Incompressible Limit of Compressible Ideal MHD Flows inside a Perfectly Conducting Wall

Jiawei WANG*   Junyan ZHANG†

January 24, 2024

Abstract

We prove the incompressible limit of compressible ideal magnetohydrodynamic(MHD) flows in a reference domain where the magnetic field is tangential to the boundary. Unlike the case of transversal magnetic fields, the linearized problem of our case is not well-posed in standard Sobolev space H^m ($m \geq 2$), while the incompressible problem is still well-posed in H^m. The key observation to overcome the difficulty is a hidden structure contributed by Lorentz force in the vorticity analysis, which reveals that one should trade one normal derivative for two tangential derivatives together with a gain of Mach number weight ε^2. Thus, the energy functional should be defined by using suitable anisotropic Sobolev spaces. The weights of Mach number should be carefully chosen according to the number of tangential derivatives, such that the energy estimates are uniform in Mach number. Besides, part of the proof is similar to the study of compressible water waves, so our result opens the possibility to study the incompressible limit of free-boundary problems in ideal MHD.

Contents

1   Introduction

In this paper, we consider the compressible ideal magnetohydrodynamics(MHD) equations

$$
\begin{cases}
D_t \rho + \rho(\nabla \cdot u) = 0 & \text{in } [0,T] \times \Omega, \\
\rho D_t u = B \cdot \nabla B - \nabla P, \quad P := p + \frac{1}{2}|B|^2 & \text{in } [0,T] \times \Omega, \\
D_t B = B \cdot \nabla u - B(\nabla \cdot u) & \text{in } [0,T] \times \Omega, \\
\nabla \cdot B = 0 & \text{in } [0,T] \times \Omega, \\
D_t S = 0 & \text{in } [0,T] \times \Omega,
\end{cases}
\tag{1.1}
$$

---
*Institute of Applied Physics and Computational Mathematics, Beijing, P.R. China. Email: `wangjiawei19@gscaep.ac.cn`
†Department of Mathematics, National University of Singapore, Singapore. Email: `zhangjy@nus.edu.sg`

describing the motion of a compressible conducting fluid in an electro-magnetic field. Here $\Omega := \mathbb{T}^{d-1} \times (-1, 1)$ is the reference domain in $\mathbb{R}^d$ ($d = 2, 3$) with boundary $\Sigma := \Sigma_+ \cup \Sigma_-$ and $\Sigma_\pm := \{x_d = \pm 1\}$. $\nabla := (\partial_{x_1}, \cdots, \partial_{x_d})$ is the standard spatial derivative. $D_t := \partial_t + u \cdot \nabla$ is the material derivative. The fluid velocity, the magnetic field, the fluid density, the fluid pressure and the entropy are denoted by $u = (u_1, \cdots, u_d)$, $B = (B_1, \cdots, B_d)$, ρ, p and S respectively. Note that the last equation of (1.1) is derived from the equation of total energy and Gibbs relation. See [6, Ch. 4.3] for more details. Throughout this paper, we will use follow the convention of Einstein's summation, that is, repeated indices represent taking sum over the indices.

We assume the fluid pressure $p = p(\rho, S)$ to be a given smooth function of ρ, S which satisfies

$$\rho > 0, \quad \frac{\partial p}{\partial \rho} > 0, \quad \text{in } \bar{\Omega}. \tag{1.2}$$

These two conditions also guarantee the hyperbolicity of system (1.1).

The initial and boundary conditions of system (1.1) are

$$(u, B, \rho, S)|_{t=0} = (u_0, B_0, \rho_0, S_0) \quad \text{in } [0, T] \times \Omega, \tag{1.3}$$

$$u_d = B_d = 0 \quad \text{on } [0, T] \times \Sigma, \tag{1.4}$$

where the boundary condition for u_d is the slip boundary condition, and the boundary condition for B_d shows that $\Sigma_\pm$ are perfectly conducting walls.

Remark 1.1. The conditions $\nabla \cdot B = 0$ in Ω and $B_d = 0$ on Σ are both constraints for initial data so that the MHD system is not over-determined. One can show that they propagate within the lifespan of the solution. Using the theory of hyperbolic system with charateristic boundary condition [22], one can show that the correct number of boundary conditions when assuming $u_d|_\Sigma = 0$ is 1 (see [31]). So, $B_d|_\Sigma = 0$ has to be an initial constraint.

The initial-boundary value problem (1.1)-(1.4) is used to characterize the motion of a plasma confined in a rigid perfectly conducting wall, which is an important model in the study of laboratory plasma confinement problems. See [6, Ch. 4.6] for more details. To make the initial-boundary value problem (1.1)-(1.4) solvable, we need to require the initial data satisfying the compatibility conditions up to certain order. For $m \in \mathbb{N}$, we define the m-th order compatibility conditions to be

$$\partial_t^j u_d|_{t=0} = 0, \quad 0 \le j \le m. \tag{1.5}$$

Let $\mathcal{F} := \log \rho$. Since $\frac{\partial p}{\partial \rho} > 0$ indicates $\frac{\partial \mathcal{F}}{\partial p} > 0$, using $D_t S = 0$, the first equation of (1.1) is equivalent to

$$\frac{\partial \mathcal{F}}{\partial p} D_t p + \nabla \cdot u = 0. \tag{1.6}$$

Thus the compressible MHD system is now reformulated as follows

$$\begin{cases} \frac{\partial \mathcal{F}}{\partial p} D_t p + \nabla \cdot u = 0 & \text{in } [0, T] \times \Omega, \\ \rho D_t u = B \cdot \nabla B - \nabla P, \quad P := p + \frac{1}{2}|B|^2 & \text{in } [0, T] \times \Omega, \\ D_t B = B \cdot \nabla u - B(\nabla \cdot u) & \text{in } [0, T] \times \Omega, \\ \nabla \cdot B = 0 & \text{in } [0, T] \times \Omega, \\ D_t S = 0 & \text{in } [0, T] \times \Omega, \\ p = p(\rho, S), \frac{\partial p}{\partial \rho} > 0 & \text{in } [0, T] \times \bar{\Omega}, \\ u_d = B_d = 0 & \text{on } [0, T] \times \Sigma, \\ (u, B, \rho, S)|_{t=0} = (u_0, B_0, \rho_0, S_0) & \text{on } \{t = 0\} \times \Omega. \end{cases} \tag{1.7}$$

1.1 The equation of states and sound speed

This paper is devoted to studying the behavior of the solution of (1.7) as the sound speed goes to infinity, which is known to be the incompressible limit. Physically, the sound speed is defined by $c_s := \sqrt{\frac{\partial p}{\partial \rho}}$. Mathematically, Mathematically, it is convenient to view the sound speed as a parameter λ. This being said, a

typical choice of the equation of states $p_\lambda(\rho, S)$ would be the polytropic gas

$$p_\lambda(\rho, S) = \lambda^2 \left(\rho^\gamma \exp(S/C_V) - 1 \right), \quad \gamma \geq 1, \quad C_V > 0. \tag{1.8}$$

When viewing the density as a function of the pressure and the entropy, this indicates

$$\rho_\lambda(p/\lambda^2, S) = \left(\left(1 + \frac{p}{\lambda^2}\right) e^{-\frac{S}{C_V}} \right)^{\frac{1}{\gamma}}, \quad \text{and} \quad \log\left(\rho_\lambda(p/\lambda^2, S) \right) = \gamma^{-1} \log\left(\left(1 + \frac{p}{\lambda^2}\right) e^{-\frac{S}{C_V}} \right). \tag{1.9}$$

Hence, we can view $\mathcal{F} = \log \rho$ as a parametrized family $\{\mathcal{F}_\varepsilon(p, S)\}$ as well, where $\varepsilon = \frac{1}{\lambda}$. Indeed, we have

$$\frac{\partial \mathcal{F}_\varepsilon}{\partial p} = \gamma^{-1} \log\left((1 + \varepsilon^2 p) e^{-\frac{S}{C_V}} \right). \tag{1.10}$$

Since we work on the case when the entropy and velocity are both bounded (later we will assume $u, S \in H^4(\Omega)$), we again slightly abuse the terminology, and call λ the sound speed and call ε the Mach number and thus $M = O(\varepsilon)$. Furthermore, there exists $C > 0$ such that

$$C^{-1} \varepsilon^2 \leq \frac{\partial \mathcal{F}_\varepsilon}{\partial p}(p, S) = \frac{1}{\rho} \frac{\partial \rho}{\partial p}(p, S) \leq C \varepsilon^2. \tag{1.11}$$

Also, we assume

$$|\partial_p^s \mathcal{F}_\varepsilon(p, S)| \leq C, \quad |\partial_p^s \mathcal{F}_\varepsilon(p, S)| \leq C |\partial_p \mathcal{F}_\varepsilon(p, S)|^s \leq C \mathcal{F}_\varepsilon'(p) \tag{1.12}$$

holds for $0 \leq s \leq 8$. Thus, when considering the incompressible limit, that is, when $\varepsilon > 0$ is sufficiently small, it is more convenient to reformulate the compressible MHD system by replacing $\frac{\partial \mathcal{F}}{\partial p}$ with ε^2 as follows

$$\begin{cases} \varepsilon^2 D_t p + \nabla \cdot u = 0 & \text{in } [0, T] \times \Omega, \\ \rho D_t u = B \cdot \nabla B - \nabla P, \quad P := p + \frac{1}{2}|B|^2 & \text{in } [0, T] \times \Omega, \\ D_t B = B \cdot \nabla u - B(\nabla \cdot u) & \text{in } [0, T] \times \Omega, \\ \nabla \cdot B = 0 & \text{in } [0, T] \times \Omega, \\ D_t S = 0 & \text{in } [0, T] \times \Omega, \\ p = p(\rho, S), \frac{\partial p}{\partial \rho} > 0 & \text{in } [0, T] \times \bar{\Omega}, \\ u_d = B_d = 0 & \text{on } [0, T] \times \Sigma, \\ (u, B, \rho, S)|_{t=0} = (u_0, B_0, \rho_0, S_0) & \text{on } \{t = 0\} \times \Omega. \end{cases} \tag{1.13}$$

1.2   An overview of previous results

The incompressible limit of compressible inviscid fluids is considered to be a type of singular limit of hyperbolic system: the pressure for compressible fluids is a variable of hyperbolic system whereas the pressure for incompressible fluids is a Lagrangian multiplier and the equation of state is no longer valid. Early works about compressible Euler equations can be dated back to Klainerman-Majda [14, 15] when the domain is the whole space $\mathbb{R}^d$ or the periodic domain $\mathbb{T}^d$, Ebin [5] and Schochet [23, 24] when the domain is bounded, and Isozaki [9] when considering an exterior domain. The abovementioned papers consider the case of "well-prepared" initial data, which means the compressible initial data is exactly a small perturbation of a given incompressible initial data. When the initial data is "ill-prepared", that is, the compressible initial data is the small perturbation of incompressible initial data plus a highly oscillatory part, we refer to [34, 2, 9, 25, 8, 29, 18, 1] and references therein. The precise definitions of "well-prepared data" and "ill-prepared data" can be found in [29, 18, 1].

For compressible ideal MHD, the incompressible limit in the whole space was studied by Jiang-Ju-Li [10]. However, when the domain has a boundary, the study of incompressible limit for ideal compressible MHD is much more subtle. The first difficulty is that the vorticity estimate for ideal compressible MHD cannot be closed when using standard Sobolev spaces, and thus the method presented in [23, 24] is no longer valid. When the magnetic field is not tangential to the boundary, one can use such transversality to

3

compensate the loss of normal derivative arising in the vorticity analysis. See Yanagisawa [35] for the well-posedness result under the condition $B \times N|_{\partial\Omega} = \mathbf{0}$. As for the singular limits, we refer to Ju-Schochet-Xu [12] for the singular limits when both Mach number and Alfvén number converge to zero and Jiang-Ju-Xu [11] for the low Alfvén number limit of incompressible ideal MHD whose local existence was proved via a low-Mach-number regime.

Unfortunately, when the magnetic field is tangential to the boundary, that is, $B \cdot N|_{\partial\Omega} = 0$, Ohno-Shirota [20] proved that the linearized problem near a non-zero magnetic field is ill-posed in standard Sobolev spaces. In this case, one has to use suitable anisotropic Sobolev spaces introduced by Chen [3] to proved the well-posedness. See Yanagisawa-Matsumura [36] and Secchi [26, 28].

Therefore, an essential difficulty in establishing the incompressible limit for ideal MHD with $B \cdot N|_{\partial\Omega} = 0$ is the "incompatibility of function spaces for the existence": compressible ideal MHD flow with the perfectly conducting wall condition may not be well-posed in standard Sobolev spaces, while the corresponding incompressible problem is well-posed (see, for example, Gu-Wang [7]). So far, there is only a partial answer for this incompressible limit problem. Ju and the first author [13] proved the incompressible limit in a suitable closed subspace of standard Sobolev space, introduced by Secchi [27], by adding more restrictive constraints for the boundary value of initial data. Specifically, [27, 13] require the initial data to satisfy

$$\partial_3^{2k}(u_3, B_3)|_\Sigma = 0, \ \ \partial_3^{2k+1}(p, u_1, u_2, B_1, B_2, S)|_\Sigma = 0, \ \ k = 0, 1, \cdots, \lfloor \frac{m-1}{2} \rfloor.$$

However, the physical interpretation of these "extra constraints" is still unclear. In other words, it is still unknown how to thoroughly overcome the difficulty caused by the abovementioned "incompatibility".

The aim of this paper is to give a definitive answer to the incompressible limit problem of ideal MHD with the perfectly conducting wall condition when the initial data is "well-prepared". Indeed, in the vorticity analysis, there is a hidden structure brought by the Lorentz force term. To the best of our knowledge, such an observation never appears in any previous works, but it can really help us easily prove the uniform estimates in Mach number and also illustrates why the anisotropic Sobolev spaces are naturally introduced and necessary for compressible MHD but are not needed for incompressible MHD.

1.3 The main theorems

Before stating our results, we should first define the anisotropic Sobolev space $H_*^m(\Omega)$ for $m \in \mathbb{N}$ and $\Omega = \mathbb{T}^{d-1} \times [-1, 1]$. Let $\omega = \omega(x_d)$ be a cutoff function[1] on $[-1, 1]$ defined by $\omega(x_d) = (1 - x_d)(1 + x_d)$. Then we define $H_*^m(\Omega)$ for $m \in \mathbb{N}^*$ as follows

$$H_*^m(\Omega) := \left\{ f \in L^2(\Omega) \middle| (\omega \partial_d)^{\alpha_{d+1}} \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} f \in L^2(\Omega), \ \ \forall \alpha \text{ with } \sum_{j=1}^{d-1} \alpha_j + 2\alpha_d + \alpha_{d+1} \leq m \right\},$$

equipped with the norm

$$\|f\|_{H_*^m(\Omega)}^2 := \sum_{\sum_{j=1}^{d-1} \alpha_j + 2\alpha_d + \alpha_{d+1} \leq m} \|(\omega \partial_d)^{\alpha_{d+1}} \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} f\|_{L^2(\Omega)}^2. \tag{1.14}$$

For any multi-index $\alpha := (\alpha_0, \alpha_1, \cdots, \alpha_d, \alpha_{d+1}) \in \mathbb{N}^{d+2}$, we define

$$\partial_*^\alpha := \partial_t^{\alpha_0} (\omega \partial_3)^{\alpha_{d+1}} \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}, \ \ \langle \alpha \rangle := \sum_{j=0}^{d-1} \alpha_j + 2\alpha_d + \alpha_{d+1},$$

and define the **space-time anisotropic Sobolev norm** $\| \cdot \|_{m,*}$ to be

$$\|f\|_{m,*}^2 := \sum_{\langle \alpha \rangle \leq m} \|\partial_*^\alpha f\|_{L^2(\Omega)}^2 = \sum_{\alpha_0 \leq m} \|\partial_t^{\alpha_0} f\|_{H_*^{m-\alpha_0}(\Omega)}^2. \tag{1.15}$$

We also denote the interior Sobolev norm to be $\|f\|_s := \|f(t, \cdot)\|_{H^s(\Omega)}$ for any function $f(t, x)$ on $[0, T] \times \Omega$ and denote the boundary Sobolev norm to be $|f|_s := |f(t, \cdot)|_{H^s(\Sigma)}$ for any function $f(t, x)$ on $[0, T] \times \Sigma$. From

---

[1]The choice of $\omega(x_d)$ is not unique. We just need $\omega(x_d)$ vanishes on Σ and is comparable to the distance function near Σ.

now on, we assume the dimension to be $d = 3$, that is, $\Omega = \mathbb{T}^2 \times (-1, 1)$ and $\Sigma_\pm = \{x_3 = \pm 1\}$. In the proof, we will see the 2D case follow in the same manner as the 3D case up to a slight modification in the vorticity analysis.

First, we establish a local-in-time estimate that is uniform in Mach number ε.

Theorem 1.1 (**Uniform estimate in ε**). Let $\varepsilon \in (0, 1)$ be fixed. Let $(u_0, B_0, \rho_0, S_0) \in H_*^8(\Omega) \times H_*^8(\Omega) \times H_*^8(\Omega) \times H_*^8(\Omega)$ be the initial data of (1.13) satisfying the compatibility conditions (1.5) up to 7-th order and

$$E(0) \le M \tag{1.16}$$

for some $M > 0$ independent of ε. Then there exists $T > 0$ depending only on M, such that (1.13) admits a unique solution $(p(t), u(t), B(t), S(t))$ verifies the energy estimate

$$\sup_{t \in [0,T]} E(t) \le P(E(0)), \tag{1.17}$$

where $P(\cdots)$ is a generic polynomial in its arguments, and the energy $E(t)$ is defined to be

$$E(t) = E_4(t) + E_5(t) + E_6(t) + E_7(t) + E_8(t),$$

$$E_4(t) = \sum_{k=0}^{4} \left\| \left(\varepsilon^{(k-1)_+} \partial_t^k u, \varepsilon^{(k-1)_+} \partial_t^k B, \varepsilon^{(k-1)_+} \partial_t^k S, \varepsilon^k \partial_t^k p \right) \right\|_{4-k}^2$$

$$E_{4+l}(t) = \sum_{\substack{\langle \alpha \rangle = 2l, \\ \alpha_0 < 2l}} \sum_{k=0}^{4-l} \left\| \varepsilon^{(k-1)_+ + 2l} \mathcal{T}^\alpha \partial_t^k (u, B, S, p) \right\|_{4-k-l}^2 \tag{1.18}$$

$$+ \sum_{k=0}^{4-l} \left\| \varepsilon^{(k-1)_+ + 2l} \partial_t^{k+2l} (u, B, S), \varepsilon^{k+2l} \partial_t^{k+2l} p \right\|_{4-k-l}^2, \quad 1 \le l \le 4,$$

where $K_+ := \max\{K, 0\}$ and we denote $\mathcal{T}^\alpha := (\omega(x_3) \partial_3)^{\alpha_4} \partial_t^{\alpha_0} \partial_1^{\alpha_1} \partial_2^{\alpha_2}$ to be a high-order tangential derivative for the multi-index $\alpha = (\alpha_0, \alpha_1, \alpha_2, 0, \alpha_4)$ with length (for the anisotropic Sobolev spaces) $\langle \alpha \rangle = \alpha_0 + \alpha_1 + \alpha_2 + 2 \times 0 + \alpha_4$. In the rest of this paper, we sometimes write $\mathcal{T}^k$ to represent a tangential derivative $\mathcal{T}^\alpha$ with order $\langle \alpha \rangle = k$ when we do not need to specify what the derivative $\mathcal{T}^\alpha$ contains.

