

CONVERGENCE RATE ESTIMATES FOR THE LOW MACH AND ALFVÉN NUMBER THREE-SCALE SINGULAR LIMIT OF COMPRESSIBLE IDEAL MAGNETO-HYDRODYNAMICS

BIN CHENG, QIANGCHANG JU, AND STEVE SCHOCHET

ABSTRACT. Singular limits of the compressible ideal magneto-hydrodynamics equations at small Mach and Alfvén numbers are considered, in which those two parameters are allowed to tend to zero at different rates. After showing that solutions exist in a fixed time interval and developing explicit formulas for the limit system, convergence-rate estimates are obtained using an improvement of the uniform bounds shown previously by the authors, the time-integration method, and a detailed analysis of exact and approximate fast, intermediate, and slow modes.

35B25, 35L45.

1. INTRODUCTION

A uniform existence theorem and a convergence theorem as the small parameters tend to zero were recently developed [CJS18] for singular limits of symmetric hyperbolic systems of the form

$$A_0(\varepsilon_M \mathbf{V}) \mathbf{V}_t + \sum_{i=1}^d A_i(\mathbf{V}) \mathbf{V}_{x_i} = \frac{1}{\varepsilon_A} \mathcal{L}_A \mathbf{V} + \frac{1}{\varepsilon_M} \mathcal{L}_M \mathbf{V}, \quad (1.1)$$

where ε_A and ε_M are small positive parameters and \mathcal{L}_A and \mathcal{L}_M are skew-adjoint constant-coefficient first-order differential operators. If $\frac{\varepsilon_A}{\varepsilon_M}$ tends to zero as the parameters tend to zero then systems of the form (1.1) have three time scales: $O(\frac{1}{\varepsilon_A})$, $O(\frac{1}{\varepsilon_M})$, and $O(1)$.

In this paper we begin the study of the rate of convergence of solutions of three-scale singular limits to corresponding solutions of their limit equations, an issue that was not considered in [CJS18]. The convergence rate in the general case will undoubtedly be very complicated, since in general many different limit systems are obtained for different power-law relations between the two small parameters as they both tend to zero [CJS18, §4]. In this paper we study the particular case of the low Mach and Alfvén number limit of the compressible ideal magneto-hydrodynamics (MHD) equations in the presence of a large uniform magnetic field, which was the motivating example for our work. As we will show, that system has essentially only one limit system, although the limit

$$\mu_{\text{lim}} := \lim_{\varepsilon_A, \varepsilon_M \rightarrow 0} \mu \quad (1.2)$$

of the ratio

$$\mu := \frac{\varepsilon_A}{\varepsilon_M} \quad (1.3)$$

Date: 14-03-2019

B. Cheng

Department of Mathematics, University of Surrey, Guildford, GU2 7XH, United Kingdom

E-mail: b.cheng@surrey.ac.uk

Q. Ju

Institute of Applied Physics and Computational Mathematics, P.O. Box 8009, Beijing 100088, China

E-mail: ju_qiangchang@iapcm.ac.cn

S. Schochet

School of Mathematical Sciences, Tel-Aviv University, Tel Aviv 69978, Israel

E-mail: schochet@tauex.tau.ac.il

appears in that limit system as a parameter. Even for the MHD system the study of the convergence rate is much more intricate than for standard two-scale singular limits [KM81, KM82, Maj84].

The MHD system in three spatial dimensions that we study, derived in Appendix A from a standard formulation of that system, is

$$a(\varepsilon_M r) (\partial_t r + (\mathbf{u} \cdot \nabla) r) + \frac{a(\varepsilon_M r) \rho(\varepsilon_M r)}{\varepsilon_M} \nabla \cdot \mathbf{u} = 0 \quad (1.4a)$$

$$\rho(\varepsilon_M r) (\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u}) + \frac{a(\varepsilon_M r) \rho(\varepsilon_M r)}{\varepsilon_M} \nabla r + \nabla \frac{|\mathbf{b}|^2}{2} - (\mathbf{b} \cdot \nabla) \mathbf{b} = \varepsilon_A^{-1} (\partial_z \mathbf{b} - \nabla b_3), \quad (1.4b)$$

$$\partial_t \mathbf{b} + (\mathbf{u} \cdot \nabla) \mathbf{b} + (\nabla \cdot \mathbf{u}) \mathbf{b} - (\mathbf{b} \cdot \nabla) \mathbf{u} = \varepsilon_A^{-1} (\partial_z \mathbf{u} - \mathbf{e}_z \nabla \cdot \mathbf{u}), \quad (1.4c)$$

$$\nabla \cdot \mathbf{b} = 0, \quad (1.4d)$$

where ε_M has been rescaled to make the pressure law $p(\rho)$ satisfy

$$p'(1) = 1 \quad (1.5)$$

for notational convenience, and

$$a(s) := \frac{p'(1+s)}{1+s}, \quad \rho(s) := 1+s. \quad (1.6)$$

Note that in (1.4) ‘‘symmetric terms’’ are grouped into the same vertical column.

As noted in Appendix A, the divergence-free condition (1.4d) on the magnetic field is preserved by the dynamics of (1.4c), and so is just a restriction on the initial data. Hence straightforward calculations show that the system (1.4) has the form (1.1), with $\mathbf{V} = (r, \mathbf{u}, \mathbf{b})$.

When μ is fixed so that only two time scales are present, uniform existence and convergence results were obtained in [KM81] for the limit of smooth solutions of initial value problems of systems (1.1) as the small parameter tends to zero. Moreover, whenever $\mu_{\text{lim}} > 0$ corresponding results can be proven via cosmetic changes to the theory in [KM81].

We therefore focus on the more challenging case when $\mu \rightarrow 0$, although our results will be phrased so as to remain valid when $\mu_{\text{lim}} > 0$. We prove the following two theorems about the behavior of solutions to the MHD system (1.4) as the small parameters ε_A and ε_M tend to zero, of which the first is an application to system (1.4) of an improvement to be shown in §2 of the results of [CJS18] plus an extension of results from [CJS18] itself, and the second is our main result.

Before stating the theorems we define some notations and three operators that will appear in the limit equations or the convergence-rate estimates. First, we let $\|\cdot\|_k$ denote the H^k norm. Next, for any vector \mathbf{w} let its ‘‘horizontal’’ part w_h denote its x, y components (w_1, w_2) , and let $\mathcal{P}_h^{\text{div}}$ denote the *two-dimensional* Leray-Helmholtz projection onto divergence-free velocity fields in the x, y plane or 2-torus, i.e.

$$\mathcal{P}_h^{\text{div}} w_h := w_h - \nabla_h \Delta_h^{-1} \nabla_h \cdot w_h, \quad \text{where } \nabla_h := \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}, \quad \Delta_h := \partial_x^2 + \partial_y^2. \quad (1.7)$$

Define the Alfvén and Mach operators

$$\mathcal{L}_A \mathbf{V} := \begin{pmatrix} 0 \\ (-\nabla_h b_3 + \partial_z b_h) \\ 0 \\ \begin{pmatrix} \partial_z u_h \\ -\nabla_h \cdot u_h \end{pmatrix} \end{pmatrix}, \quad \mathcal{L}_M \mathbf{V} := \begin{pmatrix} -\nabla \cdot \mathbf{u} \\ \begin{pmatrix} -\nabla_h r \\ -\partial_z r \end{pmatrix} \\ 0_3 \end{pmatrix}. \quad (1.8)$$

Finally, the full average $\text{av } f$ and the vertical average $\mathfrak{a}_z f$ of any function f defined on the 3-torus \mathbb{T}^3 are

$$\text{av } f := \frac{\iiint f \, dx \, dy \, dz}{\iiint 1 \, dx \, dy \, dz}, \quad (\mathfrak{a}_z f)(x, y) := \frac{\int f(x, y, z) \, dz}{\int 1 \, dz}. \quad (1.9)$$

Throughout this paper c and C denote positive constants that are independent of ε_A and ε_M , which may take different values in each appearance.

Theorem 1.1. *Let $n \geq 3$ be an integer, and assume that the small parameters ε_A and ε_M satisfy*

$$0 < \varepsilon_M \leq \varepsilon_M^0 \quad \text{and} \quad \varepsilon_A \geq c \varepsilon_M^{1+\frac{1}{n-1}}. \quad (1.10)$$

Assume that the spatial domain is \mathbb{R}^3 or \mathbb{T}^3 and that the initial data $\mathbf{V}^0 := (r^0, \mathbf{u}^0, \mathbf{b}^0)$, which may depend on the parameters ε_A and ε_M , are uniformly bounded in H^n and satisfy the constraint (1.4d) and the “well-preparedness” condition

$$\|(\varepsilon_A^{-1} \mathcal{L}_A + \varepsilon_M^{-1} \mathcal{L}_M) \mathbf{V}^0\|_{n-1} \leq c. \quad (1.11)$$

Then there exist fixed positive T and K such that for $(\varepsilon_A, \varepsilon_M)$ satisfying (1.10) the solution to (1.4) having the initial data \mathbf{V}^0 exists for $0 \leq t \leq T$ and satisfies

$$\sup_{0 \leq t \leq T} \left[\|\mathbf{V}\|_n + \|\mathbf{V}_t\|_0 + \sum_{j=1}^n \varepsilon_M^{j-1} \left(\min\left(\frac{\varepsilon_A}{\varepsilon_M}, 1\right) \right)^{n-1} \|\partial_t^j \mathbf{V}\|_{n-j} \right] \leq K; \quad (1.12)$$

in particular the solution is uniformly bounded in H^n .

Assume in addition that as $(\varepsilon_A, \varepsilon_M)$ satisfying (1.10) tend to zero, their ratio $\frac{\varepsilon_A}{\varepsilon_M}$ converges to some value μ_{lim} and the initial data \mathbf{V}^0 converges in H^n to $\bar{\mathbf{V}}^0$. Then the solution $\mathbf{V} = (r, \mathbf{u} = (u_h, u_3), \mathbf{b} = (b_h, b_3))$ of the MHD system (1.4) with initial data \mathbf{V}^0 converges in $C^0([0, T]; H^{n-\alpha})$ for every $\alpha > 0$. Its limit is independent of z , and is identically zero if the spatial domain is \mathbb{R}^3 , while if the spatial domain is \mathbb{T}^3 it is the unique solution $\bar{\mathbf{V}} = (\bar{r}, (\bar{u}_h, \bar{u}_3), (\bar{b}_h, \bar{b}_3))$ of the limit system

$$(1 + \mu_{\text{lim}}^2) [\partial_t \bar{r} + (\bar{u}_h \cdot \nabla_h) \bar{r}] + \mu_{\text{lim}} (\bar{b}_h \cdot \nabla_h) \bar{u}_3 = 0, \quad (1.13a)$$

$$\mathcal{D}_h^{\text{div}} (\partial_t \bar{u}_h + (\bar{u}_h \cdot \nabla_h) \bar{u}_h - (\bar{b}_h \cdot \nabla_h) \bar{b}_h) = 0, \quad \nabla_h \cdot \bar{u}_h = 0, \quad (1.13b)$$

$$\partial_t \bar{u}_3 + (\bar{u}_h \cdot \nabla_h) \bar{u}_3 + \mu_{\text{lim}} (\bar{b}_h \cdot \nabla_h) \bar{r} = 0, \quad (1.13c)$$

$$\partial_t \bar{b}_h + (\bar{u}_h \cdot \nabla_h) \bar{b}_h - (\bar{b}_h \cdot \nabla_h) \bar{u}_h = 0, \quad \nabla_h \cdot \bar{b}_h = 0, \quad (1.13d)$$

$$\bar{b}_3 = \text{av} \bar{b}_3^0 - \mu_{\text{lim}} (\bar{r} - \text{av} \bar{r}^0) \quad (1.13e)$$

having initial data $\bar{\mathbf{V}}^0$.

Note that if $\mu_{\text{lim}} > 0$ then the limit equation (1.13a) for the density perturbation includes a term involving the magnetic field, and the limit equation (1.13c) for the vertical fluid velocity contains a term involving the product of the magnetic field and derivatives of the density, even though no such terms appear in the corresponding equations of the original MHD system (1.4). Those terms arise on account of the coupling between the density perturbation and the vertical magnetic field that occurs in (1.13e) when μ_{lim} is nonzero.

Although the above uniform existence and convergence theorem does not require any assumption about the rate at which the initial data converge to their limit, such an assumption is obviously required in order to obtain a rate of convergence of solutions of the PDE. In the following convergence rate theorem we assume that the initial data has the form developed in §3, which is a specialization of the general form of initial data satisfying (1.11). Specifically, after expanding the initial data in powers of the small parameters and their ratio, the leading-order terms are assumed to be independent of the small parameters in order to avoid degrading the convergence rate. Nevertheless, full generality is retained as far as convergence rates are concerned because their proofs will *not* rely on the compactness of sequences of solutions, a feature that allows us to consider one solution to the MHD system with fixed $\varepsilon_A, \varepsilon_M$ and initial data and compare it to the limit system(s) equipped with initial data that are the leading order terms of the (fixed) initial data for the MHD system.

It will be convenient to write the convergence-rate results using just powers of ε_M , by defining a parameter ν determined by

$$\varepsilon_A = \varepsilon_M^{1+\nu}, \quad \text{or, equivalently,} \quad \mu = \varepsilon_M^\nu, \quad \text{i.e.,} \quad \nu := \frac{\ln(\frac{1}{\varepsilon_A})}{\ln(\frac{1}{\varepsilon_M})} - 1 = \frac{\ln(\frac{1}{\mu})}{\ln(\frac{1}{\varepsilon_M})}. \quad (1.14)$$

Since the case when $\varepsilon_A \geq \varepsilon_M$ is covered by the theory of two-scale singular limits or a trivial generalization of that theory, in developing the convergence rate we will assume for notational simplicity that

$$\varepsilon_A < \varepsilon_M, \quad \text{i.e., } \mu < 1 \text{ and } \nu > 0. \quad (1.15)$$

Theorem 1.2. *Assume that both the original and additional assumptions of Theorem 1.1 hold, that the initial data for the MHD system have the more specific form (3.30), (3.32), (3.36), that (1.15) holds, and that the spatial domain is \mathbb{T}^3 . Then there is a constant c independent of ε_A and ε_M such that for all $t \in [0, T]$,*

$$\begin{aligned} & \| (1 - \mathcal{P}_h^{\text{div}} \mathbf{a}_z) \mathbf{u}_h \|_j + \| (1 - \mathbf{a}_z) \mathbf{b}_h \|_j \\ & + \| b_3 - \text{av} \bar{b}_3^0 + \mu (\mathbf{a}_z r - \text{av} \bar{r}^0) \|_j \leq c \varepsilon_M^{1-(j-1)\nu}, \quad j = 0, \dots, n-1, \end{aligned} \quad (1.16a)$$

$$\| \mathcal{P}_h^{\text{div}} \mathbf{a}_z \mathbf{u}_h - \bar{\mathbf{u}}_h \|_{n-2} + \| \mathbf{a}_z \mathbf{b}_h - \bar{\mathbf{b}}_h \|_{n-2} \leq c \varepsilon_M, \quad (1.16b)$$

$$\| (1 - \mathbf{a}_z) r \|_j + \| (1 - \mathbf{a}_z) u_3 \|_j \leq c \varepsilon_M^{1-(j-1)\nu}, \quad j = 1, \dots, n-1, \quad (1.16c)$$

$$\| \mathbf{a}_z r - \bar{r} \|_{n-2} + \| \mathbf{a}_z u_3 - \bar{u}_3 \|_{n-2} \leq c \left[\varepsilon_M^{1-\max(n-5,0)\nu} + |\mu - \mu_{\text{lim}}| \right]. \quad (1.16d)$$

Moreover, there exist $O(1)$ correctors $(r^{(\text{cor})}, u_3^{(\text{cor})})$ defined in (5.88)–(5.89) such that

$$\begin{aligned} & \| \mathbf{a}_z r - (\bar{r} + \frac{\mu - \mu_{\text{lim}}}{1 + \mu^2} r^{(\text{cor})}) \|_{n-2} + \| \mathbf{a}_z u_3 - (\bar{u}_3 + (\mu - \mu_{\text{lim}}) u_3^{(\text{cor})}) \|_{n-2} \\ & \leq c \varepsilon_M^{1-\max(n-5,0)\nu}. \end{aligned} \quad (1.16e)$$

Under the scaling of (1.10) with ε_M^0 and c suitably chosen, $\nu \leq \frac{1}{n-1} + \ln \frac{1}{\varepsilon_M} c < \frac{1}{n-2}$ so all powers of ε_M appearing in (1.16) are positive, which means that a nontrivial rate of convergence is obtained over the full range of allowed values of ν .

When the initial data are chosen so that the first time derivative of solutions are uniformly bounded at time zero then the convergence rate for solutions of two-scale singular limit systems is typically first order in the single small parameter appearing in those systems [Sch88, Sch94]. However, that result depends crucially on the uniform boundedness of the first time derivative of solutions being propagated for positive time. That propagation result does *not* generally hold for three-scale systems of the form (1.1) satisfying $\frac{\varepsilon_A}{\varepsilon_M} \rightarrow 0$, as the example in [CJS18, §2] shows. Hence for three-scale singular limits the rate of convergence obtained in (1.16) is slower than for two-scale singular limits, and the proofs are more complicated. Nevertheless, the convergence rates (1.16a) and (1.16c) tend to the standard two-scale $O(\varepsilon_M)$ convergence rate as the parameter ν tends to zero, which corresponds to the three-scale system tending to a two-scale system. In order to obtain such a rate of convergence, asymptotically consistent with the two-scale rate, it is necessary to improve the estimate $\varepsilon_M \|\mathbf{V}_t\|_{n-1} \leq c$ obtained in [CJS18], since using that bound in our estimates would have led to a bound $c \varepsilon_M^\nu$ in (1.16a), which tends to the trivial $O(1)$ estimate as the three-scale limit tends to the two-scale case in which $\nu = 0$. The required improved bound (1.12) will be proven for general systems (1.1) in §2.

Moreover, an $O(\varepsilon_M)$ rate of convergence *has* been obtained in (1.16b), and for the H^1 norm in (1.16a) and (1.16c). The latter estimates are obtained by using the fact that the first time derivative of the solution *does* remain uniformly bounded in L^2 even though, as noted above, it does not remain uniformly bounded in H^{n-1} . The H^j estimates (1.16a) and (1.16c) for intermediate values of j are obtained using interpolation. Those estimates are the starting point for an improved estimate for z -averages of products derived in Appendix C, which is combined with techniques described below to yield (1.16b).

One of the main techniques used in the proof of the estimates (1.16) is partitioning the solution into fast, intermediate, and slow modes, and then analyzing each mode separately. Although projections of the solution onto the nonzero and zero eigenspaces of the large operator are often utilized in the theory of two-scale singular limits (e.g. [Che14, §2]), in the two scale case those projections are fixed operators, while in the three-scale

case the exact eigenspaces of the large operator $\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M$ having eigenvalues of sizes strictly $O(\frac{1}{\varepsilon_A})$, $O(\frac{1}{\varepsilon_M})$, and $o(\frac{1}{\varepsilon_M})$, respectively, depend on the parameter μ . For simplicity we will therefore define the fast, intermediate, and slow modes to be the fixed projections onto the limits as $\mu \rightarrow 0$ of the exact eigenspaces. Formulas for all the modes and estimates for the fast and intermediate modes in terms of the operator $\mathcal{L}_A + \mu \mathcal{L}_M$ applied to \mathbf{V} will be determined in §3. However, certain estimates in §5, in particular those for the intermediate modes, will require use of the exact μ -dependent modes, which makes their derivations more complicated than in the two-scale case. Even so, the partitioning here is much simpler than for general systems of the form (1.1), for which $\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M$ also has eigenvalues of sizes $O(\frac{\mu^j}{\varepsilon_A})$ for $2 \leq j \leq \frac{1}{\nu} + 1$ [CJS18, §4]; we will show in §3 that such eigenvalues are not present for the MHD system. The projections defined in §3 will also be used there to determine the appropriate form of the initial data used in Theorem 1.2.

For the two-scale singular limit of the compressible Euler equations without magnetic fields, the time-integration method developed by one of the authors of this paper together with the special structure of those equations yields an $O(\varepsilon_M^2)$ estimate for the difference between the incompressible component of solutions to the original equations and corresponding solutions of the incompressible limit equations [Che12, Che14]. In those papers the solution is separated into a slow part that satisfies the constraints obeyed by the limit system and a fast part that does not, and different estimates are given for each part. In similar fashion, (1.16) gives separate estimates for the slow, intermediate, and fast parts of the solution. Although estimate (1.16b) is not quadratic in the size of the fast parts of the horizontal fluid velocity and horizontal magnetic fields, the time-integration method will nevertheless play an important role in its derivation, because that method helps to mitigate the effect of the lack of uniform boundedness of the first time derivatives.

The corrector estimate (1.16e) has been included because estimate (1.16d) is rather weak in the main case of interest in which $\mu_{\text{lim}} = 0$. The constraint (1.10) together with the relation (1.14) between the parameters implies that $\varepsilon_M^{1-\max(n-5,0)} \ll \varepsilon_M^\nu = \mu$, so when $\mu_{\text{lim}} = 0$ the bound in (1.16d) reduces to $O(\varepsilon_M^\nu)$, which is much larger than the other error bounds in (1.16). The improved estimate (1.16e) involving the corrector shows that (1.16d) is in fact sharp in this case, and gives a formula for principal error term.

