Calc. Var. (2022) 61:136
https://doi.org/10.1007/s00526-022-02242-3

Calculus of Variations

Optimal decay for the 3D anisotropic Boussinesq equations near the hydrostatic balance

Ruihong Ji[1] · Li Yan[1] · Jiahong Wu[2]

Received: 10 September 2021 / Accepted: 6 April 2022

Abstract
This paper focuses on the three-dimensional (3D) incompressible anisotropic Boussinesq system with horizontal dissipation. The goal here is to assess the stability property and pinpoint the precise large-time behavior of perturbations near the hydrostatic balance. Important tools such as Schonbek's Fourier splitting method have been developed to understand the large-time behavior of PDE systems with full dissipation, but these tools may not apply directly when the systems are only partially dissipated. This paper solves the stability problem and designs an effective approach to obtain the optimal decay rates for the anisotropic Boussinesq system concerned here. The tool developed in this paper may be useful for many other partially dissipated systems.

Mathematics Subject Classification 35B35 · 35B40 · 35Q35 · 76D03

1 Introduction

This paper intends to understand the stability of the hydrostatic balance or hydrostatic equilibrium and provide optimal estimates on the large-time behavior of perturbations near the hydrostatic balance. There are two distinct motivations for this study. The first is physical. Hydrostatic balance is an important equilibrium of many geophysical fluids. In fact, our atmosphere is mainly in hydrostatic balance, between the upward-directed pressure gradient force and the downward-directed force of gravity. Understanding the stability of pertur-

Communicated by L. Szekelyhidi.

✉ Li Yan
yanli2014@cdut.edu.cn

Ruihong Ji
jiruihong09@cdut.cn

Jiahong Wu
jiahong.wu@okstate.edu

1 Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu 610059, People's Republic of China

2 Department of Mathematics, Oklahoma State University, Stillwater, OK 74078, USA

bations near the hydrostatic equilibrium may help gain insight into certain severe weather phenomena (see, e.g., [17, 24]). The second is mathematical. The partial differential equation (PDE) system concerned here models anisotropic fluids and involve only partial dissipation. Although significant progress has been made on the large-time behavior of fully dissipated PDE systems (see, e.g., [25–27]), but the large-time behavior of anisotropic PDE systems is generally not well-understood and is a very active research topic. This paper offers new ideas and presents a successful story on a partially dissipated Boussinesq system.

The most frequently employed PDE model for geophysical fluids is the Boussinesq system for buoyancy-driven fluids (see, e.g., [8, 11, 20, 24, 34]). The Boussinesq system studied here is for anisotropic fluids and involves only horizontal dissipation,

$$\begin{cases} \partial_t u + u\cdot\nabla u = -\nabla P + \nu\,\Delta_h u + \Theta e_3, & x\in\mathbb{R}^3,\ t>0,\\ \partial_t\Theta + u\cdot\nabla\Theta = \eta\,\Delta_h\Theta, & x\in\mathbb{R}^3,\ t>0,\\ \nabla\cdot u = 0, & x\in\mathbb{R}^3,\ t>0, \end{cases} \tag{1.1}$$

where $u=(u_1,u_2,u_3)$ denotes the velocity field, P the pressure, Θ the temperature, $e_3=(0,0,1)$, and $\nu>0$ and $\eta>0$ are the viscosity and the thermal diffusivity, respectively. Here $\Delta_h=\partial_{x_1x_1}+\partial_{x_2x_2}$ stands for the horizontal Laplacian. For notational convenience, we shall write ∂_i for ∂_{x_i} for $i=1,2,3$, and $\nabla_h=(\partial_1,\partial_2)$. (1.1) arises naturally in the modeling of anisotropic fluids such as the rotating fluids in Ekman layers. A standard reference is Chapter 4 of Pedlosky's book [24].

The hydrostatic balance given by

$$u^{(0)}\equiv(0,0,0),\quad \Theta^{(0)}=x_3,\quad P^{(0)}=\frac{1}{2}x_3^2 \tag{1.2}$$

is a very special steady-state solution of (1.1) with great geophysical and astrophysical importance (see, e.g., [17, 23, 24, 33]). To understand the stability and large-time behavior of perturbations near the hydrostatic balance in (1.2), we consider the equations governing the perturbation (u,θ,p) with $\theta=\Theta-x_3$, $p=P-P^{(0)}$,

$$\begin{cases} \partial_t u + u\cdot\nabla u = -\nabla p + \nu\,\Delta_h u + \theta e_3, & x\in\mathbb{R}^3,\ t>0,\\ \partial_t\theta + u\cdot\nabla\theta + u_3 = \eta\,\Delta_h\theta, & x\in\mathbb{R}^3,\ t>0,\\ \nabla\cdot u = 0, & x\in\mathbb{R}^3,\ t>0,\\ u(x,0)=u_0(x), & \theta(x,0)=\theta_0(x),\quad x\in\mathbb{R}^3. \end{cases} \tag{1.3}$$

We remark that here buoyancy acts in the direction of gravity e_3 and that the gravitational constant is rescaled to 1. Our goal here is to understand the stability problem and give a precise account of the large time behavior of the solutions to (1.3). The large-time behavior problem is not trivial. Due to the presence of the buoyancy forcing term θe_3 in the velocity equation, Sobolev norms and even the L^2-norm of the velocity in (1.1) can grow in time. Brandolese and Schonbek have shown in [6] that the L^2-norm of the velocity to the Boussinesq system with full viscous dissipation and thermal diffusion can grow in time even for very nice initial data (say, data that are smooth, fast spatial decaying and small in some strong norm). Therefore the original system (1.1) is not even stable due to the explicit examples of Brandolese and Schonbek [6]. Perturbing near the hydrostatic balance generates the new term u_3 in (1.3), which helps balance the buoyancy force. In fact, the buoyancy term is canceled by the new term u_3 in the process of estimating $\|u\|_{L^2}^2+\|\theta\|_{L^2}^2$ or Sobolev norms. We caution that, if we take $\Theta^{(0)}=-x_3$, then we may have instability [12]. However, due to the lack of the vertical dissipation, the system (1.3) is degenerate. Indeed, the wave equations in (1.7) converted from this system are degenerate wave equations. A quick inspection on the spectra of the linearized system of (1.3) would shed light on the nature of this degeneracy.

To separate the linear parts in (1.3) from the nonlinear parts, we apply the Helmholtz-Leray projection $\mathbb{P} = I - \nabla\Delta^{-1}\nabla\cdot$ to the velocity equation in (1.3) to obtain

$$\partial_t u = \nu \Delta_h u + \mathbb{P}(\theta e_3) - \mathbb{P}(u \cdot \nabla u). \tag{1.4}$$

By the definition of $\mathbb{P}$,

$$\mathbb{P}(\theta e_3) = \theta e_3 - \nabla\Delta^{-1}\nabla \cdot (\theta e_3) = \begin{bmatrix} -\partial_1\partial_3\Delta^{-1}\theta \\ -\partial_2\partial_3\Delta^{-1}\theta \\ \theta - \partial_3^2\Delta^{-1}\theta \end{bmatrix}. \tag{1.5}$$

Alternatively we can write $\theta - \partial_3^2\Delta^{-1}\theta = \Delta_h\Delta^{-1}\theta$. Inserting (1.5) in (1.4) yields

$$\begin{cases} \partial_t u = \nu\Delta_h u + \begin{bmatrix} -\partial_1\partial_3\Delta^{-1}\theta \\ -\partial_2\partial_3\Delta^{-1}\theta \\ \Delta_h\Delta^{-1}\theta \end{bmatrix} - \mathbb{P}(u \cdot \nabla u), \\ \partial_t\theta = \eta\Delta_h\theta - u_3 - u \cdot \nabla\theta, \end{cases} \tag{1.6}$$

which separates the linear parts from the nonlinear parts. Furthermore, by differentiating (1.6) in time and making suitable substitutions, we discover that (1.6) can be converted into a system of anisotropic and degenerate wave equations

$$\begin{cases} \partial_{tt}u_1 - (\nu+\eta)\Delta_h\partial_t u_1 + \nu\eta\Delta_h^2 u_1 + \partial_1^2\Delta^{-1}u_1 + \partial_1\partial_2\Delta^{-1}u_2 = N_1, \\ \partial_{tt}u_2 - (\nu+\eta)\Delta_h\partial_t u_2 + \nu\eta\Delta_h^2 u_2 + \partial_1\partial_2\Delta^{-1}u_1 + \partial_2^2\Delta^{-1}u_2 = N_2, \\ \partial_{tt}u_3 - (\nu+\eta)\Delta_h\partial_t u_3 + \nu\eta\Delta_h^2 u_3 + \Delta_h\Delta^{-1}u_3 = N_3, \\ \partial_{tt}\theta - (\nu+\eta)\Delta_h\partial_t\theta + \nu\eta\Delta_h^2\theta + \Delta_h\Delta^{-1}\theta = N_4, \end{cases} \tag{1.7}$$

where

$$\begin{aligned} N_1 &= (-\partial_t + \eta\Delta_h)\,(\mathbb{P}(u\cdot\nabla u))_1 + \partial_1\partial_3\Delta^{-1}(u\cdot\nabla\theta), \\ N_2 &= (-\partial_t + \eta\Delta_h)\,(\mathbb{P}(u\cdot\nabla u))_2 + \partial_2\partial_3\Delta^{-1}(u\cdot\nabla\theta), \\ N_3 &= (-\partial_t + \eta\Delta_h)\,(\mathbb{P}(u\cdot\nabla u))_3 - \Delta_h\Delta^{-1}(u\cdot\nabla\theta), \\ N_4 &= (-\partial_t + \nu\Delta_h)(u\cdot\nabla\theta) + (\mathbb{P}(u\cdot\nabla u))_3\,. \end{aligned}$$

Clearly, u_3 and θ satisfy the same linear wave equation with different nonlinear parts. The equations for u_1 and u_2 are slightly different. The precise formula of the spectra can be obtained from (1.6) or (1.7). To avoid nonessential notation complications, we set $\nu = \eta = 1$. Taking the Fourier transform of the linear portion of (1.6), we have

$$\partial_t \begin{bmatrix} \widehat{u} \\ \widehat{\theta} \end{bmatrix} = A \begin{bmatrix} \widehat{u} \\ \widehat{\theta} \end{bmatrix}, \tag{1.8}$$

where A denotes the matrix of multipliers associated with the linear operators,

$$A = \begin{bmatrix} -|\xi_h|^2 & 0 & 0 & -\frac{\xi_1\xi_3}{|\xi|^2} \\ 0 & -|\xi_h|^2 & 0 & -\frac{\xi_2\xi_3}{|\xi|^2} \\ 0 & 0 & -|\xi_h|^2 & \frac{|\xi_h|^2}{|\xi|^2} \\ 0 & 0 & -1 & -|\xi_h|^2 \end{bmatrix}$$

with $\xi_h = (\xi_1, \xi_2)$. The corresponding characteristic polynomial is given by

$$(\lambda + |\xi_h|^2)^2 \left(\lambda^2 + 2|\xi_h|^2\lambda + |\xi_h|^4 + \frac{|\xi_h|^2}{|\xi|^2}\right) = 0,$$

which yields the spectra,

$$\lambda_1 = \lambda_2 = -|\xi_h|^2, \quad \lambda_3 = -|\xi_h|^2 - \frac{|\xi_h|}{|\xi|}i, \quad \lambda_4 = -|\xi_h|^2 + \frac{|\xi_h|}{|\xi|}i.$$

The spectra reveal that the dissipation in the linearized system is essentially horizontal. More precisely, The real part of all eigenvalues λ_j with $j = 1, 2, 3, 4$ is $-|\xi_h|^2$ and thus

$$|e^{\lambda_j t}| = e^{-|\xi_h|^2 t},$$

which is the symbol of the heat operator associated with the horizontal Laplacian. Therefore, as far as the large-time behavior is concerned, this linearized system is essentially controlled by the horizontal Laplacian. Classical tools for large-time behavior such that Schonbek's Fourier splitting method no longer directly apply to the system studied here. This paper develops a new approach to obtain the optimal decay rates for this partially dissipated system. We expect this approach to work for many other partially dissipated PDE sytems.

To gain insight on our problem, we briefly examine the 3D anisotropic heat equation with horizontal dissipation

$$\begin{cases} \partial_t u = \nu \Delta_h u, & x \in \mathbb{R}^3,\ t > 0, \\ u(x, 0) = u_0(x), & x \in \mathbb{R}^3. \end{cases} \tag{1.9}$$

In order to obtain an explicit decay rate of the solution to (1.9), the energy method is no longer sufficient and the explicit representation of the solution is necessary,

$$u(t) = e^{\nu \Delta_h t} u_0.$$

To extract the sharp decay rates for the solution u of (1.9), it is generally necessary to assume either u_0 in a suitable Lebesgue space

$$u_0 \in L^q(\mathbb{R}^3) \quad \text{with} \quad 1 \le q < 2,$$

or in a Sobolev space with negative index. Since the dissipation in (1.9) is only horizontal, the negative derivatives should also be horizontal,

$$\Lambda_h^{-\sigma} u_0 \in L^2,$$

where $\Lambda_h^{-\sigma} u_0$ is defined in terms of the Fourier transform

$$\widehat{\Lambda_h^{-\sigma} u_0}(\xi) = |\xi_h|^{-\sigma} \widehat{u_0}(\xi).$$

We can easily check that the solution u of (1.9) and its first-order derivatives obeys the following optimal decay rates, for any $t > 0$,

$$\|u(t)\|_{L^2} \le C\, (\nu t)^{-\frac{\sigma}{2}}\, \|\Lambda_h^{-\sigma} u_0\|_{L^2}, \tag{1.10}$$

$$\|\partial_3 u(t)\|_{L^2} \le C\, (\nu t)^{-\frac{\sigma}{2}}\, \|\partial_3 \Lambda_h^{-\sigma} u_0\|_{L^2}, \tag{1.11}$$

$$\|\nabla_h u(t)\|_{L^2} \le C\, (\nu t)^{-\frac{\sigma+1}{2}}\, \|\Lambda_h^{-\sigma} u_0\|_{L^2}. \tag{1.12}$$

The estimates above follow from the solution formula $\widehat{u}(\xi, t) = e^{-\nu|\xi_h|^2 t}\widehat{u_0}(\xi)$. The power decay in t is due to arbitrarily small frequencies.

With these helpful hints from the heat equation, our approach starts with solving the linearized system (1.8) and representing the nonlinear system (1.6) in an integral form via the Duhamel principle

$$\begin{bmatrix} \widehat{u}(t) \\ \widehat{\theta}(t) \end{bmatrix} = e^{At} \begin{bmatrix} \widehat{u_0} \\ \widehat{\theta_0} \end{bmatrix} + \int_0^t e^{A(t-\tau)} \begin{bmatrix} \widehat{M}_1(\tau) \\ \widehat{M}_2(\tau) \end{bmatrix} d\tau,$$

where M_1 and M_2 are the two nonlinear terms in (1.6),

$$M_1 = -\mathbb{P}(u \cdot \nabla u), \quad M_2 = -u \cdot \nabla \theta.$$

To avoid possible notational confusion, we remark that $M_1 \in \mathbb{R}^3$ and $M_2 \in \mathbb{R}$ are not components. In order to obtain an explicit formula for the fundamental matrix e^{At}, we diagonalize A via its eigenvalues and eigenvectors, and break e^{At} down to explicit kernel functions. The detailed derivation and the precise integral representation of (1.6) are given in Sect. 2. Alternatively we could have also achieved the same formula by solving the wave equations in (1.7).

As a preparation for the optimal decay rates on the nonlinear system, we first examine the linearized system of (1.6), namely

$$\begin{cases} \partial_t u = \nu \Delta_h u + \begin{bmatrix} -\partial_1 \partial_3 \Delta^{-1} \theta \\ -\partial_2 \partial_3 \Delta^{-1} \theta \\ \Delta_h \Delta^{-1} \theta \end{bmatrix}, \\ \partial_t \theta = \eta \Delta_h \theta - u_3. \end{cases}$$

The analysis is performed on its corresponding explicit solution representation,

$$\widehat{u}_h = e^{\lambda_1 t} \widehat{u}_{0h} + \left(\frac{\xi_h \xi_3}{|\xi_h|^2} e^{\lambda_1 t} + \frac{\xi_h \xi_3}{|\xi_h|^2} G_2 + \xi_h \xi_3 G_1 \right) \widehat{u}_{03} - \frac{\xi_h \xi_3}{|\xi|^2} G_1 \widehat{\theta}_0 \tag{1.13}$$

$$\widehat{u}_3 = \left(-G_2 - |\xi_h|^2 G_1 \right) \widehat{u}_{03} + \frac{|\xi_h|^2}{|\xi|^2} G_1 \widehat{\theta}_0 \tag{1.14}$$

$$\widehat{\theta} = -G_1 \widehat{u}_{03} + (G_3 + |\xi_h|^2 G_1) \widehat{\theta}_0, \tag{1.15}$$

where G_1, G_2 and G_3 are given by (see (2.4) below)

$$\begin{aligned} G_1 &= \frac{e^{\lambda_4 t} - e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = e^{-|\xi_h|^2 t} \left(\frac{|\xi_h|}{|\xi|} \right)^{-1} \sin \frac{|\xi_h|}{|\xi|} t, \\ G_2 &= \frac{\lambda_3 e^{\lambda_4 t} - \lambda_4 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 - e^{\lambda_3 t}, \\ G_3 &= \frac{\lambda_4 e^{\lambda_4 t} - \lambda_3 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 + e^{\lambda_4 t}. \end{aligned}$$

We are able to obtain the stability and optimal decay rates stated in the following proposition. The results in this proposition and their proofs are part of our program for optimal decay rates on the nonlinear system (1.6), and will be used in the proof of our main result, Theorem 1.3 below.