Remark 1.2 (**Correction of $E_4(t)$**). We note that the norm $\|p\|_4^2$ in $E_4(t)$ defined by (1.18) should be replaced by $\|\varepsilon p\|_0^2 + \|\nabla p\|_3^2$ because we do not have L^2 estimates of p without ε weight. We still write $\|p\|_4^2$ as above for simplicity of notations.

Remark 1.3 (**"Prepared" initial data**). The above estimate only requires $\nabla \cdot u_0 = O(\varepsilon)$ and $\partial_t u|_{t=0} = O(1)$. In this case, the compressible data u_0 is a small perturbation of an incompressible data u_0^0, and this perturbation is *completely contributed by the compressibility*. Such compressible data are usually called "well-prepared initial data"[2].

Remark 1.4 (**Relations with anisotropic Sobolev space**). The energy functional $E(t)$ above is considered as a variant of $\| \cdot \|_{8,*}$ norm at time $t > 0$. For different multi-index α, we impose different Mach number weights according to the number of tangential derivatives that appear in ∂_*^α, such that the energy estimate for the slightly modified $\| \cdot \|_{8,*}$ norm is uniform in $\varepsilon > 0$.

The next main theorem concerns the incompressible limit. We consider the incompressible MHD equations together with a transport equation satisfied by (u^0, B^0, π, S^0) with incompressible initial data (u_0^0, B_0^0) and S_0^0:

$$\begin{cases} \varrho(\partial_t u^0 + u^0 \cdot \nabla u^0) - B^0 \cdot \nabla B^0 + \nabla(\pi + \frac{1}{2}|B^0|^2) = 0 & \text{in } [0, T] \times \Omega, \\ \partial_t B^0 + u^0 \cdot \nabla B^0 - B^0 \cdot \nabla u^0 = 0 & \text{in } [0, T] \times \Omega, \\ \partial_t S^0 + u^0 \cdot \nabla S^0 = 0 & \text{in } [0, T] \times \Omega, \\ \nabla \cdot u^0 = \nabla \cdot B^0 = 0 & \text{in } [0, T] \times \Omega, \\ u_3^0 = B_3^0 = 0 & \text{on } [0, T] \times \Sigma, \\ (u_0, B_0, S_0)|_{t=0} = (u_0^0, B_0^0, S_0^0) & \text{on } \{t = 0\} \times \Omega. \end{cases} \tag{1.19}$$

---

[2]One can find the definitions of "well-prepared" and "ill-prepared" in [18, 1] for rescaled Euler system, which is equivalent to the statement in our paper.

Theorem 1.2 (**Incompressible limit**). Under the hypothesis of Theorem 1.1, we assume that $(u_0, B_0, S_0) \to (u_0^0, B_{03}^0, S_0^0)$ in $H^4(\Omega)$ as $\varepsilon \to 0$ for $\nabla \cdot u^0 = \nabla \cdot B^0 = 0$ in Ω with $u_{03}^0 = B_{03}^0 = 0$ on Σ. Then it holds that

$$(u, B, S) \to (u^0, B^0, S^0) \quad \text{weakly-* in } L^\infty([0, T]; H^4(\Omega)) \text{ and strongly in } C([0, T]; H^{4-\delta}(\Omega))$$

for $\delta > 0$. (u^0, B^0, S^0) solves (1.19), that is, the incompressible MHD equations together with a transport equation of S^0. Here ϱ satisfies the transport equation

$$\partial_t \varrho + u^0 \cdot \nabla \varrho = 0, \quad \varrho|_{t=0} = \rho(0, S_0^0),$$

where we consider the equation of state (1.9) as $\rho = \rho(\varepsilon^2 p, S)$, that is, a function of $\varepsilon^2 p$ and S. The function π satisfying $\nabla \pi \in C([0, T]; H^3(\Omega))$ represents the fluid pressure for incompressible MHD system (1.19).

Remark 1.5 (Choice of regularity). We propose H^4 regularity for the energy functional because lots of commutator estimates require the bound for $\|\bar{\partial}^2(u, B)\|_{L^\infty}$. Recall that $H^{\frac{d}{2}+\delta} \hookrightarrow L^\infty$, so it would be convenient to choose $H^{2+\lceil \frac{d}{2}+\delta \rceil}$ regularity, that is, H^4 for $d = 2, 3$. The initial data are set in $H^8(\Omega)$ in order to guarantee they satisfy the compatibility conditions up to 7-th order and the uniform-in-ε bound $E(0) \leq M$. Our results can be easily generalized to the case when H^4 and H_*^8 are replaced by H^m and H_*^{2m} $(m \geq 4)$ respectively.

Remark 1.6 (The space for the convergence of initial data). The compressible initial data converges to the incompressible data in $H^4(\Omega)$ instead of $H_*^8(\Omega)$ because the higher-order energy $E_5(0) \sim E_8(0)$ automatically vanishes as $\varepsilon \to 0$. Thus, it suffices to require the convergence in $H^4(\Omega)$.

Remark 1.7 (Boundedness of the domain). In the proof of the main theorem, we do not use the boundedness of the domain, such as Poincaré's inequality, to close the energy estimate. Thus, after making necessary modifications (including setting $\rho_0 - 1 \in H_*^8(\Omega)$ instead of ρ_0 itself and assuming the initial data to be "localized"), our proof is still valid for the case of unbounded reference domains such as $\mathbb{R}^{d-1} \times (-1, 1)$ and the half space. Indeed, using a partition of unity as in [4], the proof can be directly generalized to the domains which are diffeomorphic to the reference domains via $H_*^8(\Omega)$-diffeomorphisms.

1.4 Organization of the paper

This paper is organized as follows. In section 2, we discuss the main difficulties and briefly introduce our strategies to tackle the problem. Then section 3 is devoted to the proof of uniform estimates in Mach number. Combined with the compactness argument, we conclude the incompressible limit in section 4.

List of Notations

- (Tangential derivatives) $\mathcal{T}_0 = \partial_t$ denotes the time derivative, $\mathcal{T}_j = \bar{\partial}_j$ $(1 \leq j \leq d-1)$ denotes the tangential spatial derivatives and $\mathcal{T}_{d+1} := \omega(x_d)\partial_d$ with $\omega(x_d) = (1 - x_d)(1 + x_d)$.
- (Tangential components) $\bar{u} := (u_1, \cdots, u_{d-1})$ and $\bar{B} := (B_1, \cdots, B_{d-1})$. We also write $\bar{\nabla} := (\bar{\partial}_1, \bar{\partial}_2)$ as the tangential components of the operator ∇.
- (L^∞-norm) $\|\cdot\|_\infty := \|\cdot\|_{L^\infty(\Omega)}$, $|\cdot|_\infty := \|\cdot\|_{L^\infty(\Sigma)}$.
- (Interior Sobolev norm) $\|\cdot\|_s$: We denote $\|f\|_s := \|f(t, \cdot)\|_{H^s(\Omega)}$ for any function $f(t, y)$ on $[0, T] \times \Omega$.
- (Boundary Sobolev norm) $|\cdot|_s$: We denote $|f|_s := |f(t, \cdot)|_{H^s(\Sigma)}$ for any function $f(t, y)$ on $[0, T] \times \Sigma$.
- (Anisotropic Sobolev norms) $\|\cdot\|_{m,*}$: For any function $f(t, x)$ on $[0, T] \times \Omega$, $\|f\|_{m,*}^2 := \sum_{\langle \alpha \rangle \leq m} \|\partial_*^\alpha f(t, \cdot)\|_0^2$ denotes the m-th order space-time anisotropic Sobolev norm of f.
- (Polynomials) $P(\cdots)$ denotes a generic polynomial in its arguments.
- (Commutators) $[T, f]g = T(fg) - f(Tg)$, $[f, T]g = -[T, f]g$, $[T, f, g] := T(fg) - T(f)g - fT(g)$ where T is a differential operator and f, g are functions.
- (Equality modulo lower order terms) $A \overset{L}{=} B$ means $A = B$ modulo lower order terms.

2 Difficulties and strategies

Before going to the detailed proofs, we will discuss the main difficulties in this incompressible limit problem and briefly introduce our strategies to derive energy estimates that are uniform in the Mach number. Note that our initial data is assumed to be well-prepared, so the uniform estimates together with compactness argument are enough to derive the incompressible limit.

2.1 Choice of function spaces

Compressible ideal MHD with perfectly conducting wall conditions is a first-order quasilinear hyperbolic system with *characteristic* boundary conditions of constant multiplicity, for which there is a potential of normal derivative loss. In fact, inside a rigid wall with the slip boundary condition, Euler equations and elastodynamic equations with degenerate deformation tensor on the boundary (see [38]) also have characteristic boundary conditions, but their vorticity can be controlled in the setting of standard Sobolev spaces because there are no other quantities involved in the pressure part. The loss of normal derivatives is then compensated via the control of vorticity and divergence. However, the vorticity estimates for compressible ideal MHD cannot be closed in the setting of standard Sobolev spaces, which was explicitly presented in [37].

On the one hand, people observed that such normal derivative loss can be compensated if the vertical component of the magnetic field does not vanish on the boundary. See Yanagisawa [35] for the well-posedness and [19, 32, 33] for the study of corresponding free-boundary problems (MHD contact discontinuities). But the transversality of magnetic fields violates the perfectly conducting wall condition. On the other hand, Chen [3] first introduced the anisotropic Sobolev spaces defined in (1.14)-(1.15) in the study of compressible gas dynamics inside a rigid wall. Under this setting, a normal derivative is considered as a second-order derivative, which exactly compensates the derivative loss. Detailed analysis for MHD in such setting is referred to the second author's paper [16, Sect. 2.5]. Using such function spaces, Yanagisawa-Matsumura [36] and Secchi [26, 28] proved the local well-posedness and see also [30, 31, 16] for the study of free-boundary problems, but the estimates obtained in these previous works are not uniform in Mach number.

As stated in Section 1.2, the essential difficulty in proving the incompressible limit is that the function spaces for well-posedness of compressible problem are not "compatible" with the ones for the incompressible problems. Since our initial data is well-prepared, that is, a given incompressible data plus a slight perturbation, we shall still start with the incompressible counterpart and try to find out the relationships between the incompressible problem and the compressible problem.

2.2 Key observation: hidden structure of Lorentz force in the vorticity analysis

First, the entropy is easy to control thanks to $D_t S = 0$, so it suffices to analyze the relations between (u, B) and p. We start with the control of $\|(u, B)\|_4$. Using div-curl decomposition, the H^4 Sobolev norms are bounded by $\|\nabla \times u, \nabla \times B\|_3$, $\|\nabla \cdot u, \nabla \cdot B\|_3$ and the normal traces $|u_3, B_3|_{3.5}$. The boundary conditions (1.4) eliminate the normal traces, and the divergence part is reduced to tangential derivatives $\|\varepsilon^2 D_t p\|_3$ thanks to the continuity equation and the divergence constraint $\nabla \cdot B = 0$. As for the vorticity, taking curl in the momentum equation yields the evolution equation of vorticity

$$\rho D_t(\nabla \times u) - (B \cdot \nabla)(\nabla \times B) = \cdots , \tag{2.1}$$

where we find that the term $\nabla(p + \frac{1}{2}|B|^2)$ is already eliminated. In the H^3-control of $\nabla \times u$, we invoke the evolution equation of B to get

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \rho |\partial^3(\nabla \times u)|^2 + |\partial^3(\nabla \times B)|^2 \, \mathrm{d}x = -\int_\Omega \partial^3 \nabla \times (B(\nabla \cdot u)) \cdot \partial^3(\nabla \times B) \, \mathrm{d}x + \text{ controllable terms}, \tag{2.2}$$

where we find that there are 5 derivatives falling on u and thus the vorticity estimates cannot be closed in the setting of 4-th order standard Sobolev spaces. Such derivative loss in the vorticity analysis must appear because taking curl operator eliminates the term $\nabla(\frac{1}{2}|B|^2)$ in the momentum equation but does not eliminate the term $B(\nabla \cdot u)$. However, in the proof of L^2 estimate, the contribution of $\nabla(\frac{1}{2}|B|^2)$ should be cancelled with the contribution of $B(\nabla \cdot u)$.

7

If we further analyze this problematic term by using the structure of MHD system, we will find that the anisotropic Sobolev space is naturally introduced in order to close the energy estimates. Using the continuity equation $\varepsilon^2 D_t p = -\nabla \cdot u$, we find that the highest order term in (2.2) is $\varepsilon^2 B \times (\partial^3 \nabla D_t p)$. Then commuting ∇ with D_t and invoking the momentum equation $\rho D_t u + B \times (\nabla \times B) = -\nabla p$, we get

$$\varepsilon^2 B \times (\partial^3 \nabla D_t p) \overset{L}{=} \varepsilon^2 B \times (\partial^3 D_t \nabla p) \overset{L}{=} -\varepsilon^2 \rho B \times (\partial^3 D_t^2 u) - \varepsilon^2 B \times D_t(B \times \partial^3 (\nabla \times B)). \tag{2.3}$$

On the right side of the above identity, the second term still contains 4 normal derivatives $\partial^3 \nabla \times$ and 1 tangential derivative D_t which exhibits a loss of derivative. The key observation here is that the contribution of this term in (2.2) gives an energy term and thus avoid the loss of derivative. Specifically, using $\nabla \times (B(\nabla \cdot u)) = (\nabla \times B)(\nabla \cdot u) - B \times \nabla(\nabla \cdot u)$ and omitting the contribution of $(\nabla \times B)(\nabla \cdot u)$ in (2.2) (this part has no loss of derivative), we have

$$-\int_\Omega \partial^3 \nabla \times (B(\nabla \cdot u)) \cdot \partial^3 (\nabla \times B)\, dx \overset{L}{=} \int_\Omega B \times \partial^3 \nabla \underbrace{(\nabla \cdot u)}_{=-\varepsilon^2 D_t p} \cdot \partial^3 (\nabla \times B)\, dx$$

$$\overset{L}{=} \int_\Omega \varepsilon^2 \left(\rho B \times (\partial^3 D_t^2 u)\right) \cdot \partial^3 (\nabla \times B)\, dx + \int_\Omega \varepsilon^2 \left(B \times D_t \left(B \times \partial^3 (\nabla \times B)\right)\right) \cdot \partial^3 (\nabla \times B)\, dx, \tag{2.4}$$

where we insert (2.3) in the second equality. Then setting $\mathbf{u} = B$, $\mathbf{v} = D_t\left(B \times \partial^3(\nabla \times B)\right)$ and $\mathbf{w} = \partial^3(\nabla \times B)$ in the vector identity $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = -(\mathbf{u} \times \mathbf{w}) \cdot \mathbf{v}$, we find that the last term in (2.6) contributes to an energy term

$$\int_\Omega \varepsilon^2 \left(B \times D_t \left(B \times \partial^3(\nabla \times B)\right)\right) \cdot \partial^3(\nabla \times B)\, dx$$

$$= -\int_\Omega \varepsilon^2 \left(B \times \partial^3(\nabla \times B)\right) \cdot D_t \left(B \times \partial^3(\nabla \times B)\right)\, dx \tag{2.5}$$

$$= -\frac{1}{2} \frac{d}{dt} \int_\Omega \varepsilon^2 \left|B \times \partial^3(\nabla \times B)\right|^2\, dx + \text{ controllable terms.}$$

The above analysis for compressible ideal MHD shows that the vorticity estimates cannot be closed in standard Sobolev spaces, but in fact we *trade one normal derivative* (in $\nabla \times$) *for two tangential derivatives* $\varepsilon^2 D_t^2$ thanks to the special structure of Lorentz force $B \times (\nabla \times B)$ whose contribution in the vorticity analysis is an ε-weighted energy term instead of introducing loss of derivatives. This exactly illustrates why the anisotropic Sobolev space is naturally introduced to study compressible ideal MHD.

It should also be noted that this difficulty never occurs for incompressible ideal MHD because the divergence-free condition automatically eliminates the bad term above. Indeed, taking the limit $\varepsilon \to 0_+$, the right side of (2.2) automatically vanishes and the vorticity estimates for the incompressible problem can be established in standard Sobolev spaces. Besides, to prove the incompressible limit of ideal MHD with (1.4), an extra Mach number weight ε^2 must be added together with the two tangential derivatives. So, we find that the "anisotropic part" of the energy, namely the terms $E_5 \sim E_8$ in (1.18), will automatically vanish when we take the incompressible limit $\varepsilon \to 0_+$. The remaining term $E_4(t)$ exactly consists of the standard Sobolev norms, which coincides with the energy for incompressible ideal MHD.

Remark 2.1 (Slight modifications in 2D). When dimension $d = 2$, we have to replace the curl operator $\nabla \times$ by $\nabla^\perp \cdot$ where $\nabla^\perp := (-\partial_2, \partial_1)$. Then (2.2) reads

$$\frac{d}{dt} \int_\Omega \rho |\partial^3(\nabla^\perp \cdot u)|^2 + |\partial^3(\nabla^\perp \cdot B)|^2\, dx = -\int_\Omega \left((B \cdot \nabla^\perp)\partial^3 \nabla \cdot u\right)(\partial^3 \nabla^\perp \cdot B)\, dx + \text{ controllable terms,} \tag{2.6}$$

where the highest order term can be reduced as follows after invoking the first two equations of (1.13)

$$-(B \cdot \nabla^\perp)\partial^3 \nabla \cdot u \overset{L}{=} \varepsilon^2 B \cdot \partial^3 D_t \nabla^\perp p \overset{L}{=} \varepsilon^2 (B_2 \partial^3 D_t \partial_1 p - B_1 \partial^3 D_t \partial_2 p)$$

$$\overset{L}{=} -\varepsilon^2 \rho (B_2 \partial^3 D_t^2 u_1 - B_1 \partial^3 D_t^2 u_2) - \varepsilon^2 (B_2^2 + B_1^2) D_t \partial^3 (\nabla^\perp \cdot B),$$

where we note that the momentum equations can be written as $\partial_1 p = -\rho D_t u_1 - B_2(\nabla^\perp \cdot B)$ and $\partial_2 p = -\rho D_t u_2 + B_1(\nabla^\perp \cdot B)$. The first term can be treated as in the 3D case, and the contribution of the second term gives an energy term $-\frac{1}{2}\frac{d}{dt}\int_\Omega \varepsilon^2 |B|^2 |\partial^3(\nabla^\perp \cdot B)|^2\, dx$ which also automatically vanishes when we take the incompressible limit $\varepsilon \to 0_+$.

It should be emphasized that our analysis is very different from the previous works about well-posedness [36, 26, 16]. Indeed, these previous works [36, 26, 16] avoided using div-curl analysis to reduce the normal derivatives of u and B. Instead, they controlled the normal derivatives by directly computing the corresponding L^2-type estimates which yields non-zero boundary integrals such as $-\int_\Sigma \partial_3^4 P \, \partial_3^4 v_3 \, dx'$ and only gives the estimates for $(u, B, \varepsilon p)$, not (u, B, p). Since taking normal derivatives does not preserve the boundary conditions, one has to use the MHD equations to replace $\partial_3 P$ and $\partial_3 u_3$ by tangential derivatives of the other variables, and then uses Gauss-Green formula to rewrite the boundary integrals into the interior. During this process, more time derivatives might be produced without extra weights of Mach number (e.g., the product of the underlined terms in $-\nabla P \to \underline{D_t v} - (\bar{B} \cdot \overline{\nabla}) B$ and $\partial_3 v_3 \to -\varepsilon^2 D_t p - \underline{\overline{\nabla} \cdot \bar{v}}$). However, the continuity equation indicates that one may have to add more Mach number weights to higher-order time derivatives. Thus, following the methods in [36, 26, 16] might cause a potential of loss of Mach number weight, and one can find related details in [16, Sect. 2.5]. In other words, these previous works [36, 26, 16] only revealed that one should trade one normal derivative for two derivatives but ignored the necessity of adding extra weight of Mach number ε^2. One of the advantanges of our method is that all normal derivatives are reduced via div-curl analysis which does not involve any boundary estimates; and every time when we reduce a normal derivative, we never lose weights of Mach number because the continuity equation must be used in the reduction.

