

Asymptotic Analysis for Randomly Forced MHD

J. Földes, S. Friedlander, N. Glatt-Holtz, G. Richards

emails: juraj.foldes@ulb.ac.be, susanfri@usc.edu, negh@math.vt.edu, g.richards@rochester.edu

April 21, 2016

Abstract

We consider the three-dimensional magnetohydrodynamics (MHD) equations in the presence of a spatially degenerate stochastic forcing as a model for magnetostrophic turbulence in the Earth's fluid core. We examine the multi-parameter singular limit of vanishing Rossby number ε and magnetic Reynold's number δ , and establish that: (i) the limiting stochastically driven active scalar equation (with $\varepsilon = \delta = 0$) possesses a unique ergodic invariant measure, and (ii) any suitable sequence of statistically invariant states of the full MHD system converge weakly, as $\varepsilon, \delta \rightarrow 0$, to the unique invariant measure of the limit equation. This latter convergence result does not require any conditions on the relative rates at which ε, δ decay.

Our analysis of the limit equation relies on a recently developed theory of hypo-ellipticity for infinite-dimensional stochastic dynamical systems. We carry out a detailed study of the interactions between the nonlinear and stochastic terms to demonstrate that a Hörmander bracket condition is satisfied, which yields a contraction property for the limit equation in a suitable Wasserstein metric. This contraction property reduces the convergence of invariant states in the multi-parameter limit to the convergence of solutions at finite times. However, in view of the phase space mismatch between the small parameter system and the limit equation, and due to the multi-parameter nature of the problem, further analysis is required to establish the singular limit. In particular, we develop methods to lift the contraction for the limit equation to the extended phase space, including the velocity and magnetic fields. Moreover, for the convergence of solutions at finite times we make use of a probabilistic modification of the Grönwall inequality, relying on a delicate stopping time argument.

Keywords: *Ergodic Theory of Stochastic Partial Differential Equations (SPDEs), Magnetohydrodynamics (MHD), Singular Perturbation Theory, Wasserstein Distance, Degenerate Stochastic Forcing, Hypo-ellipticity*

MSC 2010 Classifications: *35Q86, 35R60, 35B25, 60H15*

Contents

1	Introduction	2
2	Preliminaries	6
3	Finite Time Convergence Analysis	8
4	Analysis of the Markovian Dynamics of the Formal Limit System	13
5	Convergence of Statistically Steady States	17
A	Uniform Moment Bounds	20

1 Introduction

It has long been appreciated that invariant measures of the equations of fluid dynamics provide a mathematically consistent framework for studying robust statistical quantities in turbulent flows. An ongoing challenge is therefore to investigate properties of these states such as the existence, uniqueness, ergodicity, and dependence on parameters in a variety of specific contexts. While one may certainly pose these questions for deterministic equations (cf. [FMRT01]), the stochastic setting, apart from its physical relevance, can be more accessible due to the regularizing effect of noise on the associated probability distribution functions.

In this article we will analyze a stochastically forced system of equations from magnetohydrodynamics that was presented by Moffat and Loper [ML94] as a model for magnetostrophic turbulence in the Earth’s fluid core. The stochastic forcing terms are to be interpreted as a source of heat which “continuously regenerates the statistically stationary temperature distribution throughout the core”; cf. [Mof08]. Our analysis will focus on the convergence of such statistically stationary states in a certain singular parameter limit suggested by [ML94] and which bears some formal similarities to a large Prandtl limit considered in our recent work [FGHR15]. See also [Wan04, Wan05, Wan07, Wan08, Par10] for other formally analogous limits considered within a deterministic framework.

The Stochastic Magnetohydrodynamics (MHD) Equations

The three-dimensional MHD equations govern the motion of a rotating, density stratified, electrically conducting fluid under the Boussinesq approximation. See e.g. [Mof78, Dav01]. In a rotating frame of reference these equations are written in terms of dimensionless variables as

$$\varepsilon(\partial_t U + U \cdot \nabla U) + \hat{\Omega} \times U = -\nabla P + \hat{B}_0 \cdot \nabla B + \delta B \cdot \nabla B - \Theta \hat{g} + \nu \Delta U, \quad \nabla \cdot U = 0, \quad (1.1)$$

$$\delta(\partial_t B + U \cdot \nabla B - B \cdot \nabla U) = \hat{B}_0 \cdot \nabla U + \Delta B, \quad \nabla \cdot B = 0, \quad (1.2)$$

$$d\Theta + U \cdot \nabla \Theta dt = \kappa \Delta \Theta dt + \sigma dW. \quad (1.3)$$

The unknowns are $U = U(x, t)$ the velocity, $B = B(x, t)$ the magnetic field (both vector-valued) and $\Theta = \Theta(x, t)$ the temperature (a scalar) of the fluid. Here $\hat{\Omega}$ denotes a unit vector along the axis of rotation of a sphere, \hat{g} is a unit vector in the local direction of gravity, which points radially inwards to the sphere, and \hat{B}_0 denotes a unit vector in the direction of a constant “applied” underlying magnetic field. We will work in Cartesian coordinates centered on a local tangent plane to the sphere at a co-latitude angle λ . We choose the Cartesian frame $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ such that $\hat{g} = -\hat{e}_3$ and $\hat{\Omega} = \cos \lambda \hat{e}_3 - \sin \lambda \hat{e}_2$.

The physical forces governing this system are the Coriolis force, Lorentz force, and gravity acting via buoyancy, while the equation for the temperature Θ is driven by a white-in-time, spatially correlated, Gaussian noise σdW . Specifically we consider

$$\sigma dW = \sum_{\substack{k \in \mathbb{Z}_0^3 \\ m \in \{0,1\}}} \alpha_{k,m} \sigma_k^m dW^{k,m}, \quad (1.4)$$

where $\{\sigma_k^m\}$ is a basis of $L^2(\mathbb{T}^3)$ consisting of eigenfunctions of Laplacian on the cubic torus, namely $\sigma_k^0(x) := \cos(k \cdot x)$, $\sigma_k^1(x) := \sin(k \cdot x)$, $\alpha_k \in \mathbb{R}$ are amplitudes, and $\{W^{k,m}\}$ is a collection of independent standard, 1D Brownian motions. We are mainly interested in the situation when (1.4) is ‘degenerate’ in the sense that only a few selected amplitudes α_k are non-zero. Nevertheless many of the results established below remain true (and in some cases are easier to prove) if infinitely many frequencies are activated ($\alpha_k \neq 0$) so long as the amplitudes decay sufficiently rapidly.

The non-dimensional parameters in (1.1)–(1.3) are ε the Rossby number, δ the magnetic Reynolds number, ν a (non-dimensional) viscosity and κ a (non-dimensional) thermal diffusivity. The orders of magnitude of the non-dimensional parameters are motivated by the physical postulates of the Moffatt and Loper model. In particular, the parameters ε , δ , ν , and κ are all small. Note that the ratio of the Coriolis to Lorentz forces in the model we are studying is of order 1, so for notational simplicity we have set this parameter, denoted by N^2 in [ML94], equal to 1.

For mathematical tractability we consider both (1.1)–(1.3) and (1.5)–(1.7) below on the periodic domain \mathbb{T}^3 , with all of the fields being mean free, a condition which is preserved by the equation.¹ Notice that we have followed the usual convention of including the divergence-free condition on the magnetic field B in (1.2). This is really a condition on the initial data, $\nabla \cdot B|_{t=0} = 0$, as this property is preserved by the dynamics of (1.1)–(1.3).

The analysis in Moffat and Loper [ML94] and Moffat [Mof08] assumes that the statistics of the temperature field Θ , averaged over scales characteristic of the turbulent regions, are prescribed. Further simplifications are made in [ML94, Mof08] based on the assumption that ε and δ are both small, and for their purposes can be neglected. In our work we carry the analysis to a higher level of complexity and examine the three-dimensional coupled system (1.1)–(1.3) in which a back reaction of the temperature field on the flow is allowed. We give a rigorous justification of the reduction of the full evolution system in the limit $\varepsilon, \delta \rightarrow 0$ to a stochastically driven active scalar equation by deriving suitable asymptotics in ε and δ at both finite and infinite time. The limit system when $\varepsilon = \delta = 0$ is the following active scalar equation:

$$d\theta + u \cdot \nabla \theta dt = \kappa \Delta \theta dt + \sigma dW, \quad (1.5)$$

where u, b are related to θ via

$$\hat{\Omega} \times u = -\nabla p + \hat{B}_0 \cdot \nabla b + \theta \hat{e}_3 + \nu \Delta u, \quad \nabla \cdot u = 0, \quad (1.6)$$

$$0 = \hat{B}_0 \cdot \nabla u + \Delta b. \quad (1.7)$$

Observe that the vector fields u and b are computed from the temperature θ and are not prescribed independent initial conditions. Using (1.6)–(1.7) we can define a linear operator $M = (M_u, M_b)$ with $(u, b) = (M_u(\theta), M_b(\theta))$ and it is crucially important in our analysis below that M produces two degrees of smoothing in space. Vector manipulations as in [ML94, FS15], taking the curl of (1.6) three times and using $\nabla \cdot u = 0 = \nabla \cdot b$, yield the expression:

$$\{[\nu \Delta^2 - (\hat{B}_0 \cdot \nabla)^2]^2 + (\hat{\Omega} \cdot \nabla)^2 \Delta\} u = -[\nu \Delta^2 - (\hat{B}_0 \cdot \nabla)^2] \nabla \times (\hat{e}_3 \times \nabla \theta) + (\hat{\Omega} \cdot \nabla) \Delta (\hat{e}_3 \times \nabla \theta). \quad (1.8)$$

Hence the explicit expression for the Fourier multiplier symbols $\mathcal{M}_u(k)$ and $\mathcal{M}_b(k)$ as functions of the Fourier variable $k = (k_1, k_2, k_3) \in \mathbb{Z}^3$ are

$$\mathcal{M}_u(k) = \frac{1}{D(k)} \left((\nu |k|^4 + (\hat{B}_0 \cdot k)^2) (k \times (\hat{e}_3 \times k)) + (\hat{\Omega} \cdot k) |k|^2 (\hat{e}_3 \times k) \right), \quad \mathcal{M}_b(k) = \frac{i \hat{B}_0 \cdot k}{|k|^2} \mathcal{M}_u(k), \quad (1.9)$$

where

$$D(k) = |k|^2 (\hat{\Omega} \cdot k)^2 + ((\hat{B}_0 \cdot k)^2 + \nu |k|^4)^2. \quad (1.10)$$

Since D is of order $|k|^8$ whereas the numerator of \mathcal{M}_u is of order $|k|^6$, the two degrees of smoothing of M is evident from (1.9)–(1.10).

It is worth emphasizing that the singular limit of (1.1)–(1.3) to (1.5)–(1.7) in this work bears some significant formal similarities to the infinite Prandtl limit for the stochastic Boussinesq equations which we have recently considered in [FGHR15]. Here however our problem involves multiple small parameters and both the linear and non-linear structure of the governing equations is quite different. On the other hand the parameters ν, κ in (1.1)–(1.3) are also small in practice suggesting other challenging singular perturbations problems of interest cf. [FRV14, FS15, FV11a, FV11b, FV12]. Moreover a variety of other interesting singular limit problems arise from (1.1)–(1.3) but most of the rigorous work so far is restricted to a deterministic setting. See e.g. [Wu97, DGG99, Mas07, CDGG06, CW11, FMRR14, CJTA15].

¹To verify this property, one first confirms that the mean zero condition is preserved in the Θ and B component, (1.3), as a straightforward consequence of Gauss' theorem, noting that σ is taken to be mean free in (1.4). Then, by integrating (1.1) and using Gauss' theorem, recalling that the buoyancy term has zero mean as derived already, one is left with the equation $\frac{d}{dt} V = \hat{\Omega} \times V$, where $V = \int_{\mathbb{T}^3} U dx$. But this ODE has only the trivial solution for initial condition $V(0) = 0$, yielding the preservation of the mean free condition for U .

Summary of Main Results

We now give a preview of the main results of this work. Precise statements are given below in the body of the work.

- (i) Consider any initial conditions $U_{\varepsilon,\delta}(0), B_{\varepsilon,\delta}(0), \Theta_{\varepsilon,\delta}(0)$, where $U_{\varepsilon,\delta}(0), B_{\varepsilon,\delta}(0)$ are uniformly bounded in $L^2(\mathbb{T}^3)$ independently of $\varepsilon, \delta > 0$ and

$$\|\Theta_{\varepsilon,\delta}(0) - \theta(0)\| \rightarrow 0, \quad \text{as } \varepsilon, \delta \rightarrow 0, \quad (1.11)$$

for some $\theta(0) \in L^2(\mathbb{T}^3)$. Here and in all that follows $\|\cdot\|$ denotes the usual L^2 -norm. Suppose that $(U_{\varepsilon,\delta}, B_{\varepsilon,\delta}, \Theta_{\varepsilon,\delta})$ and θ are corresponding solutions of (1.1)–(1.3) and (1.5) respectively. Then for any $t > 0$, there exists a sufficiently small $\gamma > 0$ such that for $p > \gamma$,

$$\mathbb{E} \sup_{s \in [0,t]} \|\Theta_{\varepsilon,\delta}(s) - \theta(s)\|^p \leq C(\|\Theta_{\varepsilon,\delta}(0) - \theta(0)\| + \varepsilon + \delta)^\gamma \rightarrow 0, \quad \text{as } \varepsilon, \delta \rightarrow 0, \quad (1.12)$$

for a constant $C > 0$, where γ and C are both independent of ε, δ and \mathbb{E} denotes the expected value (statistical mean). Furthermore we show that, for any $t > 0$,

$$\mathbb{E} \int_0^t \|(U_{\varepsilon,\delta}(s), B_{\varepsilon,\delta}(s)) - M(\theta)(s)\|^2 ds \leq C(\|\Theta_{\varepsilon,\delta}(0) - \theta(0)\| + \varepsilon + \delta)^\gamma \rightarrow 0, \quad \text{as } \varepsilon, \delta \rightarrow 0. \quad (1.13)$$

For the precise statement see Theorem 3.1 below.

- (ii) Regarding the active scalar equation (1.5) we show that, provided

$$\alpha_{e_1,m}, \alpha_{e_2,m}, \text{ and } \alpha_{e_3,m} \text{ are non-zero for each } m \in \{0, 1\}, \quad (1.14)$$

then (1.5) satisfies a form of the Hörmander bracket condition. See (4.4) and Theorem 4.3 below.² Invoking recent results of Hairer and Mattingly [HM08, HM11] (see also [HM06, FGHRT15]) we infer that, in a suitably chosen Wasserstein metric \mathfrak{W} ,

$$\mathfrak{W}(\mu^1 P_t, \mu^2 P_t) \leq C e^{-\gamma' t} \mathfrak{W}(\mu^1, \mu^2), \quad \text{for } t \geq 0, \quad (1.15)$$

for any probability measures μ^1, μ^2 , where the constants C, γ' are independent of $t \geq 0$ and μ^1, μ^2 . Here P_t is the Markov semigroup associated to (1.5) so that $\mu^i P_t$ represents the law of solutions to (1.5) at time $t \geq 0$ initially distributed as μ^i . In particular (1.15) immediately implies that, under the given condition (1.14), the equation (1.5) has a unique statistically stationary state μ . See Theorem 4.1 for further details.