Proposition 1.1 *Let s be non-negative and $\sigma > 0$. Assume the initial velocity field $u_0 = (u_{01}, u_{02}, u_{03})$ satisfies $\nabla \cdot u_0 = 0$.*

(1) *If (u_0, θ_0) satisfies*

$$u_0,\ \theta_0,\ \Lambda_h^{-\sigma} u_0,\ \Lambda_h^{-\sigma} \theta_0 \in \dot{H}^s,$$

then the solution (u,θ) *given by* (1.13), (1.14) *and* (1.15) *satisfies*

$$\|u(t)\|_{\dot H^s},\ \|\theta(t)\|_{\dot H^s} \le C\,\left(\|(u_0,\theta_0)\|_{\dot H^s} + \|(\Lambda_h^{-\sigma}u_0,\Lambda_h^{-\sigma}\theta_0)\|_{\dot H^s}\right)(1+t)^{-\frac{\sigma}{2}}. \quad (1.16)$$

(2) *If* (u_0,θ_0) *satisfies*

$$\partial_3 u_0,\ \partial_3\theta_0,\ \Lambda_h^{-\sigma}\partial_3 u_0,\ \Lambda_h^{-\sigma}\partial_3\theta_0 \in \dot H^s,$$

then

$$\begin{aligned}&\|\partial_3 u(t)\|_{\dot H^s},\ \|\partial_3\theta(t)\|_{\dot H^s}\\ &\quad\le C\,\left(\|(\partial_3 u_0,\partial_3\theta_0)\|_{\dot H^s} + \|(\Lambda_h^{-\sigma}\partial_3 u_0,\Lambda_h^{-\sigma}\partial_3\theta_0)\|_{\dot H^s}\right)(1+t)^{-\frac{\sigma}{2}}.\end{aligned} \quad (1.17)$$

(3) *If* (u_0,θ_0) *satisfies*

$$\Lambda_h^{-\sigma}u_0,\ \Lambda_h^{-\sigma}\theta_0 \in \dot H^s,$$

then, for any $t>0$,

$$\|\nabla_h u(t)\|_{\dot H^s},\ \|\nabla_h\theta(t)\|_{\dot H^s} \le C\,\|(\Lambda_h^{-\sigma}u_0,\Lambda_h^{-\sigma}\theta_0)\|_{\dot H^s}\, t^{-\frac{\sigma+1}{2}}. \quad (1.18)$$

If, in addition, (u_0,θ_0) *satisfies*

$$\Lambda_h^{-\sigma}\nabla_h u_0,\ \Lambda_h^{-\sigma}\nabla_h\theta_0 \in \dot H^s,$$

then

$$\begin{aligned}&\|\nabla_h u(t)\|_{\dot H^s},\ \|\nabla_h\theta(t)\|_{\dot H^s}\\ &\quad\le C\,\left(\|(\Lambda_h^{-\sigma}\nabla_h u_0,\Lambda_h^{-\sigma}\nabla_h\theta_0)\|_{\dot H^s} + \|(\Lambda_h^{-\sigma}u_0,\Lambda_h^{-\sigma}\theta_0)\|_{\dot H^s}\right)(1+t)^{-\frac{\sigma+1}{2}}.\end{aligned} \quad (1.19)$$

The bounds for the linearized problem are explicit and thus easily seen to be optimal. It is also clear that the vertical derivatives have the same decay rate as that for the solution itself, but the horizontal derivatives increase the decay rate by $-1/2$. We also remark that (1.18) is suitable for large $t>0$. For $t>0$ close to 0, (1.18) is an over-estimate and should be replaced by (1.19).

The second preparation is a small data global well-posedness and stability result on the nonlinear system (1.6).

Proposition 1.2 *Consider the nonlinear system in* (1.3) *with* $\nu>0$ *and* $\eta>0$. *Assume* $(u_0,\theta_0)\in H^m(\mathbb{R}^3)$ *with* $m\ge 2$ *satisfies* $\nabla\cdot u_0=0$. *Then there exists* $\varepsilon=\varepsilon(\nu,\eta)>0$ *such that, if*

$$\|u_0\|_{H^m} + \|\theta_0\|_{H^m} \le \varepsilon,$$

then (1.3) *has a unique global solution* $(u,\theta)\in L^\infty(0,\infty;H^m)$ *satisfying, for a constant* $C>0$ *and for all* $t\ge 0$,

$$\|u(t)\|_{H^m}^2 + \|\theta(t)\|_{H^m}^2 + \nu\int_0^t \|\nabla_h u\|_{H^m}^2\,d\tau + \eta\int_0^t \|\nabla_h\theta\|_{H^m}^2\,d\tau \le C\,\varepsilon^2.$$

We are now ready to state our main results presenting the stability and optimal decay rates for perturbations near the hydrostatic balance.

Theorem 1.3 *Consider the nonlinear system in* (1.3) *with* $\nu > 0$ *and* $\eta > 0$. *Let* $\frac{3}{4} \leq \sigma < 1$. *Assume* $(u_0, \theta_0) \in H^4(\mathbb{R}^3)$ *satisfies* $\nabla \cdot u_0 = 0$,

$$\|u_0\|_{H^4} + \|\theta_0\|_{H^4} \leq \varepsilon, \tag{1.20}$$

$$\|\Lambda_h^{-\sigma} u_0\|_{L^2} + \|\Lambda_h^{-\sigma} \theta_0\|_{L^2} \leq \varepsilon, \tag{1.21}$$

$$\|\partial_3 \Lambda_h^{-\sigma} u_0\|_{L^2} + \|\partial_3 \Lambda_h^{-\sigma} \theta_0\|_{L^2} \leq \varepsilon \tag{1.22}$$

for some sufficiently small $\varepsilon > 0$. *Then* (1.3) *has a unique global solution* (u, θ) *satisfying, for a constant* $C > 0$ *and for all* $t \geq 0$,

$$\begin{aligned}
&\|u(t)\|_{H^4} + \|\theta(t)\|_{H^4} \leq C\,\varepsilon, \\
&\|\Lambda_h^{-\sigma} u(t)\|_{L^2} + \|\Lambda_h^{-\sigma} \theta(t)\|_{L^2} \leq C\,\varepsilon, \\
&\|u(t)\|_{L^2}, \|\partial_3 u(t)\|_{L^2}, \|\theta(t)\|_{L^2}, \|\partial_3 \theta(t)\|_{L^2} \leq C\,\varepsilon(1+t)^{-\frac{\sigma}{2}}, \\
&\|\nabla_h u(t)\|_{L^2},\ \|\nabla_h \theta(t)\|_{L^2} \leq C\,\varepsilon(1+t)^{-\frac{\sigma}{2}-\frac{1}{2}}.
\end{aligned}$$

The regularity requirement $(u_0, \theta_0) \in H^4$ and the condition $\sigma \geq \frac{3}{4}$ are needed in order to handle the most challenging term $\partial_{33} u$ when we estimate $\|\partial_3 u\|_{L^2}$. More technical details are provided on pages 28-29 in the proof of Theorem 1.3. The decay rates for the solution are the same as those in (1.10), (1.11) and (1.12), and are thus optimal. The sharp decay result presented here appears to be the first such result on the 3D anisotropic Boussinesq equations. It is hoped that this result together with its proof helps chart a new path to the stability and large-time behavior problems involving anisotropic fluids.

We thank the referee and the editor for bringing to our attention the work of Shang and Xu [28]. Shang and Xu [28] examined the stability of two Boussinesq systems with dissipation and thermal diffusion in two directions as well as the decay of the corresponding linearized systems. Their decay results, stated in their Theorems 1.3 and 1.6, are for the linearized systems (1.10) and (1.17) in [28], which involve no nonlinear terms. Their paper doesn't provide any decay result for the nonlinear systems.

The main goal of our paper is to obtain the optimal decay rates of the nonlinear Boussinesq system involving only horizontal dissipation. It is generally much more difficult to obtain the large-time behavior of nonlinear PDE systems. The anisotropic dissipation here makes the optimal decay problem even more challenging. Since only dissipation in the horizontal directions is available, the nonlinear effects require much more delicate analysis. In particular, we need to exploit cancellations and other properties such as the incompressibility in order to control terms involving vertical derivatives.

Since the energy method and other classical tools such as the Fourier-splitting scheme [25] no longer work for the nonlinear Boussinesq system considered here, this paper proposes and implements an innovative approach. We derive and make use of the integral representation of the nonlinear Boussinesq system. Our approach consists of three main steps. The first is to solve the linearized system explicitly and use this explicit solution formula to derive decay rates for the solution itself as well as its derivatives. The main result of this step is presented in Proposition 1.1. In comparison with the decay results for the linearized systems in [28], Proposition 1.1 contains much more information. Besides the decay rate for the H^s-norm, Proposition 1.1 also features decay rates for the horizontal and the vertical derivatives, which are optimal and reveal a faster decay rate for the horizontal derivative. The second main step is to establish the small data global well-posedness and stability for the nonlinear system. The result of this step is stated in Proposition 1.2. In particular, Proposition 1.2 guarantees that the solution of the nonlinear system is global and it is legitimate to study its precise large-time behavior. The third step, the main thrust of our work, is to establish the optimal decay rates for the nonlinear Boussinesq system, as stated in Theorem 1.3. By the explicit solution formula of the linearized system and Duhamel's principle, we convert the nonlinear Boussinesq system into an integral representation. Then the bootstrapping argument is applied to this integral form. Due to the lack of dissipation in the vertical direction, the analysis on the nonlinear effects is very difficult and involved. In particular, we need to exploit cancellations and other properties such as the incompressibility in order to control terms involving vertical derivatives.

The framework of the proof is to apply the bootstrapping argument to the integral representation of the nonlinear system given by (2.1), (2.2) and (2.3). A very useful abstract version of the bootstrap principle can be found in [31,p. 21]. We assume the initial datum (u_0, θ_0) satisfies the assumptions (1.20), (1.21) and (1.22), and make the ansatz that the solution (u, θ) satisfies, for a suitably selected constant $C_0 > 0$,

$$
\begin{aligned}
&\|u(t)\|_{H^4},\ \ \|\theta(t)\|_{H^4} \le C_0\,\varepsilon,\\
&\|\Lambda_h^{-\sigma} u(t)\|_{L^2},\ \ \|\Lambda_h^{-\sigma} \theta(t)\|_{L^2} \le C_0\,\varepsilon,\\
&\|u(t)\|_{L^2}, \|\partial_3 u(t)\|_{L^2}, \|\theta(t)\|_{L^2}, \|\partial_3\theta(t)\|_{L^2} \le C_0\,\varepsilon(1+t)^{-\frac{\sigma}{2}},\\
&\|\nabla_h u(t)\|_{L^2},\ \|\nabla_h\theta(t)\|_{L^2} \le C_0\,\varepsilon(1+t)^{-\frac{\sigma}{2}-\frac{1}{2}},
\end{aligned}
$$

for $t \in [0, T]$ with $T > 0$. The initial time $T > 0$ exists by local well-posedness. By imposing the smallness conditions on (u_0, θ_0) as in (1.20), (1.21) and (1.22), we then show via (2.1), (2.2) and (2.3) that (u, θ) actually satisfies the following improved inequalities,

$$
\|u(t)\|_{H^4},\ \ \|\theta(t)\|_{H^4} \le \frac{C_0}{2}\,\varepsilon, \tag{1.23}
$$

$$
\|\Lambda_h^{-\sigma} u(t)\|_{L^2},\ \ \|\Lambda_h^{-\sigma} \theta(t)\|_{L^2} \le \frac{C_0}{2}\,\varepsilon, \tag{1.24}
$$

$$
\|u(t)\|_{L^2}, \|\partial_3 u(t)\|_{L^2}, \|\theta(t)\|_{L^2}, \|\partial_3\theta(t)\|_{L^2} \le \frac{C_0}{2}\,\varepsilon(1+t)^{-\frac{\sigma}{2}}, \tag{1.25}
$$

$$
\|\nabla_h u(t)\|_{L^2},\ \|\nabla_h\theta(t)\|_{L^2} \le \frac{C_0}{2}\,\varepsilon(1+t)^{-\frac{\sigma}{2}-\frac{1}{2}}. \tag{1.26}
$$

The bootstrapping argument then implies that the maximal time T with this property is given by $T = \infty$. Thus, the four inequalities above indeed hold for all $t < \infty$. In particular, they yield the global in time bounds and decay rates.

Our main efforts are devoted to proving (1.23), (1.24), (1.25) and (1.26). The initial time $T > 0$ exists by local well-posedness. Proving these improved inequalities is very hard due to the lack of full dissipation. As aforementioned, some of the nonlinear terms such as $\partial_{33}u$ in the expression of $\partial_3 u$ require extremely careful analysis. Various cancellations and other properties are exploited. This is a long and nontrivial process. Various anisotropic inequalities are invoked to fully make use of the anisotropic dissipation in the system. In order to obtain suitable upper bounds for some of the terms, we have to exploit the structure of the kernel function together with the corresponding term it acts on. To explain this point, we take two terms from the representation of u_h in (2.1) as an example,

$$
\left(\frac{\xi_h\xi_3}{|\xi_h|^2} e^{\lambda_1 t} + \frac{\xi_h\xi_3}{|\xi_h|^2} G_2 + \xi_h\xi_3 G_1\right)\widehat{u}_{03} \tag{1.27}
$$

and

$$
\int_0^t \left(\frac{\xi_h\xi_3}{|\xi_h|^2} e^{\lambda_1 (t-\tau)} + \frac{\xi_h\xi_3}{|\xi_h|^2} G_2(t-\tau) + \xi_h\xi_3 G_1(t-\tau)\right)(\mathbb{P}\widehat{(u\cdot\nabla u)})_3(\tau)\,d\tau. \tag{1.28}
$$

The kernel function in (1.27) has a factor $\frac{1}{|\xi_h|}$, which has to be canceled in order to obtain a bound of the form $e^{-|\xi_h|^2 t}$. The idea here is to combine ξ_3 with $\widehat{u_{03}}$ and use the divergence-free condition $\xi_3\widehat{u_{03}} = -\xi_h \cdot \widehat{u_{0h}}$ to generate a factor ξ_h. To deal with the nonlinear term (1.28), we have also managed to generate factor ξ_h in $(\mathbb{P}\widehat{(u \cdot \nabla u)})_3$. By applying the definition of the projection operator and the divergence-free condition $\nabla \cdot u = 0$, and invoking some cancellations via combination, we find the identity

$$
\begin{aligned}
(\mathbb{P}\widehat{(u \cdot \nabla u)})_3 &= \nabla_h \widehat{\cdot (u_h u_3)} - \nabla_h \cdot \partial_3 \widehat{\Delta^{-1}\nabla \cdot} (u \otimes u_h) \\
&\quad - \nabla_h \cdot \widehat{\Delta^{-1}\partial_{33}(u_h u_3)} + \Delta^{-1} \widehat{\Delta_h \partial_3 (u_3 u_3)}
\end{aligned}
$$

and the Fourier transform of the right-hand side involves ξ_h, which allows us to cancel the factor $\frac{1}{|\xi_h|}$ in the kernel. More technical details can be found in the proofs of the two propositions and Theorem 1.3.

Finally, we mention some of the closely related work. Due to their practical applications and mathematical significance, the stability and large-time properties of perturbations near the hydrostatic balance have recently attracted considerable mathematical interests. The work of Doering et al. [12] investigated the stability of the hydrostatic equilibrium to the 2D Boussinesq system with only kinematic dissipation (without thermal diffusion) and rigorously proved the global asymptotic stability of any perturbation near the hydrostatic equilibrium [12]. In addition, extensive numerical simulations are performed in [12] to corroborate the analytical results and predict some phenomena that are not proven. The work of Tao et al. [30] resolves several important issues left open in [12]. In particular, [30] provides a precise description of the final buoyancy distribution in case of general initial conditions and the explicit decay rate of the velocity field or the total mechanical energy. The paper of Castro, Córdoba and Lear successfully established the stability and large time behavior on the 2D Boussinesq equations with velocity damping instead of dissipation [9]. The stabilizing effect of the temperature on the buoyancy-driven fluids and the stability of the hydrostatic equilibrium were discovered for several partially dissipated 2D Boussinesq systems [2, 15, 16]. There are very significant recent developments on the stability of shear flow to the fluid equations with various partial dissipation [3–5, 10, 13, 14, 18, 22, 29, 32, 37, 39–41].

