude the proof as follows. Given any $\epsilon > 0$ and any $M > 0$, and with λ_N given as in (4.9), choose t_* as in (4.10), that is, such that $e^{-\lambda_N t_*}M \leq \frac{\epsilon}{2}$. Fix any $t \geq t_*$ and by (4.10) we have $a = a(M,\epsilon,t) := \inf_{\|\theta_0\|\leq M}\mathbb{P}(\|\bar{\theta}(t,\theta_0)\| < \epsilon) > 0$. Now, by (4.15), we can pick K sufficiently large so that $\sup_{\|\theta_0\|\leq M}\mathbb{P}(\tau_K(\theta_0) < t) \leq a/4$, and with K, M, t fixed, we can by (4.14) choose $\daleth > 0$ small enough so that $\sup_{\|\theta_0\|\leq M}\mathbb{P}\left(\mathcal{M}_{K,\theta_0}(t) < \daleth\right) \leq a/4$. Finally, by combining these choices with (4.13) we obtain that

$$\inf_{\|\theta_0\|\leq M}\mathbb{P}(\|\theta(t,\theta_0)\| < \epsilon) \geq \frac{\daleth a}{2} > 0.$$

The proof of Proposition 4.2 is thus complete. □

5. Finite time asymptotics

In this section we prove Theorem 2.1. We also derive a 'corrector' system which we show approximates the velocity component of the full system (2.1)–(2.2) up to the initial time (see Theorem 5.1 below).

5.1. Preliminaries: the Stokes operator

Before proceeding further we recall (see e.g. [9,40]) some properties of solutions of the Stokes equation

$$-\Delta\mathbf{u} = \nabla p + \mathbf{f}, \quad \nabla\cdot\mathbf{u} = 0, \tag{5.1}$$

supplemented with the mixed periodic-Dirichlet boundary conditions as in (2.3). We can express (5.1) more abstractly as $A\mathbf{u} = P\mathbf{f}$, where $A = -P\Delta$ is the Stokes operator. Here, P is the Leray projection on divergence free vector fields $P : (L^2(\mathcal{D}))^3 \to H_1$ with H_1, the space of L^2 divergence free vector fields, defined in Sect. 2.1.[9] As in the classical elliptic theory we have that for any $\mathbf{f} \in (L^2(\mathcal{D}))^3$, there exists a unique $\mathbf{u} \in D(A) = V_1 \cap (H^2(\mathcal{D}))^3$ which satisfies

$$\|\mathbf{u}\|_{H^2} \leq C\|\mathbf{f}\|, \tag{5.2}$$

where C is independent of $\mathbf{f}$. In what follows we frequently denote $\mathbf{u} = A^{-1}P\mathbf{f}$.

Since A is a positive, self-adjoint operator which is unbounded on the space H_1 with a compact inverse, by Hilbert's theorem there is a complete orthonormal basis of eigenfunctions $\{\mathbf{e}_k\}_{k\geq 1}$ of A with the associated non-decreasing sequence of eigenvalues λ_k^* diverging to infinity. Take

$$\mathcal{P}_N \text{ to be the projection onto the subspace } H_N := \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}. \tag{5.3}$$

Here the regularity theory as found in, say [9,40], show that each $\mathbf{e}_k$ is smooth and hence in particular $H_N \subset V$.

We also consider the associated linear evolution given as

$$\partial_t \mathbf{u} - \mu\Delta\mathbf{u} = \nabla p + \mathbf{f}, \quad \nabla \cdot \mathbf{u} = 0, \quad \mathbf{u}(0) = \mathbf{u}_0, \tag{5.4}$$

for any parameter $\mu > 0$ and relative to the (sufficiently regular) data $\mathbf{f}, \mathbf{u}_0$ supplemented with the boundary conditions (2.3). Here, for any $\mathbf{f} \in L^2_{loc}([0,\infty); H_1)$ and $\mathbf{u}_0 \in H_1$ there exists a unique solution $\mathbf{u}$ of (5.4) with $\mathbf{u} \in L^2_{loc}([0,\infty); V_1) \cap C([0,\infty); H_1)$. Moreover, A is the generator of an analytic semigroup which we denote as $\{\exp(-\mu At)\}_{t\geq 0}$.

5.2. Finite time convergence estimates

We next turn to the proof of Theorem 2.1:

Proof of Theorem 2.1. Take $\phi^\varepsilon = \theta^\varepsilon - \theta^0$ and $\mathbf{v}^\varepsilon = \mathbf{u}^\varepsilon - \mathbf{u}^0$ with $\mathbf{u}^0 = M(\theta^0)$, where M is defined by (2.9). Referring to (2.1)–(2.3) and (2.4)–(2.5) we see that ϕ^ε satisfies

$$\partial_t \phi^\varepsilon - \Delta\phi^\varepsilon = \tilde{Ra} \cdot v_3^\varepsilon - \mathbf{v}^\varepsilon \cdot \nabla\theta^0 - \mathbf{u}^\varepsilon \cdot \nabla\phi^\varepsilon, \qquad \phi^\varepsilon(0) = \theta^\varepsilon(0) - \theta^0(0) := \phi_0^\varepsilon .$$

Therefore, taking an L^2 inner product with ϕ^ε and using that $\nabla \cdot \mathbf{v}^\varepsilon = 0$ we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\phi^\varepsilon\|^2 + \|\nabla\phi^\varepsilon\|^2 &= \int (\tilde{Ra} \cdot v_3^\varepsilon - \mathbf{v}^\varepsilon \cdot \nabla\theta^0)\phi^\varepsilon dx \\ &\leq \tilde{Ra}\|\mathbf{v}^\varepsilon\|\|\phi^\varepsilon\| + \|\mathbf{v}^\varepsilon\|_{L^6}\|\nabla\phi^\varepsilon\|\|\theta^0\|_{L^3} .\end{aligned}$$

Hence from standard Sobolev embeddings, Young's inequality, and the Poincaré inequality we obtain

$$\frac{d}{dt}\|\phi^\varepsilon\|^2 \leq C\left(\|\theta^0\|_{L^3}^2 + \tilde{Ra}^2\right)\|\nabla\mathbf{v}^\varepsilon\|^2 .$$

[9] Equivalently $A\mathbf{u} = -\Delta\mathbf{u} - \nabla p$, where $p = p(\mathbf{u})$ the 'pressure' is the unique H^1 function satisfying $\Delta p = -\operatorname{div}(\Delta u)$ in the weak sense.

Integrating in time we infer that

$$\sup_{s\in[0,t]} \|\phi^\varepsilon(s\wedge\tau)\|^2 \\ \leq \|\phi_0^\varepsilon\|^2 + \sup_{s\in[0,t\wedge\tau]}\left(\|\theta^0(s)\|_{L^3}^2 + \tilde{Ra}^2\right)\int_0^{t\wedge\tau}\|\nabla\mathbf{v}^\varepsilon(t')\|^2 dt' \tag{5.5}$$

for any $t > 0$ and any stopping time $\tau \geq 0$.

We now turn to derive an evolution equation for $\mathbf{v}^\varepsilon$. Recalling that $\mathbf{u}^\varepsilon$ and $\mathbf{u}^0$ satisfy respectively, (2.1) and (2.4) we find

$$\varepsilon(\partial_t\mathbf{u}^\varepsilon + \mathbf{u}^\varepsilon\cdot\nabla\mathbf{u}^\varepsilon) - \Delta\mathbf{v}^\varepsilon = \nabla q^\varepsilon + Ra\hat{\mathbf{k}}\phi^\varepsilon, \tag{5.6}$$

where $q^\varepsilon = \tilde{p}^\varepsilon - \tilde{p}$ is the difference in the pressures. On the other hand, recalling that $\mathbf{u}^0 = RaA^{-1}(P(\hat{\mathbf{k}}\theta^0))$, we have

$$d\mathbf{u}^0 = -RaA^{-1}P\Big(\hat{\mathbf{k}}\Big(\mathbf{u}^0\cdot\nabla\theta^0 - \Delta\theta^0 - \tilde{Ra}\cdot u_3^0\Big)\Big)dt \\ + Ra\sum_{k=1}^N A^{-1}P(\hat{\mathbf{k}}\sigma_k)dW^k\,. \tag{5.7}$$

Multiplying (5.6) by ε^{-1}, subtracting the resulting system from (5.7) and rearranging we obtain

$$d\mathbf{v}^\varepsilon - \frac{1}{\varepsilon}\Delta\mathbf{v}^\varepsilon dt = \frac{1}{\varepsilon}\left(\nabla q^\varepsilon + Ra\hat{\mathbf{k}}\phi^\varepsilon\right)dt \\ +\Big(RaA^{-1}P\Big(\hat{\mathbf{k}}\Big(\mathbf{u}^0\cdot\nabla\theta^0 - \Delta\theta^0 + \tilde{Ra}\cdot u_3^0\Big)\Big) - \mathbf{u}^\varepsilon\cdot\nabla\mathbf{u}^\varepsilon\Big)dt \\ - Ra\sum_{k=1}^N A^{-1}P(\hat{\mathbf{k}}\sigma_k)dW^k\,, \tag{5.8}$$

with $\nabla\cdot\mathbf{v}^\varepsilon = 0$.