- (iii) Assume the noise satisfies (1.14). We prove that, for every $\varepsilon, \delta > 0$, the system (1.1)–(1.3) possesses at least one statistically stationary state $\mu_{\varepsilon,\delta}$ which satisfies certain exponential moment bounds independently of ε, δ . Moreover, we show that any such collection $\mu_{\varepsilon,\delta}$ converges to μ in the limit as $\varepsilon, \delta \rightarrow 0$, at an algebraic rate, in a suitable Wasserstein metric; see Theorem 5.1 below for the precise statement. In particular, this implies that for any sufficiently regular observable ϕ ,

$$\left| \int \phi(U, B, \Theta) d\mu_{\varepsilon,\delta} - \int \phi(M(\theta), \theta) d\mu \right| \rightarrow 0, \quad \text{in the limit as } \varepsilon, \delta \rightarrow 0. \quad (1.16)$$

We now sketch our method of proof for the results described in (i)–(iii), highlighting some of the particular challenges we encountered. Observe that the convergence to the formal limit (1.5) as $\varepsilon, \delta \rightarrow 0$ is singular in the sense that there is a phase space mismatch. It is therefore apparent that our analysis must involve multiple time scales: an initial ‘layer’, that is, short times depending on ε and δ influenced by the difference in initial conditions, an intermediate time scale, and finally a large time scale where the loss of nonlinear

²While we provide complete details only under the assumption (1.14), in principle many other noise configurations could be addressed with our methods.

terms in the momentum and magnetic equations is especially apparent. This long time scale is embodied in the statistically stationary states of (1.1)–(1.3) and of (1.5) which are the primary focus of this work.

Our analysis builds on a method developed recently in [FGHR15], which draws on a simple but powerful observation from [HM08]: if one can establish a contraction property for the limit equation as in (1.15), then the question of convergence of statistically stationary states can be reduced to the convergence of solutions on finite time scales (cf. (1.12)–(1.13)). Nevertheless proving the results herein requires significant novel developments. This is due in part to the specific structural challenges inherent in (1.1)–(1.3). However, in the course of our work here, we develop general techniques to tackle multi-parameter problems as well as methods for lifting the convergence of statistics to an extended phase space.

As a first step we establish the convergence of solutions on finite time scales without imposing any relative rates for which ε, δ vanish. To the best of our knowledge, such convergence results are new for the MHD system, even in the deterministic setting, where our methods apply with straight-forward modifications. Here one of our key observations is that a difference in initial conditions for the velocity and magnetic components has a negligible effect on the temperature, namely of an algebraic order of $\varepsilon + \delta$. We also use non-trivial cancelations in the nonlinear terms to take advantage of the improved regularity properties of the limit equation; cf. (3.20) in Section 3.

Additional challenges arise due to the probabilistic nature of (1.1)–(1.3) and (1.4). For example, we find that estimates on the difference of solutions to the temperature equation lead to an integral inequality that requires certain almost sure bounds. This leads us to develop a probabilistic modification of the Grönwall inequality which relies on a delicate stopping time argument. See (3.22) and Lemma 3.4 below. In particular this approach leads to algebraic rates of convergence in temperature, as in (1.12). We are then able to transfer these algebraic rates to the convergence of the velocity and magnetic fields, see (1.13). Note that, due to the existence of the initial layer, one should not expect the convergence as $\varepsilon, \delta \rightarrow 0$ to be uniform in the U and B components up to the initial time.

In order to obtain more precise asymptotics, one could seek to extend classical approaches involving multiple time scales to our setting; see for example [O’M74, KP03]. Such an approach would yield a ‘corrector’, or in our multi-parameter case a series of correctors, whose structure depends on the relative sizes of ε and δ . These correctors provide well-posed effective dynamics which accurately approximate the full system (1.1)–(1.3) up to time zero when ε and δ are small. See Remark 3.5 below. While the accuracy of these correctors can be rigorously established with estimates similar to those carried out in detail in Section 3, our approach, which is focused on the convergence of long time statistics, does not require such ‘intermediate’ systems, so we omit these details. For results in a similar setting see [FGHR15] and e.g. [Wan04] for an analogous situation in a deterministic context.

Having established suitable finite time convergence results we turn to analyze the limit equation (1.5) in pursuit of (1.15). Here we underline that a proof of the estimate (1.15) becomes tractable by including a stochastic forcing. The estimate (1.15) immediately implies that statistically invariant states of (1.5) are unique and ergodic, providing a rigorous foundation for some basic assumptions in statistical theories of turbulence; cf. [Fri95] or [Mof08] in our geophysical setting. Such considerations from turbulence motivate the study of unique ergodicity in a variety of infinite-dimensional stochastic systems arising in fluids; see e.g. [FM95, HM06, BDP07, Deb11, KS12, FGHRT15, CGHV13, FGHV14] and references therein.

Motivated by the fundamental postulates of turbulence where energy cascades from large to small spatial scales, we consider (1.5) driven by a ‘spectrally degenerate’ stochastic forcing, namely where the noise acts through a narrow range of frequencies. While this degenerate setting may be attractive physically, it makes the analysis leading to (1.15) substantially more difficult compared to the case when noise is active on all spatial scales. To tackle this situation we implement a strategy from the recent groundbreaking works of Hairer and Mattingly [HM06, HM08, HM11]. These works develop a theory of hypo-ellipticity for infinite-dimensional systems which allows one to establish unique ergodicity and contractivity for certain semilinear stochastic PDEs with spectrally degenerate stochastic forcing. Application of this theory requires demonstrating that the stochastic excitation can propagate to higher frequencies through the nonlinearity. This is done by verifying that a Hörmander bracket condition is satisfied by the set of activated modes. Such an analysis requires a detailed understanding of the nonlinear structure in the governing equations, as observed in previous works [EM01, Rom04, HM06, HM11, FGHRT15, FGHV14]. In our setting we rely on the explicit formulation of M in (1.9)–(1.10) and a non-trivial construction to avoid degeneracies in frequency space.

The final stage of our analysis demonstrates that the unique invariant statistics for the limit equation (1.5) approximate any reasonable invariant statistics for the original system (1.1)–(1.3) for ε, δ small. Indeed, as we already identified above, the bounds (1.12)–(1.13) and (1.15) provide a means to transfer the algebraic rates in (1.12) to the convergence of invariant states in a suitable Wasserstein metric following [HM08, FGHR15]. In particular, such Wasserstein metrics are closely related to the topology of weak convergence and hence our results imply that observations from a reasonable class of invariant states for (1.1)–(1.3) converge to observations relative to the unique invariant measure for (1.5) in the limit as $\varepsilon, \delta \rightarrow 0$; see main result (iii) above.

It should be emphasized that in our previous work [FGHR15] on the stochastic Boussinesq system, convergence of statistically invariant states on the temperature component (in the infinite Prandtl number limit) was established by taking a similar approach. In contrast, in this current work we prove convergence of invariant states on the extended phase space, including the velocity and magnetic fields, which requires us to overcome a novel challenge arising due to the phase space mismatch between the small parameter system (1.1)–(1.3) and the limit equation (1.5). This is accomplished by exploiting the structure of a Wasserstein distance and estimates on the constitutive law (1.9)–(1.10) in order to ‘lift’ the contractive property for (1.5) to the extended phase space. This strategy should prove useful for other systems with analogous singular limits, in particular for the Boussinesq system.

Organization of the Paper

The rest of the manuscript is organized as follows. In Section 2 we establish our notational conventions and recall some mathematical foundations of the equations (1.1)–(1.3) and (1.5); existence and uniqueness of solutions, and the associated Markovian framework. In Section 3 we prove the finite time convergence of solutions to (1.1)–(1.3) to those of (1.5) in the limit as $\varepsilon, \delta \rightarrow 0$, as described in point (i) above. Section 4 contains some background on Wasserstein metrics, provides a proof of the contractive property (1.15) as discussed in point (ii), and the uniqueness of statistically stationary states for (1.5) follows as a consequence. To prove (1.15) we carry out the relevant infinite-dimensional Lie bracket analysis to demonstrate that a form of the Hörmander condition is satisfied. The proof of weak convergence of statistically stationary states, as described in point (iii) above, is given in Section 5. Finally, we include an appendix with parameter independent exponential moment bounds which will be utilized throughout the manuscript.

2 Preliminaries

We begin by recalling some details about the functional setting of (1.1)–(1.3) and the formal limit equation (1.5).

The main function spaces are

$$H := \left\{ U \in L^2(\mathbb{T}^3)^3 : \nabla \cdot U = 0, \int_{\mathbb{T}^3} U \, dx = 0 \right\} \times \left\{ B \in L^2(\mathbb{T}^3)^3 : \nabla \cdot B = 0, \int_{\mathbb{T}^3} B \, dx = 0 \right\} \\ \times \left\{ \Theta \in L^2(\mathbb{T}^3) : \int_{\mathbb{T}^3} \Theta \, dx = 0 \right\},$$

and

$$V := H \cap \left\{ U \in H^1(\mathbb{T}^3)^3 \right\} \times \left\{ B \in H^1(\mathbb{T}^3)^3 \right\} \times \left\{ \Theta \in H^1(\mathbb{T}^3) \right\},$$

where $H^s(\mathbb{T}^3)$ are the usual Sobolev spaces. For the limit problem (1.5) we take $H' = L^2(\mathbb{T}^3)$ and $V' = H^1(\mathbb{T}^3)$ with the mean free condition imposed. We denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the usual L^2 -norm and inner product, respectively. For any $p \geq 2$ we set

$$\|\sigma\|_{L^p}^p := \int_{\mathbb{T}^2} \left(\sum_{\substack{k \in \mathbb{Z}_0^3 \\ m=0,1}} |\alpha_{k,m} \sigma_k^m(x)|^2 \right)^{p/2} dx,$$

so that $\|\sigma\|_{L^2} = \|\sigma\|$ reduces to the usual Hilbert-Schmidt norm. Note that $\|\sigma\|$ is controlled by $\|\sigma\|_{L^p}$ for every $p \geq 2$.

We have the following existence results for (1.1)–(1.3):

Proposition 2.1 (Martingale solutions for the full system). *Consider (1.1)–(1.3) with any $\varepsilon, \delta > 0$ and where (1.4) holds with only finitely many $\alpha_{k,m} \neq 0$ (or where the $\alpha_{k,m}$ decay sufficiently rapidly as $|k| \rightarrow \infty$),*

- (i) *Given any initial probability distribution $\mu \in \text{Pr}(H)$ there exists a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0}, W)$ ³ and a corresponding stochastic process (U, B, Θ) with*

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

which is weakly continuous, adapted to the filtration provided by the stochastic basis, (weakly) solves (1.1)–(1.3), and such that $\text{Law}(U(0), B(0), \Theta(0)) = \mu$. We say that $(\mathcal{S}, (U, B, \Theta))$ is a Martingale solution of (1.1)–(1.3).

- (ii) *Moreover, for every $\varepsilon, \delta > 0$, there exists a stationary Martingale solution $(\mathcal{S}, (U_S, B_S, \Theta_S))$ such that this family of solutions satisfies the uniform moment bound*

$$\sup_{\varepsilon, \delta \in (0, N]} \mathbb{E} \exp(\eta(\varepsilon \|U_S\|^2 + \delta \|B_S\|^2 + \|\Theta_S\|_{L^3}^2)) \leq C_N < \infty, \quad (2.1)$$

for every $N > 0$, where $\eta, C_N > 0$ are independent of $\varepsilon, \delta > 0$.

Proposition 2.1, (i) can be established in a similar manner as for the 3D Navier stokes equations, namely through a Galerkin regularization procedure, see e.g. [AFS08]. Regarding the existence of stationary Martingale solutions, Proposition 2.1, (ii), the proof proceeds exactly as in [FGHRW15] using Lemmas A.1, A.6 in the Appendix below. Here again the existence of stationary solutions follows as a limit of invariant states of a regularization of the governing equations following [AFS08], but with an additional twist. The idea is to consider a Galerkin truncation only in the velocity and magnetic components of (1.1)–(1.3) (cf. [FGHRW15]). This allows the regularization procedure to preserve the advection-diffusion structure in (1.3) which in turn permits the ε, δ -uniform bound on $L^3(\mathbb{T}^3)$ norms of Θ in (2.1) (or for that matter on $L^p(\mathbb{T}^3)$ for any $p \geq 2$). This uniform bound is needed for Theorems 3.1 and 5.1 below.

The formal limit equation is much more tractable analytically, and possesses unique, pathwise solutions.

Proposition 2.2 (Well-Posedness for the limit equation). *Consider (1.5) supplemented with (1.6), (1.7) and subject to the same condition on σ given by (1.4) as in Proposition 2.1.*

- (i) *Fix a stochastic basis \mathcal{S} and any initial condition $\theta_0 \in L^2(\Omega; H')$ which is \mathcal{F}_0 measurable. Then there exists a unique*

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

which is \mathcal{F}_t -adapted, solves (1.5) and satisfies the initial condition $\theta(0) = \theta_0$. We say that θ is a Pathwise solution of (1.5).

- (ii) *Writing $\theta(t, \theta_0)$ for the solution with initial condition $\theta_0 \in H'$ at time $t \geq 0$, we have for any $t \geq 0$*

$$\theta(t, \theta_0^n) \rightarrow \theta(t, \theta_0) \text{ a.s. in } H' \text{ whenever } \theta_0^n \rightarrow \theta_0 \text{ in } H'.$$

- (iii) *Pathwise solutions θ of (1.5) satisfy the exponential moment bounds (A.1)–(A.3).*

This well-posedness result for (1.5) is again a straightforward extension of known results in the deterministic case, cf. [FS15], particularly since we are working with an additive noise. For discussion of the exponential moment bounds, see Lemma A.1 in the Appendix.

³Recall that a stochastic basis consists of a probability space, a right-continuous, complete filtration $\{\mathcal{F}_t\}_{t \geq 0}$ along with a collection $W = \{W^{k,m}\}$ of independent, identically distributed 1d Brownian motions adapted to this filtration.

We finally recall the Markovian setting of (1.5) as follows. Denote by $\text{Pr}(H')$ the space of Borel probability measures on H' . Also denote by $M_b(H')$ and $C_b(H')$ the bounded measurable and bounded continuous (real valued) functions on H' , respectively. The Markov transition functions are defined by

$$P_t(\theta_0, A) = \mathbb{P}(\theta(t, \theta_0) \in A) \quad \text{for any } t \geq 0, \theta_0 \in H', A \in \mathcal{B}(H'),$$

where $\mathcal{B}(H')$ is the set of Borel subsets of H' and the associated semigroup $\{P_t\}_{t \geq 0}$ acts on ‘observables’ $\phi \in M_b(H')$ according to

$$P_t \phi(\theta_0) := \int_{H'} \phi(\theta) P_t(\theta_0, d\theta) = \mathbb{E} \phi(\theta(t, \theta_0)),$$

and on ‘initial distributions’ $\mu \in \text{Pr}(H')$ by

$$\mu P_t(A) := \int_{H'} P_t(\theta, A) d\mu(\theta).$$

Recall that $\mu \in \text{Pr}(H')$ is an invariant measure for P_t , i.e. an invariant measure for (1.5), if $\mu P_t = \mu$ for all $t \geq 0$. Since $\theta_0 \mapsto \theta(t, \theta_0)$ is continuous by Proposition 2.2 for each $t \geq 0$, we have that P_t is *Feller*, mapping $C_b(H')$ to itself.