The uniform existence part of Theorem 1.1, which concerns solutions to (1.4), will be deduced in §4 from the uniform existence result for (1.1) in §2. When $\mu_{\text{lim}} = 0$ the convergence of solutions to (1.4) can be deduced from the convergence theorem in [CJS18, §4] for the general system (1.1), as will be indicated in §4; in particular, the general convergence result simplifies considerably for the MHD equations thanks to the absence shown in §3 of eigenvalues of the large operator of size $O(\frac{\mu^j}{\varepsilon_A})$ for $j \geq 2$. However, in order to treat the cases $\mu_{\text{lim}} = 0$ and $\mu_{\text{lim}} > 0$ together, a direct proof of the convergence part of Theorem 1.1 will be presented in §4. The estimates (1.16) of Theorem 1.2 are direct consequences of the estimates (5.4), (5.16), (5.46), (5.69), and (5.90) for the fast, intermediate, and slow modes proven in §5.

The restriction (1.10) on the relation between ε_A and ε_M needed to obtain uniform existence does not allow taking the limit when ε_A tends to zero but ε_M is fixed. As shown in [BK82] and [CJS18], if the initial data for a system of the form (1.1) imply that the first $n \geq \lfloor \frac{d}{2} \rfloor + 2$ time derivatives of the solution are uniformly bounded at time zero rather than only the first time derivative being uniformly bounded, then the restriction (1.10) is not needed, so it is possible to let ε_A tend to zero with ε_M fixed. The limit equations for a slight variant of the MHD system (1.4) for that case were obtained in [BK82]. However, the additional conditions placed on the initial data are very restrictive: while condition (1.11) *only* requires that the fast and intermediate modes of the initial data be of size ε_A and ε_M , respectively, the “super-well-posed” condition that time derivatives through order n be bounded initially determines the expansion of the fast and intermediate modes of the initial data up to the n^{th} powers of the small parameters, and each term in the expansion depends nonlinearly on the slow mode of the initial data. Furthermore, that restriction cannot be removed by the perturbative technique that estimates in H^{n-1} the difference between solutions with super-well-posed initial data and solutions having perturbed initial

data that only yield one bounded time derivative at time zero, because the example in [CJS18, §2] implies that no such three-scale perturbation theorem can possibly hold. However, the estimates in §5 make use of a perturbation result concerning both initial data and inhomogeneous terms for symmetric-hyperbolic systems without large terms, which is proven in Appendix B.

2. IMPROVED UNIFORM BOUND

In the notation in this paper, in which the smaller parameter is denoted ε_A rather than δ and the larger small parameter is denoted ε_M rather than ε , Theorem 3.6 of [CJS18] says that solutions to the three-scale symmetric hyperbolic system (1.1) satisfy the uniform estimate

$$\max_{0 \leq t \leq T} \left[\sum_{j=0}^n \varepsilon_M^j \|\partial_t^j \mathbf{V}\|_{n-j} + \|\mathbf{V}_t\|_0 \right] \leq K \quad (2.1)$$

for some positive T and finite K provided that:

- Assumption 2.1.** (1) *the operators \mathcal{L}_A and \mathcal{L}_M are constant-coefficient differential operators of order at most one and are skew-adjoint on L^2 ,*
(2) *$n \geq s_0 + 1$, where $s_0 := \lfloor \frac{d}{2} \rfloor + 1$ is the Sobolev embedding exponent, i.e., the smallest integer s for which $\|f\|_{L^\infty} \leq c\|f\|_{H^s}$,*
(3) *the matrices A_i are smooth symmetric functions for $j \geq 0$ and the matrix A_0 is positive definite,*
(4) *the small parameters are restricted to the region (1.10).*
(5) *the initial data \mathbf{V}^0 are uniformly bounded in H^n and satisfy (1.11).*

Actually, [CJS18, Theorem 3.6] is only stated for $n = s_0 + 1$, but the proof remains valid in the more general case and, as noted in [CJS18, §2], the value of n can be increased arbitrarily by artificially adding extra dimensions that the system and initial data do not depend on. In this section we show the improvement (1.12) of (2.1).

Theorem 2.2. *When Assumption 2.1 holds there exist a positive T and finite K such that (1.12) holds.*

Remark 2.3. *Assumption (1.10) ensures that the new estimate (1.12) is at least as strong as the old estimate (2.1), and is stronger when $\limsup_{\varepsilon_M \rightarrow 0} v < \frac{1}{n-1}$. In particular, for the two scale case in which $v \rightarrow 0$, (1.12) recovers the uniform H^{n-1} estimate for \mathbf{V}_t which is not guaranteed by (2.1).*

Proof. Theorem 2.2 differs from Theorem 3.6 of [CJS18] only by having different weights multiplying the norms of time derivatives. Hence it suffices to show that in all places in the proof of [CJS18, Theorem 3.6] where the use of the weights in (2.1) was justified the use of the weights in (1.12) is also justified. There are only two such places, namely where it was shown that the weighted sum of norms is bounded at time zero and where it was shown that the small parameters scale out of the estimate for the time derivative of an appropriately weighted energy.

As noted in the proof of Lemma 3.5 of [CJS18], assumption (1.11) ensures that $\|\mathbf{V}_t|_{t=0}\|_{n-1}$ is bounded uniformly in the small parameters, and hence the PDE (1.1) yields $\|\partial_t^j \mathbf{V}|_{t=0}\|_{n-j} \leq c \varepsilon_A^{1-j}$ for $1 \leq j \leq n$. Therefore, for $1 \leq j \leq n$,

$$\begin{aligned} \varepsilon_M^{j-1} \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1} \|\partial_t^j \mathbf{V}|_{t=0}\|_{n-j} &\leq \varepsilon_M^{j-1} \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1} \left(c \varepsilon_A^{1-j} \right) \\ &= c \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1} \left(\frac{\varepsilon_A}{\varepsilon_M} \right)^{1-j} \leq c \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1} \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{1-j} \\ &= c \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-j} \leq c, \end{aligned}$$

which shows that the weighted sum of norms in (1.12) is also bounded at time zero uniformly in the small parameters.

The energy estimate both in [CJS18] and here makes use of the norms

$$\|f\|_{\ell, A_0} := \sqrt{\sum_{0 \leq |\alpha| \leq \ell} \int (D^\alpha f)^T A_0 (\varepsilon_M \mathbf{V}) D^\alpha f dx}, \quad (2.2)$$

where \mathbf{V} is a solution to (1.1) and $D^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}$. As shown in [CJS18], in order to prove a weighted energy estimate like (2.1) or (1.12) it suffices to obtain a uniform bound for

$$E := \|\mathbf{V}\|_{n, A_0}^2 + \|\mathbf{V}_t\|_{0, A_0}^2 + \sum_{j=1}^n w_j^2 \|\partial_t^j \mathbf{V}\|_{n-j, A_0}^2, \quad (2.3)$$

where the weights w_j are ε_M^j for the estimate (2.1) or

$$w_j = \varepsilon_M^{j-1} \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1}, \quad 1 \leq j \leq n \quad (2.4)$$

for the estimate (1.12). Moreover, in the estimates [CJS18, (3.12), (3.24)] for $\frac{d}{dt}E$, the only facts used about the weights w_j to prove a uniform bound for E are that for some finite c that may be different in each appearance

$$\varepsilon_M \leq c w_1, \quad (2.5a)$$

$$\varepsilon_M w_j \leq c w_{j+1} \quad \text{for } 1 \leq j \leq n-1, \quad (2.5b)$$

$$w_k \leq c \prod_{j=1}^J w_{k_j} \quad \text{whenever } \sum_{j=1}^J k_j = k, \quad (2.5c)$$

and

$$\varepsilon_M w_k \leq c \prod_{j=1}^J w_{k_j} \quad \text{whenever } \sum_{j=1}^J k_j = k+1. \quad (2.5d)$$

Since (2.5d) can be obtained by substituting (2.5c) with k replaced by $k+1$ into (2.5b) with j set equal to k , it suffices to prove (2.5a)–(2.5c). The definitions (2.4) imply that (2.5b) holds provided that c there is at least one, while both (2.5a) and (2.5c) reduce to the condition $\varepsilon_M \leq c \left(\min \left(\frac{\varepsilon_A}{\varepsilon_M}, 1 \right) \right)^{n-1}$ that is equivalent to (1.10). \square \square

Combining the estimate (1.12) with the standard Sobolev-space interpolation inequality

$$\|f\|_r \leq C_{r,s} \|f\|_s^{\frac{r}{s}} \|f\|_0^{1-\frac{r}{s}} \quad \text{for } 0 \leq r \leq s \quad (2.6)$$

(e.g. [Maj84, (2.32)]), yields the following result.

Corollary 2.4. *When the conditions of Theorem 2.2 hold and $\mu \leq 1$ then*

$$\|\mathbf{V}_t\|_j \leq c \mu^{-j} = c \varepsilon_M^{-j\nu} \quad j = 0, \dots, n-1. \quad (2.7)$$

3. ANALYSIS OF THE LARGE OPERATOR

Following [CJS18, §4] but without treating each Fourier mode separately, let \mathbb{P}^0 denote the L^2 -orthogonal projection operator onto the nullspace of \mathcal{L}_A , and let \mathbb{P}^1 denote the L^2 -orthogonal projection operator onto the nullspace of $\mathbb{P}^0 \mathcal{L}_M \mathbb{P}^0$. Since \mathbb{P}^0 , being a projection, satisfies $\mathbb{P}^0(I - \mathbb{P}^0) = 0$, $R(I - \mathbb{P}^0) \subseteq N(\mathbb{P}^0 \mathcal{L}_M \mathbb{P}^0) = R(\mathbb{P}^1) = N(I - \mathbb{P}^1)$. Hence $R(I - \mathbb{P}^0) \perp R(I - \mathbb{P}^1)$, so

$$(I - \mathbb{P}^0)(I - \mathbb{P}^1) = 0 = (I - \mathbb{P}^1)(I - \mathbb{P}^0). \quad (3.1)$$

Expanding the factors in (3.1) shows that $\mathbb{P}^0\mathbb{P}^1 = \mathbb{P}^1\mathbb{P}^0$, which implies that $\mathbb{P} := \mathbb{P}^0\mathbb{P}^1$ is an orthogonal projection operator satisfying $\mathbb{P}(I - \mathbb{P}^j) = 0 = (I - \mathbb{P}^j)\mathbb{P}$ for $j \in \{0, 1\}$. Moreover, (3.1) together with the definition of \mathbb{P} yield

$$(I - \mathbb{P}^0) + (I - \mathbb{P}^1) + \mathbb{P} = I + (I - \mathbb{P}^0)(I - \mathbb{P}^1) = I,$$

which shows that the sum of the fast, intermediate, and slow modes defined by

$$\mathbf{V}^F := (I - \mathbb{P}^0)\mathbf{V}, \quad \mathbf{V}^I := (I - \mathbb{P}^1)\mathbf{V}, \quad \mathbf{V}^S := \mathbb{P}\mathbf{V} \quad (3.2)$$

satisfy

$$\mathbf{V}^F + \mathbf{V}^I + \mathbf{V}^S = \mathbf{V}. \quad (3.3)$$

Moreover, the mutual L^2 orthogonality of the projections defining the modes together with the fact that the \mathbb{P}^j are projections onto the null spaces of constant-coefficient differential operators and so commute with derivatives implies that they are also orthogonal in any H^k . Therefore combining estimates for different modes to obtain an estimate for the full solution yields estimates that are as sharp up to a constant factor as the component estimates.

These general facts about projections were developed in [Che12, CM13] for the two-scale case. Since the operators \mathcal{L}_A and \mathcal{L}_M used here have constant coefficients, our projections \mathbb{P}^j operate separately on each Fourier mode, so the results for them also follow from perturbation theory of (anti)symmetric operators [Kat82, CJS18]. Since the modes as defined above do not depend on the parameter μ they are in general *not* the direct sums of the eigenspaces of $\mathcal{L}_A + \mu\mathcal{L}_M$ whose eigenvalues are strictly $O(1)$, $O(\mu)$, and $o(\mu)$, respectively, but are rather the limits as $\mu \rightarrow 0$ of those direct sums of eigenspaces [Kat82], [CJS18, §4].

We now calculate the projections, the fast, intermediate, and slow modes, and some elementary results about them for the specific case of the MHD system (1.4). For brevity, we restrict consideration to $\mathbf{V} = (r, \mathbf{u}, \mathbf{b})$ satisfying $\nabla \cdot \mathbf{b} = 0$, and since we only consider initial data and solutions satisfying that constraint that restriction causes no difficulties.

Lemma 3.1. *Assume that the spatial domain is \mathbb{T}^3 and that $\mathbf{V} = (r, \mathbf{u}, \mathbf{b})$ where \mathbf{b} satisfies $\nabla \cdot \mathbf{b} = 0$. Then*

- (1) $\mathcal{L}_A \mathbf{V} = \mathbf{0}$ iff $\partial_z \mathbf{u}_h = \mathbf{0} = \partial_z \mathbf{b}_h$, $\nabla_h \cdot \mathbf{u}_h = 0$, and $b_3 = \text{av} b_3$.
- (2) $\left(\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M \right) \mathbf{V} = \mathbf{0}$ iff

$$\partial_z \mathbf{V} = \mathbf{0}, \quad \nabla_h \cdot \mathbf{u}_h = 0, \quad b_3 = \text{av} b_3 - \frac{\varepsilon_A}{\varepsilon_M} (r - \text{av} r), \quad (3.4)$$

where the operator av was defined in (1.9).

- (3) *The formulas for the projections are*

$$\mathbb{P}^0 = \begin{pmatrix} 1 & & & \\ & \left(\mathcal{P}_h^{\text{div}} \mathbf{a}_z I_{2 \times 2} \right) & & \\ & & \left(\mathbf{a}_z I_{2 \times 2} \right) & \\ & & & \text{av} \end{pmatrix}, \quad \mathbb{P}^1 = \begin{pmatrix} \mathbf{a}_z & & & \\ & \left(I_{2 \times 2} \right) & & \\ & & \mathbf{a}_z & \\ & & & I_{3 \times 3} \end{pmatrix}, \quad (3.5)$$

and

$$\mathbb{P} = \begin{pmatrix} \mathbf{a}_z & & & \\ & \left(\mathcal{P}_h^{\text{div}} \mathbf{a}_z I_{2 \times 2} \right) & & \\ & & \mathbf{a}_z & \\ & & & \left(\mathbf{a}_z I_{2 \times 2} \right) \\ & & & & \text{av} \end{pmatrix}, \quad (3.6)$$

where all missing entries vanish, and $\mathcal{P}_h^{\text{div}}$ and \mathbf{a}_z are defined in (1.7) and (1.9).

- (4) *All eigenvalues of $\mathcal{L}_A + \mu\mathcal{L}_M$ that are $o(\mu)$ are identically zero.*

(5) Using the notations $\mathbf{V}^\ell = (r^\ell, \mathbf{u}^\ell, \mathbf{b}^\ell)$ for $\ell \in \{F, I, S\}$ and $\mathbf{w}^\ell = (w_h^\ell, w_3^\ell)$ for $w \in \{\mathbf{u}, \mathbf{b}\}$, the formulas for the fast, intermediate and slow modes are

$$\begin{pmatrix} r^F \\ \mathbf{u}^F \\ \mathbf{b}^F \end{pmatrix} = \begin{pmatrix} 0 \\ (1 - \alpha_z \mathcal{P}_h^{\text{div}} \mathbf{u}_h) \\ 0 \\ (1 - \alpha_z) \mathbf{b}_h \\ (1 - \alpha_v) b_3 \end{pmatrix}, \quad \begin{pmatrix} r^I \\ \mathbf{u}^I \\ \mathbf{b}^I \end{pmatrix} = \begin{pmatrix} (1 - \alpha_z) r \\ 0_2 \\ (1 - \alpha_z) u_3 \\ \mathbf{0}_3 \end{pmatrix}, \quad (3.7)$$

$$\begin{pmatrix} r^S \\ \mathbf{u}^S \\ \mathbf{b}^S \end{pmatrix} = \begin{pmatrix} \alpha_z r \\ \mathcal{P}_h^{\text{div}} \alpha_z \mathbf{u}_h \\ \alpha_z u_3 \\ \alpha_z \mathbf{b}_h \\ \alpha_v b_3 \end{pmatrix}.$$

In particular,

$$\nabla \cdot \mathbf{b}^F = 0, \quad \nabla \cdot \mathbf{b}^I = 0, \quad \nabla \cdot \mathbf{b}^S = 0, \quad (3.8)$$

$$\mathbf{V}^S \text{ is independent of } z, \quad (3.9)$$

and

$$\nabla_h \cdot \mathbf{u}_h^S = 0 = \nabla_h \cdot \mathbf{b}_h^S. \quad (3.10)$$

(6) For any nonnegative integer j , there are constants c_1 and c_2 such that

$$\begin{aligned} c_1 \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j &\leq \|\partial_z \mathbf{u}_h^F\|_j + \|\nabla_h \cdot \mathbf{u}_h^F\|_j + \|\partial_z (\mathbf{b}_h^F - \mu \Delta^{-1} \nabla_h \partial_z r^I)\|_j \\ &\quad + \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1} + \mu [\|\partial_z r^I\|_j + \|\partial_z u_3^I\|_j] \\ &\leq c_2 \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j. \end{aligned} \quad (3.11)$$

Proof. Applying \mathcal{L}_A to \mathbf{V} yields $\left(0, \begin{pmatrix} \partial_z \mathbf{b}_h - \nabla_h b_3 \\ 0 \end{pmatrix}, \begin{pmatrix} \partial_z \mathbf{u}_h \\ -\nabla_h \cdot \mathbf{u}_h \end{pmatrix}\right)$. Hence the last part of $\mathcal{L}_A \mathbf{V}$ vanishes iff $\partial_z \mathbf{u}_h = \mathbf{0}$ and $\nabla_h \cdot \mathbf{u}_h = 0$. Furthermore, solving the divergence-free condition $\nabla \cdot \mathbf{b} = 0$ on the magnetic field for its horizontal component to obtain $\nabla_h \cdot \mathbf{b}_h = -\partial_z b_3$ and substituting that formula into the horizontal divergence of the component $\partial_z \mathbf{b}_h - \nabla_h b_3$ of $\mathcal{L}_A \mathbf{V}$ yields

$$\nabla_h \cdot (\partial_z \mathbf{b}_h - \nabla_h b_3) = -\Delta b_3. \quad (3.12)$$

Hence if $\mathcal{L}_A \mathbf{V} = 0$ then b_3 is a constant, i.e., $b_3 = \alpha_v b_3$. That implies further that $\partial_z \mathbf{b}_h = \nabla_h b_3 = \mathbf{0}$. On the other hand, when those conditions hold then each term in the second part of $\mathcal{L}_A \mathbf{V}$ vanishes.

Similarly,

$$\left(\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M\right) \mathbf{V} = \begin{pmatrix} -\frac{1}{\varepsilon_M} \nabla \cdot \mathbf{u} \\ \left(\frac{1}{\varepsilon_A} (-\nabla_h b_3 + \partial_z \mathbf{b}_h - \frac{\varepsilon_A}{\varepsilon_M} \nabla_h r)\right) \\ -\frac{1}{\varepsilon_M} \partial_z r \\ \frac{1}{\varepsilon_A} \begin{pmatrix} \partial_z \mathbf{u}_h \\ -\nabla_h \cdot \mathbf{u}_h \end{pmatrix} \end{pmatrix}. \quad (3.13)$$

The last component of the second part vanishes iff $\partial_z r = 0$ and, since $\partial_z u_3 = \nabla \cdot \mathbf{u} - \nabla_h \cdot \mathbf{u}_h$, the first and last parts vanish iff $\nabla_h \cdot \mathbf{u}_h = 0$ and $\partial_z \mathbf{u} = \mathbf{0}$. Next, use (3.12) to write the horizontal divergence of the second part as

$$\nabla_h \cdot \left(-\nabla_h b_3 + \partial_z \mathbf{b}_h - \frac{\varepsilon_A}{\varepsilon_M} \nabla_h r\right) = -(\Delta b_3 + \frac{\varepsilon_A}{\varepsilon_M} \Delta_h r), \quad (3.14)$$

and since it has been shown that $\partial_z r = 0$, (3.14) vanishes iff $\Delta(b_3 + \frac{\varepsilon_A}{\varepsilon_M} r) = 0$, which is equivalent to the last equation of (3.4). That equation together with the vanishing of $\partial_z r$ implies that $\partial_z b_3 = 0$. Finally, the last equation of (3.4) also shows that the horizontal components of the second part of (3.13) vanish iff $\partial_z \mathbf{b}_h = \mathbf{0}$.

The formula for \mathbb{P}^0 follows directly from the conditions for $\mathcal{L}_A \mathbf{V}$ to vanish in the first part of the lemma, and that formula together with the formula for \mathcal{L}_M that can be seen from the $O(\varepsilon_M^{-1})$ terms in (3.13) implies that

$$\mathbb{P}^0 \mathcal{L}_M \mathbb{P}^0 \mathbf{V} = - \left(\partial_z u_3, \begin{pmatrix} 0 \\ \partial_z r \end{pmatrix}, 0 \right). \quad (3.15)$$

In turn, formula (3.15) implies the formula for \mathbb{P}^1 , and the formulas for the \mathbb{P}^j imply (3.6).

