Rigorous justification of the hydrostatic approximation limit of viscous compressible flows

Xin Liu* and Edriss S. Titi†

May 14, 2023

Abstract

This paper considers the asymptotic limit of small aspect ratio between vertical and horizontal spatial scales for viscous isothermal compressible flows. In particular, it is observed that fast vertical acoustic waves arise and induce an averaging mechanism of the density in the vertical variable, which at the limit leads to the hydrostatic approximation of compressible flows, i.e., the compressible primitive equations of atmospheric dynamics. We justify the hydrostatic approximation for general as well as "well-prepared" initial data. The initial data is called well-prepared when it is close to the hydrostatic balance in a strong topology. Moreover, the convergence rate is calculated in the well-prepared initial data case in terms of the aspect ratio, as the latter goes to zero.

Mathematics Subject Classification 2020: 35B25, 35B40, 76N10, 76N30.

Keywords: Hydrostatic approximation, Compressible primitive equations, Viscous isothermal compressible flows.

Contents

1 Introduction **2**

1.1 Hydrostatic approximation of compressible flows 2

1.2 Preliminaries 8

1.3 Main theorems 9

*Department of Mathematics, Texas A&M University, College Station, TX 77843, USA. Email: xliu23@tamu.edu

†Department of Mathematics, Texas A&M University, College Station, TX 77843, USA. Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK. Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: titi@math.tamu.edu and Edriss.Titi@maths.cam.ac.uk

2 Uniform-in-ε estimates **11**
2.1 L^2–estimates . . . 11
2.2 Estimates for the vertical velocity and the density . . . 12
2.3 Higher order uniform-in-ε estimates . . . 16
2.4 Uniform-in-ε estimates for the nonlinearities . . . 19
2.5 Summary of uniform-in-ε estimates . . . 22

3 Convergence to the compressible primitive equations **23**

4 The rates of convergence in terms of ε **24**
4.1 Additional uniform-in-ε bounds . . . 24
4.2 Convergence rates in terms of ε . . . 27
4.3 Proofs of Proposition 3 and Proposition 4 . . . 32

1 Introduction

1.1 Hydrostatic approximation of compressible flows

The compressible primitive equations of atmospheric dynamics are the hydrostatic approximation of the compressible hydrodynamic equations, which are obtained by replacing the evolutionary vertical momentum equation with the hydrostatic balance equation (see $(5)_3$, below). A formal derivation can be found, e.g., in [17]. The small aspect ratio in the atmosphere between the vertical scale and the horizontal planetary scale plays an essential role in this derivation and is the main factor behind the hydrostatic approximation. The hydrostatic approximation is commonly used in the atmospheric science models, and has successfully simplified the complex hydrodynamic equations and their computational aspects. In this work, our goal is to rigorously justify the hydrostatic approximation for viscous compressible flows and to provide a mathematical foundation for its application. For more backgrounds about the hydrostatic approximation, see, e.g., [37, 38, 41, 42]

Recall that the dynamics of compressible isothermal flow in a thin periodic channel domain is governed, after omitting the viscosities, by the isothermal compressible Euler equations in the domain $\Omega_\varepsilon = 2\mathbb{T}^2 \times (0, \varepsilon)$, for small $\varepsilon \in (0, 1)$:

$$\begin{cases} \partial_t \rho + \mathrm{div}_h(\rho v) + \partial_z(\rho w) = 0, & \text{in } \Omega_\varepsilon, \\ \rho \partial_t v + \rho v \cdot \nabla_h v + \rho w \partial_z v + \nabla_h \rho = 0, & \text{in } \Omega_\varepsilon, \\ \rho \partial_t w + \rho v \cdot \nabla_h w + \rho w \partial_z w + \partial_z \rho = 0, & \text{in } \Omega_\varepsilon, \end{cases} \tag{EQ}$$

subject to the impenetrable boundary condition at the wall boundaries

$$w\big|_{z=0,\varepsilon} = 0,$$

where ρ, v, w represent the unknown density, the horizontal velocity, and the vertical velocity, respectively. Here div_h, ∇_h represent the horizontal divergence and the horizontal gradient, respectively (see section 1.2, below). $2\mathbb{T}^2 \subset \mathbb{R}^2$ is the periodic horizontal domain (flat torus) with period 2 in each direction.

Alternatively, set $\sigma := \log \rho$, which, for simplicity, we will still refer to as the density. Then, in the region where $\rho > 0$, (EQ) is equivalent to

$$\begin{cases} \partial_t \sigma + v \cdot \nabla_h \sigma + w \partial_z \sigma + \mathrm{div}_h v + \partial_z w = 0, & \text{in } \Omega_\varepsilon, \\ \partial_t v + v \cdot \nabla_h v + w \partial_z v + \nabla_h \sigma = 0, & \text{in } \Omega_\varepsilon, \\ \partial_t w + v \cdot \nabla_h w + w \partial_z w + \partial_z \sigma = 0, & \text{in } \Omega_\varepsilon, \end{cases} \tag{EQ'}$$

with $w\big|_{z=0,\varepsilon} = 0$. As $\varepsilon \to 0^+$, it is expected that

$$\Big(\frac{1}{\varepsilon}\int_0^\varepsilon \sigma(\cdot, z')\,dz', \frac{1}{\varepsilon}\int_0^\varepsilon v(\cdot, z')\,dz'\Big)$$

will converge to solutions to the two-dimensional compressible Euler equations. This has been verified in the viscous case (i.e., the compressible Navier-Stokes equations) in [2] (see also [45]). In view of multi-scale analysis, this is to say that, if one considers the multi-scale expansion of the solutions to (EQ') by writing

$$\begin{aligned} \sigma(x,y,z,t) =& \sigma^0(x,y,z/\varepsilon,t) + \varepsilon\sigma^1(x,y,z/\varepsilon,t) + \cdots, \\ v(x,y,z,t) =& v^0(x,y,z/\varepsilon,t) + \varepsilon v^1(x,y,z/\varepsilon,t) + \cdots, \\ w(x,y,z,t) =& w^0(x,y,z/\varepsilon,t) + \varepsilon w^1(x,y,z/\varepsilon,t) + \cdots, \end{aligned}$$

it should be expected that $(\int_0^1 \sigma^0(\cdot, z')\,dz', \int_0^1 v^0(\cdot, z')\,dz')$ solves the two-dimensional compressible Euler equations and $w^0 \equiv 0$. Therefore, the non-trivial leading order of the vertical velocity w is w^1, and one can easily check that the equations satisfied by (σ^0, v^0, w^1) are exactly given by the inviscid compressible primitive equations, i.e., equations (5), below, without the viscosities. To capture the above scale analysis rigorously, we consider the ansatz

$$\begin{aligned} \sigma(x,y,z) :=& \sigma_\varepsilon(x,y,z/\varepsilon), \\ v(x,y,z) :=& v_\varepsilon(x,y,z/\varepsilon), \\ w(x,y,z) :=& \varepsilon w_\varepsilon(x,y,z/\varepsilon), \end{aligned} \tag{1}$$

and denote by $z' := z/\varepsilon \in (0,1)$ and $\Omega := 2\mathbb{T}^2 \times (0,1)$. Then we write down the equations satisfied by $(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$ in Ω, taking into account the eddy viscosities:

$$\begin{cases} \partial_t \sigma_\varepsilon + v_\varepsilon \cdot \nabla_h \sigma_\varepsilon + w_\varepsilon \partial_z \sigma_\varepsilon + \mathrm{div}_h v_\varepsilon + \partial_z w_\varepsilon = 0, & \text{in } \Omega, \\ \partial_t v_\varepsilon + v_\varepsilon \cdot \nabla_h v_\varepsilon + w_\varepsilon \partial_z v_\varepsilon + \nabla_h \sigma_\varepsilon = \Delta_h v_\varepsilon + \partial_{zz} v_\varepsilon, & \text{in } \Omega, \\ \varepsilon^2\big(\partial_t w_\varepsilon + v_\varepsilon \cdot \nabla_h w_\varepsilon + w_\varepsilon \partial_z w_\varepsilon\big) + \partial_z \sigma_\varepsilon = \varepsilon^2\big(\Delta_h w_\varepsilon + \partial_{zz} w_\varepsilon\big), & \text{in } \Omega, \end{cases} \tag{2}$$

subject to the impenetrable and stress-free boundary conditions at the wall boundaries

$$w_\varepsilon\big|_{z=0,1} = 0, \ \partial_z v_\varepsilon\big|_{z=0,1} = 0, \tag{3}$$

where we have dropped the prime sign for the vertical variable.

We remark that (2) can also be obtained by considering ansatz (1) in the viscous version of (EQ') with viscosities $\Delta_h v_\varepsilon + \varepsilon^{-2}\partial_{zz} v_\varepsilon$ and $\Delta_h w_\varepsilon + \varepsilon^{-2}\partial_{zz} w_\varepsilon$ on the right-hand side of the horizontal momentum equation $(\mathrm{EQ'})_2$ and the vertical momentum equation $(\mathrm{EQ'})_3$, respectively.

Let us observe that the solutions to (2) is invariant with respect to the following symmetry:

$$\sigma_\varepsilon, v_\varepsilon, w_\varepsilon \text{ are even, even, and odd in the } z\text{-variable, respectively.} \tag{SYM}$$

For this reason, in this work, we equivalently consider the following viscous compressible hydrodynamic system with turbulence eddy viscosities: for small $\varepsilon \in (0,1)$,

$$\begin{cases} \partial_t \sigma_\varepsilon + v_\varepsilon \cdot \nabla_h \sigma_\varepsilon + w_\varepsilon \partial_z \sigma_\varepsilon + \mathrm{div}_h v_\varepsilon + \partial_z w_\varepsilon = 0, & \text{in } 2\mathbb{T}^3, \\ \partial_t v_\varepsilon + v_\varepsilon \cdot \nabla_h v_\varepsilon + w_\varepsilon \partial_z v_\varepsilon + \nabla_h \sigma_\varepsilon = \Delta_h v_\varepsilon + \partial_{zz} v_\varepsilon, & \text{in } 2\mathbb{T}^3, \\ \varepsilon^2 \big(\partial_t w_\varepsilon + v_\varepsilon \cdot \nabla_h w_\varepsilon + w_\varepsilon \partial_z w_\varepsilon \big) + \partial_z \sigma_\varepsilon = \varepsilon^2 \big(\Delta_h w_\varepsilon + \partial_{zz} w_\varepsilon \big), & \text{in } 2\mathbb{T}^3, \end{cases} \tag{4}$$

satisfying the symmetry in (SYM) and the periodic boundary condition in all directions. Notice that the boundary conditions in (3) are automatically satisfied by regular solutions to (4) owing to symmetry (SYM), and one can obtain a solution to (2) with (3) by restricting any regular solution of (4) to the domain Ω.

The formal limit system of (4), as $\varepsilon \to 0^+$, is

$$\begin{cases} \partial_t \sigma_p + v_p \cdot \nabla_h \sigma_p + w_p \partial_z \sigma_p + \mathrm{div}_h v_p + \partial_z w_p = 0, & \text{in } 2\mathbb{T}^3, \\ \partial_t v_p + v_p \cdot \nabla_h v_p + w_p \partial_z v_p + \nabla_h \sigma_p = \Delta_h v_p + \partial_{zz} v_p, & \text{in } 2\mathbb{T}^3, \\ \partial_z \sigma_p = 0, & \text{in } 2\mathbb{T}^3, \end{cases} \tag{5}$$

satisfying the same symmetry as in (SYM), with $\sigma_\varepsilon, v_\varepsilon, w_\varepsilon$ replaced by σ_p, v_p, w_p, respectively.

Before we move on to the discussion of our strategy to rigorously establish the aforementioned limiting problem, we would like to mention some relevant previous results.

As for the compressible hydrodynamic equations for viscous flows, i.e., the compressible Navier-Stokes equations, the derivation of the system can be found in [18, 39]. In [18, 40], global weak solutions to the compressible Navier-Stokes equations are constructed. Recently, the authors in [33] and [55] independently construct global weak solutions to the compressible Navier-Strokes equations with degenerate viscosities. See also [3] and [5] for relevant developments. As for the strong solutions of the compressible Navier-Stokes equations, [13–15, 31] establish the local well-posedness of strong and classical solutions with vacuum. Without vacuum, the local well-posedness theory can be dated back to [29, 51, 53]. The first global well-posedness result is established in [46, 47], where the asymptotic stability of constant states is studied with respect to small perturbations. A global existence theorem with vacuum and small energy is given in [28] (see also [27]). We refer readers, for other developments of the compressible Navier-Stokes equations to, e.g., [4, 23, 24, 26, 56, 59].

As for the compressible primitive equations (PE) (5), one can find the meteorological discussion and applications of the system in, e.g., [50] and [58]. In two dimensions, the global weak solutions are constructed in [16, 20]. The stability of weak solutions is established in [17]. Uniqueness of the weak solutions in two dimensions is studied in [30]. As for the three-dimensional dynamics, the existence of global weak solutions for the compressible primitive equations with degenerate viscosities is established in [42] and [57], independently. The local well-posedness of strong solutions for the compressible primitive equations with constant viscosities is established in [41]. We introduce the PE diagram and study the small Mach number limit of the compressible primitive equations in [43] and [44]. For readers' convenience, results concerning the incompressible primitive equations can be found in [1, 6–12, 21, 22, 25, 32, 34, 35, 48, 49].

We would like to remark that, the well-posedness of local strong or classical solutions to (4) for any fixed $\varepsilon \in (0,1)$, and (5), follows easily, using the similar arguments as in [29] and [41], respectively. We only mention that w_p in (5) can be calculated by

$$w_p(\cdot_h, z) = -e^{-\sigma_p(\cdot_h)} \int_0^z e^{\sigma_p(\cdot_h)} \Big(\widetilde{v}_p(\cdot_h, z') \cdot \nabla_h \sigma_p(\cdot_h) + \mathrm{div}_h \widetilde{v}_p(\cdot_h, z') \Big) \, dz', \quad (6)$$

using the continuity equation $(5)_1$, while the vertical average part of the continuity equation yields an evolutionary equation for σ_p, i.e.,

$$\partial_t \sigma_p + \overline{v}_p \cdot \nabla_h \sigma_p + \mathrm{div}_h \overline{v}_p = 0.$$

Here we have used the fact that σ_p is independent of the z-variable due to the hydrostatic equation $(5)_3$, $\cdot_h$ denotes the horizontal variables, i.e., x, y, and $\overline{\cdot}$ and $\widetilde{\cdot}$ are the vertical average and fluctuation, respectively, defined explicitly in (11) in section 1.2, below.

To establish the asymptotic limit as $\varepsilon \to 0^+$, we will need to establish some uniform-in-ε estimates for $(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$. While in the case of the incompressible flows, e.g., [1, 35, 36], the vertical velocity can be represented by the horizontal velocity, using the incompressibility condition, this benefit no longer exists in the compressible case. Instead, for the compressible primitive equations, as in (6), the vertical velocity is represented by the density σ_p combined with the horizontal velocity v_p, which does not involve any time derivative. This allows us to consider $(5)_1$ and $(5)_2$ as closed evolutionary equations of (σ_p, v_p) only, and thus allows us to construct the local strong solutions in [41]. However, all these structures do not work for equations (4). Indeed, the vertical momentum equation $(4)_3$ can only provide estimates for $\varepsilon w_\varepsilon$, while using the continuity equation $(4)_1$, the representation of w_ε in terms of σ_ε and v_ε involves the time derivative $\partial_t \sigma_\varepsilon$ (see (26), below). This may involve high oscillations, as $\varepsilon \to 0^+$, and is the main obstacle to overcome in this work. It is worth noticing that, in [19, 52], by assuming the uniform-in-ε existence of, and bounds on, global weak solutions to fluid system (2) with and without viscosity, and with the additional assumption that the density is independent of the z-variable, the authors justify the hydrostatic approximation in compressible fluid. However, these assumptions in [19, 52] can not be justified physically for atmospheric dynamics. They are mathematically assumed in order to essentially avoid dealing with the obstacle mentioned above. In this paper, our goal is to overcome these difficulties by showing the uniform-in-ε existence, bounds, and convergence.

To motivate our strategy and to shed light on the treatment of this obstacle, let us consider the following linear model system:

$$\begin{cases} \partial_t \eta + \mathrm{div}_h \psi^h + \partial_z \psi^z = 0, \\ \partial_t \psi^h + \nabla_h \eta = \Delta \psi^h, \\ \varepsilon^2 \partial_t \psi^z + \partial_z \eta = \varepsilon^2 \Delta \psi^z, \end{cases} \tag{7}$$

with unknown scalar functions η, ψ^h, and ψ^z in $2\mathbb{T}^3$, subject to the symmetry that η, ψ^h, ψ^z are even, even, and odd in the z-variable, respectively, similar to (SYM). In particular, $\psi^z|_{z=0} = 0$. Then the standard H^s estimate of (7), for every $s \in \mathbb{Z}^+$, yields

$$\sup_{0\leq t\leq T} \|\eta(t), \psi^h(t), \varepsilon\psi^z(t)\|_{H^s}^2 + \int_0^T \|\psi^h(t), \varepsilon\psi^z(t)\|_{H^{s+1}}^2 \, dt < \infty,$$

for any $T \in (0, \infty)$, which allows us to pass the limit $\varepsilon \to 0^+$ in $(7)_2$ and $(7)_3$, but not $(7)_1$, due to the lack of compactness of the ψ^z sequence.

To overcome this obstacle, we focus instead on the hyperbolic structure of the system

$$\begin{cases} \partial_t \eta + \partial_z \psi^z + \cdots = 0, \\ \partial_t \psi^z + \dfrac{1}{\varepsilon^2} \partial_z \eta + \cdots = 0. \end{cases} \tag{8}$$

Then from (8), one can obtain a wave equation for η, i.e.,

$$\partial_{tt} \eta - \frac{1}{\varepsilon^2} \partial_{zz} \eta = \cdots ,$$

which can provide uniform-in-ε estimates for $\partial_t \eta$. In the end, using $(8)_1$ again, one can write $\psi^z(z) = -\int_0^z (\partial_t \eta(z') + \cdots) \, dz'$, from which one can obtain the required uniform-in-ε estimates of ψ^z, and thus the missing (weak) compactness of ψ^z is obtained. In other words, one has to take advantage of the oscillatory nature of the underlying system in order to obtain the required uniform-in-ε estimates.

Back to (7), the contribution of the viscosities (i.e., $\Delta\psi^h$ and $\Delta\psi^z$) should also be taken into consideration. One can calculate, similarly,

$$\partial_{tt} \eta = -\mathrm{div}_h \partial_t \psi^h - \partial_z \partial_t \psi^z = \Delta_h \eta + \frac{1}{\varepsilon^2} \partial_{zz} \eta - \Delta(\mathrm{div}_h \psi^h + \partial_z \psi^z),$$

where, using $(7)_1$, we have $\Delta(\mathrm{div}_h \psi^h + \partial_z \psi^z) = -\partial_t \Delta \eta$. Therefore, we arrive at

$$\partial_t(\partial_t \eta - \Delta \eta) - \Delta_h \eta - \frac{1}{\varepsilon^2} \partial_{zz} \eta = 0,$$

which is a strongly damped wave equation. Using such a structure in the nonlinear problem (4) (see (38), below), one can obtain the required uniform-in-ε estimates for $\partial_t \sigma_\varepsilon$ and hence those for w_ε.

In order to deal with the nonlinearities, we complement (4) with H^3 initial data. However, we remark that this may not be the optimal regularity for the initial data.