Recall that the Krylov-Bogolyubov procedure and the compact embedding of H^1 into L^2 immediately produce the existence of an invariant measure for (1.5)–(1.7), while the uniqueness, a much more delicate issue, is addressed below in Section 4. Using the moment bound (A.3) we infer that, for any invariant measure and any $p \geq 2$

$$\int \|\theta\|_{L^p} d\mu(\theta) \leq C < \infty, \quad (2.2)$$

where the constant C is as in (2.1) above.

3 Finite Time Convergence Analysis

In this section we establish the finite time convergence of solutions of (1.1)–(1.3) to solutions of (1.5) in the limit as $\varepsilon, \delta \rightarrow 0$, see (i) in the summary of our main results. We again emphasize that no relative rates between ε and δ are required here. In Section 5 these results will be used to establish the convergence of stationary solutions (see Theorem 5.1 below). The section concludes with some remarks concerning ‘corrector systems’ which provide a different well-posed approximation to (1.1)–(1.3), namely an approximation that is accurate up to time zero.

Note that we are working only with zero mean functions and vector fields. This restriction justifies the use of the Poincaré inequality which we apply below without further comment.

Theorem 3.1. *For every $\varepsilon, \delta \in (0, 1]$ let (U, B, Θ) be a Martingale solution of (1.1)–(1.3) in the sense of Proposition 2.1 and let θ be a solution of (1.5) as in Proposition 2.2. Suppose that there is $\eta_0 > 0$ such that*

$$\sup_{\varepsilon, \delta \in (0, 1]} \mathbb{E} \exp(\eta_0(\varepsilon \|U(0)\|^2 + \delta \|B(0)\|^2 + \|\Theta(0)\|_{L^3}^2 + \|\theta(0)\|_{L^3}^2)) \leq C_0 < \infty. \quad (3.1)$$

Then there exists $\eta_1 = \eta_1(\eta_0, \nu, \kappa, C_0)$ such that for each $T > 0$, $\eta \in (0, \eta_1]$, $\gamma \in (0, \eta/(\eta + CT)]$, and $p > \gamma$, there is a constant $C = C(p, \eta_0, \nu, \kappa, C_0)$ such that

$$\mathbb{E} \sup_{t \in [0, T]} \|\Theta(t) - \theta(t)\|^p \quad (3.2)$$

$$\leq C \exp(\eta T \|\sigma\|_{L^3}^2) \left((\varepsilon + \delta)^\gamma + (\mathbb{E} \|\Theta(0) - \theta(0)\|^2 + \varepsilon \|U(0) - M_u(\theta)(0)\|^2 + \delta \|B(0) - M_b(\theta)(0)\|^2)^\gamma \right)^{1/2},$$

and

$$\mathbb{E} \int_0^T (\|U(t) - M_u(\theta)(t)\|_{H^1}^2 + \|B(t) - M_b(\theta)(t)\|_{H^1}^2) dt \quad (3.3)$$

$$\leq CT \exp(\eta T \|\sigma\|_{L^3}^2) \left((\varepsilon + \delta)^\gamma + (\mathbb{E} \|\Theta(0) - \theta(0)\|^2 + \varepsilon \|U(0) - M_u(\theta)(0)\|^2 + \delta \|B(0) - M_b(\theta)(0)\|^2)^\gamma \right)^{1/2}.$$

Remark 3.2. *As identified in Proposition 2.1, (3.1) holds for a collection of stationary Martingale solutions of (1.1)–(1.3). The bound (3.1) holds for the unique stationary state of (1.5) due to (2.2). These observations will play a crucial role below in Section 5.*

Before proceeding to the proof we set notations for several operators that are used extensively. We denote the Coriolis-Stokes operator by A which acts on sufficiently smooth, divergence free vector fields according to

$$Au := -\nu\Delta u + P\hat{\Omega} \times u, \quad (3.4)$$

where P is the Leray projector onto divergence free vector fields. Note that P commutes with Δ , and therefore we omit P in front of the Laplacian term in (3.4). Since $u \mapsto \hat{\Omega} \times u$ is a skew symmetric operator, then $\langle P\hat{\Omega} \times u, u \rangle = \langle \hat{\Omega} \times u, u \rangle = 0$, and therefore A is positive definite on zero mean functions with

$$\langle Au, u \rangle = \nu\|\nabla u\|^2, \quad \text{for all } u \in H^1(\mathbb{T}^3)^3 \text{ with } \nabla \cdot u = 0 \quad (3.5)$$

which we use below without further comment. We also make use of the operator Q defined as

$$Qu := Au - (-\Delta)^{-1}(\hat{B}_0 \cdot \nabla)^2 u, \quad (3.6)$$

where $(-\Delta)^{-1}$ has the range of zero mean functions. Note that

$$\langle Qu, u \rangle \geq \langle Au, u \rangle = \nu\|\nabla u\|^2, \quad \text{for all } u \in H^1(\mathbb{T}^3)^3 \text{ with } \nabla \cdot u = 0,$$

so that, in particular, Q is positive definite. Thus A and Q are invertible operators, and we may rewrite the limit equation (1.6)–(1.7) as

$$u = Q^{-1}P\hat{e}_3\theta, \quad b = (-\Delta)^{-1}(\hat{B}_0 \cdot \nabla u). \quad (3.7)$$

Remark 3.3. *Notice that by (3.7), u and b are respectively two and three degrees smoother than θ . By this we mean that for any $s \in \mathbb{R}$ if $\theta \in H^s$, then $u \in H^{s+2}$ and $b \in H^{s+3}$, as can be seen from corresponding symbols given explicitly in (1.9).*

Proof of Theorem 3.1. In what follows C denotes any constant which depends on κ, ν , and C_0 , but which is independent of $\varepsilon, \delta, K, \sigma$, and $t > 0$.

Working from (3.7), (1.5) we have

$$du = Q^{-1}P\hat{e}_3d\theta = Q^{-1}P\hat{e}_3(\kappa\Delta\theta - u \cdot \nabla\theta)dt + Q^{-1}P\hat{e}_3\sigma dW. \quad (3.8)$$

Observe (1.6) takes the form $Au = P\hat{B}_0 \cdot \nabla b + P\hat{e}_3\theta$ and adding it to an ε multiple of (3.8) we obtain

$$\varepsilon du + Audt = (P\hat{B}_0 \cdot \nabla b + P\hat{e}_3\theta)dt + \varepsilon Q^{-1}P\hat{e}_3(\kappa\Delta\theta - u \cdot \nabla\theta)dt + \varepsilon Q^{-1}P\hat{e}_3\sigma dW. \quad (3.9)$$

Next, we derive an evolution equation for b . From (3.7), (1.5) we obtain that

$$db = (-\Delta)^{-1}\hat{B}_0 \cdot \nabla du = R(\kappa\Delta\theta - u \cdot \nabla\theta)dt + R\sigma dW, \quad (3.10)$$

where

$$R := (-\Delta)^{-1}(\hat{B}_0 \cdot \nabla)Q^{-1}P\hat{e}_3. \quad (3.11)$$

Adding a δ -multiple of (3.10) to (1.7) yields

$$\delta db - \Delta bdt = (\hat{B}_0 \cdot \nabla u)dt + \delta R(\kappa\Delta\theta - u \cdot \nabla\theta)dt + \delta R\sigma dW. \quad (3.12)$$

Set

$$v := U - u, \quad f := B - b, \quad \phi := \Theta - \theta, \quad (3.13)$$

where (U, B, Θ) is solution of (1.1)–(1.3). Comparing (3.9) and (1.1) we obtain

$$\varepsilon dv + Avdt = (P\hat{B}_0 \cdot \nabla f + P\hat{e}_3\phi)dt + P(\delta B \cdot \nabla B - \varepsilon U \cdot \nabla U)dt - \varepsilon Q^{-1}P\hat{e}_3(\kappa\Delta\theta - u \cdot \nabla\theta)dt - \varepsilon Q^{-1}P\hat{e}_3\sigma dW. \quad (3.14)$$

Similarly by subtracting (3.12) from (1.2) we have

$$\delta df - \Delta f dt = \hat{B}_0 \cdot \nabla v dt - \delta(U \cdot \nabla B - B \cdot \nabla U) dt - \delta R(\kappa \Delta \theta - u \cdot \nabla \theta) dt - \delta R \sigma dW. \quad (3.15)$$

Next, comparing (1.3) and (1.5) we obtain

$$\partial_t \phi - \kappa \Delta \phi = -U \cdot \nabla \phi - v \cdot \nabla \theta. \quad (3.16)$$

Applying the Itô formula to (3.14) and recalling (3.5) yields

$$\begin{aligned} \frac{\varepsilon}{2} d\|v\|^2 + \nu \|\nabla v\|^2 dt &= \left(\langle \hat{B}_0 \cdot \nabla f, v \rangle + \langle \hat{e}_3 \phi, v \rangle \right) dt - \varepsilon \langle U \cdot \nabla U, v \rangle dt + \delta \langle B \cdot \nabla B, v \rangle dt \\ &\quad - \varepsilon \langle Q^{-1} P \hat{e}_3 (\kappa \Delta \theta - u \cdot \nabla \theta), v \rangle dt + \frac{\varepsilon}{2} \|Q^{-1} P \hat{e}_3 \sigma\|^2 dt - \varepsilon \langle Q^{-1} P \hat{e}_3 \sigma, v \rangle dW. \end{aligned} \quad (3.17)$$

Similarly (3.15) yields

$$\begin{aligned} \frac{\delta}{2} d\|f\|^2 + \|\nabla f\|^2 dt &= \langle \hat{B}_0 \cdot \nabla v, f \rangle dt - \delta \langle U \cdot \nabla B - B \cdot \nabla U, f \rangle dt - \delta \langle R(\kappa \Delta \theta - u \cdot \nabla \theta), f \rangle dt \\ &\quad + \frac{\delta}{2} \|R \sigma\|^2 dt - \delta \langle R \sigma, f \rangle dW. \end{aligned} \quad (3.18)$$

Finally from (3.16) and the fact that $\nabla \cdot v = \nabla \cdot U = 0$,

$$\frac{1}{2} \frac{d}{dt} \|\phi\|^2 + \kappa \|\nabla \phi\|^2 = \langle v \cdot \nabla \phi, \theta \rangle. \quad (3.19)$$

Addition of (3.17) and (3.18), and integration in time up to $t \wedge \tau$, where τ is any stopping time, gives

$$\begin{aligned} \frac{1}{2} (\varepsilon \|v(t \wedge \tau)\|^2 + \delta \|f(t \wedge \tau)\|^2) + \int_0^{t \wedge \tau} (\nu \|\nabla v\|^2 + \|\nabla f\|^2) ds &= \frac{1}{2} (\varepsilon \|v(0)\|^2 + \delta \|f(0)\|^2) \\ &\quad + \int_0^{t \wedge \tau} \langle \hat{e}_3 \phi, v \rangle ds - \varepsilon \int_0^{t \wedge \tau} \langle U \cdot \nabla U, v \rangle ds + \delta \int_0^{t \wedge \tau} (\langle B \cdot \nabla B, v \rangle - \langle U \cdot \nabla B - B \cdot \nabla U, f \rangle) ds \\ &\quad + \varepsilon \int_0^{t \wedge \tau} \left(\frac{1}{2} \|Q^{-1} P \hat{e}_3 \sigma\|^2 - \langle Q^{-1} P \hat{e}_3 (\kappa \Delta \theta - u \cdot \nabla \theta), v \rangle \right) ds \\ &\quad + \delta \int_0^{t \wedge \tau} \left(\frac{1}{2} \|R \sigma\|^2 - \langle R(\kappa \Delta \theta - u \cdot \nabla \theta), f \rangle \right) ds - \int_0^{t \wedge \tau} (\varepsilon \langle Q^{-1} P \hat{e}_3 \sigma, v \rangle + \delta \langle R \sigma, f \rangle) dW. \end{aligned}$$

Here we note that W is the same process in both (3.17), (3.18), and we have used that $(\hat{B}_0 \cdot \nabla)$ is anti-symmetric to cancel terms. Next, we estimate terms on the right hand side. First,

$$|\langle \hat{e}_3 \phi, v \rangle| \leq \|\phi\| \|v\| \leq \frac{\nu}{8} \|\nabla v\|^2 + \frac{C}{\nu} \|\phi\|^2.$$

Since $\langle U \cdot \nabla v, v \rangle = 0$, u is two degrees smoother than θ (by (3.7)), and using the embedding $H^2 \hookrightarrow L^\infty$, we have

$$\begin{aligned} |\langle U \cdot \nabla U, v \rangle| &= |\langle U \cdot \nabla u, v \rangle| = |\langle U \cdot \nabla v, u \rangle| \leq \|U\| \|\nabla v\| \|u\|_{L^\infty} \leq C \|U\| \|\nabla v\| \|u\|_{H^2} \leq C \|U\| \|\nabla v\| \|\theta\| \\ &\leq \frac{\nu}{8\varepsilon} \|\nabla v\|^2 + \frac{C\varepsilon}{\nu} \|U\|^2 \|\theta\|^2. \end{aligned} \quad (3.20)$$

Next observe, again using (3.13) and that all of the vector field are divergence free,

$$\begin{aligned} \langle B \cdot \nabla B, v \rangle - \langle U \cdot \nabla B - B \cdot \nabla U, f \rangle &= \langle B \cdot \nabla f, v \rangle + \langle B \cdot \nabla b, v \rangle - \langle U \cdot \nabla b - B \cdot \nabla v - B \cdot \nabla u, f \rangle \\ &= \langle B \cdot \nabla b, v \rangle - \langle U \cdot \nabla b - B \cdot \nabla u, f \rangle \\ &= -\langle B \cdot \nabla v, b \rangle + \langle U \cdot \nabla f, b \rangle - \langle B \cdot \nabla f, u \rangle, \end{aligned}$$

and consequently, using again that both u and b are two degrees smoother than θ , similarly to (3.20), we have

$$\begin{aligned} |\langle B \cdot \nabla B, v \rangle - \langle U \cdot \nabla B - B \cdot \nabla U, f \rangle| &\leq C(\|B\| \|\nabla v\| \|b\|_{H^2} + \|U\| \|\nabla f\| \|b\|_{H^2} + \|B\| \|\nabla f\| \|u\|_{H^2}) \\ &\leq C(\|B\| \|\nabla v\| \|\theta\| + \|U\| \|\nabla f\| \|\theta\| + \|B\| \|\nabla f\| \|\theta\|) \\ &\leq \frac{1}{8\delta}(\nu \|\nabla v\|^2 + \|\nabla f\|^2) + \frac{C\delta}{\nu \wedge 1} \|\theta\|^2 (\|U\|^2 + \|B\|^2). \end{aligned}$$