Using (5.8) we estimate $\mathbf{v}^\varepsilon$ as follows. The Itō formula and (5.9) yields

$$d\|\mathbf{v}^\varepsilon\|^2 + \frac{2}{\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 dt \\ =\frac{2}{\varepsilon}Ra\langle\phi^\varepsilon, v_3^\varepsilon\rangle dt \\ + 2\left\langle RaA^{-1}P\Big(\hat{\mathbf{k}}\left(\mathbf{u}^0\cdot\nabla\theta^0 - \Delta\theta^0 - \tilde{Ra}\cdot u_3^0\right)\Big) - \mathbf{u}^\varepsilon\cdot\nabla\mathbf{u}^\varepsilon, \mathbf{v}^\varepsilon\right\rangle dt \\ + Ra^2\sum_{k=1}^N |A^{-1}P(\hat{\mathbf{k}}\sigma_k)|^2 dt - 2Ra\sum_{k=1}^N\langle A^{-1}P(\hat{\mathbf{k}}\sigma_k), \mathbf{v}^\varepsilon\rangle dW^k \\ :=(T_1 + T_2 + T_3 + T_4 + T_5 + T_6)dt + SdW. \tag{5.9}$$

With the Young and Poincaré inequalities we have

$$|T_1| \leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + \frac{4Ra^2}{\varepsilon}\|\phi^\varepsilon\|^2. \tag{5.10}$$

For T_2 we use that A^{-1} is self-adjoint on H, $D(A) \subset H$ and that $\mathbf{u}^0, \mathbf{v}^\varepsilon$ are divergence free to obtain

$$|T_2| = 2Ra \left| \int \mathbf{u}^0 \cdot \nabla \theta^0 (A^{-1}\mathbf{v}^\varepsilon)_3 dx \right| = 2Ra \left| \int \mathbf{u}^0 \cdot \nabla (A^{-1}\mathbf{v}^\varepsilon)_3 \, \theta^0 dx \right| ,$$

where $(A^{-1}\mathbf{v}^\varepsilon)_3$ represents the third component of the vector field $A^{-1}\mathbf{v}^\varepsilon$. Hence (5.2) and the imbedding $H^2 \hookrightarrow L^\infty$ imply

$$\begin{aligned} |T_2| &\leq 2Ra\|\mathbf{u}^0\|\|\theta^0\|\|\nabla(A^{-1}\mathbf{v}^\varepsilon)\|_{L^\infty} \leq CRa^2\|\theta^0\|^2\|\nabla\mathbf{v}^\varepsilon\| \\ &\leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + \varepsilon CRa^4\|\theta^0\|^4. \end{aligned} \tag{5.11}$$

For the terms T_3 and T_4 we use the regularity of the Stokes operator to obtain

$$\begin{aligned} |T_3| + |T_4| &\leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + 4\varepsilon Ra^2(\|\theta^0\|^2 + \tilde{Ra}^2\|\mathbf{u}^0\|^2) \\ &\leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + \varepsilon Ra^2(\tilde{Ra}^2 Ra^2 + 1)C\|\theta^0\|^2 . \end{aligned} \tag{5.12}$$

The most delicate term is T_5. Here we take advantage of an additional cancellation to obtain extra regularity. Since $\mathbf{u}^\varepsilon = \mathbf{v}^\varepsilon + \mathbf{u}^0$ we find

$$\begin{aligned} |T_5| = 2|\langle \mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^0, \mathbf{v}^\varepsilon\rangle| &= 2|\langle \mathbf{u}^\varepsilon \cdot \nabla \mathbf{v}^\varepsilon, \mathbf{u}^0\rangle| \leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + 4\varepsilon\|\mathbf{u}^0\|^2_{L^\infty}\|\mathbf{u}^\varepsilon\|^2 \\ &\leq \frac{1}{4\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 + \varepsilon CRa^2(\|\theta^0\|^4 + \|\mathbf{u}^\varepsilon\|^4), \end{aligned} \tag{5.13}$$

where we used the imbedding $H^2 \hookrightarrow L^\infty$ and (5.2) for the final bound. Finally we observe $|T_6| \leq CRa^2$.

Combining the bounds (5.10)–(5.13) and rearranging in (5.9) we find

$$\begin{aligned} d\|\mathbf{v}^\varepsilon\|^2 &+ \frac{1}{\varepsilon}\|\nabla\mathbf{v}^\varepsilon\|^2 dt \\ &\leq \frac{4Ra^2}{\varepsilon}\|\phi^\varepsilon\|^2 dt + \varepsilon C(1 + Ra^4)(1 + \tilde{Ra}^2)(\|\theta^0\|^4 + \|\mathbf{u}^\varepsilon\|^4 + 1)dt \\ &\quad + CRa^2 dt - 2Ra\sum_{k=1}^N \langle A^{-1}P(\hat{\mathbf{k}}\sigma_k), \mathbf{v}^\varepsilon\rangle dW^k , \end{aligned} \tag{5.14}$$

where the constant $C > 0$ is independent of $Ra, \tilde{Ra}$, and $\varepsilon > 0$. Consequently, for any $t \geq 0$ and any stopping time τ, we have

$$\begin{aligned} \int_0^{t\wedge\tau} \|\nabla\mathbf{v}^\varepsilon\|^2 dt' &\leq \varepsilon\|\mathbf{v}^\varepsilon(0)\|^2 + 4Ra^2\int_0^{t\wedge\tau}(\|\phi^\varepsilon\|^2 + \varepsilon C)dt' \\ &\quad + \varepsilon^2 C(1 + \mathrm{Ra}^4)(1 + \tilde{Ra}^2)\int_0^{t\wedge\tau}\left(\|\theta^0\|^4 + \|\mathbf{u}^\varepsilon\|^4 + 1\right) dt' \\ &\quad - \varepsilon Ra\sum_{k=1}^N \int_0^{t\wedge\tau} \langle A^{-1}P(\hat{\mathbf{k}}\sigma_k), \mathbf{v}^\varepsilon\rangle dW^k, \end{aligned} \tag{5.15}$$

where C is independent of $\varepsilon > 0$, $Ra, \tilde{Ra}$, and τ.

Next for any $\kappa > 0$ define the stopping times

$$\tau_\kappa := \inf_{t \geq 0} \left\{ \|\theta^0(t)\|_{L^3}^2 \geq \kappa \right\} . \tag{5.16}$$

From this definition and the bounds (5.5), (5.15) we now infer

$$\begin{aligned}
&\mathbb{E} \sup_{s\in[0,t]} \|\phi^\varepsilon(s\wedge\tau_\kappa)\|^2 \\
&\quad \leq \mathbb{E}\|\phi_0^\varepsilon\|^2 + 4Ra^2\left(\kappa + \tilde{Ra}^2\right)\int_0^t \mathbb{E}\left(\sup_{s\in[0,t']}\|\phi^\varepsilon(s\wedge\tau_\kappa)\|^2\right)dt' \\
&\qquad + \varepsilon\left(\kappa+\tilde{Ra}^2\right)\left(\mathbb{E}\|\mathbf{v}^\varepsilon(0)\|^2 + Ra^2Ct\right) \\
&\qquad + \varepsilon^2 C(\kappa + \tilde{Ra}^4 + 1)(\mathrm{Ra}^4+1)\int_0^t \mathbb{E}\left(\|\theta^0\|^4 + \|\mathbf{u}^\varepsilon\|^4 + 1\right)dt',
\end{aligned}$$

which implies with the Gronwall inequality that

$$\begin{aligned}
&\mathbb{E} \sup_{s\in[0,t]} \|\phi^\varepsilon(s\wedge\tau_\kappa)\|^2 \\
&\quad \leq \exp\left(C(Ra^4+1)\left(\kappa+\tilde{Ra}^4+1\right)(t+1)\right)\left(\varepsilon\mathcal{M}_\varepsilon(t) + \mathbb{E}\|\phi_0^\varepsilon\|^2\right),
\end{aligned} \tag{5.17}$$

where

$$\mathcal{M}_\varepsilon(t) := \mathbb{E}\|\mathbf{v}^\varepsilon(0)\|^2 + \int_0^t \left[\varepsilon\mathbb{E}\left(\|\mathbf{u}^\varepsilon\|^4 + \|\theta^0\|^4\right) + 1\right]dt',$$

and the constant C is independent of $\kappa, \varepsilon, Ra, \tilde{Ra}$, and t. By (2.26) and our standing assumption that (2.12) holds, we observe that $\mathcal{M}_\varepsilon$ is bounded independently of $\varepsilon > 0$ and we obtain

$$\begin{aligned}
\mathbb{E}\left(\sup_{s\in[0,t]}\|\phi^\varepsilon(s)\|^2 \mathbb{1}_{\tau_\kappa > t}\right) &\leq \mathbb{E}\sup_{s\in[0,t]}\|\phi^\varepsilon(s\wedge\tau_\kappa)\|^2 \\
&\leq C_1(\varepsilon + \mathbb{E}\|\phi_0^\varepsilon\|^2)\exp\left(C_1\kappa\right),
\end{aligned} \tag{5.18}$$

where the constant $C_1 = C(Ra, \tilde{Ra}, t)$ is independent of $\varepsilon > 0$ and $\kappa > 0$.

Set $X_\varepsilon(t) := \sup_{s\in[0,t]}\|\phi^\varepsilon(s)\|^2$ and for each $t \geq 0$, $\kappa > 0$, $\varepsilon > 0$ define the sets

$$E_{t,\kappa,\varepsilon} := \left\{\sup_{s\in[0,t]}\|\theta^0(s)\|_{L^3}^2 \geq \kappa\right\} = \{\tau_\kappa \leq t\}.$$

For each $t > 0$ one finds by the Markov inequality, Proposition A.1, and the assumption (2.26) that for sufficiently small $\eta_1 = \eta_1(\tilde{Ra}, \eta) > 0$,

$$\mathbb{P}(E_{t,\kappa,\varepsilon}) \leq e^{-\eta\kappa}\mathbb{E}\exp\left(\eta_1 \sup_{s\in[0,t]}\|\theta^0(s)\|_{L^3}^2\right) \leq C_2 e^{-\eta\kappa}, \tag{5.19}$$

where $C_2 = C_2(C_0, Ra, \tilde{Ra}, \|\sigma\|_{L^3}, \eta, t) > 0$ is independent of $\varepsilon > 0$ and $\kappa > 0$. On the other hand, for any $\gamma \in (0,1)$ we have

$$\begin{aligned}\mathbb{E}X_\varepsilon(t)^\gamma &= \sum_{k=0}^{\infty} \mathbb{E}\left(X_\varepsilon(t)^\gamma \mathbb{1}_{k \le \left(\sup_{s\in[0,t]} \|\theta^0(s)\|^2_{L^3}\right) < k+1} \right) \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left(X_\varepsilon(t)^\gamma \mathbb{1}_{\tau_k \le t} \mathbb{1}_{\tau_{k+1} > t} \right) \\ &\le \sum_{k=0}^{\infty} \left(\mathbb{E}(X_\varepsilon(t) \mathbb{1}_{\tau_{k+1} > t}) \right)^\gamma \left(\mathbb{P}(E_{t,\kappa,\varepsilon}) \right)^{1-\gamma} \\ &\le C(\varepsilon + \mathbb{E}\|\phi_0^\varepsilon\|^2)^\gamma \sum_{k=0}^{\infty} \exp(\gamma C_1(k+1) - (1-\gamma)\eta k), \end{aligned} \tag{5.20}$$

where we have used (5.18) and (5.19) for the final bound. Here, $C = C(C_0, Ra, \tilde{Ra}, \|\sigma\|_{L^3}, \eta, t)$ is independent of $\varepsilon > 0$, $\mathbb{E}\|\phi_0^\varepsilon\|^2$ and C_1 is the constant appearing in (5.18). Thus when $\gamma < \frac{\eta}{C_1+\eta}$ the series in (5.20) converges. Then for any $p > 0$ and any $\gamma < (\frac{\eta}{C_1+\eta}) \wedge p$ we find

$$\begin{aligned} &\mathbb{E} \sup_{s\in[0,t]} \|\phi^\varepsilon(s)\|^p \\ &\quad \le C \left(\mathbb{E} \sup_{s\in[0,t]} (\|\theta^\varepsilon(s)\|^{2(p-\gamma)} + \|\theta^0(s)\|^{2(p-\gamma)}) \right)^{1/2} \left(\mathbb{E} \sup_{s\in[0,t]} \|\phi^\varepsilon(s)\|^{2\gamma} \right)^{1/2}. \end{aligned}$$

Combing this bound with (2.26) and (5.20) we now obtain the first part of (2.27).