The rest of this paper is divided into four sections. Section 2 details how we convert the nonlinear Boussinesq system (1.6) into an integral form stated in Proposition 2.1. Section 3 presents the linear stability theory and the optimal decay rates for the linearized system. In particular, we prove Proposition 1.1. Section 4 proves the nonlinear stability result stated in Proposition 1.2. The optimal decay rates, our main result stated in Theorem 1.3, are established in Sect. 5. For the sake of clarity, Sect. 5 is further divided into four subsections.

2 Spectra and integral representation

This section separates the linear and the nonlinear parts in (1.3), solves the linearized system and represents the nonlinear system in an integral form via Duhamel's principle. More precisely, we prove the following proposition.

Proposition 2.1 *The system in (1.3) can be converted into the following integral form*

$$
\begin{aligned}
\widehat{u}_h = {} & e^{\lambda_1 t}\widehat{u}_{0h} + \left(\frac{\xi_h\xi_3}{|\xi_h|^2}e^{\lambda_1 t} + \frac{\xi_h\xi_3}{|\xi_h|^2}G_2 + \xi_h\xi_3 G_1\right)\widehat{u}_{03} - \frac{\xi_h\xi_3}{|\xi|^2}G_1\widehat{\theta}_0 \\
& - \int_0^t e^{\lambda_1(t-\tau)}(\mathbb{P}\widehat{(u\cdot\nabla u)})_h(\tau)\,d\tau \\
& - \int_0^t \left(\frac{\xi_h\xi_3}{|\xi_h|^2}e^{\lambda_1(t-\tau)} + \frac{\xi_h\xi_3}{|\xi_h|^2}G_2 + \xi_h\xi_3 G_1\right)(\mathbb{P}\widehat{(u\cdot\nabla u)})_3(\tau)\,d\tau \\
& + \int_0^t \frac{\xi_h\xi_3}{|\xi|^2}G_1(t-\tau)\widehat{(u\cdot\nabla\theta)}(\tau)\,d\tau
\end{aligned} \tag{2.1}
$$

$$
\begin{aligned}
\widehat{u}_3 = {} & \left(-G_2 - |\xi_h|^2 G_1\right)\widehat{u}_{03} + \frac{|\xi_h|^2}{|\xi|^2}G_1\widehat{\theta}_0 \\
& + \int_0^t (G_2 + |\xi_h|^2 G_1)(t-\tau)\mathbb{P}\widehat{(u\cdot\nabla u)}_3(\tau)\,d\tau \\
& - \int_0^t \frac{|\xi_h|^2}{|\xi|^2}G_1(t-\tau)\,\widehat{(u\cdot\nabla\theta)}(\tau)\,d\tau
\end{aligned} \tag{2.2}
$$

$$
\begin{aligned}
\widehat{\theta} = {} & -G_1\widehat{u}_{03} + (G_3 + |\xi_h|^2 G_1)\widehat{\theta}_0 + \int_0^t G_1(t-\tau)\mathbb{P}\widehat{(u\cdot\nabla u)}_3(\tau)\,d\tau \\
& - \int_0^t (G_3(t-\tau) + |\xi_h|^2 G_1(t-\tau))\widehat{(u\cdot\nabla\theta)}(\tau)\,d\tau
\end{aligned} \tag{2.3}
$$

where

$$
\begin{aligned}
G_1 &= \frac{e^{\lambda_4 t} - e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = e^{-|\xi_h|^2 t}\left(\frac{|\xi_h|}{|\xi|}\right)^{-1}\sin\frac{|\xi_h|}{|\xi|}t, \\
G_2 &= \frac{\lambda_3 e^{\lambda_4 t} - \lambda_4 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 - e^{\lambda_3 t}, \\
G_3 &= \frac{\lambda_4 e^{\lambda_4 t} - \lambda_3 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 + e^{\lambda_4 t}.
\end{aligned} \tag{2.4}
$$

with $\lambda_1, \lambda_2, \lambda_3$ and λ_4 given by

$$
\lambda_1 = \lambda_2 = -|\xi_h|^2, \quad \lambda_3 = -|\xi_h|^2 - \frac{|\xi_h|}{|\xi|}i, \quad \lambda_4 = -|\xi_h|^2 + \frac{|\xi_h|}{|\xi|}i. \tag{2.5}
$$

Proof of Proposition 2.1 We have separated the linear parts from the nonlinear ones in (1.3) and obtained (1.6). Taking the Fourier transform of (1.6), we find

$$
\partial_t \begin{bmatrix} \widehat{u}(t) \\ \widehat{\theta}(t) \end{bmatrix} = A \begin{bmatrix} \widehat{u}(t) \\ \widehat{\theta}(t) \end{bmatrix} + \begin{bmatrix} \widehat{M}_1 \\ \widehat{M}_2 \end{bmatrix},
$$

where A represents the multiplier matrix of the linear operators, and M_1 and M_2 are the nonlinear terms,

$$
A = \begin{bmatrix} -|\xi_h|^2 & 0 & 0 & -\frac{\xi_1\xi_3}{|\xi|^2} \\ 0 & -|\xi_h|^2 & 0 & -\frac{\xi_2\xi_3}{|\xi|^2} \\ 0 & 0 & -|\xi_h|^2 & \frac{|\xi_h|^2}{|\xi|^2} \\ 0 & 0 & -1 & -|\xi_h|^2 \end{bmatrix}
$$

$$
M_1 = -\mathbb{P}(u\cdot\nabla u), \quad M_2 = -u\cdot\nabla\theta.
$$

By the Duhamel principle,

$$\begin{bmatrix} \widehat{u}(t) \\ \widehat{\theta}(t) \end{bmatrix} = e^{At} \begin{bmatrix} \widehat{u_0} \\ \widehat{\theta_0} \end{bmatrix} + \int_0^t e^{A(t-\tau)} \begin{bmatrix} \widehat{M}_1(\tau) \\ \widehat{M}_2(\tau) \end{bmatrix} d\tau. \tag{2.6}$$

We compute the fundamental matrix e^{At} explicitly. The characteristic polynomial associated with A is given by

$$(\lambda + |\xi_h|^2)^2 \left(\lambda^2 + 2|\xi_h|^2\lambda + |\xi_h|^4 + \frac{|\xi_h|^2}{|\xi|^2} \right) = 0$$

and thus the spectra of A are

$$\lambda_1 = \lambda_2 = -|\xi_h|^2, \quad \lambda_3 = -|\xi_h|^2 - \frac{|\xi_h|}{|\xi|}i, \quad \lambda_4 = -|\xi_h|^2 + \frac{|\xi_h|}{|\xi|}i.$$

The two eigenvectors corresponding to $\lambda_1 = \lambda_2$ are

$$\overrightarrow{V}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{V}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

The eigenvectors $\overrightarrow{V}_m$ associated with the eigenvalues $\lambda_m (m = 3, 4)$ satisfy

$$\begin{aligned} &(\lambda_m I - A)\overrightarrow{V}_m \\ &= \begin{bmatrix} \lambda_m + |\xi_h|^2 & 0 & 0 & \frac{\xi_1\xi_3}{|\xi|^2} \\ 0 & \lambda_m + |\xi_h|^2 & 0 & \frac{\xi_2\xi_3}{|\xi|^2} \\ 0 & 0 & \lambda_m + |\xi_h|^2 & -\frac{|\xi_h|^2}{|\xi|^2} \\ 0 & 0 & 1 & \lambda_m + |\xi_h|^2 \end{bmatrix} \begin{bmatrix} V_{m1} \\ V_{m2} \\ V_{m3} \\ V_{m4} \end{bmatrix} = 0 \end{aligned}$$

and thus

$$\overrightarrow{V}_3 = \begin{bmatrix} -\xi_1\xi_3 \\ -\xi_2\xi_3 \\ -(\lambda_3 + |\xi_h|^2)^2\,|\xi|^2 \\ (\lambda_3 + |\xi_h|^2)|\xi|^2 \end{bmatrix}, \quad \overrightarrow{V}_4 = \begin{bmatrix} -\xi_1\xi_3 \\ -\xi_2\xi_3 \\ -(\lambda_4 + |\xi_h|^2)^2\,|\xi|^2 \\ (\lambda_4 + |\xi_h|^2)|\xi|^2 \end{bmatrix}.$$

Thus the eigen-matrix is given by

$$V = \begin{bmatrix} 1 & 0 & -\xi_1\xi_3 & -\xi_1\xi_3 \\ 0 & 1 & -\xi_2\xi_3 & -\xi_2\xi_3 \\ 0 & 0 & -(\lambda_3 + |\xi_h|^2)^2|\xi|^2 & -(\lambda_4 + |\xi_h|^2)^2|\xi|^2 \\ 0 & 0 & (\lambda_3 + |\xi_h|^2)|\xi|^2 & (\lambda_4 + |\xi_h|^2)|\xi|^2 \end{bmatrix}$$

and

$$V^{-1} = \begin{bmatrix} 1 & 0 & \frac{\xi_1\xi_3}{|\xi_h|^2} & 0 \\ 0 & 1 & \frac{\xi_2\xi_3}{|\xi_h|^2} & 0 \\ 0 & 0 & \frac{\lambda_4+|\xi_h|^2}{|\xi_h|^2(\lambda_4-\lambda_3)} & \frac{(\lambda_4+|\xi_h|^2)^2}{|\xi_h|^2(\lambda_4-\lambda_3)} \\ 0 & 0 & -\frac{\lambda_3+|\xi_h|^2}{|\xi_h|^2(\lambda_4-\lambda_3)} & -\frac{(\lambda_3+|\xi_h|^2)^2}{|\xi_h|^2(\lambda_4-\lambda_3)} \end{bmatrix}.$$

As a consequence, the fundamental matrix is given by

$$e^{At} = V \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & 0 \\ 0 & e^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & e^{\lambda_3 t} & 0 \\ 0 & 0 & 0 & e^{\lambda_4 t} \end{bmatrix} V^{-1}$$

$$= \begin{bmatrix} e^{\lambda_1 t} & 0 & \frac{\xi_1\xi_3}{|\xi_h|^2}e^{\lambda_1 t} + \frac{\xi_1\xi_3}{|\xi_h|^2}G_2 + \xi_1\xi_3 G_1 & -\frac{\xi_1\xi_3}{|\xi|^2}G_1 \\ 0 & e^{\lambda_2 t} & \frac{\xi_2\xi_3}{|\xi_h|^2}e^{\lambda_2 t} + \frac{\xi_2\xi_3}{|\xi_h|^2}G_2 + \xi_2\xi_3 G_1 & -\frac{\xi_2\xi_3}{|\xi|^2}G_1 \\ 0 & 0 & -G_2 - |\xi_h|^2 G_1 & \frac{|\xi_h|^2}{|\xi|^2}G_1 \\ 0 & 0 & -G_1 & G_3 + |\xi_h|^2 G_1 \end{bmatrix},$$

where we have written

$$G_1 = \frac{e^{\lambda_4 t} - e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = e^{-|\xi_h|^2 t}\frac{\sin\frac{|\xi_h|}{|\xi|}t}{\frac{|\xi_h|}{|\xi|}},$$

$$G_2 = \frac{\lambda_3 e^{\lambda_4 t} - \lambda_4 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 - e^{\lambda_3 t},$$

$$G_3 = \frac{\lambda_4 e^{\lambda_4 t} - \lambda_3 e^{\lambda_3 t}}{\lambda_4 - \lambda_3} = \lambda_3 G_1 + e^{\lambda_4 t}.$$

Inserting e^{At} in (2.6) yields the desired representations (2.1), (2.2) and (2.3). This completes the proof of Proposition 2.1. □

3 Linear stability and optimal decay

This section focuses on the stability of the linearized system of (1.3) and the optimal decay rates. This result serves as the first step for the nonlinear stability and optimal decay rates presented in the next two sections. Recall that, by (1.6), the linearized portion of (1.3) can be written as

$$\begin{cases} \partial_t u = \nu\Delta_h u + \begin{bmatrix} -\partial_1\partial_3\Delta^{-1}\theta \\ -\partial_2\partial_3\Delta^{-1}\theta \\ \Delta_h\Delta^{-1}\theta \end{bmatrix}, \\ \partial_t\theta = \eta\Delta_h\theta - u_3. \end{cases} \tag{3.1}$$

According to Proposition 2.1, (3.1) can be solved explicitly as (1.3) can be represented as

$$\widehat{u_h} = e^{\lambda_1 t}\widehat{u}_{0h} + \left(\frac{\xi_h\xi_3}{|\xi_h|^2}e^{\lambda_1 t} + \frac{\xi_h\xi_3}{|\xi_h|^2}G_2 + \xi_h\xi_3 G_1\right)\widehat{u}_{03} - \frac{\xi_h\xi_3}{|\xi|^2}G_1\widehat{\theta}_0 \tag{3.2}$$

$$\widehat{u}_3 = \left(-G_2 - |\xi_h|^2 G_1\right)\widehat{u}_{03} + \frac{|\xi_h|^2}{|\xi|^2}G_1\widehat{\theta}_0 \tag{3.3}$$

$$\widehat{\theta} = -G_1\widehat{u}_{03} + (G_3 + |\xi_h|^2 G_1)\widehat{\theta}_0, \tag{3.4}$$

where G_1, G_2, G_3 and λ_1 through λ_4 are given in (2.4) and (2.5), respectively. Our goal here is to prove Proposition 1.1 given in the introduction.

In order to prove Proposition 1.1, we first state a lemma that provides upper bounds for G_1, G_2 and G_3 defined in (2.4).

Lemma 3.1 *There are two constants $C > 0$ and $c_0 > 0$ such that, for any $\xi \in \mathbb{R}^3$ and $t \geq 0$,*

$$|G_1| \leq t\, e^{-|\xi_h|^2 t}, \quad |\xi_h|^2 |G_1| \leq C\, e^{-c_0|\xi_h|^2 t},$$
$$|G_2| \leq C\, e^{-c_0|\xi_h|^2 t}, \quad |G_3| \leq C\, e^{-c_0|\xi_h|^2 t}.$$

Proof of Lemma 3.1 The upper bounds for G_1 follow directly from the definition of G_1 and the simple fact that $|\sin y| \leq |y|$ for any real number y. To bound G_2, we notice that

$$G_2 = \lambda_3 G_1 - e^{\lambda_3 t} = -|\xi_h|^2 G_1 - i\,\frac{|\xi_h|}{|\xi|} G_1 - e^{\lambda_3 t}$$

and thus

$$|G_2| \leq |\xi_h|^2 t\, e^{-|\xi_h|^2 t} + 2e^{-|\xi_h|^2 t} \leq C\, e^{-c_0\,|\xi_h|^2 t},$$

where we have used the fact that $y^a\, e^{-y} \leq e^{-c_0 y}$ for any $a \geq 0$ and $y \geq 0$. G_3 can be bounded similarly. □

Proof of Proposition 1.1 Due to the frequency decoupling in the solution representation in (3.2), (3.3) and (3.4), it suffices to set $s = 0$ and consider the L^2-norm. We start with the estimate of u_h. The first term in (3.2) is easily bounded. For any $0 \leq t < 1$,

$$\|e^{\lambda_1 t}\widehat{u}_{0h}\|_{L^2} \leq \|\widehat{u}_{0h}\|_{L^2} = \|u_{0h}\|_{L^2}. \tag{3.5}$$

For $t \geq 1$,

$$\|e^{\lambda_1 t}\widehat{u}_{0h}\|_{L^2} = \||\xi_h|^\sigma\, e^{-|\xi_h|^2 t}|\xi_h|^{-\sigma}\widehat{u}_{0h}\|_{L^2} \leq C\, t^{-\frac{\sigma}{2}}\,\|\Lambda_h^{-\sigma} u_{0h}\|_{L^2}. \tag{3.6}$$

Here we have used the following inequality

$$\sup_{\xi_h} |\xi_h|^\sigma\, e^{-|\xi_h|^2 t} = t^{-\frac{\sigma}{2}}\, \sup_{\xi_h}(|\xi_h|^2 t)^{\frac{\sigma}{2}}\, e^{-|\xi_h|^2 t} = C\, t^{-\frac{\sigma}{2}},$$

where $C = \sup_{b \geq 0} b^{\frac{\sigma}{2}}\, e^{-b} < \infty$. Combining (3.5) and (3.6) yields

$$\|e^{\lambda_1 t}\widehat{u}_{0h}\|_{L^2} \leq C\, \left(\|u_{0h}\|_{L^2} + \|\Lambda_h^{-\sigma}\, u_{0h}\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