After obtaining the aforementioned uniform-in-ε estimates, we will be able to establish the limit of (4) as $\varepsilon \to 0^+$. This is done in section 3. However, this is not enough to establish the convergence rates as $\varepsilon \to 0^+$. To explain why, consider the difference $(\delta\sigma, \delta v, \delta w) := (\sigma_\varepsilon - \sigma_p, v_\varepsilon - v_p, w_\varepsilon - w_p)$. Then $(\delta\sigma, \delta v, \delta w)$ satisfies

$$\begin{cases} \partial_t \delta\sigma + v_\varepsilon \cdot \nabla_h \delta\sigma + w_\varepsilon \partial_z \delta\sigma + \delta v \cdot \nabla_h \sigma_p + \mathrm{div}_h \delta v + \partial_z \delta w = 0, \\ \partial_t \delta v + v_\varepsilon \cdot \nabla_h \delta v + w_\varepsilon \partial_z \delta v + \delta v \cdot \nabla_h v_p + \delta w \partial_z v_p \\ \qquad + \nabla_h \delta\sigma = \Delta_h \delta v + \partial_{zz} \delta v, \\ \partial_t \delta w + v_\varepsilon \cdot \nabla_h \delta w + w_\varepsilon \partial_z \delta w + \delta v \cdot \nabla_h w_p + \delta w \partial_z w_p \\ \qquad + \frac{1}{\varepsilon^2} \partial_z \delta\sigma = \Delta_h \delta w + \partial_{zz} \delta w + \big(\Delta_h w_p + \partial_{zz} w_p - \partial_t w_p \\ \qquad - v_p \cdot \nabla_h w_p - w_p \partial_z w_p\big). \end{cases} \tag{9}$$

While the uniform-in-ε estimates work well for $(\delta\sigma, \delta v, \delta w)$, because of the terms

$$\Delta_h w_p + \partial_{zz} w_p - \partial_t w_p - v_p \cdot \nabla_h w_p - w_p \partial_z w_p$$

on the right-hand side of $(9)_3$, the uniform-in-ε estimates are not comparable to ε, with the exception, however, only of the estimate

$$\partial_z \delta\sigma \sim \mathcal{O}(\varepsilon) \quad \text{in some norm.} \tag{10}$$

See (58), below. For comparison with the case of incompressible flows, we refer the reader to [35, 36], where such an issue does not exist.

Writing $\delta\sigma = \overline{\delta\sigma} + \widetilde{\delta\sigma}$, see (11), below, we notice that (10) implies that the fluctuation of $\delta\sigma$ is of order ε, i.e., $\widetilde{\delta\sigma} \sim \mathcal{O}(\varepsilon)$ in some norm. Inspired by the study of (5) in [41], one can separate $(9)_1$ into the (vertical) average part and the fluctuation part. Then the average part is nothing but an evolutionary equation of $\overline{\delta\sigma}$ and does not involve δw. Therefore, using the average part and $(9)_2$, one can show that $\overline{\delta\sigma} \sim \mathcal{O}(\varepsilon)$ and $\delta v \sim \mathcal{O}(\varepsilon)$ in some norm. On the other hand, the fluctuation part yields that the estimates for δw, in fact, only involve $\partial_t \widetilde{\delta\sigma}$. Therefore, one will only need to obtain an estimate of the order $\partial_z \partial_t \delta\sigma \sim \mathcal{O}(\varepsilon)$ in some norm in order to close the estimates of converging rates. This is established by the use of some additional uniform-in-ε estimates in section 4.1. That is, by assuming a well-prepared initial data, i.e., initial data that is close to the hydrostatic approximation (see assumption (19), below).

The rest of this paper is organized as follows. In section 1.2, we introduce the notations that have been and will be used in this paper, as well as the energy and dissipation functionals. In section 1.3, we summarize the main theorems. In section 2, we establish the uniform-in-ε estimates, which imply the necessary weak and strong compactness in order to pass the limit $\varepsilon \to 0^+$ in section 3. In section 4, we focus on the study of convergence rates, which is established for a restricted class of well-prepared initial data, see (19), below. Readers who are more interested in the converging rates can skip directly to section 4.2.

1.2 Preliminaries

In this paper, we consistently use $t \in [0, \infty)$ to represent the temporal variable, $x, y \in 2\mathbb{T}$ to represent the horizontal spatial variables, and $z \in 2\mathbb{T}$ to represent the vertical spatial variable. We have and will use ∇_h, div_h, and Δ_h to represent the horizontal gradient, the horizontal divergence, and the horizontal Laplace operator, respectively; that is,

$$\nabla_h := \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix}, \ \mathrm{div}_h := \nabla_h \cdot, \ \Delta_h := \mathrm{div}_h \nabla_h.$$

Using such notations, we have

$$\Delta = \Delta_h + \partial_{zz}.$$

Also, for any function f, we denote the vertical average and fluctuation of f as

$$\overline{f}(x, y) := \int_0^1 f(x, y, z)\, dz, \ \widetilde{f} := f - \overline{f}, \tag{11}$$

also known as the barotropic and baroclinic modes, respectively. $|f(z,t)|_X, \|f(t)\|_X$ are used to denote the X-norm in the horizontal domain $2\mathbb{T}^2$ for any fixed z, t, and in the three-dimensional domain $2\mathbb{T}^3$ for any fixed t, respectively.

The following functionals will be used in this paper:

$$\begin{aligned} E = E(t) :=& \|v_\varepsilon, \varepsilon w_\varepsilon\|_{H^3} + \|\partial_t \sigma_\varepsilon, \nabla_h \sigma_\varepsilon, \frac{\partial_z \sigma_\varepsilon}{\varepsilon}\|_{H^2} + \|\sigma_\varepsilon\|_{H^4} \\ & + \|w_\varepsilon, \partial_z w_\varepsilon\|_{H^2}, \end{aligned} \tag{12}$$

$$\begin{aligned} D = D(t) :=& \|\partial_t v_\varepsilon, \partial_t(\varepsilon w_\varepsilon)\|_{H^2} + \|v_\varepsilon, \varepsilon w_\varepsilon\|_{H^4} + \|\partial_t \sigma_\varepsilon, \nabla_h \sigma_\varepsilon, \frac{\partial_z \sigma_\varepsilon}{\varepsilon}\|_{H^3} \\ & + \|w_\varepsilon, \partial_z w_\varepsilon\|_{H^3}. \end{aligned} \tag{13}$$

Correspondingly, we denote the functionals of temporal derivatives as,

$$\begin{aligned} E_1 = E_1(t) :=& \|\partial_t v_\varepsilon, \partial_t(\varepsilon w_\varepsilon)\|_{H^1} + \|\partial_t^2 \sigma_\varepsilon, \partial_t \nabla_h \sigma_\varepsilon, \frac{\partial_z \partial_t \sigma_\varepsilon}{\varepsilon}\|_{L^2} \\ & + \|\partial_t \sigma_\varepsilon\|_{H^2} + \|\partial_t w_\varepsilon, \partial_z \partial_t w_\varepsilon\|_{L^2}, \end{aligned} \tag{14}$$

$$
\begin{aligned}
D_1 = D_1(t) := {} & \|\partial_t^2 v_\varepsilon, \partial_t^2(\varepsilon w_\varepsilon)\|_{L^2} + \|\partial_t v_\varepsilon, \partial_t(\varepsilon w_\varepsilon)\|_{H^2} \\
& + \|\partial_t^2 \sigma_\varepsilon, \partial_t \nabla_h \sigma_\varepsilon, \frac{\partial_z \partial_t \sigma_\varepsilon}{\varepsilon}\|_{H^1} + \|\partial_t w_\varepsilon, \partial_z \partial_t w_\varepsilon\|_{H^1}.
\end{aligned} \tag{15}
$$

We will use $\mathfrak{N}(\cdot)$ to denote a locally Lipschitz nonlinear function of its argument(s), which can be different from line to line. For any two quantities Q_1 and Q_2, $Q_1 \lesssim Q_2$ is used to represent $Q_1 \leq CQ_2$ for some constant $C \in (0, \infty)$, whose value will be different from line to line.

1.3 Main theorems

We consider (4) with initial data

$$
(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)\big|_{t=0} = (\sigma_{0,\varepsilon}, v_{0,\varepsilon}, w_{0,\varepsilon}). \tag{16}
$$

Then the initial data for the time derivatives $\partial_t \sigma_\varepsilon, \partial_t v_\varepsilon, \partial_t w_\varepsilon, \partial_t^2 \sigma_\varepsilon$ are given through compatibility by employing the equations in (4), inductively, i.e.,

$$
(\partial_t \sigma_\varepsilon, \partial_t v_\varepsilon, \partial_t w_\varepsilon, \partial_t^2 \sigma_\varepsilon)\big|_{t=0} = (\sigma_{1,\varepsilon}, v_{1,\varepsilon}, w_{1,\varepsilon}, \sigma_{2,\varepsilon}),
$$

where $\sigma_{1,\varepsilon}, v_{1,\varepsilon}, w_{1,\varepsilon}, \sigma_{2,\varepsilon}, v_{2,\varepsilon}, w_{2,\varepsilon}$ are given by

$$
\begin{gathered}
\sigma_{1,\varepsilon} + v_{0,\varepsilon} \cdot \nabla_h \sigma_{0,\varepsilon} + w_{0,\varepsilon} \partial_z \sigma_{0,\varepsilon} + \mathrm{div}_h v_{0,\varepsilon} + \partial_z w_{0,\varepsilon} = 0, \\
v_{1,\varepsilon} + v_{0,\varepsilon} \cdot \nabla_h v_{0,\varepsilon} + w_{0,\varepsilon} \partial_z v_{0,\varepsilon} + \nabla_h \sigma_{0,\varepsilon} = \Delta_h v_{0,\varepsilon} + \partial_{zz} v_{0,\varepsilon}, \\
w_{1,\varepsilon} + v_{0,\varepsilon} \cdot \nabla_h w_{0,\varepsilon} + w_{0,\varepsilon} \partial_z w_{0,\varepsilon} + \frac{1}{\varepsilon^2} \partial_z \sigma_{0,\varepsilon} = \Delta_h w_{0,\varepsilon} + \partial_{zz} w_{0,\varepsilon}, \\
\sigma_{2,\varepsilon} + v_{0,\varepsilon} \cdot \nabla_h \sigma_{1,\varepsilon} + w_{0,\varepsilon} \partial_z \sigma_{1,\varepsilon} + v_{1,\varepsilon} \cdot \nabla_h \sigma_{0,\varepsilon} \\
+ w_{1,\varepsilon} \partial_z \sigma_{0,\varepsilon} + \mathrm{div}_h v_{1,\varepsilon} + \partial_z w_{1,\varepsilon} = 0.
\end{gathered}
$$

The first theorem in this paper is concerning the uniform-in-ε estimates and the justification of hydrostatic approximation limit:

Theorem 1 (Hydrostatic approximation)**.** *Suppose that the initial data in* (16) *satisfies $E(0) < \infty$. Then there is a $T^* \in (0, \infty)$, independent of ε, such that the solution to* (4) *with initial data* (16) *satisfies*

$$
\sup_{0 \leq t \leq T^*} E^2(t) + \int_0^{T^*} D^2(t)\, dt < C_0 < \infty, \tag{17}
$$

where C_0 is some positive constant depending only on $E(0)$ and is independent of ε. Moreover, there exist (σ_p, v_p, w_p) with

$$
\begin{gathered}
\sigma_p \in L^\infty(0, T^*; H^4), \quad \partial_t \sigma_p \in L^\infty(0, T^*; H^2) \cap L^2(0, T^*; H^3), \\
v_p \in L^\infty(0, T^*; H^3) \cap L^2(0, T^*; H^4), \quad \partial_t v_p \in L^2(0, T^*; H^2), \\
w_p, \partial_z w_p \in L^\infty(0, T^*; H^2) \cap L^2(0, T^*; H^3),
\end{gathered}
$$

such that for a subsequence of $\{(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)\}$, as $\varepsilon \to 0^+$,

$$
\sigma_\varepsilon \overset{*}{\rightharpoonup} \sigma_p \qquad \textit{weak-}{*} \textit{ in } L^\infty(0, T^*; H^4),
$$

$$
\begin{aligned}
\sigma_\varepsilon &\to \sigma_p && \text{in } L^\infty(0,T^*;H^3)\cap C([0,T^*];H^3),\\
\partial_t\sigma_\varepsilon, w_\varepsilon, \partial_z w_\varepsilon &\overset{*}{\rightharpoonup} \partial_t\sigma_p, w_p, \partial_z w_p && \text{weak-}* \text{ in } L^\infty(0,T^*;H^2),\\
\partial_t\sigma_\varepsilon, w_\varepsilon, \partial_z w_\varepsilon &\rightharpoonup \partial_t\sigma_p, w_p, \partial_z w_p && \text{weakly in } L^2(0,T^*;H^3),\\
v_\varepsilon &\overset{*}{\rightharpoonup} v_p && \text{weak-}* \text{ in } L^\infty(0,T^*;H^3),\\
v_\varepsilon &\to v_p && \text{in } L^\infty(0,T^*;H^2)\cap C([0,T^*];H^2),\\
v_\varepsilon &\rightharpoonup v_p && \text{weakly in } L^2(0,T^*;H^4),\\
\partial_t v_\varepsilon &\rightharpoonup \partial_t v_p && \text{weakly in } L^2(0,T^*;H^2),
\end{aligned}
$$

and (σ_p, v_p, w_p) *is a solution to* (5).

In addition, suppose that the initial data in (16) *satisfies* $E_1(0) < \infty$. *Then there is a* $T^{**} \in (0, T^*]$, *independent of* ε, *such that*

$$
\sup_{0\leq t\leq T^{**}} (E^2(t) + E_1^2(t)) + \int_0^{T^{**}} (D^2(t) + D_1(t))\,dt < C_1 < \infty, \tag{18}
$$

where C_1 *is some positive constant depending only on* $E(0)$ *and* $E_1(0)$ *and is independent of* ε.

We summarize the convergence rates in the following theorem:

Theorem 2 (Rates of convergence)**.** *Suppose that the solution* (σ_p, v_p, w_p) *to* (5) *given by the limit in Theorem 1 satisfies*

$$
\begin{gathered}
\|\sigma_p\|_{L^\infty(0,T^{**};H^4)} + \|\partial_t\sigma_p\|_{L^\infty(0,T^{**};H^2)} + \|v_p\|_{L^\infty(0,T^{**};H^3)}\\
+\|v_p\|_{L^2(0,T^{**};H^4)} + \|w_p\|_{L^\infty(0,T^{**};H^2)} < C_p < \infty,
\end{gathered}
$$

for some constant $C_p \in (0,\infty)$, *and the initial data in* (16) *satisfy*

$$
\|\sigma_{0,\varepsilon} - \sigma_p\big|_{t=0}, v_{0,\varepsilon} - v_p\big|_{t=0}\|_{L^2} \lesssim \varepsilon. \tag{19}
$$

Then under the conditions in Theorem 1, i.e., $E(0), E_1(0) < \infty$, *we have*

$$
\begin{gathered}
\|\sigma_\varepsilon - \sigma_p, v_\varepsilon - v_p\|_{L^\infty(0,T^{**};L^2)} + \|v_\varepsilon - v_p\|_{L^2(0,T^{**};H^1)} \leq C_2\varepsilon,\\
\|w_\varepsilon - w_p\|_{L^\infty(0,T^{**};L^2)} \leq C_2\varepsilon^{2/3}, \quad \|w_\varepsilon - w_p\|_{L^2(0,T^{**};L^2)} \leq C_2\varepsilon^{3/4},
\end{gathered} \tag{20}
$$

where $C_2 \in (0,\infty)$ *is a constant depending only on* $E(0), E_1(0)$ *and* C_p, *and is independent of* ε.

Proof of Theorem 1. The uniform estimates in (17) and (18) are shown in section 2.5 and section 4.1, given by (58) and (76), respectively. The convergence is given in section 3, below. □

Proof of Theorem 2. (20) is the consequence of (93) and (98), in section 4.2, below. □

Remark 1. Assumption (19) on the initial data above is the definition of well-prepared initial data, i.e., it essentially assumes that the initial data is close to the hydrostatic approximation.

Remark 2. Theorem 1 guarantees the convergence of a subsequence to a solution of the limit equations. However, since the strong solution to the limit equations is well-posed, and in particular, is unique (see, e.g., [41]), the convergence is actually of the full sequence. Theorem 2 states the convergence rate of the full sequence for well-prepared initial data.

2 Uniform-in-ε estimates

In this section, we shorten the notations by dropping the subscript ε in $(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$, i.e., $(\sigma, v, w) = (\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$. Recall that we already have short time existence and uniqueness of solutions to (4) on a time interval that might depend on ε (see, e.g., [13]). The main goal of this section is to obtain the existence time and corresponding estimates that are independent of ε.

2.1 L^2–estimates

Take the L^2-inner product of $(4)_2$ with $\partial_t v - \Delta v$, and $(4)_3$ with $\partial_t w - \Delta w$, respectively. One obtains,

$$\begin{aligned} &\|\partial_t v - \Delta v\|_{L^2}^2 = - \int \big(\partial_t v - \Delta v\big) \cdot \big(\nabla_h \sigma + v \cdot \nabla_h v + w \partial_z v\big) \, d\vec{x} \\ &\quad \leq \frac{1}{2} \|\partial_t v - \Delta v\|_{L^2}^2 + C \big(\|\nabla_h \sigma\|_{L^2}^2 + \|v \cdot \nabla_h v\|_{L^2}^2 + \|w \partial_z v\|_{L^2}^2 \big), \end{aligned} \tag{21}$$

$$\begin{aligned} &\varepsilon^2 \|\partial_t w - \Delta w\|_{L^2}^2 = -\varepsilon^2 \int \big(\partial_t w - \Delta w\big) \big(\frac{1}{\varepsilon^2} \partial_z \sigma + v \cdot \nabla_h w + w \partial_z w \big) \, d\vec{x} \\ &\quad \leq \frac{\varepsilon^2}{2} \|\partial_t w - \Delta w\|_{L^2}^2 + \frac{C}{\varepsilon^2} \|\partial_z \sigma\|_{L^2}^2 + C \varepsilon^2 \big(\|v \cdot \nabla_h w\|_{L^2}^2 + \|w \partial_z w\|_{L^2}^2 \big), \end{aligned} \tag{22}$$

where we have applied the Hölder and Young inequalities. Notice that, for $\eta = v, w$, applying integration by parts yields

$$\|\partial_t \eta - \Delta \eta\|_{L^2}^2 = \frac{d}{dt} \|\nabla \eta\|_{L^2}^2 + \|\partial_t \eta\|_{L^2}^2 + \|\nabla^2 \eta\|_{L^2}^2. \tag{23}$$

Thus (21) and (22) can be written as,

$$\begin{aligned} &\frac{d}{dt} \|\nabla v, \varepsilon \nabla w\|_{L^2}^2 + \|\partial_t v, \nabla^2 v, \varepsilon \partial_t w, \varepsilon \nabla^2 w\|_{L^2}^2 \leq C \|\nabla_h \sigma, \frac{\partial_z \sigma}{\varepsilon}\|_{L^2}^2 \\ &\qquad + C \sum_{i=1}^{2} \|\mathcal{I}_i\|_{L^2}^2, \end{aligned} \tag{24}$$

where

$$\mathcal{I}_1 = \big(v \cdot \nabla_h v, v \cdot \nabla_h (\varepsilon w)\big), \quad \mathcal{I}_2 = \big(w \partial_z v, w \partial_z (\varepsilon w)\big). \tag{25}$$

We postpone the estimates for $\|\mathcal{I}_i\|_{L^2}$ to section 2.4. However, even without detailed calculations, one can see that, a uniform-in-ε estimate for w is necessary, for instance, to get an estimate for $\mathcal{I}_2$. This indeed is the main challenge to be addressed in this work. In the next section, we discuss our strategy to overcome this obstacle.