To address the second term in (2.27) we return to (5.15). Taking expected values we obtain

$$\mathbb{E} \int_0^t \|\nabla \mathbf{v}^\varepsilon\|^2 dt \le \varepsilon(\mathbb{E}\|\mathbf{v}^\varepsilon(0)\|^2 + C) + C\mathbb{E}\left(\sup_{s\in[0,t]} \|\phi^\varepsilon(s)\|^2 \right),$$

where $C = C(Ra, \tilde{Ra}, t, C_0)$ is independent of $\varepsilon > 0$. Combining this observation with the previous bound, the proof of Theorem 2.1 is now complete. □

5.3. Approximation up to initial conditions: the corrector

We next formally derive and then rigorously analyze a refined approximation of (2.1)–(2.2). By Theorem 2.1, the velocity component $M(\theta^0)$ of the limit system (2.4)–(2.5) well approximates the velocity field $\mathbf{u}^\varepsilon$ of the full system (2.1)–(2.2) in the norm $L^2([0,t], H^1(\mathcal{D}))$ for each fixed $t > 0$. Also Theorem 2.2, (ii) shows that the invariant measure of the limit system approximates any invariant state of the full system, which can be interpreted as a approximation of laws of solutions as $t \to \infty$. On the other hand, we do not expect (2.4)–(2.5) to accurately describe the behavior of (2.1)–(2.2) up to $t = 0$ due to the presence of a (initial time) boundary layer.

We next derive the so called 'corrector equation' which provides effective dynamics for (2.1)–(2.2) and which is globally well-posed and whose velocity component remains close to the dynamics of (2.1)–(2.2) in $L^\infty([0,T],L^2(\mathcal{D}))$, that is, even up to time zero. Note that similar considerations motivate the analysis in [45] which treats such small time approximations in the deterministic setting.

Formal derivation. In order to identify multiple time scales in (2.1)–(2.2) we introduce an additional 'slow time' variable $\varsigma = \varepsilon t$. We then replace

$$\partial_t \to \partial_t + \frac{1}{\varepsilon}\partial_\varsigma.$$

Under this ansatz the momentum equation (2.1) becomes

$$\varepsilon(\partial_t \mathbf{u}^\varepsilon + \mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^\varepsilon) + \partial_\varsigma \mathbf{u}^\varepsilon - \Delta \mathbf{u}^\varepsilon = \nabla \tilde{p}^\varepsilon + Ra\hat{\mathbf{k}}\theta^\varepsilon$$

Dropping the terms of order ε and using Duhamel's formula we obtain

$$\mathbf{u}^\varepsilon(\varsigma) = e^{-A\varsigma}\mathbf{u}^\varepsilon(0) + \int_0^\varsigma e^{-A(\varsigma - r)} P(Ra\hat{\mathbf{k}}\theta^\varepsilon)dr \tag{5.21}$$

where as in Sect. 5.1, $e^{-A\varsigma}$ denotes the semigroup whose generator is the Stokes operator A.

The form of (2.1) suggests that $\mathbf{u}^\varepsilon$ fluctuates rapidly in comparison to θ^ε. Under the further ansatz that there is a clear separation of time scales between the motion of $\mathbf{u}^\varepsilon$ and that of θ^ε we suppose that θ^ε is independent of ς. From (5.21) this yields

$$\begin{aligned}\mathbf{u}^\varepsilon(\varsigma) &= e^{-A\varsigma}\mathbf{u}^\varepsilon(0) + A^{-1}(P(Ra \cdot \hat{\mathbf{k}}\tilde{\theta}^\varepsilon)) - e^{-A\varsigma}A^{-1}(P(Ra\hat{k}\theta^\varepsilon)) \\ &:= A^{-1}(P(Ra \cdot \hat{\mathbf{k}}\tilde{\theta}^\varepsilon)) + \mathbf{w}^\varepsilon(\varsigma),\end{aligned} \tag{5.22}$$

where $\mathbf{w}^\varepsilon$ solves

$$\begin{aligned}\partial_\varsigma \mathbf{w}^\varepsilon - \Delta \mathbf{w}^\varepsilon = \nabla q^\varepsilon, \quad \mathbf{w}^\varepsilon(0) = \mathbf{u}^\varepsilon(0) - \mathbf{y}^\varepsilon \\ \text{and} \quad -\Delta \mathbf{y}^\varepsilon = \nabla p^\varepsilon + Ra\hat{k}\theta^\varepsilon(0)\end{aligned} \tag{5.23}$$

and we have made the further approximation that $\theta^\varepsilon(t) \approx \theta^\varepsilon(0)$ relative to the slow time scale ς.

Next, we return to the original time scale t and obtain the effective dynamics for (2.1)–(2.2) starting from any initial condition $(\theta_0^\varepsilon, \mathbf{u}_0^\varepsilon) \in \mathbf{H}$,

$$-\Delta \tilde{\mathbf{u}}^\varepsilon = \nabla p^\varepsilon + Ra \cdot \hat{\mathbf{k}}\tilde{\theta}^\varepsilon + \Delta \mathbf{w}^\varepsilon(t), \quad \nabla \cdot \tilde{\mathbf{u}}^\varepsilon = 0, \tag{5.24}$$

$$d\tilde{\theta}^\varepsilon + \left(\tilde{\mathbf{u}}^\varepsilon \cdot \nabla\tilde{\theta}^\varepsilon - \Delta\tilde{\theta}^\varepsilon\right)dt = \tilde{Ra} \cdot \tilde{u}_3^\varepsilon dt + \sum_{k=1}^N \sigma_k dW^k, \quad \tilde{\theta}^\varepsilon(0) = \theta_0^\varepsilon, \tag{5.25}$$

where $\mathbf{w}^\varepsilon$ solves

$$\begin{aligned}\partial_t \mathbf{w}^\varepsilon - \frac{1}{\varepsilon}\Delta \mathbf{w}^\varepsilon = \frac{1}{\varepsilon}\nabla q^\varepsilon, \quad \nabla \cdot \mathbf{w}^\varepsilon = 0, \quad \mathbf{w}^\varepsilon(0) = \mathcal{P}_{N^\varepsilon}\mathbf{u}_0^\varepsilon + \mathbf{y}^\varepsilon, \\ \text{and} \quad -\Delta \mathbf{y}^\varepsilon = \nabla p^\varepsilon + Ra \cdot \hat{\mathbf{k}}\theta_0^\varepsilon, \quad \nabla \cdot \mathbf{y}^\varepsilon = 0.\end{aligned} \tag{5.26}$$

We supplement (5.24)–(5.26) with boundary conditions (2.3). Note that for technical reasons we slightly modify the initial condition for $\mathbf{w}^\varepsilon$ compared to

(5.23) by taking $\tilde{\mathbf{u}}^\varepsilon(0) = \mathcal{P}_{N^\varepsilon}\mathbf{u}_0^\varepsilon$, where we recall that $\mathcal{P}_{N^\varepsilon}$ is the projection onto the first N^ε modes of the Stokes operator A as in (5.27) and N^ε satisfies

$$\varepsilon(\lambda_{N^\varepsilon}^*)^2 \sim 1. \tag{5.27}$$

This specification $\tilde{\mathbf{u}}^\varepsilon(0)$ is only used to avoid regularity issues at the initial time in (5.24) and as such, a number of other modifications can be employed.

5.3.1. Rigorous error estimates for the corrector. The following theorem asserts that (5.24)–(5.26) approximates (2.1)–(2.2) in the desired norms.

Theorem 5.1. *Fix any $\varepsilon > 0$ choose N_ε satisfying (5.27). Suppose we are given a sequence $\{\mu^{0,\varepsilon}\}_{\varepsilon\in(0,1)} \subset \mathrm{Pr}(\mathbf{H})$ such that*

$$\sup_{0<\varepsilon\leq 1}\int_{\mathbf{H}} \left(\|\nabla \mathbf{u}\|^2 + \exp(\eta\|\mathbf{u}\|^2 + \|\theta\|_{L^3}^2)\right) d\mu^{0,\varepsilon}(\mathbf{u},\theta) < \infty. \tag{5.28}$$

For each $\varepsilon > 0$ we consider a martingale solutions $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ of (2.1)–(2.3) as in Proposition 2.1, (i). We suppose that each $(\mathbf{u}^\varepsilon, \theta^\varepsilon)$ has initial conditions distributed according to the distribution $\mu^{0,\varepsilon}$ and satisfies the uniform moment bound (2.12). In particular, for each $\varepsilon > 0$ the corresponding martingale solution fixes a stochastic basis $\mathcal{S}$ and defines $(\theta_0^\varepsilon, \mathbf{u}_0^\varepsilon) := (\theta^\varepsilon(0), \mathbf{u}^\varepsilon(0))$. Then,

(i) *up to the specification of the stochastic basis $\mathcal{S}$ and the initial conditions $(\mathbf{u}_0^\varepsilon, \theta_0^\varepsilon)$, there exists a unique, adapted*

$$\tilde{\theta}^\varepsilon \in L^2(\Omega; L^2_{loc}([0,\infty); V) \cap C([0,\infty); H))$$

solving (5.24)–(5.26).