2.3   Reduction of pressure and design of energy functional

The main idea of proving the uniform-in-ε estimates is to repeatedly reduce normal derivatives to tangential derivatives until all derivatives are tangential (with suitable Mach number weights); and reduce spatial derivatives of p to tangential derivatives of u, B until all derivatives on p are time derivatives. Once we only have tangential derivatives, it suffices to mimic the L^2 estimates to close the tangential estimates because taking tangential derivatives preserves the boundary conditions.

From the definition of D_t and $u_3|_\Sigma = B_3|_\Sigma = 0$, we know both D_t and $B \cdot \nabla$ are tangential derivatives. Based on this fact and the analysis in section 2.2, we have the following reductions:

a. Vorticity: $\nabla \times (u, B) \to \varepsilon^2 \mathcal{T}^2 u$.
b. Divergence: $\nabla \cdot (u, B) \to \varepsilon^2 \mathcal{T} p$.
c. Reduction of pressure: The momentum equation reads $-\nabla(p + |B|^2/2) = \rho D_t u - B \cdot \nabla B$ gives the relation $\nabla P \to \mathcal{T}(u, B)$ and thus $\nabla p \to \mathcal{T}(u, B)$ plus ∇B where ∇B is further reduced by using (a) and (b). This indicates that $\partial_t^k \nabla p$ should have the same Mach number weight as $\partial_t^{k+1}(u, B, S)$.
d. Tangential estimates: When estimating $E_{4+l}(t)$ (defined in (1.18)), $\mathcal{T}^\alpha(u, B)$ is controlled together with $\varepsilon \mathcal{T}^\alpha p$ in the estimates of full tangential derivatives, i.e., when $\langle \alpha \rangle = 4 + l$.

In the above subsection we start with E_4 and reduce $\|u, B\|_4$ (part of E_4) to $\|\varepsilon^2 \mathcal{T} p\|_3$ (still a part of E_4) and $\|\varepsilon^2 \mathcal{T}^2 u\|_3$ (part of E_5) via the div-curl analysis. The divergence part introduces time derivative, but the number of derivatives does not exceed 4, so we can repeatedly use (c) to reduce the spatial derivatives of p to tangential derivatives of u, B until there is no spatial derivative falling on p. Finally, it suffices to control 4-th order time derivatives of u, B and p with suitable Mach number weights. The Mach number weights of these quantities are determined by the relation (d) and the "preparedness" of initial data. See also diagram 2 below.

For the control of terms that have 5 derivatives, for example the term $\|\varepsilon^2 \mathcal{T}^2 u\|_3$, we repeat the div-curl analysis to get $\|\varepsilon^4 \mathcal{T}^3 p\|_2$ (still a part of E_5) from the divergence part and $\|\varepsilon^4 \mathcal{T}^4 u\|_2$ (part of E_6) from the curl part. Again, the divergence part is reduced to $E_5(t)$ which is controlled in a similar way as $E_4(t)$; and then we do further div-curl decomposition for $\|\varepsilon^4 \mathcal{T}^4 u\|_2$ which then requires the control of $E_7(t)$. Repeatedly, we finally need to control the L^2 norms of $\varepsilon^8 \mathcal{T}^8(u, B)$, which is controlled together with $\varepsilon^8 \mathcal{T}^8 p$ when $\mathcal{T}^8$ are not purely time derivatives, and $\varepsilon^9 \partial_t^8 p$ when $\mathcal{T}^8$ are purely time derivatives. Then the control of $\mathcal{T}^8$ is completely parallel to the proof of L^2 energy conservation because there isn't any normal derivative.

We present the above complicated reduction procedure in the following diagrams. For more details of the reduction procedures, we refer to section 3.2.3. We also list an example of reduction of $E_{4+l}(t)$ by repeatedly using (c) and (d) when $l = 0$. The other cases follow in the same manner.
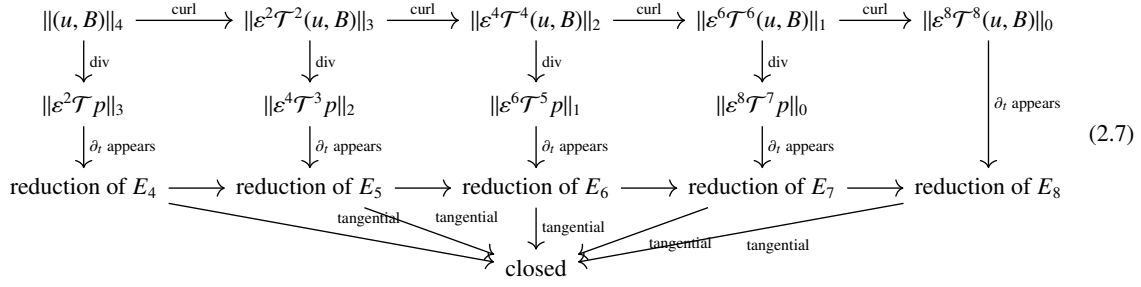
$$\begin{array}{ccccccccc}
\|(u,B)\|_4 & \xrightarrow{\text{curl}} & \|\varepsilon^2\mathcal{T}^2(u,B)\|_3 & \xrightarrow{\text{curl}} & \|\varepsilon^4\mathcal{T}^4(u,B)\|_2 & \xrightarrow{\text{curl}} & \|\varepsilon^6\mathcal{T}^6(u,B)\|_1 & \xrightarrow{\text{curl}} & \|\varepsilon^8\mathcal{T}^8(u,B)\|_0 \\
\downarrow{\scriptstyle\text{div}} & & \downarrow{\scriptstyle\text{div}} & & \downarrow{\scriptstyle\text{div}} & & \downarrow{\scriptstyle\text{div}} & & \\
\|\varepsilon^2\mathcal{T}p\|_3 & & \|\varepsilon^4\mathcal{T}^3p\|_2 & & \|\varepsilon^6\mathcal{T}^5p\|_1 & & \|\varepsilon^8\mathcal{T}^7p\|_0 & & \downarrow{\scriptstyle\partial_t\text{ appears}} \\
\downarrow{\scriptstyle\partial_t\text{ appears}} & & \downarrow{\scriptstyle\partial_t\text{ appears}} & & \downarrow{\scriptstyle\partial_t\text{ appears}} & & \downarrow{\scriptstyle\partial_t\text{ appears}} & & \\
\text{reduction of } E_4 & \longrightarrow & \text{reduction of } E_5 & \longrightarrow & \text{reduction of } E_6 & \longrightarrow & \text{reduction of } E_7 & \longrightarrow & \text{reduction of } E_8
\end{array} \tag{2.7}$$

reduction of E_4, E_5, E_6, E_7, E_8 → tangential → closed

Diagram (2.7): From standard Sobolev space to anisotropic Sobolev space.

$$\begin{array}{ccccccccc}
 & & & & E_5(t) & & & & \\
 & {\scriptstyle\text{curl}}\nearrow & {\scriptstyle\text{curl}}\nearrow & {\scriptstyle\text{curl}}\uparrow & & {\scriptstyle\text{curl}}\nwarrow & & & \\
\|(u,B)\|_4 & & \|\partial_t(u,B)\|_3 & & \|\varepsilon\partial_t^2(u,B)\|_2 & & \|\varepsilon^2\partial_t^3(u,B)\|_1 & & \|\varepsilon^3\partial_t^4(u,B)\|_0 \\
{\scriptstyle(c)}\nearrow & & {\scriptstyle(c)}\nearrow & \downarrow{\scriptstyle(d)} & & {\scriptstyle(c)}\nearrow & \downarrow{\scriptstyle(d)} & & {\scriptstyle(c)}\nearrow\ \downarrow{\scriptstyle(d)} \\
\|p\|_4 & & \|\varepsilon\partial_t p\|_3 & & \|\varepsilon^2\partial_t^2 p\|_2 & & \|\varepsilon^3\partial_t^3 p\|_1 & & \|\varepsilon^4\partial_t^4 p\|_0 \xrightarrow{\text{tangential}} \text{closed}
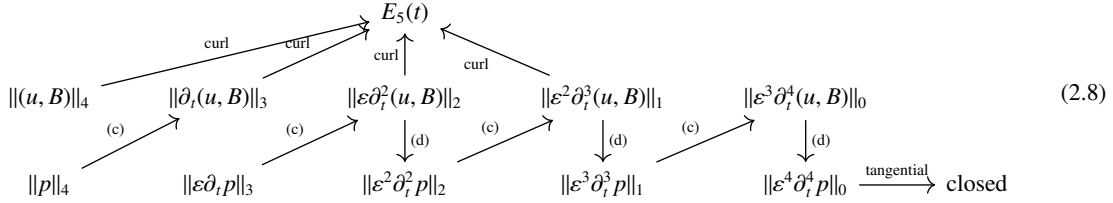\end{array} \tag{2.8}$$

Diagram (2.8): Reduction of $E_4(t)$ via (c) and (d).

Recall that the vorticity estimates for compressible Euler equations can be closed under the setting of standard Sobolev spaces. As presented in [17], the reduction scheme for Euler equations is merely the control of $E_4(t)$ but does not involve any higher-order energies. In other words, the reduction scheme is finished within the first column of diagram (2.7), or equivalently diagram (2.8) with $E_5(t)$ replaced by $E_4(t)$. Compared with Euler equations, the extra thing for compressible ideal MHD is that the vorticity analysis for $E_{4+l}(t)$ requires the control of $E_{4+l+1}(t)$ until there is no normal derivative on u, B. That's why we have the first row in diagram (2.7). Then, as presented in each column of diagram (2.7), we need to mimic the scheme in [17], that is, run the scheme presented in diagram (2.8), for each $E_{4+l}(t)$ to close the energy estimates.

2.4 Further applications of our method

Based on the above analysis, we give a complete and definitive answer to the incompressible limit problem of ideal MHD under the perfectly conducting wall condition when the initial data is well-prepared. Besides, compared with the previous work [13] by Ju and the first author, we give a clear illustration on the "incompatibility" of the function spaces of well-posedness and thoroughly resolve this issue. It should be noted that, in the uniform-in-ε estimates, the "preparedness" of initial data is only used to guarantee $E(0) < \infty$. If the initial data is not "prepared", that is, $\nabla \cdot u_0 = O(1)$ is not small and $\partial_t u|_{t=0} = O(1/\varepsilon)$ is unbounded, we believe that the uniform-in-ε estimates can be proven in a similar way after doing suitable technical modifications. Hence, one can expect to generalize our result to the case when the initial data is not "prepared".

It should also be noted that the reduction procedure presented in diagram (2.8) is parallel to the case of compressible water wave, which is presented in the paper [17, Sect. 2.1.3] by Luo and the second author. Indeed, it is exactly the appearance of the magnetic field and the perfectly conducting wall condition that force us to do further vorticity analysis as stated in section 2.2 and presented in diagram (2.7). Therefore, one can expect to extend our results to free-boundary problems. In fact, we believe that our method together with the techniques about free-boundary problems presented in [17] can be generalized to study the current-vortex sheets in ideal compressible MHD, which will be presented in a forthcoming paper by the second author.

3 Uniform energy estimates

3.1 L^2 estimate

First, we establish L^2-energy estimate for (1.13). Invoking the momentum equation and integrate by parts, we have:

$$\frac{1}{2}\frac{d}{dt}\int_\Omega \varepsilon^2 p^2 + \rho|u|^2 + |B|^2 + S^2 \, dx$$

$$= \frac{1}{2}\int_\Omega \nabla \cdot u(\varepsilon^2 p^2 + |B|^2 + S^2) \, dx - \int_\Omega \nabla \cdot (pu) \, dx$$

$$+ \int_\Omega (B \cdot \nabla B) \cdot u + (B \cdot \nabla u) \cdot B \, dx - \int_\Omega \frac{1}{2}\nabla|B|^2 \cdot u + (B \cdot B)\nabla \cdot u \, dx$$

$$= \frac{1}{2}\int_\Omega \nabla \cdot u(\varepsilon^2 p^2 + |B|^2 + S^2) \, dx, \tag{3.1}$$

which give the energy estimate:

$$\frac{d}{dt}\|(\varepsilon p, u, B, S)\|_0^2 \leq P(E_4(t)). \tag{3.2}$$

3.2 Reduction of normal derivatives

3.2.1 Reduction of pressure

We show how to reduce the control of the pressure to that of the velocity and magnetic field when there is at least one spatial derivative on P. This follows from using the momentum equation

$$-\nabla(p + |B|^2/2) = \rho D_t u - B \cdot \nabla B. \tag{3.3}$$

Then

$$\|\nabla(p + |B|^2/2)\|_0 \lesssim \|\rho\|_\infty \|D_t u\|_0 + \|B \cdot \nabla B\|_0, \tag{3.4}$$

where $D_t u = \partial_t u + \bar{u} \cdot \overline{\nabla} u + u_3\partial_3 u$ and $B \cdot \nabla B = \bar{B} \cdot \overline{\nabla} B + B_3\partial_3 B$. Similarly, using (3.3), we also have the reduction for L^∞ norm

$$\|\nabla(p + |B|^2/2)\|_\infty \lesssim \|\rho\|_\infty \|D_t u\|_\infty + \|B \cdot \nabla B\|_\infty. \tag{3.5}$$

Recall that $D_t = \partial_t + \bar{u} \cdot \overline{\nabla} + u_3\partial_3$ and $B \cdot \nabla = \bar{B} \cdot \overline{\nabla} + B_3\partial_3$ are both tangential derivatives thanks to $u_3 = B_3 = 0$ on Σ. The above inequalities (3.4)-(3.5) show that ∇P is reduced to $\partial_t u$, $\bar{\partial} u$, $\omega(x_3)\partial_3 u$, $\bar{\partial} B$ and $\omega(x_3)\partial_3 B$ for some weight function $\omega(x_3)$ vanishing on Σ. Thus, the momentum equation reduces a *normal* derivative of P to a *tangential* derivative of u, B.

Also, one can reduce the Sobolev norm of ∇P in the same way. Let $\mathcal{T} = \partial_t$ or $\bar{\partial}$ or $\omega(x_3)\partial_3$ and $D = \partial$ or ∂_t. The above estimate yields the control of $\|\mathcal{T}^\alpha D^k \nabla p\|_0$ after taking $\mathcal{T}^\alpha D^k$, $k + \langle \alpha \rangle \geq 1$, to (3.3). Specifically, at the leading order, $\|\mathcal{T}^\alpha D^k \nabla(p + |B|^2/2)\|_0$ is controlled by

$$\|\rho\|_\infty \|\mathcal{T}^\alpha D^k \mathcal{T} u\|_0 + \|\mathcal{T}^\alpha D^k \mathcal{T} B\|_0. \tag{3.6}$$

As for the fluid pressure p, by definition we have $\partial_i p = \partial_i P - \partial_i B \cdot B$. Thus, to control the Sobolev norms of ∇p, we still need to do further analysis for ∇B, which will be controlled together with ∇u.

3.2.2 Div-Curl analysis

From the analysis of section 3.2.1, in order to control the Sobolev norms defined in (1.18), we shall reduce the normal derivatives of u and B and then analyze the tangential derivatives. We use the following lemma to reduce the normal derivatives of u, B to their divergence and curl.

Lemma 3.1 (Hodge elliptic estimates). For any sufficiently smooth vector field X and any real number $s \geq 1$, one has

$$\|X\|_s^2 \lesssim \|X\|_0^2 + \|\nabla \cdot X\|_{s-1}^2 + \|\nabla \times X\|_{s-1}^2 + |\bar{\partial}X \cdot N|_{s-\frac{3}{2}}^2. \tag{3.7}$$

We start from the control of $E_4(t)$. We apply Lemma 3.1 to $\|\varepsilon^{(k-1)_+}\partial_t^k(u,B)\|_{4-k}$ for $0 \le k \le 3$ to get

$$\|\varepsilon^{(k-1)_+}\partial_t^k u\|_{4-k}^2 \lesssim \|\varepsilon^{(k-1)_+}\partial_t^k u\|_0^2 + \|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k u\|_{3-k}^2 + \|\varepsilon^{(k-1)_+}\nabla \cdot \partial_t^k u\|_{3-k}^2, \tag{3.8}$$

$$\|\varepsilon^{(k-1)_+}\partial_t^k B\|_{4-k}^2 \lesssim \|\varepsilon^{(k-1)_+}\partial_t^k B\|_0^2 + \|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k B\|_{3-k}^2, \tag{3.9}$$

where we already use $\nabla \cdot B = 0$ and the boundary condition $u_3 = B_3 = 0$ on Σ.

The L^2 norm of u and B has been controlled in (3.2), while the control of $\|\varepsilon^{(k-1)_+}\partial_t^k(u,B)\|_0$ ($1 \le k \le 3$) is parallel to the case of $k = 0$ and will be postponed to later sections. The divergence part is reduced to the estimates of p by using the continuity equation

$$\|\varepsilon^{(k-1)_+}\nabla \cdot \partial_t^k u\|_{3-k}^2 = \left\|\varepsilon^{(k-1)_++2}\partial_t^k(D_t p)\right\|_{3-k}^2, \tag{3.10}$$

which will be further reduced to the tangenital estimates of (u,B) by using the argument in Section 3.2.1.

According to the reduction procedure in section 3.2.1 and (b)-(d) in section 2.3, the control of $E_4(t)$ will be reduced to the control of $\mathcal{T}^\alpha(u,B)$ and $\partial_t^4 p$ with suitable Mach number weights, plus the curl part. Next, we analyze the curl part via the evolution equation of vorticity.