2.2 Estimates for the vertical velocity and the density

We use $(4)_1$ to represent w in terms of σ and v. Indeed, after multiplying $(4)_1$ with e^{σ}, it follows that

$$\partial_t e^{\sigma} + \mathrm{div}_h(e^{\sigma} v) + \partial_z(e^{\sigma} w) = 0.$$

Thus, recalling that $\cdot_h$ represents the horizontal variables x, y, one has that,

$$\begin{aligned} w(\cdot_h, z) = -e^{-\sigma(\cdot_h, z)} \int_0^z e^{\sigma(\cdot_h, z')} \Big(& \partial_t \sigma(\cdot_h, z') + v(\cdot_h, z') \cdot \nabla_h \sigma(\cdot_h, z') \\ & + \mathrm{div}_h v(\cdot_h, z') \Big) \, dz'. \end{aligned} \tag{26}$$

Similarly, we remark that we can represent δw in terms of $\delta\sigma$, δv, σ_p, v_p, and w_p, by using $(9)_1$, even thought we don't need it now. One can obtain, after multiplying $(9)_1$ with $e^{\delta\sigma}$, that

$$\partial_t e^{\delta\sigma} + v_p \cdot \nabla_h e^{\delta\sigma} + w_p \partial_z e^{\delta\sigma} + e^{\delta\sigma} \delta v \cdot \nabla_h \sigma_p + \mathrm{div}_h(e^{\delta\sigma} \delta v) + \partial_z(e^{\delta\sigma} \delta w) = 0,$$

and

$$\begin{aligned} \delta w(\cdot_h, z) = -e^{-\delta\sigma(\cdot_h, z)} \int_0^z e^{\delta\sigma(\cdot_h, z')} \Big(& \partial_t \delta\sigma(\cdot_h, z') + v(\cdot_h, z') \cdot \nabla_h \delta\sigma(\cdot_h, z') \\ + \delta v(\cdot_h, z') \cdot \nabla_h \sigma_p(\cdot_h, z') + w_p(\cdot_h, z') \partial_z \delta\sigma(\cdot_h, z') & + \mathrm{div}_h \delta v(\cdot_h, z') \Big) \, dz'. \end{aligned} \tag{27}$$

Eventually, we will return to (27). Then directly using the representation of w in (26), we have the following:

Proposition 1. *Assuming w is given by* (26)*, one can show that*

$$\|w, \partial_z w\|_{H^2} \leq C e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^3}^3 + 1 \big) \big(\|\partial_t \sigma\|_{H^2} + \|v\|_{H^3} + \|\sigma\|_{H^3} \|v\|_{H^2} \big), \tag{28}$$

$$\|w, \partial_z w\|_{H^3} \leq C e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^4}^4 + 1 \big) \big(\|\partial_t \sigma\|_{H^3} + \|v\|_{H^4} + \|\sigma\|_{H^4} \|v\|_{H^3} \big), \tag{29}$$

for some generic positive constant $C \in (0, \infty)$, independent of ε.

Proof. To simplify the notations, denote by

$$\Xi := \partial_t \sigma + v \cdot \nabla_h \sigma + \mathrm{div}_h v. \tag{30}$$

Next, we calculate the derivatives of w:

$$\partial_z w = e^{-\sigma} \partial_z \sigma \int_0^z \big(e^{\sigma} \Xi \big) \, dz' - \Xi,$$

$$\partial_h w = e^{-\sigma} \partial_h \sigma \int_0^z \big(e^{\sigma} \Xi \big) \, dz' - e^{-\sigma} \int_0^z \big(e^{\sigma} \partial_h \sigma \Xi + e^{\sigma} \partial_h \Xi \big) \, dz',$$

$$\partial_{zz} w = e^{-\sigma} \big(\partial_{zz} \sigma - (\partial_z \sigma)^2 \big) \int_0^z \big(e^{\sigma} \Xi \big) \, dz' + \partial_z \sigma \Xi - \partial_z \Xi,$$

$$
\begin{aligned}
\partial_{hz} w &= e^{-\sigma}\big(\partial_{hz}\sigma - \partial_h\sigma\partial_z\sigma\big)\int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&\quad + e^{-\sigma}\partial_z\sigma\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' - \partial_h\Xi, \\
\partial_{hh} w &= e^{-\sigma}\big(\partial_{hh}\sigma - (\partial_h\sigma)^2\big)\int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&\quad + 2e^{-\sigma}\partial_h\sigma\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&\quad - e^{-\sigma}\int_0^z \big[e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\Xi + 2e^\sigma\partial_h\sigma\partial_h\Xi + e^\sigma\partial_{hh}\Xi\big]\,dz', \\
\partial_{zzz} w &= e^{-\sigma}\big(\partial_{zzz}\sigma - 3\partial_z\sigma\partial_{zz}\sigma + (\partial_z\sigma)^3\big)\int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&\quad + \big(2\partial_{zz}\sigma - (\partial_z\sigma)^2\big)\Xi + \partial_z\sigma\partial_z\Xi - \partial_{zz}\Xi, \\
\partial_{hzz} w &= e^{-\sigma}\big(\partial_{hzz}\sigma - 2\partial_{hz}\sigma\partial_z\sigma - \partial_h\sigma\partial_{zz}\sigma + \partial_h\sigma(\partial_z\sigma)^2\big) \\
&\qquad \times\int_0^z \big(e^\sigma\Xi\big)\,dz' + e^{-\sigma}\big(\partial_{zz}\sigma - (\partial_z\sigma)^2\big)\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&\quad + \partial_{hz}\sigma\Xi + \partial_z\sigma\partial_h\Xi - \partial_{hz}\Xi, \\
\partial_{hhz} w &= e^{-\sigma}\big(\partial_{hhz}\sigma - 2\partial_h\sigma\partial_{hz}\sigma - \partial_{hh}\sigma\partial_z\sigma + (\partial_h\sigma)^2\partial_z\sigma\big)\int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&\quad + 2e^{-\sigma}\big(\partial_{hz}\sigma - \partial_h\sigma\partial_z\sigma\big)\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&\quad + e^{-\sigma}\partial_z\sigma\int_0^z \big[e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\Xi + 2e^\sigma\partial_h\sigma\partial_h\Xi + e^\sigma\partial_{hh}\Xi\big]\,dz' \\
&\quad - \partial_{hh}\Xi, \\
\partial_{hhh} w &= e^{-\sigma}\big(\partial_{hhh}\sigma - 3\partial_h\sigma\partial_{hh}\sigma + (\partial_h\sigma)^3\big)\int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&\quad + 3e^{-\sigma}\big(\partial_{hh}\sigma - (\partial_h\sigma)^2\big)\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&\quad + 3e^{-\sigma}\partial_h\sigma\int_0^z \big[e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\Xi + 2e^\sigma\partial_h\sigma\partial_h\Xi + e^\sigma\partial_{hh}\Xi\big]\,dz' \\
&\quad - e^{-\sigma}\int_0^z \big[e^\sigma\big(\partial_{hhh}\sigma + 3\partial_h\sigma\partial_{hh}\sigma + (\partial_h\sigma)^3\big)\Xi + 3e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\partial_h\Xi \\
&\qquad + 3e^\sigma\partial_h\sigma\partial_{hh}\Xi + e^\sigma\partial_{hhh}\Xi\big]\,dz', \\
\partial_{zzzz} w &= e^{-\sigma}\big[\partial_{zzzz}\sigma - 3(\partial_{zz}\sigma)^2 - 4\partial_z\sigma\partial_{zzz}\sigma + 6(\partial_z\sigma)^2\partial_{zz}\sigma - (\partial_z\sigma)^4\big] \\
&\qquad \times\int_0^z \big(e^\sigma\Xi\big)\,dz' + \big(3\partial_{zzz}\sigma - 5\partial_z\sigma\partial_{zz}\sigma + (\partial_z\sigma)^3\big)\Xi \\
&\quad + \big(3\partial_{zz}\sigma - (\partial_z\sigma)^2\big)\partial_z\Xi + \partial_z\sigma\partial_{zz}\Xi - \partial_{zzz}\Xi, \\
\partial_{hzzz} w &= e^{-\sigma}\big[\partial_{hzzz}\sigma - 3\partial_{hzz}\sigma\partial_z\sigma - 3\partial_{hz}\sigma\partial_{zz}\sigma - \partial_h\sigma\partial_{zzz}\sigma
\end{aligned}
$$

$$
\begin{aligned}
&\quad + 3\partial_{hz}\sigma(\partial_z\sigma)^2 + 3\partial_h\sigma\partial_z\sigma\partial_{zz}\sigma - \partial_h\sigma(\partial_z\sigma)^3\big] \times \int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&+ e^{-\sigma}\big[\partial_{zzz}\sigma - 3\partial_z\sigma\partial_{zz}\sigma + (\partial_z\sigma)^3\big]\int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&+ \big(2\partial_{hzz}\sigma - 2\partial_{hz}\sigma\partial_z\sigma\big)\Xi + \big(2\partial_{zz}\sigma - (\partial_z\sigma)^2\big)\partial_h\Xi + \partial_{hz}\sigma\partial_z\Xi \\
&+ \partial_z\sigma\partial_{hz}\Xi - \partial_{hzz}\Xi,
\end{aligned}
$$

$$
\begin{aligned}
\partial_{hhzz}w = {}& e^{-\sigma}\big[\partial_{hhzz}\sigma - 2(\partial_{hz}\sigma)^2 - 2\partial_h\sigma\partial_{hzz}\sigma - 2\partial_{hhz}\sigma\partial_z\sigma \\
&\quad - \partial_{hh}\sigma\partial_{zz}\sigma + 4\partial_h\sigma\partial_z\sigma\partial_{hz}\sigma + (\partial_h\sigma)^2\partial_{zz}\sigma + \partial_{hh}\sigma(\partial_z\sigma)^2 \\
&\quad - (\partial_h\sigma)^2(\partial_z\sigma)^2\big] \times \int_0^z \big(e^\sigma\Xi\big)\,dz' + 2e^{-\sigma}\big[\partial_{hzz}\sigma - 2\partial_{hz}\sigma\partial_z\sigma \\
&\quad - \partial_h\sigma\partial_{zz}\sigma + \partial_h\sigma(\partial_z\sigma)^2\big] \times \int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' \\
&+ e^{-\sigma}\big[\partial_{zz}\sigma - (\partial_z\sigma)^2\big]\int_0^z \big[e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\Xi + 2e^\sigma\partial_h\sigma\partial_h\Xi \\
&\quad + e^\sigma\partial_{hh}\Xi\big]\,dz' + \partial_{hhz}\sigma\Xi + 2\partial_{hz}\sigma\partial_h\Xi + \partial_z\sigma\partial_{hh}\Xi - \partial_{hhz}\Xi,
\end{aligned}
$$

$$
\begin{aligned}
\partial_{hhhz}w = {}& e^{-\sigma}\big[\partial_{hhhz}\sigma - 3\partial_{hz}\sigma\partial_{hh}\sigma - 3\partial_h\sigma\partial_{hhz}\sigma + 3(\partial_h\sigma)^2\partial_{hz}\sigma \\
&\quad - \partial_{hhh}\sigma\partial_z\sigma + 3\partial_h\sigma\partial_z\sigma\partial_{hh}\sigma - (\partial_h\sigma)^3\partial_z\sigma\big] \times \int_0^z \big(e^\sigma\Xi\big)\,dz' \\
&+ 3e^{-\sigma}\big[\partial_{hhz}\sigma - 2\partial_h\sigma\partial_{hz}\sigma - \partial_{hh}\sigma\partial_z\sigma + (\partial_h\sigma)^2\partial_z\sigma\big] \\
&\quad \times \int_0^z \big(e^\sigma\partial_h\sigma\Xi + e^\sigma\partial_h\Xi\big)\,dz' + 3e^{-\sigma}\big[\partial_{hz}\sigma - \partial_h\sigma\partial_z\sigma\big] \\
&\quad \times \int_0^z \big[e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\Xi + 2e^\sigma\partial_h\sigma\partial_h\Xi + e^\sigma\partial_{hh}\Xi\big]\,dz' \\
&+ e^{-\sigma}\partial_z\sigma\int_0^z \big[e^\sigma\big(\partial_{hhh}\sigma + 3\partial_h\sigma\partial_{hh}\sigma + (\partial_h\sigma)^3\big)\Xi \\
&\quad + 3e^\sigma\big(\partial_{hh}\sigma + (\partial_h\sigma)^2\big)\partial_h\Xi + 3e^\sigma\partial_h\sigma\partial_{hh}\Xi + e^\sigma\partial_{hhh}\Xi\big]\,dz' - \partial_{hhh}\Xi.
\end{aligned}
$$

Next, after applying the Hölder, Minkowski, the Sobolev embedding inequalities, and that H^m is an algebra for $m \geq 2$, we obtain

$$
\begin{aligned}
\|w\|_{L^2} &\lesssim \|e^{-\sigma}\|_{L^\infty}\|e^\sigma\big(\partial_t\sigma + v\cdot\nabla_h\sigma + \mathrm{div}_h v\big)\|_{L^2} \\
&\lesssim \|e^{-\sigma}\|_{L^\infty}\|e^\sigma\|_{L^\infty}\big(\|\partial_t\sigma\|_{L^2} + \|v\|_{L^6}\|\nabla_h\sigma\|_{L^3} + \|\nabla_h v\|_{L^2}\big) \\
&\lesssim e^{2\|\sigma\|_{H^2}}\big(\|\partial_t\sigma\|_{L^2} + \|v\|_{H^1}\|\sigma\|_{H^2} + \|v\|_{H^1}\big).
\end{aligned} \tag{31}
$$

Similarly, with details omitted, one can also show,

$$
\begin{aligned}
&\|\partial_z w, \nabla_h w, \partial_{zz}w, \nabla_h\partial_z w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}}\big(\|\sigma\|_{H^3}^2 + 1\big) \\
&\qquad \times \big(\|\partial_t\sigma\|_{H^1} + \|\sigma\|_{H^3}\|v\|_{H^1} + \|v\|_{H^2}\big),
\end{aligned} \tag{32}
$$

$$
\|\nabla_h^2 w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}}\big(\|\sigma\|_{H^3}^2 + 1\big)\big(\|\partial_t\sigma\|_{H^2} + \|\sigma\|_{H^3}\|v\|_{H^2} + \|v\|_{H^3}\big), \tag{33}
$$

$$
\|\nabla^2\partial_z w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}}\big(\|\sigma\|_{H^3}^3 + 1\big)\big(\|\partial_t\sigma\|_{H^2} + \|\sigma\|_{H^3}\|v\|_{H^2} + \|v\|_{H^3}\big), \tag{34}
$$

$$\|\nabla_h^3 w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^3}^3 + 1\big)\big(\|\partial_t\sigma\|_{H^3} + \|\sigma\|_{H^4}\|v\|_{H^3} + \|v\|_{H^4}\big), \tag{35}$$

$$\|\nabla^3 \partial_z w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^4}^4 + 1\big)\big(\|\partial_t\sigma\|_{H^3} + \|\sigma\|_{H^4}\|v\|_{H^3} + \|v\|_{H^4}\big). \tag{36}$$

We summarize inequalities (28) and (29) from inequalities (31)–(36). □

In view of (28) and (29), the estimates for w can be bounded by the estimates for $\partial_t\sigma$, which may not be bounded and induce fast oscillation as $\varepsilon \to 0^+$. Therefore, we will need to obtain the uniform estimates for $\partial_t\sigma$. To achieve this goal, we formulate a strongly damped wave equation for σ in the following (see (38), below). In fact, after applying ∂_t to $(4)_1$, div_h to $(4)_2$, and ∂_z to $(4)_3$, and combining the resultant equations, we obtain, with direct calculations,

$$\begin{aligned}
&\partial_{tt}\sigma + v\cdot\nabla_h\partial_t\sigma + w\partial_z\partial_t\sigma + \partial_t v\cdot\nabla_h\sigma + \partial_t w\partial_z\sigma \\
&\quad = -\mathrm{div}_h\partial_t v - \partial_z\partial_t w = \Delta_h\sigma + \frac{1}{\varepsilon^2}\partial_{zz}\sigma - \Delta(\mathrm{div}_h v + \partial_z w) \\
&\qquad + \mathrm{div}_h\big(v\cdot\nabla_h v + w\partial_z v\big) + \partial_z\big(v\cdot\nabla_h w + w\partial_z w\big).
\end{aligned} \tag{37}$$

On the other hand, from $(4)_1$, we have

$$\begin{aligned}
&-\Delta(\mathrm{div}_h v + \partial_z w) = \Delta(\partial_t\sigma + v\cdot\nabla_h\sigma + w\partial_z\sigma) \\
&= \partial_t\Delta\sigma + (v\cdot\nabla_h + w\partial_z)\Delta\sigma + \Delta v\cdot\nabla_h\sigma + 2\nabla v : \nabla_h\nabla\sigma \\
&\quad + \Delta w\partial_z\sigma + 2\nabla w\cdot\partial_z\nabla\sigma.
\end{aligned}$$

Hence, after substituting the above identity to (37), we arrive at the following equation:

$$\partial_t\big(\partial_t\sigma - \Delta\sigma\big) + (v\cdot\nabla_h + w\partial_z)\big(\partial_t\sigma - \Delta\sigma\big) - \Delta_h\sigma - \frac{1}{\varepsilon^2}\partial_{zz}\sigma = \sum_{j=1}^{4}\mathcal{J}_j, \tag{38}$$

where

$$\begin{aligned}
&\mathcal{J}_1 := \partial_t v\cdot\nabla_h\sigma - \Delta v\cdot\nabla_h\sigma - 2\nabla v : \nabla_h\nabla\sigma, \\
&\mathcal{J}_2 := \partial_t w\partial_z\sigma - \Delta w\partial_z\sigma - 2\nabla w\cdot\partial_z\nabla\sigma, \\
&\mathcal{J}_3 := \mathrm{div}_h\big(v\cdot\nabla_h v + w\partial_z v\big), \\
&\mathcal{J}_4 := \partial_z\big(v\cdot\nabla_h w + w\partial_z w\big).
\end{aligned} \tag{39}$$

Next, we take the L^2–inner product of (38) with $\partial_t\sigma - \Delta\sigma$. Noticing that,

$$\begin{aligned}
&\int \big(-\Delta_h\sigma - \frac{1}{\varepsilon^2}\partial_{zz}\sigma\big)\big(\partial_t\sigma - \Delta\sigma\big)\,d\vec{x} \\
&\quad = \frac{1}{2}\frac{d}{dt}\|\nabla_h\sigma, \frac{\partial_z\sigma}{\varepsilon}\|_{L^2}^2 + \|\nabla\nabla_h\sigma, \frac{\nabla\partial_z\sigma}{\varepsilon}\|_{L^2}^2,
\end{aligned}$$

we arrive at:

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\partial_t\sigma-\Delta\sigma,\nabla_h\sigma,\frac{\partial_z\sigma}{\varepsilon}\|_{L^2}^2+\|\nabla\nabla_h\sigma,\frac{\nabla\partial_z\sigma}{\varepsilon}\|_{L^2}^2\\
&\quad=\frac{1}{2}\int\big(\mathrm{div}_h v+\partial_z w\big)|\partial_t\sigma-\Delta\sigma|^2\,d\vec{x}+\int\big(\partial_t\sigma-\Delta\sigma\big)\big(\sum_{j=1}^4\mathcal{J}_j\big)\,d\vec{x}\\
&\quad\lesssim\|\partial_t\sigma-\Delta\sigma\|_{L^2}^2\|\nabla_h v,\partial_z w\|_{H^2}+\|\partial_t\sigma-\Delta\sigma\|_{L^2}\sum_{j=1}^4\|\mathcal{J}_j\|_{L^2}.
\end{aligned}
\tag{40}
$$

