STABILITY OF COUETTE FLOW FOR 2D BOUSSINESQ SYSTEM WITH VERTICAL DISSIPATION

WEN DENG[1], JIAHONG WU[2] AND PING ZHANG[3]

ABSTRACT. This paper establishes the nonlinear stability of the Couette flow for the 2D Boussinesq equations with only vertical dissipation. The Boussinesq equations concerned here model buoyancy-driven fluids such as atmospheric and oceanographic flows. Due to the presence of the buoyancy forcing, the energy of the standard Boussinesq equations could grow in time. It is the enhanced dissipation created by the linear non-self-adjoint operator $y\partial_x - \nu\partial_{yy}$ in the perturbation equation that makes the nonlinear stability possible. When the initial perturbation from the Couette flow $(y, 0)$ is no more than the viscosity to a suitable power (in the Sobolev space H^b with $b > \frac{4}{3}$), we prove that the solution of the 2D Boussnesq system with only vertical dissipation on $\mathbb{T} \times \mathbb{R}$ remains close to the Couette at the same order. A special consequence of this result is the stability of the Couette for the 2D Navier-Stokes equations with only vertical dissipation.

1. INTRODUCTION

The Boussinesq system reflects the basic physics laws obeyed by buoyancy-driven fluids. It is one of the most frequently used models for atmospheric and oceanographic flows and serves as the centerpiece in the study of the Rayleigh-Bénard convection (see, e.g., [11, 15, 20, 25]). The Boussinesq equations are mathematically significant. The 2D Boussinesq equations serve as a lower dimensional model of the 3D hydrodynamics equations. In fact, the 2D Boussinesq equations retain some key features of the 3D Euler and Navier-Stokes equations such as the vortex stretching mechanism. The inviscid 2D Boussinesq equations can be identified as the Euler equations for the 3D axisymmetric swirling flows [21]. Furthermore, the Boussinesq equations have some special characteristics of their own and offer many opportunities for new discoveries.

Due to their broad physical applications and mathematical significance, the Boussinesq equations have recently attracted considerable interests. Two fundamental problems, the global regularity problem and the stability problem, have been among the main driving forces in advancing the mathematical theory on the Boussinesq equations. Significant progress has been made on the global regularity of the 2D Boussinesq equations, especially those with only partial or fractional dissipation or no dissipation at all. Our attention here will be focused on the stability problem. The study of the stability problem on two physically important steady states has gained strong momentum. The first steady state is the hydrostatic equilibrium,

which is a prominent topic in fluid dynamics and astrophysics. Understanding this stability problem may help gain insight into some weather phenomena. Important progress has been made on the stability and large-time behavior ([9, 16, 28, 31]). The second steady state is the shear flow, which is the focus of this paper. The aim here is to fully understand the stability of perturbations near the Couette flow and their large-time behavior. Our consideration will cover both the Boussinesq equations with full dissipation and the Boussinesq equations with only vertical dissipation. Our emphasis is on the case when the dissipation is degenerate and only in the vertical direction.

The 2D Boussinesq system with full dissipation is given by

$$\begin{cases} \partial_t u + (u\partial_x + v\partial_y)u = -\partial_x p + \nu\Delta u, \\ \partial_t v + (u\partial_x + v\partial_y)v = -\partial_y p + \nu\Delta v + \theta, \\ \partial_x u + \partial_y v = 0, \\ \partial_t \theta + (u\partial_x + v\partial_y)\theta = \mu\Delta\theta, \end{cases} \tag{1.1}$$

where $\mathbf{u} = (u, v)$ denotes the 2D velocity field, p the pressure, θ the temperature, ν the viscosity and μ the thermal diffusivity. The first three equations in (1.1) are the incompressible Navier-Stokes equation with buoyancy forcing in the vertical direction. The last equation is a balance of the temperature convection and diffusion. The spatial domain Ω here is taken to be