Since the bound for $0 \leq t < 1$ is quite simple, we shall only present the estimates for $t \geq 1$ in the rest of the proof. We consider the second term in (3.2). For notational convenience, we write

$$I = \left(\frac{\xi_h \xi_3}{|\xi_h|^2} e^{\lambda_1 t} + \frac{\xi_h \xi_3}{|\xi_h|^2} G_2 + \xi_h \xi_3 G_1\right)\widehat{u}_{03}.$$

By $\nabla \cdot u_0 = 0$ or $\xi_1 \widehat{u}_{01} + \xi_2 \widehat{u}_{02} + \xi_3 \widehat{u}_{03} = 0$ and Lemma 3.1,

$$\begin{aligned}
\|I\|_{L^2} &= \left\|\xi_h \left(\frac{e^{\lambda_1 t}}{|\xi_h|^2} + \frac{G_2}{|\xi_h|^2} + G_1\right)\xi_3 \widehat{u}_{03}\right\|_{L^2} \\
&= \left\|\xi_h \left(\frac{e^{\lambda_1 t}}{|\xi_h|^2} + \frac{G_2}{|\xi_h|^2} + G_1\right)(-\xi_h \cdot \widehat{u}_{0h})\right\|_{L^2} \\
&= \|(e^{\lambda_1 t} + G_2 + |\xi_h|^2 G_1)\widehat{u}_{0h}\|_{L^2} \\
&\leq \|e^{-|\xi_h|^2 t}\, \widehat{u}_{0h}\|_{L^2}.
\end{aligned}$$

Therefore,

$$\|I\|_{L^2} \leq C\, \left(\|u_{0h}\|_{L^2} + \|\Lambda_h^{-\sigma}\, u_{0h}\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

We now turn to the last term in (3.2). Since

$$\left|\frac{\xi_h\xi_3}{|\xi|^2}G_1\right| = \left|\frac{\xi_h\xi_3}{|\xi|^2}\, e^{-|\xi_h|^2 t}\,\frac{\sin\frac{|\xi_h|}{|\xi|}t}{\frac{|\xi_h|}{|\xi|}}\right| \le e^{-|\xi_h|^2 t}.$$

Therefore,

$$\left\|\frac{\xi_h\xi_3}{|\xi|^2}G_1\widehat{\theta_0}\right\|_{L^2} \le \|e^{-|\xi_h|^2 t}\widehat{\theta_0}\|_{L^2} \le C\,\left(\|\theta_0\|_{L^2} + \|\Lambda_h^{-\sigma}\theta_0\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

Combining the estimates for the three terms above yields

$$\|u(t)\|_{L^2} \le C\,\left(\|(u_{0h},\theta_0)\|_{L^2} + \|(\Lambda_h^{-\sigma}u_{0h},\Lambda_h^{-\sigma}\theta_0)\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

Using Lemma 3.1 and noticing that

$$\left|\frac{|\xi_h|^2}{|\xi|^2}G_1\right| = \left|\frac{|\xi_h|^2}{|\xi|^2}\, e^{-|\xi_h|^2 t}\,\frac{\sin\frac{|\xi_h|}{|\xi|}t}{\frac{|\xi_h|}{|\xi|}}\right| \le e^{-|\xi_h|^2 t},$$

we have from (3.3) that

$$\begin{aligned}\|u_3\|_{L^2} &\le C\,\|e^{-|\xi_h|^2 t}\widehat{u}_{03}\|_{L^2} + \|e^{-|\xi_h|^2 t}\widehat{\theta_0}\|_{L^2}\\ &\le C\,\left(\|(u_{03},\theta_0)\|_{L^2} + \|(\Lambda_h^{-\sigma}u_{03},\Lambda_h^{-\sigma}\theta_0)\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.\end{aligned}$$

The estimate of the first term in (3.4) needs some attention. By $\nabla\cdot u_0 = 0$ or $\xi_h\cdot\widehat{u}_{0h}+\xi_3\widehat{u}_{03} = 0$,

$$\begin{aligned}\|G_1\widehat{u}_{03}\|_{L^2} &= \left\|e^{-|\xi_h|^2 t}\frac{\sin\frac{|\xi_h|}{|\xi|}t}{\frac{|\xi_h|}{|\xi|}}\widehat{u}_{03}\right\|_{L^2} \le \left\|e^{-|\xi_h|^2 t}\frac{|\xi|}{|\xi_h|}\widehat{u}_{03}\right\|_{L^2}\\ &\le \left\|e^{-|\xi_h|^2 t}\frac{|\xi_h|+|\xi_3|}{|\xi_h|}\widehat{u}_{03}\right\|_{L^2}\\ &\le \|e^{-|\xi_h|^2 t}\widehat{u}_{03}\|_{L^2} + \left\|e^{-|\xi_h|^2 t}\frac{|\xi_h\cdot\widehat{u_{0h}}|}{|\xi_h|}\right\|_{L^2}\\ &\le C\,\left(\|u_0\|_{L^2} + \|\Lambda_h^{-\sigma}u_0\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.\end{aligned}$$

The second term is easily bounded,

$$\|(G_3+|\xi_h|^2G_1)\widehat{\theta_0}\|_{L^2} \le C\,\left(\|\theta_0\|_{L^2} + \|\Lambda_h^{-\sigma}\theta_0\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

Therefore,

$$\|\theta\|_{L^2} \le C\,\left(\|(u_0,\theta_0)\|_{L^2} + \|(\Lambda_h^{-\sigma}u_0,\Lambda_h^{-\sigma}\theta_0)\|_{L^2}\right)(1+t)^{-\frac{\sigma}{2}}.$$

Therefore (1.16) is proven. The proof of (1.17) is similarly to that for (1.16). In fact, (1.17) can be shown by repeating the process for (1.16) with $\partial_3 u$ and $\partial_3\theta$ replacing u and θ, respectively. We now turn to (1.18). Noticing that the upper bound for each term in (3.2), (3.3) and (3.4) contains the factor $e^{-|\xi_h|^2 t}$, we can easily obtain the extra decay factor via the inequality, for any $t>0$,

$$\||\xi_h|e^{-|\xi_h|^2 t}f\|_{L^2} \le C\,t^{-\frac{\sigma+1}{2}}\,\|\Lambda_h^{-\sigma}f\|_{L^2}.$$

This explains (1.18). Combining (1.18) with $t \geq 1$ and the basic inequality with $0 \leq t < 1$,

$$\| |\xi_h| e^{-|\xi_h|^2 t} f \|_{L^2} \leq C \, \| |\xi_h| f \|_{L^2}$$

leads to (1.19). This completes the proof of Proposition 1.1. □

4 Nonlinear stability

This section is devoted to proving Proposition 1.2, which establishes the nonlinear stability. This proposition serves as a preparation for our main result on the optimal decay proven in the next section.

The proof uses the following lemma that provides anisotropic upper bounds for the integral of a triple product. It is a very powerful tool in dealing with anisotropic equations.

Lemma 4.1 *The following estimates hold when the right-hand sides are all bounded.*

$$\int_{\mathbb{R}^3} |fgh| dx \leq C \, \|f\|_{L^2}^{\frac{1}{2}} \|\partial_1 f\|_{L^2}^{\frac{1}{2}} \|g\|_{L^2}^{\frac{1}{2}} \|\partial_2 g\|_{L^2}^{\frac{1}{2}} \|h\|_{L^2}^{\frac{1}{2}} \|\partial_3 h\|_{L^2}^{\frac{1}{2}},$$

$$\int_{\mathbb{R}^3} |fgh| dx \leq C \, \|f\|_{L^2}^{\frac{1}{4}} \|\partial_1 f\|_{L^2}^{\frac{1}{4}} \|\partial_2 f\|_{L^2}^{\frac{1}{4}} \|\partial_1 \partial_2 f\|_{L^2}^{\frac{1}{4}} \|g\|_{L^2}^{\frac{1}{2}} \|\partial_3 g\|_{L^2}^{\frac{1}{2}} \|h\|_{L^2}.$$

A simple proof of this lemma can be found in [36]. The 2D version of such anisotropic upper bounds can be found in [7]. We will not reproduce a proof of Lemma 4.1 here, but instead begin with the proof of Proposition 1.2.

Proof of Proposition 1.2 Since the local (in time) well-posedness of (1.3) can be established via a standard approach (see [21]), our attention is focused on the global bound of (u, θ). The framework of the proof is the bootstrapping argument. Define the energy functional $E(t)$ by

$$E(t) = \sup_{0 \leq \tau \leq t} \|(u, \theta)(\tau)\|_{H^m}^2 + \nu \int_0^t \|\nabla_h u\|_{H^m}^2 \, d\tau + \eta \int_0^t \|\nabla_h \theta\|_{H^m}^2 d\tau.$$

Our main efforts are devoted to showing that, for a constant $C > 0$ and for $t > 0$,

$$E(t) \leq E(0) + C \, E(t)^{\frac{3}{2}}. \tag{4.1}$$

Once (4.1) is shown, then a direct application of the bootstrapping argument implies that, if

$$E(0) = \|(u_0, \theta_0)\|_{H^m}^2 \leq \frac{1}{16C^2} \quad \text{or} \quad \|(u_0, \theta_0)\|_{H^m} \leq \varepsilon := \frac{1}{4C}, \tag{4.2}$$

then,

$$E(t) \leq \frac{1}{8C^2} \quad \text{for all } t > 0. \tag{4.3}$$

In fact, if we make the ansatz that

$$E(t) \leq \frac{1}{4C^2}. \tag{4.4}$$

Inserting (4.4) in (4.1) and invoking (4.2) yields

$$E(t) \leq E(0) + \frac{1}{2} E(t) \quad \text{or} \quad E(t) \leq 2E(0) \leq \frac{1}{8C^2},$$

which is only half of the bound in the ansatz in (4.4). The bootstrapping argument then implies (4.3). It remains to prove (4.1). Due to the norm equivalence

$$\|f\|_{H^m}^2 \sim \|f\|_{L^2}^2 + \sum_{i=1}^{3} \|\partial_i^m f\|_{L^2}^2,$$

it suffices to bound $\|(u, \theta)\|_{L^2}$ and $\sum_{i=1}^{3} \|(\partial_i^m u, \partial_i^m \theta)\|_{L^2}$. First of all, we have the global L^2-bound. Dotting the equations in (1.3) by (u, θ) and integrating by parts, we find

$$\|(u, \theta)(t)\|_{L^2}^2 + 2\nu \int_0^t \|\nabla_h u\|_{L^2}^2 \, d\tau + 2\eta \int_0^t \|\nabla_h \theta\|_{L^2}^2 d\tau = \|(u_0, \theta_0)\|_{L^2}^2. \tag{4.5}$$

Applying the differential operator ∂_i^m to the equations in (1.3), dotting the resulting equations by $(\partial_i^m u, \partial_i^m \theta)$, and integrating by parts, we have

$$\begin{aligned}
&\frac{d}{dt} \sum_{i=1}^{3} \left(\|\partial_i^m u\|_{L^2}^2 + \|\partial_i^m \theta\|_{L^2}^2\right) + 2\nu \sum_{i=1}^{3} \|\nabla_h \partial_i^m u\|_{L^2}^2 + 2\eta \sum_{i=1}^{3} \|\nabla_h \partial_i^m \theta\|_{L^2}^m \\
&\quad = J_1 + J_2,
\end{aligned} \tag{4.6}$$

where J_1 and J_2 are given by

$$J_1 = -\sum_{i=1}^{3} \int \partial_i^m (u \cdot \nabla u) \cdot \partial_i^m u \, dx,$$

$$J_2 = -\sum_{i=1}^{3} \int \partial_i^m (u \cdot \nabla \theta) \cdot \partial_i^m \theta \, dx.$$

Here we have used the fact that

$$\int (\partial_i^m \theta e_3 \cdot \partial_i^m u - \partial_i^m u_3 \partial_i^m \theta) \, dx = 0.$$

We decompose J_1 as

$$\begin{aligned}
J_1 &= -\sum_{i=1}^{2} \int \partial_i^m (u \cdot \nabla u) \cdot \partial_i^m u \, dx - \sum_{k=1}^{2} \int \partial_3^m (u_k \cdot \partial_k u) \cdot \partial_3^m u \, dx \\
&\quad - \int \partial_3^m (u_3 \cdot \partial_3 u) \cdot \partial_3^m u \, dx \\
&:= J_{11} + J_{12} + J_{13}.
\end{aligned}$$

J_{11} is easy to deal with. Due to $\nabla \cdot u = 0$,

$$\begin{aligned}
J_{11} &= \sum_{i=1}^{2} \int \partial_i^m (u \otimes u) \cdot \nabla \partial_i^m u \, dx \\
&\le C \, \|u\|_{L^\infty} \, \|\nabla_h u\|_{H^m}^2 \le C \, \|u\|_{H^m} \, \|\nabla_h u\|_{H^m}^2.
\end{aligned}$$

Here we have used the calculus inequality for a product. For any nonnegative integer m, there exists $C > 0$ such that, for any $u, v \in L^\infty \cap H^m$,

$$\|D^m(uv)\|_{L^2} \le C \left(\|u\|_{L^\infty} \|D^m v||_{L^2} + \|D^m u\|_{L^2} \|v\|_{L^\infty}\right).$$

This inequality can be found in [21,p.98]. By the Leibniz Formula and

$$\int (u \cdot \nabla \partial_i^m u) \cdot \partial_i^m u \, dx = 0, \quad i = 1, 2, 3,$$

we have

$$J_{12} = -\sum_{k=1}^{2} \sum_{l=1}^{m} \mathcal{C}_m^l \int \partial_3^l u_k \cdot \partial_3^{m-l} \partial_k u \cdot \partial_3^m u \, dx,$$

where $\mathcal{C}_m^l$ denotes the combinatorial number,

$$\mathcal{C}_m^l = \frac{m!}{l!(m-l)!}.$$

By Lemma 4.1,

$$\begin{aligned}
|J_{12}| &\leq C \sum_{k=1}^{2} \sum_{l=1}^{m} \|\partial_3^l u_k\|_{L^2}^{\frac{1}{2}} \|\partial_1 \partial_3^l u_k\|_{L^2}^{\frac{1}{2}} \|\partial_3^{m-l} \partial_k u\|_{L^2}^{\frac{1}{2}} \|\partial_3 \partial_3^{m-l} \partial_k u\|_{L^2}^{\frac{1}{2}} \\
&\quad \times \|\partial_3^m u\|_{L^2}^{\frac{1}{2}} \|\partial_2 \partial_3^m u\|_{L^2}^{\frac{1}{2}} \\
&\leq C \|u\|_{H^m} \|\nabla_h u\|_{H^m}^2.
\end{aligned}$$

By $\nabla \cdot u = 0$ or $\partial_3 u_3 = -\nabla_h \cdot u_h$ and Lemma 4.1,

$$\begin{aligned}
J_{13} &= -\sum_{l=1}^{m} \mathcal{C}_m^l \int \partial_3^l u_3 \cdot \partial_3^{m-l} \partial_3 u \cdot \partial_3^m u \, dx \\
&= \sum_{l=1}^{m} \mathcal{C}_m^l \int \partial_3^{l-1} \nabla_h \cdot u_h \cdot \partial_3^{m-l} \partial_3 u \cdot \partial_3^m u \, dx \\
&\leq C \sum_{l=1}^{m} \|\partial_3^{l-1} \nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \|\partial_3 \partial_3^{l-1} \nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \|\partial_3^{m-l} \partial_3 u\|_{L^2}^{\frac{1}{2}} \|\partial_1 \partial_3^{m-l} \partial_3 u\|_{L^2}^{\frac{1}{2}} \\
&\quad \times \|\partial_3^m u\|_{L^2}^{\frac{1}{2}} \|\partial_2 \partial_3^m u\|_{L^2}^{\frac{1}{2}} \\
&\leq C \|u\|_{H^m} \|\nabla_h u\|_{H^m}^2.
\end{aligned}$$

Therefore,

$$|J_1| \leq C \|u\|_{H^m} \|\nabla_h u\|_{H^m}^2. \tag{4.7}$$

Using $\nabla \cdot u = 0$, we decompose J_2 as

$$\begin{aligned}
J_2 &= -\sum_{i=1}^{2} \int \partial_i^m \nabla \cdot (u\theta) \cdot \partial_i^m \theta \, dx \\
&\quad - \sum_{k=1}^{2} \int \partial_3^m (u_k \cdot \partial_k \theta) \cdot \partial_3^m \theta \, dx - \int \partial_3^m (u_3 \partial_3 \theta) \cdot \partial_3^m \theta \, dx \\
&:= J_{21} + J_{22} + J_{23}.
\end{aligned}$$