Lemma 3.2.

$$\sum_{k=0}^{3}\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k u\right\|_{3-k}^2 + \left\|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k B\right\|_{3-k}^2 + \left\|\varepsilon^{(k-1)_++1}B \times (\nabla \times \partial_t^k B)\right\|_{3-k}^2\right) \le P(E_4(t), E_5(t)). \tag{3.11}$$

Proof. First, we take $\nabla\times$ in the momentum equation $\rho D_t u - (B \cdot \nabla)B = -\nabla P$ to get the evolution equation of vorticity

$$\rho D_t(\nabla \times u) - (B \cdot \nabla)\nabla \times B = \rho[D_t, \nabla\times]u - [B \cdot \nabla, \nabla\times]B - (\nabla\rho) \times (D_t u), \tag{3.12}$$

where we notice that the right side only contains the first-order derivatives and does not lose Mach number weight. Note that the equation of state is smooth, so $\partial_S \rho$ is bounded.

$$[D_t, \partial_i](\cdot) = -\partial_i u_j \partial_j(\cdot), \quad [B \cdot \nabla, \partial_i](\cdot) = -\partial_i B_k \partial_k(\cdot), \quad \nabla\rho = \frac{\partial\rho}{\partial p}\nabla p + \partial_S \rho \nabla S = O(\varepsilon^2)\nabla p + \partial_S \rho \nabla S.$$

We first prove the case when $k = 0$, and the cases for $1 \le k \le 3$ follow in the same manner. Considering the structure of the evolution equation (3.12) and ρ is bounded below by a positive constant, it suffices to prove that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho\left|\partial^3\nabla \times u\right|^2 + \left|\partial^3\nabla \times B\right|^2 + \left|\varepsilon B \times (\partial^3\nabla \times B)\right|^2 \,\mathrm{d}x \le P(E_4(t), E_5(t)).$$

We take ∂^3 in (3.12) to get

$$\rho D_t\partial^3(\nabla \times u) - (B \cdot \nabla)\partial^3\nabla \times B = \underbrace{\partial^3(\text{RHS of (3.12)}) + [\rho D_t, \partial^3](\nabla \times u) - [B \cdot \nabla, \partial^3](\nabla \times B)}_{\mathcal{R}_1}, \tag{3.13}$$

where the order of derivatives on the right side must be ≤ 4. Now, standard L^2-type estimate yields that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho|\partial^3\nabla \times u|^2 \,\mathrm{d}x = \int_\Omega \rho(\partial_t\partial^3\nabla \times u) \cdot \partial^3\nabla \times u \,\mathrm{d}x + \frac{1}{2}\int_\Omega \partial_t\rho|\partial^3\nabla \times u|^2 \,\mathrm{d}x$$

$$= \int_\Omega \rho D_t\partial^3\nabla \times u \cdot \partial^3\nabla \times u \,\mathrm{d}x$$

$$\underbrace{- \int_\Omega \rho(u \cdot \nabla)\partial^3\nabla \times u \cdot \partial^3\nabla \times u \,\mathrm{d}x - \frac{1}{2}\int_\Omega (\frac{\partial\rho}{\partial p}\partial_t p + \partial_S\rho\partial_t S)|\partial^3\nabla \times u|^2 \,\mathrm{d}x}_{:=\mathcal{I}_1}. \tag{3.14}$$

12

Then invoking (3.13) gives us the following terms

$$\int_\Omega (\rho D_t \partial^3 \nabla \times u) \cdot (\partial^3 \nabla \times u) \, dx$$

$$= \int_\Omega \left((B \cdot \nabla)\partial^3 \nabla \times B\right) \cdot (\partial^3 \nabla \times u) \, dx + \underbrace{\int_\Omega \mathcal{R}_1 \cdot (\partial^3 \nabla \times u) \, dx}_{:=\mathcal{I}_2}$$

$$\overset{B \cdot \nabla}{=} - \int_\Omega (\partial^3 \nabla \times B) \cdot \left(\partial^3 \nabla \times (B \cdot \nabla u)\right) dx - \underbrace{\int_\Omega (\partial^3 \nabla \times B) \cdot \left([B \cdot \nabla, \partial^3 \nabla \times] u\right) dx}_{:=\mathcal{I}_3} + \mathcal{I}_2. \qquad (3.15)$$

Next, we insert the evolution equation of B, that is, $D_t B = (B \cdot \nabla)u - B(\nabla \cdot u)$ to get

$$- \int_\Omega (\partial^3 \nabla \times B) \cdot \left(\partial^3 \nabla \times (B \cdot \nabla u)\right) dx$$

$$= - \int_\Omega (\partial^3 \nabla \times B) \cdot (\partial^3 \nabla \times D_t B) \, dx - \int_\Omega (\partial^3 \nabla \times B) \cdot \left(\partial^3 \nabla \times (B \nabla \cdot u)\right) dx$$

$$= - \frac{1}{2} \frac{d}{dt} \int_\Omega |\partial^3 \nabla \times B|^2 \, dx - \underbrace{\int_\Omega \partial^3 \nabla \times B \cdot \left([\partial^3 \nabla \times, D_t]B + (u \cdot \nabla)\partial^3 \nabla \times B\right) dx}_{:=\mathcal{I}_4}$$

$$\underbrace{+ \int_\Omega (\partial^3 \nabla \times B) \cdot \partial^3 \nabla \times (\varepsilon^2 B D_t p) \, dx}_{:=\mathcal{K}} . \qquad (3.16)$$

Based on the concrete form of the commutators, a straightforward product estimate for $\mathcal{I}_j$ ($1 \leq j \leq 4$) gives us

$$\sum_{j=1}^4 \mathcal{I}_j \leq P(E_4(t)). \qquad (3.17)$$

The most difficult term is $\mathcal{K}$ because the highest-order term has 5-th order derivative and thus cannot be controlled by $E_4(t)$:

$$\varepsilon^2 \partial^3 \nabla \times (B D_t p) = \varepsilon^2 \partial^3 ((\nabla D_t p) \times B + D_t p(\nabla \times B))$$

$$= - \varepsilon^2 B \times (\partial^3 \nabla D_t p) \underbrace{- \varepsilon^2 [\partial^3, B \times](\nabla D_t p) + \partial^3 (\varepsilon^2 D_t p(\nabla \times B))}_{\mathcal{R}_2},$$

where straightforward calculation shows that $\|\mathcal{R}_2\|_0 \leq P(E_4(t))$. We then further analyze the problematic term $-\varepsilon^2 B \times (\partial^3 \nabla D_t p)$. We invoke a simple vector identity (in 3D) $(B \cdot \nabla)B - \frac{1}{2}\nabla|B|^2 = -B \times (\nabla \times B)$ to rewrite the momentum equation to be

$$\rho D_t u + B \times (\nabla \times B) = -\nabla p.$$

So, we commute D_t with ∇ and insert this equation into $\varepsilon^2 B \times (\partial^3 \nabla D_t p)$ to get

$$- \varepsilon^2 B \times (\partial^3 \nabla D_t p) = -\varepsilon^2 B \times (\partial^3 D_t \nabla p) - \varepsilon^2 B \times (\partial^3 (\nabla u_j \partial_j p))$$

$$= \varepsilon^2 B \times \partial^3 D_t (\rho D_t u) + \varepsilon^2 B \times \partial^3 D_t (B \times (\nabla \times B)) - \varepsilon^2 B \times (\partial^3 (\nabla u_j \partial_j p))$$

$$= \varepsilon^2 \rho B \times (\partial^3 D_t^2 u) + \underbrace{\varepsilon^2 B \times D_t (B \times \partial^3 (\nabla \times B))}_{=\mathcal{K}_1}$$

$$\underbrace{+ \varepsilon^2 B \times ([\partial^3 D_t, \rho]D_t u) + \varepsilon^2 B \times \left([\partial^3, D_t](\nabla \times B) + D_t \left([\partial^3, B \times](\nabla \times B)\right)\right) - \varepsilon^2 B \times (\partial^3 (\nabla u_j \partial_j p))}_{\mathcal{R}_3}, \qquad (3.18)$$

where $\|\mathcal{R}_3\|_0 \leq P(E_4(t))$ can be proved by using again the concrete forms of these commutators as the order of derivatives does not exceed 4. In particular, in the commutator $D_t \left([\partial^3, B \times](\nabla \times B)\right)$, the highest-order

terms have the form $D_t\left(\partial B \times \partial^2(\nabla \times B)\right)$ or $D_t\left(\partial^3 B \times (\nabla \times B)\right)$ and contain a time derivative, in which there is no loss of either derivative or ε-weights because there is already an ε^2 factor in $\mathcal{R}_3$. Now, we analyze the contribution of $\mathcal{K}_1$ in the integral $\mathcal{K}$, which reads

$$\int_\Omega \varepsilon^2 (\partial^3 \nabla \times B) \cdot \left(B \times D_t \left(B \times \partial^3(\nabla \times B)\right)\right) \, \mathrm{d}x. \tag{3.19}$$

Then we let $\mathbf{u} = B$, $\mathbf{v} = D_t\left(B \times \partial^3(\nabla \times B)\right)$ and $\mathbf{w} = \partial^3(\nabla \times B)$ in the vector identity $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = -(\mathbf{u} \times \mathbf{w}) \cdot \mathbf{v}$ to get

$$\int_\Omega \varepsilon^2 (\partial^3 \nabla \times B) \cdot \left(B \times D_t \left(B \times \partial^3(\nabla \times B)\right)\right) \mathrm{d}x = -\int_\Omega \varepsilon^2 D_t\left(B \times (\partial^3 \nabla \times B)\right) \cdot \left(B \times (\partial^3 \nabla \times B)\right) \mathrm{d}x, \tag{3.20}$$

in which the right side contributes to an ε-weighted energy terms instead of introducing a loss of derivative

$$-\int_\Omega \varepsilon^2 D_t\left(B \times (\partial^3 \nabla \times B)\right) \cdot \left(B \times (\partial^3 \nabla \times B)\right) \mathrm{d}x$$
$$= -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \varepsilon^2 \left|B \times (\partial^3 \nabla \times B)\right|^2 \mathrm{d}x + \underbrace{\frac{1}{2}\int_\Omega (\nabla \cdot u) \left|B \times (\partial^3 \nabla \times B)\right|^2 \mathrm{d}x}_{\mathcal{R}_4}. \tag{3.21}$$

The term $\mathcal{R}_4$ is directly controlled by $\|u\|_{W^{1,\infty}}\|B\|_{L^\infty}^2\|B\|_4^2 \leq P(E_4(t))$. Therefore, we control the term $\mathcal{K}$ with both $E_4(t)$ and $E_5(t)$ as follows:

$$\mathcal{K} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \varepsilon^2 \left|B \times (\partial^3 \nabla \times B)\right|^2 \mathrm{d}x \leq \|B\|_4 \left(P(E_4(t)) + \|\rho B\|_\infty \|\varepsilon^2 D_t^2 u\|_3\right) \leq P(E_4(t), E_5(t)), \tag{3.22}$$

which gives us the energy estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\nabla \times u\|_3^2 + \|\nabla \times B\|_3^2 + \|\varepsilon B \times (\nabla \times B)\|_3^2\right) \leq P(E_4(t), E_5(t)). \tag{3.23}$$

Recall that $D_t = \partial_t + u \cdot \nabla = \partial_t + \bar{u} \cdot \overline{\nabla} + u_3 \partial_3$ and $u_3|_\Sigma = 0$ imply that D_t is a tangential derivative. So, the above analysis shows that the curl estimate for compressible ideal MHD cannot be closed in standard Sobolev space, but in fact we can trade one normal derivative $(\nabla \times)$ for two tangential derivatives together with square Mach number weight, that is, $\varepsilon^2 \mathcal{T}^2$. This step naturally introduces the so-called anisotropic Sobolev space and also explains why we add ε^2 weight to $E_5(t)$.

Similarly, we can prove the same conclusion for $\partial^\alpha \partial_t^k$ with $k + |\alpha| = 3$ by replacing ∂^3 with $\varepsilon^{(k-1)_+}\partial^\alpha \partial_t^k$. Indeed, the highest order derivatives in the above commutators do not exceed 4-th order, and there is no loss of Mach number weight because none of the above steps creates negative power of Mach number. Besides, the problematic term can still be analyzed in the same manner:

$$\varepsilon^{(k-1)_++2} B \times (\partial^\alpha \partial_t^k \nabla D_t p) \stackrel{L}{=} -\varepsilon^{(k-1)_+}\rho B \times \partial^\alpha \partial_t^k(\varepsilon^2 D_t^2 u) - \varepsilon^{(k-1)_+}\varepsilon^2 B \times D_t\left(B \times (\partial^\alpha \partial_t^k(\nabla \times B))\right),$$

where the L^2 norms of the omitted lower-order terms are controlled by $P(E_4(t))$. The contribution of the last term in the above equality is again the energy term $-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \varepsilon^2 \left|B \times (\partial^\alpha \partial_t^k \nabla \times B)\right|^2 \mathrm{d}x$. The proof of this follows in the same way if we replace ∂^3 above by $\varepsilon^{(k-1)_+}\partial^\alpha \partial_t^k$ (that is, using the vector identity $(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w} = -(\mathbf{u} \times \mathbf{w}) \cdot \mathbf{v}$. Hence, we can conclude the vorticity analysis for $E_4(t)$ with the following inequality

$$\sum_{k=0}^3 \frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k u\right\|_{3-k}^2 + \left\|\varepsilon^{(k-1)_+}\nabla \times \partial_t^k B\right\|_{3-k}^2 + \left\|\varepsilon^{(k-1)_++1}B \times \left(\partial_t^k \nabla \times B\right)\right\|_{3-k}^2\right) \leq P(E_4(t), E_5(t)). \tag{3.24}$$

$\square$

The above div-curl analysis shows that it is necessary to control $\|\varepsilon^{2+(k-1)_+}\mathcal{T}^2\partial_t^k u\|_{3-k}$, which should be analyzed together with $\|\varepsilon^{2+(k-1)_+}\mathcal{T}^2\partial_t^k B\|_{3-k}$. When $k = 3$, $\mathcal{T}^2\partial_t^3$ is a purely tangential derivative and we

14

would like to postpone the tangential estimates to the next subsection. When $0 \leq k \leq 2$, we again apply Lemma 3.1 to $\|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2(u,B)\|_{3-k}$ to obtain

$$\|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2 u\|_{3-k}^2 \lesssim \|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2 u\|_0^2 + \|\varepsilon^{2+(k-1)_+}\partial_t^k\nabla\cdot\mathcal{T}^2 u\|_{2-k}^2 + \|\varepsilon^{2+(k-1)_+}\partial_t^k\nabla\times\mathcal{T}^2 u\|_{2-k}^2, \qquad (3.25)$$

$$\|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2 B\|_{3-k}^2 \lesssim \|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2 B\|_0^2 + \|\varepsilon^{2+(k-1)_+}[\mathcal{T}^2,\nabla\cdot]\partial_t^k B\|_{2-k}^2, \qquad (3.26)$$

where we already use the boundary conditions and the divergence constraint to eliminate the boundary normal traces and $\nabla\cdot B$. The divergence part is reduced to the estimates of p by using the continuity equation

$$\|\varepsilon^{2+(k-1)_+}\partial_t^k\nabla\cdot\mathcal{T}^2 u\|_{2-k}^2 \leq \|\varepsilon^{2+(k-1)_+}\partial_t^k\mathcal{T}^2\nabla\cdot u\|_{2-k}^2 + \|\varepsilon^{2+(k-1)_+}[\mathcal{T}^2,\nabla\cdot]\partial_t^k u\|_{2-k}^2$$
$$= \left\|\varepsilon^{4+(k-1)_+}\partial_t^k\mathcal{T}^2(D_t p)\right\|_{2-k}^2 + \left\|\varepsilon^{2+(k-1)_+}[\mathcal{T}^2,\nabla\cdot]\partial_t^k u\right\|_{2-k}^2, \qquad (3.27)$$

where $\|\varepsilon^{2+(k-1)_+}[\mathcal{T}^2,\nabla\cdot]\partial_t^k u\|_{2-k}^2$ can be controlled by $\varepsilon^4 E_4(t) \leq \varepsilon^4 E_4(0) + \int_0^t E_5(\tau)\mathrm{d}\tau$. Indeed, such commutators only appear when we commute the third component of $\nabla\cdot$ with $\omega\partial_3$. We have the following identity which can be shown by induction:

$$[(\omega\partial_3)^k,\partial_3](\cdot) = \sum_{\ell\leq k-1} c_{k,\ell}(\omega\partial_3)^\ell\partial_3(\cdot) = \sum_{\ell\leq k-1} d_{k,\ell}\partial_3(\omega\partial_3)^\ell(\cdot), \qquad (3.28)$$

where $c_{k,\ell}$ and $d_{k,\ell}$ are smooth functions that depend on k, ℓ and the derivatives (up to order k) of ω. For example, we assume $k = 0$ and $\mathcal{T}^2 = (\omega\partial_3)\bar\partial$, for which the highest-order term in the commutator is $\|\varepsilon^2(\partial_3\omega(x_3))\partial_3\bar\partial u_3\|_2 \leq \varepsilon^4 E_4(t)$. This term can be either controlled by $\varepsilon^4 E_4(0) + \int_0^t E_5(\tau)\mathrm{d}\tau$ or absorbed by $E_4(t)$ when ε is sufficiently small. Similar estimates apply to the commutator $\|\varepsilon^{2+(k-1)_+}[\mathcal{T}^2,\nabla\cdot]\partial_t^k B\|_{2-k}^2$ so we omit the details.

The term $\|\varepsilon^{4+(k-1)_+}\partial_t^k\mathcal{T}^2(D_t p)\|_{2-k}^2$ will be further reduced to the tangenital estimates of (u,B) by using the argument in Section 3.2.1. Finally, we need to control the tangential derivatives of u and B (including time derivatives) in $E_5(t)$ and the L^2 norm of $\partial_t^5 p$ with suitable Mach number weights.

Next, we analyze the vorticity term. The proof is parallel to Lemma 3.2.

Lemma 3.3.

$$\sum_{k=0}^2 \frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^{2+(k-1)_+}\nabla\times\mathcal{T}^2\partial_t^k u\right\|_{2-k}^2 + \left\|\varepsilon^{2+(k-1)_+}\nabla\times\mathcal{T}^2\partial_t^k B\right\|_{2-k}^2 + \left\|\varepsilon^{2+(k-1)_++1}B\times\left(\nabla\times\mathcal{T}^2\partial_t^k B\right)\right\|_{2-k}^2\right) \qquad (3.29)$$
$$\leq P(E_4(t),E_5(t),E_6(t)).$$

Proof. For the case of $k = 0$, $\alpha_3 = \alpha_4 = 0$, $\mathcal{T}^\alpha = \partial_t^{\alpha_0}\partial_1^{\alpha_1}\partial_2^{\alpha_2}$ with $\alpha_0 + \alpha_1 + \alpha_2 = 2$, we write $\mathcal{T}^\alpha$ to be $\mathcal{T}^2$ for convenience and take $\partial^2\nabla\times\mathcal{T}^2$ in the momentum equation $\rho D_t u - (B\cdot\nabla)B = -\nabla(p + \frac{1}{2}|B|^2)$ to get

$$\rho D_t(\partial^2\nabla\times\mathcal{T}^2 u) - (B\cdot\nabla)(\partial^2\nabla\times\mathcal{T}^2 B) = [\rho D_t, \partial^2\nabla\times\mathcal{T}^2]u - [B\cdot\nabla, \partial^2\nabla\times\mathcal{T}^2]B. \qquad (3.30)$$

Again, it suffices to prove that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho\left|\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u\right|^2 + \left|\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 B\right|^2 + \left|\varepsilon^3 B\times(\partial^2\nabla\times\mathcal{T}^2 B)\right|^2 \mathrm{d}x \leq P(E_4(t),E_5(t),E_6(t)).$$

We compute that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho|\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u|^2\,\mathrm{d}x$$

$$=\int_\Omega \rho D_t(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\cdot(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\,\mathrm{d}x$$

$$\underbrace{-\int_\Omega \rho(u\cdot\nabla)(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\cdot(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\,\mathrm{d}x-\int_\Omega(\frac{\partial\rho}{\partial p}\partial_t p+\partial_S\rho\partial_t S)|\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u|^2\,\mathrm{d}x}_{:=\mathcal{I}_1'}$$

$$=\int_\Omega \varepsilon^2(B\cdot\nabla)(\partial^2\nabla\times\mathcal{T}^2 B)\cdot(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\,\mathrm{d}x$$

$$\underbrace{+\int_\Omega \varepsilon^2\left([\rho D_t,\partial^2\nabla\times\mathcal{T}^2]u-[B\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]B\right)\cdot(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\,\mathrm{d}x}_{:=\mathcal{I}_2'}+\mathcal{I}_1'. \tag{3.31}$$