(ii) *For any $t > 0$ there is a $\gamma_0 = \gamma_0(\tilde{Ra}, Ra, t)$ such that, for any $0 < \gamma \leq \gamma_0$, $p \geq \gamma$, and $\varepsilon > 0$,*

$$\begin{aligned} \mathbb{E}\left(\sup_{s\in[0,t]} \|\tilde{\theta}^\varepsilon(s) - \theta^\varepsilon(s)\|^p\right) &\leq C\varepsilon^\gamma\,, \\ \mathbb{E}\left(\sup_{s\in[0,t]} \|\tilde{\mathbf{u}}^\varepsilon(s) - \mathbf{u}^\varepsilon(s)\|^p\right) &\leq C\varepsilon^{\gamma/4}\,, \end{aligned} \tag{5.29}$$

where the constants $C = C(\tilde{Ra}, Ra, t, p)$ and γ_0 and are both independent of $\varepsilon > 0$.

Proof. As in Proposition 2.1, (iii), (iv) the well posedness of (5.24)–(5.26) is standard and can be established along similar lines as one would for the 2D Stochastic Navier–Stokes equations. To see this observe that, although we are working in 3D, we have one more degree of smoothing in the constitutive law, (5.24), producing $\tilde{\mathbf{u}}^\varepsilon$ from $\tilde{\theta}^\varepsilon$ compared to Biot-Savart in the Navier–Stokes equation. We omit further details here again referring the reader to e.g. [12,13,17].

To prove (5.29) we reuse many of the estimates from the proof of Theorem 2.1. Taking $\tilde{\mathbf{v}}^\varepsilon = \mathbf{u}^\varepsilon - \tilde{\mathbf{u}}^\varepsilon$ and $\tilde{\phi}^\varepsilon = \theta^\varepsilon - \tilde{\theta}^\varepsilon$, we have

$$\partial_t\tilde{\phi}^\varepsilon - \Delta\tilde{\phi}^\varepsilon = \tilde{Ra}\cdot\tilde{v}_3^\varepsilon - \tilde{\mathbf{v}}^\varepsilon\cdot\nabla\tilde{\theta}^\varepsilon - \mathbf{u}^\varepsilon\cdot\nabla\tilde{\phi}^\varepsilon, \qquad \tilde{\phi}^\varepsilon(0) = 0\,,$$

and hence repeating the arguments leading to (5.5) we obtain the estimate

$$\sup_{s\in[0,t]} \|\tilde{\phi}^\varepsilon(s\wedge\tau)\|^2 \le \sup_{s\in[0,t\wedge\tau]} \left(\|\tilde{\theta}^\varepsilon(s)\|^2_{L^3} + \tilde{Ra}^2\right) \int_0^{t\wedge\tau} \|\nabla\tilde{\mathbf{v}}^\varepsilon(t')\|^2 dt' \quad (5.30)$$

for any $t > 0$ and any stopping time $\tau \ge 0$. By (5.24) and (2.1), $\tilde{\mathbf{v}}^\varepsilon$ satisfies

$$\varepsilon(\partial_t \mathbf{u}^\varepsilon + \mathbf{u}^\varepsilon \cdot \nabla \mathbf{u}^\varepsilon) - \Delta\tilde{\mathbf{v}}^\varepsilon = \nabla q^\varepsilon + Ra\hat{\mathbf{k}}\tilde{\phi}^\varepsilon - \Delta\mathbf{w}^\varepsilon, \quad (5.31)$$

where $\mathbf{w}^\varepsilon$ solves (5.26). Referring to (5.24) we have

$$\tilde{\mathbf{u}}^\varepsilon = A^{-1}(RaP(\hat{\mathbf{k}}\tilde{\theta}^\varepsilon)) - \mathbf{w}^\varepsilon\,, \quad (5.32)$$

and consequently (5.25) and (5.26) yield

$$\begin{aligned} d\tilde{\mathbf{u}}^\varepsilon &= -d\mathbf{w}^\varepsilon + RaA^{-1}P(\hat{\mathbf{k}}d\tilde{\theta}^\varepsilon) \\ &= -\frac{1}{\varepsilon}(\Delta\mathbf{w}^\varepsilon + \nabla q^\varepsilon)dt \\ &\quad -RaA^{-1}P\Big(\hat{\mathbf{k}}\Big(\tilde{\mathbf{u}}^\varepsilon\cdot\nabla\tilde{\theta}^\varepsilon - \Delta\tilde{\theta}^\varepsilon - \tilde{Ra}\cdot\tilde{u}_3^\varepsilon\Big)\Big)\,dt \\ &\quad + Ra\sum_{k=1}^N A^{-1}P(\hat{\mathbf{k}}\sigma_k)dW^k\,. \end{aligned} \quad (5.33)$$

Multiplying (5.31) by ε^{-1}, subtracting (5.33) and rearranging we obtain

$$\begin{aligned} d\tilde{\mathbf{v}}^\varepsilon - \frac{1}{\varepsilon}\Delta\tilde{\mathbf{v}}^\varepsilon dt &= \frac{1}{\varepsilon}\left(\nabla\tilde{q}^\varepsilon + Ra\hat{\mathbf{k}}\tilde{\phi}^\varepsilon\right)dt \\ &\quad +\Big(RaA^{-1}P\Big(\hat{\mathbf{k}}\Big(\tilde{\mathbf{u}}^\varepsilon\cdot\nabla\tilde{\theta}^\varepsilon - \Delta\tilde{\theta}^\varepsilon - \tilde{Ra}\cdot\tilde{u}_3^\varepsilon\Big)\Big) - \mathbf{u}^\varepsilon\cdot\nabla\mathbf{u}^\varepsilon\Big)dt \\ &\quad - Ra\sum_{k=1}^N A^{-1}P(\hat{\mathbf{k}}\sigma_k)dW^k, \quad \tilde{\mathbf{v}}^\varepsilon(0) = (I-\mathcal{P}_N)\mathbf{u}_0^\varepsilon, \end{aligned} \quad (5.34)$$

with $\nabla\cdot\tilde{\mathbf{v}}^\varepsilon = 0$. Here note the close similarity between (5.8) and (5.34); in view of (5.32), (5.26) the primary distinction here is in the initial condition.

As above in (5.9), the Itō formula implies

$$\begin{aligned} d\|\tilde{\mathbf{v}}^\varepsilon\|^2 + \frac{2}{\varepsilon}\|\nabla\tilde{\mathbf{v}}^\varepsilon\|^2 dt \\ &= \frac{2}{\varepsilon}Ra\langle\tilde{\phi}^\varepsilon, \tilde{v}_3^\varepsilon\rangle dt \\ &\quad + 2\left\langle RaA^{-1}P\Big(\hat{\mathbf{k}}\left(\tilde{\mathbf{u}}^\varepsilon\cdot\nabla\tilde{\theta}^\varepsilon - \Delta\tilde{\theta}^\varepsilon - \tilde{Ra}\cdot\tilde{u}_3^\varepsilon\right)\Big) - \mathbf{u}^\varepsilon\cdot\nabla\mathbf{u}^\varepsilon, \tilde{\mathbf{v}}^\varepsilon\right\rangle dt \\ &\quad + Ra^2\sum_{k=1}^N |A^{-1}P(\hat{\mathbf{k}}\sigma_k)|^2 dt - 2Ra\sum_{k=1}^N\langle A^{-1}P(\hat{\mathbf{k}}\sigma_k), \tilde{\mathbf{v}}^\varepsilon\rangle dW^k \\ &:= (T_1+T_2+T_3+T_4+T_5+T_6)dt + SdW. \end{aligned} \quad (5.35)$$

We now estimate (5.35) with bounds similar to (5.10)–(5.13). Here, bounds on $\mathbf{u}^0$ need to be replaced with appropriate estimates for $\tilde{\mathbf{u}}^\varepsilon$. For the terms T_2, T_4 we simply treat $\tilde{\mathbf{u}}^\varepsilon$ terms as

$$\|\tilde{\mathbf{u}}^\varepsilon(t)\|^2 \le C(Ra^2\|\tilde{\theta}^\varepsilon(t)\|^2 + \|\mathbf{w}^\varepsilon(t)\|^2)$$
$$\le C(Ra^2\|\tilde{\theta}^\varepsilon(t)\|^2 + \|\mathbf{u}_0^\varepsilon\|^2 + Ra^2\|\theta_0^\varepsilon\|^2)\,. \tag{5.36}$$

The estimate (5.13) on the term T_5 involves an L^∞ bound on $\tilde{\mathbf{u}}^\varepsilon$ and thus requires a bit more care. In this case

$$\|\tilde{\mathbf{u}}^\varepsilon(t)\|_{L^\infty}^2 \le C\|\tilde{\mathbf{u}}^\varepsilon(t)\|_{H^2}^2 \le C(Ra^2\|\tilde{\theta}^\varepsilon(t)\|^2 + \|\mathbf{w}^\varepsilon(t)\|_{H^2}^2).$$

By standard properties of analytic semigroups, the inverse Poincaré inequality, and the form of $\mathbf{w}(0)$ one has

$$\|\mathbf{w}^\varepsilon(t)\|_{H^2}^2 = \|e^{At/\varepsilon}\mathbf{w}^\varepsilon(0)\|_{H^2}^2 \le C\|\mathbf{w}^\varepsilon(0)\|_{H^2}^2 \le CRa^2\|\theta_0^\varepsilon\|^2 + C\|\mathcal{P}_{N^\varepsilon}\tilde{\mathbf{u}}_0^\varepsilon\|_{H^2}^2$$
$$\le CRa^2\|\theta_0^\varepsilon\|^2 + C(\lambda_{N^\varepsilon}^*)^2\|\tilde{\mathbf{u}}_0^\varepsilon\|^2\,, \tag{5.37}$$

where C is independent of $\varepsilon \in (0,1)$. Combining these two estimate and again taking advantage of the cancelation from $\mathbf{u}^\varepsilon = \tilde{\mathbf{v}}^\varepsilon + \tilde{\mathbf{u}}^\varepsilon$

$$|T_5| \le \frac{1}{4\varepsilon}\|\nabla\tilde{\mathbf{v}}^\varepsilon\|^2 + 4\varepsilon\|\tilde{\mathbf{u}}^\varepsilon\|_{L^\infty}^2\|\mathbf{u}^\varepsilon\|^2$$
$$\le \frac{1}{4\varepsilon}\|\nabla\tilde{\mathbf{v}}^\varepsilon\|^2 + C(Ra^2+1)(\|\theta_0^\varepsilon\|^4 + \|\mathbf{u}_0^\varepsilon\|^4 + \|\mathbf{u}^\varepsilon\|^4). \tag{5.38}$$

Observe that in comparison to (5.13), the additional power of ε is used to cancel $(\lambda_{N^\varepsilon}^*)^2$.