Next observe that the range of Q^{-1} consists of divergence free vector fields, and consequently

$$\langle Q^{-1} P \hat{e}_3 (u \cdot \nabla \theta), v \rangle = \langle u \cdot \nabla \theta, \hat{e}_3 \cdot Q^{-1} v \rangle = -\langle u \cdot \nabla (\hat{e}_3 \cdot Q^{-1} v), \theta \rangle$$

Recalling that Q^{-1} provides two degrees of smoothing, and using again $H^2 \hookrightarrow L^\infty$, we thus obtain

$$|\langle Q^{-1} P \hat{e}_3 (\kappa \Delta \theta - u \cdot \nabla \theta), v \rangle| \leq C \|\theta\| \|v\| + C \|v\| \|\theta\| \|u\|_{H^2} \leq \frac{\nu}{8\varepsilon} \|\nabla v\|^2 + \frac{C\varepsilon}{\nu} (\|\theta\|^2 + \|\theta\|^4).$$

Analogously since R has also two (in fact three, cf. (3.11)) degrees of smoothing,

$$|\langle R(\kappa \Delta \theta - u \cdot \nabla \theta), f \rangle| \leq C \|\theta\| \|f\| + C \|f\| \|\theta\| \|u\|_{H^2} \leq \frac{1}{8\delta} \|\nabla f\|^2 + C\delta (\|\theta\|^2 + \|\theta\|^4).$$

Finally, we trivially have

$$\|R\sigma\|^2 \leq C\|\sigma\|^2, \quad \|Q^{-1} P \hat{e}_3 \sigma\|^2 \leq C\|\sigma\|^2.$$

Combining all of the preceding estimates we infer

$$\begin{aligned} \frac{1}{2} \int_0^{t \wedge \tau} (\nu \|\nabla v\|^2 + \|\nabla f\|^2) ds &\leq \frac{1}{2} (\varepsilon \|v(0)\|^2 + \delta \|f(0)\|^2) + \frac{C}{\nu} \int_0^{t \wedge \tau} \|\phi\|^2 ds \\ &+ C(\varepsilon^2 + \delta^2) \left(1 + \frac{1}{\nu}\right) \int_0^{t \wedge \tau} \|\theta\|^2 (\|U\|^2 + \|B\|^2 + \|\theta\|^2 + 1) ds + \frac{t \wedge \tau}{2} \|\sigma\|^2 (\varepsilon + \delta) + M_{t \wedge \tau}, \end{aligned} \quad (3.21)$$

where M_t represents a Martingale term. Working from (3.19) we have

$$\frac{1}{2} \frac{d}{dt} \|\phi\|^2 + \kappa \|\nabla \phi\|^2 \leq \|v\|_{L^6} \|\nabla \phi\| \|\theta\|_{L^3} \leq \kappa \|\nabla \phi\|^2 + \frac{C}{\kappa} \|\nabla v\|^2 \|\theta\|_{L^3}^2,$$

and therefore

$$\sup_{s \in [0, t \wedge \tau]} \|\phi(s)\|^2 \leq \|\phi(0)\|^2 + \frac{C}{\kappa} \left(\sup_{s \in [0, t \wedge \tau]} \|\theta(s)\|_{L^3}^2 \right) \int_0^{t \wedge \tau} \|\nabla v\|^2 ds. \quad (3.22)$$

We will now implement a probabilistic modification of the Grönwall inequality.

Lemma 3.4. *Assume that an \mathcal{F}_t -adapted stochastic process $(\mathcal{X}(t))_{t \geq 0}$ satisfies, for each $K > 0$,*

$$\mathbb{E} \mathcal{X}(t \wedge \tau_K) \leq K\mathcal{T} + CK \int_0^t \mathbb{E} \mathcal{X}(s \wedge \tau_K) ds + K^2 C(\varepsilon + \delta)(1 + \|\sigma\|^2 t), \quad (3.23)$$

where $\mathcal{T}, \varepsilon, \delta > 0$ are constants, and $\{\tau_K\}$ is a collection of stopping times which satisfies

$$\mathbb{P}(\tau_K \leq t) \leq C' e^{-\eta K},$$

for some $\eta > 0$. Then for any $t > 0$ and $\gamma \leq \frac{\eta}{\eta + Ct}$ one has

$$\mathbb{E}(\mathcal{X}(t))^\gamma \leq C_1 (\mathcal{T}^\gamma + (\varepsilon + \delta)^\gamma), \quad (3.24)$$

where $C_1 = C_1(C', t, \|\sigma\|, \gamma) > 0$.

We postpone the proof of Lemma 3.4 to the end of this section. To implement Lemma 3.4, we define the stopping times

$$\tau_K := \inf_{s \geq 0} \{ \|\theta(s)\|_{L^3}^2 \geq K \}, \quad (3.25)$$

for any $K > 0$. Note that for each $t > 0$, and sufficiently small $0 < \eta \leq \eta_1(\kappa, \|\sigma\|_{L^3}, \eta_0)$, one finds by the Markov inequality, (A.1) and the assumption (3.1),

$$\mathbb{P}(\tau_K \leq t) = \mathbb{P}\left(\sup_{s \in [0, t]} \|\theta(s)\|_{L^3}^2 \geq K\right) \leq e^{-\eta K} \mathbb{E} \exp\left(\eta \sup_{s \in [0, t]} \|\theta(s)\|_{L^3}^2\right) \leq C e^{-\eta K} \exp(\eta t \|\sigma\|_{L^3}^2).$$

With $\tau = \tau_K$, standard energy estimates, (A.2), (A.9), and the assumption (3.1), we can estimate a term on the right hand side of (3.21) as

$$\begin{aligned} \mathbb{E} \int_0^{t \wedge \tau_K} \|\theta\|^2 (\|U\|^2 + \|B\|^2 + \|\theta\|^2) ds &\leq CK \mathbb{E} \int_0^t (\|\nabla U\|^2 + \|\nabla B\|^2 + \|\theta\|^2) ds \\ &\leq CK \mathbb{E} \left(\frac{\varepsilon}{2} \|U(0)\|^2 + \frac{\delta}{2} \|B(0)\|^2 + \frac{C}{\kappa\nu} \|\Theta(0)\|^2 + \frac{1}{2\kappa} \|\theta(0)\|^2 \right) + CK \|\sigma\|^2 t \\ &\leq CK(1 + \|\sigma\|^2 t). \end{aligned} \quad (3.26)$$

Hence, by combining (3.21) with (3.22) and using (3.26) we find for each $K \geq 1$, the estimate (3.23) is satisfied, where

$$\mathcal{X}(t) := \sup_{s \in [0, t]} \|\phi(s)\|^2, \quad \mathcal{T} := \mathbb{E} (\|\phi(0)\|^2 + C(\varepsilon\|v(0)\|^2 + \delta\|f(0)\|^2)). \quad (3.27)$$

By Lemma 3.4 the estimate (3.24) is satisfied with such $\mathcal{X}(t)$ and \mathcal{T} .

We can now prove (3.2) by combining (3.24), (A.2), Corollary A.5, and the assumption (3.1). For any $p > \gamma$, we find

$$\begin{aligned} \mathbb{E} \sup_{s \in [0, t]} \|\phi(s)\|^p &\leq C \mathbb{E} \sup_{s \in [0, t]} \|\phi(s)\|^\gamma (\|\Theta(s)\|^{p-\gamma} + \|\theta(s)\|^{p-\gamma}) \\ &\leq C \left(\mathbb{E} \left(\sup_{s \in [0, t]} \|\phi(s)\|^{2\gamma} \right) \right)^{1/2} \left[\left(\mathbb{E} \sup_{s \in [0, t]} \|\Theta(s)\|^{2(p-\gamma)} \right)^{1/2} + \left(\mathbb{E} \sup_{s \in [0, t]} \|\theta(s)\|^{2(p-\gamma)} \right)^{1/2} \right] \\ &\leq C \left(\mathbb{E} \left(\sup_{s \in [0, t]} \|\phi(s)\|^{2\gamma} \right) \right)^{1/2} \left[\left(\mathbb{E} \exp \left(\eta \sup_{s \in [0, t]} \|\Theta(s)\|^2 \right) \right)^{1/2} + \left(\mathbb{E} \exp \left(\eta \sup_{s \in [0, t]} \|\theta(s)\|^2 \right) \right)^{1/2} \right] \\ &\leq C \exp(\eta t \|\sigma\|_{L^3}^2) \left[\mathcal{T}^\gamma + (\varepsilon + \delta)^\gamma \right]^{1/2}. \end{aligned} \quad (3.28)$$

Having established (3.2), it remains to prove (3.3). Note that from (3.21) we have

$$\begin{aligned} \mathbb{E} \int_0^t (\nu \|\nabla v\|^2 + \|\nabla f\|^2) ds &\leq \mathbb{E}(\varepsilon\|v(0)\|^2 + \delta\|f(0)\|^2) + Ct \mathbb{E} \sup_{s \in [0, t]} \|\phi(s)\|^2 \\ &\quad + C(\varepsilon^2 + \delta^2) \left(\mathbb{E} \left(\sup_{s \in [0, t]} \|\theta(s)\|^2 \right) \right)^{1/2} \left(\mathbb{E} \left(\int_0^t (\|U\|^2 + \|B\|^2 + \|\theta\|^2) ds \right) \right)^{1/2} \\ &\quad + C(\varepsilon^2 + \delta^2) \mathbb{E} \int_0^t \|\theta\|^2 ds + (\varepsilon + \delta) C \|\sigma\|^2 t. \end{aligned} \quad (3.29)$$

Combining (3.2) and (3.29), and once more invoking (A.2), Corollary A.5, and (3.1), we find

$$\begin{aligned} \mathbb{E} \int_0^t (\nu \|\nabla v\|^2 + \|\nabla f\|^2) ds &\leq \mathbb{E}(\varepsilon\|v(0)\|^2 + \delta\|f(0)\|^2) + Ct \exp(\eta t \|\sigma\|_{L^3}^2) \left[\mathcal{T}^\gamma + (\varepsilon + \delta)^\gamma \right]^{1/2} \\ &\quad + C(\varepsilon^2 + \delta^2) \exp(\eta t \|\sigma\|^2) + C(\varepsilon + \delta) \|\sigma\|^2 t, \end{aligned} \quad (3.30)$$

completing the proof of (3.3), and thus of Theorem 3.1. \square

Proof of Lemma 3.4. For fixed $K > 0$, applying the Grönwall lemma to (3.23), we have

$$\mathbb{E}\left(\mathcal{X}(t)\mathbb{1}_{\tau_K > t}\right) \leq \mathbb{E}\mathcal{X}(t \wedge \tau_K) \leq \left(K\mathcal{T} + CK^2(\varepsilon + \delta)(1 + \|\sigma\|^2 t)\right) \exp(CKt). \quad (3.31)$$

Then, for any $0 < \gamma < 1$,

$$\begin{aligned} \mathbb{E}(\mathcal{X}(t))^\gamma &= \sum_{k=1}^{\infty} \mathbb{E}\left((\mathcal{X}(t))^\gamma \mathbb{1}_{\tau_k > t} \mathbb{1}_{\tau_{k-1} \leq t}\right) \leq \sum_{k=1}^{\infty} \left(\mathbb{E}\left(\mathcal{X}(t)\mathbb{1}_{\tau_k > t}\right)\right)^\gamma (\mathbb{P}(\tau_{k-1} \leq t))^{1-\gamma} \\ &\leq (C')^{1-\gamma} \sum_{k=1}^{\infty} \left[k\mathcal{T} + Ck^2(\varepsilon + \delta)(1 + \|\sigma\|^2 t)\right]^\gamma \exp(\gamma Ctk - (1-\gamma)\eta k) \\ &\leq C_1(\mathcal{T}^\gamma + (\varepsilon + \delta)^\gamma), \end{aligned} \quad (3.32)$$

where we have assumed in the last line $\gamma < \frac{\eta}{\eta + Ct}$ to guarantee convergence of the series. \square

Remark 3.5. We note that the formal limit system (1.5)–(1.7) does not well approximate (1.1)–(1.3) in its velocity and magnetic components at time $t = 0$. This is due to the singular nature of the limit as $\varepsilon, \delta \rightarrow 0$ and is reflected in the phase space mismatch between the two systems. Such considerations are behind the derivation and analysis of the so called ‘corrector systems’ for analogous singular perturbation problems arising in large Prandtl number convection carried out in [Wan04, FGHR15].

To obtain a more accurate approximation for (1.1)–(1.3) in our situation we may expand solutions in formal asymptotic series involving multiple time scales for instance following the method of inner and outer expansions as in e.g. [O’M74, KP03]. If $\delta \ll \varepsilon \ll 1$ this leads to a two stage approximation. By first treating $\delta \approx 0$ in comparison to ε we obtain

$$\begin{aligned} u_\varepsilon &= e^{-tQ/\varepsilon}U(0) + Q^{-1}\hat{e}_3\theta_\varepsilon(t) - e^{-tQ/\varepsilon}Q^{-1}\hat{e}_3\Theta(0), \\ b_\varepsilon &= (-\Delta)^{-1}(\hat{B}_0 \cdot \nabla)u_\varepsilon, \\ d\theta_\varepsilon &= (-u_\varepsilon \cdot \nabla\theta_\varepsilon + \kappa\Delta\theta_\varepsilon)dt + \sigma dW. \end{aligned}$$

By next treating $\varepsilon \sim 1$, in comparison to δ we find

$$\varepsilon(\partial_t u_{\varepsilon,\delta} + u_\varepsilon \cdot \nabla u_{\varepsilon,\delta}) = -Au_{\varepsilon,\delta} + \hat{B}_0 \cdot \nabla b_{\varepsilon,\delta} + \hat{e}_3\theta_{\varepsilon,\delta} \quad (3.33)$$

$$b_{\varepsilon,\delta}(t) = e^{t\Delta/\delta}B(0) + (-\Delta)^{-1}\hat{B}_0 \cdot \nabla u_{\varepsilon,\delta}(t) - e^{t\Delta/\delta}(-\Delta)^{-1}\hat{B}_0 \cdot \nabla U(0), \quad (3.34)$$

$$d\theta_{\varepsilon,\delta} = (-u_{\varepsilon,\delta} \cdot \nabla\theta_{\varepsilon,\delta} + \kappa\Delta\theta_{\varepsilon,\delta})dt + \sigma dW. \quad (3.35)$$

This collection of equations provides a well posed approximation of (1.1)–(1.3). The accuracy of these approximations can be rigorously established using variations on the analysis carried out in the proof of Theorem 3.1 (and see also [FGHR15]). Note also that analogous approximate systems can be found in the cases $\varepsilon \ll \delta \ll 1$ and when $\varepsilon \sim \delta$. In any case such approximate equations are not required for our analysis below which focuses on long time asymptotics, and we omit further details.

4 Analysis of the Markovian Dynamics of the Formal Limit System

In this section we establish conditions guaranteeing the contractive property (1.15), see point (ii) in the summary of main results. Recall that (1.15) implies that the limit equation (1.5) has at most one invariant measure which is thus ergodic. Moreover, this contractive property plays a crucial role in Section 5, where it is used to establish the convergence of statistically stationary states for (1.1)–(1.3) as $\varepsilon, \delta \rightarrow 0$.