Then we integrate $B\cdot\nabla$ by parts and invoke the evolution equation of B, namely $D_t B=(B\cdot\nabla)u-B(\nabla\cdot u)$ to get

$$\int_\Omega \varepsilon^2(B\cdot\nabla)(\partial^2\nabla\times\mathcal{T}^2 B)\cdot(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 u)\,\mathrm{d}x$$

$$=-\int_\Omega \varepsilon^2(\partial^2\nabla\times\mathcal{T}^2 B)\cdot\varepsilon^2(\partial^2\nabla\times\mathcal{T}^2(B\cdot\nabla u))\,\mathrm{d}x-\int_\Omega \varepsilon^2(\partial^2\nabla\times\mathcal{T}^2 B)\cdot\varepsilon^2([B\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]u)\,\mathrm{d}x$$

$$=-\int_\Omega \varepsilon^2(\partial^2\nabla\times\mathcal{T}^2 B)\cdot\varepsilon^2(\partial^2\nabla\times\mathcal{T}^2(D_t B))\,\mathrm{d}x-\int_\Omega \varepsilon^2(\partial^2\nabla\times\mathcal{T}^2 B)\cdot\varepsilon^2(\partial^2\nabla\times\mathcal{T}^2(B\nabla\cdot u))\,\mathrm{d}x$$

$$-\int_\Omega \varepsilon^2(\partial^2\nabla\times\mathcal{T}^2 B)\cdot\varepsilon^2([B\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]u)\,\mathrm{d}x$$

$$=-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega|\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 B|^2\,\mathrm{d}x+\underbrace{\int_\Omega\left(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 B\right)\cdot\left(\varepsilon^4\partial^2\nabla\times\mathcal{T}^2(BD_t p)\right)\,\mathrm{d}x}_{:=\mathcal{K}'}$$

$$\underbrace{-\int_\Omega \varepsilon^2\partial^2\nabla\times\mathcal{T}^2 B\cdot\varepsilon^2\left([\partial^2\nabla\times\mathcal{T}^2,D_t]B+(u\cdot\nabla)(\varepsilon^2\partial^2\nabla\times\mathcal{T}^2 B)+[B\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]u\right)\,\mathrm{d}x}_{:=\mathcal{I}_3'} \tag{3.32}$$

A straightforward product estimate for $\mathcal{I}_1'$ gives us

$$\mathcal{I}_1'\le P(E_4(t),E_5(t)). \tag{3.33}$$

For $\mathcal{I}_2'$, we first consider the commutator $\varepsilon^2[\rho u\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]u$. Since,

$$[\rho u\cdot\nabla,\partial^2\nabla\times\mathcal{T}^2]u=\partial(\rho u_j)\partial^2\mathcal{T}^2\partial_j u_i+\mathcal{T}(\rho u_j)\partial^3\mathcal{T}\partial_j u_i+\text{ lower order terms}, \tag{3.34}$$

we obtain

$$\varepsilon^2\|\partial(\rho u_j)\partial^2\mathcal{T}^2\partial_j u_i\|_0\le\|\partial(\rho u)\|_\infty\|\varepsilon^2\mathcal{T}^2 u\|_3\le P(E_4(t),E_5(t)), \tag{3.35}$$

$$\varepsilon^2\|\mathcal{T}(\rho u_j)\partial^3\mathcal{T}\partial_j u_i\|_0\le\|\mathcal{T}(\rho u)\|_\infty\|\varepsilon^2\mathcal{T}\partial_j u\|_3\le P(E_4(t),E_5(t)),\ \ j=1,2, \tag{3.36}$$

Since $u_3|_\Sigma=0$, using fundamental theorem of calculus, we have $u_3(x_3)=u_3(-1)+\int_{-1}^{x_3}\partial_3 u_3(\xi_3)\mathrm{d}\xi_3$ (W.L.O.G. $x_3<0$), so the length of the interval $[-1,x_3]$ is comparable to the weight function $\omega(x_3)=(1-x_3)(1+x_3)$. We have

$$|\mathcal{T}(\rho u_3)(x_3)|\le 0+\int_{-1}^{x_3}\|\partial_3\mathcal{T}(\rho u_3)(\xi_3)\|_\infty\mathrm{d}\xi_3\lesssim\omega(x_3)\|\partial_3\mathcal{T}(\rho u_3)\|_\infty.$$

Then we can combine this $\omega(x_3)$ with ∂_3 to convert a normal derivative to a tangential derivative.

$$\varepsilon^2\|\mathcal{T}(\rho u_3)(x_3)\partial^3\mathcal{T}\partial_3 u_i\|_0\lesssim\|\partial_3\mathcal{T}(\rho u_3)\|_\infty\|\varepsilon^2\omega(x_3)\partial^3\mathcal{T}\partial_3 u\|_0$$

$$\le\|\partial_3\mathcal{T}(\rho u_3)\|_\infty\|\varepsilon^2\partial^3\mathcal{T}\omega(x_3)\partial_3 u\|_0+P(E_4(t),E_5(t))$$

$$\le P(E_4(t),E_5(t)). \tag{3.37}$$

16

Thus,

$$\|\varepsilon^2[\rho u \cdot \nabla, \partial^2 \nabla \times \mathcal{T}^2]u\|_0 \leq P(E_4(t), E_5(t)). \tag{3.38}$$

Similarly, since $B_3|_\Sigma = 0$,

$$\|\varepsilon^2[\rho \partial_t, \partial^2 \nabla \times \mathcal{T}^2]u\|_0 + \|\varepsilon^2[B \cdot \nabla, \partial^2 \nabla \times \mathcal{T}^2]B\|_0 \leq P(E_4(t), E_5(t)), \tag{3.39}$$

then

$$\mathcal{I}_2' \leq P(E_4(t), E_5(t)). \tag{3.40}$$

We can also get the estimate of $\mathcal{I}_3'$ in a similar way:

$$\mathcal{I}_3' \leq P(E_4(t), E_5(t)). \tag{3.41}$$

For $\mathcal{K}'$, we consider the estimate of $\varepsilon^4 \partial^2 \nabla \times \mathcal{T}^2 (BD_t p)$. Since,

$$
\begin{aligned}
&\varepsilon^4 \partial^2 \nabla \times \mathcal{T}^2 (BD_t p) \\
&= \varepsilon^4 \partial^2 \mathcal{T}^2 (D_t \nabla p \times B) \underbrace{+\varepsilon^4 \partial^2 \mathcal{T}^2 (\nabla u \cdot \nabla p \times B) + \varepsilon^4 \partial^2 \mathcal{T}^2 (D_t p \nabla \times B)}_{:=\mathcal{R}_1'} \\
&= -B \times \varepsilon^4 \partial^2 \mathcal{T}^2 D_t \nabla p \underbrace{-\varepsilon^4 [\partial^2 \mathcal{T}^2, B\times](D_t \nabla p)}_{\mathcal{R}_2'} + \mathcal{R}_1' \\
&= B \times (\varepsilon^4 \partial^2 \mathcal{T}^2 D_t(\rho D_t u)) + B \times \varepsilon^4 \partial^2 \mathcal{T}^2 D_t(B \times (\nabla \times B)) + \mathcal{R}_1' + \mathcal{R}_2' \\
&= B \times (\varepsilon^4 \rho \partial^2 \mathcal{T}^2 D_t^2 u) + \underbrace{\varepsilon^4 B \times D_t \left(B \times \partial^2 \mathcal{T}^2 (\nabla \times B)\right)}_{\mathcal{K}_1'} + \mathcal{R}_1' + \mathcal{R}_2' \\
&\quad + \underbrace{B \times \varepsilon^4([\partial^2 \mathcal{T}^2 D_t, \rho]D_t u) + \varepsilon^2 B \times \left([\partial^2 \mathcal{T}^2, D_t](\nabla \times B) + D_t\left([\partial^2 \mathcal{T}^2, B\times](\nabla \times B)\right)\right)}_{\mathcal{R}_3'}, \tag{3.42}
\end{aligned}
$$

we get that

$$\|\mathcal{R}_1' + \mathcal{R}_2' + \mathcal{R}_3'\|_0 \lesssim P(E_4(t))(\|\varepsilon^2 \mathcal{T}^2 B\|_3 + \|\varepsilon^2 \mathcal{T}^3 B\|_2 + \|\varepsilon^4 \mathcal{T}^2 D_t p\|_2 + \|\varepsilon^4 \mathcal{T}^2 u\|_3). \tag{3.43}$$

It should be noted that when commuting D_t with $\bar{\partial}$ or ∂_t, no extra normal derivative is generated as D_t itself is also a tangential derivative. Indeed, when $j = 3$ in $[D_t, \bar{\partial}_i](\cdot) = -\bar{\partial}_i u_j \partial_j(\cdot)$ for $i = t, 1, 2$, we have $\bar{\partial}_i u_3|_\Sigma = 0$ and so this is still a tangential derivative.

The analysis of $\mathcal{K}_1'$ is also similar to the analysis of $\mathcal{K}_1$. That is, the contribution of $\mathcal{K}_1'$ in $\mathcal{K}'$ is an energy term plus controllable error terms

$$
\begin{aligned}
&\int_\Omega \left(\varepsilon^2 \partial^2 \nabla \times \mathcal{T}^2 B\right) \cdot \left(\varepsilon^4 B \times D_t\left(B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right)\right) \, \mathrm{d}x \\
&= -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\varepsilon^3 B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right|^2 \, \mathrm{d}x + \frac{1}{2}\int_\Omega (\nabla \cdot u)\left|\varepsilon^3 B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right|^2 \, \mathrm{d}x \\
&\lesssim -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\varepsilon^3 B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right|^2 \, \mathrm{d}x + P(E_4(t))E_5(t), \tag{3.44}
\end{aligned}
$$

and thus

$$\mathcal{K}' + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\varepsilon^3 B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right|^2 \, \mathrm{d}x \leq P(E_4(t))\|\varepsilon^4 \mathcal{T}^2 D_t^2 u\|_2 + P(E_4(t), E_5(t)). \tag{3.45}$$

Here we use the fact that $\mathcal{T}^2$ commutes with ∇ when $\alpha_4 = 0$. So we obtain the vorticity estimates for $k = 0$ and $\alpha_4 = 0$

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^2 \nabla \times \mathcal{T}^2 u\right\|_2^2 + \left\|\varepsilon^2 \nabla \times \mathcal{T}^2 B\right\|_2^2 + \left\|\varepsilon^3 B \times (\nabla \times \mathcal{T}^2 B)\right\|_2^2\right) \\
&\leq P(E_4(t), E_5(t))\|\varepsilon^4 \mathcal{T}^2 D_t^2 u\|_2 \leq P(E_4(t), E_5(t), E_6(t)). \tag{3.46}
\end{aligned}
$$

For the case of $\alpha_4 \geq 1$, since $[\partial_3, \omega\partial_3] \neq 0$, we need to reconsider the estimates of commutators. First, we consider commutators: $[\mathcal{T}^2, \nabla\times](\cdot)$ and $[\mathcal{T}^2, \nabla\cdot](\cdot)$. According to the identity (3.28), such commutators contain at most two derivatives and at most one of them is ∂_3. Thus, we have

$$\|\varepsilon^2[\mathcal{T}^2, \nabla\times]u\|_2 + \|\varepsilon^2[\mathcal{T}^2, \nabla\times]B\|_2 + \|\varepsilon^2[\mathcal{T}^2, \nabla\cdot]u\|_2 \leq \varepsilon^2\left(E_4(0) + \int_0^t P(E_4(\tau))\mathrm{d}\tau\right). \tag{3.47}$$

Next, for example, we consider the commutator $\varepsilon^2[B \cdot \nabla, \partial^2\nabla \times (\omega\partial_3)\mathcal{T}]B$ in $\mathcal{I}_2'$ (we take $\alpha_4 = 1$, $\mathcal{T}^\alpha = (\omega\partial_3)\partial_t^{\alpha_0}\partial_1^{\alpha_1}\partial_2^{\alpha_2}$ with $\alpha_0 + \alpha_1 + \alpha_2 = 1$ without loss of generality),

$$[B \cdot \nabla, \partial^2\nabla \times (\omega\partial_3)\mathcal{T}]B = [B_3\partial_3, \omega\partial_3^4\mathcal{T}]B + \text{ lower order terms}$$
$$\overset{L}{=} B_3\partial_3(\omega\partial_3^4\mathcal{T}B) - \omega\partial_3^4\mathcal{T}(B_3\partial_3 B)$$
$$\overset{L}{=} B_3(\partial_3\omega)\partial_3^3\mathcal{T}B, \tag{3.48}$$

where the "lower order terms" represent the terms that do not have the leading order derivative and the terms that $\partial^2\nabla\times$ contributes to tangential derivatives instead of ∂_3^3 as above.

Since $B_3|_\Sigma = 0$, we again use the fundamental theorem of calculus to get

$$\varepsilon^2\|B_3(\partial_3\omega)\partial_3^3\mathcal{T}B\|_0 = \|\varepsilon^2|B_3(x_3)|\partial_3^3\mathcal{T}B\|_0$$
$$\lesssim \|\partial_3 B_3\|_\infty\|\varepsilon^2\omega(x_3)\partial_3^3\mathcal{T}B\|_0$$
$$\leq P(E_4(t)). \tag{3.49}$$

All the other terms in $\mathcal{I}_2'$ and $\mathcal{I}_3'$ can be treated similarly. Finally, we need to consider the term $\mathcal{K}_1'$ because we shall control an extra commutator arising from

$$D_t\left(B \times \partial^2\mathcal{T}^2(\nabla \times B)\right) = D_t\left(B \times \partial^2(\nabla \times \mathcal{T}^2 B)\right) + D_t\left(B \times \partial^2\left([\mathcal{T}^2, \nabla\times]B\right)\right).$$

Again, using (3.28), $[\mathcal{T}^2, \nabla\times]B$ contains at most two derivatives and at most one of these two derivatives is ∂_3. Thus, the highest-order part in the term $D_t\left(B \times \partial^2\left([\mathcal{T}^2, \nabla\times]B\right)\right)$ contains at most 3 normal derivatives and two tangential derivatives, whose $L^2(\Omega)$ can still be bounded by $E_5(t)$ aftering multiplying the weight ε^4.

At this point, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^2\nabla \times \mathcal{T}^2 u\right\|_2^2 + \left\|\varepsilon^2\nabla \times \mathcal{T}^2 B\right\|_2^2 + \left\|\varepsilon^3 B \times (\nabla \times \mathcal{T}^2 B)\right\|_2^2\right)$$
$$\leq P(E_4(t), E_5(t))\|\varepsilon^4\mathcal{T}^4 u\|_2 \leq P(E_4(t), E_5(t), E_6(t)). \tag{3.50}$$

For $k = 1, 2$, we obtain the following energy estimates in similar way as in the proof of Lemma 3.2

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left\|\varepsilon^{2+(k-1)_+}\nabla \times \mathcal{T}^2\partial_t^k u\right\|_{2-k}^2 + \left\|\varepsilon^{2+(k-1)_+}\nabla \times \mathcal{T}^2\partial_t^k B\right\|_{2-k}^2 + \left\|\varepsilon^{2+(k-1)_++1}B \times (\nabla \times \mathcal{T}^2\partial_t^k B)\right\|_{2-k}^2\right)$$
$$\leq P(E_4(t), E_5(t), E_6(t)). \tag{3.51}$$

The proof of Lemma 3.3 is completed. $\qquad\square$

The proof of Lemma 3.3 shows that we need to trade one normal derivative for two more tangentials with Mach number weight ε^2, that is, we need to control $\|\varepsilon^4\mathcal{T}^4(\partial_t^k u, \partial_t^k B)\|_{2-k}$ ($k = 0, 1, 2$). We again do the analogous div-curl analysis for the case $k = 0, 1$ and reduce it to $\|\varepsilon^6\mathcal{T}^6(\partial_t^k u, \partial_t^k B)\|_{1-k}$, and then repeat once more such that we finally need to control $\|\varepsilon^8\mathcal{T}^8(u, B)\|_0$ which is a purely tangential estimate. To sum up, we can obtain the following estimates by mimicking the proof of Lemma 3.2 and Lemma 3.3.

Lemma 3.4. The div-curl analysis and vorticity estimates for $\|\varepsilon^4\mathcal{T}^4(u, B)\|_2$, $\|\varepsilon^4\partial_t\mathcal{T}^4(u, B)\|_1$ and $\|\varepsilon^6\mathcal{T}^6(u, B)\|_1$

are listed below:

$$\|\varepsilon^4 \mathcal{T}^4 \partial_t^k u\|_{2-k}^2 \lesssim \|\varepsilon^4 \mathcal{T}^4 \partial_t^k u\|_0^2 + \|\varepsilon^4 \nabla \times \mathcal{T}^4 \partial_t^k u\|_{1-k}^2 + \|\varepsilon^4 \nabla \cdot \mathcal{T}^4 \partial_t^k u\|_{1-k}^2, \ k = 0, 1, \tag{3.52}$$

$$\|\varepsilon^4 \mathcal{T}^4 \partial_t^k B\|_{2-k}^2 \lesssim \|\varepsilon^4 \mathcal{T}^4 \partial_t^k B\|_0^2 + \|\varepsilon^4 \nabla \times \mathcal{T}^4 \partial_t^k B\|_{1-k}^2 + \|\varepsilon^{4+(k-1)_+}[\mathcal{T}^4, \nabla\cdot] \partial_t^k B\|_{1-k}^2, \ k = 0, 1, \tag{3.53}$$

$$\sum_{k=0,1} \frac{\mathrm{d}}{\mathrm{d}t} \left(\left\|\varepsilon^4 \nabla \times \partial_t^k \mathcal{T}^4 u\right\|_{1-k}^2 + \left\|\varepsilon^4 \nabla \times \partial_t^k \mathcal{T}^4 B\right\|_{1-k}^2 + \left\|\varepsilon^5 B \times (\nabla \times \partial_t^k \mathcal{T}^4 B)\right\|_{1-k}^2 \right)$$

$$\leq P(E_4(t), E_5(t), E_6(t)) \|\varepsilon^6 \mathcal{T}^6 \partial_t^k u\|_{1-k}, \tag{3.54}$$

$$\|\varepsilon^6 \mathcal{T}^6 u\|_1^2 \lesssim \|\varepsilon^6 \mathcal{T}^6 u\|_0^2 + \|\varepsilon^6 \nabla \times \mathcal{T}^6 u\|_0^2 + \|\varepsilon^6 \nabla \cdot \mathcal{T}^6 u\|_0^2, \tag{3.55}$$

$$\|\varepsilon^6 \mathcal{T}^6 B\|_1^2 \lesssim \|\varepsilon^6 \mathcal{T}^6 B\|_0^2 + \|\varepsilon^6 \nabla \times \mathcal{T}^6 B\|_0^2 + \|\varepsilon^6 [\mathcal{T}^6, \nabla\cdot] B\|_0^2, \tag{3.56}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\left\|\varepsilon^6 \nabla \times \mathcal{T}^6 u\right\|_0^2 + \left\|\varepsilon^6 \nabla \times \mathcal{T}^6 B\right\|_0^2 + \left\|\varepsilon^7 B \times (\nabla \times \mathcal{T}^6 B)\right\|_0^2 \right) \leq P(E_4(t) E_5(t), E_6(t), E_7(t)) \|\varepsilon^8 \mathcal{T}^8 u\|_0. \tag{3.57}$$

Besides, the divergence part is reduced to the estimates of p by using the continuity equation

$$\|\varepsilon^4 \partial_t^k \nabla \cdot \mathcal{T}^4 u\|_{1-k}^2 = \left\|\varepsilon^6 \partial_t^k \mathcal{T}^4 (D_t p)\right\|_{1-k}^2 + \|\varepsilon^4 [\mathcal{T}^4, \nabla\cdot] \partial_t^k u\|_{1-k}^2, \quad k = 0, 1,$$

$$\|\varepsilon^6 \nabla \cdot \mathcal{T}^6 u\|_0^2 = \left\|\varepsilon^8 \mathcal{T}^6 (D_t p)\right\|_0^2 + \|\varepsilon^6 [\mathcal{T}^6, \nabla\cdot] u\|_0^2, \tag{3.58}$$

where the commutators can be controlled

$$\sum_{k=0,1} \|\varepsilon^4 [\mathcal{T}^4, \nabla\cdot] \partial_t^k u\|_{1-k}^2 \leq \varepsilon^4 E_5(t) \lesssim \varepsilon^4 E_5(0) + \int_0^t E_6(\tau) \mathrm{d}\tau, \tag{3.59}$$

$$\|\varepsilon^6 [\mathcal{T}^6, \nabla\cdot] u\|_0^2 \leq \varepsilon^4 E_6(t) \lesssim \varepsilon^4 E_6(0) + \int_0^t E_7(\tau) \mathrm{d}\tau. \tag{3.60}$$

by using the identity (3.28). Moreover, $\left\|\varepsilon^6 \partial_t^k \mathcal{T}^4 (D_t p)\right\|_{1-k}^2$ and $\left\|\varepsilon^8 \mathcal{T}^6 (D_t p)\right\|_0^2$ will be further reduced to the tangenital estimates of (u, B) by using the argument in Section 3.2.1.