By integration by parts and Sobolev's inequality, for $m \geq 2$,

$$
\begin{aligned}
J_{21} &= \sum_{i=1}^{2} \int \partial_i^m (u\theta) \cdot \nabla \partial_i^m \theta \, dx \\
&\leq C \sum_{i=1}^{2} \left(\|\partial_i^m u\|_{L^2} \|\theta\|_{L^\infty} + \|u\|_{L^\infty} \|\partial_i^m \theta\|_{L^2} \right) \|\nabla \partial_i^m \theta\|_{L^2} \\
&\leq C \left(\|u\|_{H^m} + \|\theta\|_{H^m} \right) \left(\|\nabla_h u\|_{H^m}^2 + \|\nabla_h \theta\|_{H^m}^2 \right).
\end{aligned}
$$

By the Leibniz Formula and

$$
\int (u \cdot \nabla \partial_i^m \theta) \cdot \partial_i^m \theta \, dx = 0, \quad i = 1, 2, 3,
$$

we have

$$
\begin{aligned}
J_{22} + J_{23} = &- \sum_{k=1}^{2} \sum_{l=1}^{m} C_m^l \int \partial_3^l u_k \cdot \partial_3^{m-l} \partial_k \theta \cdot \partial_3^m \theta \, dx \\
&- \sum_{l=1}^{m} C_m^l \int \partial_3^l u_3 \cdot \partial_3^{m-l} \partial_3 \theta \cdot \partial_3^m \theta \, dx.
\end{aligned}
$$

By $\nabla \cdot u = 0$ or $\partial_3 u_3 = -\nabla_h \cdot u_h$ and Lemma 4.1,

$$
\begin{aligned}
|J_{22}| + |J_{23}| \leq\ & C \sum_{k=1}^{2} \sum_{l=1}^{m} \|\partial_3^l u_k\|_{L^2}^{\frac{1}{2}} \|\partial_1 \partial_3^l u_k\|_{L^2}^{\frac{1}{2}} \|\partial_3^{m-l} \partial_k \theta\|_{L^2}^{\frac{1}{2}} \\
& \times \|\partial_3 \partial_3^{m-l} \partial_k \theta\|_{L^2}^{\frac{1}{2}} \|\partial_3^m \theta\|_{L^2}^{\frac{1}{2}} \|\partial_2 \partial_3^m \theta\|_{L^2}^{\frac{1}{2}} \\
& + C \sum_{l=1}^{m} \|\partial^{l-1} \nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \|\partial_3 \partial^{l-1} \nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \|\partial_3^{m-l} \partial_3 \theta\|_{L^2}^{\frac{1}{2}} \\
& \times \|\partial_1 \partial_3^{m-l} \partial_3 \theta\|_{L^2}^{\frac{1}{2}} \|\partial_3^m \theta\|_{L^2}^{\frac{1}{2}} \|\partial_2 \partial_3^m \theta\|_{L^2}^{\frac{1}{2}} \\
& \leq C \left(\|u\|_{H^m} + \|\theta\|_{H^m} \right) \left(\|\nabla_h u\|_{H^m}^2 + \|\nabla_h \theta\|_{H^m}^2 \right).
\end{aligned}
$$

Collecting the bounds for J_2, we obtain

$$
|J_2| \leq C \left(\|u\|_{H^m} + \|\theta\|_{H^m} \right) \left(\|\nabla_h u\|_{H^m}^2 + \|\nabla_h \theta\|_{H^m}^2 \right). \tag{4.8}
$$

Inserting (4.7) and (4.8) in (4.6), integrating in time over $[0, t]$ and adding to (4.5), we deduce

$$
\begin{aligned}
E(t) \leq\ & E(0) + C \int_0^t \Big(\|u\|_{H^m} \|\nabla_h u\|_{H^m}^2 \\
& + \left(\|u\|_{H^m} + \|\theta\|_{H^m} \right) \left(\|\nabla_h u\|_{H^m}^2 + \|\nabla_h \theta\|_{H^m}^2 \right) \Big) \, d\tau \\
\leq\ & E(0) + C\, E(t)^{\frac{3}{2}},
\end{aligned}
$$

which is the desired inequality (4.1). This completes the proof of Proposition 1.2. □

5 Optimal decays for the nonlinear system

This section proves our main result, Theorem 1.3. We need several tools, which are stated in the following lemmas. The first lemma provides an upper bound for the L^p-norm of a one-dimensional function, which serves as a basic ingredient for anisotropic upper bounds. A proof can be found in [38].

Lemma 5.1 *Let $2 \le p \le \infty$. Let $s > \frac{1}{2} - \frac{1}{p}$. Then, there exists a constant $C = C(p, s)$ such that, for any 1D function $f \in H^s(\mathbb{R})$,*

$$\|f\|_{L^p(\mathbb{R})} \le C\, \|f\|_{L^2(\mathbb{R})}^{1-\frac{1}{s}\left(\frac{1}{2}-\frac{1}{p}\right)}\, \|\Lambda^s f\|_{L^2(\mathbb{R})}^{\frac{1}{s}\left(\frac{1}{2}-\frac{1}{p}\right)}.$$

In particular, if $p = \infty$ and $s = 1$, then $f = f(x_3)$,

$$\|f\|_{L^\infty} \le C\, \|f\|_{L^2(\mathbb{R})}^{\frac{1}{2}}\, \|\partial_3 f\|_{L^2(\mathbb{R})}^{\frac{1}{2}}.$$

The second lemma states Minkowski's inequality. It is an elementary tool that allows us to estimate the Lebesgue norm with larger index first followed by the Lebesgue norm with a smaller index. The following version is taken from [1,p. 4] and a more general statement can be found in [19, p. 47].

Lemma 5.2 *For a nonnegative measurable function f over $\mathbb{R}^m \times \mathbb{R}^n$ and for $1 \le p \le q \le \infty$,*

$$\big\| \|f\|_{L^p(\mathbb{R}^m)} \big\|_{L^q(\mathbb{R}^n)} \le \big\| \|f\|_{L^q(\mathbb{R}^n)} \big\|_{L^p(\mathbb{R}^m)}.$$

For convenience, we introduce the notation

$$L^q_h(\mathbb{R}^3) := L^q_{x_1,x_2}(\mathbb{R}^3), \quad \|f\|_{L^p_h L^q_{x_3}} := \Big\| \|f\|_{L^q_{x_3}} \Big\|_{L^p_h},$$

which is frequently used in the context.

The next lemma provides an exact $L^p - L^q$ decay estimate for the generalized heat operator associated with a fractional Laplacian (see, e.g, [35]).

Lemma 5.3 *Let $\beta \ge 0$, $\alpha > 0$, $\nu > 0$, $1 \le p \le q \le \infty$. Then*

$$\|\Lambda^\beta e^{-\nu(-\Delta)^\alpha t} f\|_{L^q(\mathbb{R}^d)} \le C\, t^{-\frac{\beta}{2\alpha} - \frac{d}{2\alpha}\left(\frac{1}{p}-\frac{1}{q}\right)} \|f\|_{L^p(\mathbb{R}^d)}.$$

Proof of Theorem 1.3 The bootstrapping argument is suitable for our purpose. We assume the initial datum (u_0, θ_0) satisfies (1.20), (1.21) and (1.22) for sufficiently small $\varepsilon > 0$. The bootstrapping argument starts with the ansatz that, for a suitably selected $C_0 > 0$,

$$\|u(t)\|_{H^4},\ \|\theta(t)\|_{H^4} \le C_0\, \varepsilon, \tag{5.1}$$

$$\|\Lambda_h^{-\sigma} u(t)\|_{L^2},\ \|\Lambda_h^{-\sigma}\theta(t)\|_{L^2} \le C_0\, \varepsilon, \tag{5.2}$$

$$\|u(t)\|_{L^2}, \|\partial_3 u(t)\|_{L^2}, \|\theta(t)\|_{L^2}, \|\partial_3\theta(t)\|_{L^2} \le C_0\, \varepsilon(1+t)^{-\frac{\sigma}{2}}, \tag{5.3}$$

$$\|\nabla_h u(t)\|_{L^2},\ \|\nabla_h\theta(t)\|_{L^2} \le C_0\, \varepsilon(1+t)^{-\frac{\sigma}{2}-\frac{1}{2}}. \tag{5.4}$$

for $t \in [0, T]$ with some $T > 0$. These inequalities hold on the initial time interval $[0, T]$ guaranteed by local well-posedness. We then show that (5.1), (5.2), (5.3) and (5.4) remain true with C_0 replaced by $C_0/2$, namely

$$\|u(t)\|_{H^4},\ \|\theta(t)\|_{H^4} \le \frac{C_0}{2}\, \varepsilon, \tag{5.5}$$

$$\|\Lambda_h^{-\sigma} u(t)\|_{L^2}, \ \|\Lambda_h^{-\sigma} \theta(t)\|_{L^2} \leq \frac{C_0}{2}\,\varepsilon, \tag{5.6}$$

$$\|u(t)\|_{L^2}, \|\partial_3 u(t)\|_{L^2}, \|\theta(t)\|_{L^2}, \|\partial_3 \theta(t)\|_{L^2} \leq \frac{C_0}{2}\,\varepsilon(1+t)^{-\frac{\sigma}{2}}, \tag{5.7}$$

$$\|\nabla_h u(t)\|_{L^2}, \ \|\nabla_h \theta(t)\|_{L^2} \leq \frac{C_0}{2}\,\varepsilon(1+t)^{-\frac{\sigma}{2}-\frac{1}{2}}. \tag{5.8}$$

The bootstrapping argument then asserts that (5.5), (5.6), (5.7) and (5.8) hold for all $t > 0$.

It then suffices to prove (5.5) through (5.8). (5.5) follows directly from Proposition 1.2 with $m = 4$. By Proposition 1.2,

$$\|u(t)\|_{H^4}, \ \|\theta(t)\|_{H^4} \leq C_1\varepsilon.$$

Then (5.5) clearly holds when we take $C_0 \geq 2C_1$. The rest of this section is divided into four subsections. The first subsection estimates $\|\Lambda_h^{-\sigma} u(t)\|_{L^2}$ and $\|\Lambda_h^{-\sigma} \theta(t)\|_{L^2}$ and verifies (5.6). The second subsection estimates $\|u(t)\|_{L^2}$ and $\|\theta(t)\|_{L^2}$ and verifies part of (5.7). The third subsection bounds $\|\partial_3 u(t)\|_{L^2}$ and $\|\partial_3 \theta(t)\|_{L^2}$ and completes verifying (5.7). The last subsection works on $\|\nabla_h u(t)\|_{L^2}$ and $\|\nabla_h \theta(t)\|_{L^2}$ and proves (5.8).

5.1 Verification of (5.6)

This subsection estimates $\|\Lambda_h^{-\sigma} u(t)\|_{L^2}$ and $\|\Lambda_h^{-\sigma} \theta(t)\|_{L^2}$ and verifies (5.6). Applying $\Lambda_h^{-\sigma}$ to (1.3) and dotting with $(\Lambda_h^{-\sigma} u, \Lambda_h^{-\sigma} \theta)$, we obtain

$$\begin{aligned}
&\frac{d}{dt}\left(\|\Lambda_h^{-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{-\sigma} \theta\|_{L^2}^2\right) + 2\left(\|\Lambda_h^{1-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{1-\sigma} \theta\|_{L^2}^2\right) \\
&\quad = -2\int \Lambda_h^{-\sigma}(u\cdot\nabla u)\cdot \Lambda_h^{-\sigma} u\,dx - 2\int \Lambda_h^{-\sigma}(u\cdot\nabla\theta)\cdot\Lambda_h^{-\sigma}\theta\,dx \\
&\quad := N_1 + N_2,
\end{aligned} \tag{5.9}$$

where we have used

$$\int (\Lambda_h^{-\sigma}(\theta e_3)\cdot \Lambda_h^{-\sigma} u - \Lambda_h^{-\sigma} u_3\, \Lambda_h^{-\sigma}\theta)\,dx = 0.$$

We distinguish the horizontal derivatives from the vertical ones and write N_1 as

$$\begin{aligned}
N_1 &= -2\int \Lambda_h^{-\sigma}(u_h\cdot\nabla_h u)\cdot\Lambda_h^{-\sigma} u\,dx - 2\int \Lambda_h^{-\sigma}(u_3\partial_3 u_h)\cdot \Lambda_h^{-\sigma} u_h\,dx \\
&\quad - 2\int \Lambda_h^{-\sigma}(u_3\partial_3 u_3)\cdot\Lambda_h^{-\sigma} u_3\,dx \\
&:= N_{11} + N_{12} + N_{13}.
\end{aligned} \tag{5.10}$$

N_{12} involves the unfavorable derivative ∂_3 and may potentially generate the worst upper bound. We deal with this term first. By Hölder's inequality, the Hardy-Littlewood-Sobolev inequality and Lemmas 5.1, 5.2 and 5.3,

$$\begin{aligned}
|N_{12}| &\leq \|\Lambda_h^{-\sigma}(u_3\partial_3 u_h)\|_{L^2}\,\|\Lambda_h^{-\sigma} u_h\|_{L^2} \\
&= \left\|\|\Lambda_h^{-\sigma}(u_3\partial_3 u_h)\|_{L_h^2}\right\|_{L_{x_3}^2}\,\|\Lambda_h^{-\sigma} u_h\|_{L^2} \\
&\leq \left\|\|u_3\partial_3 u_h\|_{L_h^q}\right\|_{L_{x_3}^2}\,\|\Lambda_h^{-\sigma} u_h\|_{L^2}
\end{aligned}$$

$$
\begin{aligned}
&\leq \Big\| \|u_3 \partial_3 u_h\|_{L^2_{x_3}} \Big\|_{L^q_h} \, \|\Lambda_h^{-\sigma} u_h\|_{L^2} \\
&\leq \Big\| \|u_3\|_{L^\infty_{x_3}} \|\partial_3 u_h\|_{L^2_{x_3}} \Big\|_{L^q_h} \, \|\Lambda_h^{-\sigma} u_h\|_{L^2} \\
&\leq \|u_3\|_{L_h^{\frac{2}{\sigma}} L^\infty_{x_3}} \|\partial_3 u_h\|_{L^2_h L^2_{x_3}} \, \|\Lambda_h^{-\sigma} u_h\|_{L^2},
\end{aligned}
$$

where

$$
\frac{1}{2} + \frac{\sigma}{2} = \frac{1}{q}.
$$

Clearly, for $\frac{3}{4} \leq \sigma < 1$, we have

$$
1 < q < 2.
$$

The first part on the right-hand side can be further bounded as follows. By Lemma 5.1 and Hölder's inequality with $\frac{\sigma}{2} = \frac{1}{4} + \frac{2\sigma - 1}{4}$,

$$
\begin{aligned}
\|u_3\|_{L_h^{\frac{2}{\sigma}} L^\infty_{x_3}} &\leq C \left\| \|u_3\|_{L^2_{x_3}}^{\frac{1}{2}} \, \|\partial_3 u_3\|_{L^2_{x_3}}^{\frac{1}{2}} \right\|_{L_h^{\frac{2}{\sigma}}} \\
&\leq C \left\| \|u_3\|_{L^2_{x_3}}^{\frac{1}{2}} \right\|_{L_h^{\frac{4}{2\sigma-1}}} \left\| \|\partial_3 u_3\|_{L^2_{x_3}}^{\frac{1}{2}} \right\|_{L^4_h} \\
&\leq C \, \|\partial_3 u_3\|_{L^2}^{\frac{1}{2}} \, \|u_3\|^{\frac{1}{2}}_{L_h^{\frac{2}{2\sigma-1}} L^2_{x_3}} \\
&\leq C \, \|\nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \, \|u_3\|^{\frac{1}{2}}_{L^2_{x_3} L_h^{\frac{2}{2\sigma-1}}} \\
&\leq C \, \|\nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \, \|u_3\|_{L^2}^{\sigma - \frac{1}{2}} \, \|\nabla_h u_3\|_{L^2}^{1-\sigma}.
\end{aligned}
$$

Thus we have obtained the following bound

$$
|N_{12}| \leq C \, \|\nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \, \|u_3\|_{L^2}^{\sigma - \frac{1}{2}} \, \|\nabla_h u_3\|_{L^2}^{1-\sigma} \, \|\partial_3 u_h\|_{L^2} \, \|\Lambda_h^{-\sigma} u_h\|_{L^2}.
$$

The other two terms in (5.10) can be estimated similarly,

$$
\begin{aligned}
N_{11} &\leq C \, \|\partial_3 u_h\|_{L^2}^{\frac{1}{2}} \, \|u_h\|_{L^2}^{\sigma - \frac{1}{2}} \, \|\nabla_h u_h\|_{L^2}^{1-\sigma} \, \|\nabla_h u\|_{L^2} \, \|\Lambda_h^{-\sigma} u\|_{L^2}, \\
N_{13} &\leq C \, \|\nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}} \, \|u_3\|_{L^2}^{\sigma - \frac{1}{2}} \, \|\nabla_h u_3\|_{L^2}^{1-\sigma} \, \|\nabla_h \cdot u_h\|_{L^2} \, \|\Lambda_h^{-\sigma} u_3\|_{L^2}.
\end{aligned}
$$