$$\Omega = \mathbb{T} \times \mathbb{R}$$

with $\mathbb{T} = [0, 2\pi]$ being the periodic box and $\mathbb{R}$ being the whole line. In suitable physical regimes or under suitable scaling, the Boussinesq equations may involve only vertical dissipation ([22]), namely

$$\begin{cases} \partial_t u + (u\partial_x + v\partial_y)u = -\partial_x p + \nu\partial_{yy}u, \\ \partial_t v + (u\partial_x + v\partial_y)v = -\partial_y p + \nu\partial_{yy}v + \theta, \\ \partial_x u + \partial_y v = 0, \\ \partial_t \theta + (u\partial_x + v\partial_y)\theta = \mu\partial_{yy}\theta. \end{cases} \tag{1.2}$$

Cao and Wu previously examined the 2D Boussinesq system with only vertical dissipation and established its global regularity [10]. The Couette flow,

$$\mathbf{u}_{sh} = (y, 0), \quad p_{sh} = 0, \quad \theta_{sh} = 0,$$

is clearly a stationary solution of (1.1) and also of (1.2). Our goal is to understand the stability and large-time behavior of perturbations near the Couette flow. The perturbations

$$\widetilde{u} = u - y, \qquad \widetilde{v} = v, \qquad \widetilde{p} = p, \qquad \widetilde{\theta} = \theta,$$

satisfy, in the case of full dissipation,

$$\begin{cases} \partial_t \widetilde{u} + y\partial_x\widetilde{u} + \widetilde{v} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{u} = -\partial_x\widetilde{p} + \nu\Delta\widetilde{u}, \\ \partial_t \widetilde{v} + y\partial_x\widetilde{v} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{v} = -\partial_y\widetilde{p} + \nu\Delta\widetilde{v} + \widetilde{\theta}, \\ \partial_x\widetilde{u} + \partial_y\widetilde{v} = 0, \\ \partial_t \widetilde{\theta} + y\partial_x\widetilde{\theta} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{\theta} = \mu\Delta\widetilde{\theta}. \end{cases}$$

The corresponding perturbed vorticity near the steady vorticity $\omega_{sh} = -1$

$$\widetilde{\omega} = \partial_x \widetilde{v} - \partial_y \widetilde{u}$$

verifies, together with $\widetilde{\theta}$, the following system

$$\begin{cases} \partial_t \widetilde{\omega} + y\partial_x \widetilde{\omega} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{\omega} = \nu\Delta\widetilde{\omega} + \partial_x \widetilde{\theta}, \\ \partial_t \widetilde{\theta} + y\partial_x \widetilde{\theta} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{\theta} = \mu\Delta\widetilde{\theta}, \\ \widetilde{\mathbf{u}} = -\nabla^\perp(-\Delta)^{-1}\widetilde{\omega}. \end{cases} \qquad (1.3)$$

In the case when there is only vertical dissipation, the vorticity perturbation $\widetilde{\omega}$ and the temperature perturbation $\widetilde{\theta}$ satisfy

$$\begin{cases} \partial_t \widetilde{\omega} + y\partial_x \widetilde{\omega} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{\omega} = \nu\partial_{yy}\widetilde{\omega} + \partial_x \widetilde{\theta}, \\ \partial_t \widetilde{\theta} + y\partial_x \widetilde{\theta} + (\widetilde{\mathbf{u}} \cdot \nabla)\widetilde{\theta} = \mu\partial_{yy}\widetilde{\theta}, \\ \widetilde{\mathbf{u}} = -\nabla^\perp(-\Delta)^{-1}\widetilde{\omega}. \end{cases} \qquad (1.4)$$