Combining the analogues of (5.10)–(5.12) with (5.36) and using (5.36) with (5.35) we obtain

$$d\|\tilde{\mathbf{v}}^\varepsilon\|^2 + \frac{1}{\varepsilon}\|\nabla\tilde{\mathbf{v}}^\varepsilon\|^2 dt$$
$$\le \frac{CRa^2}{\varepsilon}\|\tilde{\phi}^\varepsilon\|^2 dt$$
$$+ C(1+Ra^4)(1+\tilde{Ra}^2)(\|\tilde{\theta}^\varepsilon\|^4 + \|\theta_0^\varepsilon\|^4 + \|\mathbf{u}_0^\varepsilon\|^4 + \|\mathbf{u}^\varepsilon\|^4 + 1)dt$$
$$- 2Ra\sum_{k=1}^N \langle A^{-1}(P(\hat{\mathbf{k}}\sigma_k)), \tilde{\mathbf{v}}^\varepsilon\rangle dW^k\,, \tag{5.39}$$

where the constant $C > 0$ is independent of $Ra, \tilde{Ra}$, and $\varepsilon > 0$. We now use (5.39) with (5.30) and repeat the stopping time argument as in (5.17)–(5.20) to infer the first part of the (5.29).[10]

We turn next to the the convergence of the velocity fields, the second part of (5.29). We obtain from (5.35) and the pointwise bounds yielding the drift terms in (5.39) that

$$\|\tilde{\mathbf{v}}^\varepsilon(t)\|^2 \le \exp\left(-\frac{t}{\varepsilon}\right)\|\tilde{\mathbf{v}}_0^\varepsilon\|^2$$
$$+ C\int_0^t \exp\left(-\frac{t-s}{\varepsilon}\right)\left(\frac{Ra^2}{\varepsilon}\|\tilde{\phi}^\varepsilon(s)\|^2 + \mathcal{R}_\varepsilon(s)\right)ds + \mathcal{X}_\varepsilon(t),$$

[10]Note that the loss of the ε in front of the second term after the inequality in (5.39) in comparison (5.14) does not charge the ultimate outcome of this bound as we only required an ε-independent upper bound for $\mathcal{M}_\varepsilon$ in (5.18).

$$\leq \exp\left(-\frac{t}{\varepsilon}\right)\|\tilde{\mathbf{v}}_0^\varepsilon\|^2 + CRa^2 \sup_{s\in[0,t]}\|\tilde{\phi}^\varepsilon(s)\|^2 + \varepsilon \sup_{s\in[0,t]}\mathcal{R}_\varepsilon(s) + \mathcal{X}_\varepsilon(t)\,, \tag{5.40}$$

where

$$\mathcal{R}_\varepsilon(t) := C(1+Ra^4)(1+\tilde{Ra}^2)(\|\tilde{\theta}^\varepsilon\|^4 + \|\theta_0^\varepsilon\|^4 + \|\mathbf{u}_0^\varepsilon\|^4 + \|\mathbf{u}^\varepsilon\|^4 + 1)$$

and

$$\begin{aligned}\mathcal{X}_\varepsilon(t) &:= -2Ra\int_0^t \exp\left(-\frac{t-s}{\varepsilon}\right)\sum_{k=1}^N \langle A^{-1}(P(\hat{\mathbf{k}}\sigma_k)), \tilde{\mathbf{v}}^\varepsilon\rangle dW^k \\ &=: \int_0^t \exp\left(-\frac{t-s}{\varepsilon}\right) g(s)dW.\end{aligned}$$

Using the inverse Poincaré inequality and (5.27) one has

$$\begin{aligned}\exp\left(-\frac{t}{\varepsilon}\right)\|\tilde{\mathbf{v}}_0^\varepsilon\|^2 &\leq \|(I-P_{N^\varepsilon})\mathbf{u}_0^\varepsilon\|^2 \leq C(\lambda_{N^\varepsilon}^*)^{-1}\|\nabla\mathbf{u}_0^\varepsilon\|^2 \\ &\leq C\varepsilon\|\nabla\mathbf{u}_0^\varepsilon\|^2\,.\end{aligned} \tag{5.41}$$

Therefore combining (5.40) with (5.41), using the bound already obtained for $\|\tilde{\phi}^\varepsilon(s)\|$ in (5.29) and the uniform bounds (2.12), (5.28)

$$\mathbb{E}\sup_{s\in[0,t]}\|\tilde{\mathbf{v}}^\varepsilon(t)\|^p \leq C(\varepsilon^{p/2}+\varepsilon^\gamma) + C\mathbb{E}\sup_{s\in[0,t]}|\mathcal{X}_\varepsilon(s)|^{p/2} \tag{5.42}$$

for any $p>0$, where $\gamma = \min\{p,\gamma_0\}$ is obtained from the bound $\|\tilde{\phi}^\varepsilon(s)\|$ and the constant $C = C(Ra, \tilde{Ra}, t, p)$ is independent of $\varepsilon > 0$.

In order to estimate $\mathcal{X}_\varepsilon$ observe that this process satisfies

$$d\mathcal{X}_\varepsilon + \frac{1}{\varepsilon}\mathcal{X}_\varepsilon dt = gdW, \quad \mathcal{X}_\varepsilon(0) = 0,$$

and hence, by the Itō lemma,

$$d\mathcal{X}_\varepsilon^2 + \frac{2}{\varepsilon}\mathcal{X}_\varepsilon^2 dt = g^2 dt + 2g\mathcal{X}_\varepsilon dW\,.$$

Consequently,

$$\mathbb{E}\sup_{s\in[0,t]}|\mathcal{X}_\varepsilon(s)|^{p/2} \leq C\mathbb{E}\sup_{s\in[0,t]}\left|\int_0^s g\mathcal{X}_\varepsilon dW\right|^{p/4} + C\mathbb{E}\left(\int_0^t g^2 ds\right)^{p/4}. \tag{5.43}$$

With the Burkholder–David–Gundy inequality and Young's inequality we have

$$\begin{aligned}\mathbb{E}\sup_{s\in[0,t]}\left|\int_0^s g\mathcal{X}_\varepsilon dW\right|^{p/4} &\leq C\mathbb{E}\left(\int_0^t g^2\mathcal{X}_\varepsilon^2 ds\right)^{p/8} \\ &\leq \frac{1}{2}\mathbb{E}\sup_{s\in[0,t]}|\mathcal{X}_\varepsilon|^{p/2} + C\mathbb{E}\left(\int_0^t \|\tilde{\mathbf{v}}^\varepsilon\|^2 ds\right)^{p/4}.\end{aligned} \tag{5.44}$$

With (5.43), (5.44), the bound (5.45) now yields

$$\mathbb{E}\sup_{s\in[0,t]}\|\tilde{\mathbf{v}}^\varepsilon(t)\|^p \le C(\varepsilon^{p/2}+\varepsilon^\gamma)+C\mathbb{E}\left(\int_0^t\|\tilde{\mathbf{v}}^\varepsilon\|^2ds\right)^{p/4}. \tag{5.45}$$

On the other hand, from (5.39), using for a second time the existing bounds on $\|\tilde{\phi}^\varepsilon(s)\|$ in (5.29) and (2.12), we have

$$\begin{aligned}\mathbb{E}\int_0^t\|\nabla\tilde{\mathbf{v}}^\varepsilon\|^2dt &\le \varepsilon\mathbb{E}\|\tilde{\mathbf{v}}_0^\varepsilon\|^2 + CRa^2\int_0^t\mathbb{E}\|\tilde{\phi}^\varepsilon\|^2ds+\varepsilon\mathbb{E}\int_0^t\mathcal{R}_\varepsilon ds\\ &\le C(\varepsilon+\varepsilon^\gamma),\end{aligned} \tag{5.46}$$

where the constant C depends on t, Ra, $\tilde{Ra}$ but again is independent of $\varepsilon>0$. When $p\le 4$ the second portion of the desired inequality (5.29) now follows from (5.45), (5.46), and Hölder's inequality. On the other hand, when $p>4$ then we estimate

$$\begin{aligned}\mathbb{E}\left(\int_0^t\|\tilde{\mathbf{v}}^\varepsilon\|^2ds\right)^{p/4} &\le \mathbb{E}\sup_{s\in[0,t]}\|\tilde{\mathbf{v}}^\varepsilon\|^{p-2}+\mathbb{E}\int_0^t\|\tilde{\mathbf{v}}^\varepsilon\|^2ds\\ &\le \sup_{s\in[0,t]}\|\tilde{\mathbf{v}}^\varepsilon\|^{p-2}+C(\varepsilon+\varepsilon^\gamma),\end{aligned}$$

so that the second part of (5.29) follows in this later case with (5.46) and an iterative argument. This completes the proof of Theorem 5.1, (ii). □

6. Convergence of the Nusselt number

In this final section we prove Corollary 1.1, note that this is not an immediate consequence of Theorem 2.2. Indeed, the Nusselt number Nu (cf. (1.8)) is defined as a statistical average of the observable

$$\phi_{Nu}(\mathbf{u},\theta)=\int_{\mathcal{D}}u_2\theta dx, \tag{6.1}$$

but (2.30) is not satisfied since $\phi_{Nu}\notin V(\mathbf{H})$. Nevertheless, Corollary 1.1 follows from a combination of Theorem 2.2 with exponential moment bounds on the unique invariant measures.

Proof of Corollary 1.1. In two space dimensions, with $N=\infty$ and $Pr=Pr(Ra,\tilde{Ra})$ large enough (or $\varepsilon=Pr^{-1}$ small enough), [21, Theorem 1.3] and Theorem 2.2 yields respectively unique invariant measures μ_ε of (2.1)–(2.2) and μ_0 for (2.4)–(2.5). We aim to show that $\lim_{\varepsilon\to0}(Nu)_\varepsilon=(Nu)_0$. To simplify the notation, write ϕ instead of ϕ_{Nu} defined by (6.1).