As we have already noted in the introduction associating the contraction (1.15) with a version of the Hörmander bracket condition does not require any significant new ideas in comparison to [HM08]. Rather our contribution here is to show that Hörmander’s condition holds for (1.5) with some specific noise configurations. Indeed, in view of the complex nature of the constitutive law (1.9)–(1.10) this verification requires a non-trivial analysis.

Before proceeding to the main results in the section, we first recall some preliminaries related to the Hörmander bracket condition and to Wasserstein metrics.

Hörmander's Condition

In our setting and in view of [HM06, HM08] (see also [Doo48, Kha60]), the proof of (1.15) reduces to establishing smoothing properties of the Markov semigroup $\{P_t\}_{t \geq 0}$ associated to (1.5). Following [HM06] and in the spirit of [H67] this leads to a Hörmander-type bracket condition as follows.

Define, for each $k \in \mathbb{Z}^3$ and $x \in \mathbb{T}^3$,

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

to be the eigenfunctions of the Laplacian \mathbb{T}^3 , and recall that

$$\sigma dW = \sum_{k \in \mathbb{Z}_+^3, m \in \{0,1\}} \alpha_{k,m} \sigma_k^m dW^{k,m},$$

where $\mathbb{Z}_+^3 = \{k = (k_1, k_2, k_3) \in \mathbb{Z}^3 : k \neq 0, k_1 \geq 0\}$ is the integer lattice half-space and $\alpha_{k,m} \in \mathbb{R}$. Let

$$\begin{aligned} \mathcal{W}_0 &:= \text{span}\{\sigma_j^m : \alpha_{j,m} \neq 0\}, \\ \mathcal{W}_n &:= \text{span}\left\{ \{[F, \psi], \sigma\} = M_u(\psi) \cdot \nabla \sigma + M_u(\sigma) \cdot \nabla \psi : \psi \in \mathcal{W}_{n-1}, \sigma \in \mathcal{W}_0\} \cup \mathcal{W}_{n-1} \right\}, \end{aligned} \quad (4.1)$$

where

$$F(\theta) = -\kappa \Delta \theta + M_u(\theta) \cdot \nabla \theta, \quad (4.2)$$

and where, for any Fréchet differentiable functions (vector fields) $E_1, E_2 : H \rightarrow H$, one defines the Lie bracket $[E_1, E_2] : H \rightarrow H$ as

$$[E_1, E_2](\theta) = \nabla E_2(\theta) E_1(\theta) - \nabla E_1(\theta) E_2(\theta). \quad (4.3)$$

The Hörmander-type condition is now stated as follows:

$$\text{For every } N > 0 \text{ there exists an } n = n(N) \text{ such that } \mathcal{W}_n \supseteq H_N, \quad (4.4)$$

where $H_N = \text{span}\{\sigma_k^m : k \in \mathbb{Z}_+^3, |k| \leq N, m \in \{0, 1\}\}$.

Wasserstein Metrics

In order to clarify how (4.4) implies contractivity in the Markovian dynamics associated to (1.5), we proceed to discuss a few generalities about Wasserstein metrics. Recall that, given a metric space (X, \mathbf{m}) , the associated Wasserstein distance $\mathfrak{W}_{\mathbf{m}}$ on $\text{Pr}(X)$ is defined by the following equivalent formulations:

$$\mathfrak{W}_{\mathbf{m}}(\mu, \nu) = \inf_{\Gamma \in \mathcal{C}(\mu, \nu)} \int \mathbf{m}(u, v) d\Gamma(u, v) = \sup_{\|\phi\|_{\text{Lip}, \mathbf{m}} \leq 1} \left| \int \phi d\mu - \int \phi d\nu \right|. \quad (4.5)$$

See e.g. [Vil08]. Here $\mathcal{C}(\mu, \nu)$ denotes the set of couplings between μ and ν , namely

$$\mathcal{C}(\mu, \nu) = \{\Gamma \in \text{Pr}(X \times X) : \Gamma(A \times X) = \mu(A), \Gamma(X \times B) = \nu(B) \text{ for any } A, B \in \mathcal{B}(X)\},$$

and

$$\|\phi\|_{\text{Lip}, \mathbf{m}} = \sup_{u \neq v} \frac{|\phi(u) - \phi(v)|}{\mathbf{m}(u, v)}, \quad (4.6)$$

is the Lipschitz semi-norm corresponding to (X, \mathbf{m}) .

In our setting we consider the metric $\rho = \rho_\eta$ on the phase space $H' = L^2(\mathbb{T}^3)$ as

$$\rho(\theta, \psi) = \inf_{p \in P(\theta, \psi)} \int_0^1 \exp(\eta \|p\|^2) \|p'\| dt \quad (4.7)$$

for a suitable $\eta > 0$, where $P(\theta, \psi) = \{p \in C^1([0, 1], H') : p(0) = \theta, p(1) = \psi\}$. Notice that

$$\|\theta - \psi\| \leq \rho(\theta, \psi) \leq \exp(2\eta(\|\theta\|^2 + \|\psi\|^2)) \|\theta - \psi\|. \quad (4.8)$$

Criteria for Contractivity

Invoking the abstract machinery developed in [HM08, HM11], we now establish contractivity for the Markovian dynamics whenever Hörmander's condition (4.4) is satisfied.

Theorem 4.1. *Let $\{P_t\}_{t \geq 0}$ be the Markov semigroup associated to (1.5) and assume that the Hörmander condition (4.4) holds. Then $\{P_t\}_{t \geq 0}$ is contractive in \mathfrak{W}_ρ :*

$$\mathfrak{W}_\rho(\mu_1 P_t, \mu_2 P_t) \leq C e^{-\gamma' t} \mathfrak{W}_\rho(\mu_1, \mu_2) \quad \text{for every } t \geq 0, \quad (4.9)$$

and hence $\{P_t\}_{t \geq 0}$ possesses a unique ergodic invariant measure. Here the metric ρ is defined relative to an $0 < \eta \leq \eta_0(\|\sigma\|, \kappa, \nu)$. This number η_0 and the constants $C = C(\|\sigma\|, \kappa, \nu) > 0$, $\gamma' = \gamma'(\|\sigma\|, \kappa, \nu) > 0$ are all independent of $t \geq 0$ and μ_1, μ_2 .

Remark 4.2. *Under the conditions of Theorem 4.1 the unique invariant measure μ satisfies certain additional attraction properties: μ is exponentially mixing, and it obeys a strong law of large numbers and a central limit theorem. See e.g. [FGHV14] for more precise analogous statements.*

Proof. The proof essentially follows ideas for the 2D Navier-Stokes equations in vorticity formulation developed in [HM06, HM08, HM11], so we will be sparing in details. Indeed, while we are working in 3D, the constitutive law in (1.5)–(1.7) (that is, the functional M , cf. (1.9)) has two degrees of smoothing as compared to the one degree in the Biot-Sawart law.

The bound (4.9) follows directly from [HM08, Theorem 3.4] which requires us to verify three conditions. The first condition, and the most difficult step in our setting, is a bound for $\nabla P_t \phi(u)$ closely related to the asymptotic strong Feller condition and is thus a form of smoothing. This property follows from [HM11, Theorem 8.4] by invoking (4.4) along with the Lemmata A.1–A.3 established below. Note that this is the only point in the argument where we make use of (4.4).

The second and third conditions needed for [HM08, Theorem 3.4] are a form of irreducibility and certain exponential moment bounds. The irreducibility condition is straightforward in our setting due essentially to the fact that zero is a stable fixed point of the unforced dynamics. See e.g. [EM01, CGHV13] for further details. The second moment condition holds due to the Lemmata A.1–A.3 below. With (4.9) verified, we immediately infer the uniqueness and ergodicity of invariant measures. A standard application of the Krylov-Bogoliubov averaging procedure and compactness yields the existence of invariant measures and thus the proof is now complete. \square

Lie Bracket Computations

We now demonstrate one situation in which (4.4) is satisfied, but a variety of other configurations can be handled with the approach given below.

Theorem 4.3. *The Hörmander bracket condition (4.4) holds if $\sigma_k^m \in \mathcal{W}_0$, or equivalently $\alpha_{k,m} \neq 0$ for all $k \in \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$, $m \in \{0, 1\}$.*

A preliminary step in the proof of Theorem 4.3 is to establish an iterative process for producing new directions that generates the subspaces \mathcal{W}_n specified by (4.1).

Lemma 4.4.

(i) *Suppose that $\sigma_j^{m'} \in \mathcal{W}_0$ and $\sigma_k^m \in \mathcal{W}_n$ for some $j, k \in \mathbb{Z}_+^3$ and all $m, m' \in \{0, 1\}$. Then $\sigma_{k+j}^m, \sigma_{k-j}^m \in \mathcal{W}_{n+1}$ for $m \in \{0, 1\}$ whenever*

$$|\mathcal{M}_u(k) \cdot j| \neq |\mathcal{M}_u(j) \cdot k|. \quad (4.10)$$

(ii) *Assume that $\sigma_{\hat{e}_1}^m \in \mathcal{W}_0$ and $\sigma_k^{m'} \in \mathcal{W}_n$ for some $k = (k_1, k_2, k_3) \in \mathbb{Z}_+^3$ and all $m, m' \in \{0, 1\}$. Then $\sigma_{k \pm \hat{e}_1}^m \in \mathcal{W}_{n+1}$ for $m \in \{0, 1\}$, provided*

$$|k_2(\hat{\Omega} \cdot k)|k|^2 + k_1 k_3((\hat{B}_0 \cdot k)^2 + \nu|k|^4)| \neq \frac{|k_3|D(k)}{\nu + (\hat{B}_0)_1^2} \quad (4.11)$$

(iii) Similarly if $\sigma_{\hat{e}_2}^m \in \mathcal{W}_0$ and $\sigma_k^{m'} \in \mathcal{W}_n$ for $m, m' \in \{0, 1\}$ then $\sigma_{k \pm \hat{e}_2}^{m'} \in \mathcal{W}_{n+1}$ for $m' \in \{0, 1\}$ if

$$|-k_1(\hat{\Omega} \cdot k)|k|^2 + k_2 k_3 ((\hat{B}_0 \cdot k)^2 + \nu|k|^4) \neq \frac{D(k)}{(\nu + (\hat{B}_0)_2^2)^2 + \hat{\Omega}_2^2} |(\nu + (\hat{B}_0)_2^2)k_3 - \hat{\Omega}_2 k_1| \quad (4.12)$$

(iv) Finally whenever $\sigma_{\hat{e}_3}^m \in \mathcal{W}_0$ and $\sigma_k^{m'} \in \mathcal{W}_n$ for $m, m' \in \{0, 1\}$, then $\sigma_{k \pm \hat{e}_3}^{m'} \in \mathcal{W}_{n+1}$ for $m' \in \{0, 1\}$ as long as

$$k_1^2 + k_2^2 \neq 0. \quad (4.13)$$

Proof. Observe that from (1.8) and the definition of the operator M one obtains for each $k \in \mathbb{Z}_+^3$ and $m \in \{0, 1\}$,

$$M_u(\sigma_k^m) = \mathcal{M}_u(k)\sigma_k^m, \quad (4.14)$$

where \mathcal{M}_u is defined in (1.9) and is independent of m . To simplify notation we drop the subscripts u of M_u , \mathcal{M}_u and instead write M , \mathcal{M} for the remainder of the proof.

For the first item, (i), observe that by definition $[[F, \sigma_k^m], \sigma_j^{m'}] \in \mathcal{W}_{n+1}$ for the specified j, k and $m, m' \in \{0, 1\}$. Here recall that the superscript in σ_k^m is understood modulo 2, and F and the Lie bracket are defined by (4.2), (4.3). With (4.14) we observe that:

$$\begin{aligned} [[F, \sigma_k^m], \sigma_j^{m'}] &= M(\sigma_k^m) \cdot \nabla \sigma_j^{m'} + M(\sigma_j^{m'}) \cdot \nabla \sigma_k^m \\ &= (-1)^{m'+1} (\mathcal{M}(k) \cdot j) \sigma_k^m \sigma_j^{m'+1} + (-1)^{m+1} (\mathcal{M}(j) \cdot k) \sigma_k^{m+1} \sigma_j^{m'}, \end{aligned} \quad (4.15)$$

Using standard trigonometric identities we have

$$\sigma_k^n \sigma_j^{n'} = \frac{1}{2} \left((-1)^{nn'} \sigma_{k+j}^{n+n'} + (-1)^{(n+1)n'} \sigma_{k-j}^{n+n'} \right),$$

and consequently (4.15) implies

$$\begin{aligned} &[[F, \sigma_k^m], \sigma_j^{m'}] \\ &= \frac{1}{2} \left((-1)^{(m+1)(m'+1)} (\mathcal{M}(k) \cdot j + \mathcal{M}(j) \cdot k) \sigma_{k+j}^{m+m'+1} + (-1)^{m(m'+1)} (\mathcal{M}(k) \cdot j - \mathcal{M}(j) \cdot k) \sigma_{k-j}^{m+m'+1} \right). \end{aligned} \quad (4.16)$$

Then with $m, m' \in \{0, 1\}$ replaced by $m+1, m'+1$,

$$\begin{aligned} &[[F, \sigma_k^{m+1}], \sigma_j^{m'+1}] \\ &= \frac{1}{2} \left((-1)^{mm'} (\mathcal{M}(k) \cdot j + \mathcal{M}(j) \cdot k) \sigma_{k+j}^{m+m'+1} + (-1)^{(m+1)m'} (\mathcal{M}(k) \cdot j - \mathcal{M}(j) \cdot k) \sigma_{k-j}^{m+m'+1} \right). \end{aligned} \quad (4.17)$$

Combining (4.16), (4.17) we find that $\sigma_{k+j}^{m+m'+1}, \sigma_{k-j}^{m+m'+1} \in \mathcal{W}_{n+1}$, provided that the matrix

$$\begin{pmatrix} (-1)^{(m+1)(m'+1)} (\mathcal{M}(k) \cdot j + \mathcal{M}(j) \cdot k) & (-1)^{mm'} (\mathcal{M}(k) \cdot j + \mathcal{M}(j) \cdot k) \\ (-1)^{m(m'+1)} (\mathcal{M}(k) \cdot j - \mathcal{M}(j) \cdot k) & (-1)^{(m+1)m'} (\mathcal{M}(k) \cdot j - \mathcal{M}(j) \cdot k) \end{pmatrix} \quad (4.18)$$

is invertible, which is true exactly when (4.10) holds, and (i) follows.

The remaining items (ii)–(iv) now follow from (i) and (1.9)–(1.10) after some direct computations. Here in particular note (4.11) uses the fact that $\hat{\Omega}_1 = 0$. The proof of Lemma 4.4 is complete. \square

Proof of Theorem 4.3. We begin by making the preliminary observation that

$$\forall N > 0, \exists K = K(N, \nu, \hat{\Omega}) \text{ such that (4.11), (4.12) both hold whenever } |k_1|, |k_2| \leq N, k_3 \geq K. \quad (4.19)$$

Indeed, for any fixed $|k_1|, |k_2| \leq N$, the left hand side of (4.11), (4.12) grows as k_3^5 while the right hand side grows as k_3^9 as $k_3 \rightarrow \infty$.