The conditions (3.4) for $(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}$ to vanish differ from the conditions to belong to the null space of \mathbb{P} in (3.6) only by adding an $O(\mu)$ term to the formula for b_3 . Hence the rank of the restriction to any Fourier mode of $N(\mathcal{L}_A + \mu \mathcal{L}_M)$ equals the rank of the restriction to that mode of \mathbb{P} , which in turn equals the dimension of the direct sum of all eigenspaces of $\mathcal{L}_A + \mu \mathcal{L}_M$ in that Fourier mode having eigenvalues of size $o(\mu)$ [Kat82], [CJS18, §4]. Hence all such eigenvalues vanish identically.

Alternatively, the eigenvalues and eigenvectors of the restriction of $\mathcal{L}_A + \mu \mathcal{L}_M$ to any Fourier mode (k, l, m) with $m = 0$ and $(k, l) \neq (0, 0)$ can be easily calculated, showing that the zero eigenvalue has multiplicity five and the remaining eigenvalues are $\pm i \sqrt{1 + \mu^2 \sqrt{k^2 + l^2}}$. For Fourier modes with $m \neq 0$, $(0, 0_3, (k, l, m))$ is an eigenvector of $\mathcal{L}_A + \mu \mathcal{L}_M$ having eigenvalue zero, while the formulas (1.8) for \mathcal{L}_A and (3.15) for $\mathbb{P}^0 \mathcal{L}_M \mathbb{P}^0$ imply their restriction to the Fourier mode have rank four and two respectively, so perturbation theory [Kat82] shows that the sum of the dimensions of the eigenspaces of $\mathcal{L}_A + \mu \mathcal{L}_M$ in such Fourier modes of sizes strictly $O(1)$ and strictly $O(\mu)$ remain four and two, respectively, for $|\mu|$ sufficiently small. In each case the sum of the dimensions of the eigenspaces corresponding to eigenvalues of size strictly $O(1)$, strictly $O(\mu)$, and identically zero add up to the full dimension seven of the restriction of $\mathcal{L}_A + \mu \mathcal{L}_M$ to a Fourier mode, so there can be no nonzero eigenvalues of size $o(\mu)$.

The formulas (3.7) for the modes follow from their definition (3.2) and the formulas (3.5)–(3.6) for the projections \mathbb{P}^j and \mathbb{P} . Since $\alpha_z \mathbf{b}$ is independent of z , $\nabla \cdot \mathbf{b}^S = \nabla \cdot (\alpha_z \mathbf{b}_h, av b_3) = \nabla \cdot (\alpha_z \mathbf{b}) = \alpha_z \nabla \cdot \mathbf{b} = 0$, and trivially $\nabla \cdot \mathbf{b}^I = 0$, which implies the rest of (3.8). Each component of \mathbf{V}^S contains α_z or av , so (3.9) holds. The left equation in (3.10) follows from the presence of the operator $\mathcal{P}_h^{\text{div}}$ in the formula for u_h^S , while the right equation there follows from (3.8)–(3.9).

The formula for u_h^F in (3.7) and the formula (3.13) for $(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}$ yield

$$\|\partial_z u_h^F\|_j = \|\partial_z u_h\|_j \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j, \quad (3.16)$$

$$\|\nabla_h \cdot u_h^F\|_j = \|\nabla_h \cdot u_h\|_j \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j. \quad (3.17)$$

Next, by the ellipticity of Δ together with the formula for b_3^F in (3.7), the identity (3.14) and formula (3.13),

$$\begin{aligned} \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1} &= \|\Delta^{-1} [\Delta(1 - av) b_3 + \mu \Delta_h r]\|_{j+1} \leq c \|\Delta b_3 + \mu \Delta_h r\|_{j-1} \\ &= c \|\nabla_h \cdot (\nabla_h b_3 + \mu \nabla_h r - \partial_z \mathbf{b}_h)\|_{j-1} \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j \end{aligned} \quad (3.18)$$

Also, combining the formula (3.7) for b_h^F , formula (3.13) for $\mathcal{L}_A + \mu \mathcal{L}_M$, and (3.18) yields

$$\begin{aligned} \|\partial_z (b_h^F - \mu \Delta^{-1} \nabla_h \partial_z r^I)\|_j &= \|\partial_z (b_h - \mu \Delta^{-1} \nabla_h \partial_z r)\|_j \\ &= \|\partial_z b_h - \mu \nabla_h \Delta^{-1} (\Delta r - \Delta_h r)\|_j \\ &= \|(\partial_z b_h - \mu \nabla_h r - \nabla_h b_3) + \nabla_h (b_3 + \mu \Delta^{-1} \Delta_h r)\|_j \\ &\leq \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j + \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1} \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j. \end{aligned} \quad (3.19)$$

By the formula (3.7) defining the modes and formula (3.13) for $\mathcal{L}_A + \mu \mathcal{L}_M$,

$$\mu \|\partial_z r^I\|_j = \mu \|\partial_z r\|_j \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j, \quad (3.20)$$

$$\mu \|\partial_z u_3^I\|_j = \mu \|\partial_z u_3\|_j = \mu \|\nabla \cdot \mathbf{u} - \nabla_h \cdot u_h\|_j \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}\|_j \quad (3.21)$$

Combining (3.16), (3.17), (3.18), (3.19), (3.20), and (3.21) yields the right inequality in (3.11), and since the expressions estimated there can be combined to yield all the terms in $(\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V}$ the left inequality also holds. \square \square

Since it is possible to obtain a bound for $(\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V}$ in H^{n-1} by solving the PDE for that expression and using the bounds (1.12) to estimate the result, the estimate (3.11) can be used to obtain “static” estimates for the fast and intermediate modes, i.e., estimates for those modes at any given time in terms of just \mathbf{V} and $\partial_t\mathbf{V}$ at that time. Those estimates will be obtained in §5. However, we will also need to obtain “dynamic” estimates for the intermediate and slow modes via differential inequalities. For that purpose it does not suffice to use the modes defined above because they are not direct sums of eigenspaces for the operator $\mathcal{L}_A + \mu\mathcal{L}_M$ but only limits as $\mu \rightarrow 0$ of such sums. The exact slow eigenspace of the zero eigenvalue was determined in Lemma 3.1. In the next lemma we obtain formulas for the 7-vectors of Fourier multiplier operators \mathbf{V}_α and \mathbf{V}_β and the self-adjoint Fourier multiplier operator \mathcal{Q} , such that

$$(\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V}_\alpha \cdot \tilde{\mathbf{V}} + \mu\mathcal{Q}\partial_z\mathbf{V}_\beta \cdot \tilde{\mathbf{V}} = 0 = (\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V}_\beta \cdot \tilde{\mathbf{V}} + \mu\mathcal{Q}\partial_z\mathbf{V}_\alpha \cdot \tilde{\mathbf{V}} \quad (3.22)$$

for every vector-valued function $\tilde{\mathbf{V}}$. Equations (3.22) imply that for every Fourier mode (k, l, m) the linear combinations $(\mathbf{V}_\alpha \pm \mathbf{V}_\beta)e^{i(kx+ly+mz)}$ are eigenfunctions of the operator $\mathcal{L}_A + \mu\mathcal{L}_M$ with purely imaginary or zero eigenvalues $\mp im\mu\widehat{\mathcal{Q}}$ where

$$\widehat{\mathcal{Q}} := e^{-i(kx+ly+mz)} \left[\mathcal{Q}e^{i(kx+ly+mz)} \right], \quad (3.23)$$

i.e., they yield the exact μ -dependent intermediate eigenspaces. In other words, in the equations for $\partial_t((\mathbf{V}_\alpha \pm \mathbf{V}_\beta) \cdot \mathbf{V})$ the largest terms will be, by the MHD system itself, the skew-adjointness of $\mathcal{L}_A, \mathcal{L}_M$ and (3.23),

$$\begin{aligned} \frac{1}{\varepsilon_A}(\mathbf{V}_\alpha \pm \mathbf{V}_\beta) \cdot ((\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V}) &= -\frac{1}{\varepsilon_A}((\mathcal{L}_A + \mu\mathcal{L}_M)(\mathbf{V}_\alpha \pm \mathbf{V}_\beta)) \cdot \mathbf{V} \\ &= \frac{1}{\varepsilon_M}\mathcal{Q}\partial_z(\mathbf{V}_\beta \pm \mathbf{V}_\alpha) \cdot \mathbf{V}. \end{aligned} \quad (3.24)$$

However, in §5 we will obtain dynamic estimates for $(1 - \alpha_z)\mathbf{V}_\alpha \cdot \mathbf{V}$ and $(1 - \alpha_z)\mathbf{V}_\beta \cdot \mathbf{V}$ rather than $(\mathbf{V}_\alpha \pm \mathbf{V}_\beta) \cdot \mathbf{V}$ to reduce disruption to the structure of the rest of the PDE. The operator $1 - \alpha_z$ is applied because the eigenvalues $\pm im\mu\mathcal{Q}$ are only of size μ when $m \neq 0$. Also since we only consider initial data for which $\nabla \cdot \mathbf{b} = 0$ we can and will use the variant $\mathbf{V}_\alpha^{\text{no-div}}$ defined in (3.25) that omits the term in \mathbf{V}_α containing a gradient in the magnetic field component.

Since the limits as $\mu \rightarrow 0$ of \mathbf{V}_α and \mathbf{V}_β should belong to the intermediate mode defined in (3.7), trying various perturbations leads to the ansatz

$$\begin{aligned} \mathbf{V}_\alpha &:= \mathbf{V}_\alpha^{\text{div}} + \mathbf{V}_\alpha^{\text{no-div}} = \begin{pmatrix} 0 \\ 0_3 \\ \mu\mathcal{A}\Delta^{-1}\partial_z\nabla \end{pmatrix} + \begin{pmatrix} 1 \\ 0_3 \\ -\mu\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \mu^3\begin{pmatrix} 0 \\ 0 \\ \mathcal{B} \end{pmatrix} \end{pmatrix}, \\ \mathbf{V}_\beta &:= \begin{pmatrix} 0 \\ \left(\mu^2\mathcal{C}\Delta^{-1}\partial_z\nabla_h \right) \\ 0_3 \end{pmatrix}, \end{aligned} \quad (3.25)$$

where the vectors have been normalized by setting the first component of \mathbf{V}_α to the identity operator 1, and factors of Δ^{-1} have been included so that if $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ are all homogeneous order zero Fourier multipliers then all components of \mathbf{V}_α and \mathbf{V}_β will also be Fourier multipliers of order zero. Furthermore, in the limit as $\mu \rightarrow 0$ the operator $-\partial_z\mathcal{Q}$ should tend to the operator $-\partial_z$ appearing in $\mathbb{P}^0\mathcal{L}_M\mathbb{P}^0$ in (3.15), i.e., \mathcal{Q} should tend to one.

Lemma 3.2. *The vectors \mathbf{V}_α and \mathbf{V}_β defined in (3.25) satisfy (3.22) with*

$$\begin{aligned} \mathcal{Q} &= \sqrt{2 \left(1 + \mu^2 + \sqrt{(1 + \mu^2)^2 - 4\mu^2 \partial_z^2 \Delta^{-1}} \right)^{-1}}, \\ \text{i.e., } \widehat{\mathcal{Q}} &= \frac{\sqrt{2}}{\sqrt{1 + \mu^2 + \sqrt{(1 + \mu^2)^2 - 4\mu^2 \frac{m^2}{k^2 + l^2 + m^2}}}}, \end{aligned} \quad (3.26)$$

provided that

$$\mathcal{C} = -\mathcal{Q} (1 - \mu^2 \partial_z^2 \Delta^{-1} \mathcal{Q}^2)^{-1}, \quad \mathcal{A} = -\mathcal{C} \mathcal{Q}^{-1}, \quad \mathcal{B} = \partial_z^2 \Delta^{-1} \mathcal{C} \mathcal{Q}, \quad \mathcal{D} = \mathcal{Q}^{-1}. \quad (3.27)$$

Proof. Substituting (3.25) into (3.22) yields

$$\begin{aligned} &\begin{pmatrix} 0 \\ (-\mu^3 (\mathcal{B} - \mathcal{C} \mathcal{Q} \partial_z^2 \Delta^{-1}) \nabla_h) \\ -\mu \partial_z (1 - \mathcal{Q} \mathcal{Q}) \\ 0_3 \end{pmatrix} = \mathbf{0}, \\ &\begin{pmatrix} -\mu \partial_z (\mathcal{D} - \mathcal{Q} + \mu^2 \mathcal{C} \Delta_h \Delta^{-1}) \\ 0_3 \\ \mu^2 \partial_z^2 \Delta^{-1} (\mathcal{C} + \mathcal{A} \mathcal{Q}) \nabla - \begin{pmatrix} 0 \\ \mu^2 \partial_z (\mathcal{C} + \mathcal{Q} - \mu^2 \mathcal{B} \mathcal{Q}) \end{pmatrix} \end{pmatrix} = \mathbf{0}, \end{aligned} \quad (3.28)$$

which will hold provided that

$$\begin{aligned} \mathcal{C} + \mathcal{A} \mathcal{Q} &= 0, \quad \mathcal{D} \mathcal{Q} = 1, \quad \mathcal{B} = \mathcal{C} \mathcal{Q} \partial_z^2 \Delta^{-1}, \quad \mathcal{C} + \mathcal{Q} = \mu^2 \mathcal{B} \mathcal{Q}, \\ \mathcal{D} - \mathcal{Q} + \mu^2 \Delta_h \Delta^{-1} \mathcal{C} &= 0. \end{aligned} \quad (3.29)$$

Solving the first three equations in (3.29) for \mathcal{A} , \mathcal{D} , and \mathcal{B} yields the formulas for those operators claimed in (3.27). Substituting those formulas into the fourth equation in (3.29) and solving the result for \mathcal{C} yields the formula for that operator in (3.27). Substituting the formulas obtained so far into the last equation in (3.29) yields $\mathcal{Q}^{-1} - \mathcal{Q} - \mu^2 \Delta_h \Delta^{-1} \mathcal{Q} (1 - \mu^2 \partial_z^2 \Delta^{-1} \mathcal{Q}^2)^{-1} = 0$, whose only solution tending to one as $\mu \rightarrow 0$ is (3.26). \square \square

We now turn to explicating the form that initial data satisfying (1.11) takes.

Lemma 3.3. *Initial data $\mathbf{V}^0 = (r^0, \mathbf{u}^0, \mathbf{b}^0)$ will be uniformly bounded in H^n and satisfy the constraint $\nabla \cdot \mathbf{b}^0 = 0$ and the condition (1.11) iff it has the form*

$$\begin{aligned} r^0 &= r^{0,S} + \varepsilon_M r^{0,I}, \quad \mathbf{u}_h^0 = \mathbf{u}_h^{0,S} + \varepsilon_A \mathbf{u}_h^{0,F}, \quad u_3^0 = u_3^{0,S} + \varepsilon_M u_3^{0,I}, \\ \mathbf{b}_h^0 &= \mathbf{b}_h^{0,S} + \varepsilon_A \mathbf{b}_h^{0,F}, \quad b_3^0 = \text{av} b_3^0 - \mu (1 - \text{av}) r^{0,S} + \varepsilon_A b_3^{0,F}, \end{aligned} \quad (3.30)$$

where every term $w^{0,\ell}$ with $w \in \{r, \mathbf{u}_h, u_3, \mathbf{b}_h, b_3\}$ and $\ell \in \{F, I, S\}$ has the form specified for the w component of the ℓ mode in (3.7), and may depend on $(\varepsilon_A, \varepsilon_M)$ but satisfies

$$\begin{aligned} &\|r^{0,S}\|_n + [\|\partial_z r^{0,I}\|_{n-1} + \varepsilon_M \|r^{0,I}\|_n] \\ &\quad + \|\mathbf{u}_h^{0,S}\|_n + [\|\partial_z \mathbf{u}_h^{0,F}\|_{n-1} + \|\nabla_h \cdot \mathbf{u}_h^{0,F}\|_n + \varepsilon_A \|\mathbf{u}_h^{0,F}\|_n] \\ &\quad + \|u_3^{0,S}\|_n + [\|\partial_z u_3^{0,I}\|_{n-1} + \varepsilon_M \|u_3^{0,I}\|_n] \\ &\quad + \|\mathbf{b}_h^{0,S}\|_n + [\|\partial_z \mathbf{b}_h^{0,F}\|_{n-1} + \varepsilon_A \|\mathbf{b}_h^{0,F}\|_n] \\ &\quad + |\text{av} b_3^0| + \|b_3^{0,F} + \Delta^{-1} \Delta_h r^{0,I}\|_n \leq c \end{aligned} \quad (3.31)$$

uniformly in those parameters, and

$$\nabla_h \cdot \mathbf{b}_h^{0,S} = 0 = \nabla_h \cdot \mathbf{b}_h^{0,F} + \partial_z b_3^{0,F}. \quad (3.32)$$

Proof. Since the terms appearing in (3.30) are allowed to depend on ε_A and ε_M , that formula simply expresses the separation of the initial data into fast, intermediate, and slow modes, with the factors of ε_A and ε_M being purely for later convenience, as is the inclusion of the specific term $-\mu(1 - \text{av})r^{0,S}$ as part of the fast part of b_3^0 . By Lemma 3.1, the condition $\nabla \cdot \mathbf{b} = 0$ implies that each mode is divergence-free, and the conditions $\nabla \cdot \mathbf{b}^\ell = 0$ for $\ell \in \{F, I, S\}$ clearly imply that $\nabla \cdot \mathbf{b} = 0$, so for initial data of the form (3.30) the conditions (3.32) are equivalent to the assumed condition that $\nabla \cdot \mathbf{b}^0 = 0$.

Since, as shown above, the square of the H^n norm of \mathbf{V}^0 equals the sum of the squares of the H^n norms of its modes, the assumed uniform boundedness of $\|\mathbf{V}^0\|_n$ is equivalent to

$$\begin{aligned} & \|r^{0,S}\|_n + \varepsilon_M \|r^{0,I}\|_n + \|u_h^{0,S}\|_n + \varepsilon_A \|u_h^{0,F}\|_n + \|u_3^{0,S}\|_n + \varepsilon_M \|u_3^{0,I}\|_n \\ & + \|b_h^{0,S}\|_n + \varepsilon_A \|b_h^{0,F}\|_n + |\text{av} b_3^0| + \|-\mu(1 - \text{av})r^{0,S} + \varepsilon_A b_3^{0,F}\|_n \leq c. \end{aligned} \quad (3.33)$$

When k is set equal to $n - 1$ the sum of terms estimated in (3.11) is equivalent to the H^{n-1} norm of $(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}$. Hence the assumed uniform boundedness of $\frac{1}{\varepsilon_A} \|(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}^0\|_{n-1}$ becomes, for initial data \mathbf{V}^0 of the form (3.30) satisfying (3.32),

$$\begin{aligned} & \|\partial_z u_h^{0,F}\|_{n-1} + \|\nabla_h \cdot u_h^{0,F}\|_n + \|\partial_z (b_h^{0,F} - \Delta^{-1} \nabla_h \partial_z r^{0,I})\|_{n-1} \\ & + \|\partial_z r^{0,I}\|_{n-1} + \|\partial_z u_3^{0,I}\|_{n-1} + \|-\varepsilon_M^{-1} (1 - \text{av})r^{0,S} + b_3^{0,F}\|_n \\ & + \varepsilon_M^{-1} \Delta^{-1} \Delta_h r^{0,S} + \Delta^{-1} \Delta_h r^{0,I}\|_n \leq c. \end{aligned} \quad (3.34)$$

Since the expression $-\mu(1 - \text{av})r^{0,S}$ appearing in the last term in (3.33) can be estimated by the first term there it can be omitted from that last term, leaving $\varepsilon_A \|b_3^{0,F}\|_n$. Similarly, the expression $-\partial_z (\Delta^{-1} \Delta_h) \partial_z r^{0,I}$ in the third term of (3.34) can be estimated by the fourth term there, and so can be omitted from the third term. Also the expression $-\varepsilon_M^{-1} (1 - \text{av})r^{0,S} + \varepsilon_M^{-1} \Delta^{-1} \Delta_h r^{0,S}$ in the last term of (3.34) vanishes identically because $r^{0,S}$ is independent of z , leaving $\|b_3^{0,F} + \Delta^{-1} \Delta_h r^{0,I}\|_n$. The uniform bounds for that term and for $\varepsilon_M \|r^{0,I}\|_n$ from (3.33) imply the uniform boundedness of $\varepsilon_M \|b_3^{0,F}\|_n$, so the modified term $\varepsilon_A \|b_3^{0,F}\|_n$ for (3.33) can be omitted. However, since (3.34) only contains an $O(1)$ estimate for $\partial_z r^{0,I}$ not $r^{0,I}$ itself, it is not possible to omit the expression $\Delta^{-1} \Delta_h r^{0,I}$ from the term $\|b_3^{0,F} + \Delta^{-1} \Delta_h r^{0,I}\|_n$. Adding (3.33) and (3.34) and making the above-mentioned modifications shows that (3.31) is equivalent to the uniform boundedness of $\|\mathbf{V}^0\|_n + \frac{1}{\varepsilon_A} \|(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}^0\|_{n-1}$ for initial data of the form (3.30). \square \square

We note that, by a slight notational exception, the fast part of b_3^0 in (3.30) satisfies

$$(b_3^0)^F = -\mu(1 - \text{av})r^{0,S} + \varepsilon_A b_3^{0,F}, \quad (3.35)$$

showing that, to the leading $O(\mu)$ order, $(b_3^0)^F$ only depends on the slow part of r^0 .