Consider a smooth cut-off function $\psi:[0,\infty)\to[0,1]$ satisfying $\psi\equiv1$ for $x\in[0,1]$, $\psi\equiv0$ for $x\ge2$, and $|\psi'|\le2$ everywhere. Denote $\psi_R(\cdot):=\psi(\cdot/R)$ and write

$$\begin{aligned}&\left|\int_{\mathbf{H}}\phi(\mathbf{u},\theta)d\mu_\varepsilon(\mathbf{u},\theta)-\int_H\phi(L(\theta))d\mu_0(\theta)\right|\\ &\quad\le\left|\int_{\mathbf{H}}\phi(\mathbf{u},\theta)\psi_R(\|\mathbf{u}\|^2+\|\theta\|^2)d\mu_\varepsilon(\mathbf{u},\theta)-\int_H\phi(L(\theta))\psi_R(\|L(\theta)\|^2)d\mu_0(\theta)\right|\end{aligned}$$

$$
\begin{aligned}
&+ \left| \int_{\mathbf{H}} \phi(\mathbf{u},\theta)(1-\psi_R(\|\mathbf{u}\|^2+\|\theta\|^2)) d\mu_\varepsilon(\mathbf{u},\theta) \right| \\
&+ \left| \int_{H} \phi(L(\theta))(1-\psi_R(\|L(\theta)\|^2)) d\mu_0(\theta) \right| \\
=: &\ I_1 + I_2 + I_3.
\end{aligned}
$$

We bound the first term by observing that $\phi_R(\mathbf{u},\theta) := \phi(\mathbf{u},\theta)\psi_R(\|\mathbf{u}\|^2+\|\theta\|^2)$ satisfies

$$
[\phi_R]_\eta \leq C(R+1), \tag{6.2}
$$

for a constant $C > 0$. It follows from (6.2) and Theorem 2.2 that $I_1 \leq \tilde{C}(R+1)\varepsilon^{\tilde{q}}$. We control I_2 using Markov's inequality and (2.13), which yields, for any $\eta > 0$ sufficiently small

$$
\mu_\varepsilon(E_R^c) \leq e^{-\eta R} \int_{\mathbf{H}} \exp(\eta(\|\mathbf{u}\|^2+\|\theta\|^2)) d\mu_\varepsilon(\mathbf{u},\theta) \leq C_0 e^{-\eta R}.
$$

uniformly in $\varepsilon > 0$, where $E_R = \{\|\mathbf{u}\|^2 + \|\theta\|^2 \leq R\}$. Notice that $1 - \psi_R(\|u\|^2 + \|\theta\|^2) \leq \mathbb{1}_{E_R^c}$, and observe that for fixed $\eta > 0$, on the set E_R^c we have $|\phi_{Nu}(\mathbf{u},\theta)| \leq \exp(\eta(\|\mathbf{u}\|^2+\|\theta\|^2)/2)$ for $R = R(\eta)$ sufficiently large. By Hölder's inequality we obtain

$$
I_2 \leq C \left(\int_{\mathbf{H}} \exp(\eta(\|\mathbf{u}\|^2+\|\theta\|^2)) d\mu_\varepsilon(\mathbf{u},\theta) \right)^{1/2} (\mu_\varepsilon(E_R^c))^{1/2} \leq C C_0^{1/2} e^{-\eta R/2}.
$$

We can control I_3 similarly by using exponential moment bounds for the invariant measure μ_0, and the estimate $\|L(\theta)\| \leq C\|\theta\|$. Thus,

$$
\begin{aligned}
&\left| \int_{\mathbf{H}} \phi_{Nu(\varepsilon)}(\mathbf{u},\theta) d\mu_\varepsilon(\mathbf{u},\theta) - \int_{H} \phi_{Nu(0)}(L(\theta)) d\mu_0(\theta) \right| \\
&\quad \leq C((R+1)\varepsilon^{\tilde{q}} + e^{-\eta R/2}),
\end{aligned} \tag{6.3}
$$

and we can make this expression less than any $\delta > 0$ by taking $R = R(\delta)$ sufficiently large and then $\varepsilon = \varepsilon(R,\delta)$ sufficiently small. □

Remark 6.1. From the proof above we can easily infer that Corollary 1.1 will also apply to any observable ϕ on the extended phase space that is locally Lipschitz and sub-exponential at infinity (for a sufficiently small exponential power $\eta > 0$ dictated by (2.13)). We have chosen the convergence of the Nusselt number to emphasize the physical significance of our results.

Acknowledgements

The majority of this work was carried out during a Research in Peace (RIP) fellowship at the Institut Mittag-Leffler, Stockholm, Sweden. We are grateful to this institution for providing us with a unique working environment to continue our ongoing collaborations. JF and GR would like to acknowledge Virginia Tech for facilitating a research visit during which this work was finalized. We would also like to extend our warm thanks to P. Constantin, C. Doering, S. Friedlander, J. Mattingly, E. Thomann, and J. Whitehead for stimulating feedback on this work. JF was partially supported in this work under the Grant NSF-DMS-1816408. NEGH was partially supported in this work under the Grants NSF-DMS-1313272, NSF-DMS-1816551, and Simons-515990.

Appendix A: Moment bounds for stochastic drift-diffusion equations

In this appendix we collect some moment bounds proved in [21] which have been used in the analysis above.

As in [21] we consider the following class of stochastic divergence-free drift diffusion systems

$$d\xi + \mathbf{v} \cdot \nabla \xi dt = (\tilde{Ra} \cdot v_3 + \Delta \xi)dt + \sum_{k=1}^{N} \sigma_k dW^k, \quad \xi(0) = \xi_0 \tag{A.1}$$

evolving on the three dimensional domain $\mathcal{D} = [0, L]^2 \times [0, 1]$. Here $\tilde{Ra} > 0$ is a fixed parameter and $\mathbf{v} = (v_1, v_2, v_3)$ is any sufficiently regular and adapted, divergence free vector field. Both $\mathbf{v}$ and ξ are supposed to satisfy the boundary condition (2.3). Recall that by the change of variable $T = \xi + \tilde{Ra}(1 - z)$ we may reformulate (A.1) as

$$\begin{aligned} dT + \mathbf{v} \cdot \nabla T dt &= \Delta T dt + \sum_{k=1}^{N} \sigma_k dW^k, \\ T(0) &= T_0 = \xi_0 + \tilde{Ra}(1 - z)\,, \end{aligned} \tag{A.2}$$

where $\mathbf{v}$ and T satisfy boundary conditions (1.3). As such, bounds for ξ solving (A.1) immediately translate to bounds for T.

In [21] we prove:

Proposition A.1. *Suppose that* $\mathbf{v} \in L^2_{loc}([0, \infty); V_1 \cap (H^2(\mathcal{D}))^3) \cap C([0, \infty); H_1)$ *a.s. and is* $\mathcal{F}_t$*-adapted. Fix any* $p \geq 2$ *and any initial condition* $\xi_0 \in H \cap L^p(\mathcal{D})$ *which is* $\mathcal{F}_0$*-measurable with*

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

for some $\eta > 0$*. Then there exists* $\eta_0 = \eta_0(\sigma, \tilde{Ra}, p) > 0$ *such that for any* $t \geq 0$ *and any positive* $\eta \leq \eta_0$*,*

$$\begin{aligned} &\mathbb{E} \exp\left(\frac{\eta}{2^{p/2+2}} \sup_{s \in [0,t]} \|\xi\|^2_{L^p} \right) \\ &\quad \leq C_1 \mathbb{E} \exp\left(\eta \|\xi_0\|^2_{L^p} + \eta p t (\|\sigma\|^2_{L^p} + 2^{p/2}(4\tilde{Ra}^2 + 1)) \right) \end{aligned} \tag{A.3}$$

for a constant $C = C(\tilde{Ra}, p)$ *independent of* t*,* η*,* ξ_0*, and* $\mathbf{v}$*. Furthermore,*

$$\mathbb{E} \exp\left(\frac{\eta}{2^{p/2+2}} \|\xi(t)\|^2_{L^p} \right) \leq C \mathbb{E} \exp\left(\eta (e^{-\kappa t} \|\xi_0\|^2) \right)\,, \tag{A.4}$$

where again $C = C(\tilde{Ra}, p, \|\sigma\|_{L^p}, \mathcal{D})$ *and* $\kappa = \kappa(\tilde{Ra}, \mathcal{D}) > 0$ *are independent of* t, η, ξ_0, *and* $\mathbf{v}$.

We now return to the infinite Prandtl system (2.4)–(2.5) and recall a bound analogous to (A.4) but which uses more of the specific structure of the velocity equation.

Proposition A.2. *Fix an initial condition* $\theta_0^0 \in H$ *which is* $\mathcal{F}_0$*-measurable, and let* $\theta^0 = \theta^0(t, \theta_0^0)$ *denote the corresponding solution to* (2.4)–(2.5)*. There is a universal constant* $\eta^* > 0$ *such that for any* $t > 0$ *and* $\eta \in (0, \eta^*]$*, there exists* $C = C(Ra, \tilde{Ra}) > 0$ *such that*

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

The proof of Proposition A.2 can be found in [21].

Appendix B: Gradient estimates on the Markov semigroup

In this section we establish the gradient bound for the Markov semigroup generated by (2.4)–(2.5) in order to prove (4.2). For this purpose we begin by briefly recalling how (4.2) is translated to a control problem through the use of Malliavin calculus. We refer to e.g. [36] or [35] for further general background on this subject and to [20,26,28] for the application of this formalism in a setting close to ours.

Define the random operators

$$\mathcal{J}_{0,t}\xi := \lim_{\epsilon\to 0}\frac{\theta^0(t, \theta_0 + \epsilon\xi, W) - \theta^0(t, \theta_0, W)}{\epsilon} \tag{B.1}$$

for any $\xi \in H$ and

$$\mathcal{A}_{0,t}w := \lim_{\epsilon\to 0}\frac{\theta^0(t, \theta_0, W + \epsilon\int_0^\cdot w) - \theta^0(t, \theta_0, W)}{\epsilon} \tag{B.2}$$

for any $w \in L^2(\Omega; L^2([0,t]; \mathbb{R}^N))$. Here $\mathcal{A}_{0,t}w = \langle \mathfrak{D}\theta^0, w\rangle$, where the unbounded operator $\mathfrak{D} : L^2(\Omega; H) \mapsto L^2(\Omega; L^2(0, t, \mathbb{R}^N) \otimes H)$ is the Malliavin derivative and w is any element in the domain of the dual operator δ of $\mathfrak{D}$.