On the other hand, we take the L^2-inner product of (38) with $\partial_t\sigma$, and arrive at:

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\partial_t\sigma,\nabla_h\sigma,\frac{\partial_z\sigma}{\varepsilon}\|_{L^2}^2+\|\nabla\partial_t\sigma\|_{L^2}^2\\
&\quad=\frac{1}{2}\int(\mathrm{div}_h v+\partial_z w)\big(|\partial_t\sigma|^2-2\Delta\sigma\partial_t\sigma\big)\,d\vec{x}\\
&\quad-\int(v\cdot\nabla_h+w\partial_z)\partial_t\sigma\Delta\sigma\,d\vec{x}+\int\partial_t\sigma\big(\sum_{j=1}^4\mathcal{J}_j\big)\,d\vec{x}\\
&\quad\lesssim\big(\|\partial_t\sigma\|_{L^2}^2+\|\Delta\sigma\|_{L^2}^2\big)\|\nabla_h v,\partial_z w\|_{H^2}\\
&\quad+\|v,w\|_{H^2}\|\nabla\partial_t\sigma\|_{L^2}\|\Delta\sigma\|_{L^2}+\|\partial_t\sigma\|_{L^2}\sum_{j=1}^4\|\mathcal{J}_j\|_{L^2},
\end{aligned}
\tag{41}
$$

where we have substituted, after applying integration by parts,

$$
\begin{aligned}
\int\bigg((v\cdot\nabla_h+w\partial_z)\Delta\sigma\bigg)\partial_t\sigma\,d\vec{x}&=-\int(\mathrm{div}_h v+\partial_z w)\Delta\sigma\partial_t\sigma\,d\vec{x}\\
&-\int\bigg((v\cdot\nabla_h+w\partial_z)\partial_t\sigma\bigg)\Delta\sigma\,d\vec{x}.
\end{aligned}
$$

Thus, (40) and (41) yield

$$
\begin{aligned}
&\frac{d}{dt}\big\{\|\partial_t\sigma,\partial_t\sigma-\Delta\sigma\|_{L^2}^2+2\|\nabla_h\sigma,\frac{\partial_z\sigma}{\varepsilon}\|_{L^2}^2\big\}+2\|\nabla\nabla_h\sigma,\frac{\nabla\partial_z\sigma}{\varepsilon},\nabla\partial_t\sigma\|_{L^2}^2\\
&\leq C\big(\|\partial_t\sigma\|_{L^2}^2+\|\partial_t\sigma-\Delta\sigma\|_{L^2}^2\big)\|\nabla_h v,\partial_z w\|_{H^2}\\
&\quad+C\|v,w\|_{H^2}\|\nabla\partial_t\sigma\|_{L^2}\|\Delta\sigma\|_{L^2}\\
&\quad+C\big(\|\partial_t\sigma\|_{L^2}+\|\partial_t\sigma-\Delta\sigma\|_{L^2}\big)\sum_{j=1}^4\|\mathcal{J}_j\|_{L^2}.
\end{aligned}
\tag{42}
$$

We postpone the estimates for $\|\mathcal{J}_j\|_{L^2}$, $j=1,2,3,4$, to section 2.4.

2.3 Higher order uniform-in-ε estimates

In this section, we will derive higher order estimates, i.e., estimates for higher order spatial derivatives, corresponding to (24) and (42). These are the essential parts for closing the estimates for the nonlinearities.

Take the L^2-inner product of $(4)_2$ with $\Delta^2(\partial_t v-\Delta v)$, and $(4)_3$ with $\Delta^2(\partial_t w-\Delta w)$, respectively. This yields, after applying integration by parts,

$$\begin{aligned}\|\Delta(\partial_t v-\Delta v)\|_{L^2}^2 = -\int \Delta(\partial_t v-\Delta v)\cdot\Delta\big(\nabla_h\sigma + v\cdot\nabla_h v \\ + w\partial_z v\big)\,d\vec{x} \leq \frac{1}{2}\|\Delta(\partial_t v-\Delta v)\|_{L^2}^2 \\ + C\big(\|\Delta\nabla_h\sigma\|_{L^2}^2 + \|\Delta(v\cdot\nabla_h v)\|_{L^2}^2 + \|\Delta(w\partial_z v)\|_{L^2}^2\big),\end{aligned} \tag{43}$$

$$\begin{aligned}\varepsilon^2\|\Delta(\partial_t w-\Delta w)\|_{L^2}^2 = -\varepsilon^2\int \Delta(\partial_t w-\Delta w)\times\Delta\big(\frac{1}{\varepsilon^2}\partial_z\sigma \\ + v\cdot\nabla_h w + w\partial_z w\big)\,d\vec{x} \leq \frac{\varepsilon^2}{2}\|\Delta(\partial_t w-\Delta w)\|_{L^2}^2 \\ + \frac{C}{\varepsilon^2}\|\Delta\partial_z\sigma\|_{L^2}^2 + C\varepsilon^2\big(\|\Delta(v\cdot\nabla_h w)\|_{L^2}^2 + \|\Delta(w\partial_z w)\|_{L^2}^2\big).\end{aligned} \tag{44}$$

Notice, similarly to (23), for $\eta = v, w$, after applying integration by parts, we have the identity

$$\|\Delta(\partial_t\eta-\Delta\eta)\|_{L^2}^2 = \frac{d}{dt}\|\nabla^3\eta\|_{L^2}^2 + \|\nabla^2\partial_t\eta\|_{L^2}^2 + \|\nabla^4\eta\|_{L^2}^2. \tag{45}$$

Therefore, (43) and (44) imply, after applying the Hölder inequality, that

$$\begin{aligned}\frac{d}{dt}\|\nabla^3 v, \varepsilon\nabla^3 w\|_{L^2}^2 + \|\nabla^2\partial_t v, \nabla^4 v, \varepsilon\nabla^2\partial_t w, \varepsilon\nabla^4 w\|_{L^2}^2 \\ \leq C\|\nabla^2\nabla_h\sigma, \frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2 + C\sum_{i=1}^{2}\|\mathcal{I}_i\|_{H^2}^2.\end{aligned} \tag{46}$$

Notice, again, that after applying integration by parts, we have the following identities:

$$\begin{aligned}&\int \partial_t(\partial_t\sigma-\Delta\sigma)\Delta^2(\partial_t\sigma-\Delta\sigma)\,d\vec{x} = \frac{1}{2}\frac{d}{dt}\|\nabla^2(\partial_t\sigma-\Delta\sigma)\|_{L^2}^2,\\ &\int(-\Delta_h\sigma - \frac{1}{\varepsilon^2}\partial_{zz}\sigma)\Delta^2(\partial_t\sigma-\Delta\sigma)\,d\vec{x} = \frac{1}{2}\frac{d}{dt}\|\nabla^2\nabla_h\sigma, \frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2\\ &\quad + \|\nabla^3\nabla_h\sigma, \frac{\nabla^3\partial_z\sigma}{\varepsilon}\|_{L^2}^2,\\ &\int(v\cdot\nabla_h + w\partial_z)\big(\partial_t\sigma-\Delta\sigma\big)\Delta^2(\partial_t\sigma-\Delta\sigma)\,d\vec{x} = \int\Delta(\partial_t\sigma-\Delta\sigma)\\ &\qquad\times\Delta\big[(v\cdot\nabla_h + w\partial_z)\big(\partial_t\sigma-\Delta\sigma\big)\big]\,d\vec{x}\\ &\quad = \underbrace{\int(v\cdot\nabla_h + w\partial_z)\Delta(\partial_t\sigma-\Delta\sigma)\Delta(\partial_t\sigma-\Delta\sigma)\,d\vec{x}}_{=-\frac{1}{2}\int(\mathrm{div}_h v+\partial_z w)|\Delta(\partial_t\sigma-\Delta\sigma)|^2\,d\vec{x}}\\ &\quad + 2\sum_{\partial\in\{\partial_h,\partial_z\}}\int(\partial v\cdot\nabla_h + \partial w\partial_z)\big(\partial_t\partial\sigma - \Delta\partial\sigma\big)\Delta(\partial_t\sigma-\Delta\sigma)\,d\vec{x}\end{aligned}$$

$$+ \int (\Delta v \cdot \nabla_h + \Delta w \partial_z)\big(\partial_t \sigma - \Delta\sigma\big)\Delta(\partial_t\sigma - \Delta\sigma)\, d\vec{x}.$$

Thus, taking the L^2-inner product of (38) with $\Delta^2(\partial_t\sigma - \Delta\sigma)$ yields, after substituting the identities above, that

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\nabla^2(\partial_t\sigma - \Delta\sigma), \nabla^2\nabla_h\sigma, \frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2 + \|\nabla^3\nabla_h\sigma, \frac{\nabla^3\partial_z\sigma}{\varepsilon}\|_{L^2}^2 \\
&\quad = \sum_{j=1}^{4}\int \Delta\mathcal{J}_j \times \Delta(\partial_t\sigma - \Delta\sigma)\, d\vec{x} + \frac{1}{2}\int (\mathrm{div}_h v + \partial_z w)|\Delta(\partial_t\sigma - \Delta\sigma)|^2\, d\vec{x} \\
&\quad - 2\sum_{\partial\in\{\partial_h,\partial_z\}}\int (\partial v \cdot \nabla_h + \partial w\partial_z)\big(\partial_t\partial\sigma - \Delta\partial\sigma\big)\Delta(\partial_t\sigma - \Delta\sigma)\, d\vec{x} \\
&\quad - \int (\Delta v \cdot \nabla_h + \Delta w\partial_z)\big(\partial_t\sigma - \Delta\sigma\big)\Delta(\partial_t\sigma - \Delta\sigma)\, d\vec{x} \\
&\quad \lesssim \|\nabla v, \nabla w\|_{L^\infty}\|\nabla^2(\partial_t\sigma - \Delta\sigma)\|_{L^2}^2 + \|\nabla^2 v, \nabla^2 w\|_{L^3}\|\nabla(\partial_t\sigma - \Delta\sigma)\|_{L^6} \\
&\quad \times \|\nabla^2(\partial_t\sigma - \Delta\sigma)\|_{L^2} + \sum_{j=1}^{4}\|\nabla^2\mathcal{J}_j\|_{L^2}\|\nabla^2(\partial_t\sigma - \Delta\sigma)\|_{L^2} \\
&\quad \lesssim \|v, w\|_{H^3}\|\partial_t\sigma - \Delta\sigma\|_{H^2}^2 + \sum_{j=1}^{4}\|\mathcal{J}_j\|_{H^2}\|\partial_t\sigma - \Delta\sigma\|_{H^2},
\end{aligned} \tag{47}$$

where $\mathcal{J}_j$, $j = 1, 2, 3, 4$, are as in (39).

On the other hand, we apply integration by parts to obtain the following integral identities:

$$\begin{aligned}
&\int \partial_t(\partial_t\sigma - \Delta\sigma)\Delta^2\partial_t\sigma\, d\vec{x} = \frac{1}{2}\frac{d}{dt}\|\nabla^2\partial_t\sigma\|_{L^2}^2 + \|\nabla^3\partial_t\sigma\|_{L^2}^2, \\
&\int (-\Delta_h\sigma - \frac{1}{\varepsilon^2}\partial_{zz}\sigma)\Delta^2\partial_t\sigma\, d\vec{x} = \frac{1}{2}\frac{d}{dt}\|\nabla^2\nabla_h\sigma, \frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2, \\
&\int (v \cdot \nabla_h + w\partial_z)\big(\partial_t\sigma - \Delta\sigma\big)\Delta^2\partial_t\sigma\, d\vec{x} \\
&\quad = -\sum_{\partial\in\{\partial_h,\partial_z\}}\int (\partial v \cdot \nabla_h + \partial w\partial_z)(\partial_t\sigma - \Delta\sigma)\partial\Delta\partial_t\sigma\, d\vec{x} \\
&\quad - \sum_{\partial\in\{\partial_h,\partial_z\}}\int (v \cdot \nabla_h + w\partial_z)(\partial_t\partial\sigma - \Delta\partial\sigma)\partial\Delta\partial_t\sigma\, d\vec{x}.
\end{aligned}$$

Therefore, taking the L^2-inner product of (38) with $\Delta^2\partial_t\sigma$ yields, after substituting the identities above, that

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\nabla^2\partial_t\sigma, \nabla^2\nabla_h\sigma, \frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2 + \|\nabla^3\partial_t\sigma\|_{L^2}^2 \\
&= \sum_{j=1}^{4}\int \Delta\mathcal{J}_j \times \Delta\partial_t\sigma\, d\vec{x} + \sum_{\partial\in\{\partial_h,\partial_z\}}\int (v\cdot\nabla_h + w\partial_z)\partial(\partial_t\sigma - \Delta\sigma)\partial\Delta\partial_t\sigma\, d\vec{x} \\
&\quad + \sum_{\partial\in\{\partial_h,\partial_z\}}\int (\partial v\cdot\nabla_h + \partial w\partial_z)(\partial_t\sigma - \Delta\sigma)\partial\Delta\partial_t\sigma\, d\vec{x} \\
&\quad \lesssim \|\nabla v, \nabla w\|_{L^3}\|\nabla(\partial_t\sigma - \Delta\sigma)\|_{L^6}\|\nabla\Delta\partial_t\sigma\|_{L^2} \\
&\quad + \|v, w\|_{L^\infty}\|\nabla^2(\partial_t\sigma - \Delta\sigma)\|_{L^2}\|\nabla\Delta\partial_t\sigma\|_{L^2} + \sum_{j=1}^{4}\|\nabla^2\mathcal{J}_j\|_{L^2}\|\nabla^2\partial_t\sigma\|_{L^2} \\
&\quad \lesssim \|v, w\|_{H^2}\|\partial_t\sigma - \Delta\sigma\|_{H^2}\|\partial_t\sigma\|_{H^3} + \sum_{j=1}^{4}\|\mathcal{J}_j\|_{H^2}\|\partial_t\sigma\|_{H^2},
\end{aligned}
\tag{48}
$$

where $\mathcal{J}_j$, $j = 1, 2, 3, 4$, are as in (39).

2.4 Uniform-in-ε estimates for the nonlinearities

In this section, we will estimate $\|\mathcal{I}_i\|_{H^2}, \|\mathcal{J}_j\|_{H^2}$, $i = 1, 2, j = 1, 2, 3, 4$.

Proposition 2. *I_i, J_j, $i = 1, 2$, $j = 1, 2, 3, 4$, given by* (25) *and* (39)*, satisfy the following estimates:*

$$
\|\mathcal{I}_i\|_{H^2} \leq CE^2, \qquad i = 1, 2, \tag{49}
$$

$$
\|\mathcal{J}_j\|_{H^2} \leq CE(E + D), \qquad j = 1, 2, 3, 4, \tag{50}
$$

where E and D are as in (12) *and* (13)*, respectively, and $C \in (0, \infty)$ is some generic constant, independent of ε.*

Proof. We will only sketch the estimates for

$$
\|\nabla^2\mathcal{I}_i, \nabla^2\mathcal{J}_j\|_{L^2}, \quad i = 1, 2, \; j = 1, 2, 3, 4.
$$

The rests of (49) and (50) can be derived via similar arguments. We organize the proof in the following order:

$$
\nabla^2\mathcal{J}_4, \nabla^2\mathcal{J}_3, \nabla^2\mathcal{J}_1, \nabla^2\mathcal{J}_2, \nabla^2\mathcal{I}_1, \nabla^2\mathcal{I}_2.
$$

Estimate for $\|\nabla^2\mathcal{J}_4\|_{L^2}$: From the definition, one can write

$$
\mathcal{J}_4 = \underbrace{\partial_z(v\cdot\nabla_h w)}_{\mathcal{J}_{41}} + \underbrace{\partial_z(w\partial_z w)}_{\mathcal{J}_{42}}.
$$

Then applying the Hölder and Sobolev embedding inequalities yields

$$
\|\nabla^2\mathcal{J}_{41}\|_{L^2} \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}}\Big(\|\partial^2\partial_z v\cdot\nabla_h w\|_{L^2} + \|\partial\partial_z v\cdot\nabla_h\partial w\|_{L^2}
$$

$$
\begin{aligned}
&+ \|\partial_z v \cdot \nabla_h \partial^2 w\|_{L^2} + \|v \cdot \nabla_h \partial_z \partial^2 w\|_{L^2} + \|\partial v \cdot \nabla_h \partial_z \partial w\|_{L^2} \\
&+ \|\partial^2 v \cdot \nabla_h \partial_z w\|_{L^2}\Big) \lesssim \sum_{\partial \in \{\partial_h, \partial_z\}} \Big(\|\partial^2 \partial_z v\|_{L^2} \|\nabla_h w\|_{L^\infty} \\
&+ \|\partial \partial_z v\|_{L^6} \|\nabla_h \partial w\|_{L^3} + \|\partial_z v\|_{L^\infty} \|\nabla_h \partial^2 w\|_{L^2} \\
&+ \|v\|_{L^\infty} \|\nabla_h \partial_z \partial^2 w\|_{L^2} + \|\partial v\|_{L^3} \|\nabla_h \partial_z \partial w\|_{L^6} \\
&+ \|\partial^2 v\|_{L^2} \|\nabla_h \partial_z w\|_{L^\infty}\Big) \lesssim \|v\|_{H^3} \|w\|_{H^3} + \|v\|_{H^2} \|\partial_z w\|_{H^3} \\
&\lesssim ED,
\end{aligned}
$$

$$
\begin{aligned}
\|\nabla^2 \mathcal{J}_{42}\|_{L^2} \lesssim & \sum_{\partial \in \{\partial_h, \partial_z\}} \Big(\|\partial^2 \partial_z w \partial_z w\|_{L^2} + \|\partial \partial_z w \partial_z \partial w\|_{L^2} \\
&+ \|\partial_z w \partial_z \partial^2 w\|_{L^2} + \|\partial^2 w \partial_{zz} w\|_{L^2} + \|\partial w \partial_{zz} \partial w\|_{L^2} \\
&+ \|w \partial_{zz} \partial^2 w\|_{L^2}\Big) \lesssim \sum_{\partial \in \{\partial_h, \partial_z\}} \Big(\|\partial^2 \partial_z w\|_{L^2} \|\partial_z w\|_{L^\infty} \\
&+ \|\partial \partial_z w\|_{L^3} \|\partial_z \partial w\|_{L^6} + \|\partial_z w\|_{L^\infty} \|\partial_z \partial^2 w\|_{L^2} \\
&+ \|\partial^2 w\|_{L^2} \|\partial_{zz} w\|_{L^\infty} + \|\partial w\|_{L^3} \|\partial_{zz} \partial w\|_{L^6} \\
&+ \|w\|_{L^\infty} \|\partial_{zz} \partial^2 w\|_{L^2}\Big) \lesssim \|\partial_z w\|_{H^2} \|\partial_z w\|_{H^2} + \|w\|_{H^2} \|\partial_z w\|_{H^3} \\
&\lesssim E^2 + ED.
\end{aligned}
$$