3.2.3 The remaining tangential estimates

Recall that the entropy S is directly controlled via $D_t S = 0$, so we can temporarily ignore that. Let us summarize what kind of tangential derivatives we shall control in order to close the estimates that are uniform in Mach number ε. Again we start with $\|\varepsilon^{(k-1)_+} (\partial_t^k u, \partial_t^k B)\|_{4-k}$ in $E_4(t)$. We set such Mach number weights based on the following facts:

a. H^4 regularity: Several commutator estimates require the bound for $\|\bar{\partial}^2 (u, B)\|_{L^\infty}$. Recall that $H^{\frac{d}{2}+\delta} \hookrightarrow L^\infty$, so it would be convenient to choose $H^{2+\lceil \frac{d}{2}+\delta \rceil}$ regularity, that is, H^4 for $d = 2, 3$.
b. The initial data is well-prepared, so $\|\nabla p(0)\|_3$ should be bounded with respect to ε and we shall add Mach number weights to $\partial_t^k p$ when $k \geq 1$.
c. The L^2 estimate in section 3.1 suggests that $u, B, \varepsilon p$ should share the same weights of Mach number ε because taking tagential derivatives preserves the boundary conditions.
d. The reduction procedure in section 3.2.1 shows that ∇p is converted to $\mathcal{T}(u, B)$ and ∇B. When $\mathcal{T}$ is a spatial derivative, $\mathcal{T}(u, B)$ and ∇B should be controlled via div-curl analysis as shown in Lemma 3.2. When $\mathcal{T} = \partial_t$, we find that $\partial_t^k p$, $\partial_t^{k+1}(u, B)$ should share the same weights of Mach number.

Thus, $E_4(t)$ consists of the following quantities.

$$\|u, B, S, p\|_4, \ \|\partial_t u, \partial_t B, \partial_t S, \varepsilon \partial_t p\|_3, \ \|\varepsilon \partial_t^2 u, \varepsilon \partial_t^2 B, \varepsilon \partial_t^2 S, \varepsilon^2 \partial_t^2 p\|_2,$$
$$\|\varepsilon^2 \partial_t^3 u, \varepsilon^2 \partial_t^3 B, \varepsilon^2 \partial_t^3 S, \varepsilon^3 \partial_t^3 p\|_1, \ \|\varepsilon^3 \partial_t^4 u, \varepsilon^3 \partial_t^4 B, \varepsilon^3 \partial_t^4 S, \varepsilon^4 \partial_t^4 p\|_0.$$

The Sobolev norms of $\|\varepsilon^{(k-1)_+} (\partial_t^k u, \partial_t^k B)\|_{4-k}$ for $0 \leq k \leq 3$ are controlled via div-curl analysis. The divergence part can be absorbed by $E_4(t)$ itself because the continuity equation produces extra ε^2 weight and $\varepsilon^2 < \varepsilon$ since we assume the Mach number is small. Even if we do not assume ε is small, we can alternatively control the divergence by repeatedly using the reduction of pressure such that there are only tangential derivatives. The curl part, thanks to the special structure of Lorentz force, is reduced to $\|\varepsilon^{(k-1)_+ +2} D_t^2 \partial_t^k u\|_{3-k}$ which

is a part of $E_5(t)$. So, it remains to control $E_5(t)$ and the full time derivatives, namely $\|\varepsilon^3\partial_t^4 u, \varepsilon^3\partial_t^4 B, \varepsilon^4\partial_t^4 p\|_0$. See also diagram (2.8).

Then $E_5(t)$ is designed in the same manner as $E_4(t)$:

$\|\varepsilon^2\mathcal{T}^2(u,B,S,p)\|_3$ (produced from the vorticity analysis of $\|u,B\|_4$),

$\|\varepsilon^2\mathcal{T}^2(\partial_t u,\partial_t B,\partial_t S,\varepsilon\partial_t p)\|_2$ (produced from the vorticity analysis of $\|\partial_t u,\partial_t B\|_3$),

$\|\varepsilon^2\mathcal{T}^2(\varepsilon\partial_t^2 u,\varepsilon\partial_t^2 B,\varepsilon\partial_t^2 S,\varepsilon^2\partial_t^2 p)\|_1$ (produced from the vorticity analysis of $\|\varepsilon\partial_t^2 u,\varepsilon\partial_t^2 B\|_2$),

$\|\varepsilon^2\mathcal{T}^2(\varepsilon^2\partial_t^3 u,\varepsilon^2\partial_t^3 B,\varepsilon^2\partial_t^3 S,\varepsilon^3\partial_t^3 p)\|_0$ (produced from the vorticity analysis of $\|\varepsilon^2\partial_t^3 u,\varepsilon^2\partial_t^3 B\|_1$).

The div-curl analysis of $E_5(t)$ produces $\|\varepsilon^{(k-1)_++4}\mathcal{T}^2 D_t^2\partial_t^k u\|_{2-k}$ which is part of $E_6(t)$. In order to close the energy estimates, we now need to control $E_6(t)$ and also the full time derivatives $\|\varepsilon^2\mathcal{T}^2(\varepsilon\partial_t^3 u,\varepsilon\partial_t^3 B,\varepsilon^2\partial_t^3 p)\|_0$.

Repeat the above procedure once more, we can see $E_6(t)$ should be designed as

$\|\varepsilon^4\mathcal{T}^4(u,B,S,p)\|_2$ (produced from the vorticity analysis of $\|\varepsilon^2\mathcal{T}^2(u,B)\|_3$),

$\|\varepsilon^4\mathcal{T}^4(\partial_t u,\partial_t B,\partial_t S,\varepsilon\partial_t p)\|_1$ (produced from the vorticity analysis of $\|\varepsilon^2\mathcal{T}^2(\partial_t u,\partial_t B)\|_2$,

$\|\varepsilon^4\mathcal{T}^4(\varepsilon\partial_t^2 u,\varepsilon\partial_t^2 B,\varepsilon\partial_t^2 S,\varepsilon^2\partial_t^2 p)\|_0$ (produced from the vorticity analysis of $\|\varepsilon^2\mathcal{T}^2(\varepsilon\partial_t^2 u,\varepsilon\partial_t^2 B)\|_1$).

Combining the result of vorticity estimate, it remains to control $\|\varepsilon^6\mathcal{T}^4 D_t^2\partial_t^k u\|_{1-k}$ (part of $E_7(t)$) and also the full time derivatives $\|\varepsilon^4\mathcal{T}^4(\varepsilon\partial_t^2 u,\varepsilon\partial_t^2 B,\varepsilon\partial_t^2 S,\varepsilon^2\partial_t^2 p)\|_0$.

Repeat the above procedure once more, we can see $E_7(t)$ should be designed as

$\|\varepsilon^6\mathcal{T}^6(u,B,S,p)\|_1$ (produced from the vorticity analysis of $\|\varepsilon^4\mathcal{T}^4(u,B)\|_2$),

$\|\varepsilon^6\mathcal{T}^6(\partial_t u,\partial_t B,\partial_t S,\varepsilon\partial_t p)\|_0$ (produced from the vorticity analysis of $\|\varepsilon^4\mathcal{T}^4(\partial_t u,\partial_t B)\|_1$ and (c).

Discussion about the weights of p. For the pressure p, we only have the control of $\|\varepsilon^{k+2l}\mathcal{T}^\alpha\partial_t^k p\|_{4-k-l}$ at this step, which has one more ε-weight than what we defined in (1.18) when $k \geq 1$, $l \geq 1$ and $\alpha_0 < 2l$. To replace ε^{k+2l} by $\varepsilon^{(k-1)_++2l}$ when $l \geq 1$ and $\alpha_0 < 2l$, we just need to notice that $\mathcal{T}^\alpha$ now must contain at least one tangential spatial derivative, and then we can use the momentum equation

$$-\partial_i p = B_k\partial_i B_k + \rho D_t u_i - (B\cdot\nabla)B_i, \quad -\omega\partial_3 p = B_k\underbrace{\omega\partial_3 B_k}_{\text{tangential}} +\omega\rho D_t u_3 - (\omega B\cdot\nabla)B_3$$

to reduce the control of $\|\varepsilon^{(k-1)_++2l}\mathcal{T}^\alpha\partial_t^k p\|_{4-k-l}$ ($\langle\alpha\rangle = 2l$, $\alpha_0 < 2l$) to $\|\varepsilon^{(k-1)_++2l}\mathcal{T}^\beta\partial_t^k(u,B)\|_{4-k-l}$ for some β satisfying $\langle\beta\rangle = 2l$ and $\alpha_0 \leq \beta_0 \leq \alpha_0 + 1$, which has been included in $E_{4+l}(t)$.

Combining the result of vorticity estimate, it remains to control $\|\varepsilon^6\mathcal{T}^\alpha(\partial_t u,\partial_t B,\partial_t p)\|_0$ for $\langle\alpha\rangle = 6$, $\alpha_0 < 6$, $\|\varepsilon^6(\partial_t^6 u,\partial_t^6 B,\varepsilon\partial_t^7 p)\|_0$ and $\|\varepsilon^8\mathcal{T}^\alpha D_t^2 u\|_0$ for $\langle\alpha\rangle = 6$. This indicates us to define $E_8(t)$ to be

$$E_8(t) = \sum_{\langle\alpha\rangle=8,\alpha_0<8}\left\|\varepsilon^8\mathcal{T}^\alpha(u,B,S,p)\right\|_0^2 + \left\|\varepsilon^8\partial_t^8(u,B,S,\varepsilon p)\right\|_0^2.$$

To sum up, the remaining estimates are all tangential estimates, which can be proved by analyzing the L^2 estimates after taking the following tangential derivatives to MHD system (1.13)

$$\varepsilon^3\partial_t^4, \ \varepsilon^4\mathcal{T}^2\partial_t^3, \ \varepsilon^5\mathcal{T}^4\partial_t^2, \ \varepsilon^6\mathcal{T}^6\partial_t, \ \varepsilon^8\mathcal{T}^8. \tag{3.61}$$

3.3 Tangential estimates

3.3.1 The $\mathcal{T}^\alpha$-differentiated equations

By the div-curl analysis, the crucial step is to study the higher order tangential energy estimate of (1.13). In particular, we define the following tangential derivatives

$$\mathcal{T}_0 = \partial_t, \quad \mathcal{T}_1 = \partial_1, \quad \mathcal{T}_2 = \partial_2, \quad \mathcal{T}_4 = \omega(x_3)\partial_3. \tag{3.62}$$

We take $\mathcal{T}^\alpha$ in the momentum equation $\rho D_t u = -\nabla p - B \times (\nabla \times B)$ to get

$$\rho D_t(\mathcal{T}^\alpha u) - (B \cdot \nabla)(\mathcal{T}^\alpha B) + \mathcal{T}^\alpha \nabla(p + |B|^2/2) = [\rho D_t, \mathcal{T}^\alpha]u - [B \cdot \nabla, \mathcal{T}^\alpha]B, \qquad (3.63)$$

where $\mathcal{T}^\alpha := \mathcal{T}_0^{\alpha_0} \mathcal{T}_1^{\alpha_1} \mathcal{T}_2^{\alpha_2} \mathcal{T}_4^{\alpha_4}$ and $\langle\alpha\rangle = \alpha_0 + \alpha_1 + \alpha_2 + 0 \times 2 + \alpha_4 \leq 8$. Accoding to section 3.2.3, it suffices to consider the following cases

- $\langle\alpha\rangle = \alpha_0 = 4$ with ε^3 weight,
- $\langle\alpha\rangle = 5$, $\alpha_0 \geq 3$ with ε^4 weight,
- $\langle\alpha\rangle = 6$, $\alpha_0 \geq 2$ with ε^5 weight,
- $\langle\alpha\rangle = 7$, $\alpha_0 \geq 1$ with ε^6 weight,
- $\langle\alpha\rangle = 8$ with ε^8 weight.

Since these derivatives are all tangential, taking either of them still preserves the boundary conditions. For simplicity, we will only show the details of the following cases (either full spatial derivatives or full time derivatives)

$$\varepsilon^8 \mathcal{T}^\alpha \text{ with } \langle\alpha\rangle = 8 \text{ and } \alpha_0 = 0, \quad \varepsilon^8 \partial_t^8, \quad \varepsilon^{k-1}\partial_t^k \ (4 \leq k \leq 7).$$

and the other cases (space-time mixed derivatives) can be proven in the same manner.

3.3.2 Tangential energy estimate with full spatial derivatives

In this subsection we study the spatially-differentiated equations, i.e., the equations obtained by commuting $\mathcal{T}^\alpha$, $\alpha_0 = 0$, and $\langle\alpha\rangle = 8$, with (1.13). We aim to prove the following estimate

Proposition 3.5. For $\mathcal{T}^\alpha$ with multi-index α satisfying $\alpha_0 = 0$ and $\langle\alpha\rangle = 8$, we have the energy inequality:

$$\sum_{\langle\alpha\rangle=8,\alpha_0=0} \frac{d}{dt} \left\| \varepsilon^8 \mathcal{T}^\alpha(u, B, p) \right\|_0^2 \leq P(E(t)). \qquad (3.64)$$

Proof. We first consider the case of $\alpha_4 = 0$, that is, $\mathcal{T}^\alpha = \overline{\partial}_1^{\alpha_1} \overline{\partial}_2^{\alpha_2}$ with $\alpha_1 + \alpha_2 = 8$. To simplify our notation, we will write $\overline{\partial}^8$ to represent such $\mathcal{T}^\alpha$.

$$\frac{1}{2}\frac{d}{dt} \int_\Omega \rho|\varepsilon^8\overline{\partial}^8 u|^2 \, dx$$

$$= \int_\Omega \varepsilon^{16}\rho\overline{\partial}^8\partial_t u \cdot \overline{\partial}^8 u \, dx + \frac{1}{2} \int_\Omega \partial_t\rho|\varepsilon^8\overline{\partial}^8 u|^2 \, dx$$

$$= \int_\Omega \varepsilon^8\overline{\partial}^8(\rho D_t u) \cdot (\varepsilon^8\overline{\partial}^8 u) \, dx - \underbrace{\int_\Omega \varepsilon^{16}\left(\rho\overline{\partial}^8(u\cdot\nabla u) + [\overline{\partial}^8, \rho]D_t u + (\frac{\partial\rho}{\partial p}\partial_t p + \partial_S\rho\partial_t S)\overline{\partial}^8 u\right) \cdot \overline{\partial}^8 u \, dx}_{:=\mathcal{J}_1}. \quad (3.65)$$

The first integral above, after invoking the momentum equation and integrating $B \cdot \nabla$ by parts, is equal to

$$\int_\Omega \varepsilon^8\overline{\partial}^8(B\cdot\nabla B) \cdot (\varepsilon^8\overline{\partial}^8 u) \, dx - \int_\Omega \varepsilon^{2k}\overline{\partial}^8(\nabla(p + |B|^2/2)) \cdot (\varepsilon^{2k}\overline{\partial}^{4+k} u) \, dx. \qquad (3.66)$$

For the first term, we integrate $B \cdot \nabla$ by parts

$$\int_\Omega \varepsilon^8\overline{\partial}^8(B\cdot\nabla B) \cdot (\varepsilon^8\overline{\partial}^8 u) \, dx$$

$$\overset{B\cdot\nabla}{=} -\int_\Omega \varepsilon^8\overline{\partial}^8 B \cdot (\varepsilon^8\overline{\partial}^8(B\cdot\nabla u)) \, dx + \underbrace{\int_\Omega \varepsilon^8[\overline{\partial}^8, B\cdot\nabla]B \cdot (\varepsilon^8\overline{\partial}^8 u) \, dx - \int_\Omega \varepsilon^8\overline{\partial}^8 B \cdot (\varepsilon^8[B\cdot\nabla, \overline{\partial}^8]u) \, dx}_{:=\mathcal{J}_2}. \quad (3.67)$$

For the second term, we integrate by parts and invoke the continuity equation to get

$$-\int_\Omega \varepsilon^8\overline{\partial}^8(\nabla(p + |B|^2/2)) \cdot (\varepsilon^8\overline{\partial}^8 u) \, dx = \int_\Omega \varepsilon^8\overline{\partial}^8(p + |B|^2/2) \cdot (\varepsilon^8\overline{\partial}^8\nabla \cdot u) \, dx$$

$$= \int_\Omega \varepsilon^8\overline{\partial}^8(|B|^2/2) \cdot (\varepsilon^8\overline{\partial}^8\nabla \cdot u) \, dx - \int_\Omega \varepsilon^8\overline{\partial}^8 p \cdot (\varepsilon^8\overline{\partial}^8(\varepsilon^2 D_t p)) \, dx, \qquad (3.68)$$

where the boundary intergral on Σ vanishes thanks to the boundary condition $u_3 = B_3 = 0$ on Σ.