For our purposes it is sufficient to recall that any $\mathcal{F}_t$-adapted process in $L^2(\Omega; L^2([0,t]; \mathbb{R}^N))$ belongs to the domain of δ and $\delta(w)$ corresponds to the Itō integral of w so that

$$\mathbb{E}\langle \mathfrak{D}X, w\rangle = \mathbb{E}\left(X\int_0^t w dW\right) \tag{B.3}$$

for any $X \in \mathrm{Dom}(\mathfrak{D})$ and any $\mathcal{F}_t$-adapted w. This is a special case of the *Malliavin integration by parts formula*. We furthermore recall that $\mathfrak{D}$ satisfies a chain rule namely that if $\phi \in C^1(H)$ and $\theta \in \mathrm{Dom}(\mathfrak{D})$ then $\phi(\theta) \in \mathrm{Dom}(\mathfrak{D})$ and

$$\mathfrak{D}\phi(\theta) = \nabla\phi(\theta)\mathfrak{D}\theta. \tag{B.4}$$

Combining (B.3)–(B.4) and making use of the Itō isometry we infer that,

$$\begin{aligned}\nabla P_t^0\phi(\theta_0)\xi &= \mathbb{E}\left(\nabla\phi(\theta^0(t,\theta_0))\mathcal{J}_{0,t}\xi\right)\\ &= \mathbb{E}\left(\phi(\theta^0(t,\theta_0))\int_0^t wdW\right) + \mathbb{E}\left(\nabla\phi(\theta^0(t,\theta_0))\left(\mathcal{J}_{0,t}\xi - \mathcal{A}_{0,t}w\right)\right)\\ &\leq \sqrt{P_t^0(|\phi(\theta)|^2)}\left(\mathbb{E}\int_0^t |w|^2 ds\right)^{1/2}\\ &\quad + \sqrt{P_t^0(\|\nabla\phi(\theta)\|^2)}\left(\mathbb{E}\|\mathcal{J}_{0,t}\xi - \mathcal{A}_{0,t}w\|^2\right)^{1/2} \end{aligned} \tag{B.5}$$

for any $\phi \in C_b^1(H)$, $\theta_0 \in H$ and any (adapted) $w \in L^2(\Omega; L^2([0,t];\mathbb{R}^N))$.

Our desired bound (4.2) follows from (B.5) if, for every $\xi \in H$ with $\|\xi\| = 1$ there is (adapted) $w = w(\xi) \in L^2([0,\infty);\mathbb{R}^N)$ such that

$$\mathbb{E}\|\mathcal{J}_{0,t}\xi - \mathcal{A}_{0,t}w(\xi)\|^2 \leq C\exp(2\eta\|\theta_0\|^2)\delta(t)\,, \tag{B.6}$$

$$\sup_{\|\xi\|=1}\mathbb{E}\int_0^\infty |w(\xi)|^2 dt \leq C\exp(2\eta\|\theta_0\|^2)\,, \tag{B.7}$$

where $\delta(t) \to 0$ as $t \to \infty$ and C, η, and δ are independent of θ_0.

To solve the control problem (B.6)–(B.7) we observe that (B.1) and (B.2) admit explicit characterizations as linearizations of (2.4)–(2.5). For any $\xi \in H^0$ we let $\rho(t) = \rho(t,\xi) := \mathcal{J}_{0,t}\xi$, which satisfies

$$\begin{aligned} &\partial_t\rho + \mathbf{u}^0\cdot\nabla\rho + \mathbf{v}^0\cdot\nabla\theta^0 = \tilde{Ra}\cdot v_3^0 + \Delta\rho,\\ &-\Delta\mathbf{v}^0 = \nabla p + \mathrm{Ra}\hat{\mathbf{k}}\rho, \quad \nabla\cdot\mathbf{v}^0 = 0\,, \quad \rho(0) = \xi, \end{aligned} \tag{B.8}$$

supplemented by boundary conditions as in (2.3).[11] On the other hand, setting $\tilde{\rho} := \mathcal{A}_{0,t}w$ for any $w \in L^2([0,t],\mathbb{R}^N)$ we have

$$\begin{aligned} &\partial_t\tilde{\rho} + \mathbf{u}^0\cdot\nabla\tilde{\rho} + \tilde{\mathbf{v}}^0\cdot\nabla\theta^0 = \tilde{Ra}\cdot v_3^0 + \Delta\tilde{\rho} + \sum_{k=1}^N \sigma_k w_k,\\ &-\Delta\tilde{\mathbf{v}}^0 = \nabla p + \mathrm{Ra}\hat{\mathbf{k}}\tilde{\rho},\ \nabla\cdot\tilde{\mathbf{v}}^0 = 0\,,\ \tilde{\rho}(0) = 0, \end{aligned} \tag{B.9}$$

again with boundary conditions as in (2.3).

Denote $\bar{\rho}(t) = \bar{\rho}(t,\xi,w) = \rho - \tilde{\rho}$ and $\bar{\mathbf{v}} := \mathbf{v} - \tilde{\mathbf{v}}$ for any $w \in L^2([0,\infty);\mathbb{R}^N)$ and $\xi \in H$. We now choose w as a function of ξ as follows. Let P_N be the projection on the first N eigenfunctions of the Laplacian with boundary conditions as in (2.3). Set $w(t) := \sigma^{-1}\lambda P_N\bar{\rho}$, where $\lambda > 0$ and N will be selected below.[12]

[11] Notice that (B.8) can also be written as

$$\partial_t\rho + (L\theta^0)\cdot\nabla\rho + (L\rho)\cdot\nabla\theta^0 = \tilde{Ra}(L\rho) + \Delta\rho, \quad \rho(0) = \xi,$$

where $L = RaA^{-1}P\hat{k}$ and A is the Stokes operator, P; cf. (5.1) and (2.6) above. Similar formulations can also be given for (B.9), (B.10).

[12] Of course the choice of N will determine the number of modes subject to stochastic perturbation. Observe that w is well defined as $\{\sigma_k\}_{k=1}^N$ is the set of the first N (nonzero) eigenvectors of the Laplacian.

Relative to this choice of $w = w(\xi)$, $\bar\rho$ satisfies

$$\begin{aligned} &\partial_t \bar\rho + \mathbf{u}^0 \cdot \nabla \bar\rho + \bar{\mathbf{v}}^0 \cdot \nabla \theta^0 = \tilde{Ra} \cdot \bar v_3^0 + \Delta \bar\rho - \lambda P_N \bar\rho, \\ &\Delta \bar{\mathbf{v}}^0 = \nabla p + \mathrm{Ra}\hat{\mathbf{k}}\bar\rho, \quad \nabla \cdot \bar{\mathbf{v}}^0 = 0, \quad \bar\rho(0) = \xi. \end{aligned} \tag{B.10}$$

Testing (B.10) with $\bar\rho$ and $\bar{\mathbf{v}}^0$ respectively, and using that both $\mathbf{u}^0$ and $\bar{\mathbf{v}}^0$ are divergence free vector fields, we obtain

$$\frac{d}{dt}\|\bar\rho\|^2 + 2\|\nabla\bar\rho\|^2 + 2\lambda\|P_N\bar\rho\|^2 = 2\int_{\mathcal{D}} \left(\tilde{Ra}\bar v_3^0 - \bar{\mathbf{v}}^0 \cdot \nabla\theta^0\right)\bar\rho dx \tag{B.11}$$

and

$$\|\nabla \bar{\mathbf{v}}^0\| \le \mathrm{Ra}\|\bar\rho\|. \tag{B.12}$$

With standard Sobolev embeddings and (B.12) we have, for any $\eta > 0$,

$$\begin{aligned} \left|\int_{\mathcal{D}} \left(\tilde{Ra}\bar v_3^0 - \bar{\mathbf{v}}^0 \cdot \nabla\theta^0\right)\bar\rho dx\right| &\le \|\bar{\mathbf{v}}^0\|_{L^6}\|\nabla\theta^0\|\|\bar\rho\|_{L^3} + \tilde{Ra}\|\bar{\mathbf{v}}^0\|\|\bar\rho\| \\ &\le \tilde C\|\nabla\bar{\mathbf{v}}^0\|\|\nabla\theta^0\|\|\bar\rho\|^{1/2}\|\nabla\bar\rho\|^{1/2} + \tilde{Ra}\|\nabla\bar{\mathbf{v}}^0\|\|\bar\rho\| \\ &\le \tilde C Ra\|\nabla\theta^0\|\|\bar\rho\|^{3/2}\|\nabla\bar\rho\|^{1/2} + \tilde C Ra\tilde{Ra}\|\bar\rho\|^2 \\ &\le \|\nabla\bar\rho\|^2 + (\tilde C(Ra)^{4/3}\|\nabla\theta^0\|^{4/3} + \tilde C Ra\tilde{Ra})\|\bar\rho\|^2 \\ &\le \|\nabla\bar\rho\|^2 + (\eta\|\nabla\theta^0\|^2 + C)\|\bar\rho\|^2, \end{aligned} \tag{B.13}$$

where $C = C(\mathrm{Ra}, \tilde{Ra}, \eta) = \frac{\tilde C Ra^4}{\eta^2} + \tilde C Ra\tilde{Ra}$ and $\tilde C$ is a universal constant. Also since P_N and $-\Delta$ commute we have for $Q_N := I - P_N$

$$\begin{aligned} \|\nabla\bar\rho\|^2 &= -\langle P_N\bar\rho, \Delta P_N\bar\rho\rangle - \langle Q_N\bar\rho, \Delta Q_N\bar\rho\rangle = \|\nabla P_N\bar\rho\|^2 + \|\nabla Q_N\bar\rho\|^2 \\ &\ge \|\nabla Q_N\bar\rho\|^2 \ge \lambda_N\|Q_N\bar\rho\|^2, \end{aligned} \tag{B.14}$$

where the last inequality follows from the generalized Poincaré inequality. Choose $2\lambda = \lambda_N$ (with N to be chosen below) and combine (B.11) and (B.13) to infer

$$\frac{d}{dt}\|\bar\rho\|^2 + (\lambda_N - (\eta_0\|\nabla\theta^0\|^2 + C))\|\bar\rho\|^2 \le 0,$$

and hence, since $\bar\rho(0) = \xi$,

$$\|\bar\rho(t)\|^2 \le \|\xi\|^2 \exp\left(\eta_0\int_0^t \|\nabla\theta^0\|^2\, dr + (C - \lambda_N)t\right). \tag{B.15}$$

Applying Proposition A.1 we conclude that, for any $\theta_0^0 \in H$, and $\eta \in (0, \eta_0]$,

$$\mathbb{E}\|\bar\rho(t)\|^2 \le C\|\xi\|^2 \exp\left(\eta\|\theta_0^0\|^2 + (C + \eta - \lambda_N)t\right),$$

where $C = C(\mathrm{Ra}, \tilde{Ra})$ is independent of ξ and θ_0^0 and $t \ge 0$. By now choosing N large enough such that $\lambda_N > 2(C + \eta\|\sigma\|^2)$ we obtain

$$\mathbb{E}\|\bar\rho(t)\|^2 \le C\|\xi\|^2 \exp\left(\eta\|\theta_0^0\|^2 - \frac{\lambda_N}{2}t\right), \tag{B.16}$$

where $C = C(\mathrm{Ra}, \tilde{Ra})$ is independent of ξ and θ_0^0 and $t \ge 0$. This yields the first bound (B.6).