We proceed to establish the desired result by showing that

$$\text{for any } N > 0 \text{ there is an } n \text{ such that } \sigma_k^m \in \mathcal{W}_n \text{ for any } |k_1|, |k_2|, |k_3| \leq N \text{ and any } m \in \{0, 1\}. \quad (4.20)$$

Fix $N > 0$. Starting from $k = (1, 0, 0)$ and repeatedly applying Lemma 4.4, (iv), we obtain an n_1 such that $\sigma_k^m \in \mathcal{W}_{n_1}$ for $k = (1, 0, K)$ and each of $m \in \{0, 1\}$. Here, $K > 0$ is as in (4.19) corresponding to the given $N > 0$.

Next, by Lemma 4.4, (ii) and (iii), we obtain an $n_2 > n_1$ such that $\sigma_k^m \in \mathcal{W}_{n_2}$ for every $k = (k_1, k_2, K)$ with $|k_1|, |k_2| \leq N$ and $k_3 = K$. From here, applying Lemma 4.4, (iv), we obtain $n_3 > n_2$ so that $\sigma_k^m \in \mathcal{W}_{n_3}$, whenever $|k_1|, |k_2|, |k_3| \leq N$, $m \in \{0, 1\}$ and $k \notin \text{span}\{\hat{e}_3\}$.

Finally, to obtain the directions corresponding to $\text{span}\{\hat{e}_3\}$, consider, for any $l \in \mathbb{Z} \setminus \{0\}$, elements $\sigma_k^m \in \mathcal{W}_{n_3}$ of the form $k = (\pm 1, 0, l)$ with $m = 0, 1$. For any such k , noting that $k_2 = 0$, the condition (4.11) reduces to

$$(\hat{B}_0 \cdot k)^2 + \nu |k|^4 \neq \frac{D(k)}{\nu + (\hat{B}_0)_1^2}. \quad (4.21)$$

Now observe that the sign in k_1 may be chosen such that $(\hat{B}_0)_1^2 \leq (\hat{B}_0 \cdot k)^2$. Indeed, one may take k_1 so that $\text{sign}(k_1(\hat{B}_0)_1) = \text{sign}(k_3(\hat{B}_0)_3)$ or, in the case that either $(\hat{B}_0)_1 = 0$ or $(\hat{B}_0)_3 = 0$, one may choose any sign for k_1 . With this choice, (4.21) holds since

$$(\hat{B}_0 \cdot k)^2 + \nu |k|^4 < \frac{1}{\nu + (\hat{B}_0)_1^2} ((\hat{B}_0 \cdot k)^2 + \nu |k|^4)^2 \leq \frac{1}{\nu + (\hat{B}_0)_1^2} D(k).$$

Thus, we obtain an $n \geq n_3$ such that (4.20) holds for the given N . The proof of Theorem 4.3 is complete. \square

5 Convergence of Statistically Steady States

In this final section we combine the finite time bounds obtained from Section 3 with the contraction estimate (4.9) of Section 4, to establish the convergence of statistically invariant states for (1.1)–(1.3) to the unique invariant measure of (1.5) (see (iv) in the summary of main results). As we discussed in the introduction, one of the novelties here is that our convergence analysis applies to the extend phase space H , not just to the temperature component, as was the case in our previous work [FGHR15].

In order to precisely state the main result of this section we introduce some notation. Recall that the metric ρ on $L^2(\mathbb{T}^3)$ is defined in (4.7). Here the parameter $\eta > 0$ appearing in (4.7) is taken to be a suitably small value so that (4.9) applies. Using ρ we define a new distance on H according to

$$\tilde{\rho}((U, B, \Theta), (\tilde{U}, \tilde{B}, \tilde{\Theta})) = \|U - \tilde{U}\|_{H^1} + \|B - \tilde{B}\|_{H^1} + \rho(\Theta, \tilde{\Theta}).$$

Following the notation in (4.5), $\tilde{\rho}$ induces a Wasserstein metric on H which we denote by $\mathfrak{W}_{\tilde{\rho}}$.

In what follows it will be useful to consider the set of ‘observables’

$$V(H) = V_\eta(H) := \{\phi \in C^1(H) : [\phi]_\eta < \infty\}, \quad (5.1)$$

where $[\cdot]_\eta$ denotes

$$[\phi]_\eta := \sup_{(u, b, \theta) \in (H^1)^2 \times H'} \left[\sup_{\zeta_1 \in H^1, \|\zeta_1\|_{H^1} = 1} |\nabla_u \phi(u, b, \theta) \cdot \zeta_1| + \sup_{\zeta_2 \in H^1, \|\zeta_2\|_{H^1} = 1} |\nabla_b \phi(u, b, \theta) \cdot \zeta_2| + \exp(-\eta \|\theta\|) \sup_{\xi \in H', \|\xi\| = 1} |\nabla_\theta \phi(u, b, \theta) \cdot \xi| \right].$$

It can be verified with a straight-forward proof (see [HM08, Proposition 4.1]) that, for any $\phi \in C^1(H)$,

$$\|\phi\|_{Lip, \tilde{\rho}} \leq C[\phi]_\eta. \quad (5.2)$$

This bound is useful for translating the convergence of measures relative to $\mathfrak{W}_{\tilde{\rho}}$ to the convergence of observables from $V(H)$. See (5.4) below.

Let $L : H' \rightarrow H$ be the extension operator associated to (1.6)–(1.7) given by

$$L(\theta) = (M(\theta), \theta).$$

Also, the projection operator $\Pi : H \rightarrow H'$ will associate elements in H to their θ component,

$$\Pi(u, b, \theta) = \theta.$$

Recall that for any measure μ and function F , the push-forward of μ under F is given by $F_{\#}\mu = \mu \circ F^{-1}$. To simplify notation, for a measure μ_1 on H' , we will use $L\mu_1 = L_{\#}\mu_1$ to denote the extended measure on H , and similarly for μ_2 on H we will write $\Pi\mu_2 = \Pi_{\#}\mu_2$ for the projected measure on H' (i.e. its marginal in the θ component).

Theorem 5.1. *Suppose that the conditions imposed in Theorem 4.1 are satisfied. For any $\varepsilon, \delta > 0$, let $\mu_{\varepsilon, \delta}$ be a statistically invariant state of (1.1)–(1.3) satisfying the uniform moment condition (2.1) and let μ_0 be the unique invariant measure of (1.5). Then, there exists $\tilde{\gamma} = \tilde{\gamma}(\nu, \kappa, \|\sigma\|)$, $C = C(C_0, \nu, \kappa, \|\sigma\|)$ which are independent of $\varepsilon, \delta > 0$, such that*

$$\mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L\mu_0) \leq C(\varepsilon + \delta)^{\tilde{\gamma}} \quad (5.3)$$

for every $\varepsilon, \delta > 0$. This implies that for statistically invariant states (U, B, Θ) and θ distributed respectively as $\mu_{\varepsilon, \delta}$ and μ_0 ,

$$|\mathbb{E}(\phi(U, B, \Theta) - \phi(L\theta))| \leq C[\phi]_{\eta}(\varepsilon + \delta)^{\tilde{\gamma}} \quad (5.4)$$

for any $\phi \in V(H)$.

Remark 5.2. *The estimate (5.4) indicates that, at statistical equilibrium, observations of Θ are well approximated by observations of θ , which in turn provide good estimates on observations of (U, B) through $M(\theta)$. That is, using (5.4) we can approximate (U, B) using observations of Θ from statistically stationary states.*

The proof of Theorem 5.1 makes use of a second metric on $L^2(\mathbb{T}^3)$ defined as

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

and its associated Wasserstein distance \mathfrak{W}_{ρ^*} .

Lemma 5.3. *The metric ρ^* is equivalent to ρ and hence \mathfrak{W}_{ρ^*} and \mathfrak{W}_{ρ} are also equivalent. Moreover,*

$$\mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2) \leq \mathfrak{W}_{\rho^*}(\mu_1, \mu_2) \leq C\mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2), \quad (5.5)$$

for any $\mu_1, \mu_2 \in Pr(L^2(\mathbb{T}^3))$.

Proof of Lemma 5.3. Invoking the smoothing properties of the constitutive law $M = (M_u, M_b)$ along with (4.8), we find

$$\rho(\theta, \tilde{\theta}) \leq \rho^*(\theta, \tilde{\theta}) = \|M_u(\theta - \tilde{\theta})\|_{H^1} + \|M_b(\theta - \tilde{\theta})\|_{H^1} + \rho(\theta, \tilde{\theta}) \leq C\|\theta - \tilde{\theta}\| + \rho(\theta, \tilde{\theta}) \leq C\rho(\theta, \tilde{\theta}). \quad (5.6)$$

It follows directly from the definition that the corresponding Wasserstein distances \mathfrak{W}_{ρ} and \mathfrak{W}_{ρ^*} are equivalent as well; cf. (4.5).

To establish (5.5) for each $k > 0$, we consider

$$\begin{aligned} T_1^k &:= \{\phi : L^2(\mathbb{T}^3) \rightarrow \mathbb{R} : \|\phi\|_{Lip, \rho^*} \leq k\}, \\ T_2^k &:= \{\phi : L^2(\mathbb{T}^3) \rightarrow \mathbb{R} : \phi(\theta) = \psi(L\theta) \text{ for some } \psi : H \rightarrow \mathbb{R} \text{ with } \|\psi\|_{Lip, \tilde{\rho}} \leq k\}. \end{aligned}$$

We claim that for each $k > 0$ one has $T_2^k \subset T_1^k \subset T_2^{Ck}$, where C is as in (5.6). Indeed, if $\phi \in T_2^k$, then

$$\frac{|\phi(\theta) - \phi(\tilde{\theta})|}{\rho^*(\theta, \tilde{\theta})} = \frac{|\psi(L(\theta)) - \psi(L(\tilde{\theta}))|}{\tilde{\rho}(L(\theta), L(\tilde{\theta}))} \leq k, \quad \text{for each } \theta \neq \tilde{\theta},$$

and therefore $\phi \in T_1^k$. On the other hand if $\phi \in T_1^k$ we define $\psi := \phi \circ \Pi$ (clearly $\phi = \psi \circ L$) and calculate

$$\frac{|\psi(u, b, \theta) - \psi(\tilde{u}, \tilde{b}, \tilde{\theta})|}{\tilde{\rho}((u, b, \theta), (\tilde{u}, \tilde{b}, \tilde{\theta}))} = \frac{|\phi(\theta) - \phi(\tilde{\theta})|}{\tilde{\rho}((u, b, \theta), (\tilde{u}, \tilde{b}, \tilde{\theta}))} \leq \frac{|\phi(\theta) - \phi(\tilde{\theta})|}{\rho(\theta, \tilde{\theta})} \leq C \frac{|\phi(\theta) - \phi(\tilde{\theta})|}{\rho^*(\theta, \tilde{\theta})} = Ck,$$

and we showed $\phi \in T_2^{Ck}$. Finally, for any $\mu_1, \mu_2 \in Pr(L^2(\mathbb{T}^3))$,

$$\begin{aligned} \mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2) &= \sup_{\|\psi\|_{L^{ip}, \tilde{\rho}} \leq 1} \left| \int \psi(L\theta) d\mu_1(\theta) - \int \psi(L\theta) d\mu_2(\theta) \right| \\ &= \sup_{\phi \in T_2^1} \left| \int \phi(\theta) d\mu_1(\theta) - \int \phi(\theta) d\mu_2(\theta) \right| \leq \sup_{\phi \in T_1^1} \left| \int \phi(\theta) d\mu_1(\theta) - \int \phi(\theta) d\mu_2(\theta) \right| = \mathfrak{W}_{\rho^*}(\mu_1, \mu_2), \end{aligned}$$

and

$$\begin{aligned} \mathfrak{W}_{\rho^*}(\mu_1, \mu_2) &= \sup_{\phi \in T_1^1} \left| \int \phi(\theta) d\mu_1(\theta) - \int \phi(\theta) d\mu_2(\theta) \right| \leq \sup_{\tilde{\phi} \in T_2^C} \left| \int \tilde{\phi}(\theta) d\mu_1(\theta) - \int \tilde{\phi}(\theta) d\mu_2(\theta) \right| \\ &= C \sup_{\tilde{\phi} \in T_2^1} \left| \int \tilde{\phi}(\theta) d\mu_1(\theta) - \int \tilde{\phi}(\theta) d\mu_2(\theta) \right| = C\mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2), \end{aligned}$$

as desired. \square

By combining Lemma 5.3 with Theorem 4.1 we obtain the following ‘lifted’ contraction property.

Corollary 5.4. *Let $\{P_t\}_{t \geq 0}$ be the Markov semigroup associated to (1.5) and assume that the Hörmander condition (4.4) holds. Then $\{P_t\}_{t \geq 0}$ satisfies the following contractive property in $\mathfrak{W}_{\tilde{\rho}}$:*

$$\mathfrak{W}_{\tilde{\rho}}(L(\mu_1 P_t), L(\mu_2 P_t)) \leq C e^{-\gamma' t} \mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2) \quad \text{for every } t \geq 0, \quad (5.7)$$

where the constants $C = C(\|\sigma\|, \kappa, \nu) > 0$, $\eta = \eta(\|\sigma\|, \kappa, \nu) > 0$, $\gamma' = \gamma'(\|\sigma\|, \kappa, \nu) > 0$ are all independent of $t \geq 0$ and μ_1, μ_2 .

Proof of Corollary 5.4. By combining Lemma 5.3 with (4.9), we find that

$$\begin{aligned} \mathfrak{W}_{\tilde{\rho}}(L(\mu_1 P_t), L(\mu_2 P_t)) &\leq C \mathfrak{W}_{\rho^*}(\mu_1 P_t, \mu_2 P_t) \leq C \mathfrak{W}_{\rho}(\mu_1 P_t, \mu_2 P_t) \\ &\leq C e^{-\gamma' t} \mathfrak{W}_{\rho}(\mu_1, \mu_2) \leq C e^{-\gamma' t} \mathfrak{W}_{\rho^*}(\mu_1, \mu_2) \leq C e^{-\gamma' t} \mathfrak{W}_{\tilde{\rho}}(L\mu_1, L\mu_2), \end{aligned}$$

for any $\mu_1, \mu_2 \in Pr(L^2(\mathbb{T}^3))$, where C is independent of t and μ_1, μ_2 . \square

We now proceed with the proof of Theorem 5.1.