Although Lemma 3.3 determines the most general initial data satisfying the conditions needed for Theorem 1.1, as noted in the introduction we need to take more specific data in order to obtain a rate of convergence. In order to allow the initial data to contain all modes but not interfere with the convergence rate, we will still assume that the initial data have the form (3.30), and that (3.32) holds, but will assume that

$$\text{all terms } w^{0,\ell} \text{ in (3.30) are uniformly bounded in } H^n \text{ and each term } w^{0,S} \text{ is fixed,} \quad (3.36)$$

which automatically implies that (3.31) holds. We allow the $w^{0,F}, w^{0,I}$ to depend on $(\varepsilon_A, \varepsilon_M)$ because we only bound the distance of the fast and intermediate modes to zero (i.e., estimate their Sobolev norms), not their distance to any non-trivial limits.

4. UNIFORM EXISTENCE AND LIMIT

Remark 4.1. When $\mu_{\text{lim}} = 0$ then the convergence part of Theorem 1.1 follows from [CJS18, Theorem 4.6]. To see this note that the fact shown in Lemma 3.1 that all $o(\mu)$ eigenvalues of $\mathcal{L}_A + \mu\mathcal{L}_M$ are identically zero implies that for the MHD system (1.4) the projection \mathbb{P} of [CJS18, Theorem 4.6] equals the projection \mathbb{P} defined here, and the limit term T_{lim} in the limit equation there vanishes, without making any assumptions on the relation between ε_A and ε_M as both tend to zero. Together with the uniform bounds on the solution shown in the proof below, that implies that the solution of the MHD system (1.4) with the specified initial data converge as $(\varepsilon_A, \varepsilon_M)$ tend to zero while satisfying (1.10) and $\frac{\varepsilon_A}{\varepsilon_M} \rightarrow 0$ to the unique solution $\bar{\mathbf{V}}(t, x)$ of the limit equations [CJS18, (4.14)–(4.15)] having the limit initial data. When the spatial domain is \mathbb{R}^3 then the fact that the limit lies in the range of \mathbb{P} implies by Lemma 3.1 that it is independent of z and hence vanishes identically. When the spatial domain is \mathbb{T}^3 a calculation shows that for the MHD equations and projection \mathbb{P} from (3.6) those limit equations reduce to (1.13) with μ_{lim} set to zero.

of Theorem 1.1. As noted in the introduction, the condition $\nabla \cdot \mathbf{b} = 0$ included in (1.4) will be satisfied automatically provided that it is satisfied at time zero, so in view of the assumption that the initial data satisfy that condition it may be omitted from (1.4). The remaining part of (1.4) has the form (1.1) to which Theorem 2.2 and [CJS18, Theorem 4.6] apply, and the form of the MHD equations (1.4) together with the assumptions of Theorem 1.1 ensure that Assumption 2.1 holds. In particular, the boundedness of r in H^n will imply its boundedness in L^∞ , so the coefficients of the time-derivative terms in (1.4) will be positive when ε_M is sufficiently small. Hence by Theorem 2.2 the solution of (1.4) having the initial data \mathbf{V}^0 exists and is bounded in H^n for a fixed time T , and (1.12) holds.

The uniform bounds on \mathbf{V} and \mathbf{V}_t show that for any sequence of values $(\varepsilon_A)_j$ and $(\varepsilon_M)_j$ tending to zero, a subsequence of the corresponding solutions converges in $C^0([0, T]; H^{n-\alpha})$ for any $\alpha > 0$, and after extracting a further subsequence the solutions also converge weak-* in $L^\infty([0, T]; H^n)$ and their time derivatives converge weak-* in $L^\infty([0, T]; H^{n-1})$. Since a standard L^2 energy estimate shows that solutions of the limiting equations (1.13) with initial data in H^n are unique, convergence to the same limit $\bar{\mathbf{V}}$ holds for all ε_A and ε_M converging to zero with their ratio μ converging to the same limiting value μ_{lim} .

Multiplying (1.1) by ε_A or applying \mathbb{P}^0 to it and multiplying the result by ε_M yields

$$\begin{pmatrix} (\mathcal{L}_A + \mu\mathcal{L}_M)\mathbf{V} \\ \mathbb{P}^0 \mathcal{L}_M \end{pmatrix} = \begin{pmatrix} \varepsilon_A \\ \varepsilon_M \mathbb{P}^0 \end{pmatrix} [A_0(\varepsilon_M \mathbf{V}) \mathbf{V}_t + \sum_{i=1}^d A_i(\mathbf{V}) \mathbf{V}_{x_i}], \quad (4.1)$$

and since the uniform bounds (1.12) show that the expression in brackets on the right side of (4.1) is bounded in L^2 , taking the limit yields

$$(\mathcal{L}_A + \mu_{\text{lim}}\mathcal{L}_M)\bar{\mathbf{V}} = 0 = \mathbb{P}^0 \mathcal{L}_M \bar{\mathbf{V}}. \quad (4.2)$$

When $\mu_{\text{lim}} = 0$ then the left equation in (4.2) implies that $\bar{\mathbf{V}} = \mathbb{P}^0 \bar{\mathbf{V}}$, and hence the right equation there implies that $\bar{\mathbf{V}} = \mathbb{P} \bar{\mathbf{V}}$. Since the vanishing of $\nabla \cdot \mathbf{b}^0$ implies that \mathbf{b} remains divergence free, Lemma 3.1 then shows that $\bar{\mathbf{V}}$ is independent of z , the horizontal parts of its velocity and magnetic field are divergence free, and (1.13e) holds. When $\mu_{\text{lim}} > 0$ then Lemma 3.1 shows that the same conclusion follows from just the left equation of (4.2). When the spatial domain is \mathbb{R}^3 then the fact that the limit is independent of z implies that it vanishes, so from now on we consider the case when that domain is \mathbb{T}^3 .

We next note from (3.6) and (3.13) that multiplying the large operator $\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M$ by \mathbb{P} on the left yields $\mathbb{P} \left(\frac{1}{\varepsilon_A} \mathcal{L}_A + \frac{1}{\varepsilon_M} \mathcal{L}_M \right) = \left(\frac{1}{\varepsilon_M} \mathbf{a}_z \nabla_h \cdot \mathbf{u}_h, 0_3, 0_3 \right)$. Hence the limits of the velocity equation and the horizontal magnetic field equation can be determined by taking the limit of \mathbb{P} applied to (1.4), since the resulting equations contain no large terms and each $O(1)$ term not containing a time derivative converges strongly, and each term containing a time derivative converges weakly, to the term obtained by replacing each variable with its limit.

Moreover, since each equation includes only one term with a time derivative, the equations together with the strong convergence of the terms without time derivatives show that the terms containing time derivatives also converge strongly. This yields (1.13b), (1.13c), and (1.13d); the term in (1.13c) containing a factor of μ_{lim} arises from the limit of the term $\mathbf{a}_z(\mathbf{b}_h \cdot \nabla_h b_3)$ on account of (1.13e).

Finally, in order to determine the limit equation for the time evolution of the density we will add to the equation (1.4a) for r an appropriate multiple of the last component of the equation (1.4c) so as to eliminate the large term that remains in \mathbb{P} applied to (1.4). However, it will be convenient to divide (1.4a) by $a(\varepsilon_M r)$ to recover the conservation form

$$r_t + \nabla \cdot (r \mathbf{u}) + \frac{1}{\varepsilon_M} \nabla \cdot \mathbf{u} = 0, \quad (4.3)$$

which is advantageous since the operator \mathbf{a}_z from the definition of the slow modes annihilates z derivatives. Applying \mathbf{a}_z to (4.3) yields

$$\partial_t(\mathbf{a}_z r) + \mathbf{a}_z[\nabla_h \cdot (r \mathbf{u}_h)] + \frac{1}{\varepsilon_M} \mathbf{a}_z(\nabla_h \cdot \mathbf{u}_h) = 0. \quad (4.4)$$

The third component of (1.4c) can be written in conservation form as

$$\partial_t b_3 + \nabla \cdot (b_3 \mathbf{u}) - \nabla \cdot (u_3 \mathbf{b}) + \varepsilon_A^{-1} \nabla_h \cdot \mathbf{u}_h = 0 \quad (4.5)$$

in view of the fact that \mathbf{b} is divergence free. Applying \mathbf{a}_z to (4.5), then multiplying by μ , and subtracting the resulting equation from (4.4) yields

$$\partial_t[\mathbf{a}_z(r - \mu b_3)] + \mathbf{a}_z\{\nabla_h \cdot [(r - \mu b_3) \mathbf{u}_h]\} + \mu \mathbf{a}_z[\nabla_h \cdot (u_3 \mathbf{b}_h)] = 0, \quad (4.6)$$

which contains no large terms. Hence we can take the limit of (4.6) to obtain

$$\partial_t[\bar{r} - \mu_{\text{lim}} \bar{b}_3] + (\bar{\mathbf{u}}_h \cdot \nabla_h)(\bar{r} - \mu_{\text{lim}} \bar{b}_3) + \mu_{\text{lim}}(\bar{\mathbf{b}}_h \cdot \nabla_h) \bar{u}_3 = 0, \quad (4.7)$$

where we have used the facts that the limit variables are independent of z and that $\bar{\mathbf{u}}_h$ and $\bar{\mathbf{b}}_h$ are divergence free. Substituting (1.13e) into (4.7) yields (1.13a). \square \square

5. CONVERGENCE RATE ESTIMATES

The estimates will be obtained by estimating separately the size of the fast modes, the size of the intermediate modes, and the difference between the slow modes and the solution of the limit system, with earlier estimates used when deriving later ones. The estimates for the fast modes will be purely static, i.e., will be obtained simply by solving for certain terms in the PDE system (1.4) and estimating their norms. In contrast, the estimates for difference between the slow modes and the limit solution will be obtained dynamically, i.e., by obtaining energy estimates for the time evolution of norms of that difference. Estimates for the intermediate modes will first be obtained via static estimates and will then be improved via dynamic estimates that use those static estimates.

Recall that $\mu = \frac{\varepsilon_A}{\varepsilon_M} = \varepsilon_M^v$ is assumed here to be less than one. In order to obtain sharp dynamic estimates we will make use of the fact that smaller static estimates can be obtained by using weaker norms, on account of the interpolation estimate (2.7). In order to see easily the estimate expected to be obtained as a function of the norm used, we introduce an increasing geometric sequence

$$\varepsilon_j := \varepsilon_M^{1+v-jv} \quad \text{so that by (1.10), (1.14)} \quad \varepsilon_0 = \varepsilon_A, \varepsilon_1 = \varepsilon_M, \varepsilon_j = \mu \varepsilon_{j+1}, \varepsilon_n \leq c. \quad (5.1)$$

5.1. Static estimates. Static estimates will be obtained by solving the PDE system for $(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}$ and using the uniform bounds (1.12) together with the standard Sobolev product and function estimates

$$\|fg\|_j \leq c\|f\|_{n-1}\|g\|_j, \quad j = 0, \dots, n-1 \quad (5.2)$$

$$\|F(g)\|_{n-1} \leq C(\|g\|_{n-1}), \quad (5.3)$$

which all will be used henceforth without mention, plus the interpolation estimate (2.7). We will also need an estimate for the time integral of certain fast terms, which will be obtained similarly from the time integral of the PDE.

Theorem 5.1. *Assume that the conditions of Theorem 1.1 hold. Then the fast component \mathbf{V}^F satisfies the estimates*

$$\begin{aligned} \sup_{t \in [0, T]} [\|\mathbf{u}_h^F\|_j + \|\nabla_h \cdot \mathbf{u}_h^F\|_j + \|\partial_z \mathbf{u}^F\|_j + \|\partial_z \mathbf{b}^F\|_j + \|\mathbf{b}_h^F\|_j + \|(1 - \mathbf{a}_z)b_3^F\|_j \\ + \|\mathbf{a}_z \mathbf{u}_h^F\|_{j+1} + \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1}] \leq C \varepsilon_j, \quad j = 0 \dots n-1, \end{aligned} \quad (5.4)$$

$$\begin{aligned} \sup_{t \in [0, T]} [\|\int_0^t \mathbf{a}_z \mathbf{u}_h^F dt'\|_n + \|\int_0^t \mathbf{u}_h^F dt'\|_{n-1} + \|\int_0^t \nabla_h \cdot \mathbf{u}_h^F dt'\|_{n-1} \\ + \|\int_0^t \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S) dt'\|_n] \leq C(T) \varepsilon_M^{1+\nu}, \end{aligned} \quad (5.5)$$

and the intermediate component $(r^I, u_3^I) = ((1 - \mathbf{a}_z)r, (1 - \mathbf{a}_z)u_3)$ satisfies the **static estimates**

$$\|(r^I, u_3^I)\|_j \leq c\|(\partial_z r^I, \partial_z u_3^I)\|_j \leq c \varepsilon_{j+1} = \frac{c \varepsilon_j}{\mu}, \quad j = 0 \dots n-1. \quad (5.6)$$

Note that the bounds in (5.5) for the time integrals of fast components are smaller than the bounds in (5.4) for those components themselves, a phenomenon that seems to have first been utilized in the two-scale case in [KLN91] and is of independent interest, although only the estimates for the first and last expressions on the left side of (5.5) will be used below. Also note that while (5.4) is in the fashion of ‘‘interpolative estimates’’, the time-integral estimates (5.5) can not be improved by reducing the Sobolev norm index, to the best of our knowledge.

Proof. Taking the H^k norm of both sides of the first component of (4.1) and using the interpolation bounds (2.7) to estimate the right side shows that

$$\|(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}\|_j \leq c \varepsilon_M^{1+\nu-j\nu} = c \varepsilon_j. \quad (5.7)$$

Combining (5.7), the estimate (3.11) for intermediate components of the solution in terms of $(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}$ and the Poincaré inequality

$$\|(1 - \mathbf{a}_z)f\|_j \leq c\|\partial_z f\|_j, \quad (5.8)$$

yields (5.6). By (3.11), (5.7), and the fact that $u_3^F \equiv 0$,

$$\|\partial_z \mathbf{u}^F\|_j + \|\nabla_h \cdot \mathbf{u}_h^F\|_j + \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1} \leq c \varepsilon_j, \quad j = 0, \dots, n-1. \quad (5.9)$$

Combining the definition (3.7) of the fast modes, the Poincaré inequality (5.8), the second inequality of (3.11), and (5.7) yields

$$\begin{aligned} [\|\partial_z \mathbf{b}^F\|_j + \|\mathbf{b}_h^F\|_j + \|(1 - \mathbf{a}_z)b_3^F\|_j] &\leq c\|\partial_z \mathbf{b}^F\|_j \\ &\leq c[\|\partial_z (\mathbf{b}_h^F - \mu \Delta^{-1} \nabla_h \partial_z r^I)\|_j + \mu\|\partial_z r^I\|_j + \|\partial_z (b_3^F + \mu \Delta^{-1} \Delta_h r)\|_j + \mu\|\partial_z r^I\|_j] \\ &\leq c[\|\partial_z (\mathbf{b}_h^F - \mu \Delta^{-1} \nabla_h \partial_z r^I)\|_j + \mu\|\partial_z r^I\|_j + \|b_3^F + \mu \Delta^{-1} \Delta_h r\|_{j+1} + \mu\|\partial_z r^I\|_j] \\ &\leq c\|(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}\|_j \leq c \varepsilon_j, \quad j = 0, \dots, n-1. \end{aligned} \quad (5.10)$$

By the definition (3.7) of the fast modes

$$\mathbf{u}_h^F = (1 - \alpha_z)\mathbf{u}_h^F + \alpha_z \mathbf{u}_h^F = (1 - \alpha_z)\mathbf{u}_h^F + \nabla_h \Delta_h^{-1} \alpha_z (\nabla_h \cdot \mathbf{u}_h^F). \quad (5.11)$$

Note that Δ_h is elliptic when applied to functions independent of z so

$$\|\nabla_h \Delta_h^{-1} \alpha_z f\|_{j+1} \leq c \|f\|_j. \quad (5.12)$$

By (5.11), (5.8), (5.12), (3.11), and (5.7),

$$\begin{aligned} \|\mathbf{u}_h^F\|_j + \|\alpha_z \mathbf{u}_h^F\|_{j+1} &\leq \|(1 - \alpha_z)\mathbf{u}_h^F\|_j + \|\nabla_h \Delta_h^{-1} \alpha_z (\nabla_h \cdot \mathbf{u}_h^F)\|_{j+1} \\ &\leq c \|\partial_z \mathbf{u}_h^F\|_j + c \|\nabla_h \cdot \mathbf{u}_h^F\|_j \leq c \|(\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V}\|_j \\ &\leq c \varepsilon_j \quad j = 0, \dots, n-1. \end{aligned} \quad (5.13)$$

Combining (5.9), (5.10), and (5.13) yields (5.4).

The bounds on time integrals in (5.5) are obtained by integrating (1.1) with respect to time, which yields

$$\int_0^t (\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V} = \varepsilon_A \left(A_0(\varepsilon_M \mathbf{V})\mathbf{V}|_0^t - \int_0^t (\varepsilon_M \mathbf{V}_t \cdot \nabla_{\mathbf{V}} A_0)\mathbf{V} + \int_0^t \sum_{i=1}^d A_i(\mathbf{V})\mathbf{V}_{x_i} \right). \quad (5.14)$$

For the MHD system (1.4), only the variable r appears in the argument of A_0 , and (1.4a) shows that the time derivative of r is $O(\varepsilon_M^{-1})$, so the term $\varepsilon_M \mathbf{V}_t \cdot \nabla_{\mathbf{V}} A_0$ is uniformly bounded. This yields the estimate

$$\left\| \int_0^t (\mathcal{L}_A + \mu \mathcal{L}_M)\mathbf{V} dt' \right\|_{n-1} \leq c \varepsilon_A = c \varepsilon_M^{1+\nu}. \quad (5.15)$$

Since spatial operators commute with time integration, replacing every solution component in the second inequality of (3.11) with its time-integral from 0 to t and combining the result with the bound (5.15) yields (5.5) since $\alpha_z r = r^S$. \square

5.2. Intermediate system dynamic estimates.

Theorem 5.2. *Assume that the conditions of Theorem 1.2 hold, and let $(r, \mathbf{u}, \mathbf{b})$ be the solution obtained for the MHD system (1.4). Then*

$$\|r^J\|_j + \|u_3^J\|_j \leq c \varepsilon_{\max(j,1)}, \quad j = 0, \dots, n-1 \quad (5.16)$$

and

$$\|\partial_z r^J\|_{n-1} + \|\partial_z u_3^J\|_{n-1} \leq c \varepsilon_n. \quad (5.17)$$

The case $j = 0$ in (5.16) and estimate (5.17) were already proven in Theorem 5.1. The remaining cases in (5.16) are an improvement over the corresponding cases in (5.6) by one factor of μ .

As shown in Lemma 3.2 and the discussion preceding it, and in particular formula (3.25), the appropriate expressions to use in the dynamic estimate for the intermediate modes are not the fixed intermediate modes defined in (3.7) but rather

$$\begin{aligned} \alpha &:= (1 - \alpha_z)\mathbf{V}_\alpha^{\text{no-div}} \cdot \mathbf{V} = (1 - \alpha_z)(r - \mu b_3 + \mu^3 \mathcal{C} \mathcal{Q} \Delta^{-1} \partial_z^2 b_3), \\ \beta &:= (1 - \alpha_z)\mathbf{V}_\beta \cdot \mathbf{V} = (1 - \alpha_z)(\mathcal{D} u_3 + \mu^2 \mathcal{C} \Delta^{-1} \partial_z \nabla_h \cdot \mathbf{u}_h), \end{aligned} \quad (5.18)$$

where the operators \mathcal{C} , \mathcal{D} , and \mathcal{Q} were defined in (3.27).