Combining the upper bounds for N_1 yields

$$
|N_1| \leq C \, \|u\|_{L^2}^{\sigma - \frac{1}{2}} \|\nabla u\|_{L^2} \|\nabla_h u\|_{L^2}^{\frac{3}{2} - \sigma} \| \|\Lambda_h^{-\sigma} u\|_{L^2}. \tag{5.11}
$$

The estimate of N_2 shares some similarities with that for N_1 and starts by writing

$$
\begin{aligned}
N_2 &= -2 \int \Lambda_h^{-\sigma} (u_h \cdot \nabla_h \theta) \cdot \Lambda_h^{-\sigma} \theta \, dx - 2 \int \Lambda_h^{-\sigma} (u_3 \partial_3 \theta) \cdot \Lambda_h^{-\sigma} \theta \, dx \\
&= N_{21} + N_{22}.
\end{aligned}
$$

The two terms N_{21} and N_{22} can be estimated similarly as N_{11} and N_{12},

$$
\begin{aligned}
N_{21} &\le C\, \|\partial_3 u_h\|_{L^2}^{\frac{1}{2}}\, \|u_h\|_{L^2}^{\sigma-\frac{1}{2}}\, \|\nabla_h u_h\|_{L^2}^{1-\sigma}\, \|\nabla_h \theta\|_{L^2}\, \|\Lambda_h^{-\sigma}\theta\|_{L^2},\\
N_{22} &\le C\, \|\nabla_h \cdot u_h\|_{L^2}^{\frac{1}{2}}\, \|u_3\|_{L^2}^{\sigma-\frac{1}{2}}\, \|\nabla_h u_3\|_{L^2}^{1-\sigma}\, \|\partial_3 \theta\|_{L^2}\, \|\Lambda_h^{-\sigma}\theta\|_{L^2}.
\end{aligned}
$$

Therefore,

$$
|N_2| \le C\, \|u\|_{L^2}^{\sigma-\frac{1}{2}}\, \|(\nabla u, \nabla\theta)\|_{L^2} \left(\|\nabla_h u\|_{L^2}^{\frac{3}{2}-\sigma} + \|\nabla_h \theta\|_{L^2}^{\frac{3}{2}-\sigma} \right) \|\Lambda_h^{-\sigma}\theta\|_{L^2}. \tag{5.12}
$$

Inserting (5.11) and (5.12) in (5.9) and integrating in time, we obtain

$$
\begin{aligned}
&\|\Lambda_h^{-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{-\sigma}\theta\|_{L^2}^2 + 2\int_0^t \left(\|\Lambda_h^{1-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{1-\sigma}\theta\|_{L^2}^2 \right) d\tau\\
&\quad \le C \int_0^t \|u\|_{L^2}^{\sigma-\frac{1}{2}}\, \|(\nabla u, \nabla\theta)\|_{L^2} \left(\|\nabla_h u\|_{L^2}^{\frac{3}{2}-\sigma} + \|\nabla_h \theta\|_{L^2}^{\frac{3}{2}-\sigma} \right) \|\Lambda_h^{-\sigma}\theta\|_{L^2}\, d\tau.
\end{aligned}
$$

Invoking the ansatz in (5.1) through (5.4), we obtain

$$
\begin{aligned}
\|\Lambda_h^{-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{-\sigma}\theta\|_{L^2}^2 &\le C\, C_0^3\, \varepsilon^3 \int_0^t (1+\tau)^{-\frac{\sigma}{2}(\sigma+\frac{1}{2})}\, (1+\tau)^{(\frac{3}{2}-\sigma)(-\frac{\sigma}{2}-\frac{1}{2})}\, d\tau\\
&\le C\, C_0^3\, \varepsilon^3 \int_0^t (1+\tau)^{-\frac{\sigma}{2}-\frac{3}{4}}\, d\tau\\
&\le C\, C_0^3\, \varepsilon^3
\end{aligned}
$$

for any $\sigma > \frac{1}{2}$. If we choose $\varepsilon > 0$ to be sufficiently small such that

$$
C C_0 \varepsilon \le \frac{1}{4},
$$

then

$$
\|\Lambda_h^{-\sigma} u\|_{L^2}^2 + \|\Lambda_h^{-\sigma}\theta\|_{L^2}^2 \le \frac{1}{4} C_0^2 \varepsilon^2,
$$

which, in particular, verifies (5.6).

5.2 Estimates of $\|u\|_{L^2}$ and $\|\theta\|_{L^2}$ and verification of (5.7)

This subsection verifies part of (5.7). We take advantage of the integral representation formula in (2.1), (2.2) and (2.3). Since the linear terms in these formula have been estimated in Proposition 1.1 and its proof, it suffices to bound the time integral parts. For notational convenience, we denote the time integral terms in (2.1), (2.2) and (2.3) as

$$
\begin{aligned}
K_1 &= \int_0^t e^{\lambda_1(t-\tau)} (\mathbb{P}\widehat{(u\cdot\nabla u)})_h(\tau)\, d\tau,\\
K_2 &= \int_0^t \left(\frac{\xi_h \xi_3}{|\xi_h|^2} e^{\lambda_1(t-\tau} + \frac{\xi_h \xi_3}{|\xi_h|^2} G_2(t-\tau) + \xi_h \xi_3 G_1(t-\tau) \right) (\mathbb{P}\widehat{(u\cdot\nabla u)})_3(\tau)\, d\tau,\\
K_3 &= \int_0^t \frac{\xi_h \xi_3}{|\xi|^2} G_1(t-\tau) \widehat{(u\cdot\nabla\theta)}(\tau)\, d\tau,\\
K_4 &= \int_0^t (G_2 + |\xi_h|^2 G_1)(t-\tau) \mathbb{P}\widehat{(u\cdot\nabla u)}_3(\tau)\, d\tau,
\end{aligned}
$$

$$
\begin{aligned}
K_5 &= \int_0^t \frac{|\xi_h|^2}{|\xi|^2} G_1(t-\tau)\,(\widehat{u\cdot\nabla\theta})(\tau)\,d\tau,\\
K_6 &= \int_0^t G_1(t-\tau)\mathbb{P}(\widehat{u\cdot\nabla u})_3(\tau)\,d\tau,\\
K_7 &= \int_0^t (G_3(t-\tau)+|\xi_h|^2 G_1(t-\tau))(\widehat{u\cdot\nabla\theta})(\tau)\,d\tau.
\end{aligned}
$$

We first bound u_h and start with K_1. By the definition of the projection operator $\mathbb{P} = I-\nabla\Delta^{-1}\nabla\cdot$, we have

$$
\begin{aligned}
\mathbb{P}(u\cdot\nabla u))_h &= u\cdot\nabla u_h - \nabla_h\Delta^{-1}\nabla\cdot(u\cdot\nabla u)\\
&= u_h\cdot\nabla_h u_h + u_3\partial_3 u_h - \Delta^{-1}\nabla\cdot\nabla\cdot\nabla_h(u\otimes u).
\end{aligned}
\tag{5.13}
$$

Correspondingly the upper bound consists of three parts,

$$
\begin{aligned}
\|K_1\|_{L^2} \le & \int_0^t \|e^{\Delta_h(t-\tau)} u_h\cdot\nabla_h u_h(\tau)\|_{L^2}\,d\tau\\
&+ C\int_0^t \|e^{\Delta_h(t-\tau)}\nabla_h(u\otimes u)(\tau)\|_{L^2}\,d\tau\\
&+ \int_0^t \|e^{\Delta_h(t-\tau)} u_3\partial_3 u_h(\tau)\|_{L^2}\,d\tau\\
=& K_{11}+K_{12}+K_{13},
\end{aligned}
$$

where we have used the bounedness of the Riesz transform on L^2,

$$
\|\Delta^{-1}\nabla\cdot\nabla\cdot F\|_{L^2} \le C\,\|F\|_{L^2}.
$$

K_{11} and K_{12} involve the good derivative ∇_h and are easier to control while K_{13} is harder due to the bad derivative ∂_3. By Lemmas 5.1, 5.2 and 5.3,

$$
\begin{aligned}
K_{11} &\le \int_0^t \Big\| \|e^{\Delta_h(t-\tau)} u_h\cdot\nabla_h u(\tau)\|_{L^2_h}\Big\|_{L^2_{x_3}}\,d\tau\\
&\le \int_0^t \Big\| (t-\tau)^{-\frac12}\|u_h\cdot\nabla_h u(\tau)\|_{L^1_h}\Big\|_{L^2_{x_3}}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\Big\| \|u_h(\tau)\|_{L^2_h}\|\nabla_h u(\tau)\|_{L^2_h}\Big\|_{L^2_{x_3}}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_h(\tau)\|_{L^\infty_{x_3}L^2_h}\|\nabla_h u(\tau)\|_{L^2_{x_3}L^2_h}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_h(\tau)\|_{L^2_h L^\infty_{x_3}}\|\nabla_h u(\tau)\|_{L^2}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_h(\tau)\|^{\frac12}_{L^2_h L^2_{x_3}}\|\partial_3 u_h(\tau)\|^{\frac12}_{L^2_h L^2_{x_3}}\|\nabla_h u(\tau)\|_{L^2}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_h\|^{\frac12}_{L^2}\|\partial_3 u_h\|^{\frac12}_{L^2}\|\nabla_h u(\tau)\|_{L^2}\,d\tau.
\end{aligned}
$$

Invoking (5.1) through (5.4), we have,

$$
\begin{aligned}
K_{11} &\le C_0^2\,\varepsilon^2 \int_0^t (t-\tau)^{-\frac12}(1+\tau)^{-\frac{\sigma}{4}}\,(1+\tau)^{-\frac{\sigma}{4}}(1+\tau)^{-\frac{\sigma}{2}-\frac12}\,d\tau\\
&= C_0^2\,\varepsilon^2 \int_0^t (t-\tau)^{-\frac12}(1+\tau)^{-\sigma-\frac12}\,d\tau\\
&\le \begin{cases} C_0^2\,\varepsilon^2\,(1+t)^{-\sigma} & \text{if } \sigma < \frac12\\ C_0^2\,\varepsilon^2\,(1+t)^{-\frac12} & \text{if } \sigma > \frac12\\ C_0^2\,\varepsilon^2\,(1+t)^{-\frac12}\ln(1+t) & \text{if } \sigma = \frac12. \end{cases}
\end{aligned}
$$

Therefore, for $\frac12 < \sigma < 1$,

$$K_{11} \le C_0^2\,\varepsilon^2\,(1+t)^{-\frac{\sigma}{2}}.$$

If ε is taken to be sufficiently small such that

$$C_0\,\varepsilon \le \frac{1}{128}, \tag{5.14}$$

then

$$K_{11} \le \frac{C_0}{128}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

K_{12} contains the good derivative ∇_h and admits the same upper bound as the one for K_{11}. We now turn to K_{13}.

$$
\begin{aligned}
K_{13} &\le \int_0^t \Big\| \|e^{\Delta_h(t-\tau)}u_3\partial_3 u_h(\tau)\|_{L_h^2}\Big\|_{L_{x_3}^2}\,d\tau\\
&\le \int_0^t \|(t-\tau)^{-\frac12}\|u_3\partial_3 u_h(\tau)\|_{L_h^1}\|_{L_{x_3}^2}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|\|u_3(\tau)\|_{L_h^2}\|\partial_3 u_h(\tau)\|_{L_h^2}\|_{L_{x_3}^2}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_3(\tau)\|_{L_h^2L_{x_3}^2}^{\frac12}\,\|\partial_3 u_3(\tau)\|_{L_h^2L_{x_3}^2}^{\frac12}\,\|\partial_3 u_h(\tau)\|_{L^2}\,d\tau\\
&\le \int_0^t (t-\tau)^{-\frac12}\|u_3(\tau)\|_{L^2}^{\frac12}\|\nabla_h\cdot u_h(\tau)\|_{L^2}^{\frac12}\|\partial_3 u_h(\tau)\|_{L^2}\,d\tau
\end{aligned}
$$

Invoking (5.1) through (5.4) yields

$$
\begin{aligned}
&\int_0^t \|e^{\Delta_h(t-\tau)}u_3\partial_3 u_h(\tau)\|_{L^2}\,d\tau\\
&\le C_0^2\,\varepsilon^2 \int_0^t (t-\tau)^{-\frac12}(1+\tau)^{-\frac{\sigma}{4}}\,(1+\tau)^{-\frac{\sigma}{4}-\frac14}(1+\tau)^{-\frac{\sigma}{2}}\,d\tau\\
&= C_0^2\,\varepsilon^2 \int_0^t (t-\tau)^{-\frac12}(1+\tau)^{-\sigma-\frac14}\,d\tau\\
&\le C_0^2\,\varepsilon^2\,(1+t)^{-\frac{\sigma}{2}}
\end{aligned}
$$

for any $\frac12 \le \sigma < 1$. If ε is taken to satisfy (5.14), then

$$\int_0^t \|e^{\Delta_h(t-\tau)}u_3\partial_3 u_h(\tau)\|_{L^2}\,d\tau \le \frac{C_0}{128}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

Combining the upper bounds, we obtain

$$\|K_1\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

We now bound $\|K_2\|_{L^2}$. We take a quick inspection of the integrand in K_2. In order to bound

$$\frac{\xi_h\xi_3}{|\xi_h|^2}e^{\lambda_1(t-\tau)} + \frac{\xi_h\xi_3}{|\xi_h|^2}G_2(t-\tau) + \xi_h\xi_3 G_1(t-\tau)$$

suitably, we need to generate the factor ξ_h from $\mathbb{P}\widehat{(u\cdot\nabla u)}_3$. By the definition of $\mathbb{P}$,

$$\begin{aligned}
\mathbb{P}(u\cdot\nabla u)_3 &= u\cdot\nabla u_3 - \partial_3\Delta^{-1}\nabla\cdot(u\cdot\nabla u)\\
&= \partial_1(u_1u_3) + \partial_2(u_2u_3) + \partial_3(u_3u_3)\\
&\quad - \partial_3\Delta^{-1}(\partial_1(u\cdot\nabla u_1) + \partial_2(u\cdot\nabla u_2) + \partial_3(u\cdot\nabla u_3))\\
&= \partial_1(u_1u_3) + \partial_2(u_2u_3) - \partial_3\Delta^{-1}(\partial_1\nabla\cdot(uu_1) + \partial_2\nabla\cdot(uu_2)\\
&\quad + \partial_3(u_3u_3)) - \partial_3\Delta^{-1}\partial_3\nabla\cdot(uu_3)\\
&= \partial_1(u_1u_3) + \partial_2(u_2u_3) - \partial_3\Delta^{-1}(\partial_1\nabla\cdot(uu_1) + \partial_2\nabla\cdot(uu_2))\\
&\quad - \partial_3\Delta^{-1}\partial_3\partial_1(u_1u_3) - \partial_3\Delta^{-1}\partial_3\partial_2(u_2u_3)\\
&\quad + \partial_3(u_3u_3) - \partial_3\Delta^{-1}\partial_3\partial_3(u_3u_3)\\
&= \partial_1(u_1u_3) + \partial_2(u_2u_3) - \partial_3\Delta^{-1}(\partial_1\nabla\cdot(uu_1) + \partial_2\nabla\cdot(uu_2))\\
&\quad - \partial_3\Delta^{-1}\partial_3\partial_1(u_1u_3) - \partial_3\Delta^{-1}\partial_3\partial_2(u_2u_3)\\
&\quad + \Delta^{-1}(\Delta\partial_3(u_3u_3) - \partial_3\partial_3\partial_3(u_3u_3))\\
&= \partial_1(u_1u_3) + \partial_2(u_2u_3) - \partial_3\Delta^{-1}(\partial_1\nabla\cdot(uu_1) + \partial_2\nabla\cdot(uu_2))\\
&\quad - \partial_3\Delta^{-1}\partial_3\partial_1(u_1u_3) - \partial_3\Delta^{-1}\partial_3\partial_2(u_2u_3)\\
&\quad + \Delta^{-1}\Delta_h\partial_3(u_3u_3),
\end{aligned} \tag{5.15}$$

where we have combined two terms to generate the desirable factor

$$\Delta\partial_3(u_3u_3) - \partial_3\partial_3\partial_3(u_3u_3) = \Delta_h\partial_3(u_3u_3).$$