Proof of Theorem 5.1. We invoke the invariance of μ_0 under P_t and apply (5.7) to infer for any $t, t_0 \geq 0$,

$$\begin{aligned} \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L\mu_0) &= \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L(\mu_0 P_{t+t_0})) \\ &\leq \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_{t+t_0})) + \mathfrak{W}_{\tilde{\rho}}(L((\Pi\mu_{\varepsilon, \delta})P_{t+t_0}), L(\mu_0 P_{t+t_0})) \\ &\leq \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_{t+t_0})) + C e^{-\gamma' t_0} \mathfrak{W}_{\tilde{\rho}}(L((\Pi\mu_{\varepsilon, \delta})P_t), L(\mu_0 P_t)) \\ &\leq \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_{t+t_0})) + C e^{-\gamma' t_0} [\mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_t)) + \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L\mu_0)]. \end{aligned}$$

By taking $t_0 = t_0(\|\sigma\|, \kappa, \nu)$ large enough such that $C e^{-\gamma' t_0} = \frac{1}{2}$, we find that for $t \geq 0$,

$$\mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L\mu_0) \leq 2\mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_{t+t_0})) + C\mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_t)). \quad (5.8)$$

Using the coupling definition of the Wasserstein metric in (4.5) it follows that for any $s \geq 0$,

$$\begin{aligned} \mathfrak{W}_{\tilde{\rho}}(\mu_{\varepsilon, \delta}, L((\Pi\mu_{\varepsilon, \delta})P_s)) &\leq \mathbb{E} \tilde{\rho}((U, B, \Theta)(s), L(\theta(s))) \\ &= \underbrace{\mathbb{E}(\|U(s) - M_u(\theta(s))\|_{H^1} + \|B(s) - M_b(\theta(s))\|_{H^1})}_{I_1(s)} + \underbrace{\mathbb{E}(\Theta(s), \theta(s))}_{I_2(s)}, \end{aligned} \quad (5.9)$$

where $(U, B, \Theta) \sim \mu_{\varepsilon, \delta}$ denotes the stationary solution to (1.1)–(1.3) provided by Proposition 2.1, and θ denotes the solution to (1.5) with initial condition distributed according to the stationary state $\Theta = \Theta(\varepsilon, \delta) \sim \Pi\mu_{\varepsilon, \delta}$. By combining the estimates (5.8)–(5.9) and integrating in t over the interval $[0, t_0]$ we infer that

$$\mathfrak{W}_{\bar{\rho}}(\mu_{\varepsilon, \delta}, L\mu_0) \leq \frac{C}{t_0} \int_0^{t_0} (I_1(t+t_0) + I_2(t+t_0) + I_1(t) + I_2(t)) dt = C \int_0^{2t_0} (I_1(t) + I_2(t)) dt. \quad (5.10)$$

where we have absorbed t_0 into the constant $C = C(\|\sigma\|, \kappa, \nu)$ in the last line. Applying the Cauchy-Schwarz inequality, (3.3) with $\theta(0)$ distributed as $\Theta(0)$, and (2.1), we obtain a bound for I_1 as

$$\begin{aligned} \int_0^{2t_0} I_1(t) dt &\leq C \left(\mathbb{E} \int_0^{2t_0} (\|U(t) - M_u(\theta(t))\|_{H^1}^2 + \|B(t) - M_b(\theta(t))\|_{H^1}^2) dt \right)^{1/2} \\ &\leq C \exp(\eta t_0 \|\sigma\|_{L^3}^2) \left((\varepsilon + \delta)^\gamma + (\mathbb{E}(\varepsilon \|U(0) - M_u(\theta(0))\|^2 + \delta \|B(0) - M_b(\theta(0))\|^2))^\gamma \right)^{1/4} \\ &\leq C \exp(\eta t_0 \|\sigma\|_{L^3}^2) (\varepsilon + \delta)^{\gamma/4} (\mathbb{E}(\|U(0)\|^2 + \|B(0)\|^2 + \|\Theta(0)\|^2 + 1))^{\gamma/4} \\ &\leq C \exp(\eta t_0 \|\sigma\|_{L^3}^2) (\varepsilon + \delta)^{\gamma/4}, \end{aligned} \quad (5.11)$$

where C in the last line depends on C_0 given in (3.1), but is independent of t , ε , and δ . Similarly, using a property of the metric ρ , (4.8), and the bounds (3.2), (2.1) we have

$$\begin{aligned} \int_0^{2t_0} I_2(t) dt &\leq C \mathbb{E} \int_0^{2t_0} \exp(2\eta(\|\Theta(t)\|^2 + \|\theta(t)\|^2)) \|\Theta(t) - \theta(t)\| dt \\ &\leq C \left(\mathbb{E} \exp\left(4\eta \sup_{t \in [0, 2t_0]} (\|\Theta(t)\|^2 + \|\theta(t)\|^2)\right) \right)^{1/2} \left(\mathbb{E} \sup_{t \in [0, 2t_0]} \|\Theta(t) - \theta(t)\|^2 \right)^{1/2} \\ &\leq C \exp(10\eta t_0 \|\sigma\|_{L^3}^2) (\varepsilon + \delta)^{\gamma/4} (\mathbb{E}(\|U(0)\|^2 + \|B(0)\|^2 + \|\Theta(0)\|^2))^{\gamma/4} \\ &\leq C \exp(10\eta t_0 \|\sigma\|_{L^3}^2) (\varepsilon + \delta)^{\gamma/4}, \end{aligned} \quad (5.12)$$

where the constant C depends on C_0 , but is independent of t_0, ε , and δ . Combining (5.10) with (5.11)–(5.12) yields (5.3). The convergence of observables as in (5.4) follows from (5.3) by using the definition of $\mathfrak{W}_{\bar{\rho}}$ (see (4.5)) and applying (5.2). \square

A Uniform Moment Bounds

This section collects various moment bounds related to the full system (1.1)–(1.3) as well as the limit active scalar equation (1.5) which we use repeatedly in the analysis above.

We begin with some L^p -bounds on the limit equation.

Lemma A.1. *Let θ be a solution of (1.5) with \mathcal{F}_0 adapted initial condition θ_0 . Then, for any $p \geq 2$ there exists constants $\eta_0 = \eta_0(p, \kappa, \|\sigma\|_{L^p}) > 0$, $C = C(p, \kappa)$ and $\alpha = \alpha(p, \kappa)$, such that for each $\eta \leq \eta_0$*

$$\mathbb{E} \exp \left(\eta \sup_{s \in [0, t]} \|\theta\|_{L^p}^2 + \eta \int_0^t \|\theta\|_{L^p}^2 ds \right) \leq C \mathbb{E} \exp(\eta(\|\theta_0\|_{L^p}^2 + t\|\sigma\|_{L^p}^2)). \quad (\text{A.1})$$

Moreover,

$$\mathbb{E} \exp \left(\eta \left(\sup_{s \in [0, t]} \|\theta\|^2 + \kappa \int_0^t \|\nabla \theta\|^2 ds \right) \right) \leq C \mathbb{E} \exp(\eta(\|\theta_0\|^2 + t\|\sigma\|^2)). \quad (\text{A.2})$$

Finally, for any $t \geq 0$

$$\mathbb{E} \exp(\eta \|\theta(t)\|_{L^p}^2) \leq C \mathbb{E} \exp(\eta(e^{-\alpha t} \|\theta_0\|_{L^p}^2 + C \|\sigma\|_{L^p}^2)). \quad (\text{A.3})$$

The L^p -estimates (A.1) and (A.3) are established in [FGHRW15, Proposition 3.1] for a more general divergence-free drift diffusion equation supplemented with mixed Dirichlet-periodic boundary conditions. Indeed, recall that $\xi = \theta$ satisfies

$$d\xi + (v \cdot \nabla \xi - \kappa \Delta \xi) dt = \sigma dW, \quad \xi(0) = \xi_0, \quad (\text{A.4})$$

with $\|v\|_{H^2} \leq C\|\xi\|$ (see (1.9)). Then (A.1) and (A.3) follow precisely as in the proof of [FGHRW15, Proposition 3.1] with $\bar{R}a = 0$, making obvious adjustments to work on the domain $\mathcal{D} = \mathbb{T}^3$.

Similarly, an estimate of the form (A.2) follows immediately from [FGHRW15, Proposition 3.3] for the active scalar equation obtained in the infinite Prandtl limit of the Boussinesq system. This equation bears structural similarity to (1.5), and the proof of (A.2) for (1.5) can be obtained with essentially the same argument, under a simplifying assumption that $\bar{R}a = 0$, and with obvious changes for the domain $\mathcal{D} = \mathbb{T}^3$.

Next, we need an estimate for the first variation $J_{s,t}\xi$ of (1.5). That is, for the solution of

$$\partial_t \zeta + M_u(\zeta) \cdot \nabla \theta + M_u(\theta) \cdot \nabla \zeta - \kappa \Delta \zeta = 0, \quad \zeta(s) = \xi, \quad (\text{A.5})$$

where $\xi \in L^2(\mathbb{T}^3)$, $s < t$, and M_u being as above (see (1.9)).

Lemma A.2. *For each $p > 0$ and $\eta > 0$ there exists C such that*

$$\mathbb{E} \sup_{s,t \in [0,1]} \|J_{s,t}\xi\|^p \leq C \exp(\eta \|\theta_0\|^2) \|\xi\|^p.$$

We also need a similar estimate for the second variation $J_{s,t}^2(\xi, \xi')$ of (1.5), that is, for the solution of the problem

$$\partial_t \zeta + M_u(\zeta) \cdot \nabla \theta + M_u(\theta) \nabla \zeta + M_u(J_{0,t}\xi) \cdot \nabla J_{0,t}\xi' + M_u(J_{0,t}\xi') \cdot \nabla J_{0,t}\xi - \kappa \Delta \zeta = 0, \quad \zeta(s) = 0, \quad (\text{A.6})$$

where $\xi, \xi' \in L^2(\mathbb{T}^3)$ and $s < t$.

Lemma A.3. *For each $p > 0$ and $\eta > 0$ there exists C such that*

$$\mathbb{E} \sup_{s,t \in [0,1]} \|J_{s,t}^{(2)}(\xi, \xi')\|^p \leq C \exp(\eta \|\theta_0\|^2) \|\xi\|^p \|\xi'\|^p.$$

The proofs of Lemma A.2 and Lemma A.3 follow ideas developed for the 2D Navier-Stokes equations in vorticity formulation (see e.g. [HM06]). The difference here is that we work in 3D, but have superior smoothing properties of the constitutive law (1.9). The required modifications are straightforward, and we omit details.

Next, we prove for the system (1.1)–(1.3) analogous results to those in Lemma A.1.

Lemma A.4. *For any $\gamma \leq \kappa^2 \nu / (4C \|\sigma\|^2)$ and $K > 0$,*

$$\begin{aligned} \mathbb{P} \left(\sup_{t \geq 0} \left(\frac{\varepsilon}{2} \|U(t)\|^2 + \frac{\delta}{2} \|B(t)\|^2 + \frac{C}{2\kappa\nu} \|\Theta(t)\|^2 + \int_0^t \frac{\nu}{2} \|\nabla U\|^2 + \|\nabla B\|^2 + \frac{C}{2\nu} \|\nabla \Theta\|^2 ds \right. \right. \\ \left. \left. - \frac{\varepsilon}{2} \|U(0)\|^2 - \frac{\delta}{2} \|B(0)\|^2 - \frac{C}{\kappa\nu} \|\Theta(0)\|^2 - \frac{C\|\sigma\|^2}{\kappa\nu} t \geq K \middle| \mathcal{F}_0 \right) \leq e^{-\gamma K}, \end{aligned}$$

where $C = C(\mathbb{T}^3)$.

We have the following corollary as an immediate consequence:

Corollary A.5. *For any $T \geq 0$, $\delta, \varepsilon \leq 1$ and γ, C as in Lemma A.4, there exists $\eta_0 = \eta_0(\kappa, \nu, \|\sigma\|)$ such that for any $0 < \eta < \eta_0$ and any \mathcal{F}_0 measurable $U(0), B(0), \Theta(0)$ one has*

$$\begin{aligned} \mathbb{E} \exp \left(\eta \sup_{t \in [0,T]} \left(\frac{\varepsilon}{2} \|U(t)\|^2 + \frac{\delta}{2} \|B(t)\|^2 + \frac{C}{2\kappa\nu} \|\Theta(t)\|^2 \right) + \eta \int_0^T \frac{\nu}{2} \|\nabla U\|^2 + \|\nabla B\|^2 + \frac{C}{2\nu} \|\nabla \Theta\|^2 ds \right) \\ \leq C \mathbb{E} \exp \left(\eta \left(\frac{\varepsilon}{2} \|U(0)\|^2 + \frac{\delta}{2} \|B(0)\|^2 + \frac{C}{\kappa\nu} \|\Theta(0)\|^2 + \frac{C\|\sigma\|^2}{\kappa\nu} T \right) \right). \end{aligned}$$

Proof of Lemma A.4. Testing (1.1), (1.2) by U and B respectively and adding them we obtain

$$\frac{d}{dt} \left(\frac{\varepsilon}{2} \|U\|^2 + \frac{\delta}{2} \|B\|^2 \right) + \nu \|\nabla U\|^2 + \|\nabla B\|^2 = \langle \Theta, U_3 \rangle \leq \frac{\nu}{2} \|\nabla U\|^2 + \frac{C}{\nu} \|\nabla \Theta\|^2. \quad (\text{A.7})$$

Itô formula applied to (1.3) yields

$$\frac{1}{2} d\|\Theta\|^2 + \kappa \|\nabla \Theta\|^2 dt = \frac{1}{2} \|\sigma\|^2 dt + \langle \sigma, \Theta \rangle dW. \quad (\text{A.8})$$

By adding $2C/(\kappa\nu)$ multiple of (A.8) to (A.7) we obtain

$$d \left(\frac{\varepsilon}{2} \|U\|^2 + \frac{\delta}{2} \|B\|^2 + \frac{C}{\kappa\nu} \|\Theta\|^2 \right) + \frac{\nu}{2} \|\nabla U\|^2 + \|\nabla B\|^2 + \frac{C}{\nu} \|\nabla \Theta\|^2 \leq \frac{C}{\kappa\nu} \|\sigma\|^2 dt + \frac{2C}{\kappa\nu} \langle \sigma, \Theta \rangle dW. \quad (\text{A.9})$$

Using that for every $\gamma \leq \kappa^2\nu/(4C\|\sigma\|^2)$ one has for the quadratic variation of the Martingale

$$\frac{\gamma}{2} \frac{4C^2}{\kappa^2\nu^2} |\langle \sigma, \Theta \rangle|^2 \leq \frac{\gamma}{2} \frac{4C^2}{\kappa^2\nu^2} \|\sigma\|^2 \|\Theta\|^2 \leq \frac{C}{2\nu} \|\nabla \Theta\|^2,$$

and consequently the following exponential Martingale inequality: given a continuous Martingale $N = N(t)$, for any $K \geq 0$,

$$\mathbb{P} \left(\sup_{t \geq 0} \left(N(t) - \frac{\gamma}{2} \langle N(t) \rangle \right) \geq K \mid \mathcal{F}_0 \right) \leq \exp(-\gamma K), \quad (\text{A.10})$$

where $\langle N \rangle$ is the quadratic variation of N implies the desired result. \square