Estimate for $\|\nabla^2 \mathcal{J}_3\|_{L^2}$: We write

$$
\mathcal{J}_3 = \underbrace{\operatorname{div}_h(v \cdot \nabla_h v)}_{\mathcal{J}_{31}} + \underbrace{\operatorname{div}_h(w \partial_z v)}_{\mathcal{J}_{32}}.
$$

Then applying the Hölder and Sobolev embedding inequalities yields

$$
\begin{aligned}
\|\nabla^2 \mathcal{J}_{31}\|_{L^2} \lesssim & \sum_{\partial \in \{\partial_h, \partial_z\}} \Big(\|\partial^3 v \cdot \nabla_h v\|_{L^2} + \|\partial^2 v \cdot \nabla_h \partial v\|_{L^2} \\
&+ \|\partial v \cdot \nabla_h \partial^2 v\|_{L^2} + \|v \cdot \nabla_h \partial^3 v\|_{L^2}\Big) \\
\lesssim & \sum_{\partial \in \{\partial_h, \partial_z\}} \Big(\|\partial^3 v\|_{L^2} \|\nabla_h v\|_{L^\infty} + \|\partial^2 v\|_{L^6} \|\nabla_h \partial v\|_{L^3} \\
&+ \|\partial v\|_{L^\infty} \|\nabla_h \partial^2 v\|_{L^2} + \|v\|_{L^\infty} \|\nabla_h \partial^3 v\|_{L^2}\Big) \\
\lesssim & \|v\|_{H^3} \|v\|_{H^3} + \|v\|_{H^2} \|v\|_{H^4} \lesssim E^2 + ED, \\
\|\nabla^2 \mathcal{J}_{32}\|_{L^2} \lesssim & \|w\|_{H^3} \|v\|_{H^3} + \|w\|_{H^2} \|v\|_{H^4} \lesssim ED.
\end{aligned}
$$

Estimate for $\|\nabla^2 \mathcal{J}_1\|_{L^2}$:

$$
\mathcal{J}_1 = \underbrace{\partial_t v \cdot \nabla_h \sigma}_{\mathcal{J}_{11}} - \underbrace{\Delta v \cdot \nabla_h \sigma}_{\mathcal{J}_{12}} - 2 \underbrace{\nabla v : \nabla_h \nabla \sigma}_{\mathcal{J}_{13}}.
$$

We list the estimates in the following:

$$
\begin{aligned}
\|\nabla^2 \mathcal{J}_{11}\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \partial_t v \cdot \nabla_h \sigma\|_{L^2} + \|\partial \partial_t v \cdot \nabla_h \partial \sigma\|_{L^2} \\
& + \|\partial_t v \cdot \nabla_h \partial^2 \sigma\|_{L^2} \Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \partial_t v\|_{L^2} \|\nabla_h \sigma\|_{L^\infty} \\
& + \|\partial \partial_t v\|_{L^3} \|\nabla_h \partial \sigma\|_{L^6} + \|\partial_t v\|_{L^\infty} \|\nabla_h \partial^2 \sigma\|_{L^2} \Big) \\
\lesssim & \|\partial_t v\|_{H^2} \|\sigma\|_{H^3} \lesssim ED, \\
\|\nabla^2 \mathcal{J}_{12}\|_{L^2} \lesssim & \|\Delta v\|_{H^2} \|\sigma\|_{H^3} \lesssim ED, \\
\|\nabla^2 \mathcal{J}_{13}\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \nabla v : \nabla_h \nabla \sigma\|_{L^2} + \|\partial \nabla v : \nabla_h \nabla \partial \sigma\|_{L^2} \\
& + \|\nabla v : \nabla_h \nabla \partial^2 \sigma\|_{L^2} \Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \nabla v\|_{L^2} \|\nabla_h \nabla \sigma\|_{L^\infty} \\
& + \|\partial \nabla v\|_{L^3} \|\nabla_h \nabla \partial \sigma\|_{L^6} + \|\nabla v\|_{L^\infty} \|\nabla_h \nabla \partial^2 \sigma\|_{L^2} \Big) \\
\lesssim & \|v\|_{H^3} \|\sigma\|_{H^4} \lesssim E^2.
\end{aligned}
$$

Estimate for $\|\nabla^2 \mathcal{J}_2\|_{L^2}$: We write $\mathcal{J}_2$ as

$$
\mathcal{J}_2 = \underbrace{\partial_t(\varepsilon w) \frac{\partial_z \sigma}{\varepsilon}}_{\mathcal{J}_{21}} - \underbrace{\Delta(\varepsilon w) \frac{\partial_z \sigma}{\varepsilon}}_{\mathcal{J}_{22}} - 2 \underbrace{\nabla(\varepsilon w) \cdot \frac{\nabla \partial_z \sigma}{\varepsilon}}_{\mathcal{J}_{23}}.
$$

Then applying the Hölder and Sobolev embedding inequalities implies,

$$
\begin{aligned}
\|\nabla^2 \mathcal{J}_{21}\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \partial_t(\varepsilon w) \frac{\partial_z \sigma}{\varepsilon}\|_{L^2} + \|\partial \partial_t(\varepsilon w) \frac{\partial_z \partial \sigma}{\varepsilon}\|_{L^2} \\
& + \|\partial_t(\varepsilon w) \frac{\partial_z \partial^2 \sigma}{\varepsilon}\|_{L^2} \Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \partial_t(\varepsilon w)\|_{L^2} \|\frac{\partial_z \sigma}{\varepsilon}\|_{L^\infty} \\
& + \|\partial \partial_t(\varepsilon w)\|_{L^6} \|\frac{\partial_z \partial \sigma}{\varepsilon}\|_{L^3} + \|\partial_t(\varepsilon w)\|_{L^\infty} \|\frac{\partial_z \partial^2 \sigma}{\varepsilon}\|_{L^2} \Big) \\
\lesssim & \|\partial_t(\varepsilon w)\|_{H^2} \|\frac{\partial_z \sigma}{\varepsilon}\|_{H^2} \lesssim ED, \\
\|\nabla^2 \mathcal{J}_{22}\|_{L^2} \lesssim & \|\Delta(\varepsilon w)\|_{H^2} \|\frac{\partial_z \sigma}{\varepsilon}\|_{H^2} \lesssim ED, \\
\|\nabla^2 \mathcal{J}_{23}\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \nabla(\varepsilon w) \cdot \frac{\nabla \partial_z \sigma}{\varepsilon}\|_{L^2} + \|\partial \nabla(\varepsilon w) \cdot \frac{\nabla \partial_z \partial \sigma}{\varepsilon}\|_{L^2} \\
& + \|\nabla(\varepsilon w) \cdot \frac{\nabla \partial_z \partial^2 \sigma}{\varepsilon}\|_{L^2} \lesssim \Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}} \Big(\|\partial^2 \nabla(\varepsilon w)\|_{L^2} \|\frac{\nabla \partial_z \sigma}{\varepsilon}\|_{L^\infty}
\end{aligned}
$$

$$+ \|\partial\nabla(\varepsilon w)\|_{L^6}\|\frac{\nabla\partial_z\partial\sigma}{\varepsilon}\|_{L^3} + \|\nabla(\varepsilon w)\|_{L^\infty}\|\frac{\nabla\partial_z\partial^2\sigma}{\varepsilon}\|_{L^2}\Big)$$

$$\lesssim \|\varepsilon w\|_{H^3}\|\frac{\partial_z\sigma}{\varepsilon}\|_{H^3} \lesssim ED.$$

Estimate for $\|\nabla^2\mathcal{I}_1\|_{L^2}, \|\nabla^2\mathcal{I}_2\|_{L^2}$: The estimates are similar. They are,

$$\begin{aligned}
\|\nabla^2\mathcal{I}_1\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}}\Big(\|\partial^2 v\cdot\nabla_h(v,\varepsilon w)\|_{L^2} + \|\partial v\cdot\nabla_h\partial(v,\varepsilon w)\|_{L^2} \\
& + \|v\cdot\nabla_h\partial^2(v,\varepsilon w)\|_{L^2}\Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}}\Big(\|\partial^2 v\|_{L^2}\|\nabla_h(v,\varepsilon w)\|_{L^\infty} \\
& + \|\partial v\|_{L^3}\|\nabla_h\partial(v,\varepsilon w)\|_{L^6} + \|v\|_{L^\infty}\|\nabla_h\partial^2(v,\varepsilon w)\|_{L^2}\Big) \\
& \lesssim \|v\|_{H^2}\|v,\varepsilon w\|_{H^3} \lesssim E^2, \\
\|\nabla^2\mathcal{I}_2\|_{L^2} \lesssim & \sum_{\partial\in\{\partial_h,\partial_z\}}\Big(\|\partial^2 w\partial_z(v,\varepsilon w)\|_{L^2} + \|\partial w\partial_z\partial(v,\varepsilon w)\|_{L^2} \\
& + \|w\partial_z\partial^2(v,\varepsilon w)\|_{L^2}\Big) \lesssim \sum_{\partial\in\{\partial_h,\partial_z\}}\Big(\|\partial^2 w\|_{L^2}\|\partial_z(v,\varepsilon w)\|_{L^\infty} \\
& + \|\partial w\|_{L^3}\|\partial_z\partial(v,\varepsilon w)\|_{L^6} + \|w\|_{L^\infty}\|\partial_z\partial^2(v,\varepsilon w)\|_{L^2}\Big) \\
& \lesssim \|w\|_{H^2}\|v,\varepsilon w\|_{H^3} \lesssim E^2.
\end{aligned}$$

□

2.5 Summary of uniform-in-ε estimates

To summarize the estimates in the previous sections, let

$$\begin{aligned}
\mathcal{E} := & \|v,\nabla v,\varepsilon w,\nabla(\varepsilon w),\sigma,\partial_t\sigma-\Delta\sigma,\partial_t\sigma\|_{L^2}^2 + 2\|\nabla_h\sigma,\frac{\partial_z\sigma}{\varepsilon}\|_{L^2}^2 \\
& + \|\nabla^3 v,\nabla^3(\varepsilon w),\nabla^2\partial_t\sigma,\nabla^2(\partial_t\sigma-\Delta\sigma)\|_{L^2}^2 + 2\|\nabla^2\nabla_h\sigma,\frac{\nabla^2\partial_z\sigma}{\varepsilon}\|_{L^2}^2,
\end{aligned} \tag{51}$$

$$\begin{aligned}
\mathcal{D} := & \|\partial_t v,\nabla^2 v,\partial_t(\varepsilon w),\nabla^2(\varepsilon w)\|_{L^2}^2 + 2\|\nabla\partial_t\sigma,\nabla\nabla_h\sigma,\frac{\nabla\partial_z\sigma}{\varepsilon}\|_{L^2}^2 \\
& + \|\nabla^2\partial_t v,\nabla^4 v,\nabla^2\partial_t(\varepsilon w),\nabla^4(\varepsilon w)\|_{L^2}^2 + 2\|\nabla^3\partial_t\sigma,\nabla^3\nabla_h\sigma,\frac{\nabla^3\partial_z\sigma}{\varepsilon}\|_{L^2}^2.
\end{aligned} \tag{52}$$

Then, from the estimates in (28), (29), the definitions in (12), (13), (51), and (52), one concludes, after applying the triangle inequality and the interpolation inequality for Sobolev spaces, that

$$\begin{aligned}
\mathcal{E} \leq E^2 \leq \mathfrak{N}(\mathcal{E}), \\
D \leq \mathfrak{N}(\mathcal{E},1)\sqrt{\mathcal{D}} + \mathfrak{N}(\mathcal{E}).
\end{aligned} \tag{53}$$

Let us recall once again that $\mathfrak{N}(\cdot)$ is a locally Lipschitz in its argument.

Applying the Cauchy–Schwarz inequality implies that

$$\frac{d}{dt}\|v,\varepsilon w,\sigma\|_{L^2}^2 \leq 2\|v,\varepsilon w,\sigma\|_{L^2}\|\partial_t v,\partial_t(\varepsilon w),\partial_t\sigma\|_{L^2}. \tag{54}$$

Thus, (24), (42), (46), (47), (48), and (54) yield

$$\begin{aligned} \frac{d}{dt}\mathcal{E} + \mathcal{D} &\leq \mathfrak{N}(E) + C(E+E^2)D + C\sum_{i=1}^{2}\|\mathcal{I}_i\|_{H^2}^2 + CE\sum_{j=1}^{4}\|\mathcal{J}_j\|_{H^2} \\ &\leq \mathfrak{N}(E) + \mathfrak{N}(E)D \leq \mathfrak{N}(\mathcal{E}) + \mathfrak{N}(\mathcal{E},1)\sqrt{\mathcal{D}}, \end{aligned} \tag{55}$$

where the second and the last inequalities are consequences of (49), (50), and (53). Therefore, (55) implies, after applying the Cauchy–Schwarz inequality, that

$$\frac{d}{dt}\mathcal{E} + \frac{1}{2}\mathcal{D} \leq \mathfrak{N}(\mathcal{E},1). \tag{56}$$

Then for some $T^* \in (0,\infty)$, independent of ε, provided that $\mathcal{E}(0)$ is finite, we have

$$\sup_{0\leq t\leq T^*}\mathcal{E}(t) + \int_0^{T^*}\mathcal{D}(t)\,dt < \infty. \tag{57}$$

Together with (53), we have

$$\sup_{0\leq t\leq T^*}E^2(t) + \int_0^{T^*}D^2(t)\,dt \leq C < \infty, \tag{58}$$

where $C \in (0,\infty)$ depends only on the initial data and is independent of ε.

3 Convergence to the compressible primitive equations

The definitions of E, D in (12), (13), respectively, and estimate (58) imply that there are a time $T^* \in (0,\infty)$ and a constant $C \in (0,\infty)$, which are independent of ε, such that

$$\begin{aligned} &\|v_\varepsilon,\varepsilon w_\varepsilon\|_{L^\infty(0,T^*;H^3)} + \|\partial_t v_\varepsilon,\partial_t(\varepsilon w_\varepsilon)\|_{L^2(0,T^*;H^2)} + \|v_\varepsilon,\varepsilon w_\varepsilon\|_{L^2(0,T^*;H^4)} \\ &+\|\partial_t\sigma_\varepsilon,\nabla_h\sigma_\varepsilon,\frac{\partial_z\sigma_\varepsilon}{\varepsilon}\|_{L^\infty(0,T^*;H^2)} + \|\sigma_\varepsilon\|_{L^\infty(0,T^*;H^4)} \\ &+\|\partial_t\sigma_\varepsilon,\nabla_h\sigma_\varepsilon,\frac{\partial_z\sigma_\varepsilon}{\varepsilon}\|_{L^2(0,T^*;H^3)} + \|w_\varepsilon,\partial_z w_\varepsilon\|_{L^\infty(0,T^*;H^2)} \\ &+\|w_\varepsilon,\partial_z w_\varepsilon\|_{L^2(0,T^*;H^3)} < C. \end{aligned} \tag{59}$$

Then from (59), one can conclude that there exist

$$\begin{gathered} \sigma^* \in L^\infty(0,T^*;H^4), \quad \partial_t\sigma^* \in L^\infty(0,T^*;H^2)\cap L^2(0,T^*;H^3), \\ v^* \in L^\infty(0,T^*;H^3)\cap L^2(0,T^*;H^4), \quad \partial_t v^* \in L^2(0,T^*;H^2), \\ w^*,\partial_z w^* \in L^\infty(0,T^*;H^2)\cap L^2(0,T^*;H^3), \end{gathered} \tag{60}$$

such that for a subsequence of $\{(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)\}$, as $\varepsilon \to 0^+$, one has,

$$
\begin{aligned}
\sigma_\varepsilon &\overset{*}{\rightharpoonup} \sigma^* && \text{weak-}*\text{ in } L^\infty(0,T^*;H^4),\\
\sigma_\varepsilon &\to \sigma^* && \text{in } L^\infty(0,T^*;H^3)\cap C([0,T^*];H^3),\\
\partial_t\sigma_\varepsilon, w_\varepsilon, \partial_z w_\varepsilon &\overset{*}{\rightharpoonup} \partial_t\sigma^*, w^*, \partial_z w^* && \text{weak-}*\text{ in } L^\infty(0,T^*;H^2),\\
\partial_t\sigma_\varepsilon, w_\varepsilon, \partial_z w_\varepsilon &\rightharpoonup \partial_t\sigma^*, w^*, \partial_z w^* && \text{weakly in } L^2(0,T^*;H^3),\\
v_\varepsilon &\overset{*}{\rightharpoonup} v^* && \text{weak-}*\text{ in } L^\infty(0,T^*;H^3),\\
v_\varepsilon &\to v^* && \text{in } L^\infty(0,T^*;H^2)\cap C([0,T^*];H^2),\\
v_\varepsilon &\rightharpoonup v^* && \text{weakly in } L^2(0,T^*;H^4),\\
\partial_t v_\varepsilon &\rightharpoonup \partial_t v^* && \text{weakly in } L^2(0,T^*;H^2),
\end{aligned}
$$

where we have applied the Aubin compactness Lemma (see, e.g., [54]). Then it is easy to verify that (σ^*, v^*, w^*) satisfies the compressible primitive equations (5), after applying the above weak and strong convergences in each terms in system (4). We omit the details.

4 The rates of convergence in terms of ε

In this section, we aim at investigating the rates of convergence of $\delta\sigma, \delta v, \delta v$ to 0, as $\varepsilon \to 0^+$. In section 4.1, we derive some additional uniform-in-ε estimates, which will be used in section 4.2 to derive estimates for the rates of convergence. In section 4.3, we provide the proofs of some Propositions, which have been used in section 4.1.

We shorten the notations by dropping the subscript ε in $(\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$, i.e., $(\sigma, v, w) = (\sigma_\varepsilon, v_\varepsilon, w_\varepsilon)$.