It is clear that each term contains ∂_1 or ∂_2. That is, its Fourier transform has the desired factor $|\xi_h|$. Therefore, by the upper bounds for G_1 and G_2 in Lemma 3.1

$$|G_1| \le t\,e^{-|\xi_h|^2 t}, \quad |\xi_h|^2|G_1| \le C\,e^{-c_0|\xi_h|^2 t}, \quad |G_2| \le C\,e^{-c_0|\xi_h|^2 t},$$

we have

$$\begin{aligned}
&\left|\left(\frac{1}{|\xi_h|}e^{\lambda_1(t-\tau)} + \frac{1}{|\xi_h|}G_2(t-\tau) + \xi_h G_1(t-\tau)\right)\xi_3\mathbb{P}\widehat{(u\cdot\nabla u)}_3\right|\\
&\le C\,e^{-c_0|\xi_h|^2(t-\tau)}\Big(|\xi_3\widehat{u_1u_3}| + |\xi_3\widehat{u_2u_3}| + |\xi_3|^2|\xi|^{-2}|\widehat{\nabla\cdot(uu_1)}|\\
&\quad + |\xi_3|^2|\xi|^{-2}|\widehat{\nabla\cdot(uu_2)}| + |\xi_3|^2|\xi|^{-2}|\widehat{\partial_3(u_1u_3)}|\\
&\quad + |\xi_3|^2|\xi|^{-2}|\widehat{\partial_3(u_2u_3)}| + |\xi_3|^2|\xi|^{-2}|\widehat{\nabla_h\cdot(u_3u_3)}|\Big)\\
&\le C\,e^{-c_0|\xi_h|^2(t-\tau)}\left(|\widehat{\partial_3(u_hu_3)}| + |\widehat{\nabla\cdot(uu_h)}| + |\widehat{\nabla_h\cdot(u_3u_3)}|\right).
\end{aligned} \tag{5.16}$$

Therefore,

$$\begin{aligned}\|K_2\|_{L^2} \le C \int_0^t \|e^{c_0\Delta_h(t-\tau)}\partial_3(u_h u_3)(\tau)\|_{L^2}\, d\tau \\ + C\int_0^t \|e^{c_0\Delta_h(t-\tau)}\nabla_h\cdot(uu_h)(\tau)\|_{L^2}\, d\tau \\ + C\int_0^t \|e^{c_0\Delta_h(t-\tau)}\nabla_h\cdot(u_3u_3)\|_{L^2}\, d\tau.\end{aligned}$$

These terms are pretty much like the terms in K_1. Thus similar estimates lead to the same upper bound

$$\|K_2\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

To bound K_3, we first bound the kernel. By the definition of G_1,

$$\left|\frac{\xi_h\xi_3}{|\xi|^2}G_1\right| = \left|\frac{\xi_h\xi_3}{|\xi|^2}\, e^{-|\xi_h|^2 t}\,\frac{\sin\frac{|\xi_h|}{|\xi|}t}{\frac{|\xi_h|}{|\xi|}}\right| \le e^{-|\xi_h|^2 t}. \tag{5.17}$$

As in the estimate of $\|K_1\|_{L^2}$,

$$\begin{aligned}\|K_3\|_{L^2} \le C\int_0^t \|e^{c_0\Delta_h(t-\tau)}u_h\cdot\nabla_h\theta(\tau)\|_{L^2}\, d\tau \\ + \int_0^t \|e^{c_0\Delta_h(t-\tau)}u_3\partial_3\theta(\tau)\|_{L^2}\, d\tau.\end{aligned}$$

The two terms on the right-hand side can be bounded as K_{11} and K_{13} above. Thus,

$$\|K_3\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

To bound K_4, we use Lemma 3.1 and the definition of G_1,

$$\left|(G_2 + |\xi_h|^2 G_1)(t-\tau)\right| \le C\, e^{-c_0|\xi_h|^2(t-\tau)}$$

for two constants $C > 0$ and $c_0 > 0$. We also invoke (5.15). Then the estimate of $\|K_4\|_{L^2}$ can be proceeded as in $\|K_2\|_{L^2}$ and

$$\|K_4\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

K_5 behaves like K_3. Since

$$\left|\frac{|\xi_h|^2}{|\xi|^2}G_1(t-\tau)\right| \le C\, e^{-c_0|\xi_h|^2(t-\tau)},$$

$\|K_5\|_{L^2}$ obeys the same bound as $\|K_3\|_{L^2}$,

$$\|K_5\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

The estimate of K_6 is pretty much like the last term in K_2, and thus

$$\|K_6\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

K_7 behaves like K_3 and has the same bound

$$\|K_7\|_{L^2} \le \frac{C_0}{32}\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

Combining the upper bounds for the linear parts and the upper bounds for K_1 through K_5, we obtain

$$\|u(t)\|_{L^2} \le C\,\varepsilon(1+t)^{-\frac{\sigma}{2}} + \frac{1}{4}C_0\,\varepsilon(1+t)^{-\frac{\sigma}{2}} \le \frac{1}{2}C_0\,\varepsilon(1+t)^{-\frac{\sigma}{2}}$$

if we choose $C_0 \ge 4C$. Similarly, by the upper bounds on K_6 and K_7, we obtain

$$\|\theta(t)\|_{L^2} \le \frac{1}{2}C_0\,\varepsilon(1+t)^{-\frac{\sigma}{2}}.$$

5.3 Estimates of $\|\partial_3 u\|_{L^2}$ and $\|\partial_3\theta\|_{L^2}$ and verification of (5.7)

This subsection provides upper bounds for $\|\partial_3 u\|_{L^2}$ and $\|\partial_3\theta\|_{L^2}$, which allow us to complete the verification of (5.7). We again make use of the integral representation (2.1), (2.2) and (2.3). We apply ∂_3 to (2.1), (2.2) and (2.3) and then take the L^2-norm. The linear parts have been estimated in the proof of Proposition 1.1, so we focus on the bounds for $\partial_3 K_1$ through $\partial_3 K_7$ with K_1 through K_7 defined in the previous subsection.

We start with $\|\partial_3 K_1\|_{L^2}$. As in (5.13), we can write

$$\begin{aligned}
\partial_3\mathbb{P}(u\cdot\nabla u))_h &= \partial_3 u\cdot\nabla u_h + u\cdot\partial_3\nabla u_h - \partial_3\nabla_h\Delta^{-1}\nabla\cdot(u\cdot\nabla u)\\
&= \partial_3 u_h\cdot\nabla_h u_h + \partial_3 u_3\partial_3 u_h + u_h\cdot\nabla_h\partial_3 u_h\\
&\quad + u_3\partial_{33}u_h - \partial_3\Delta^{-1}\nabla\cdot\nabla_h(u\cdot\nabla u)\\
&= u_3\partial_{33}u_h + \partial_3 u_h\cdot\nabla_h u_h - \nabla_h\cdot u_h\partial_3 u_h + \nabla_h\cdot(u_h\partial_3 u_h)\\
&\quad - \nabla_h\cdot u_h\partial_3 u_h - \partial_3\Delta^{-1}\nabla\cdot\nabla_h(u\cdot\nabla u)\\
&= u_3\partial_{33}u_h + (\partial_3 u_h\cdot\nabla_h u_h - 2\nabla_h\cdot u_h\partial_3 u_h)\\
&\quad + \nabla_h\cdot(u_h\partial_3 u_h) - \partial_3\Delta^{-1}\nabla\cdot\nabla_h(u\cdot\nabla u).
\end{aligned}$$

Correspondingly $\partial_3 K_1$ is then divided into four terms,

$$\partial_3 K_1 = L_{11} + L_{12} + L_{13} + L_{14}, \tag{5.18}$$

where

$$\begin{aligned}
L_{11} &= \int_0^t e^{\Delta_h(t-\tau)}\,u_3\partial_{33}u_h(\tau)\,d\tau,\\
L_{12} &= \int_0^t e^{\Delta_h(t-\tau)}\,(\partial_3 u_h\cdot\nabla_h u_h - 2\nabla_h\cdot u_h\partial_3 u_h)(\tau)\,d\tau,\\
L_{13} &= \int_0^t e^{\Delta_h(t-\tau)}\,\nabla_h\cdot(u_h\partial_3 u_h)(\tau)\,d\tau,\\
L_{14} &= -\int_0^t e^{\Delta_h(t-\tau)}\,\partial_3\Delta^{-1}\nabla\cdot\nabla_h(u\cdot\nabla u)(\tau)\,d\tau.
\end{aligned}$$

L_{11} involves the unfavorable derivative ∂_{33} and may yield the worst decay rate. By Lemmas 5.1, 5.2 and 5.3,

$$
\begin{aligned}
\|L_{11}\|_{L^2} &\le \int_0^t \|e^{\Delta_h(t-\tau)} u_3\, \partial_{33} u_h(\tau)\|_{L^2}\, d\tau \\
&\le \int_0^t \Big\| \|e^{\Delta_h(t-\tau)} u_3\, \partial_{33} u_h(\tau)\|_{L^2_h} \Big\|_{L^2_{x_3}}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \Big\| \|u_3\, \partial_{33} u_h(\tau)\|_{L^1_h} \Big\|_{L^2_{x_3}}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \Big\| \|u_3\|_{L^2_h} \|\partial_{33} u_h\|_{L^2_h} \Big\|_{L^2_{x_3}}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \|u_3\|_{L^2_h L^\infty_{x_3}} \|\partial_{33} u_h\|_{L^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \|u_3\|_{L^2}^{\frac12} \|\partial_3 u_3\|_{L^2}^{\frac12} \|\partial_{33} u_h\|_{L^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \|u_3\|_{L^2}^{\frac12} \|\nabla_h \cdot u_h\|_{L^2}^{\frac12} \|\partial_3 u_h\|_{L^2}^{\frac23} \|\partial_3^4 u_h\|_{L^2}^{\frac13}\, d\tau,
\end{aligned}
$$

where we have used the interpolation inequality,

$$
\|\partial_{33} u_h\|_{L^2} \le \|\partial_3 u_h\|_{L^2}^{\frac23} \|\partial_3^4 u_h\|_{L^2}^{\frac13}.
$$

We now invoke the ansatz in (5.3) and (5.4) to obtain, for $\frac34 \le \sigma < 1$,

$$
\begin{aligned}
&\int_0^t \|e^{\Delta_h(t-\tau)} u_3\, \partial_{33} u_h(\tau)\|_{L^2}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2 \int_0^t (t-\tau)^{-\frac12} (1+\tau)^{-\frac{\sigma}{4}} (1+\tau)^{-\frac14-\frac{\sigma}{4}} (1+\tau)^{-\frac{\sigma}{3}}\, d\tau \\
&= C\, C_0^2\, \varepsilon^2 \int_0^t (t-\tau)^{-\frac12} (1+\tau)^{-\frac56\sigma-\frac14}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2\, (1+t)^{-\frac56\sigma+\frac14} \\
&\le C\, C_0^2\, \varepsilon^2\, (1+t)^{-\frac{\sigma}{2}} \\
&\le \frac{C_0}{128} \varepsilon\, (1+t)^{-\frac{\sigma}{2}},
\end{aligned}
$$

where $\sigma \ge \frac34$ is used in the last inequality to ensure that

$$
-\frac56\sigma + \frac14 \le -\frac{\sigma}{2}.
$$

This is exactly where we need the constraints on σ. L_{12} and L_{13} can be dealt with similarly.

$$
\begin{aligned}
\|L_{12}\|_{L^2} &\le C \int_0^t \Big\| \|e^{\Delta_h(t-\tau)} \partial_3 u_h \cdot \nabla_h u_h(\tau)\|_{L^2_h} \Big\|_{L^2_{x_3}}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \Big\| \|\partial_3 u_h \cdot \nabla_h u_h(\tau)\|_{L^1_h} \Big\|_{L^2_{x_3}}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac12} \|\partial_3 u_h\|_{L^2}^{\frac12} \|\partial_{33} u_h\|_{L^2}^{\frac12} \|\nabla_h u_h(\tau)\|_{L^2}\, d\tau
\end{aligned}
$$

$$
\begin{aligned}
&\le C \int_0^t (t-\tau)^{-\frac{1}{2}} \|\partial_3 u_h\|_{L^2}^{\frac{5}{6}} \|\partial_3^4 u_h\|_{L^2}^{\frac{1}{6}} \|\nabla_h u_h(\tau)\|_{L^2}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2 \int_0^t (t-\tau)^{-\frac{1}{2}} (1+\tau)^{-\frac{5}{12}\sigma} (1+\tau)^{-\frac{1}{2}-\frac{\sigma}{2}}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2 \int_0^t (t-\tau)^{-\frac{1}{2}} (1+\tau)^{-\frac{11}{12}\sigma-\frac{1}{2}}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2\, (1+t)^{-\frac{11}{12}\sigma} \\
&\le \frac{C_0}{128} \varepsilon\, (1+t)^{-\frac{\sigma}{2}}.
\end{aligned}
$$

$$
\begin{aligned}
\|L_{13}\|_{L^2} &\le C \int_0^t \Big\| \|e^{\Delta_h(t-\tau)} \nabla_h \cdot (u_h \partial_3 u_h)(\tau)\|_{L_h^2} \Big\|_{L_{x_3}^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac{3}{4}} \Big\| \|u_h \partial_3 u_h\|_{L_h^{4/3}} \Big\|_{L_{x_3}^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac{3}{4}} \Big\| \|u_h\|_{L_h^4} \|\partial_3 u_h\|_{L_h^2} \Big\|_{L_{x_3}^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac{3}{4}} \|u_h\|_{L_h^4 L_{x_3}^\infty} \|\partial_3 u_h\|_{L^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac{3}{4}} \|u_h\|_{L_h^2 L_{x_3}^\infty}^{\frac{1}{2}} \|\nabla_h u_h\|_{L_h^2 L_{x_3}^\infty}^{\frac{1}{2}} \|\partial_3 u_h\|_{L^2}\, d\tau \\
&\le C \int_0^t (t-\tau)^{-\frac{3}{4}} \|u\|_{L^2}^{\frac{1}{4}} \|\partial_3 u\|_{L^2}^{\frac{1}{4}} \|\nabla_h u_h\|_{L^2}^{\frac{1}{4}} \|\partial_3 \nabla_h u_h\|_{L^2}^{\frac{1}{4}} \|\partial_3 u_h\|_{L^2}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2 \int_0^t (t-\tau)^{-\frac{3}{4}} (1+\tau)^{-\frac{7}{8}\sigma-\frac{1}{8}}\, d\tau \\
&\le C\, C_0^2\, \varepsilon^2 (1+t)^{-\frac{7}{8}\sigma+\frac{1}{8}} \\
&\le \frac{C_0}{128} \varepsilon\, (1+t)^{-\frac{\sigma}{2}}.
\end{aligned}
$$

The estimate of L_{14} in L^2 is pretty much the same as those in the first three terms, so we just briefly sketch it. By the boundedness of the Riesz transform

$$
\begin{aligned}
&\|\partial_3 \Delta^{-1} \nabla \cdot f\|_{L^2} \le C\, \|f\|_{L^2}, \\
&\|L_{14}\|_{L^2} \le C \int_0^t \Big\| e^{\Delta_h(t-\tau)} \nabla_h (u \cdot \nabla u) \Big\|_{L^2}\, d\tau \le \frac{C_0}{128} \varepsilon\, (1+t)^{-\frac{\sigma}{2}}.
\end{aligned}
$$

Therefore,

$$
\|\partial_3 K_1\|_{L^2} \le \frac{C_0}{32} \varepsilon\, (1+t)^{-\frac{\sigma}{2}}.
$$

To bound $\|\partial_3 K_2\|_{L^2}$, we obtain as in (5.16) the following upper bound

$$
\begin{aligned}
&\left| \left(\left(\frac{\xi_h \xi_3}{|\xi_h|^2} e^{\lambda_1(t-\tau)} + \frac{\xi_h \xi_3}{|\xi_h|^2} G_2 + \xi_h \xi_3 G_1 \right) (t-\tau) \right) \partial_3 \widehat{\mathbb{P}(u \cdot \nabla u)_3} \right| \\
&\quad \le C\, e^{-c_0 |\xi_h|^2 (t-\tau)} \left(|\partial_{33} \widehat{(u_h u_3)}| + |\partial_3 \widehat{\nabla \cdot (u u_h)}| + |\partial_3 \widehat{\nabla_h \cdot (u_3 u_3)}| \right).
\end{aligned}
$$

Therefore,

$$\begin{aligned}
&\|\partial_3 K_2\|_{L^2} \\
&\quad \le C \int_0^t \left\| e^{c_0\Delta_h(t-\tau)} \left(|\partial_{33}(u_h u_3)| + |\partial_3 \nabla\cdot(u u_h)| + |\partial_3 \nabla_h\cdot(u_3 u_3)|\right)\right\|_{L^2} d\tau.
\end{aligned}$$

The three terms on the right-hand side are similar to those terms in (5.18) and admit the same bound as the one for $\|\partial_3 K_1\|_{L^2}$,

$$\|\partial_3 K_2\|_{L^2} \le \frac{C_0}{32}\varepsilon\,(1+t)^{-\frac{\sigma}{2}}.$$

The estimates of $\partial_3 K_3$ through $\partial_3 K_7$ are very similar and thus omitted. Combining all the upper bounds, we obtain

$$\|\partial_3 u(t)\|_{L^2},\ \|\partial_3\theta(t)\|_{L^2} \le \frac{1}{2}C_0\,\varepsilon\,(1+t)^{-\frac{\sigma}{2}},$$

which verifies (5.7).