Lemma A.6. Let $\alpha := \frac{\nu}{2\varepsilon} \wedge \frac{1}{\delta}$ and $\beta := \kappa/C$ (where $C = C(\mathbb{T}^3)$ will be determined below). There exists $\eta_1 = \eta_1(\kappa, \nu, \|\sigma\|)$ such that for any $0 < \eta < \eta_1$, $T > 0$, and $0 < \varepsilon, \delta \leq c_0(\kappa, \nu, \mathbb{T}^3)$,

$$\begin{aligned} & \mathbb{E} \exp \left(\eta \left(\frac{\varepsilon}{2} \|U(T)\|^2 + \frac{\delta}{2} \|B(T)\|^2 + \frac{C}{\kappa\nu} \|\Theta(T)\|^2 \right. \right. \\ & \quad \left. \left. + \exp(-\alpha T) \int_0^T \left(\frac{\nu}{2} \|\nabla U\|^2 + \|\nabla B\|^2 \right) ds + \exp(-\beta T) \frac{C}{2\nu} \int_0^T \|\nabla \Theta\|^2 ds \right) \right) \\ & \leq C \mathbb{E} \exp \left(\eta \left(\frac{\varepsilon}{2} \exp(-\alpha T) \|U(0)\|^2 + \frac{\delta}{2} \exp(-\alpha T) \|B(0)\|^2 + \frac{C}{\kappa\nu} \exp(-\beta T) \|\Theta(0)\|^2 + \frac{C\|\sigma\|^2}{\beta\kappa\nu} \right) \right). \end{aligned}$$

Proof. Fix $T \geq 0$ and for any real r denote $m_r(t) := \exp(r(t-T))$. Then from (A.7) we obtain

$$\begin{aligned} & \frac{d}{dt} \left(\frac{\varepsilon}{2} m_\alpha \|U\|^2 + \frac{\delta}{2} m_\alpha \|B\|^2 \right) + \nu m_\alpha \|\nabla U\|^2 + m_\alpha \|\nabla B\|^2 = m_\alpha \langle \Theta, U_3 \rangle + \frac{\varepsilon\alpha}{2} m_\alpha \|U\|^2 + \frac{\delta\alpha}{2} m_\alpha \|B\|^2 \\ & \leq \frac{\nu}{2} m_\alpha \|\nabla U\|^2 + \frac{1}{2} m_\alpha \|\nabla B\|^2 + \frac{C}{\nu} m_\alpha \|\nabla \Theta\|^2. \end{aligned} \quad (\text{A.11})$$

Also, from (A.8) with $\beta = \kappa/C$ with $C = C(\mathbb{T}^3)$ we obtain

$$\begin{aligned} & \frac{1}{2} d(m_\beta \|\Theta\|^2) + \kappa m_\beta \|\nabla \Theta\|^2 dt = \frac{1}{2} m_\beta \|\sigma\|^2 dt + \frac{\beta}{2} m_\beta \|\Theta\|^2 dt + m_\beta \langle \sigma, \Theta \rangle dW \\ & \leq \frac{1}{2} m_\beta \|\sigma\|^2 dt + \frac{\kappa}{2} m_\beta \|\nabla \Theta\|^2 dt + m_\beta \langle \sigma, \Theta \rangle dW. \end{aligned} \quad (\text{A.12})$$

If $\varepsilon, \delta \ll 1$, then $\beta \leq \alpha$ and $m_\beta \geq m_\alpha$ on $(-\infty, T]$. Then adding $2C/(\kappa\nu)$ multiple of (A.12) to (A.11) we find for any $\gamma \leq \frac{C(\mathbb{T}^3)\kappa^2\nu}{2\|\sigma\|^2}$,

$$\begin{aligned} & d \left(\frac{\varepsilon}{2} m_\alpha \|U\|^2 + \frac{\delta}{2} m_\alpha \|B\|^2 + \frac{C}{\kappa\nu} m_\beta \|\Theta\|^2 \right) \\ & \quad + \left(\frac{\nu}{2} m_\alpha \|\nabla U\|^2 + \frac{1}{2} m_\alpha \|\nabla B\|^2 + \frac{C}{2\nu} m_\beta \|\nabla \Theta\|^2 \right) dt \\ & \leq \frac{C}{\kappa\nu} m_\beta \|\sigma\|^2 dt + \frac{2}{\nu\kappa} m_\beta \langle \sigma, \Theta \rangle dW - \frac{\gamma}{2} \frac{4}{\nu^2\kappa^2} m_\beta^2 |\langle \sigma, \Theta \rangle|^2 dt. \end{aligned} \quad (\text{A.13})$$

By the exponential Martingale inequality (A.10), for any $K > 0$,

$$\begin{aligned} & \mathbb{P}\left(\frac{\varepsilon}{2}\|U(T)\|^2 + \frac{\delta}{2}\|B(T)\|^2 + \frac{C}{\kappa\nu}\|\Theta(T)\|^2\right. \\ & \quad + \exp(-\alpha T) \int_0^T \left(\frac{\nu}{2}\|\nabla U\|^2 + \|\nabla B\|^2\right) ds + \exp(-\alpha T) \frac{C}{2\nu} \int_0^T \|\nabla\Theta\|^2 ds \\ & \quad \left. - \frac{\varepsilon}{2} \exp(-\alpha T)\|U(0)\|^2 - \frac{\delta}{2} \exp(-\alpha T)\|B(0)\|^2 - \frac{C}{\kappa\nu} \exp(-\beta T)\|\Theta(0)\|^2 - \frac{C\|\sigma\|^2}{\beta\kappa\nu} \geq K \middle| \mathcal{F}_0\right) \\ & \leq e^{-\gamma K}. \end{aligned}$$

The proof follows by combining this estimate with the layer cake formula $\mathbb{E}(|X|) = \int_0^\infty \mathbb{P}(|X| \geq L) dL$ for any random variable X . \square

Acknowledgments

This work was partially supported by the National Science Foundation under the grants NSF-DMS-1207780 (SF) and NSF-DMS-1313272 (NEGH). We express our appreciation and acknowledge support from the Mathematical Sciences Research Institute (MSRI), the Mathematisches Forschungsinstitut Oberwolfach (MFO) and the WISE Program at USC. We would also like to thank Jared Whitehead for providing helpful feedback and references.

References

- [AFS08] S. Albeverio, F. Flandoli, and Y. G. Sinai, *SPDE in hydrodynamic: recent progress and prospects*, Lecture Notes in Mathematics, vol. 1942, Springer-Verlag, Berlin, 2008, Lectures given at the C.I.M.E. Summer School held in Cetraro, August 29–September 3, 2005, Edited by Giuseppe Da Prato and Michael Röckner. MR 2459087 (2009g:76120)
- [BDP07] V. Barbu and G. Da Prato, *Existence and ergodicity for the two-dimensional stochastic magneto-hydrodynamics equations*, Applied Mathematics and Optimization **56** (2007), no. 2, 145–168.
- [CJTA15] M. A. Calkins, K. Julien, S. M. Tobias, and J. M. Aurnou, *A multiscale dynamo model driven by quasi-geostrophic convection*, Journal of Fluid Mechanics **780** (2015), 143–166.
- [CW11] C. Cao and J. Wu, *Global regularity for the 2d mhd equations with mixed partial dissipation and magnetic diffusion*, Advances in Mathematics **226** (2011), no. 2, 1803–1822.
- [CDGG06] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier, *Mathematical geophysics*, Oxford Lecture Series in Mathematics and its Applications **32** (2006).
- [CGHV13] P. Constantin, N. E. Glatt-Holtz, and V. Vicol, *Unique ergodicity for fractionally dissipated, stochastically forced 2d euler equations*, Comm. Math. Phys. (2013), (to appear).
- [Dav01] P. A. Davidson, *An introduction to magnetohydrodynamics*, Cambridge Texts in Applied Mathematics, 2001.
- [Deb11] A. Debussche, *Ergodicity results for the stochastic navier-stokes equations: an introduction.*, Topics in mathematical fluid mechanics. Springer Berlin Heidelberg, (2013), 23–108.
- [DDG99] B Desjardins, E Dormy, and E Grenier, *Stability of mixed ekman-hartmann boundary layers*, Nonlinearity **12** (1999), no. 2, 181.
- [Doo48] J. L. Doob, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. **63** (1948), 393–421. MR 0025097 (9,598c)
- [EM01] W. E and J. C. Mattingly, *Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation*, Comm. Pure Appl. Math. **54** (2001), no. 11, 1386–1402. MR 1846802 (2002g:76075)
- [FMRR14] C. L. Fefferman, D. S. McCormick, J. C. Robinson, and J. L. Rodrigo, *Higher order commutator estimates and local existence for the non-resistive mhd equations and related models*, Journal of Functional Analysis **267** (2014), no. 4, 1035–1056.
- [FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2-D Navier-Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), no. 1, 119–141. MR MR1346374 (96g:35223)
- [FMRT01] C. Foias, O. Manley, R. Rosa, and R. Temam, *Navier-Stokes equations and turbulence*, Encyclopedia of Mathematics and its Applications, vol. 83, Cambridge University Press, Cambridge, 2001. MR 1855030 (2003a:76001)

- [FGHR15] J. Földes, N. E. Glatt-Holtz, and G. Richards, *Large prandtl number asymptotics in randomly forced turbulent convection*, 2015. arXiv preprint: arXiv preprint arXiv:1504.02904.
- [FGHRT15] J. Földes, N. E. Glatt-Holtz, G. Richards, and E. Thomann, *Ergodic and mixing properties of the boussinesq equations with a degenerate random forcing*, *J. Functional Analysis* **269** (2015), no. 8, arXiv preprint arXiv:1311.3620.
- [FGHRW15] J. Földes, N. E. Glatt-Holtz, G. Richards, and J. P. Whitehead, *Ergodicity in randomly forced rayleigh-bénard convection*, arXiv preprint arXiv:1511.01247 (2015).
- [FGHV14] S. Friedlander, N. E. Glatt-Holtz, and V. Vicol, *Inviscid limits for a stochastically forced shell model of turbulent flow*, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (2014), (to appear).
- [FRV14] S. Friedlander, W. Rusin, and V. Vicol, *The magneto-geostrophic equations: a survey*, *Proc. of the St. Petersburg Mathematical Society, Volume XV: Advances in Mathematical Analysis of Partial Differential Equations* (2014).
- [FS15] S. Friedlander and A. Suen, *Existence, uniqueness, regularity and instability results for the viscous magneto-geostrophic equation*, *Nonlinearity* **28** (2015), no. 9, 3193.
- [FV11a] S. Friedlander and V. Vicol, *Global well-posedness for an advection-diffusion equation arising in magneto-geostrophic dynamics*, *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, vol. 28, Elsevier, 2011, pp. 283–301.
- [FV11b] ———, *On the ill/well-posedness and nonlinear instability of the magneto-geostrophic equations*, *Nonlinearity* **24** (2011), no. 11, 3019.
- [FV12] ———, *Higher regularity of hölder continuous solutions of parabolic equations with singular drift velocities*, *Journal of Mathematical Fluid Mechanics* **14** (2012), no. 2, 255–266.
- [Fri95] U. Frisch, *Turbulence*, Cambridge University Press, Cambridge, 1995, The legacy of A. N. Kolmogorov. MR 1428905 (98e:76002)
- [HM06] M. Hairer and J. C. Mattingly, *Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing*, *Ann. of Math. (2)* **164** (2006), no. 3, 993–1032. MR 2259251 (2008a:37095)
- [HM08] ———, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations*, *Ann. Probab.* **36** (2008), no. 6, 2050–2091. MR 2478676 (2010i:35295)
- [HM11] ———, *A theory of hypoellipticity and unique ergodicity for semilinear stochastic pdes*, *Electron. J. Probab.* **16** (2011), no. 23, 658–738.
- [H67] L. Hörmander, *Hypoelliptic second order differential equations*, *Acta Math.* **119** (1967), 147–171. MR 0222474 (36 #5526)
- [KP03] Y. Kabanov and S. Pergamenschikov, *Two-scale stochastic systems: asymptotic analysis and control*, vol. 49, Springer Science & Business Media, 2003.
- [Kha60] R. Z. Khas'minskii, *Ergodic properties of recurrent diffusion processes and stabilization of the solution to the cauchy problem for parabolic equations*, *Theory of Probability & Its Applications* **5** (1960), no. 2, 179–196.
- [KS12] S. Kuksin and A. Shirikyan, *Mathematics of two-dimensional turbulence*, Cambridge Tracts in Mathematics, no. 194, Cambridge University Press, 2012.
- [Mas07] N. Masmoudi, *Examples of singular limits in hydrodynamics*, *Handbook of differential equations: evolutionary equations* **3** (2007), 195–275.
- [Mof78] H. K. Moffatt, *Field generation in electrically conducting fluids*, Cambridge University Press, 1978.
- [Mof08] ———, *Magnetostrophic turbulence and the geodynamo*, *IUTAM Symposium on Computational Physics and New Perspectives in Turbulence*, Springer, 2008, pp. 339–346.
- [ML94] H. K. Moffatt and D. E. Loper, *The magnetostrophic rise of a buoyant parcel in the earth's core*, *Geophysical Journal International* **117** (1994), no. 2, 394–402.
- [O'M74] R.E. O'Malley, *Introduction to singular perturbations*, 1974.
- [Par10] R. D. Parshad, *Asymptotic behaviour of the darcy-boussinesq system at large darcy-prandtl number*, *Discrete Contin. Dyn. Syst* **26** (2010), no. 4, 1441–1469.
- [Rom04] M. Romito, *Ergodicity of the finite dimensional approximation of the 3D Navier-Stokes equations forced by a degenerate noise*, *J. Statist. Phys.* **114** (2004), no. 1-2, 155–177. MR 2032128 (2005a:76128)
- [Vil08] C. Villani, *Optimal transport: old and new*, vol. 338, Springer Science & Business Media, 2008.
- [Wan04] X. Wang, *Infinite Prandtl number limit of Rayleigh-Bénard convection*, *Comm. Pure Appl. Math.* **57** (2004), no. 10, 1265–1282. MR 2069723 (2005d:76015)
- [Wan05] ———, *A note on long time behavior of solutions to the Boussinesq system at large Prandtl number*, *Nonlinear partial differential equations and related analysis*, *Contemp. Math.*, vol. 371, Amer. Math. Soc., Providence, RI, 2005, pp. 315–323. MR 2143874 (2006a:76105)
- [Wan07] ———, *Asymptotic behavior of the global attractors to the Boussinesq system for Rayleigh-Bénard convection at large Prandtl number*, *Comm. Pure Appl. Math.* **60** (2007), no. 9, 1293–1318. MR 2337505 (2009a:35196)
- [Wan08] ———, *Stationary statistical properties of Rayleigh-Bénard convection at large Prandtl number*, *Comm. Pure Appl. Math.* **61** (2008), no. 6, 789–815. MR 2400606 (2010b:76098)
- [Wu97] J. Wu, *Viscous and inviscid magneto-hydrodynamics equations*, *Journal d'analyse Mathématique* **73** (1997), no. 1, 251–265.

Juraj Földes

Department of Mathematics

Université Libre de Bruxelles

Web: <http://homepages.ulb.ac.be/~jfoldes/>

Email: juraj.foldes@ulb.ac.be

Susan Friedlander

Department of Mathematics

University of Southern California

Web: <http://www-bcf.usc.edu/~susanfri/>

Email: susanfri@usc.edu

Nathan Glatt-Holtz

Department of Mathematics

Virginia Polytechnic Institute and State University

Web: www.math.vt.edu/people/negh/

Email: negh@math.vt.edu

Geordie Richards

Department of Mathematics

University of Rochester

Web: www.math.rochester.edu/grichar5/

Email: g.richards@rochester.edu