As a preliminary to proving Theorem 5.2 we will derive a system of PDEs satisfied by (α, β) , with remainder terms that are consistent with the desired estimate (5.16). The general idea is to apply each of the operators $\mathbf{V}_\alpha^{\text{no-div}}$ and \mathbf{V}_β to the PDE, note that by Lemma 3.2 the large terms of the result are $\partial_z \mathcal{Q}$ applied to the other operator, and calculate the form of the order one terms. In order to simplify that calculation it will be convenient to simplify the original equations beforehand by moving to the right sides of the equations those terms whose H^j norms can be estimated by a constant times ε_j or the H^j norms of r^J and u_3^J using (5.4), (5.6), and (2.7). To

do so we will make use of the formulas in (3.7) for the fast, intermediate, and slow parts of the solution, and in particular we will use the formula

$$\mathbf{u} \cdot \nabla = (\mathbf{u}^S + \mathbf{u}^I + \mathbf{u}^F) \cdot \nabla = \mathbf{u}^S \cdot \nabla + u_3^I \partial_z + u_h^F \cdot \nabla_h. \quad (5.19)$$

Starting from the MHD equations (1.4), we replace the argument $\varepsilon_M r$ of ρ and a by $\varepsilon_M r^S$ or zero where possible and compensate by adding terms to the right sides of the equations, except that we retain a factor of $a(\varepsilon_M r^S)$ multiplying $\frac{1}{\varepsilon_M} \nabla_h \cdot \mathbf{u}_h$ because that expression will then cancel exactly when we build the time evolution equation for α , (5.23). In addition, we apply $1 - \alpha_z$ to the equations since that operator appears in all terms of the formulas (5.18) for α and β . When doing so we reap the first benefit from the transfer of the coefficients depending on intermediate and fast modes to the right side: the fact that slow modes are independent of z allows us to move the operator $1 - \alpha_z$ past most coefficients so that it applies only to differentiated factors. This yields

$$a(\varepsilon_M r^S) (\partial_t + (\mathbf{u}^S \cdot \nabla)) r^I + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)}{\varepsilon_M} \partial_z u_3^I + \frac{a(\varepsilon_M r^S)}{\varepsilon_M} (1 - \alpha_z) \nabla_h \cdot \mathbf{u}_h = (1 - \alpha_z) R_1, \quad (5.20a)$$

$$\begin{aligned} \rho(\varepsilon_M r^S) (\partial_t + (\mathbf{u}^S \cdot \nabla)) (1 - \alpha_z) \mathbf{u}_h + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S) - 1}{\varepsilon_M} \nabla_h r^I \\ + (1 - \alpha_z) [\nabla_h \frac{|\mathbf{b}|^2}{2} - (\mathbf{b} \cdot \nabla) \mathbf{b}_h] - \frac{(1 - \alpha_z)}{\varepsilon_A} (\partial_z \mathbf{b}_h^F - \nabla_h b_3^F - \mu \nabla_h r^I) = (1 - \alpha_z) R_2, \end{aligned} \quad (5.20b)$$

$$\begin{aligned} \rho(\varepsilon_M r^S) (\partial_t + (\mathbf{u}^S \cdot \nabla)) u_3^I + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)}{\varepsilon_M} \partial_z r^I - (\mathbf{b}_h^S \cdot \nabla_h) (1 - \alpha_z) b_3^F \\ = (1 - \alpha_z) R_3, \end{aligned} \quad (5.20c)$$

$$(\partial_t + (\mathbf{u}^S \cdot \nabla)) (1 - \alpha_z) b_3 - (\mathbf{b}_h^S \cdot \nabla_h) u_3^I + (1 - \alpha_z) \frac{1}{\varepsilon_A} \nabla_h \cdot \mathbf{u}_h = (1 - \alpha_z) R_4, \quad (5.20d)$$

where the equation for the horizontal magnetic field has been omitted since it does not enter into α or β , and

$$\begin{aligned} R_1 := -a(\varepsilon_M r) (u_3^I \partial_z r + (u_h^F \cdot \nabla_h) r) + (a(\varepsilon_M r^S) - a(\varepsilon_M r)) (\partial_t r + (\mathbf{u}^S \cdot \nabla) r) \\ + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S) - a(\varepsilon_M r) \rho(\varepsilon_M r)}{\varepsilon_M} (\nabla_h \cdot \mathbf{u}_h + \partial_z u_3^I) + a(\varepsilon_M r^S) \frac{1 - \rho(\varepsilon_M r^S)}{\varepsilon_M} \nabla_h \cdot \mathbf{u}_h, \end{aligned} \quad (5.21a)$$

$$\begin{aligned} R_2 := -\rho(\varepsilon_M r) (u_3^I \partial_z \mathbf{u}_h + (u_h^F \cdot \nabla_h) \mathbf{u}_h) + (\rho(\varepsilon_M r^S) - \rho(\varepsilon_M r)) (\partial_t \mathbf{u}_h + (\mathbf{u}^S \cdot \nabla) \mathbf{u}_h) \\ + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S) - a(\varepsilon_M r) \rho(\varepsilon_M r)}{\varepsilon_M} \nabla_h r, \end{aligned} \quad (5.21b)$$

$$\begin{aligned} R_3 := -\rho(\varepsilon_M r) (u_3^I \partial_z u_3 + (u_h^F \cdot \nabla_h) u_3) + (\rho(\varepsilon_M r^S) - \rho(\varepsilon_M r)) (\partial_t u_3 + (\mathbf{u}^S \cdot \nabla) u_3) \\ + \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S) - a(\varepsilon_M r) \rho(\varepsilon_M r)}{\varepsilon_M} \partial_z r - \mathbf{b}_h \cdot \partial_z \mathbf{b}_h^F + (\mathbf{b}_h^F \cdot \nabla_h) b_3^F, \end{aligned} \quad (5.21c)$$

$$R_4 = -(u_3^I \partial_z b_3 + (u_h^F \cdot \nabla_h) b_3) - (\nabla_h \cdot \mathbf{u}_h) b_3 + (\mathbf{b}_h^F \cdot \nabla_h) u_3. \quad (5.21d)$$

Wherever u_3^I appears undifferentiated in R_i , the H^j norm of the term in which it appears can be estimated by a constant times the H^j norm of u_3^I , for $j = 0, \dots, n-1$. Similarly, for any smooth function F and $j = 0, \dots, n-1$, $\|\frac{F(\varepsilon_M r^S) - F(\varepsilon_M r)}{\varepsilon_M}\|_j \leq c \|r^I\|_j$, and $\|[F(\varepsilon_M r^S) - F(\varepsilon_M r)](\partial_t \mathbf{V} + (\mathbf{u}^S \cdot \nabla) \mathbf{V})\|_j \leq c \varepsilon_M \|r^I\|_j (\|\mathbf{V}_t\|_{n-1} + \|\mathbf{V}\|_n) \leq c \|r^I\|_j$ in view of the bound for $\|\mathbf{V}_t\|_{n-1}$ in (1.12) and the constraint (1.10). Finally, terms containing u_h^F , \mathbf{b}_h^F , $\nabla_h \cdot u_h^F = \nabla_h \cdot \mathbf{u}_h$, or $\partial_z \mathbf{b}_h^F$ without further derivatives can be estimated in the H^j norm by a constant times ε_j , for $j = 0, \dots, n-1$, in view of (5.4). Since these cases cover all the terms in the R_i ,

$$\sum_{i=1}^4 \|(1 - \alpha_z) R_i\|_j \leq \sum_{i=1}^4 \|R_i\|_j \leq c (\|r^I, u_3^I\|_j + \varepsilon_j), \quad j = 0, \dots, n-1. \quad (5.22)$$

In order to obtain the evolution equation for α we apply each component of $\mathbf{V}_\alpha^{\text{no-div}}$ to the corresponding equation in (5.20), then multiply the equation involving b_3 by $a(\varepsilon_M r^S)$ and add the results. In other words we subtract $\mu a(\varepsilon_M r^S)$ times (5.20d) from (5.20a) and add $a(\varepsilon_M r^S) \mu^3 \mathcal{L} \mathcal{Q} \Delta^{-1} \partial_z^2$ applied to (5.20d) to the result. Rearranging the equation so obtained by commuting $\mathcal{L} \mathcal{Q} \Delta^{-1} \partial_z^2$ past $\mathbf{u}^S \cdot \nabla$, making the coefficient of the large

terms that do not cancel exactly be $a\rho$ everywhere while compensating via terms on the right side, forcing the function to which $\mu(\mathbf{b}_h^S \cdot \nabla_h)$ is applied to be $\frac{\beta}{\sqrt{1+\mu^2}}$ for reasons that will be explained later and again compensating on the right side, and using the identity $1 = \mathcal{D}\mathcal{D}$ from (3.27) that makes the large terms exactly involve β , as we know from (3.22) that they must, yields

$$\begin{aligned} a(\varepsilon_M r^S)(\partial_t + (\mathbf{u}^S \cdot \nabla))\alpha + \frac{\mu}{\sqrt{1+\mu^2}}(\mathbf{b}_h^S \cdot \nabla_h)\beta + \frac{a(\varepsilon_M r^S)\rho(\varepsilon_M r^S)}{\varepsilon_M} \mathcal{D}\partial_z\beta \\ = R_{\alpha,1} + R_{\alpha,2} + R_{\alpha,3} + R_{\alpha,4}, \end{aligned} \quad (5.23)$$

where

$$R_{\alpha,1} := (1 - \alpha_z)R_1 - \mu a(\varepsilon_M r^S)(1 - \alpha_z)R_4 + \mu^3 a(\varepsilon_M r^S) \mathcal{C} \mathcal{D}\Delta^{-1} \partial_z^2 R_4 \quad (5.24)$$

comes from the right sides of the modified equations (5.20),

$$R_{\alpha,2} := -\mu^3 a(\varepsilon_M r^S) [\mathcal{C} \mathcal{D}\Delta^{-1} \partial_z^2, \mathbf{u}^S] \cdot \nabla (1 - \alpha_z) b_3^F \quad (5.25)$$

comes from commuting the operator applied to (5.20d) past the coefficient u_3^S ,

$$\begin{aligned} R_{\alpha,3} := -\mu \varepsilon_M \frac{a(\varepsilon_M r^S) - 1}{\varepsilon_M} (\mathbf{b}_h^S \cdot \nabla_h) u_3^I + \mu^3 a(\varepsilon_M r^S) \mathcal{C} \mathcal{D}\Delta^{-1} \partial_z [(b_h^S \cdot \nabla_h) \partial_z u_3^I] \\ + \mu^3 a(\varepsilon_M r^S) \frac{\rho(\varepsilon_M r^S) - 1}{\varepsilon_A} \mathcal{C} \mathcal{D}\Delta^{-1} \partial_z^2 (\nabla_h \cdot \mathbf{u}_h) \end{aligned} \quad (5.26)$$

comes from adding compensating terms to the right side and moving entire terms there, and

$$R_{\alpha,4} := \mu (\mathbf{b}_h^S \cdot \nabla_h) \left(\frac{\beta}{\sqrt{1+\mu^2}} - u_3^I \right) \quad (5.27)$$

comes from forcing the term involving $\mathbf{b}_h^S \cdot \nabla_h$ to have the desired form, and will be rearranged further later. The form of the large term in both (5.23) and (5.28) below follows from (3.24) or direct calculation using (3.27).

Similarly, to obtain the evolution equation for β we begin by applying \mathbf{V}_β to (5.20), i.e., adding $\mu^2 \mathcal{C} \Delta^{-1} \partial_z \nabla_h \cdot$ applied to (5.20b) to \mathcal{D} applied to (5.20c), and rearranging terms in similar fashion as for (5.23). In addition, we force the function to which $\mathbf{b}_h^S \cdot \nabla_h$ is applied on the left side of the equation to be $\frac{\mu}{\sqrt{1+\mu^2}} \alpha$ for reasons to be explained later, and subtract appropriate constants from the factors appearing inside commutators since that does not affect their value, in order to facilitate estimate the size of the resulting terms. Also, to ensure the symmetry of the resulting system for α and β we multiply the large term $\frac{\mu^2}{\varepsilon_A} (1 - \alpha_z) \mathcal{C} \Delta^{-1} \partial_z \nabla_h \cdot (\partial_z \mathbf{b}_h - \nabla_h b_3 - \mu \nabla_h r)$ appearing on the left side of the equation by $a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)$ and compensate by adding $a(\varepsilon_M r^S) \rho(\varepsilon_M r^S) - 1$ times that term to the right side. By lemma 3.2 or by using the formulas in (3.27) there, the resulting large term exactly involves α . This yields

$$\begin{aligned} \rho(\varepsilon_M r^S)(\partial_t + (\mathbf{u}^S \cdot \nabla))\beta + \frac{\mu}{\sqrt{1+\mu^2}}(\mathbf{b}_h^S \cdot \nabla_h)\alpha + \frac{a(\varepsilon_M r^S)\rho(\varepsilon_M r^S)}{\varepsilon_M} \mathcal{D}\partial_z\alpha \\ = R_{\beta,1} + R_{\beta,2} + R_{\beta,3} + R_{\beta,4}, \end{aligned} \quad (5.28)$$

where

$$R_{\beta,1} := \mathcal{D}(1 - \alpha_z)R_3 + \mu^2 \mathcal{C} \Delta^{-1} \partial_z \nabla_h \cdot (1 - \alpha_z)R_2 \quad (5.29)$$

comes from the right sides of (5.20),

$$\begin{aligned}
R_{\beta,2} = & -\mu^2 \varepsilon_M \left[\frac{\mathcal{D}-1}{\mu^2}, \frac{\rho(\varepsilon_M r^S)-1}{\varepsilon_M} \right] (\partial_t + (\mathbf{u}^S \cdot \nabla)) u_3^I - \mu^2 \rho(\varepsilon_M r^S) \left[\frac{\mathcal{D}-1}{\mu^2}, \mathbf{u}^S \right] \cdot \nabla u_3^I \\
& - \mu^2 \left[\frac{\mathcal{D}-1}{\mu^2}, \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)-1}{\varepsilon_M} \right] \partial_z r^I \\
& - \mu^2 \varepsilon_M \left[\mathcal{C} \Delta^{-1} \partial_z \nabla_h, \frac{\rho(\varepsilon_M r^S)-1}{\varepsilon_M} \right] \cdot (\partial_t + (\mathbf{u}^S \cdot \nabla)) (1 - \alpha_z) u_h^F \\
& - \mu^2 \rho(\varepsilon_M r^S) \left[\mathcal{C} \Delta^{-1} \partial_z \nabla_h, \mathbf{u}^S \right] \nabla (1 - \alpha_z) u_h^F \\
& - \mu^2 \left[\mathcal{C} \Delta^{-1} \partial_z \nabla_h, \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)-1}{\varepsilon_M} \right] \cdot \nabla_h r^I
\end{aligned} \tag{5.30}$$

comes from commuting the operators applied to (5.20c) and (5.20b) past coefficients in those equations,

$$\begin{aligned}
R_{\beta,3} := & -\mu^2 \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)-1}{\varepsilon_M} \mathcal{C} \Delta^{-1} \Delta_h \partial_z r^I - \mu^2 \mathcal{C} \Delta^{-1} \partial_z \nabla_h \cdot (\nabla_h \frac{|\mathbf{b}|^2}{2} - (\mathbf{b} \cdot \nabla) \mathbf{b}_h) \\
& - \mu^2 \frac{a(\varepsilon_M r^S) \rho(\varepsilon_M r^S)-1}{\varepsilon_A} \mathcal{C} \partial_z (b_3^F + \mu \Delta^{-1} \Delta_h r)
\end{aligned} \tag{5.31}$$

comes from moving terms to the right side, balancing a term added on the left side, and using (3.12), and

$$R_{\beta,4} := \frac{\mu}{\sqrt{1+\mu^2}} (\mathbf{b}_h^S \cdot \nabla_h) \alpha + \mathcal{D} (\mathbf{b}_h^S \cdot \nabla_h) (1 - \alpha_z) b_3^F \tag{5.32}$$

comes from forcing the term involving $\mathbf{b}_h^S \cdot \nabla_h$ to have the desired form, and will be rearranged further later.

We now estimate the terms $R_{\alpha,i}$ and $R_{\beta,i}$. Since the operators applied to the R_i in $R_{\alpha,1}$ and $R_{\beta,1}$ are all bounded, (5.22) implies that

$$\|R_{\alpha,1}\|_j + \|R_{\beta,1}\|_j \leq c(\|(r^I, u_3^I)\|_j + \varepsilon_j), \quad j = 0, \dots, n-1. \tag{5.33}$$

The commutator terms in $R_{\alpha,2}$ and $R_{\beta,2}$ all involve commutators of a constant-coefficient zeroth-order pseudo-differential operator that is bounded as a function of μ with a function in H^n . Each of the two terms making up the commutator is therefore a bounded operator on H^j for $j \leq n$. The following lemma says that the commutator actually gains one derivative, which in many cases is a vital improvement.

Lemma 5.3. ([MS01, Lemma 2.5]) *Let $\widehat{P}(k, l, m)$ be homogeneous of degree zero and real analytic for $(k, l, m) \neq (0, 0, 0)$, and let P be the Fourier multiplier operator defined by $\widehat{P}f = \widehat{P}\widehat{f}$. Then for all $n \geq 3$, $f \in H^n(\mathbb{T}^3)$, $j \in 1, \dots, n$, and $g \in H^{j-1}$,*

$$\|[P, f]g\|_j \leq c\|f\|_n\|g\|_{j-1}. \tag{5.34}$$

The constant-coefficient pseudo-differential operators appearing in the commutators in $R_{\alpha,2}$ and $R_{\beta,2}$ are all homogeneous of degree zero and bounded uniformly in μ (in particular (5.42) below implies a bound on $\frac{\mathcal{D}-1}{\mu^2}$), and they are real analytic for $(k, l, m) \neq (0, 0, 0)$ since the denominators in the formulas for \mathcal{C} , \mathcal{D} , and \mathcal{Q} in (3.26), (3.27) are positive for $\mu < 1$, so they satisfy the conditions of Lemma 5.3. The expressions to which the commutators are applied have one of two forms: either they consist of a single spatial derivative applied to a fast or intermediate component that is estimated in (5.4) or (5.6), or they are some component of $(\partial_t + (\mathbf{u}^S \cdot \nabla)) \mathbf{V}$. In the former case the expression contains a factor μ^p with $p \geq 2$, so by Lemma 5.3 its H^j norm for $j = 0, \dots, n-1$ can be estimated by

$$\begin{aligned}
c\mu^2 \|\nabla((1 - \alpha_z) b_3^F, u_3^I, r^I, u_h^F)\|_{\max(0, j-1)} & \leq c\mu^2 \|((1 - \alpha_z) b_3^F, u_3^I, r^I, u_h^F)\|_{\max(1, j)} \\
& \leq c\mu^2 \varepsilon_{j+2} \leq c\varepsilon_j.
\end{aligned}$$

In the latter case the expression contains the factor $\mu^2 \varepsilon_M$ so by the interpolation bounds (2.7) its H^j norm for $j = 0, \dots, n-1$ can be estimated by $c\mu^2 \varepsilon_M (\|\mathbf{V}_t\|_{j-1} + c) \leq c\mu^2 \varepsilon_M \mu^{-j} \leq c\mu \varepsilon_j$. Hence

$$\|R_{\alpha,2}\|_j + \|R_{\beta,2}\|_j \leq c\varepsilon_j, \quad j = 0, \dots, n-1. \tag{5.35}$$

To estimate the terms in $R_{\alpha,3}$ and $R_{\beta,3}$ note that those terms either have a factor of $\mu^2 \varepsilon_M$, which is smaller than ε_j for any $j \in \{0, \dots, n-1\}$, have a zeroth-order pseudo-differential operator applied to $\partial_z r^j$, \mathbf{b}_z (after applying the z -derivative in the expression to some factor of \mathbf{b}), $\nabla_{\mathbf{h}} \cdot \mathbf{u}_{\mathbf{h}}$, or $\partial_z(b_3^F + \mu \Delta^{-1} \Delta_{\mathbf{h}} r)$, all of which have their H^j norms estimated in (5.4) or (5.6), or have a zeroth-order pseudo-differential operator applied to $\Delta^{-1} \partial_z[(\mathbf{b}_{\mathbf{h}} \cdot \nabla_{\mathbf{h}}) \partial_z u_3^I]$. In the first case using the uniform estimate (1.12) shows that the H^{n-1} norm is bounded by $c \varepsilon_0$. In the middle cases, since there each term also contains a factor of μ^2 the H^j norm is bounded by $c \mu^2 \varepsilon_{j+1} \leq c \varepsilon_j$. In the final case, since $\|\Delta^{-1} \partial_z f\|_j \leq c \|f\|_{\max(j-1,0)}$ and the term again contains a factor of μ^3 , it is bounded by $c \mu^3 \|\nabla \partial_z u_3^I\|_{\max(j-1,0)} \leq c \mu^3 \|\partial_z u_3^I\|_{\max(j,1)} \leq c \mu^3 \varepsilon_{j+2} \leq c \mu \varepsilon_j$. Together, these estimates yield

$$\|R_{\alpha,3}\|_j + \|R_{\beta,3}\|_j \leq c \varepsilon_j, \quad j = 0, \dots, n-1. \quad (5.36)$$

Before estimating the remainders $R_{\alpha,4}$ and $R_{\beta,4}$ we must rearrange those expressions further. For $j \leq n-2$ a simpler method suffices because (5.4) and (5.6) imply that

$$\mu \|\mathbf{V}_*^F\|_{j+1} + \mu^2 \|\mathbf{V}^I\|_{j+1} \leq c(\mu \varepsilon_{j+1} + \mu^2 \varepsilon_{j+2}) \leq c \varepsilon_j, \quad j = 0, \dots, n-2, \quad (5.37)$$

where \mathbf{V}_*^F means all components of \mathbf{V}^F except $\alpha_z b_3^F$, which is not estimated in (5.4). However, (5.37) is not valid for $j = n-1$ because (5.4) and (5.6) do not hold for $j = n$. For $j \leq n-2$ we can therefore obtain the estimate $\|R_{\alpha,4}\|_j + \|R_{\beta,4}\|_j \leq c \varepsilon_j$ by using the formulas for α and β in (5.18) together with the facts that $\frac{\mathcal{D}-1}{\mu^2}$ is a bounded zeroth-order pseudo-differential operator and that

$$b_3^F = (b_3^F + \mu \Delta^{-1} \Delta_{\mathbf{h}} r) + \mu \Delta^{-1} \partial_z^2 r^I - \mu(r - \text{av} r) \quad (5.38)$$

plus estimates similar to those used for $R_{\alpha,3}$ and $R_{\beta,3}$.

However, in order to obtain an estimate in H^{n-1} rearrangement is necessary, and it is sufficient to show that $R_{\alpha,4}$ and $R_{\beta,4}$ can be expressed as linear combinations of the terms

$$\nabla_{\mathbf{h}} \cdot \mathbf{u}_{\mathbf{h}}, \quad \partial_z \mathbf{u}^F, \quad \partial_z \mathbf{b}, \quad \nabla(b_3^F + \Delta^{-1} \Delta_{\mathbf{h}} r), \quad \partial_z r \quad \text{and} \quad \partial_z u_3 \quad (5.39)$$

whose H^j norms are estimated in (5.4) or (5.6) even though they involve a first derivative of \mathbf{V} , and in addition a factor of μ must be present multiplying the terms $\partial_z(r^I, u_3^I)$ to compensate for the extra factor of $\frac{1}{\mu}$ in (5.6).