4.1 Additional uniform-in-ε bounds

The additional uniform-in-ε estimates, in order to obtain the convergence rates, are basically the temporal derivative version of the estimates in section 2. Denote by

$$
\begin{aligned}
\mathcal{E}_1 :=& \|\partial_t v, \nabla\partial_t v, \partial_t(\varepsilon w), \nabla\partial_t(\varepsilon w), \partial_t^2\sigma - \Delta\partial_t\sigma, \partial_t^2\sigma\|_{L^2}^2\\
&+ 2\|\nabla_h\partial_t\sigma, \frac{\partial_z\partial_t\sigma}{\varepsilon}\|_{L^2}^2,
\end{aligned} \tag{61}
$$

$$
\begin{aligned}
\mathcal{D}_1 :=& \|\partial_t^2 v, \nabla^2\partial_t v, \partial_t^2(\varepsilon w), \nabla^2\partial_t(\varepsilon w)\|_{L^2}^2 + 2\|\nabla\nabla_h\partial_t\sigma, \frac{\nabla\partial_z\partial_t\sigma}{\varepsilon}\|_{L^2}^2\\
&+ 2\|\nabla\partial_t^2\sigma\|_{L^2}^2.
\end{aligned} \tag{62}
$$

After applying one temporal derivative to $(4)_2$, $(4)_3$, and (38), we obtain the following set of equations:

$$
\partial_t v_t - \Delta_h v_t - \partial_{zz} v_t = -\partial_t\nabla_h\sigma - \partial_t(v\cdot\nabla_h v) - \partial_t(w\partial_z v), \tag{63}
$$

$$\partial_t w_t - \Delta_h w_t - \partial_{zz} w_t = -\frac{1}{\varepsilon^2}\partial_z \partial_t \sigma - \partial_t (v \cdot \nabla_h w) - \partial_t (w \partial_z w), \tag{64}$$

$$\begin{aligned} &\partial_t \big(\partial_t \sigma_t - \Delta \sigma_t\big) + (v \cdot \nabla_h + w\partial_z)\big(\partial_t \sigma_t - \Delta \sigma_t\big) - \Delta_h \sigma_t - \frac{1}{\varepsilon^2}\partial_{zz}\sigma_t \\ &\quad = -(v_t \cdot \nabla_h + w_t \partial_z)\big(\partial_t \sigma - \Delta \sigma\big) + \sum_{j=1}^{4} (\mathcal{J}_j)_t, \end{aligned} \tag{65}$$

where $\mathcal{J}_j$, $j = 1, 2, 3, 4$, are as in (39). Take the L^2-inner product of (63), (64) with $\partial_t v_t - \Delta v_t$, $\varepsilon^2(\partial_t w_t - \Delta w_t)$, respectively. Then after applying integration by parts and the Hölder inequality in the resultant equations, similar arguments as in from (21) to (24) lead to

$$\begin{aligned} &\frac{d}{dt}\|\nabla \partial_t v, \varepsilon \nabla \partial_t w\|_{L^2}^2 + \|\partial_t^2 v, \nabla^2 \partial_t v, \varepsilon \partial_t^2 w, \varepsilon \nabla^2 \partial_t w\|_{L^2}^2 \\ &\quad \leq C\|\nabla_h \partial_t \sigma, \frac{\partial_z \partial_t \sigma}{\varepsilon}\|_{L^2}^2 + C\sum_{i=1}^{2}\|\partial_t \mathcal{I}_i\|_{L^2}^2, \end{aligned} \tag{66}$$

for some constant $C \in (0, \infty)$, independent of ε. In addition, applying the Cauchy-Schwarz inequality yields,

$$\frac{d}{dt}\|\partial_t v, \partial_t(\varepsilon w)\|_{L^2}^2 \leq 2\|\partial_t v, \partial_t(\varepsilon w)\|_{L^2}\|\partial_t^2 v, \partial_t^2(\varepsilon w)\|_{L^2}. \tag{67}$$

On the other hand, taking the L^2-inner product of (65) with $\partial_t \sigma_t - \Delta \sigma_t$ yields, after applying integration by parts, the Hölder and Sobolev inequalities,

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|\partial_t^2 \sigma - \Delta \partial_t \sigma, \nabla_h \partial_t \sigma, \frac{\partial_z \partial_t \sigma}{\varepsilon}\|_{L^2}^2 + \|\nabla \nabla_h \partial_t \sigma, \frac{\nabla \partial_z \partial_t \sigma}{\varepsilon}\|_{L^2}^2 \\ &= \sum_{j=1}^{4}\int \partial_t \mathcal{J}_j \times (\partial_t^2 \sigma - \Delta \partial_t \sigma)\, d\vec{x} + \frac{1}{2}\int (\mathrm{div}_h v + \partial_z w)|\partial_t^2 \sigma - \Delta \partial_t \sigma|^2\, d\vec{x} \\ &\quad - \int (v_t \cdot \nabla_h + w_t \partial_z)\big(\partial_t \sigma - \Delta \sigma\big)\big(\partial_t^2 \sigma - \Delta \partial_t \sigma\big)\, d\vec{x} \\ &\lesssim \sum_{j=1}^{4}\|\partial_t \mathcal{J}_j\|_{L^2}\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2} + \|\nabla_h v, \partial_z w\|_{L^\infty}\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2}^2 \\ &\quad + \|v_t, w_t\|_{L^3}\|\nabla \partial_t \sigma, \nabla \Delta \sigma\|_{L^6}\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2} \\ &\lesssim \sum_{j=1}^{4}\|\partial_t \mathcal{J}_j\|_{L^2}\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2} + (\|v\|_{H^3} + \|\partial_z w\|_{H^2})\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2}^2 \\ &\quad + \|v_t, w_t\|_{H^1}\|\nabla \partial_t \sigma, \nabla \Delta \sigma\|_{H^1}\|\partial_t^2 \sigma - \Delta \partial_t \sigma\|_{L^2}. \end{aligned} \tag{68}$$

In the meantime, after taking the L^2-inner product of (65) with $\partial_t^2 \sigma$ and applying integration by parts, the Hölder and Sobolev inequalities in the resultant equation, one can show that

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\partial_t^2\sigma, \nabla_h\partial_t\sigma, \frac{\partial_z\partial_t\sigma}{\varepsilon}\|_{L^2}^2 + \|\nabla\partial_t^2\sigma\|_{L^2}^2 \\
&\quad = \sum_{j=1}^{4}\int \partial_t\mathcal{J}_j \times \partial_t^2\sigma \, d\vec{x} + \frac{1}{2}\int (\mathrm{div}_h v + \partial_z w)|\partial_t^2\sigma|^2 \, d\vec{x} \\
&- \int \Big[(\mathrm{div}_h v + \partial_z w)\Delta\partial_t\sigma\Big]\partial_t^2\sigma \, d\vec{x} - \int \Big[(v\cdot\nabla_h + w\partial_z)\partial_t^2\sigma\Big]\Delta\partial_t\sigma \, d\vec{x} \\
&\quad - \int \Big[(v_t\cdot\nabla_h + w_t\partial_z)\big(\partial_t\sigma - \Delta\sigma\big)\Big]\partial_t^2\sigma \, d\vec{x} \\
&\quad \lesssim \sum_{j=1}^{4}\|\partial_t\mathcal{J}_j\|_{L^2}\|\partial_t^2\sigma\|_{L^2} + \|\mathrm{div}_h v, \partial_z w\|_{L^\infty}\|\Delta\partial_t\sigma, \partial_t^2\sigma\|_{L^2}^2 \\
&\quad + \|v, w\|_{L^\infty}\|\nabla\partial_t^2\sigma\|_{L^2}\|\Delta\partial_t\sigma\|_{L^2} \\
&\quad + \|\partial_t v, \partial_t w\|_{L^3}\|\nabla\partial_t\sigma, \nabla\Delta\sigma\|_{L^6}\|\partial_t^2\sigma\|_{L^2} \\
&\quad \lesssim \sum_{j=1}^{4}\|\partial_t\mathcal{J}_j\|_{L^2}\|\partial_t^2\sigma\|_{L^2} + (\|v\|_{H^3} + \|\partial_z w\|_{H^2})\|\Delta\partial_t\sigma, \partial_t^2\sigma\|_{L^2}^2 \\
&\quad + \|v, w\|_{H^2}\|\nabla\partial_t^2\sigma\|_{L^2}\|\Delta\partial_t\sigma\|_{L^2} \\
&\quad + \|\partial_t v, \partial_t w\|_{H^1}\|\nabla\partial_t\sigma, \nabla\Delta\sigma\|_{H^1}\|\partial_t^2\sigma\|_{L^2}.
\end{aligned} \tag{69}
$$

Next, we state the estimates for $\|\partial_t\mathcal{I}_i\|_{L^2}, \|\partial_t\mathcal{J}_j\|_{L^2}, , i = 1,2, j = 1,2,3,4$ in the following:

Proposition 3. *$\mathcal{I}_i, \mathcal{J}_j$, $i = 1,2$, $j = 1,2,3,4$, given by* (25) *and* (39)*, satisfy the following estimates:*

$$
\|\partial_t\mathcal{I}_i\|_{L^2} \leq C(E + E_1)^2, \qquad i = 1,2, \tag{70}
$$
$$
\|\partial_t\mathcal{J}_j\|_{L^2} \leq C(E + E_1)(E + E_1 + D + D_1), \qquad j = 1,2,3,4, \tag{71}
$$

where $C \in (0,\infty)$ is some generic constant, independent of ε. Here, E, D, E_1, D_1 are defined in (12), (13), (14)*, and* (15)*, respectively.*

In addition, we state the estimates for $\|\partial_t w, \partial_z\partial_t w\|_{L^2}, \|\partial_t w, \partial_z\partial_t w\|_{H^1}$ in the following:

Proposition 4. *Let w be given as in* (26)*. Then it holds*

$$
\begin{aligned}
&\|\partial_t w, \partial_z\partial_t w\|_{L^2} \leq Ce^{2\|\sigma\|_{H^2}}\big(\|\sigma\|_{H^3}^2 + \|\partial_t\sigma\|_{H^1}^2 + 1\big)\big(\|\partial_t^2\sigma\|_{L^2} \\
&\qquad + \|\partial_t v\|_{H^1} + \|\partial_t v\|_{H^1}\|\sigma\|_{H^2} + \|v\|_{H^2}\|\partial_t\sigma\|_{H^1} + \|\partial_t\sigma\|_{H^2} \\
&\qquad + \|v\|_{H^3} + \|v\|_{H^2}\|\sigma\|_{H^3}\big), \\
&\|\partial_t w, \partial_z\partial_t w\|_{H^1} \leq Ce^{2\|\sigma\|_{H^2}}\big(\|\sigma\|_{H^3}^3 + \|\partial_t\sigma\|_{H^2}^3 + 1\big)\big(\|\partial_t^2\sigma\|_{H^1} \\
&\qquad + \|\partial_t v\|_{H^2} + \|\partial_t v\|_{H^2}\|\sigma\|_{H^2} + \|v\|_{H^2}\|\partial_t\sigma\|_{H^2} + \|\partial_t\sigma\|_{H^2} \\
&\qquad + \|v\|_{H^3} + \|v\|_{H^3}\|\sigma\|_{H^3}\big),
\end{aligned} \tag{72}
$$

for some generic positive constant $C \in (0, \infty)$, independent of ε. In particular,

$$\begin{aligned} \mathcal{E}_1 \leq E_1^2 \leq \mathfrak{N}(\mathcal{E}, \mathcal{E}_1), \\ D_1 \leq \mathfrak{N}(\mathcal{E}, \mathcal{E}_1, 1)\sqrt{\mathcal{D}_1} + \mathfrak{N}(\mathcal{E}, \mathcal{E}_1). \end{aligned} \tag{73}$$

The proofs of Proposition 3 and Proposition 4 are similar to those of Proposition 2 and Proposition 1, which are postponed to section 4.3.

Now we apply the estimates in Proposition 3. Indeed, (66), (67), (68), (69), (70), and (71) yield

$$\frac{d}{dt}\mathcal{E}_1 + \mathcal{D}_1 \leq \mathfrak{N}(E, E_1)(1 + D + D_1).$$

After substituting (53) and (73) on the right-hand side of the above inequality and applying the Young inequality in the resultant, one gets

$$\frac{d}{dt}\mathcal{E}_1 + \mathcal{D}_1 \leq \mathfrak{N}(\mathcal{E}, \mathcal{E}_1, 1) + \frac{1}{2}\mathcal{D}_1 + \mathcal{D}. \tag{74}$$

Thus, there is some constant $T^{**} \in (0, T^*]$, independent of ε, provided that $\mathcal{E}(0), \mathcal{E}_1(0)$ are finite, such that

$$\sup_{0 \leq t \leq T^{**}} \mathcal{E}_1(t) + \int_0^{T^{**}} \mathcal{D}_1(t) \leq C < \infty, \tag{75}$$

where $C \in (0, \infty)$ depends on the initial data but is independent of ε. Here we have also used (57). It follows from (57), (73), and (75), that

$$\sup_{0 \leq t \leq T^{**}} E_1^2(t) + \int_0^{T^{**}} D_1^2(t)\, dt < \infty, \qquad \text{for some } T^{**} \in (0, T^*]. \tag{76}$$

4.2 Convergence rates in terms of ε

Our goal in this section is to estimate the convergence rates of $\delta\sigma, \delta v, \delta w$ to zero, as $\varepsilon \to 0^+$.

In addition to (58) and (76), let (σ_p, v_p, w_p) be a solution to the hydrostatic system (5), satisfying

$$\begin{aligned} \|\sigma_p\|_{L^\infty(0,T^{**};H^4)} + \|\partial_t \sigma_p\|_{L^\infty(0,T^{**};H^2)} + \|v_p\|_{L^\infty(0,T^{**};H^3)} \\ + \|v_p\|_{L^2(0,T^{**};H^4)} + \|w_p\|_{L^\infty(0,T^{**};H^2)} \leq C < \infty, \end{aligned} \tag{77}$$

for some constant $C \in (0, \infty)$. Such a solution can be obtained as the limit in Theorem 1, restricted to the sub-interval of time $[0, T^{**}] \subset [0, T^*]$, which was established in section 3. Alternatively, one can refer to [41] for the proof of existence of such a solution to the compressible primitive equations. Let

$\mathcal{C} \in (0,\infty)$, independent of ε, be the bound given by (58), (76), and (77), i.e.,

$$
\begin{aligned}
& \|v, \varepsilon w\|_{L^\infty(0,T^{**};H^3)} + \|\partial_t v, \partial_t(\varepsilon w)\|_{L^\infty(0,T^{**};H^1)} + \|v, \varepsilon w\|_{L^2(0,T^{**};H^4)} \\
& + \|\partial_t v, \partial_t(\varepsilon w)\|_{L^2(0,T^{**};H^2)} + \|\partial_t^2 v, \partial_t^2(\varepsilon w)\|_{L^2(0,T^{**};L^2)} \\
& + \|\partial_t \sigma, \nabla_h \sigma, \frac{\partial_z \sigma}{\varepsilon}\|_{L^\infty(0,T^{**};H^2)} + \|\sigma\|_{L^\infty(0,T^{**};H^4)} \\
& + \|\partial_t^2 \sigma, \nabla_h \partial_t \sigma, \frac{\partial_z \partial_t \sigma}{\varepsilon}\|_{L^\infty(0,T^{**};L^2)} + \|\partial_t \sigma\|_{L^\infty(0,T^{**};H^2)} \\
& + \|\partial_t \sigma, \nabla_h \sigma, \frac{\partial_z \sigma}{\varepsilon}\|_{L^2(0,T^{**};H^3)} + \|\partial_t^2 \sigma, \nabla_h \partial_t \sigma, \frac{\partial_z \partial_t \sigma}{\varepsilon}\|_{L^2(0,T^{**};H^1)} \\
& + \|w, \partial_z w\|_{L^\infty(0,T^{**};H^2)} + \|\partial_t w, \partial_z \partial_t w\|_{L^\infty(0,T^{**};L^2)} \\
& + \|w, \partial_z w\|_{L^2(0,T^{**};H^3)} + \|\partial_t w, \partial_z \partial_t w\|_{L^2(0,T^{**};H^1)} \\
& + \|\sigma_p\|_{L^\infty(0,T^{**};H^4)} + \|\partial_t \sigma_p\|_{L^\infty(0,T^{**};H^2)} + \|v_p\|_{L^\infty(0,T^{**};H^3)} \\
& + \|v_p\|_{L^2(0,T^{**};H^4)} + \|w_p\|_{L^\infty(0,T^{**};H^2)} < \mathcal{C}.
\end{aligned} \tag{78}
$$

Using (78) and the triangle inequality, we have

$$
\begin{aligned}
& \|\delta\sigma\|_{L^\infty(0,T^{**};H^4)} + \|\partial_t \delta\sigma\|_{L^\infty(0,T^{**};H^2)} + \|\delta v\|_{L^\infty(0,T^{**};H^3)} \\
& \quad + \|\delta v\|_{L^2(0,T^{**};H^4)} + \|\delta w\|_{L^\infty(0,T^{**};H^2)} < 2\mathcal{C}.
\end{aligned} \tag{79}
$$

Noticing that $\partial_z \sigma_p = 0$ and therefore $\partial_z \sigma = \partial_z \delta\sigma$, (78) implies that

$$
\begin{aligned}
& \|\partial_z \delta\sigma\|_{L^\infty(0,T^{**};H^2)} + \|\partial_z \delta\sigma\|_{L^2(0,T^{**};H^3)} \\
& \quad + \|\partial_z \partial_t \delta\sigma\|_{L^\infty(0,T^{**},L^2)} + \|\partial_z \partial_t \delta\sigma\|_{L^2(0,T^{**},H^1)} \leq 2\varepsilon\mathcal{C} = \mathcal{O}(\varepsilon).
\end{aligned} \tag{80}
$$

This gives us partial information of the convergence rates.

Furthermore, we separate $\delta\sigma$ by its vertical average and fluctuation, defined as in (11), i.e.,

$$
\delta\sigma = \overline{\delta\sigma} + \widetilde{\delta\sigma}. \tag{81}
$$

Then the following inequalities are ture:

$$
\|\widetilde{\delta\sigma}\|_{L^2} \lesssim \|\partial_z \delta\sigma\|_{L^2}, \tag{82}
$$

$$
\|\nabla_h \widetilde{\delta\sigma}\|_{L^2} \lesssim \|\partial_z \nabla_h \delta\sigma\|_{L^2}, \tag{83}
$$

$$
\|\widetilde{\partial_t \delta\sigma}\|_{L^2} \lesssim \|\partial_z \partial_t \delta\sigma\|_{L^2}. \tag{84}
$$

The proof of (82)–(84) follows directly from applying the one-dimensional Poincaré inequality in the vertical direction, i.e.,

$$
\|\widetilde{\phi}(\cdot)\|_{L^2(\mathbb{T})} = \|\phi(\cdot) - \overline{\phi}\|_{L^2(\mathbb{T})} \lesssim \|\partial_z \phi(\cdot)\|_{L^2(\mathbb{T})}.
$$

We establish the rest of the required estimates in the following steps: the estimate of $\overline{\delta\sigma}$; the estimate of δv; summary of the estimates; convergence rates via interpolation.

Step 1: estimates of $\overline{\delta\sigma}$. Taking the vertical-average of $(9)_1$ leads to

$$
\partial_t \overline{\delta\sigma} + \overline{v \cdot \nabla_h \delta\sigma} + \overline{\delta v} \cdot \nabla_h \sigma_p + \mathrm{div}_h \overline{\delta v} + \overline{w \partial_z \delta\sigma} = 0. \tag{85}
$$

Then we take the L^2-inner product of (85) with $2\overline{\delta\sigma}$, it follows,

$$
\begin{aligned}
\frac{d}{dt}|\overline{\delta\sigma}|_{L^2}^2 = & -2\int_{\Omega_h} \overline{v\cdot\nabla_h\delta\sigma}\times\overline{\delta\sigma}\,d\vec{x}_h \\
& -2\int_{\Omega_h}(\overline{\delta v}\cdot\nabla_h\sigma_p+\mathrm{div}_h\overline{\delta v}+\overline{w\partial_z\delta\sigma})\times\overline{\delta\sigma}\,d\vec{x}_h =: N_1+N_2.
\end{aligned}
\tag{86}
$$

N_2 can be estimated, after applying the Hölder, Minkowski, and Sobolev embedding inequalities, as

$$
\begin{aligned}
N_2 &\lesssim \big(\|\delta v\|_{L^2}|\nabla_h\sigma_p|_{L^\infty}+\|\nabla\delta v\|_{L^2}+\|w\|_{L^\infty}\|\partial_z\delta\sigma\|_{L^2}\big)|\overline{\delta\sigma}|_{L^2}\\
&\lesssim \big(\|\delta v\|_{L^2}|\nabla_h\sigma_p|_{H^2}+\|\nabla\delta v\|_{L^2}+\|w\|_{H^2}\|\partial_z\delta\sigma\|_{L^2}\big)|\overline{\delta\sigma}|_{L^2}\\
&\lesssim \mathcal{C}|\overline{\delta\sigma}|_{L^2}\|\delta v\|_{L^2}+\|\nabla\delta v\|_{L^2}|\overline{\delta\sigma}|_{L^2}+\varepsilon\mathcal{C}|\overline{\delta\sigma}|_{L^2},
\end{aligned}
$$

where we have substituted the bounds in (78) and (80). On the other hand, N_1 can be written as, after substituting (81) and integration by parts,

$$
\begin{aligned}
N_1 &= -2\int_{\Omega_h}\overline{v}\cdot\nabla_h\overline{\delta\sigma}\times\overline{\delta\sigma}\,d\vec{x}_h-2\int_{\Omega_h}\overline{v\cdot\nabla_h\widetilde{\delta\sigma}}\times\overline{\delta\sigma}\,d\vec{x}_h\\
&= \int_{\Omega_h}\mathrm{div}_h\overline{v}\times|\overline{\delta\sigma}|^2\,d\vec{x}_h-2\int_{\Omega_h}\overline{v\cdot\nabla_h\widetilde{\delta\sigma}}\times\overline{\delta\sigma}\,d\vec{x}_h.
\end{aligned}
$$

Thus applying the Hölder, Minkowski, and Sobolev embedding inequalities and (83) to the above identity yields

$$
\begin{aligned}
N_1 &\lesssim \|\nabla v\|_{L^\infty}|\overline{\delta\sigma}|_{L^2}^2+\|v\|_{L^\infty}\|\nabla_h\widetilde{\delta\sigma}\|_{L^2}|\overline{\delta\sigma}|_{L^2}\\
&\lesssim \|v\|_{H^3}|\overline{\delta\sigma}|_{L^2}^2+\|v\|_{H^2}\|\partial_z\delta\sigma\|_{H^1}|\overline{\delta\sigma}|_{L^2}\lesssim \mathcal{C}|\overline{\delta\sigma}|_{L^2}^2+\varepsilon\mathcal{C}|\overline{\delta\sigma}|_{L^2},
\end{aligned}
$$

where in the last inequality we have substituted the bounds in (78) and (80). Hence, (86) implies that

$$
\frac{d}{dt}|\overline{\delta\sigma}|_{L^2}^2\lesssim \mathcal{C}\big(|\overline{\delta\sigma}|_{L^2}^2+\|\delta v\|_{L^2}^2\big)+\varepsilon\mathcal{C}|\overline{\delta\sigma}|_{L^2}+\|\nabla\delta v\|_{L^2}|\overline{\delta\sigma}|_{L^2}. \tag{87}
$$

Step 2: estimates of δv.