The first integral in (3.67), after inserting the evolution equation of B, is equal to

$$-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega |\varepsilon^8\overline{\partial}^8 B|^2\,\mathrm{d}x \underbrace{-\int_\Omega \varepsilon^8\overline{\partial}^8 B\cdot((u\cdot\nabla)\overline{\partial}^8 B + [\overline{\partial}^8, u\cdot\nabla]B)\,\mathrm{d}x}_{:=\mathcal{J}_3}$$

$$\underbrace{-\int_\Omega \varepsilon^8\overline{\partial}^8 B\cdot(\varepsilon^8\overline{\partial}^8(B\nabla\cdot u))\,\mathrm{d}x}_{:=\mathcal{K}_1}, \tag{3.69}$$

which gives the energy of B. The last term on the right side of (3.68) gives the energy of εp

$$-\int_\Omega \varepsilon^8\overline{\partial}^8 p\cdot(\varepsilon^8\overline{\partial}^8(\varepsilon^2 D_t p))\,\mathrm{d}x$$

$$= -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\varepsilon^8\overline{\partial}^8(\varepsilon p)\right|^2\,\mathrm{d}x \underbrace{-\int_\Omega \varepsilon^8\overline{\partial}^8 p\cdot(\varepsilon^{10}[\overline{\partial}^8, u\cdot\nabla]p)\,\mathrm{d}x}_{:=\mathcal{J}_4}. \tag{3.70}$$

Then we analyze the first term on the right side of (3.68). Since $\overline{\partial}(\frac{1}{2}|B|^2) = \overline{\partial}B\cdot B$, we have

$$\int_\Omega \varepsilon^8\overline{\partial}^8(|B|^2/2)\cdot(\varepsilon^8\overline{\partial}^8\nabla\cdot u)\,\mathrm{d}x$$

$$= \underbrace{\int_\Omega (\varepsilon^8 B)\cdot(\overline{\partial}^8 B)(\varepsilon^8\overline{\partial}^8\nabla\cdot u)\,\mathrm{d}x}_{:=\mathcal{K}_2} + \underbrace{\sum_{0<\langle\alpha\rangle<8}\int_\Omega \varepsilon^8 C_\alpha(\overline{\partial}^\alpha B)\cdot(\overline{\partial}^{8-\alpha}B)(\varepsilon^8\overline{\partial}^8\nabla\cdot u)\,\mathrm{d}x}_{:=\mathcal{J}_5} \tag{3.71}$$

where C_α are some positive constants. Now let us control the commutators $\mathcal{J}_1 \sim \mathcal{J}_5$. For $\mathcal{J}_1 \sim \mathcal{J}_4$, since $u_3 = B_3 = 0$ on Σ, it suffices to analyze one of the following two types of commutators

$$\varepsilon^8[\overline{\partial}^8, u\cdot\nabla]f, \quad \varepsilon^8[\overline{\partial}^8, B\cdot\nabla]f.$$

For example. we expand the first one to find that

$$\varepsilon^8[\overline{\partial}^8, u\cdot\nabla]f = \varepsilon^8\overline{\partial}^8 u\cdot\nabla f + \sum_{j=1,2}\varepsilon^8(\overline{\partial}\bar{u}_j)(\overline{\nabla}_j\overline{\partial}^7 f) + \varepsilon^8(\overline{\partial}u_3)(\partial_3\overline{\partial}^7 f) + \sum_{k=2}^7\binom{8}{k}\varepsilon^8\overline{\partial}^k u_j\partial_j\overline{\partial}^{8-j}f,$$

and it is easy to see that the last term is directly controlled by $E(t)$. For the first term, we have

$$\|\varepsilon^8\overline{\partial}^8 u\cdot\nabla f + \varepsilon^8(\overline{\partial}\bar{u})\cdot(\overline{\nabla}\overline{\partial}^7 f)\|_0 \le \|\varepsilon^8\overline{\partial}^8 u\|_0\|\nabla f\|_\infty + \|\overline{\partial}u\|_{L^\infty}\|\varepsilon^8\overline{\partial}^8 f\|_0,$$

which is directly controlled by $E(t)$. For the third term, note that $\overline{\partial}u_3|_\Sigma = 0$, using fundamental theorem of calculus, we have

$$|\overline{\partial}u_3(x_3)| \lesssim \omega(x_3)\|\overline{\partial}\partial_3 u_3\|_{L^\infty},$$

and thus

$$\|\varepsilon^8(\overline{\partial}u_3)(\partial_3\overline{\partial}^7 f)\|_0 \lesssim \|\overline{\partial}\partial_3 u_3\|_{L^\infty}\|\varepsilon^8(\omega\partial_3)\overline{\partial}^7 f\|_0$$

which is again controlled by $E(t)$. As for $\mathcal{J}_5$, we integrate $\overline{\partial}$ by parts to get

$$\mathcal{J}_5 \lesssim \sum_{0<\langle\alpha\rangle<8}\|\varepsilon^8\overline{\partial}^{\alpha+1}B\cdot\overline{\partial}^{8-\alpha}B)\|_0\|\varepsilon^8\overline{\partial}^7(\nabla\cdot u)\|_0 \lesssim E(t)\|\varepsilon^{10}\overline{\partial}^7 D_t p\|_0 \le P(E(t)).$$

Thus, all the commutators $\mathcal{J}_i$ ($i = 1,\cdots,5$) are directly controlled

$$\mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \mathcal{J}_4 + \mathcal{J}_5 \le P(E(t)). \tag{3.72}$$

22

Next, we show that $\mathcal{K}_1$ has cancellation with $\mathcal{K}_2$. We have

$$\mathcal{K}_1 = -\sum_{i,j=1}^{3} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot (\varepsilon^8 \overline{\partial}^8 (B_i \partial_j u_j)) \, \mathrm{d}x$$

$$= \underbrace{-\int_\Omega \varepsilon^8 B \cdot \overline{\partial}^8 B(\varepsilon^8 \overline{\partial}^8 \nabla \cdot u) \, \mathrm{d}x}_{=-\mathcal{K}_2} - \sum_{i,j=1}^{3} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot (\varepsilon^8 \overline{\partial} B_i \overline{\partial}^7 \partial_j u_j) \, \mathrm{d}x$$

$$\underbrace{-\sum_{i,j=1}^{3} \sum_{2 \le \langle \alpha \rangle \le 8} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot (\varepsilon^8 C_\alpha \overline{\partial}^\alpha B_i \overline{\partial}^{8-\alpha} \partial_j u_j) \, \mathrm{d}x}_{\mathcal{J}_6}.\tag{3.73}$$

The last term $\mathcal{J}_6$ is again directly controlled: $\mathcal{J}_6 \le P(E(t))$. The first term cancels with $\mathcal{K}_2$. For the second term, we need to further analyze the case when $j = 3$. Thanks to the continuity equation, we have $-\partial_3 u_3 = \varepsilon^2 D_t p + \overline{\nabla} \cdot \bar{u}$ and thus

$$-\sum_{i,j=1}^{3} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot (\varepsilon^8 \overline{\partial} B_i \overline{\partial}^7 \partial_3 u_3) \, \mathrm{d}x = \sum_{i=1}^{3} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot \left(\varepsilon^8 \overline{\partial} B_i \overline{\partial}^7 (\varepsilon^2 D_t p + \overline{\nabla} \cdot \bar{u}) \right) \, \mathrm{d}x \lesssim P(E(t)).\tag{3.74}$$

When $j = 1, 2$, the second term is again directly controlled by $P(E(t))$ because all the derivatives are tangential.

$$-\sum_{i=1}^{3} \sum_{j=1,2} \int_\Omega \varepsilon^8 \overline{\partial}^8 B_i \cdot (\varepsilon^8 \overline{\partial} B_i \overline{\partial}^7 \overline{\partial}_j \bar{u}_j) \, \mathrm{d}x \le P(E(t)).\tag{3.75}$$

Thus, we conclude that

$$\mathcal{K}_1 + \mathcal{K}_2 \le P(E(t)).\tag{3.76}$$

Combining (3.66)-(3.76), we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\| \varepsilon^8 \overline{\partial}^8 (\varepsilon p, u, B) \right\|_0^2 \le P(E(t)).\tag{3.77}$$

To control $\|\varepsilon^8 \overline{\partial}^8 p\|$, it suffices to invoke the momentum equation $-\overline{\partial}_i p = B \cdot (\overline{\partial}_i B) + \rho D_t \bar{u}_i - (B \cdot \nabla) B$ and thus we convert $\varepsilon^8 \overline{\partial}^8 p$ to $\varepsilon^8 \overline{\partial}^7 \mathcal{T}(u, B)$ which is part of $E_8(t)$. We then conclude that

$$\left\| \varepsilon^8 \overline{\partial}^8 (p, u, B)(t) \right\|_0^2 \le E_8(0) + \int_0^t P(E(\tau)) \mathrm{d}\tau.\tag{3.78}$$

For the case of $\alpha_4 \ge 1$, since $[\partial_3, \omega \partial_3] \ne 0$, we need to reconsider the estimates of commutators. For the commutators of type $[\mathcal{T}^\alpha, \partial_3](\cdot)$, we can use identity (3.28). For example, we need to control the extra terms when commuting $\mathcal{T}^\alpha$ with ∇:

$$-\int_\Omega (\varepsilon^8 \mathcal{T}^\alpha u_i)(\varepsilon^8 [\mathcal{T}^\alpha, \partial_i] P) \, \mathrm{d}x \quad \text{and} \quad \int_\Omega (\varepsilon^8 [\partial_i, \mathcal{T}^\alpha] u_i)(\varepsilon^8 \mathcal{T}^\alpha P) \, \mathrm{d}x.$$

Without loss of generality, we assume $\mathcal{T}^\alpha = (\omega(x_3) \partial_3) \overline{\partial}^7$. In the first integral above, the highest-order term should be

$$-\int_\Omega (\varepsilon^8 \mathcal{T}^\alpha u_3)(\varepsilon^8 \, \partial_3 \omega \, \partial_3 \overline{\partial}^7 P) \, \mathrm{d}x.$$

Now we invoke the momentum equation to replace $-\partial_3 P$ with $\rho D_t u_3 - (B \cdot \nabla) B_3$. Recall that D_t and $B \cdot \nabla$ are both tangential derivatives thanks to $u_3|_\Sigma = B_3|_\Sigma = 0$. So the above integral is controlled by $\|\varepsilon^8 \mathcal{T}^\alpha u\|_0 (\|\varepsilon^8 \mathcal{T}^{\alpha'} u\|_0 + \|\varepsilon^8 \mathcal{T}^{\alpha''} B\|_0)$ plus lower-order terms where $\langle \alpha' \rangle = \langle \alpha'' \rangle = 8$. Similarly, for the second integral above, one can invoke the continuity equation, which reads $\partial_3 u_3 = -\overline{\nabla} \cdot \bar{u} - \varepsilon^2 D_t p$, to finish the

23

control. A slight difference is that we do not have the energy for $\|\varepsilon^8 \mathcal{T}^\alpha p\|_0$. Luckily, at this step, $\mathcal{T}^\alpha$ must contain a spatial derivative $\omega(x_3)\partial_3$. So, it suffices to again invoke the momentum equation

$$-\omega\partial_3 p = \underbrace{\omega\partial_3}_{\text{tangential}} (\frac{1}{2}|B|^2) + \omega\rho D_t u_3 - \omega(B \cdot \nabla B_3)$$

such that all terms that contain the fluid pressure are replaced by the velocity and the magnetic field.

As for the other commutators, for example, we re-consider $\varepsilon^8[\omega\partial_3\overline{\partial}^7, B \cdot \nabla]B$ in the analogue of $\mathcal{J}_2$ (W.L.O.G. $\alpha_4 = 1$). We have

$$[\omega\partial_3\overline{\partial}^7, B \cdot \nabla]B = [\omega\partial_3\overline{\partial}^7, B_3\partial_3]B + [\omega\partial_3\overline{\partial}^7, \bar{B} \cdot \overline{\nabla}]B$$
$$= \omega\partial_3\overline{\partial}^7(B_3\partial_3 B) - B_3\partial_3(\omega\partial_3\overline{\partial}^7 B) + [\omega\partial_3\overline{\partial}^7, \bar{B} \cdot \overline{\nabla}]B. \tag{3.79}$$

Compared with the case of $\alpha_4 = 0$, the potential risk is that ∂_3 may fall on $\omega(x_3)$ and then converts the tangential derivative $\omega\partial_3$ to be a normal derivative ∂_3. So, it suffices to analyze the term $-B_3(\partial_3\omega)(\partial_3\overline{\partial}^7 B)$. Luckily, there is a B_3 (replaced by u_3 if we commute $\mathcal{T}^\alpha$ with $u \cdot \nabla$). Since $B_3|_\Sigma = 0$, we again use the fundamental theorem of calculus to obtain

$$|B_3(x_3)| \leq \omega(x_3)\|\partial_3 B_3\|_\infty,$$

and thus

$$\varepsilon^8\|B_3(\partial_3\omega)(\partial_3\overline{\partial}^7 B)\|_0 \lesssim \|\partial_3 B_3\|_\infty\|\varepsilon^8(\omega\partial_3)\overline{\partial}^7 B\|_0 \leq E(t). \tag{3.80}$$

Other terms arise from commutators can be treated similarly. Therefore, we obtain the energy estimate

$$\sum_{\langle\alpha\rangle=8,\alpha_0=0} \frac{\mathrm{d}}{\mathrm{d}t} \left\|\varepsilon^8\mathcal{T}^\alpha(u, B, p)\right\|_0^2 \leq P(E(t)). \tag{3.81}$$

The proof is completed. □

3.3.3 Tangential energy estimate with full time derivatives

In this subsection we study the time-differentiated equations, i.e., the equations obtained by commuting ∂_t^k, with (1.13). We aim to prove the following proposition which gives stronger estimates for full time derivatives than what we need in (3.61).

Proposition 3.6. The following tangential estimate for full time derivatives holds:

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{k=4}^{7} \left\|\varepsilon^{k-1}\partial_t^k(u, B, \varepsilon p)\right\|_0^2 + \left\|\varepsilon^8\partial_t^8(u, B, \varepsilon p)\right\|_0^2\right) \leq P(E(t)). \tag{3.82}$$

Proof. Let us first consider the case when $k = 8$, that is, $\varepsilon^8\partial_t^8$-estimates. We just need to replace $\overline{\partial}^8$ by ∂_t^8 in the proof of Proposition 3.5 and then check if the analogues of the commutators $\mathcal{J}_1 \sim \mathcal{J}_6$ can be controlled by $P(E(t))$ or not. From (3.71) and (3.73), we know that the analogue of $\mathcal{J}_5$ and $\mathcal{J}_6$ can be controlled in the same way because $\partial_t^8(u, B)$ and $\overline{\partial}^8(u, B)$ have the same weights of Mach number. Next, let us analyze the commutators in the analogues of $\mathcal{J}_1 \sim \mathcal{J}_3$. For example, we consider

$$\varepsilon^8[\partial_t^8, B \cdot \nabla]f, \quad f = u \text{ or } B.$$

This commutator is equal to

$$\varepsilon^8(\partial_t^8 B_j\partial_j f + 8\partial_t B_j\partial_j\partial_t^7 f) + \sum_{k=2}^{7} \varepsilon^8\binom{8}{k}\partial_t^k B_j\partial_j\partial_t^{8-k} f$$

$$= (\varepsilon^8\partial_t^8 B_j)(\partial_j f) + 8(\partial_t B_j)\partial_j(\varepsilon^8\partial_t^7 f) + \sum_{k=2}^{7}\binom{8}{k}(\varepsilon^{k-1}\partial_t^k B_j)(\varepsilon^{9-k}\partial_j\partial_t^{8-k} f),$$

24

where the $L^2(\Omega)$ norms of the first term and the third term can be directly controlled by $E(t)$. For the second term, when $j = 1, 2$, it can be directly controlled; when $j = 3$, we again use $\partial_t B_3|_\Sigma = 0$ and fundamental theorem of calculus to create a weight function $\omega(x_3)$:

$$\|(\partial_t B_j)\partial_j(\varepsilon^8 \partial_t^7 f)\| \leq \|\partial_t \bar{B}\|_\infty \|\varepsilon^8 \overline{\partial} B \partial_t^7 f\|_0 + \|\partial_3 \partial_t B_3\|_\infty \|\varepsilon^8(\omega \partial_3)\partial_t^7 f\|_0 \lesssim \sqrt{E_4(t)E_8(t)} \leq E(t).$$

It remains to analyze the commutator $\mathcal{J}_4$ arising from (3.70). The difference is that $\partial_t^8 p$ needs one more ε weight. Luckily, this term is produced because we invoke the continuity equation which automatically gives us ε^2 weight. So, the analogue of $\mathcal{J}_4$ is controlled in the following way:

$$
\begin{aligned}
&-\int_\Omega \varepsilon^8 \partial_t^8 p \left(\varepsilon^{10}[\partial_t^8, u \cdot \nabla]p\right) \mathrm{d}x = -\int_\Omega \varepsilon^9 \partial_t^8 p \left(\varepsilon^9 [\partial_t^8, u \cdot \nabla]p\right) \mathrm{d}x \\
&= -\int_\Omega \varepsilon^9 \partial_t^8 p \left(\varepsilon^9(\partial_t^8 u_j \partial_j p + 8\partial_t u_j \partial_j \partial_t^8 p)\right) \mathrm{d}x \\
&\quad - \int_\Omega (\varepsilon^9 \partial_t^8 p)\left(\sum_{k=2}^{7} \binom{8}{k}(\varepsilon^{k-1}\partial_t^k u_j)(\varepsilon^{10-k}u_j \partial_j \partial_t^{8-k}p)\right) \mathrm{d}x \lesssim E(t).
\end{aligned}
\tag{3.83}
$$

Therefore, we conclude the $\varepsilon^8 \partial_t^8$-estimates as follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\varepsilon^8 \partial_t^8(u, B, \varepsilon p)\right\|_0^2 \leq P(E(t)).\tag{3.84}$$

For $4 \leq k \leq 7$, the proof is still parallel to Proposition 3.5 if we replace $\varepsilon^8 \partial_t^8$ by $\varepsilon^{k-1}\partial_t^k$ ($4 \leq k \leq 7$). We only need to re-consider the estimates of the commutators arising in the analogues of $\mathcal{J}_1 \sim \mathcal{J}_6$. For example, we look at the case $k = 7$. Similarly as above, we first consider th L^2 estimates of the following commuator

$$\varepsilon^6[\partial_t^7, B \cdot \nabla]f, \quad f = u \text{ or } B,$$

which is equal to

$$
\varepsilon^6(\partial_t^7 B_j \partial_j f + 7\partial_t B_j \partial_j \partial_t^6 f) + \sum_{k=2}^{6} \varepsilon^6\binom{7}{k}\partial_t^k B_j \partial_j \partial_t^{7-k} f
$$

$$
= (\varepsilon^6 \partial_t^7 B_j)(\partial_j f) + 7(\partial_t B_j)\partial_j(\varepsilon^6 \partial_t^7 f) + \sum_{k=2}^{6} \binom{7}{k}(\varepsilon^{k-1}\partial_t^k B_j)(\varepsilon^{7-k}\partial_j \partial_t^{7-k} f).
$$

Again, the first term and the third term are directly controlled by $E(t)$. For the second term, using $\partial_t B_3|_\Sigma = 0$ and fundamental theorem of calculus, we then obtain

$$\|(\partial_t B_j)\partial_j(\varepsilon^6 \partial_t^7 f)\|_0 \leq \|\partial_t \bar{B}\|_\infty \|\varepsilon^6 \overline{\partial}\partial_t^6 f\|_0 + \|\partial_3 \partial_t B_3\|_\infty \|\varepsilon^6(\omega \partial_3)\partial_t^6 f\|_0 \lesssim \sqrt{E_4(t)E_7(t)} \lesssim E(t).$$

As for the analogue of $\mathcal{J}_4$, the continuity equation gives us extra ε^2 weight such that the estimate is uniform in Mach number:

$$
\begin{aligned}
&-\int_\Omega \varepsilon^6 \partial_t^7 p \left(\varepsilon^8[\partial_t^7, u \cdot \nabla]p\right) \mathrm{d}x = -\int_\Omega \varepsilon^7 \partial_t^7 p \left(\varepsilon^7 [\partial_t^7, u \cdot \nabla]p\right) \mathrm{d}x \\
&= -\int_\Omega \varepsilon^7 \partial_t^7 p \left(\varepsilon^7(\partial_t^7 u_j \partial_j p + 7\partial_t u_j \partial_j \partial_t^6 p)\right) \mathrm{d}x \\
&\quad - \int_\Omega (\varepsilon^7 \partial_t^7 p)\left(\sum_{k=2}^{6} \binom{7}{k}(\varepsilon^{k-1}\partial_t^k u_j)(\varepsilon^{8-k}u_j \partial_j \partial_t^{8-k}p)\right) \mathrm{d}x \lesssim E(t).
\end{aligned}
\tag{3.85}
$$

Therefore, we obtain the energy estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\|\varepsilon^6 \partial_t^7(\varepsilon p, u, B)\right\|_0^2 \leq P(E(t)).\tag{3.86}$$

The case when $4 \leq k \leq 6$ can be proved in the same way, so we conclude the following energy estimate for full time derivatives

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=4}^{7}\left\|\varepsilon^{k-1}\partial_t^k(u, B, \varepsilon p)\right\|_0^2 \leq P(E(t)).\tag{3.87}$$

The proof is completed. $\qquad\square$

3.3.4 Space-time mixed tangential derivatives

We still need to prove the tangential estimates for $\varepsilon^8 \mathcal{T}^\alpha$ ($0 < \alpha_0 < 8$). Indeed, they can be proved in the same way as Proposition 3.5 and Proposition 3.6 because of the following reasons

- The derivatives $\varepsilon^8 \mathcal{T}^\alpha$ ($0 < \alpha_0 < 8$) are tangential, so they preserve the boundary conditions and thus all boundary integrals must vanish.
- The commutators arising in the proof of Proposition 3.5 and Proposition 3.6 do not produce extra time derivatives, so the weight ε^8 is enough to close the estimate.
- When $\alpha_4 \neq 0$, that is, $(\omega\partial_3)^{\alpha_4}$ appears in $\mathcal{T}^\alpha$, it suffices to use the same strategy as presented in Proposition 3.5. Indeed, we only commute $(\omega\partial_3)$'s with either $u \cdot \nabla$ or $B \cdot \nabla$, so there must a u_3 or B_3 appearing when ∂_3 falls on ω. Thus, we can use the boundary conditions for u_3, B_3 to reproduce the weight function $\omega(x_3)$.