Substituting (5.38) into the formula for α in (5.18) and solving the result for r^I yields

$$(1 + \mu^2) r^I = \alpha + \mu(1 - \alpha_z)(b_3^F + \mu \Delta^{-1} \Delta_{\mathbf{h}} r) + \mu^2 \Delta^{-1} \partial_z^2 r^I - \mu^3 \mathcal{C} \mathcal{D} \Delta^{-1} \partial_z^2 r^I. \quad (5.40)$$

Applying $(1 - \alpha_z)$ to (5.38), which turns the final $-\mu(r - \text{av} r)$ into $-\mu r^I$, dividing (5.40) by $1 + \mu^2$ and substituting the result for that final r^I , and substituting the result into (5.32) shows that $R_{\beta,4}$ equals

$$\mu \left(\frac{1}{\sqrt{1+\mu^2}} - \frac{1}{1+\mu^2} \mathcal{D} \right) (\mathbf{b}_{\mathbf{h}} \cdot \nabla_{\mathbf{h}}) \alpha \quad (5.41)$$

plus a sum of terms involving operators of order zero applied to the expressions in (5.39) whose H^j norms can be bounded by $c \varepsilon_j$ using (5.4) and (5.6). To estimate (5.41) we use the identity

$$\begin{aligned} \mathcal{D} &= \sqrt{1 + \mu^2} \\ &- \frac{4\mu^2 \Delta^{-1} \partial_z^2}{\left(\sqrt{\mu^2(2-4\Delta^{-1} \partial_z^2) + \mu^4 + 1 + \mu^2 + 1} \right) \left(\sqrt{2} \sqrt{\sqrt{\mu^2(2-4\Delta^{-1} \partial_z^2) + \mu^4 + 1 + \mu^2 + 1 + 2\sqrt{\mu^2 + 1}}} \right)} \end{aligned} \quad (5.42)$$

derived from (3.27). Since the constant term cancels, the fact that the other term in (5.42) contains a factor of ∂_z multiplied by an operator of order -1 together with the facts that the z -derivative of all the constituents of α are estimated in (5.4) and (5.6) and that an overall factor of μ is present in (5.42) implies that the H^j norm of $R_{\beta,4}$ is bounded by a constant times ε_j even for $j = n-1$. Similarly, upon substituting formula (5.18) for β into the definition (5.27) of $R_{\alpha,4}$ and then substituting (5.42) into the result, the constant term from (5.42)

cancel and the remaining terms all involve the expressions from (5.39) and so can be estimated by a constant times ε_j since an overall factor of μ is present. This shows that

$$\|R_{\alpha,4}\|_j + \|R_{\beta,4}\|_j \leq c\varepsilon_j, \quad j = 0, \dots, n-1. \quad (5.43)$$

of *Theorem 5.2*. The system (5.23), (5.28) has the form of the Klainerman-Majda two-scale theory. Moreover, the formulas (5.18) plus estimates similar to those used above show that

$$\|\alpha - r^I\|_j + \|\beta - u_3^I\|_j \leq c\varepsilon_j, \quad j = 0, \dots, n-1. \quad (5.44)$$

Together with the estimates (5.33), (5.35), (5.36), and (5.43), (5.44) shows that the H^j norm of the right sides of those equations is bounded by a constant times $\|(\alpha, \beta)\|_j + \varepsilon_j$. Hence the standard Klainerman-Majda energy estimates show that

$$\max_{0 \leq t \leq T} \|(\alpha, \beta)\|_j \leq c(\|(\alpha(0), \beta(0))\|_j + \varepsilon_j), \quad j = 0, \dots, n-1. \quad (5.45)$$

Since the form of the initial data from (3.30), (3.36) implies that $\|\mathbf{V}^I\|_{n-1}|_{t=0} \leq c\varepsilon_M = c\varepsilon_1$, the estimate (5.44) implies that $\|(\alpha(0), \beta(0))\|_j \leq c\varepsilon_1$ and hence (5.45) yields $\max_{0 \leq t \leq T} \|(\alpha, \beta)\|_j \leq c \max(\varepsilon_j, \varepsilon_1)$ for $j = 0, \dots, n-1$. Using (5.44) once more then yields (5.16). \square \square

5.3. Equations and estimates for horizontal components of the slow mode. In similar fashion to the intermediate mode dynamic estimates, in order to estimate the difference between the slow modes of the solution to the original system and the solution of the limit system, it is necessary to obtain PDEs for the exact zero eigenspace of the large operator $\mathcal{L}_A + \mu\mathcal{L}_M$. For the horizontal components of the velocity and magnetic field the slow modes belong to that exact zero eigenspace, so we will write the evolution equations for the slow modes in the form of the limit equations with error terms added. To do that we apply the projection \mathbb{P} onto the slow horizontal modes to the original system, expand all dependent variables into fast, intermediate, and slow modes, and move all terms except those involving purely slow modes to the right sides of the equations. The slow part of the density and vertical velocity will be treated in the following subsection.

Theorem 5.4. *Assume that the conditions of Theorem 1.2 hold. Let $(r, \mathbf{u}, \mathbf{b})$ be the solution obtained for the MHD system (1.4). In addition, let $(\bar{r}, \bar{\mathbf{u}}_h, \bar{\mathbf{b}}_h)$ be the solution of the limit system (1.13a), (1.13b), (1.13d) with the initial data equal to the $O(1)$ part $(r^{0,S}, \mathbf{u}^{0,S}, \mathbf{b}^{0,S} + \mu_{\text{lim}}(1 - \text{av})r^{0,S})$ of the initial data (3.30) of the original system. Then*

$$\|\mathbf{u}_h^S - \bar{\mathbf{u}}_h\|_{n-2} + \|\mathbf{b}_h^S - \bar{\mathbf{b}}_h\|_{n-2} \leq c\varepsilon_M. \quad (5.46)$$

Before proving Theorem 5.4 we need to derive appropriate equations. Since all slow modes contain the averaging operator \mathbf{a}_z in the z direction, it will be convenient to write the equations in conservation form, so that derivatives with respect to z disappear when \mathbf{a}_z is applied. In particular,

$$\mathbf{a}_z(\mathbf{w} \cdot \nabla f + (\nabla \cdot \mathbf{w})f) = \mathbf{a}_z \nabla \cdot (f\mathbf{w}) = \mathbf{a}_z \nabla_h \cdot (f\mathbf{w}_h) = \mathbf{a}_z(\mathbf{w}_h \cdot \nabla_h f + (\nabla_h \cdot \mathbf{w}_h)f). \quad (5.47)$$

Using the continuity equation (A.1a), the momentum equation (1.4b) can be rewritten as

$$\partial_t(\rho\mathbf{u}) + \mathbf{u} \cdot \nabla(\rho\mathbf{u}) + (\nabla \cdot \mathbf{u})(\rho\mathbf{u}) + \nabla\Phi - \mathbf{b} \cdot \nabla\mathbf{b} = \varepsilon_A^{-1}(\partial_z\mathbf{b} - \nabla b_3), \quad (5.48)$$

where Φ is some scalar-valued function defined on \mathbb{T}^3 , i.e., periodic in x, y, z . Apply $\mathcal{P}_h^{\text{div}}\mathbf{a}_z$ to the first two components of (5.48) and \mathbf{a}_z to the first two components of (1.4c), and simplify the result by using the identity (5.47) not only with $\mathbf{w} = \mathbf{u}$ but also with $\mathbf{w} = \mathbf{b}$ since the constraint $\nabla \cdot \mathbf{b} = 0$ implies that $\mathbf{b} \cdot \nabla f$ equals $\nabla \cdot (f\mathbf{b})$. The resulting equations can be further simplified by using the definitions (3.7) of the modes to obtain the

identities $\alpha_z b_h = b_h^S$ and $\alpha_z(\rho u_h) = \alpha_z(\rho u_h^S) + \alpha_z(\rho u_h^F) = (\alpha_z \rho)u_h^S + \alpha_z(\rho u_h^F)$, and then using the facts that $\rho = 1 + \varepsilon_M r$ and $\alpha_z u_h^F$ is a gradient to obtain $\mathcal{P}_h^{\text{div}} \alpha_z(\rho u_h^F) = \varepsilon_M \mathcal{P}_h^{\text{div}} \alpha_z(r u_h^F)$. This yields

$$\partial_t \mathcal{P}_h^{\text{div}}((\alpha_z \rho)u_h^S) + \mathcal{P}_h^{\text{div}} \alpha_z \nabla_h \cdot (\rho u_h \otimes u_h - b_h \otimes b_h) = -\varepsilon_M \partial_t \mathcal{P}_h^{\text{div}} \alpha_z(r u_h^F), \quad (5.49a)$$

$$\partial_t b_h^S + \alpha_z \nabla_h \cdot (u_h \otimes b_h - b_h \otimes u_h) = 0, \quad (5.49b)$$

where the tensor product \otimes follows the convention that $\nabla_h \cdot (u_h \otimes b_h) = u_h \cdot \nabla_h b_h + (\nabla_h \cdot u_h) b_h$. Since $\partial_t \mathcal{P}_h^{\text{div}}((\alpha_z \rho)u_h^S) = \partial_t u_h^S + \varepsilon_M \partial_t \mathcal{P}_h^{\text{div}}((\alpha_z r)u_h^S)$, (5.49) together with the bounds (1.12) and (1.10) and the relation (1.14) show that

$$\|\partial_t u_h^S\|_{n-1} + \|\partial_t b_h^S\|_{n-1} \leq c. \quad (5.50)$$

Recalling that u_h and b_h have no intermediate part, we separate them into their fast and slow parts in the tensor products in (5.49a):

$$\begin{aligned} \alpha_z \nabla_h \cdot (\rho u_h \otimes u_h - b_h \otimes b_h) &= \alpha_z \nabla_h \cdot (\rho u_h^S \otimes u_h^S - b_h^S \otimes b_h^S) \\ &+ \alpha_z \nabla_h \cdot \left((\rho u_h^S \otimes u_h^F - b_h^S \otimes b_h^F) + \text{trsp} \right) + \alpha_z \nabla_h \cdot (\rho u_h^F \otimes u_h^F - b_h^F \otimes b_h^F) \end{aligned} \quad (5.51)$$

where trsp denotes the tensor transpose. By (3.9)–(3.10), the slow parts (u_h^S, b_h^S) are independent of z and divergence-free, so the first term on the right side of (5.51) simplifies to $u_h^S \cdot \nabla_h((\alpha_z \rho)u_h^S) - b_h^S \cdot \nabla_h b_h^S$. Also, since b_h^S is independent of z while (3.7) shows that $\alpha_z b_h^F = 0$ (although $\alpha_z u_h^F \neq 0$ in general), the expression $-\alpha_z \nabla_h \cdot (b_h^S \otimes b_h^F)$ is identically zero and so can be omitted from the second term on the right side of (5.51). Next, we can drop the $\mathcal{P}_h^{\text{div}}$ operator from (5.49a) at the cost of adding a term $\nabla_h \theta(t, x, y)$ to that equation, since a 2-vector is in the kernel of $\mathcal{P}_h^{\text{div}}$ if and only if it is the horizontal gradient of some scalar-valued function. Thus, (5.49a) becomes

$$\begin{aligned} &\partial_t((\alpha_z \rho)u_h^S) + u_h^S \cdot \nabla_h((\alpha_z \rho)u_h^S) - b_h^S \cdot \nabla_h b_h^S \\ &= -\varepsilon_M \partial_t \alpha_z(r u_h^F) - \alpha_z \nabla_h \cdot (\rho u_h^S \otimes u_h^F + \text{trsp}) \\ &\quad - \alpha_z \nabla_h \cdot (\rho u_h^F \otimes u_h^F - b_h^F \otimes b_h^F) + \nabla_h \theta. \end{aligned}$$

The second term on the right side is a ‘‘slow-fast’’ product, which can be rewritten using time-integrated variable

$$A(t, \cdot) := \int_0^t \alpha_z u_h^F dt', \quad (5.52)$$

as the time derivative of a small term plus a small term, since (5.5) shows that

$$\|A\|_n \leq c \varepsilon_M^{1+\nu}. \quad (5.53)$$

For example,

$$\begin{aligned} \alpha_z(\rho u_h^S \otimes u_h^F) &= \alpha_z(u_h^S \otimes u_h^F) + \varepsilon_M \alpha_z(r u_h^S \otimes u_h^F) = u_h^S \otimes (\alpha_z u_h^F) + \varepsilon_M \alpha_z(r u_h^S \otimes u_h^F) \\ &= \partial_t(u_h^S \otimes A) - (\partial_t u_h^S) \otimes A + \varepsilon_M \alpha_z(r u_h^S \otimes u_h^F). \end{aligned}$$

Hence we obtain

$$(\partial_t + (u_h^S \cdot \nabla_h))((\alpha_z \rho)u_h^S) - b_h^S \cdot \nabla_h b_h^S = \partial_t \Xi_1 + \xi_1 + \nabla_h \theta, \quad \nabla_h \cdot u_h^S = 0 \quad (5.54)$$

with

$$\begin{aligned} \Xi_1 &:= -\varepsilon_M \alpha_z(r u_h^F) - \nabla_h \cdot (u_h^S \otimes A + \text{trsp}), \\ \xi_1 &:= \nabla_h \cdot \left\{ (\partial_t u_h^S) \otimes A + \text{trsp} \right\} - \varepsilon_M \alpha_z \nabla_h \cdot (r u_h^S \otimes u_h^F + \text{trsp}) \\ &\quad - \alpha_z \nabla_h \cdot (\rho u_h^F \otimes u_h^F - b_h^F \otimes b_h^F). \end{aligned}$$

The bound (5.4) together with the constraint (1.10), the relation (1.14) between the parameters and the definition (5.1) of the ε_j implies that $\|u_h^F\|_{n-1} + \|b_h^F\|_{n-1} \leq c \varepsilon_M^\nu$. Using that estimate, the time-derivative bound (5.50), estimate (5.53) for A , the formula $\rho = 1 + \varepsilon_M r$, and Corollary C.3 yields the estimates

$$\|\Xi_1\|_{n-1} + \|\xi_1\|_{n-2} \leq c \varepsilon_M^{1+\nu}. \quad (5.55)$$

Applying the same ideas to the slow magnetic equation (5.49b), and in particular noting that $\mathbf{a}_z(u_h^S \otimes \mathbf{b}_h^F + \text{trsp}) = 0$ yields

$$(\partial_t + (u_h^S \cdot \nabla_h))\mathbf{b}_h^S - \mathbf{b}_h^S \cdot \nabla_h u_h^S = \partial_t \Xi_{b_h^S} + \xi_{b_h^S}, \quad \nabla_h \cdot \mathbf{b}_h^S = 0, \quad (5.56)$$

with

$$\Xi_{b_h^S} := -\nabla_h \cdot [A \otimes \mathbf{b}_h^S - \text{trsp}], \quad \xi_{b_h^S} := \nabla_h \cdot [A \otimes \partial_t \mathbf{b}_h^S - \text{trsp}] - \mathbf{a}_z \nabla_h \cdot (u_h^F \otimes \mathbf{b}_h^F - \text{trsp}). \quad (5.57)$$

The same bounds as for (5.55) show that

$$\|\Xi_{b_h^S}\|_{n-1} + \|\xi_{b_h^S}\|_{n-2} \leq c \varepsilon_M^{1+\nu}. \quad (5.58)$$

In order to simplify the equation for the time evolution of u_h^S further, we need to use an equation for the time evolution of r^S . Using the first part of (3.10) and the facts from (3.7) that $u_h^I = 0$ and $\mathbf{a}_z r = r^S$, the vertically-averaged conservation law form (4.4) can be simplified to

$$\partial_t r^S + u_h^S \cdot \nabla_h (r^S) = -\mathbf{a}_z \nabla_h \cdot (r u_h^F) - \varepsilon_M^{-1} \nabla_h \cdot (\mathbf{a}_z u_h^F). \quad (5.59)$$

Using the time-integrated variable A from (5.52), (5.59) can be rewritten as

$$(\partial_t + (u_h^S \cdot \nabla_h))(\varepsilon_M r^S) = -\varepsilon_M \mathbf{a}_z \nabla_h \cdot (r u_h^F) - \partial_t \nabla_h \cdot A. \quad (5.60)$$

Subtracting u_h^S times (5.60) from (5.54) noting that $\mathbf{a}_z \rho = 1 + \varepsilon_M r^S$, and rewriting the term $u_h^S \partial_t \nabla_h \cdot A$ on the right side of the result as $\partial_t (u_h^S (\nabla_h \cdot A)) - (\partial_t u_h^S) \nabla_h \cdot A$ yields

$$(1 + \varepsilon_M r^S)(\partial_t + (u_h^S \cdot \nabla_h))u_h^S - \mathbf{b}_h^S \cdot \nabla_h \mathbf{b}_h^S = \partial_t \Xi_2 + \xi_2 + \nabla_h \theta_2, \quad \nabla_h \cdot u_h^S = 0, \quad (5.61)$$

where

$$\Xi_2 := \Xi_1 + u_h^S \nabla_h \cdot A, \quad \xi_2 := \xi_1 - (\partial_t u_h^S) \nabla_h \cdot A + \varepsilon_M u_h^S \mathbf{a}_z \nabla_h \cdot (r u_h^F) \quad (5.62)$$

also satisfy

$$\|\Xi_2\|_{n-1} + \|\xi_2\|_{n-2} \leq c \varepsilon_M^{1+\nu} \quad (5.63)$$

in view of the estimate (5.55) and the same bounds used to obtain that estimate.

We now move the $\varepsilon_M r^S$ term from the left to the right side and also replace the time-derivative term on the right side of (5.61) by its divergence-free part, which only changes the divergence term, obtaining

$$(\partial_t + (u_h^S \cdot \nabla_h))u_h^S - \mathbf{b}_h^S \cdot \nabla_h \mathbf{b}_h^S = \partial_t \Xi_{u_h^S} + \xi_{u_h^S} + \nabla_h \theta_{u_h^S}, \quad \nabla_h \cdot u_h^S = 0 \quad (5.64)$$

with

$$\Xi_{u_h^S} := \mathcal{P}_h^{\text{div}} \Xi_2, \quad \xi_{u_h^S} := \xi_2 - \varepsilon_M r^S (\partial_t + (u_h^S \cdot \nabla_h))u_h^S. \quad (5.65)$$

Using the estimate (5.63), the fact from (5.50) that $\partial_t u_h^S = O(1)$, and the fact that the projection $\mathcal{P}_h^{\text{div}}$ does not increase Sobolev norms yields

$$\|\Xi_{u_h^S}\|_{n-1} \leq c \varepsilon_M^{1+\nu}, \quad \|\xi_{u_h^S}\|_{n-2} \leq c \varepsilon_M \quad (5.66)$$

in view of the term that is explicitly $O(\varepsilon_M)$ in the definition (5.65) of $\xi_{u_h^S}$.

of Theorem 5.4. The functions (\bar{u}_h, \bar{b}_h) satisfy the systems

$$(\partial_t + (\bar{u}_h \cdot \nabla_h))\bar{u}_h - (\bar{b}_h \cdot \nabla_h)\bar{b}_h = \nabla_h \Phi, \quad \nabla_h \cdot \bar{u}_h = 0, \quad (5.67)$$

$$(\partial_t + (\bar{u}_h \cdot \nabla_h))\bar{b}_h - (\bar{b}_h \cdot \nabla_h)\bar{u}_h = 0. \quad (5.68)$$

We now apply Theorem B.1 to the system (5.64), (5.56) for $u := (u_h, b_h)$ and the system (5.67), (5.68) for $U := (\bar{u}_h, \bar{b}_h)$. Assumptions (3.30) and (3.36) ensure that the difference in their initial data is $O(\varepsilon_M^{1+\nu})$. On account of the operator $\mathcal{P}_h^{\text{div}}$ appearing in the definition (5.65) of $\Xi_{u_h^S}$, $\Xi_u := (\Xi_{u_h^S}, \Xi_{b_h^S})$ and $\Xi_U := 0$ satisfy $L\Xi_u = 0 = L\Xi_U$, where $L = (\nabla_h \cdot 0)$. Define $\xi_u = (\xi_{u_h^S}, \xi_{b_h^S})$ and $\xi_U = 0$. The estimates (5.66), (5.58)

together with the above-mentioned estimate on the difference of the initial data then imply that the hypotheses of Theorem B.1 hold with $k = n$, $r = 1$, and $\delta = \varepsilon_M$. Hence the conclusion of that theorem yields (5.46). \square \square

5.4. Equations and estimates for remaining slow modes. The third component of equation (1.4b) implies that there are no large terms in the PDE for $u_3^S := \alpha_z u_3$, i.e., $\left(0, \begin{pmatrix} 0_2 \\ u_3^S \end{pmatrix}, 0_3\right)$ is a zero eigenvector of the full large operator $\mathcal{L}_A + \mu \mathcal{L}_M$. However, as shown in Lemma 3.1, the μ -dependent zero eigenvector of $\mathcal{L}_A + \mu \mathcal{L}_M$ having a nontrivial projection onto the density component is not just the slow mode r^S but is a linear combination of r^S and b_3 . Specifically, (3.4) implies that $\mathbf{V}_r = \left(\alpha_z, 0_3, \begin{pmatrix} 0_2 \\ -\mu \alpha_z \end{pmatrix}\right)$ satisfies $(\mathcal{L}_A + \mu \mathcal{L}_M) \mathbf{V}_r \cdot \tilde{\mathbf{V}} = 0$ for all vector-valued functions $\tilde{\mathbf{V}}$, which by the MHD system (1.4) and the skew-adjointness of $(\mathcal{L}_A + \mu \mathcal{L}_M)$ implies that the PDE satisfied by $\alpha_z(r - \mu b_3) = \mathbf{V}_r \cdot \mathbf{V}$ will only contain terms that are at most $O(1)$. We therefore need to calculate the PDE system satisfied by $\alpha_z(r - \mu b_3)$ and $\alpha_z u_3$. It turns out that while the PDEs for those two variables are coupled by terms that are strictly $O(1)$, the coupling to other components of the solution contains only terms that are $o(1)$ and so can be considered as small perturbations.