To obtain the second desired bound, (B.7), we use (B.16) to estimate

$$\mathbb{E}\int_0^\infty |w(\xi)|^2 dt = \|\sigma^{-1}\|^2 \lambda_N^2 \mathbb{E}\int_0^\infty \|P_N\bar{\rho}\|^2\, dt \leq C\exp(\eta\|\theta_0^0\|^2)\,,$$

where $C = C(\lambda_N, Ra, \tilde{Ra})$ is independent of θ_0^0 yielding (B.7). The bound (4.2) now follows.

Remark B.1. We can use the same argument leading to (B.15) to show that

$$\|\rho(t)\|^2 \leq \|\xi\|^2 \exp\left(\eta\int_0^t \|\nabla\theta^0\|^2\, dr + Ct\right)\,,$$

that is, for any $\eta > 0$,

$$\|\mathcal{J}_{0,t}\| \leq \exp\left(\eta\int_0^t \|\nabla\theta^0\|^2\, dr + Ct\right)\,, \tag{B.17}$$

where, as above, $C = C(Ra, \tilde{Ra}) = \frac{\tilde{C}Ra^4}{\eta^2} + Ra\tilde{Ra}$.

Remark B.2. Using Proposition A.2 and (B.17) we can easily establish the Lyapunov bound (4.1) with

$$C_1 = \exp\left(\frac{CRa^4 e^{t^*/2}}{\eta^2} + Ra\tilde{Ra}\right).$$

References

[1] Agrachev, A.A., Sarychev, A.V.: Navier–Stokes equations: controllability by means of low modes forcing. J. Math. Fluid Mech. **7**(1), 108–152 (2005)

[2] Agrachev, A.A., Sarychev, A.V.: Controllability of 2D Euler and Navier–Stokes equations by degenerate forcing. Commun. Math. Phys. **265**(3), 673–697 (2006)

[3] Ahlers, G., Grossmann, S., Lohse, D.: Heat transfer and large scale dynamics in turbulent Rayleigh–Bénard convection. Rev. Mod. Phys. **81**(2), 503 (2009)

[4] Barletta, A., Nield, D.A.: On the Rayleigh–Bénard–Poiseuille problem with internal heat generation. Int. J. Therm. Sci. **57**, 1–16 (2012)

[5] Bodenschatz, E., Pesch, W., Ahlers, G.: Recent developments in Rayleigh–Bénard convection. In: Annual review of fluid mechanics, vol. 32, pp. 709–778. Annual Reviews, Palo Alto, CA (2000)

[6] Breuer, M., Hansen, U.: Turbulent convection in the zero Reynolds number limit. Euro. Phys. Lett. **86**, 24004 (2009)

[7] Cerrai, S., Glatt-Holtz, N.E.: On the convergence of stationary solutions in the Smoluchowski–Kramers approximation of infinite dimensional systems (2018). arXiv preprint arXiv:1806.05319

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

[9] Constantin, P., Foias, C.: Navier–Stokes Equations, Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL (1988)

[10] Constantin, P., Glatt-Holtz, N., Vicol, V.: Unique ergodicity for fractionally dissipated, stochastically forced 2D Euler equations. Commun. Math. Phys. **330**(2), 819–857 (2014)

[11] Da Prato, G., Zabczyk, J.: Ergodicity for infinite-dimensional systems. In: London Mathematical Society Lecture Note Series, vol. 229. Cambridge University Press, Cambridge (1996)

[12] Da Prato, G., Zabczyk, J.: Stochastic equations in infinite dimensions. In: Encyclopedia of Mathematics and its Applications, vol. 44. Cambridge University Press, Cambridge (1992)

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

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

[15] Doob, J.L.: Asymptotic properties of Markoff transition prabablities. Trans. Am. Math. Soc. **63**, 393–421 (1948)

[16] Dudley, R.M.: Real Analysis and Probability, vol. 74. Cambridge University Press, Cambridge (2002)

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

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

[19] Foias, C., Manley, O., Rosa, R., Temam, R.: Navier–Stokes equations and turbulence Encyclopedia of Mathematics and its Applications, vol. 83. Cambridge University Press, Cambridge (2001)

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

[21] Földes, J., Glatt-Holtz, N., Richards, G., Whitehead, J.P.: Ergodicity in randomly forced Rayleigh–Bénard convection. Nonlinearity **29**(11), 3309 (2016)

[22] Földes, J., Friedlander, S., Glatt-Holtz, N., Richards, G.: Asymptotic analysis for randomly forced MHD. SIAM J. Math. Anal. **49**(6), 4440–4469 (2017)

[23] Glatt-Holtz, N.E., Herzog, D.P., Mattingly, J.C.: Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations. Ann. PDE **4**(2), 16 (2018)

[24] Goluskin, D., Spiegel, E.A.: Convection driven by internal heating. Phys. Lett. A **377**(1), 83–92 (2012)

[25] Hairer, M., Majda, A.J.: A simple framework to justify linear response theory. Nonlinearity **23**(4), 909 (2010)

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

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

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

[29] Hairer, M., Mattingly, J.C., Scheutzow, M.: Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. Probab. Theory Relat. Fields **149**(1–2), 223–259 (2011)

[30] Khas'minskii, R.Z.: Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations. Theory Probab. Appl. **5**(2), 179–196 (1960)

[31] Kuksin, S., Shirikyan, A.: Mathematics of Two-Dimensional Turbulence. Number 194 in Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge (2012)

[32] Lohse, D., Xia, K.Q.: Small-scale properties of turbulent Rayleigh–Bénard convection. Ann. Rev. Fluid Mech. **42**, 335–364 (2010)

[33] Lu, L., Doering, C., Busse, F.: Bounds on convection driven by internal heating. J. Math. Phys. **45**(7), 2967–2986 (2004)

[34] Manneville, P.: Rayleigh-Bénard convection: thirty years of experimental, theoretical, and modeling work. In: Mutabazi, I., Wesfreid, J.E., Guyon, E. (eds.) Dynamics of Spatio-Temporal Cellular Structures, Springer Tracts in Modern Physics, vol. 207, pp. 41–65. Springer, New York, NY (2006)

[35] Nourdin, I., Peccati, G.: Normal Approximations with Malliavin Calculus: From Stein's Method to Universality, vol. 192. Cambridge University Press, Cambridge (2012)

[36] Nualart, D.: Malliavin Calculus and its Applications, volume 110 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC (2009)

[37] Otto, F., Seis, C.: Rayleigh–Bénard convection: improved bounds on the Nusselt number. J. Math. Phys. **52**(8), 083702 (2011)

[38] Park, J.: Dynamic bifurcation theory of Rayleigh–Bénard convection with infinite Prandtl number. Discrete Contin. Dyn. Syst. **6**(3), 591 (2006)

[39] Roberts, P.H.: Convection in horizontal layers with internal heat generation. Theory J. Fluid Mech. **30**(01), 33–49 (1967)

[40] Temam, R.: Navier–Stokes equations: Theory and Numerical Analysis. AMS Chelsea Publishing, Providence, RI, (2001). Reprint of the 1984 edition

[41] Tritton, D.J., Zarraga, M.N.: Convection in horizontal layers with internal heat generation. Exp. J. Fluid Mech. **30**(01), 21–31 (1967)

[42] Villani, C.: Optimal Transport: Old and New, vol. 338. Springer, New York (2008)

[43] Wang, X.: A note on long time behavior of solutions to the Boussinesq system at large Prandtl number. In: Nonlinear Partial Differential Equations and Related Analysis, volume 371 of Contemporary Mathematics, pp. 315–323. American Mathematical Society, Providence, RI (2005)

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

[45] Wang, X.: Large Prandtl number behavior of the Boussinesq system of Rayleigh–Bénard convection. Appl. Math. Lett. **17**(7), 821–825 (2004)

[46] Wang, X.: Asymptotic behavior of the global attractors to the Boussinesq system for Rayleigh–Bénard convection at large Prandtl number. Commun. Pure Appl. Math. **60**(9), 1293–1318 (2007)

[47] Wang, X.: Bound on vertical heat transport at large Prandtl number. Phys. D **237**(6), 854–858 (2008)

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

[49] Weinan, E., Mattingly, J.C.: Ergodicity for the Navier–Stokes equation with degenerate random forcing: finite-dimensional approximation. Commun. Pure Appl. Math. **54**(11), 1386–1402 (2001)

[50] Weinan, E., Mattingly, J.C., Sinai, Y.G.: Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation. Commun. Math. Phys. **224**(1), 83–106 (2001). Dedicated to Joel L. Lebowitz

[51] Whitehead, J.P., Doering, C.R.: Internal heating driven convection at infinite Prandtl number. J. Math. Phys. **52**(9), 093101 (2011)

Juraj Földes
Department of Mathematics
University of Virginia
Charlottesville
USA
e-mail: foldes@virginia.edu

Nathan E. Glatt-Holtz
Department of Mathematics
Tulane University
New Orleans
USA
e-mail: negh@tulane.edu

Geordie Richards
Department of Mechanical and Aerospace Engineering
Utah State University
Logan
USA
e-mail: geordie.richards@usu.edu

Received: 8 January 2019.
Accepted: 27 September 2019.