The stability problem proposed for study here on (1.3) or (1.4) is not trivial. Due to the presence of the buoyancy forcing term, the Sobolev norms or even the L^2-norm of the velocity field could grow in time if the two linear terms $y\partial_x \widetilde{\omega}$ and $y\partial_x \widetilde{\theta}$ were not included in (1.3) or (1.4). In fact, Brandolese and Schonbek have shown in [8] 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). The stability of the Couette flow on (1.3) and (1.4) is only possible because of the enhanced dissipation generated by the non-self-adjoint operator $y\partial_x - \nu\partial_{yy}$, which is the linear part of the system (1.4). Even though the linear operator $y\partial_x - \nu\partial_{yy}$ involves only vertical dissipation, the non-commutativity between its real part and imaginary part actually creates smoothing effect in the horizontal direction, a phenomenon that is called the hypoellipticity. Operators of this type are investigated by Hörmander [17]. For the standard heat equation $\partial_t f = \nu\Delta f$, the dissipation time scale is $O(\nu^{-1})$ while, for the drift diffusion equations

$$\partial_t f + y\partial_x f = \nu\Delta f \quad \text{and} \quad \partial_t f + y\partial_x f = \nu\partial_{yy}f,$$

the dissipation time scale is $O(\nu^{-\frac{1}{3}})$, which is much faster than $O(\nu^{-1})$ for small ν. A more detailed explanation will be provided later. This enhanced dissipation effect plays an extremely important role in the stability problem studied here.

The phenomenon of enhanced dissipation has been widely observed and studied in physics literature (see, e.g., [7, 18, 30, 26]). It has recently attracted enormous attention from the mathematics community and significant progress has been made. One of the earliest rigorous results on the enhanced dissipation is obtained by Constantin, Kiselev, Ryzhik and Zlatos on the enhancement of diffusive mixing [12]. Many remarkable results have since been established. In particular, the stability of the shear flows to passive scale equations and to the Navier-Stokes equations has been intensively investigated in a sequence of outstanding papers (see, e.g., [1, 2, 3, 4, 5, 6, 23, 24, 32, 33]).

The study of the stability problem on the Boussinesq system near the shear flow is very recent. The work of Tao and Wu [27] was able to establish the stability and the enhanced dissipation phenomenon for the linearized 2D Boussinesq equations with only vertical dissipation, using the method of hypocoercivity introduced by C. Villani [29]. The Boussinesq system is different from the Navier-Stokes equations. The buoyancy force in the velocity equation could drive the growth of the energy and more generally the growth of the Sobolev norms. In addition, when there is only vertical dissipation, the control of the nonlinear terms becomes much more difficult. New techniques and estimates have to be created in order to handle the degenerate dissipation. It also appears that no previous work has handled the degenerate case. Since the Boussinesq system reduces to the Navier-Stokes equation when θ is identically zero, the stability results presented in this paper fill the gap on the Navier-Stokes equations with only vertical dissipation.

1.1. **Results.** We present three main results. The first result is on the linearized Boussinesq equations with either full dissipation or with only vertical dissipation. The upper bounds are explicit and sharp. The second result assesses the nonlinear stability and large-time behavior of the Boussinesq system with full dissipation. The third stability result is for the case with only vertical dissipation. Both nonlinear stability results are presented in order to make a direct comparison between the full dissipation and the degenerate dissipation cases.

For notational convenience, we shall write ω for $\widetilde{\omega}$ and θ for $\widetilde{\theta}$ from now on. To explain the linear stability result, we rewrite the equation for both the full dissipation case and the vertical dissipation case as

$$\begin{cases} \partial_t \omega + y\partial_x \omega = \nu(\sigma\partial_{xx} + \partial_{yy})\omega + \partial_x\theta, \\ \partial_t \theta + y\partial_x \theta = \mu(\sigma\partial_{xx} + \partial_{yy})\theta, \\ \omega|_{t=0} = \omega^{(0)}, \quad \theta|_{t=0} = \theta^{(0)}. \end{cases} \tag{1.5}$$