5.4 Estimates of $\|\nabla_h u\|_{L^2}$ and $\|\nabla_h\theta\|_{L^2}$ and verification of (5.8)

This subsection proves (5.8). We again make use of the integral representation (2.1), (2.2) and (2.3). We apply ∇_h to (2.1), (2.2) and (2.3) and then take the L^2-norm. The linear parts have been estimated in the proof of Proposition 1.1, so we focus on the bounds for $\nabla_h K_1$ through $\nabla_h K_7$ with K_1 through K_7 defined in the Sect. 5.2.

We start with $\nabla_h K_1$. As in (5.13), we write

$$\mathbb{P}(u\cdot\nabla u))_h = u_3\partial_3 u_h + u_h\cdot\nabla_h u_h - \Delta^{-1}\nabla\cdot\nabla\cdot\nabla_h(u\otimes u).$$

and $\|\nabla_h K_1\|_{L^2}$ is then bounded

$$M_i := \|\nabla_h K_i\|_{L^2} \le M_{11} + M_{12} + M_{13}, \quad \text{for } i = 1,\cdots,7,$$

where

$$\begin{aligned}
M_{11} &= \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)}(u_3\partial_3 u_h)(\tau)\|_{L^2}\,d\tau,\\
M_{12} &= \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)}(u_h\cdot\nabla_h u_h)(\tau)\|_{L^2}\,d\tau,\\
M_{13} &= \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)}(\Delta^{-1}\nabla\cdot\nabla\cdot\nabla_h(u\otimes u))(\tau)\|_{L^2}\,d\tau.
\end{aligned}$$

By Lemma 5.3,

$$\begin{aligned}
M_{11} &\le \int_0^t \left\| \|\nabla_h e^{\Delta_h(t-\tau)}(u_3\partial_3 u_h)(\tau)\|_{L^2_h}\right\|_{L^2_{x_3}} d\tau\\
&\le C\int_0^t (t-\tau)^{-\frac{1}{2}}\, \|\|e^{\Delta_h(t-\tau)}(u_3\partial_3 u_h)(\tau)\|_{L^2_h}\|_{L^2_{x_3}}\,d\tau.
\end{aligned}$$

We remark that we can no longer proceed as in the estimate of K_{13}. We already have the factor $(t-\tau)^{-\frac{1}{2}}$ and an estimate as in that of K_{13} would generate another $(t-\tau)^{-\frac{1}{2}}$ and thus produce $(t-\tau)^{-1}$, which is not integrable on $(0,t)$. Instead we use a different estimate. We choose q satisfying

$$\frac{1}{q}=\frac{1}{2}+\frac{\sigma}{2} \quad \text{or} \quad q=\frac{2}{1+\sigma}.$$

For $\frac{3}{4}\le\sigma<1$, we have $1<q<2$. Then, by Lemma 5.3,

$$M_{11}\le C\int_0^t (t-\tau)^{-\frac{1+\sigma}{2}}\left\|\|u_3\partial_3 u_h(\tau)\|_{L^q_h}\right\|_{L^2_{x_3}}\,d\tau$$

The integrand can be further bounded as in Sect. 5.1,

$$\begin{aligned}
\left\|\|u_3\partial_3 u_h\|_{L^q_h}\right\|_{L^2_{x_3}} &\le \left\|\|u_3\|_{L^\infty_{x_3}}\|\partial_3 u_h\|_{L^2_{x_3}}\right\|_{L^q_h}\\
&\le \|u_3\|_{L_h^{\frac{2}{\sigma}}L^\infty_{x_3}}\|\partial_3 u_h\|_{L^2_hL^2_{x_3}}\\
&\le C\left\|\|u_3\|^{\frac12}_{L^2_{x_3}}\;\|\partial_3 u_3\|^{\frac12}_{L^2_{x_3}}\right\|_{L_h^{\frac{2}{\sigma}}}\|\partial_3 u_h\|_{L^2}\\
&\le C\left\|\|u_3\|^{\frac12}_{L^2_{x_3}}\right\|_{L_h^{\frac{4}{2\sigma-1}}}\left\|\|\partial_3 u_3\|^{\frac12}_{L^2_{x_3}}\right\|_{L^4_h}\|\partial_3 u_h\|_{L^2}\\
&\le C\,\|\partial_3 u_3\|^{\frac12}_{L^2}\,\|u_3\|^{\frac12}_{L_h^{\frac{2}{2\sigma-1}}L^2_{x_3}}\;\|\partial_3 u_h\|_{L^2}\\
&\le C\,\|\nabla_h\cdot u_h\|^{\frac12}_{L^2}\,\|u_3\|^{\frac12}_{L^2_{x_3}L_h^{\frac{2}{2\sigma-1}}}\;\|\partial_3 u_h\|_{L^2}\\
&\le C\,\|\nabla_h\cdot u_h\|^{\frac12}_{L^2}\,\|u_3\|^{\sigma-\frac12}_{L^2}\,\|\nabla_h u_3\|^{1-\sigma}_{L^2}\,\|\partial_3 u_h\|_{L^2}.
\end{aligned}$$

Therefore, for any $\frac{3}{4}\le\sigma<1$,

$$\begin{aligned}
M_{11} &\le C\int_0^t (t-\tau)^{-\frac{1+\sigma}{2}}\,\|\nabla_h\cdot u_h\|^{\frac12}_{L^2}\,\|u_3\|^{\sigma-\frac12}_{L^2}\,\|\nabla_h u_3\|^{1-\sigma}_{L^2}\,\|\partial_3 u\|_{L^2}\,d\tau\\
&\le C\,C_0^2\,\varepsilon^2\int_0^t (t-\tau)^{-\frac{1+\sigma}{2}}\,(1+\tau)^{-(\frac32-\sigma)(\frac12+\frac{\sigma}{2})}\,(1+\tau)^{-\frac{\sigma}{2}(\sigma+\frac12)}\,d\tau\\
&= C\,C_0^2\,\varepsilon^2\int_0^t (t-\tau)^{-\frac{1+\sigma}{2}}\,(1+\tau)^{-\frac34-\frac{\sigma}{2}}\,d\tau\\
&\le C\,C_0^2\,\varepsilon^2\,(1+t)^{-\frac{1+\sigma}{2}}\\
&\le \frac{1}{128}C_0\,\varepsilon\,(1+t)^{-\frac{1+\sigma}{2}}.
\end{aligned}$$

M_{12} can be bounded similarly and they admit the same upper bound. For M_{13}, we first bound it by the fact that Riesz transforms are bounded on L^2,

$$M_{13}\le C\int_0^t \|\nabla_h e^{\Delta_h(t-\tau)}\nabla_h(u\otimes u))(\tau)\|_{L^2}\,d\tau$$

and then proceed as in the estimates of M_{11} to obtain the same upper bound. Therefore,

$$\|\nabla_h M_1(t)\|_{L^2}\le\frac{C_0}{4}\,\varepsilon\,(1+t)^{-\frac{1+\sigma}{2}}.$$

According to (5.16),

$$\begin{aligned}\|\nabla_h M_2(t)\|_{L^2} \le C \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)} \partial_3(u_h u_3)\|_{L^2}\, d\tau \\ + C \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)} \nabla \cdot (u u_h)\|_{L^2}\, d\tau \\ + C \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)} \nabla_h \cdot (u_3 u_3)\|_{L^2}\, d\tau.\end{aligned}$$

These three terms can be estimated as those in $\nabla_h M_1$ and obey the same upper bound. By (5.17),

$$\begin{aligned}\|\nabla_h M_3(t)\|_{L^2} \le C \int_0^t \|\nabla_h e^{\Delta_h(t-\tau)} u_h \cdot \nabla_h \theta(\tau)\|_{L^2}\, d\tau \\ + C \int_0^t \|\nabla_h e^{c_0\Delta_h(t-\tau)} u_3 \partial_3 \theta(\tau)\|_{L^2}\, d\tau.\end{aligned}$$

The terms in $\nabla_h M_3$ can also be bounded similarly as those in $\nabla_h M_1$. The terms in $\nabla_h M_4$ through $\nabla_h M_7$ can also be bounded similarly and the details are omitted. As a consequence, we have verified (5.8). This completes the proof of Theorem 1.3. □

Finally we make some concluding remarks. We have proposed and implemented a new and effective approach to extracting the optimal decay estimates for the 3D Boussinesq system with only horizontal dissipation. It is not difficult to see that this approach also works for the Boussinesq system with full dissipation. When the full dissipation is present, the estimates of many terms (especially those with vertical derivatives) can be significantly simplified. In addition, although the approach is developed in this paper for the Boussinesq system, it is expected to be applicable to many other anisotropic PDE systems such as the magneto-hydrodynamic equations with horizontal dissipation. We are also working on anisotropic Boussinesq systems without horizontal velocity or thermal dissipation. There are many challenges. One particular difficult case is when there is no horizontal velocity dissipation in the Boussinesq system. Then the dissipation is only in the vertical direction and it is not clear in $\mathbb{R}^3$ if one can control the velocity nonlinearity by dissipation in only one direction in the Sobolev setting. We are hopeful that some progress will be made on this front in the near future.

Acknowledgements Ji is supported by the National Natural Science Foundation of China (NNSFC) under grant number 12001065. Yan is supported by the Opening Fund of Geomathematics Key Laboratory of Sichuan Province under grant number scsxdz2021yb07. Wu is partially supported by the National Science Foundation of the United States under Grant DMS 2104682 and the AT&T Foundation at Oklahoma State University. The authors are grateful to the referee for careful reading of the paper and valuable suggestions and comments.

References

1. Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren der mathematischen Wissenschaften, vol. 343. Springer, Berlin (2011)
2. Ben Said, O., Pandey, U., Wu, J.: The stabilizing effect of the temperature on buoyancy-driven fluids (2020). arXiv:2005.11661v2 [math.AP]
3. Bian, D., Dai, S., Mao, J.: Stability of Couette flow for 2D Boussinesq system in a uniform magnetic field (2020). arXiv:2012.11875 [math.AP]
4. Bian, D., Pu, X.: Stability threshold for 2D shear flows of the Boussinesq system near Couette (2020). arXiv:2012.02386 [math.AP]

5. Bianchini, R., Coti Zelati, M., Dolce, M.: Linear inviscid damping for shear flows near Couette in the 2D stably stratified regime. arXiv:2005.09058 [math.AP]
6. Brandolese, L., Schonbek, M.E.: Large time decay and growth for solutions of a viscous Boussinesq system. Trans. Am. Math. Soc. **364**, 5057–5090 (2012)
7. Cao, C., Wu, J.: Global regularity for the 2D MHD equations with mixed partial dissipation and magnetic diffusion. Adv. Math. **226**, 1803–1822 (2011)
8. Constantin, P., Doering, C.: Heat transfer in convective turbulence. Nonlinearity **9**, 1049–1060 (1996)
9. Castro, A., Córdoba, D., Lear, D.: On the asymptotic stability of stratified solutions for the 2D Boussinesq equations with a velocity damping term. Math. Models Methods Appl. Sci. **29**, 1227–1277 (2019)
10. Deng, W., Wu, J., Zhang, P.: Stability of Couette flow for 2D Boussinesq system with vertical dissipation (2020). arXiv:2004.09292v1 [math.AP]
11. Doering, C., Gibbon, J.: Applied analysis of the Navier–Stokes equations. Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge (1995)
12. Doering, C.R., Wu, J., Zhao, K., Zheng, X.: Long time behavior of the two-dimensional Boussinesq equations without buoyancy diffusion. Phys. D Nonlinear Phenom. **376**(377), 144–159 (2018)
13. Dong, B., Wu, J., Xu, X., Zhu, N.: Stability and exponential decay for the 2D anisotropic Boussinesq equations with horizontal dissipation. Calc. Var. Partial Differ. Equ. **60**, 116–136 (2021)
14. Dreyfuss, P., Houamed, H.: Uniqueness result for the 3-D Navier–Stokes–Boussinesq equations with horizontal dissipation. J. Math. Fluid Mech. **23**(1), 1–24 (2021)
15. Lai, S., Wu, J., Zhong, Y.: Stability and large-time behavior of the 2D Boussinesq equations with partial dissipation. J. Differ. Equ. **271**, 764–796 (2021)
16. Lai, S., Wu, J., Xu, X., Zhang, J., Zhong, Y.: Optimal decay estimates for the 2D Boussinesq equations with partial dissipation. J. Nonlinear Sci. **33**, 1–33 (2021)
17. Holton, J.R.: An Introduction to Dynamic Meteorology. Academic Press, New York (1972)
18. Ji, R., Wu, J., Yang, W.: Stability and optimal decay for the 3D Navier–Stokes equations with horizontal dissipation. J. Differ. Equ. **290**, 57–77 (2021)
19. Lieb, E., Loss, M.: Analysis, Graduate Studies in Mathematics, vol. 14, 2nd edn. American Mathematical Society, Providence (2001)
20. Majda, A.: Introduction to PDEs and Waves for the Atmosphere and Ocean. Courant Lecture Notes, vol. 9. Courant Institute of Mathematical Sciences and American Mathematical Society (2003)
21. Majda, A., Bertozzi, A.: Vorticity and Incompressible Flow. Cambridge University Press, Cambridge (2002)
22. Masmoudi, N., Said-Houari, B., Zhao, W.: Stability of Couette flow for 2D Boussinesq system without thermal diffusivity (2020). arXiv:2010.01612 [math.AP]
23. Pater, I., Lissauer, J.: Planetary Sciences. Cambridge University Press, Cambridge (2015)
24. Pedlosky, J.: Geophysical Fluid Dynamics, 2nd edn. Springer, Berlin (1987)
25. Schonbek, M.: L^2 decay for weak solutions of the Navier–Stokes equations. Arch. Ration. Mech. Anal. **88**, 209–222 (1985)
26. Schonbek, M., Schonbek, T.: Moments and lower bounds in the far-field of solutions to quasi-geostrophic flows. Discret. Contin. Dyn. Syst. **13**, 1277–1304 (2005)
27. Schonbek, M., Wiegner, M.: On the decay of higher-order norms of the solutions of Navier–Stokes equations. Proc. R. Soc. Edinb. Sect. A **126**, 677–685 (1996)
28. Shang, H., Xu, L.: Stability near hydrostatic equilibrium to the three-dimensional Boussinesq equations with partial dissipation. Z. Angew. Math. Phys. **72**(2), 1–14 (2021)
29. Tao, L., Wu, J.: The 2D Boussinesq equations with vertical dissipation and linear stability of shear flows. J. Differ. Equ. **267**, 1731–1747 (2019)
30. Tao, L., Wu, J., Zhao, K., Zheng, X.: Stability near hydrostatic equilibrium to the 2D Boussinesq equations without thermal diffusion. Arch. Ration. Mech. Anal. **237**, 585–630 (2020)
31. Tao, T.: Nonlinear Dispersive Equations: Local and Global Analysis. CBMS Regional Conference Series in Mathematics (2006)
32. Wang, W.: On the global regularity for a 3D Boussinesq model without thermal diffusion. Z. Angew. Math. Phys. **70**(6), 1–6 (2019)
33. Weinberg, S.: Cosmology. Oxford University Press, New York (2008)
34. Wen, B., Dianati, N., Lunasin, E., Chini, G.P., Doering, C.R.: New upper bounds and reduced dynamical modeling for Rayleigh–Bénard convection in a fluid saturated porous layer. Commun. Nonlinear Sci. Numer. Simul. **17**, 2191–2199 (2012)
35. Wu, J.: Dissipative quasi-geostrophic equations with L^p data. Electron. J. Differ. Equ. **2001**, 1–13 (2001)
36. Wu, J., Zhu, Y.: Global solutions of 3D incompressible MHD system with mixed partial dissipation and magnetic diffusion near an equilibrium. Adv. Math. **377**, 107466 (2021)

37. Yang, J., Lin, Z.: Linear inviscid damping for Couette flow in stratified fluid. J. Math. Fluid Mech. **20**, 445–472 (2018)
38. Yang, W., Jiu, Q., Wu, J.: The 3D incompressible Navier–Stokes equations with partial hyperdissipation. Math. Nach. **292**, 1823–1836 (2019)
39. Ye, Z.: On global well-posedness for the 3D Boussinesq equations with fractional partial dissipation. Appl. Math. Lett. **90**, 1–7 (2019)
40. Zillinger, C.: On enhanced dissipation for the Boussinesq equations. arXiv: 2004.08125v1 [math.AP]
41. Zillinger, C.: On the Boussinesq equations with non-monotone temperature profiles. arXiv: 2011.02316v1 [math.AP]