Theorem 5.5. *Under the conditions of Theorem 5.4,*

$$\|r^S - \bar{r}\|_{n-2} + \|u_3^S - \bar{u}_3\|_{n-2} \leq c \left[\varepsilon_M^{1-\max(n-5,0)\nu} + |\mu - \mu_{\text{lim}}| \right]. \quad (5.69)$$

Proof. Writing the variables in (4.6) in terms of fast, intermediate, and slow components and using the facts that the slow components are independent of z , u_h^S and b_h^S have zero horizontal divergence, the vertical averages of r^I , u_3^I and b_h^F vanish, and b_3^S is constant in time as well as space transforms that equation into

$$\begin{aligned} & \partial_t(r^S - \mu \alpha_z b_3^F) + (u_h^S \cdot \nabla_h)(r^S - \mu \alpha_z b_3^F) + \mu (b_h^S \cdot \nabla_h) u_3^S \\ &= -\alpha_z \{ \nabla_h \cdot [(r - \mu b_3) u_h^F] \} - \mu \alpha_z [\nabla_h \cdot (u_3^I b_h^F)] \\ &= -\nabla_h \cdot [(r^S - \mu \alpha_z b_3) \alpha_z u_h^F] - \alpha_z \nabla_h \cdot \left[(r^I - \mu(1 - \alpha_z) b_3^F) u_h^F \right] - \mu \alpha_z [\nabla_h \cdot (u_3^I b_h^F)] \end{aligned} \quad (5.70)$$

Replacing u_h^S , b_h^S , and μ on the left side of (5.70) by their limit values, and compensating by adding terms to the right side yields

$$[\partial_t + (\bar{u}_h \cdot \nabla_h)](r^S - \mu \alpha_z b_3^F) + \mu_{\text{lim}} (\bar{b}_h \cdot \nabla_h) u_3^S = \partial_t \Xi_{r^S - \mu \alpha_z b_3^F} + \xi_{r^S - \mu \alpha_z b_3^F}, \quad (5.71)$$

where

$$\begin{aligned} \Xi_{r^S - \mu \alpha_z b_3^F} &:= -\nabla_h \cdot \left[(r^S - \mu \alpha_z b_3) \int_0^t \alpha_z u_h^F \right] = -\nabla_h \cdot [(r^S - \mu \alpha_z b_3) \mathbf{A}] \\ \xi_{r^S - \mu \alpha_z b_3^F} &:= \nabla_h \cdot [(\partial_t(r^S - \mu \alpha_z b_3)) \mathbf{A}] - \alpha_z \nabla_h \cdot \left[(r^I - \mu(1 - \alpha_z) b_3^F) u_h^F \right] \\ &\quad - \mu \alpha_z [\nabla_h \cdot (u_3^I b_h^F)] + (\bar{u}_h - u_h^S) \cdot \nabla_h (r^S - \mu \alpha_z b_3^F) \\ &\quad + \mu ((\bar{b}_h - b_h^S) \cdot \nabla_h) u_3^S + (\mu_{\text{lim}} - \mu) (\bar{b}_h \cdot \nabla_h) u_3^S. \end{aligned} \quad (5.72)$$

Since $\partial_t b_3^S \equiv 0$, (5.70) implies a uniform H^{n-1} bound for $\partial_t(r^S - \mu \alpha_z b_3)$. Using in addition the uniform estimate (1.12), estimate (5.53) for \mathbf{A} , the estimate (5.46) for the convergence rate of the horizontal components, and Corollary C.3 shows that

$$\|\Xi_{r^S - \mu \alpha_z b_3^F}\|_{n-1} \leq c \varepsilon_M^{1+\nu}, \quad \|\xi_{r^S - \mu \alpha_z b_3^F}\|_{n-2} \leq c [\varepsilon_M + |\mu - \mu_{\text{lim}}|]. \quad (5.73)$$

The third component of (5.48) can be written as

$$\partial_t(\rho u_3) + \mathbf{u} \cdot \nabla(\rho u_3) + (\nabla \cdot \mathbf{u})(\rho u_3) + \partial_z \Phi - b_h \cdot \nabla_h b_3 = 0.$$

Applying the vertical averaging operator α_z and using (5.47) reduces this to

$$\partial_t(\alpha_z(\rho u_3)) + \alpha_z [u_h \cdot \nabla_h(\rho u_3)] + \alpha_z [(\nabla_h \cdot u_h)(\rho u_3)] - \alpha_z [b_h \cdot \nabla_h b_3^F] = 0.$$

In order to treat the term $\mathbf{a}_z [\mathbf{b}_h \cdot \nabla_h b_3^F]$ we write $\mathbf{a}_z b_3^F$ as

$$\begin{aligned} \mathbf{a}_z b_3^F &= -\mu \Delta^{-1} \Delta_h r^S + \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S) \\ &= -\mu (r^S - \text{av} r^S) + \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S) \end{aligned} \quad (5.74)$$

in accordance with the expression estimated in (5.4). Using (5.74) while noting that $(\mathbf{b}_h^S \cdot \nabla_h)(\text{av} r^S) = 0$, and using the facts from (1.6), (3.9), and (3.7) that $\rho = 1 + \varepsilon_M r$, r^S and u_3^S are independent of z and $\mathbf{a}_z u_3^I = 0 = u_3^F$, which imply that $\mathbf{a}_z (\rho u_3^S) = (1 + \varepsilon_M r^S) u_3^S$, $\mathbf{a}_z (\rho u_3^I) = \varepsilon_M \mathbf{a}_z (r^I u_3^I)$, and $\mathbf{a}_z r = r^S$, we obtain

$$\begin{aligned} (1 + \varepsilon_M r^S) [\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)] u_3^S + u_3^S [\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)] (\varepsilon_M r^S) + \mu (\mathbf{b}_h^S \cdot \nabla_h) r^S \\ = (\mathbf{b}_h^S \cdot \nabla_h) [\mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S)] + \mathbf{a}_z \{ (\mathbf{b}_h^F \cdot \nabla_h) (1 - \mathbf{a}_z) b_3^F \} \\ - \varepsilon_M (\mathbf{u}_h^S \cdot \nabla_h) (\mathbf{a}_z (r^I u_3^I)) - \varepsilon_M \partial_t \mathbf{a}_z (r^I u_3^I) \\ - \nabla_h \cdot [(\partial_t A) u_3^S] - \mathbf{a}_z \{ \nabla_h \cdot (\mathbf{u}_h^F u_3^I) \} - \varepsilon_M \mathbf{a}_z \{ \nabla_h \cdot (\mathbf{u}_h^F r u_3) \}, \end{aligned} \quad (5.75)$$

where the last line results from separating the various modes in $\mathbf{a}_z \{ \nabla_h \cdot (\mathbf{u}_h^F \rho u_3) \}$ and using the definition of A from (5.52). Since there are no terms of size ε_A^{-1} in the equations for the time evolution of r or u_3 , (5.75) implies that

$$\|\partial_t u_3^S\|_{n-1} \leq c. \quad (5.76)$$

Subtracting u_3^S times (5.60) from (5.75), moving the term $\varepsilon_M r^S [\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)] u_3^S$ to the right side of the result, noting that the two terms involving A partially cancel and rewriting the remaining term $(\partial_t A \cdot \nabla_h) u_3^S$ on the right side of the result as $\partial_t [(A \cdot \nabla_h) u_3^S] - (A \cdot \nabla_h) \partial_t u_3^S$ yields

$$(\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)) u_3^S + \mu (\mathbf{b}_h^S \cdot \nabla_h) r^S = \partial_t \tilde{\Xi}_{u_3^S} + \tilde{\xi}_{u_3^S} + (\mathbf{b}_h^S \cdot \nabla_h) [\mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S)] \quad (5.77)$$

where

$$\begin{aligned} \tilde{\Xi}_{u_3^S} &= (A \cdot \nabla_h) u_3^S - \varepsilon_M \mathbf{a}_z (r^I u_3^I), \\ \tilde{\xi}_{u_3^S} &= \mathbf{a}_z \{ (\mathbf{b}_h^F \cdot \nabla_h) (1 - \mathbf{a}_z) b_3^F \} - \varepsilon_M (\mathbf{u}_h^S \cdot \nabla_h) (\mathbf{a}_z (r^I u_3^I)) - \mathbf{a}_z \{ \nabla_h \cdot (\mathbf{u}_h^F u_3^I) \} \\ &\quad - \varepsilon_M \mathbf{a}_z \{ \nabla_h \cdot (\mathbf{u}_h^F r u_3) \} + \varepsilon_M u_3^S \mathbf{a}_z \nabla_h \cdot (r u_h^F) - (A \cdot \nabla_h) (\partial_t u_3^S) \\ &\quad - \varepsilon_M r^S [\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)] u_3^S. \end{aligned}$$

As a step towards symmetrizing the system consisting of (5.71), (5.77), we want to replace r^S in the last term on the left side of (5.77) by $r^S - \mu \mathbf{a}_z b_3^F$, which requires adding a balancing term involving $\mathbf{a}_z b_3^F$, which must also be rewritten using (5.74). This leads us to write $\mu (\mathbf{b}_h^S \cdot \nabla_h) r^S$ as $k_1 (\mathbf{b}_h^S \cdot \nabla_h) (r^S - \mu \mathbf{a}_z b_3^F) + k_2 (\mathbf{b}_h^S \cdot \nabla_h) (\mathbf{a}_z b_3^F + \mu r^S)$. Equating those two expressions and comparing the coefficients of $(\mathbf{b}_h^S \cdot \nabla_h) (\mathbf{a}_z b_3^F)$ shows that $k_2 = k_1 \mu$, and then comparing the coefficients of $(\mathbf{b}_h \cdot \nabla_h) r^S$ yields $k_1 = \frac{\mu}{1 + \mu^2}$, i.e.,

$$\begin{aligned} \mu (\mathbf{b}_h^S \cdot \nabla_h) r^S &= \frac{\mu}{1 + \mu^2} (\mathbf{b}_h^S \cdot \nabla_h) (r^S - \mu \mathbf{a}_z b_3^F) + \frac{\mu^2}{1 + \mu^2} (\mathbf{b}_h^S \cdot \nabla_h) (\mathbf{a}_z b_3^F + \mu r^S) \\ &= \frac{\mu}{1 + \mu^2} (\mathbf{b}_h^S \cdot \nabla_h) (r^S - \mu \mathbf{a}_z b_3^F) + \frac{\mu^2}{1 + \mu^2} (\mathbf{b}_h^S \cdot \nabla_h) [\mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S)], \end{aligned} \quad (5.78)$$

where the second equation follows as in (5.74). Substituting (5.78) into (5.77), moving the second term from (5.78) to the right side of the resulting equation and combining it with the similar term already present, and replacing \mathbf{u}_h^S , \mathbf{b}_h^S and μ on the left side of the result by their limiting values and compensating on the right side yields

$$[\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)] u_3^S + \frac{\mu_{\text{lim}}}{1 + \mu_{\text{lim}}^2} (\bar{\mathbf{b}}_h \cdot \nabla_h) (r^S - \mu \mathbf{a}_z b_3^F) = \partial_t \Xi_{u_3^S} + \xi_{u_3^S} \quad (5.79)$$

where

$$\Xi_{u_3^S} = \tilde{\Xi}_{u_3^S} + \frac{1}{1 + \mu^2} (\mathbf{b}_h^S \cdot \nabla_h) \left[\int_0^t \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S) \right] \quad (5.80)$$

$$\begin{aligned}
\xi_{u_3^S} &= \tilde{\xi}_{u_3^S} + \frac{1}{1+\mu^2} (\partial_t \mathbf{b}_h^S \cdot \nabla_h) \left[- \int_0^t \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r^S) \right] \\
&+ (\bar{\mathbf{u}}_h - \mathbf{u}_h^S) \cdot \nabla_h u_3^S + \frac{\mu}{1+\mu^2} ((\bar{\mathbf{b}}_h - \mathbf{b}_h^S) \cdot \nabla_h) (\mathbf{a}_z b_3^F + \mu r^S) \\
&+ \left(\frac{\mu_{\text{lim}}}{1+\mu_{\text{lim}}^2} - \frac{\mu}{1+\mu^2} \right) (\bar{\mathbf{b}}_h \cdot \nabla_h) (\mathbf{a}_z b_3^F + \mu r^S).
\end{aligned} \tag{5.81}$$

The system consisting of (5.71), (5.79) can be symmetrized by multiplying the latter equation by $1 + \mu_{\text{lim}}^2$. The estimates used to obtain (5.73) together with the time-derivative estimates (5.50), (5.76) and the time-integrated estimate (5.5) show that

$$\|\bar{\Xi}_{u_3^S}\|_{n-1} \leq c \varepsilon_M^{1+\nu}, \tag{5.82}$$

$$\|\xi_{u_3^S}\|_{n-2} \leq c [\varepsilon_M + |\mu - \mu_{\text{lim}}|]. \tag{5.83}$$

Using (1.13e) and the fact that $\text{av } r^{0,S}$ is a constant, the limit equations (1.13a) and (1.13c) can be rewritten as the system

$$[\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)] \{ (1 + \mu_{\text{lim}}^2) \bar{r} - \mu_{\text{lim}}^2 \text{av } r^{0,S} \} + \mu_{\text{lim}} (\bar{\mathbf{b}}_h \cdot \nabla_h) \bar{u}_3 = 0, \tag{5.84}$$

$$[\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)] \bar{u}_3 + \frac{\mu_{\text{lim}}}{1+\mu_{\text{lim}}^2} (\bar{\mathbf{b}}_h \cdot \nabla_h) \{ (1 + \mu_{\text{lim}}^2) \bar{r} - \mu_{\text{lim}}^2 \text{av } r^{0,S} \} = 0, \tag{5.85}$$

for the dependent variables $(1 + \mu_{\text{lim}}^2) \bar{r} - \mu_{\text{lim}}^2 \text{av } r^{0,S}$ and \bar{u}_3 , which has the same form as the system (5.71), (5.79) for the dependent variables $r^S - \mu \mathbf{a}_z b_3$ and u_3^S , except that the terms on the right sides are omitted. Since the evolution equation for r shows that $\text{av } r = \text{av } r^0 = \text{av } r^{0,S} + \varepsilon_M \text{av } r^{0,I} = \text{av } r^{0,S}$,

$$\begin{aligned}
r^S - \mu \mathbf{a}_z b_3^F &= (1 + \mu^2) r^S - \mu^2 \text{av } r - \mu \mathbf{a}_z (b_3^F + \mu (r - \text{av } r)) \\
&= [(1 + \mu^2) r^S - \mu^2 \text{av } r^{0,S}] - \mu \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r).
\end{aligned} \tag{5.86}$$

Hence the form of the initial data from (3.30), (3.36) shows that the difference between the initial data for the two systems is bounded in H^n by a constant times $\varepsilon_M^{1+2\nu} + |\mu - \mu_{\text{lim}}|$. In view of that bound plus the estimates (5.73), (5.82), and (5.83) for the right sides of (5.71), (5.79), Theorem B.1 shows that

$$\begin{aligned}
&\| (r^S - \mu \mathbf{a}_z b_3^F) - \{ (1 + \mu_{\text{lim}}^2) \bar{r} - \mu_{\text{lim}}^2 \text{av } r^{0,S} \} \|_{n-2} + \| u_3^S - \bar{u}_3 \|_{n-2} \\
&\leq c [\varepsilon_M + |\mu - \mu_{\text{lim}}|].
\end{aligned} \tag{5.87}$$

By (5.86), the static estimate (5.4) with $j = n - 3$ applied to the $-\mu \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r)$ term of (5.86) shows that (5.87) implies that (5.69) holds. \square \square

Remark 5.6. *As discussed in the introduction, the term $|\mu - \mu_{\text{lim}}|$ is the dominating error term in (5.87) and (5.69) whenever $\mu_{\text{lim}} = 0$, but that term will be eliminated in Theorem 5.7 below by adding corrector terms. The only term yielding an error of size $O(\varepsilon_M^{1-\max(n-5,0)\nu})$ in (5.69) is $-\mu \mathbf{a}_z (b_3^F + \mu \Delta^{-1} \Delta_h r)$ in (5.86), so that expression can be viewed as a next-order corrector. Similarly, the only terms that yield an error of size $O(\varepsilon_M)$ rather than $O(\varepsilon_M^{1+\nu})$ in (5.46) and (5.87) are the expressions $-\varepsilon_M r^S (\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)) \mathbf{u}_h^S$ in (5.65), $(\bar{\mathbf{u}}_h - \mathbf{u}_h^S) \cdot \nabla_h r^S$ in (5.71) in view of estimate (5.46), and $-\varepsilon_M r^S [\partial_t + (\mathbf{u}_h^S \cdot \nabla_h)] u_3^S$ in (5.81).*

Theorem 5.7. *Assume that the conditions of Theorem 5.4 hold, and let $(r^{(\text{cor})}, u_3^{(\text{cor})})$ be the solution of the inhomogeneous linear system*

$$\begin{aligned}
\partial_t r^{(\text{cor})} + (\bar{\mathbf{u}}_h \cdot \nabla_h) r^{(\text{cor})} + \mu (\bar{\mathbf{b}}_h \cdot \nabla_h) u_3^{(\text{cor})} \\
= -(\bar{\mathbf{b}}_h \cdot \nabla_h) \bar{u}_3 - (\mu + \mu_{\text{lim}}) (\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)) \bar{r},
\end{aligned} \tag{5.88}$$

$$\begin{aligned}
\partial_t u_3^{(\text{cor})} + (\bar{\mathbf{u}}_h \cdot \nabla_h) u_3^{(\text{cor})} + \frac{\mu}{1+\mu^2} (\bar{\mathbf{b}}_h \cdot \nabla_h) r^{(\text{cor})} \\
= -\frac{1-\mu\mu_{\text{lim}}}{(1+\mu^2)(1+\mu_{\text{lim}}^2)} (\bar{\mathbf{b}}_h \cdot \nabla_h) ((1 + \mu^2) \bar{r} - \mu^2 \text{av } r^{0,S})
\end{aligned} \tag{5.89}$$

having initial data zero. Then

$$\begin{aligned} & \|r^S - (\bar{r} + \frac{\mu - \mu_{\text{lim}}}{1 + \mu^2} r^{(\text{cor})})\|_{n-2} + \|u_3^S - (\bar{u}_3 + (\mu - \mu_{\text{lim}}) u_3^{(\text{cor})})\|_{n-2} \\ & \leq c \mathcal{E}_M^{1 - \max(n-5, 0)\nu}. \end{aligned} \quad (5.90)$$

Proof. Since $(\mu - \mu_{\text{lim}})(\mu + \mu_{\text{lim}}) = \mu^2 - \mu_{\text{lim}}^2$ and $\text{av } r^{0,S}$ is a constant, adding $\mu - \mu_{\text{lim}}$ times (5.88) to (5.84) yields

$$\begin{aligned} & [\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)] \{ (1 + \mu^2) \bar{r} - \mu^2 \text{av } r^{0,S} + (\mu - \mu_{\text{lim}}) r^{(\text{cor})} \} \\ & + \mu (\bar{\mathbf{b}}_h \cdot \nabla_h) (\bar{u}_3 + (\mu - \mu_{\text{lim}}) u_3^{(\text{cor})}) = 0, \end{aligned} \quad (5.91)$$

Similarly, since $(\mu - \mu_{\text{lim}})$ times $\frac{1 - \mu \mu_{\text{lim}}}{(1 + \mu^2)(1 + \mu_{\text{lim}}^2)}$ equals $\frac{\mu}{1 + \mu^2} - \frac{\mu_{\text{lim}}}{1 + \mu_{\text{lim}}^2}$, adding $\mu - \mu_{\text{lim}}$ times (5.89) to (5.85) yields

$$\begin{aligned} & [\partial_t + (\bar{\mathbf{u}}_h \cdot \nabla_h)] \{ \bar{u}_3 + (\mu - \mu_{\text{lim}}) u_3^{(\text{cor})} \} \\ & + \frac{\mu}{1 + \mu^2} (\bar{\mathbf{b}}_h \cdot \nabla_h) \{ (1 + \mu^2) \bar{r} - \mu^2 \text{av } r^{0,S} + (\mu - \mu_{\text{lim}}) r^{(\text{cor})} \} = 0. \end{aligned} \quad (5.92)$$

Equations (5.91)–(5.92) have the same form as as the system (5.71), (5.79) for the dependent variables $r^S - \mu \alpha_z b_3$ and u_3^S , except that the terms on the right sides are omitted and all occurrences of μ_{lim} on the left sides are replaced by μ . Omitting the step of replacing μ by μ_{lim} in the derivation of (5.71), (5.79) yields those equations with all occurrences of μ_{lim} on the left sides replaced by μ and the terms of order $\mu - \mu_{\text{lim}}$ omitted from their right sides. Since the terms of order μ^2 in (5.91) now involve μ^2 rather than μ_{lim}^2 , as in (5.86) and in contrast to (5.84), there is also no longer a term of size $O(|\mu - \mu_{\text{lim}}|)$ in the difference in the initial data. Hence applying Theorem B.1 now yields an estimate without the term involving $|\mu - \mu_{\text{lim}}|$, and by using (5.86) the estimate so obtained can be written as (5.90). \square \square

Acknowledgments. Ju is supported by the NSFC (Grants No.11571046, 11471028, 11671225). Cheng and Ju are supported by the UK Royal Society ‘‘International Exchanges’’ scheme (Award No. IE150886). Ju and Schochet are supported by the ISF-NSFC joint research program (NSFC Grant No. 11761141008 and ISF Grant No. 2519/17). Cheng is supported by the Leverhulme Trust (Award No. RPG-2017-098) and the EPSRC (Grant No. EP/R029628/1).