Taking the L^2-inner product of $(9)_2$ with $2\delta v$ leads to

$$
\begin{aligned}
&\frac{d}{dt}\|\delta v\|_{L^2}^2+2\|\nabla\delta v\|_{L^2}^2=\int(\mathrm{div}_h v+\partial_z w)|\delta v|^2\,d\vec{x}\\
&-2\int(\delta v\cdot\nabla_h v_p)\cdot\delta v\,d\vec{x}+2\int\delta\sigma(\mathrm{div}_h\delta v)\,d\vec{x}-\int\delta w(\partial_z v_p\cdot\delta v)\,d\vec{x}\\
&=: N_3+N_4+N_5+N_6.
\end{aligned}
\tag{88}
$$

Again, applying the Hölder and Sobolev embedding inequalities leads to

$$
N_3+N_4+N_5\lesssim(\|\mathrm{div}_h v\|_{L^\infty}+\|\partial_z w\|_{L^\infty}+\|\nabla_h v_p\|_{L^\infty})\|\delta v\|_{L^2}^2
$$

$$+ \|\delta\sigma\|_{L^2}\|\mathrm{div}_h\delta v\|_{L^2} \lesssim (\|v\|_{H^3} + \|\partial_z w\|_{H^2} + \|v_p\|_{H^3})\|\delta v\|_{L^2}^2$$
$$+ \|\delta\sigma\|_{L^2}\|\nabla\delta v\|_{L^2} \lesssim \mathcal{C}\|\delta v\|_{L^2}^2 + (|\overline{\delta\sigma}|_{L^2} + \varepsilon\mathcal{C})\|\nabla\delta v\|_{L^2},$$

where we have substituted (81), (82) and the bounds in (78) and (80). To estimate N_6, we first need to substitute (27) for δw, which leads to

$$N_6 = \int\Big[e^{-\delta\sigma}\Big(\int_0^z \big[e^{\delta\sigma}(\partial_t\delta\sigma + v\cdot\nabla_h\delta\sigma + \delta v\cdot\nabla_h\sigma_p + w_p\partial_z\delta\sigma + \mathrm{div}_h\delta v)\big]\,dz'\Big)$$
$$\times(\partial_z v_p\cdot\delta v)\Big]\,d\vec{x} = \int\Big[e^{-\delta\sigma}\Big(\int_0^z e^{\delta\sigma}\partial_t\delta\sigma\,dz'\Big)\times(\partial_z v_p\cdot\delta v)\Big]\,d\vec{x}$$
$$+\int\Big[e^{-\delta\sigma}\Big(\int_0^z e^{\delta\sigma}(v\cdot\nabla_h\delta\sigma)\,dz'\Big)\times(\partial_z v_p\cdot\delta v)\Big]\,d\vec{x}$$
$$+\int\Big[e^{-\delta\sigma}\Big(\int_0^z e^{\delta\sigma}(\delta v\cdot\nabla_h\sigma_p + w_p\partial_z\delta\sigma + \mathrm{div}_h\delta v)\,dz'\Big)\times(\partial_z v_p\cdot\delta v)\Big]\,d\vec{x}$$
$$=: N_6' + N_6'' + N_6'''.$$

N_6''' can be estimated, after applying the Hölder, Minkowski, and Sobolev embedding inequalities, as

$$N_6''' \lesssim e^{2\|\delta\sigma\|_{L^\infty}}\big(\|\delta v\|_{L^2}\|\nabla_h\sigma_p\|_{L^\infty} + \|w_p\|_{L^\infty}\|\partial_z\delta\sigma\|_{L^2} + \|\nabla_h\delta v\|_{L^2}\big)$$
$$\times\|\partial_z v_p\|_{L^\infty}\|\delta v\|_{L^2} \lesssim e^{2\|\delta\sigma\|_{H^2}}\big(\|\delta v\|_{L^2}\|\sigma_p\|_{H^3} + \|w_p\|_{H^2}\|\partial_z\delta\sigma\|_{L^2}$$
$$+\|\nabla\delta v\|_{L^2}\big)\times\|v_p\|_{H^3}\|\delta v\|_{L^2} \lesssim e^{2\mathcal{C}}\mathcal{C}\|\delta v\|_{L^2}^2 + \varepsilon e^{2\mathcal{C}}\mathcal{C}^2\|\delta v\|_{L^2}$$
$$+ e^{2\mathcal{C}}\mathcal{C}\|\nabla\delta v\|_{L^2}\|\delta v\|_{L^2},$$

where we have substituted the bounds in (78), (79), and (80). On the other hand, applying integration by parts in N_6'' leads to

$$N_6'' = -\int\Big[\Big(\int_0^z(e^{\delta\sigma}\delta\sigma v)\,dz'\Big)\cdot\nabla_h(e^{-\delta\sigma}\partial_z v_p\cdot\delta v)\Big]\,d\vec{x}$$
$$-\int\Big[\Big(\int_0^z\delta\sigma\,\mathrm{div}_h(e^{\delta\sigma}v)\,dz'\Big)\times(e^{-\delta\sigma}\partial_z v_p\cdot\delta v)\Big]\,d\vec{x}.$$

Then, after expanding every term in the above expression and applying the Hölder, Minkowski, and Sobolev embedding inequalities, one can derive

$$N_6'' \lesssim e^{2\|\delta\sigma\|_{L^\infty}}\|v\|_{L^\infty}\|\partial_z v_p\|_{L^\infty}\|\delta\sigma\|_{L^2}\|\nabla\delta v\|_{L^2}$$
$$+ e^{2\|\delta\sigma\|_{L^\infty}}\big(\|v\|_{L^\infty}\|\nabla_h\partial_z v_p\|_{L^\infty} + \|v\|_{L^\infty}\|\partial_z v_p\|_{L^\infty}\|\nabla_h\delta\sigma\|_{L^\infty}$$
$$+\|\nabla_h v\|_{L^\infty}\|\partial_z v_p\|_{L^\infty}\big)\|\delta\sigma\|_{L^2}\|\delta v\|_{L^2}$$
$$\lesssim e^{4\mathcal{C}}(\mathcal{C}^3 + \mathcal{C}^2 + \|v_p\|_{H^4}^2)(|\overline{\delta\sigma}|_{L^2} + \varepsilon\mathcal{C})(\|\nabla\delta v\|_{L^2} + \|\delta v\|_{L^2}),$$

where in the last inequality we have substituted (81), (82), (78), (79), and (80).

In order to estimate N_6', we first rewrite $\partial_t\delta\sigma$ into its average and fluctuation parts; that is,

$$\partial_t\delta\sigma = \overline{\partial_t\delta\sigma} + \widetilde{\partial_t\delta\sigma} = -\big(\overline{v\cdot\nabla_h\delta\sigma} + \overline{\delta v}\cdot\nabla_h\sigma_p + \mathrm{div}_h\overline{\delta v} + \overline{w\partial_z\delta\sigma}\big) + \widetilde{\partial_t\delta\sigma}, \tag{89}$$

where we have substituted (85). Thus N_6' can be written as

$$\begin{aligned} N_6' = &- \int \left[e^{-\delta\sigma} \left(\int_0^z e^{\delta\sigma} \overline{v \cdot \nabla_h \delta\sigma} \, dz' \right) \times (\partial_z v_p \cdot \delta v) \right] d\vec{x} \\ &- \int \left[e^{-\delta\sigma} \left(\int_0^z e^{\delta\sigma} (\overline{\delta v} \cdot \nabla_h \sigma_p + \mathrm{div}_h \overline{\delta v} + \overline{w \partial_z \delta\sigma}) \, dz' \right) \times (\partial_z v_p \cdot \delta v) \right] d\vec{x} \\ &+ \int \left[e^{-\delta\sigma} \left(\int_0^z e^{\delta\sigma} \widetilde{\partial_t \delta\sigma} \, dz' \right) \times (\partial_z v_p \cdot \delta v) \right] d\vec{x} =: N_{6,1}' + N_{6,2}' + N_{6,3}'. \end{aligned}$$

$N_{6,1}'$ and $N_{6,2}'$ can be estimated in the same way as that of N_6'' and N_6''', respectively. $N_{6,3}'$ can be estimated, after applying the Hölder, Minkowski, Sobolev embedding inequalities and (84), as

$$\begin{aligned} N_{6,3}' &\lesssim e^{2\|\delta\sigma\|_{L^\infty}} \|\partial_z v_p\|_{L^\infty} \|\partial_z \partial_t \delta\sigma\|_{L^2} \|\delta v\|_{L^2} \\ &\lesssim e^{2\|\delta\sigma\|_{H^2}} \|v_p\|_{H^3} \|\partial_z \partial_t \delta\sigma\|_{L^2} \|\delta v\|_{L^2} \lesssim \varepsilon e^{4\mathcal{C}} \mathcal{C}^2 \|\delta v\|_{L^2}, \end{aligned}$$

where we have substituted the bounds in (78), (79), and (80) in the last inequality. Consequently, after summing up the above estimates, (88) implies that

$$\begin{aligned} \frac{d}{dt} \|\delta v\|_{L^2}^2 + 2\|\nabla \delta v\|_{L^2}^2 &\lesssim e^{4\mathcal{C}} (1 + \mathcal{C}^3 + \|v_p\|_{H^4}^2)(\|\delta v\|_{L^2} + |\overline{\delta\sigma}|_{L^2} + \varepsilon) \\ &\times (\|\delta v\|_{L^2} + \|\nabla \delta v\|_{L^2}). \end{aligned} \tag{90}$$

Step 3: summary of the estimates and convergence rates. (87) and (90) imply, after applying the Young inequality,

$$\frac{d}{dt} \left(|\overline{\delta\sigma}|_{L^2}^2 + \|\delta v\|_{L^2}^2 \right) + \|\nabla \delta v\|_{L^2}^2 \leq C e^{4\mathcal{C}} (1 + \mathcal{C}^3 + \|v_p\|_{H^4}^2)(|\overline{\delta\sigma}|_{L^2}^2 + \|\delta v\|_{L^2}^2 + \varepsilon^2). \tag{91}$$

Therefore, applying the Grönwall inequality to (91) yields

$$\begin{aligned} \sup_{0 \leq t \leq T^{**}} \left(|\overline{\delta\sigma}(t)|_{L^2}^2 + \|\delta v(t)\|_{L^2}^2 \right) + \int_0^{T^{**}} \|\nabla \delta v(t)\|_{L^2}^2 \, dt \\ \leq \mathfrak{N}(\mathcal{C}, T^{**})(|\overline{\delta\sigma}_0|_{L^2}^2 + \|\delta v_0\|_{L^2}^2 + \varepsilon^2), \end{aligned} \tag{92}$$

where we have used the fact $\|v_p\|_{L^2(0,T^{**};H^4)} < \mathcal{C}$. Together with (80) and (82), (92) implies

$$\begin{aligned} \|\delta\sigma\|_{L^\infty(0,T^{**};L^2)} + \|\delta v\|_{L^\infty(0,T^{**};L^2)} + \|\delta v\|_{L^2(0,T^{**};H^1)} \\ \leq \mathfrak{N}(\mathcal{C}, T^{**})(\|\delta\sigma_0\|_{L^2} + \|\delta v_0\|_{L^2} + \varepsilon) = \mathcal{O}(\varepsilon), \end{aligned} \tag{93}$$

provided that $\|\delta\sigma_0\|_{L^2} + \|\delta v_0\|_{L^2} = \mathcal{O}(\varepsilon)$, which is assumption (19) for well-prepared initial data.

Step 4: convergence rates via interpolation. Notice that, the Gagliardo-Nirenberg interpolation inequality implies

$$\|\delta\sigma\|_{H^1} \lesssim \|\delta\sigma\|_{L^2}^{3/4} \|\delta\sigma\|_{H^4}^{1/4}, \quad \|\delta v\|_{H^1} \lesssim \|\delta v\|_{L^2}^{2/3} \|\delta v\|_{H^3}^{1/3}.$$

Therefore, together with (79) and (93), it follows,

$$\begin{aligned} &\|\delta\sigma\|_{L^\infty(0,T^{**},H^1)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{3/4} = \mathcal{O}(\varepsilon^{3/4}),\\ &\|\delta v\|_{L^\infty(0,T^{**},H^1)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{2/3} = \mathcal{O}(\varepsilon^{2/3}). \end{aligned} \tag{94}$$

From (89), one can derive, after applying the Hölder, Minkowski, and Sobolev embedding inequalities,

$$\begin{aligned} &\|\partial_t\delta\sigma\|_{L^2} \lesssim \|v\|_{L^\infty}\|\nabla_h\delta\sigma\|_{L^2} + \|\delta v\|_{L^2}\|\nabla_h\sigma_p\|_{L^\infty} + \|\nabla_h\delta v\|_{L^2}\\ &\quad + \|w\|_{L^\infty}\|\partial_z\delta\sigma\|_{L^2} + \|\partial_z\partial_t\delta\sigma\|_{L^2} \lesssim \|v\|_{H^2}\|\delta\sigma\|_{H^1} + \|\sigma_p\|_{H^3}\|\delta v\|_{L^2}\\ &\quad + \|\delta v\|_{H^1} + \|w\|_{H^2}\|\partial_z\delta\sigma\|_{L^2} + \|\partial_z\partial_t\delta\sigma\|_{L^2}. \end{aligned} \tag{95}$$

Thus (78), (80), (93) (94), and (95) yield,

$$\begin{aligned} &\|\partial_t\delta\sigma\|_{L^\infty(0,T^{**},L^2)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{2/3} = \mathcal{O}(\varepsilon^{2/3}),\\ &\|\partial_t\delta\sigma\|_{L^2(0,T^{**},L^2)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{3/4} = \mathcal{O}(\varepsilon^{3/4}). \end{aligned} \tag{96}$$

Similarly, from (27), after applying the Hölder, Minkowski, and Sobolev embedding inequalities, one can derive,

$$\begin{aligned} \|\delta w\|_{L^2} \lesssim e^{2\|\delta\sigma\|_{H^2}}\big(&\|\partial_t\delta\sigma\|_{L^2} + \|v\|_{H^2}\|\delta\sigma\|_{H^1} + \|\sigma_p\|_{H^3}\|\delta v\|_{L^2}\\ &+ \|\delta v\|_{H^1} + \|w_p\|_{H^2}\|\partial_z\delta\sigma\|_{L^2}\big). \end{aligned} \tag{97}$$

Therefore, substituting (78), (79), (80), (93), (94), and (96) to (97) yields,

$$\begin{aligned} &\|\delta w\|_{L^\infty(0,T^{**};L^2)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{2/3} = \mathcal{O}(\varepsilon^{2/3}),\\ &\|\delta w\|_{L^2(0,T^{**};L^2)} \lesssim \mathfrak{N}(\mathcal{C},T^{**})\varepsilon^{3/4} = \mathcal{O}(\varepsilon^{3/4}). \end{aligned} \tag{98}$$

4.3 Proofs of Proposition 3 and Proposition 4

Proof of Proposition 3. The proof is similar to that of Proposition 2. We list the estimates for readers' convenience:

$$\begin{aligned} \|\partial_t\mathcal{I}_1\|_{L^2} &\lesssim \|\partial_t v\|_{L^6}\|\nabla v, \nabla(\varepsilon w)\|_{L^3} + \|v\|_{L^\infty}\|\nabla\partial_t v, \nabla\partial_t(\varepsilon w)\|_{L^2}\\ &\lesssim \|\partial_t v, \partial_t(\varepsilon w)\|_{H^1}\|v,\varepsilon w\|_{H^2} \lesssim E_1 E,\\ \|\partial_t\mathcal{I}_2\|_{L^2} &\lesssim \|\partial_t w\|_{L^2}\|\partial_z(\varepsilon w), \partial_z v\|_{L^\infty} + \|w\|_{L^\infty}\|\partial_z\partial_t(\varepsilon w), \partial_z\partial_t v\|_{L^2}\\ &\lesssim \|\partial_t w\|_{L^2}\|\varepsilon w, v\|_{H^3} + \|w\|_{H^2}\|\partial_t(\varepsilon w), \partial_t v\|_{H^1} \lesssim E_1 E, \end{aligned}$$

where we have applied the Hölder and Sobolev embedding inequalities. Similarly, we have,

$$\begin{aligned} &\|\partial_t\mathcal{J}_1\|_{L^2} \lesssim \|\partial_t^2 v, \partial_t\Delta v\|_{L^2}\|\nabla_h\sigma\|_{L^\infty} + \|\partial_t v, \Delta v\|_{L^3}\|\nabla_h\partial_t\sigma\|_{L^6}\\ &\quad + \|\partial_t\nabla v\|_{L^2}\|\nabla_h\nabla\sigma\|_{L^\infty} + \|\nabla v\|_{L^\infty}\|\nabla_h\nabla\partial_t\sigma\|_{L^2}\\ &\quad \lesssim E(E + E_1 + D + D_1), \end{aligned}$$

$$
\begin{aligned}
&\|\partial_t \mathcal{J}_2\|_{L^2} \lesssim \|\partial_t^2(\varepsilon w), \partial_t \Delta(\varepsilon w)\|_{L^2} \|\frac{\partial_z \sigma}{\varepsilon}\|_{L^\infty} + \|\partial_t(\varepsilon w), \Delta(\varepsilon w)\|_{L^3} \|\frac{\partial_z \partial_t \sigma}{\varepsilon}\|_{L^6} \\
&\quad + \|\nabla \partial_t(\varepsilon w)\|_{L^3} \|\frac{\nabla \partial_z \sigma}{\varepsilon}\|_{L^6} + \|\nabla(\varepsilon w)\|_{L^2} \|\frac{\nabla \partial_z \partial_t \sigma}{\varepsilon}\|_{L^2} \\
&\quad \lesssim E(E + E_1 + D + D_1), \\
&\|\partial_t \mathcal{J}_3\|_{L^2} \lesssim \|\nabla \partial_t v\|_{L^2} \|\nabla_h v\|_{L^\infty} + \|\nabla v\|_{L^\infty} \|\nabla_h \partial_t v\|_{L^2} + \|\partial_t v\|_{L^\infty} \|\nabla \nabla_h v\|_{L^2} \\
&\quad + \|v\|_{L^\infty} \|\nabla \nabla_h \partial_t v\|_{L^2} + \|\nabla \partial_t w\|_{L^2} \|\partial_z v\|_{L^\infty} + \|\nabla w\|_{L^\infty} \|\partial_z \partial_t v\|_{L^2} \\
&\quad + \|\partial_t w\|_{L^3} \|\nabla \partial_z v\|_{L^6} + \|w\|_{L^\infty} \|\nabla \partial_z \partial_t v\|_{L^2} \\
&\quad \lesssim (E + E_1)(E + E_1 + D + D_1), \\
&\|\partial_t \mathcal{J}_4\|_{L^2} \lesssim \|\partial_z \partial_t v\|_{L^3} \|\nabla_h w\|_{L^6} + \|\partial_z v\|_{L^\infty} \|\nabla_h \partial_t w\|_{L^2} + \|\partial_t v\|_{L^3} \|\nabla_h \partial_z w\|_{L^6} \\
&\quad + \|v\|_{L^\infty} \|\nabla_h \partial_z \partial_t w\|_{L^2} + \|\partial_z \partial_t w\|_{L^2} \|\partial_z w\|_{L^\infty} + \|\partial_z w\|_{L^\infty} \|\partial_z \partial_t w, \|_{L^2} \\
&\quad + \|\partial_t w\|_{L^3} \|\partial_z^2 w\|_{L^6} + \|w\|_{L^\infty} \|\partial_z^2 \partial_t w\|_{L^2} \\
&\quad \lesssim (E + E_1)(E + E_1 + D + D_1).
\end{aligned}
$$