Based on the above analysis, we conclude the tangential estimates in following proposition.

Proposition 3.7. We have the energy inequality:

$$
\frac{\mathrm{d}}{\mathrm{d}t} \Big(\sum_{\langle\alpha\rangle=8,\alpha_0<8} \big\| \varepsilon^8 \mathcal{T}^\alpha(u,B,p) \big\|_0^2 + \big\| \varepsilon^8 \partial_t^8(u,B,\varepsilon p) \big\|_0^2
$$
$$
+ \sum_{l=0}^{3} \sum_{\langle\alpha\rangle=2l,\alpha_0<2l} \big\| \varepsilon^{3+l} \mathcal{T}^\alpha \partial_t^{4-l}(u,B,p) \big\|_0^2 + \sum_{l=0}^{3} \big\| \varepsilon^{3+l} \partial_t^{4+l}(u,B,\varepsilon p) \big\|_0^2 \Big) \leq P(E(t)). \tag{3.88}
$$

3.4 Uniform estimates

Since $D_t S = 0$, we can easily prove the estimates for S by directly commuting D_t with ∂_*^α and corresponding weights of Mach number. The proof does not involve any boundary term because we do not integrate by parts, so we omit the details.

$$
\frac{\mathrm{d}}{\mathrm{d}t} \sum_{l=0}^{4} \sum_{\langle\alpha\rangle=2l} \big\| \varepsilon^{(k-1)_+ + 2l} \mathcal{T}^\alpha \partial_t^k S \big\|_{4-k-l}^2 \leq P(E(t)). \tag{3.89}
$$

Combining the tangential estimates presented in Propositions 3.5-3.7 with the div-curl analysis in section 3.2.2, the reduction of pressure in section 3.2.1, the L^2 estimates and the summary in section 3.2.3, we can get the following energy inequality

$$
\frac{\mathrm{d}}{\mathrm{d}t} E(t) \leq P(E(t)), \tag{3.90}
$$

where $E(t)$ is defined by (1.18). Since the right side of the energy inequality does not rely on ε, we can use Grönwall's inequality to prove that there exists some $T > 0$ independent of ε such that

$$
\sup_{t\in[0,T]} E(t) \leq P(E(0)). \tag{3.91}
$$

Theorem 1.1 is proven.

3.5 The case of 2D MHD

The previous sections are devoted to the incompressible limit of 3D MHD. Now we explain how to modify the proof such that it is valid for 2D MHD flows. The proof for the 2D case is essentially the same as the 3D case: we only need to do a slight modification in the vorticity analysis because the vorticity in 2D is a scalar function, not a vector function. First, lemma 3.1 should be modified as below

Lemma 3.8 (Hodge elliptic estimates). For any sufficiently smooth vector field $X \in \mathbb{R}^2$ and any real number $s \geq 1$, one has

$$
\|X\|_s^2 \lesssim \|X\|_0^2 + \|\nabla \cdot X\|_{s-1}^2 + \|\nabla^\perp \cdot X\|_{s-1}^2 + |\bar{\partial}_1 X \cdot N|_{s-\frac{3}{2}}^2, \tag{3.92}
$$

where $\nabla = (\partial_1, \partial_2)$ and $\nabla^\perp = (-\partial_2, \partial_1)$.

Taking $\nabla^\perp \cdot$ in the momentum equation, we obtain the analogue of (3.12) as

$$\rho D_t(\nabla^\perp \cdot u) - (B \cdot \nabla)(\nabla^\perp \cdot B) = \rho[D_t, \nabla^\perp \cdot]u - [B \cdot \nabla, \nabla^\perp \cdot]B - (\nabla^\perp \rho) \cdot (D_t u), \tag{3.93}$$

which has the same structure as (3.12). Thus, one can follow the analysis in section 3.2.2. The only slight difference is the treatment of Lorentz force term. We take ∂^3-estimate of $\nabla^\perp \cdot (u, B)$ for an example. Following the proof of lemma 3.2, we need to control the following integral

$$\mathcal{K}'' := \int_\Omega \partial^3(\nabla^\perp \cdot B)\, \partial^3 \nabla^\perp \cdot (\varepsilon^2 B D_t p)\, \mathrm{d}x, \tag{3.94}$$

where the problematic term is

$$\varepsilon^2 (B \cdot \nabla^\perp)\partial^3 D_t p = \varepsilon^2(-B_1 \partial^3 \partial_2 D_t p + B_2 \partial^3 \partial_1 D_t p)$$

after inserting the continuity equation. Commuting $\nabla^\perp$ with D_t, we have

$$\varepsilon^2(-B_1 \partial^3 \partial_2 D_t p + B_2 \partial^3 \partial_1 D_t p)$$
$$= \varepsilon^2 \left(-B_1 \partial^3 D_t(\partial_2 p) + B_2 \partial^3 D_t(\partial_1 p) \right) - \varepsilon^2 \left(B_1 \partial^3(\partial_2 u_j \partial_j p) - B_2 \partial^3(\partial_1 u_j \partial_j p) \right),$$

where the $L^2(\Omega)$ norm of the second term is directly controlled by $E_4(t)$. For the first term, we again invoke the momentum equation, which in 2D reads

$$-\partial_1 p = \rho D_t u_1 - B_1 \partial_1 B_1 - B_2 \partial_2 B_1 + \partial_1(|B|^2/2) = \rho D_t u_1 + B_2(\partial_1 B_2 - \partial_2 B_1) = \rho D_t u_1 + B_2(\nabla^\perp \cdot B),$$
$$-\partial_2 p = \rho D_t u_2 - B_1 \partial_1 B_2 - B_2 \partial_2 B_2 + \partial_2(|B|^2/2) = \rho D_t u_1 - B_1(\partial_1 B_2 - \partial_2 B_1) = \rho D_t u_2 - B_1(\nabla^\perp \cdot B).$$

Now, we have

$$\varepsilon^2(-B_1 \partial^3 \partial_2 D_t p + B_2 \partial^3 \partial_1 D_t p)$$
$$= \varepsilon^2 \rho(B_1 \partial^3 D_t^2 u_2 - B_2 \partial^3 D_t^2 u_1) \underbrace{-\varepsilon^2(B_1^2 + B_2^2)\partial^3 D_t(\nabla^\perp \cdot B)}_{\varepsilon^2|B|^2 \text{ has definite sign}} + L^2(\Omega)\text{-controllable terms},$$

and thus

$$\mathcal{K}'' \overset{L}{=} \int_\Omega \partial^3(\nabla^\perp \cdot B)\varepsilon^2 \rho(B_1 \partial^3 D_t^2 u_2 - B_2 \partial^3 D_t^2 u_1)\, \mathrm{d}x - \int_\Omega \partial^3(\nabla^\perp \cdot B)\, \varepsilon^2|B|^2 \partial^3 D_t(\nabla^\perp \cdot B)\, \mathrm{d}x$$
$$\overset{L}{=} \int_\Omega \partial^3(\nabla^\perp \cdot B)\varepsilon^2 \rho(B_1 \partial^3 D_t^2 u_2 - B_2 \partial^3 D_t^2 u_1)\, \mathrm{d}x \tag{3.95}$$
$$- \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \left|\varepsilon|B|\partial^3(\nabla^\perp \cdot B)\right|^2\, \mathrm{d}x + \frac{1}{2}\int_\Omega (\nabla \cdot u)\left(\left|\varepsilon|B|\partial^3(\nabla^\perp \cdot B)\right|^2\right)\, \mathrm{d}x,$$

where the last term on the right side can be directly controlled, and the first term requires the control of $\|\varepsilon^2 D_t^2 u\|_3$. The conclusion of Lemma 3.2 is then modified to be

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho\left|\partial^3(\nabla^\perp \cdot u)\right|^2 + (1 + \varepsilon^2|B|^2)\left|\partial^3(\nabla^\perp \cdot B)\right|^2\, \mathrm{d}x \le P(E_4(t)) + E_5(t). \tag{3.96}$$

4 Incompressible limit

This section is devoted to showing that we can pass the solution of (1.13) to the incompressible limit (1.19) provided that we have the convergence for initial data. In other words, we study the behavior of the solution of (1.13) as the Mach number ε tends to 0. Recall that the Mach number ε is defined in Section 1.1.

The energy estimate presented in Theorem 1.1 implies the boundedness of $\sum_{k=0,1} \|\partial_t^k(u, B, S)\|_{4-k}$ uniformly in ε within the time interval $[0, T]$. Thus, by Aubin-Lions Lemma, up to a subsequence, we have

$$(u, B, S) \to (u^0, B^0, S^0) \quad \text{weakly-* in } L^\infty([0, T]; H^4(\Omega)) \text{ and strongly in } C([0, T]; H^{4-\delta}(\Omega)), \tag{4.1}$$

27

for $\delta > 0$. Moreover, by using the continuity equation $\varepsilon^2 D_t p + \nabla \cdot u = 0$, we have

$$\nabla \cdot u \to \nabla \cdot u^0 = 0 \quad \text{in } L^\infty([0, T]; H^3(\Omega)) \text{ and strongly in } C([0, T]; H^{3-\delta}(\Omega)) \tag{4.2}$$

for $\delta > 0$ because $\|\varepsilon \partial_{t,x} p\|_3$ and $\|u\|_4$ are uniformly bounded in $[0, T]$.

Now, $(u^0, B^0, S^0) \in L^\infty([0, T]; H^4(\Omega))$ solves the incompressible MHD equations together with a transport equation

$$\begin{cases} \varrho(\partial_t u^0 + u^0 \cdot \nabla u^0) - B^0 \cdot \nabla B^0 + \nabla(\pi + \frac{1}{2}|B^0|^2) = 0 & \text{in } [0, T] \times \Omega, \\ \partial_t B^0 + u^0 \cdot \nabla B^0 - B^0 \cdot \nabla u^0 = 0 & \text{in } [0, T] \times \Omega, \\ \nabla \cdot u^0 = \nabla \cdot B^0 = 0 & \text{in } [0, T] \times \Omega, \\ \partial_t S^0 + u^0 \cdot \nabla S^0 = 0 & \text{in } [0, T] \times \Omega, \\ u^0_3 = B^0_3 = 0 & \text{on } [0, T] \times \Sigma, \\ (u^0, B^0, S^0)|_{t=0} = (u^0_0, B^0_0, S^0_0), \end{cases} \tag{4.3}$$

for a suitable fluid pressure function π satisfying $\nabla \pi \in C([0, T]; H^3(\Omega))$. Here ϱ satisfies $\partial_t \varrho + u^0 \cdot \nabla \varrho = 0$, with initial data $\varrho_0 := \rho(0, S^0_0)$. Moreover, the uniqueness of the limit function implies that the convergence holds as $\varepsilon \to 0$ without restricting to a subsequence. Theorem 1.2 is then proven.

References

[1] Alazard, T. *Incompressible limit of the nonisentropic Euler equations with the solid wall boundary conditions.* Adv. Differ. Equ., 10(1):19–44, 2005.

[2] Asano, K. *On the incompressible limit of the compressible Euler equation.* Japan J. Appl. Math., 4(3):455–488, 1987.

[3] Chen, S.-X. *Initial boundary value problems for quasilinear symmetric hyperbolic systems with characteristic boundary.* Translated from Chin. Ann. Math. 3(2), 222–232 (1982). Front. Math. China 2(1), 87–102 (2007).

[4] Coutand, D., Shkoller, S. *Well-posedness of the free-surface incompressible Euler equations with or without surface tension.* J. Am. Math. Soc., 20(3): 829–930, 2007.

[5] Ebin, D. G. *Motion of slightly compressible fluids in a bounded domain. I.* Commun. Pure Appl. Math., 35(4):451–485, 1982.

[6] Goedbloed, H., Keppens, R. and Poedts, S. *Magnetohydrodynamics of Laboratory and Astrophysical plasmas.* Cambridge University Press, 2020.

[7] Gu, X., Wang, Y. *On the construction of solutions to the free-surface incompressible ideal magnetohydrodynamic equations.* J. Math. Pures Appl., Vol. 128: 1-41, 2019.

[8] Iguchi, T. *The incompressible limit and the initial layer of the compressible Euler equation in $\mathbb{R}^n_+$.* Math. Methods Appl. Sci., 20(11):945–958, 1997.

[9] Isozaki, H. *Singular limits for the compressible Euler equations in an exterior domain.* J. Reine Angew. Math., 381:1-36, 1987.

[10] Jiang, S., Ju, Q., Li, F. *Incompressible limit of the nonisentropic magnetohydrodynamic equations.* SIAM J. Math. Anal., 48(1), 302-319, 2016.

[11] Jiang, S., Ju, Q., Xu, X. *Small Alfvén number limit for incompressible magnetohydrodynamics in a domain with boundaries.* Sci. China Math., 62, 2229-2248, 2019.

[12] Ju, Q., Schochet, S., Xu, X. *Singular limits of the equations of compressible ideal magnetohydrodynamics in a domain with boundaries.* Asymptotic Anal., 113, 137-165, 2019.

[13] Ju, Q., Wang, J. *Incompressible limit of ideal magnetohydrodynamic equations with a perfectly conducting wall condition.* SIAM J. Math. Anal., to appear, 2023.

[14] Klainerman, S., Majda, A. *Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids.* Commun. Pure Appl. Math., 34(4):481–524, 1981.

[15] Klainerman, S., Majda, A. *Compressible and incompressible fluids.* Commun. Pure Appl. Math., 35(5):629–651, 1982.

[16] Lindblad, H., Zhang, J. *Anisotropic Regularity of the Free-Boundary Problem in Compressible Ideal Magnetohydrodynamics.* Arch. Rational Mech. Anal., 247(5), no.89: 1-94, 2023.

[17] Luo, C., Zhang, J. *Compressible Gravity-Capillary Water Waves: Local Well-Posedness, Incompressible and Zero-Surface-Tension Limits.* arXiv:2211.03600, preprint.

[18] Métivier, G., Schochet, S. *The incompressible limit of the non-isentropic Euler equations.* Arch. Rational Mech. Anal., 158(1):61–90, 2001.

[19] Morando, A., Trakhinin, Y., Trebeschi, P. *Local existence of MHD contact discontinuities.* Arch. Rational Mech. Anal., 228(2), 691-742, 2018.

[20] Ohno, M., Shirota, T. *On the initial-boundary-value problem for the linearized equations of magneto-hydrodynamics.* Arch. Rational Mech. Anal., 144(3), 259-299, 1998.

[21] Ohno, M., Shizuta, Y., Yanagisawa, T. *The trace theorem on anisotropic Sobolev spaces.* Tohoku Math. J., 46(3), 393-401, 1994.

[22] Rauch, J. *Symmetric Positive Systems with Boundary Characteristic of Constant Multiplicity.* Trans. Amer. Math. Soc., 291(1), 167-187, 1985.

[23] Schochet, S. *The compressible Euler equations in a bounded domain: Existence of solutions and the incompressible limit.* Commun. Math. Phys., 104(1):49–75, 1986.

[24] Schochet, S. *Singular limits in bounded domains for quasilinear symmetric hyperbolic systems having a vorticity equation.* J. Differ. Equ., 68:400-428, 1987.

[25] Schochet, S. *Fast Singular Limits of Hyperbolic PDEs.* J. Differ. Equ., 114(2):476-512, 1994.

[26] Secchi, P. *Well-posedness for Mixed Problems for the Equations of Ideal Magneto-hydrodynamics.* Archiv. der. Math. 64(3), 237–245, 1995.

[27] Secchi, P. *On an initial boundary value problem for the equations of ideal magnetohydrodynamics.* Math. Methods Appl. Sci., 18, 841-853, 1995.

[28] Secchi, P. *Well-posedness of characteristic symmetric hyperbolic systems.* Arch. Rational Mech. Anal., 134(2), 155-197, 1996.

[29] Secchi, P. *On the Singular Incompressible Limit of Inviscid Compressible Fluids.* J. Math. Fluid Mech., 2(2), 107-125, 2000.

[30] Trakhinin, Y. *The existence of current-vortex sheets in ideal compressible magnetohydrodynamics.* Arch. Rational Mech. Anal., 191(2), 245-310, 2009.

[31] Trakhinin, Y., Wang, T. *Well-posedness of Free Boundary Problem in Non-relativistic and Relativistic Ideal Compressible Magnetohydrodynamics.* Arch. Rational Mech. Anal., 239(2), 1131-1176, 2021.

[32] Trakhinin, Y., Wang, T. *Nonlinear Stability of MHD Contact Discontinuities with Surface Tension.* Arch. Rational Mech. Anal., 243(2), 1091-1149, 2022.

[33] Wang, Y., Xin, Z. *Existence of Multi-dimensional Contact Discontinuities for the Ideal Compressible Magnetohydrodynamics.* Commun. Pure Appl. Math., to appear, 2023.

[34] Ukai, S. *The incompressible limit and the initial layer of the compressible Euler equation.* J. Math. Kyoto Univ., 26(2):323–331, 1986.

[35] Yanagisawa, T. *The Initial Boundary Value Problem for the Equations of Ideal Magneto-Hydrodynamics.* Hokk. Math. J., Vol.16, 295-314, 1987.

[36] Yanagisawa, T., Matsumura, A. *The fixed boundary value problems for the equations of ideal magnetohydrodynamics with a perfectly conducting wall condition.* Commun. Math. Phys., 136(1), 119-140, 1991.

[37] Zhang, J. *Local Well-posedness of the Free-Boundary Problem in Compressible Resistive Magnetohydrodynamics.* Calc. Var. Partial Differ. Equ., 62(4), no.124: 1-60, 2023.

[38] Zhang, J. *Local Well-posedness and Incompressible Limit of the Free-Boundary Problem in Compressible Elastodynamics.* Arch. Rational Mech. Anal., 244(3), 599-697, 2022.