APPENDIX A. DERIVATION OF THE MHD SYSTEM (1.4)

Suitably scaled, the motion of an isentropic compressible, conducting, inviscid fluid is modeled by the MHD system ([Dav01, §3.8])

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (\text{A.1a})$$

$$\partial_t (\rho \mathbf{u}) + \mathbf{u} \cdot \nabla (\rho \mathbf{u}) + (\nabla \cdot \mathbf{u}) \rho \mathbf{u} + \varepsilon_M^{-2} \nabla p(\rho) + \varepsilon_A^{-2} \mathbf{B} \times (\nabla \times \mathbf{B}) = 0, \quad (\text{A.1b})$$

$$\partial_t \mathbf{B} - \nabla \times (\mathbf{u} \times \mathbf{B}) = 0, \quad \nabla \cdot \mathbf{B} = 0. \quad (\text{A.1c})$$

Here ε_M denotes the well-known Mach number, ρ is the fluid density, $p(\rho)$ is the pressure law that satisfies $p' > 0$, \mathbf{u} is the fluid velocity, and \mathbf{B} is the magnetic field. The parameter ε_A , as we call the Alfvén number in this article, is the ratio between flow velocity and speed of the magneto-sonic waves; in [KM81] the Alfvén number is the reciprocal of our version.

We consider the case in which a uniform magnetic field is applied in the direction \mathbf{e}_z parallel to the z -axis, which subjects the fluid to a large Lorenz force. To reformulate the system (A.1) into a form to which the results of [CJS18] can be applied, we begin by rescaling the magnetic field and the density via

$$\mathbf{B} = \mathbf{e}_z + \varepsilon_A \mathbf{b}, \quad \rho = 1 + \varepsilon_M r. \quad (\text{A.2})$$

Applying the well-known calculus identities (e.g., [JLL10, p. 372]) for the curl

$$\begin{aligned}\mathbf{b} \times (\nabla \times \mathbf{b}) &= \nabla \frac{|\mathbf{b}|^2}{2} - \mathbf{b} \cdot \nabla \mathbf{b}, \\ \nabla \times (\mathbf{b} \times \mathbf{u}) &= \mathbf{u} \cdot \nabla \mathbf{b} + (\nabla \cdot \mathbf{u}) \mathbf{b} - \mathbf{b} \cdot \nabla \mathbf{u}, \quad (\text{since } \nabla \cdot \mathbf{b} = 0), \\ \nabla \times (\mathbf{e}_z \times \mathbf{u}) &= \mathbf{e}_z \nabla \cdot \mathbf{u} - \partial_z \mathbf{u},\end{aligned}$$

and the similar identity

$$\mathbf{e}_z \times (\nabla \times \mathbf{b}) = \nabla b_3 - \partial_z \mathbf{b},$$

subtracting \mathbf{u} times (A.1a) from (A.1b), and multiplying (A.1a) by $a(\varepsilon_M r)$ from (1.6) yields the system (1.4).

Taking the divergence of (A.1c) and using the fact that the divergence of a curl vanishes shows that $\partial_t \nabla \cdot \mathbf{b} = 0$. Similarly, the spatial integrals $\iiint b_3 dx dy dz$ and $\iiint r dx dy dz$ are conserved by the dynamics

APPENDIX B. PERTURBATION THEOREM

The following perturbation theorem is a variant of [Che14, Lemma 3.2]).

Theorem B.1. *Suppose that u and U are solutions in $C^0([0, T]; H^k)$ of*

$$A_0(u)u_t + \sum_{i=1}^d A_i(u)u_{x_i} = F + L^*v + A_0(u)\partial_t \Xi_u + \xi_u, \quad Lu = 0, \quad (\text{B.1})$$

$$A_0(U)U_t + \sum_{i=1}^d A_i(U)U_{x_i} = F + L^*V + A_0(U)\partial_t \Xi_U + \xi_U, \quad LU = 0, \quad (\text{B.2})$$

having the same initial value $u_0 \in H^k$, where $k \geq \lfloor \frac{d}{2} \rfloor + 2$, the matrices A_i are smooth and symmetric and A_0 is positive-definite, F is a given function of t and x , L is a first-order differential operator with constant coefficients, with L^* denoting its L^2 -adjoint, and Ξ_u, Ξ_U, ξ_u , and ξ_U satisfy

$$\begin{aligned}\|\Xi_u\|_{k-r} + \|\Xi_U\|_{k-r} + \|\xi_u\|_{k-r-1} + \|\xi_U\|_{k-r-1} &\leq c\delta \quad \text{for some } 0 \leq r \leq k-1, \\ L\Xi_u = 0 = L\Xi_U,\end{aligned}$$

and

$$\|\partial_t \Xi_U\|_{k-r-1} \leq c.$$

Then $\max_{0 \leq t \leq T} \|u - U\|_{k-r-1} \leq c\delta$.

Proof. Subtracting (B.2) from (B.1) and rearranging terms yields the system

$$\begin{aligned}A_0(u)((u - U) - (\Xi_u - \Xi_U))_t + \sum_{i=1}^d A_i(u)((u - U) - (\Xi_u - \Xi_U))_{x_i} \\ = L^*(v - V) + (\xi_u - \xi_U) - \sum_i A_i(u)(\Xi_u - \Xi_U)_{x_i} \\ + [A_0(u) - A_0(U)](U - \Xi_U)_t + \sum_{i=1}^d [A_i(u) - A_i(U)]U_{x_i}\end{aligned} \quad (\text{B.3})$$

Since the A_i are smooth and u and U belong to H^k , $\|A_i(u) - A_i(U)\|_{k-r-1} = O(\|u - U\|_{k-r-1}) = O(\|(u - U) - (\Xi_u - \Xi_U)\|_{k-r-1} + \|(\Xi_u, \Xi_U)\|_{k-r-1})$. Moreover, multiplying (B.2) by U_t and integrating over the spatial

variables shows that

$$\begin{aligned}
\|U_t\|_{L^2}^2 &\leq c \int U_t A_0(U) U_t dx = - \sum_i \int U_t A_i(U) U_{x_i} dx + \int U_t L^* V dx + \int U_t (A_0(U) \partial_t \Xi_U + \xi_U) \\
&= - \sum_i \int U_t A_i(U) U_{x_i} dx + \int \partial_t(LU) \cdot V dx + \int U_t (A_0(U) \partial_t \Xi_U + \xi_U) \\
&= - \sum_i \int U_t A_i(U) U_{x_i} dx + \int U_t (A_0(U) \partial_t \Xi_U + \xi_U) \\
&\leq c \|U_t\|_{L^2} (\|U\|_1 + \|(\partial_t \Xi_U, \xi_U)\|_{L^2}),
\end{aligned}$$

which shows that $\|U_t\|_{L^2} \leq c(\|U\|_1 + 1)$ without the need for any hypothesis on the size of V . Applying a spatial derivative D^α to (B.2), multiplying the result by $D^\alpha U_t$, integrating over space, estimating $\|D^\alpha U_t\|_{L^2}$ in similar fashion, and summing over $0 \leq |\alpha| \leq k-r-1$ yields the estimate $\|U_t\|_{k-r-1} \leq C(\|U\|_{k-r} + 1)$. Since the assumption on the size of k ensures that $\|fg\|_s \leq \|f\|_s \|g\|_{k-1}$ for $s \leq k-1$, this estimate for U_t together with the assumed bounds for u , U , Ξ , and ξ show that the right side of (B.3) equals $L^*(v-V)$ plus terms whose H^{k-r-1} norm is $O(\|(u-U) - (\Xi_u - \Xi_U)\|_{r-k-1} + \delta)$. Since $L((u-U) - (\Xi_u - \Xi_U)) = 0$, and L is a constant-coefficient differential operator, the term $L^*(v-V)$ drops out of the H^{k-r-1} energy estimate. Moreover Ξ_u , Ξ_U , ξ_U , and ξ_U are $O(\delta)$, so the remaining terms on the right side are $O(\|(u-U) - (\Xi_u - \Xi_U)\| + \delta)$. The standard H^{k-r-1} energy estimate for symmetric-hyperbolic systems together with Gronwall's inequality therefore shows that $\|(u-U) - (\Xi_u - \Xi_U)\|_{k-r-1} \leq c\delta$. Using once more the fact that $\|(\Xi_u, \Xi_U)\|_{k-r-1}$ is $O(\delta)$ yields the claimed estimate. \square \square

APPENDIX C. CALCULUS INEQUALITIES FOR z -AVERAGES

The following result is sharper than what would be obtained by the standard product estimate (e.g., [Maj84, Proposition 2.1A]), because the entire product is estimate using the $W^{1,1}$ norm rather than pulling out one factor in the L^∞ norm, and the Gagliardo-Nirenberg inequalities are used in dimension two rather than three.

Lemma C.1. *For all $j \geq 1$ there exists a constant C_j such that for all $f, g \in H^j(\mathbb{T}^3)$*

$$\|\mathbf{a}_z(fg)\|_{H^{j-1}(\mathbb{T}^2)} \leq C_j \left(\|f\|_{H^j(\mathbb{T}^3)} \|g\|_{L^2(\mathbb{T}^3)} + \|f\|_{L^2(\mathbb{T}^3)} \|g\|_{H^j(\mathbb{T}^3)} \right). \quad (\text{C.1})$$

Proof. We first prove (C.1) for $j = 1$: By the Gagliardo-Nirenberg inequality $\|h\|_{L^2(\mathbb{T}^2)} \leq c\|h\|_{W^{1,1}(\mathbb{T}^2)}$, the standard estimate for integrals $|\int fh| \leq \int |h|$, and the Cauchy-Schwartz inequality,

$$\begin{aligned}
\|\mathbf{a}_z(fg)\|_{L^2(\mathbb{T}^2)} &\leq c\|\mathbf{a}_z(fg)\|_{W^{1,1}(\mathbb{T}^2)} = c \left(\|\mathbf{a}_z(fg)\|_{L^1(\mathbb{T}^2)} + \|\mathbf{a}_z(f\nabla_{x,y}g)\|_{L^1(\mathbb{T}^2)} + \|\mathbf{a}_z(g\nabla_{x,y}f)\|_{L^1(\mathbb{T}^2)} \right) \\
&\leq c \left(\|fg\|_{L^1(\mathbb{T}^3)} + \|f\nabla g\|_{L^1(\mathbb{T}^3)} + \|g\nabla f\|_{L^1(\mathbb{T}^3)} \right) \\
&\leq c \left(\|f\|_{L^2(\mathbb{T}^3)} \|g\|_{L^2(\mathbb{T}^3)} + \|f\|_{L^2(\mathbb{T}^3)} \|\nabla g\|_{L^2(\mathbb{T}^3)} + \|g\|_{L^2(\mathbb{T}^3)} \|\nabla f\|_{L^2(\mathbb{T}^3)} \right) \\
&\leq c \left(\|f\|_{H^1(\mathbb{T}^3)} \|g\|_{L^2(\mathbb{T}^3)} + \|f\|_{L^2(\mathbb{T}^3)} \|g\|_{H^1(\mathbb{T}^3)} \right).
\end{aligned}$$

Now let j be any integer greater than one. By the definition of the H^{j-1} norm, the result for the case $j = 1$, the Sobolev interpolation inequality (2.6), and Young's inequality for products $a^\sigma b^{1-\sigma} \leq a + b$ for $0 \leq \sigma \leq 1$,

$$\begin{aligned}
\|\mathbf{a}_z(fg)\|_{H^{j-1}(\mathbb{T}^2)} &\leq c \sum_{|\alpha| \leq j-1} \|\mathbf{a}_z(D_{x,y}^\alpha(fg))\|_{L^2(\mathbb{T}^2)} \leq c \sum_{|\beta|+|\gamma| \leq j-1} \|\mathbf{a}_z((D_{x,y}^\beta f)(D_{x,y}^\gamma g))\|_{L^2(\mathbb{T}^2)} \\
&\leq c \sum_{|\beta|+|\gamma| \leq j-1} \left(\|f\|_{H^{|\beta|+1}(\mathbb{T}^3)} \|g\|_{H^{|\gamma|}(\mathbb{T}^3)} + \|f\|_{H^{|\beta|}(\mathbb{T}^3)} \|g\|_{H^{|\gamma|+1}(\mathbb{T}^3)} \right) \\
&\leq c \sum_{|\beta|+|\gamma| \leq j-1} \left(\|f\|_{H^{|\beta|+|\gamma|+1}(\mathbb{T}^3)}^{|\beta|} \|f\|_{H^1(\mathbb{T}^3)}^{1-|\beta|} \|g\|_{H^{|\beta|+|\gamma|}(\mathbb{T}^3)}^{|\gamma|} \|g\|_{L^2(\mathbb{T}^3)}^{1-|\gamma|} \right. \\
&\quad \left. + \|f\|_{H^{|\beta|+|\gamma|}(\mathbb{T}^3)}^{|\beta|} \|f\|_{L^2(\mathbb{T}^3)}^{1-|\beta|} \|g\|_{H^{|\beta|+|\gamma|+1}(\mathbb{T}^3)}^{|\gamma|} \|g\|_{H^1(\mathbb{T}^3)}^{1-|\gamma|} \right) \\
&= c \sum_{|\beta|+|\gamma| \leq j-1} \left(\left(\frac{\|f\|_{H^{|\beta|+|\gamma|+1}(\mathbb{T}^3)}}{\|f\|_{H^1(\mathbb{T}^3)}} \right)^{|\beta|} \|f\|_{H^1(\mathbb{T}^3)} \left(\frac{\|g\|_{H^{|\beta|+|\gamma|}(\mathbb{T}^3)}}{\|g\|_{L^2(\mathbb{T}^3)}} \right)^{|\gamma|} \|g\|_{L^2(\mathbb{T}^3)} \right. \\
&\quad \left. + \left(\frac{\|f\|_{H^{|\beta|+|\gamma|}(\mathbb{T}^3)}}{\|f\|_{L^2(\mathbb{T}^3)}} \right)^{|\beta|} \|f\|_{L^2(\mathbb{T}^3)} \left(\frac{\|g\|_{H^{|\beta|+|\gamma|+1}(\mathbb{T}^3)}}{\|g\|_{H^1(\mathbb{T}^3)}} \right)^{|\gamma|} \|g\|_{H^1(\mathbb{T}^3)} \right) \\
&\leq c \sum_{0 \leq i \leq j-1} (\|f\|_{H^{i+1}} \|g\|_{L^2} + \|g\|_{H^i} \|f\|_{H^1}) + (\|f\|_{H^i} \|g\|_{H^1} + \|g\|_{H^{i+1}} \|f\|_{L^2}) \\
&\leq c (\|f\|_{H^j} \|g\|_{L^2} + \|g\|_{H^{j-1}} \|f\|_{H^1} + \|f\|_{H^{j-1}} \|g\|_{H^1} + \|g\|_{H^j} \|f\|_{L^2})
\end{aligned} \tag{C.2}$$

A second application of the Sobolev interpolation inequality followed by Young's inequality shows that each of the terms in the final line of (C.2) in which the H^{j-1} and H^1 norms appear is bounded by the sum of the two terms there in which the H^j and L^2 norms appear, which yields (C.1) for $j > 1$. \square \square

Corollary C.2. Consider integer $n \geq 3$ and a geometric sequence $\{\varepsilon_j\}$ with common ratio $\frac{1}{\mu} \geq 1$ and $\varepsilon_n \leq c$. Suppose $v(x, y, z), w(x, y, z) \in H^n(\mathbb{T}^3)$ satisfy the ‘‘interpolative estimates’’

$$\|(v, w)\|_{H^j(\mathbb{T}^3)} \leq c \varepsilon_j, \quad j = 0, \dots, n-1, \quad \|(v, w)\|_{H^n(\mathbb{T}^3)} \leq c.$$

Then

$$\|\mathbf{a}_z(vw)\|_{H^{n-1}(\mathbb{T}_{x,y}^2)} \leq c \varepsilon_0, \quad \|\mathbf{a}_z(vw)\|_{H^{n-2}(\mathbb{T}_{x,y}^2)} \leq c \varepsilon_0 \mu. \tag{C.3}$$

Proof. By Lemma C.1, $\|\mathbf{a}_z(vw)\|_{j-1} \leq c(\|v\|_j \|w\|_0 + \|v\|_0 \|w\|_j) \leq c \varepsilon_j \varepsilon_0$. Since $\varepsilon_n \leq c$ and $\varepsilon_{n-1} \leq \mu \varepsilon_n \leq c \mu$, this implies (C.3). \square \square

Immediately, by the uniform H^n estimate (1.12) and the static estimates (5.4), (5.6) together with the relations (1.10), (1.14) between the parameters, we have the following estimates on products amongst components of various modes.

Corollary C.3. Assume that the conditions of Theorem 1.1 hold. Let v^F, w^F be either $(1 - \mathbf{a}_z)b_3^F$ or any component of \mathbf{V}^F except b_3^F and let v^I, w^I be any component of \mathbf{V}^I . Then,

$$\sup_{0 \leq t \leq T} \left\{ \|\mathbf{a}_z(v^F w^F)\|_{n-1} + \mu \|\mathbf{a}_z(v^I v^F)\|_{n-1} + \mu^2 \|\mathbf{a}_z(v^I w^I)\|_{n-1} + \|\mathbf{a}_z \nabla \cdot (v^I u_n^F)\|_{n-2} + \varepsilon_M \|v^F\|_{n-1} \right\} \leq c \varepsilon_A,$$

and the estimates also hold when μ or μ^2 on the left side is replaced by ε_M .

The estimate on $\varepsilon_M \|v^F\|_{n-1}$ is independent of Lemma C.1, and is included here for convenience.

REFERENCES

- [BK82] G. Browning and H.-O. Kreiss. Problems with different time scales for nonlinear partial differential equations. *SIAM J. Appl. Math.*, 42(4):704–718, 1982.
- [Che12] B. Cheng. Singular limits and convergence rates of compressible Euler and rotating shallow water equations. *SIAM J. Math. Anal.*, 44(2):1050–1076, 2012.
- [Che14] B. Cheng. Improved accuracy of incompressible approximation of compressible Euler equations. *SIAM J. Math. Anal.*, 46(6):3838–3864, 2014.
- [CJS18] B. Cheng, Q. Ju, and S. Schochet. Three-scale singular limits of evolutionary PDEs. *Arch. Ration. Mech. Anal.*, 229(2):601–625, 2018.
- [CM13] B. Cheng and A. Mahalov. Euler equation on a fast rotating sphere—time-averages and zonal flows. *Eur. J. Mech. B Fluids*, 37:48–58, 2013.
- [CM19] B. Cheng and A. Mahalov. General Results on Zonation in Rotating Systems with a β -Effect and the Electromagnetic Force. In *Zonal Jets: Phenomenology, Genesis, Physics*. Boris Galperin and Peter L. Read eds. Cambridge University Press, 2019.
- [Dav01] P. A. Davidson. *An introduction to magnetohydrodynamics*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2001.
- [JL10] S. Jiang, Q. Ju, and F. Li. Incompressible limit of the compressible magnetohydrodynamic equations with periodic boundary conditions. *Comm. Math. Phys.*, 297(2):371–400, 2010.
- [Kat82] T. Kato. *A short introduction to perturbation theory for linear operators*. Springer-Verlag, New York-Berlin, 1982.
- [KM81] S. Klainerman and A. Majda. Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids. *Comm. Pure Appl. Math.*, 34(4):481–524, 1981.
- [KM82] S. Klainerman and A. Majda. Compressible and incompressible fluids. *Comm. Pure Appl. Math.*, 35(5):629–651, 1982.
- [KLN91] H.-O. Kreiss, J. Lorenz, and M. J. Naughton. Convergence of the solutions of the compressible to the solutions of the incompressible Navier-Stokes equations. *Advances in Applied Math.*, 12:187–214, 1991.
- [Maj84] A. Majda. *Compressible fluid flow and systems of conservation laws in several space variables*, volume 53 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1984.
- [MS01] G. Métivier and S. Schochet. The incompressible limit of the non-isentropic Euler equations. *Arch. Ration. Mech. Anal.*, 158:61–90, 2001.
- [Sch88] S. Schochet. Asymptotics for symmetric hyperbolic systems with a large parameter. *J. Differential Equations*, 75(1):1–27, 1988.
- [Sch94] S. Schochet. Fast singular limits of hyperbolic PDEs. *J. Differential Equations*, 114(2):476–512, 1994.