This completes the proof. □

Proof of Proposition 4. We follow similar steps as in the proof of Proposition 1. Recalling Ξ in (30), applying a temporal derivative to (26) leads to

$$
\partial_t w = -e^{-\sigma} \int_0^z e^{\sigma} (\partial_t \sigma \Xi + \partial_t \Xi) \, dz' + e^{-\sigma} \partial_t \sigma \int_0^z e^{\sigma} \Xi \, dz'.
$$

Similarly,

$$
\begin{aligned}
&\partial_z \partial_t w = e^{-\sigma} \partial_z \sigma \int_0^z e^{\sigma} (\partial_t \sigma \Xi + \partial_t \Xi) \, dz' + e^{-\sigma} (\partial_z \partial_t \sigma - \partial_z \sigma \partial_t \sigma) \int_0^z e^{\sigma} \Xi \, dz' \\
&\quad - \partial_t \Xi, \\
&\partial_h \partial_t w = e^{-\sigma} \partial_h \sigma \int_0^z e^{\sigma} (\partial_t \sigma \Xi + \partial_t \Xi) \, dz' + e^{-\sigma} (\partial_h \partial_t \sigma - \partial_h \sigma \partial_t \sigma) \int_0^z e^{\sigma} \Xi \, dz' \\
&\quad - e^{-\sigma} \int_0^z e^{\sigma} (\partial_h \sigma \partial_t \sigma \Xi + \partial_h \sigma \partial_t \Xi + \partial_h \partial_t \sigma \Xi + \partial_t \sigma \partial_h \Xi + \partial_h \partial_t \Xi) \, dz' \\
&\quad + e^{-\sigma} \partial_t \sigma \int_0^z e^{\sigma} (\partial_h \sigma \Xi + \partial_h \Xi) \, dz', \\
&\partial_{hz} \partial_t w = e^{-\sigma} \partial_z \sigma \int_0^z e^{\sigma} (\partial_h \sigma \partial_t \sigma \Xi + \partial_h \sigma \partial_t \Xi + \partial_t \sigma \partial_h \Xi + \partial_h \partial_t \sigma \Xi + \partial_h \partial_t \Xi) \, dz' \\
&\quad + e^{-\sigma} (\partial_{hz} \sigma - \partial_h \sigma \partial_z \sigma) \int_0^z e^{\sigma} (\partial_t \sigma \Xi + \partial_t \Xi) \, dz' + e^{-\sigma} (\partial_z \partial_t \sigma - \partial_z \sigma \partial_t \sigma) \\
&\quad \times \int_0^z e^{\sigma} (\partial_h \sigma \Xi + \partial_h \Xi) \, dz' + e^{-\sigma} (\partial_{hz} \partial_t \sigma - \partial_{hz} \sigma \partial_t \sigma - \partial_z \sigma \partial_h \partial_t \sigma \\
&\quad - \partial_h \sigma \partial_z \partial_t \sigma + \partial_h \sigma \partial_z \sigma \partial_t \sigma) \int_0^z e^{\sigma} \Xi \, dz' - \partial_h \partial_t \Xi.
\end{aligned}
$$

Therefore, after applying the Hölder, Minkowski, and Sobolev embedding inequalities, one obtains,

$$\begin{aligned}
&\|\partial_t w, \partial_z \partial_t w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^3} + 1\big) \big(\|\partial_t^2 \sigma\|_{L^2} + \|\partial_t v\|_{H^1} \\
&\qquad + \|\partial_t v\|_{H^1} \|\sigma\|_{H^2} + \|v\|_{H^2} \|\partial_t \sigma\|_{H^1} + \|\partial_t \sigma\|_{H^1} (\|\partial_t \sigma\|_{H^2} \qquad (99)\\
&\qquad + \|v\|_{H^3} + \|v\|_{H^2} \|\sigma\|_{H^3})\big), \\
&\|\partial_h \partial_t w, \partial_{hz} \partial_t w\|_{L^2} \lesssim e^{2\|\sigma\|_{H^2}} \big(\|\sigma\|_{H^3}^2 + 1\big) \big(\|\partial_t^2 \sigma\|_{H^1} + \|\partial_t v\|_{H^2} \\
&\qquad + \|\partial_t v\|_{H^2} \|\sigma\|_{H^2} + \|v\|_{H^2} \|\partial_t \sigma\|_{H^2} + (\|\partial_t \sigma\|_{H^2} + 1) \qquad (100)\\
&\qquad \times (\|\partial_t \sigma\|_{H^2} + \|v\|_{H^3} + \|v\|_{H^3} \|\sigma\|_{H^3})\big).
\end{aligned}$$

This completes the proof of Proposition 4. □

Acknowledgements

This work was supported in part by the Einstein Stiftung/Foundation - Berlin, through the Einstein Visiting Fellow Program. X.L. and E.S.T. would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme *"Mathematical aspects of turbulence: where do we stand?"*, where part of the work on this paper was undertaken. This work was supported in part by EPSRC grant no EP/R014604/1. X.L.'s work was partially supported by a grant from the Simons Foundation, during his visit to the Isaac Newton Institute for Mathematical Sciences. The research of E.S.T. has benefited from the inspiring environment of the CRC 1114 "Scaling Cascades in Complex Systems", Project Number 235221301, Project C06, funded by Deutsche Forschungsgemeinschaft (DFG).

References

[1] Pascal Azérad and Francisco Guillén. Mathematical justification of the hydrostatic approximation in the primitive equations of geophysical fluid dynamics. *SIAM J. Math. Anal.*, 33(4):847–859, 2001.

[2] Peter Bella, Eduard Feireisl, and Antonín Novotný. Dimension reduction for compressible viscous fluids. *Acta Appl. Math.*, 134(1):111–121, 2014.

[3] Didier Bresch and Benoît Desjardins. Stabilité de solutions faibles globales pour les équations de Navier–Stokes compressible avec température. *C. R. Acad. Sci. Paris, Ser. I*, 343(3):219–224, 2006.

[4] Didier Bresch and Pierre-Emmanuel Jabin. Global existence of weak solutions for compressible Navier–Stokes equations: Thermodynamically unstable pressure and anisotropic viscous stress tensor. *Ann. Math.*, 188(2):577–684, 2018.

[5] Didier Bresch, Alexis Vasseur, and Cheng Yu. Global existence of entropy-weak solutions to the compressible Navier–Stokes equations with non-linear density dependent viscosities. *J. Eur. Math. Soc.*, 24(5):1791–1837, 2022.

[6] Chongsheng Cao, Jinkai Li, and Edriss S. Titi. Global well-posedness of strong solutions to the 3D primitive equations with horizontal eddy diffusivity. *J. Differential Equations*, 257(11):4108–4132, 2014.

[7] Chongsheng Cao, Jinkai Li, and Edriss S Titi. Local and global well-posedness of strong solutions to the 3D primitive equations with vertical eddy diffusivity. *Arch. Rational Mech. Anal.*, 214(1):35–76, 2014.

[8] Chongsheng Cao, Jinkai Li, and Edriss S. Titi. Global well-posedness of the three-dimensional primitive equations with only horizontal viscosity and diffusion. *Communications on Pure and Applied Mathematics*, 69(8):1492–1531, 2016.

[9] Chongsheng Cao, Jinkai Li, and Edriss S. Titi. Global well-posedness of the 3D primitive equations with horizontal viscosity and vertical diffusivity. *Physica D: Nonlinear Phenomena*, 412:132606, 2020.

[10] Chongsheng Cao, Jinkai Li, and Edriss S. Titi. Strong solutions to the 3D primitive equations with only horizontal dissipation: Near H^1 initial data. *Journal of Functional Analysis*, 272(11):4606–4641, 2017.

[11] Chongsheng Cao and Edriss Titi. Global well-posedness of the three-dimensional viscous primitive equations of large scale ocean and atmosphere dynamics. *Ann. Math.*, 166(1):245–267, 2007.

[12] Chongsheng Cao and Edriss S. Titi. Global well-posedness of the 3D primitive equations with partial vertical turbulence mixing heat diffusion. *Commun. Math. Phys.*, 310(2):537–568, 2012.

[13] Yonggeun Cho, Hi Jun Choe, and Hyunseok Kim. Unique solvability of the initial boundary value problems for compressible viscous fluids. *J. Math. Pures Appl.*, 83(2):243–275, 2004.

[14] Yonggeun Cho and Hyunseok Kim. Existence results for viscous polytropic fluids with vacuum. *J. Differential Equations*, 228(2):377–411, 2006.

[15] Yonggeun Cho and Hyunseok Kim. On classical solutions of the compressible Navier–Stokes equations with nonnegative initial densities. *manuscripta Math.*, 120(1):91–129, 2006.

[16] Mehmet Ersoy and Timack Ngom. Existence of a global weak solution to compressible primitive equations. *C. R. Acad. Sci. Paris, Ser. I*, 350(7–8):379–382, 2012.

[17] Mehmet Ersoy, Timack Ngom, and Mamadou Sy. Compressible primitive equations: Formal derivation and stability of weak solutions. *Nonlinearity*, 24(1):79–96, 2011.

[18] Eduard Feireisl. *Dynamics of Viscous Compressible Fluids*. Oxford Lecture Series in Mathematics and its Applications, 26. Oxford University Press, 2004.

[19] Hongjun Gao, Šárka Nečasová, and Tong Tang. On the hydrostatic approximation of compressible anisotropic Navier–Stokes equations–rigorous justification, Journal of Mathematical Fluid Mechanics, 24(3):86, 2022.

[20] B. V. Gatapov and Aleksandr Vasilevic Kazhikhov. Existence of a global solution to one model problem of atmosphere dynamics. *Siberian Mathematical Journal*, 46(5):805–812, 2005.

[21] David Gerard-Varet, Nader Masmoudi, and Vlad Vicol. Well-posedness of the hydrostatic Navier–Stokes equations. *Anal. PDE*, 13(5):1417–1455, 2020.

[22] Francisco Guillén-González, Nader Masmoudi, and María Ángeles Rodríguez-Bellido. Anisotropic estimates and strong solutions of the primitive equations. *Differential and Integral Equations*, 14(11):1381–1408, 2001.

[23] David Hoff. Global well-posedness of the Cauchy problem for the Navier-Stokes equations of nonisentropic flow with discontinuous initial data. *J. Differential Equations*, 95(1):33–74, 1992.

[24] David Hoff and Denis Serre. The failure of continuous dependence on initial data for the Navier–Stokes equations of compressible flow. *SIAM J. Appl. Math.*, 51(4):887–898, 1991.

[25] Changbing Hu, Roger Temam, and Mohammed Ziane. The primitive equations on the large scale ocean under the small depth hypothesis. *Discrete and Continuous Dynamical Systems*, 9(1):97–131, 2003.

[26] Xiangdi Huang and Jing Li. Existence and blowup behavior of global strong solutions to the two-dimensional barotropic compressible Navier–Stokes system with vacuum and large initial data. *J. Math. Pures Appl.*, 106(1):123–154, 2016.

[27] Xiangdi Huang and Jing Li. Global classical and weak solutions to the three-dimensional full compressible Navier–Stokes system with vacuum and large oscillations. *Arch. Rational Mech. Anal.*, 227(3):995–1059, 2018.

[28] Xiangdi Huang, Jing Li, and Zhouping Xin. Global well-posedness of classical solutions with large oscillations and vacuum to the three-dimensional isentropic compressible Navier–Stokes equations. *Commun. Pure Appl. Math.*, 65(4):549–585, 2012.

[29] Nobutoshi Itaya. On the Cauchy problems for the system of fundamental equations describing the movement of compressible viscous fluid. *Kodai Math. Sem. Rep.*, 23(1):60–120, 1971.

[30] Quansen Jiu, Mingjie Li, and Fengchao Wang. Uniqueness of the global weak solutions to 2D compressible primitive equations. *J. Math. Anal. Appl.*, 461(2):1653–1671, 2018.

[31] Hi Jun Choe and Hyunseok Kim. Strong solutions of the Navier–Stokes equations for isentropic compressible fluids. *J. Differential Equations*, 190(2):504–523, 2003.

[32] Igor Kukavica and Mohammed Ziane. The regularity of solutions of the primitive equations of the ocean in space dimension three. *C. R. Acad. Sci. Paris, Ser. I*, 345(5):257–260, 2007.

[33] Jing Li and Zhouping Xin. Global existence of weak solutions to the barotropic compressible Navier–Stokes flows with degenerate viscosities. Available at arXiv:1504.06826, 2015.

[34] Jinkai Li and Edriss S Titi. Existence and uniqueness of weak solutions to viscous primitive equations for a certain class of discontinuous initial data. *SIAM J. Math. Anal.*, 49(1):1–28, 2017.

[35] Jinkai Li and Edriss S. Titi. The primitive equations as the small aspect ratio limit of the Navier–Stokes equations: Rigorous justification of the hydrostatic approximation. *J. Math. Pures Appl.*, 124:30–58, 2019.

[36] Jinkai Li, Edriss S. Titi, and Guozhi Yuan. The primitive equations approximation of the anisotropic horizontally viscous 3D Navier-Stokes equations. *J. Differential Equations*, 306:492–524, 2022.

[37] Jacques-Louis Lions, Roger Temam, and Shouhong Wang. Geostrophic asymptotics of the primitive equations of the atmosphere. *Topol. Methods Nonlinear Anal.*, 4(2):253–287, 1994.

[38] Jacques-Louis Lions, Roger Temam, and Shouhong Wang. New formulations of the primitive equations of atmosphere and applications. *Nonlinearity*, 5(2):237–288, 1992.

[39] Pierre-Louis Lions. *Mathematical Topics in Fluid Mechanics. Volume 1. Incompressible Models.* Oxford Lecture Series in Mathematics and Its Applications, 3. Oxford University Press, 1996.

[40] Pierre-Louis Lions. *Mathematical Topics in Fluid Mechanics. Volume 2. Compressible Models.* Oxford Lecture Series in Mathematics and its Applications , Vol 2, No 10. Oxford University Press, 1998.

[41] Xin Liu and Edriss S. Titi. Local well-posedness of strong solutions to the three-dimensional compressible primitive equations. *Arch. Rational Mech. Anal.*, 241:729–764, 2021.

[42] Xin Liu and Edriss S. Titi. Global existence of weak solutions to the compressible primitive equations of atmospheric dynamics with degenerate viscosities. *SIAM J. Math. Anal.*, 51(3):1913–1964, 2019.

[43] Xin Liu and Edriss S. Titi. Zero Mach number limit of the compressible primitive equations: Well-prepared initial data. *Arch. Rational Mech. Anal.*, 238:705–747, 2020.

[44] Xin Liu and Edriss S. Titi. Zero Mach Number Limit of the Compressible Primitive Equations: Ill-prepared Initial Data. *J. Differential Equations*, 356:1–58, 2023.

[45] David Maltese and Antonín Novotný. Compressible Navier-Stokes equations on thin domains. *J. Math. Fluid Mech.*, 16(3):571–594, 2014.

[46] Akitaka Matsumura and Takaaki Nishida. The initial value problem for the equations of motion of viscous and heat-conductive gases. *J. Math. Kyoto Univ.*, 20(1):67–104, 1980.

[47] Akitaka Matsumura and Takaaki Nishida. Initial boundary value problems for the equations of motion of compressible viscous and heat-conductive fluids. *Commun. Math. Phys.*, 89:445–464, 1983.

[48] Madalina Petcu and Djoko Wirosoetisno. Sobolev and Gevrey regularity results for the primitive equations in three space dimensions. *Applicable Analysis*, 84(8):769–788, 2005.

[49] Michael Renardy. Ill-posedness of the hydrostatic Euler and Navier–Stokes equations. *Arch. Rational Mech. Anal.*, 194(3):877–886, 2009.

[50] Lewis F. Richardson. *Weather Prediction by Numerical Process.* Dover Publications, Inc.,1965.

[51] James Serrin. On the uniqueness of compressible fluid motions. *Arch. Rational Mech. Anal.*, 3(1):271–288, 1959.

[52] Tong Tang and Šárka Nečasová. Derivation of the inviscid compressible primitive equations. *Applied Mathematics Letters*, 139:108534, 2023.

[53] Atusi Tani. On the first initial-boundary value problem of compressible viscous fluid motion. *Publications of the Research Institute for Mathematical Sciences*, 13(1):193–253, 1977.

[54] Roger Temam. *Navier–Stokes Equations: Theory and Numerical Analysis.* American Mathematical Soc., 2001.

[55] Alexis F. Vasseur and Cheng Yu. Existence of global weak solutions for 3D degenerate compressible Navier–Stokes equations. *Invent. Math.*, 206(3):935–974, 2016.

[56] Vladimir Andreevich Vaĭgant and Aleksandr Vasilevic Kazhikhov. On existence of global solutions to the two-dimensional Navier–Stokes equations for a compressible viscous fluid. *Siberian Mathematical Journal*, 36(6):1108–1141, 1995.

[57] Fengchao Wang, Changsheng Dou, and Quansen Jiu. Global existence of weak solutions to 3D compressible primitive equations with density-dependent viscosity. *Journal of Mathematical Physics*, 61:021507, 2020.

[58] Warren M. Washington and Claire L. Parkinson. *An Introduction to Three-Dimensional Climate Modeling*. University Science Books, 2005.

[59] Zhouping Xin. Blowup of smooth solutions to the compressible Navier–Stokes equation with compact density. *Communications on Pure and Applied Mathematics*, 51(3):229–240, 1998.