$\sigma = 1$ corresponds to the full dissipation case while $\sigma = 0$ to the vertical dissipation case. To help understand the stability results presented below, we explicitly solve the linear equation

$$\partial_t F + y\partial_x F = \nu(\sigma\partial_{xx} + \partial_{yy})F, \qquad F(x, y, 0) = F_0(x, y). \tag{1.6}$$

Taking the Fourier transform yields

$$\partial_t \widehat{F} - k\partial_\xi \widehat{F} = -\nu(\sigma k^2 + \xi^2)\widehat{F}, \qquad \widehat{F}(k, \xi, 0) = \widehat{F}_0(k, \xi),$$

where the Fourier transform is given by

$$\widehat{F}(k, \xi) = \mathcal{F}F = \frac{1}{(2\pi)^2}\int_{y\in\mathbb{R}}\int_{x\in\mathbb{T}} F(x, y)e^{-i(kx+\xi y)}\, dxdy.$$

Making the natural change of variables

$$\eta := \xi + kt, \qquad H(k, \eta, t) := \widehat{F}(k, \xi, t),$$

we find that

$$\partial_t H(k, \eta, t) = -\nu(\sigma k^2 + (\eta - kt)^2)\, H(k, \eta, t), \qquad H(k, \eta, 0) = \widehat{F}_0(k, \eta).$$

Integrating in time yields

$$H(k, \eta, t) = \widehat{F}_0(k, \eta) e^{-\nu \int_0^t (\sigma k^2 + (\eta - k\tau)^2)\, d\tau}.$$

Therefore,

$$
\begin{aligned}
\widehat{F}(k, \xi, t) &= H(k, \eta, t) = \widehat{F}_0(k, \xi + kt)\, e^{-\nu \int_0^t \sigma k^2 + (\xi + k(t-\tau))^2\, d\tau} \\
&= \widehat{F}_0(k, \xi + kt)\, e^{-\nu(\sigma k^2 + \xi^2)t}\, e^{-\frac{1}{3}\nu k^2 t^3 - \nu k \xi t^2}.
\end{aligned}
\tag{1.7}
$$

This explicit representation reflects the enhanced dissipation. Even when there is only vertical dissipation, namely $\sigma = 0$, the solution is dissipated and regularized in both directions. The dissipation time scale is $O(\nu^{-\frac{1}{3}})$, which is much faster than the standard dissipation time scale $O(\nu^{-1})$. Clearly the dissipation rate is inhomogeneous and depends on the frequencies k.

Solutions of (1.5) share the same properties as that of (1.6). The linear stability results on (1.5) are stated in Proposition 1.1 and Proposition 1.2. To make the statement precise, we define, for $f = f(x, y)$ with $(x, y) \in \mathbb{T} \times \mathbb{R}$ and $k \in \mathbb{Z}$,

$$f_k(y) := \frac{1}{2\pi} \int_{\mathbb{T}} f(x, y) e^{-ikx} dx.$$

In addition, we write $D = \frac{1}{i}\partial$. The linear stability result for (1.5) can then be stated as follows.

Proposition 1.1. *Let (ω, θ) be the solution to (1.5) with initial data $(\omega^{(0)}, \theta^{(0)})$. Assuming that $\nu \leq L\mu$ for some positive constant L, there exist constants $c > 0$, $C > 0$ such that for any $k \in \mathbb{Z}$, $t > 0$,*

$$
\begin{aligned}
\|\theta_k(t)\|_{L_y^2} &\leq C \|\theta_k^{(0)}\|_{L_y^2} e^{-\frac{1}{16}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}} t}, \\
\|\omega_k(t)\|_{L_y^2} &\leq C \big(\|\omega_k^{(0)}\|_{L_y^2} + (\nu\mu)^{-\frac{1}{6}} |k|^{\frac{1}{3}} \|\theta_k^{(0)}\|_{L_y^2} \big) e^{-c\nu^{\frac{1}{3}}|k|^{\frac{2}{3}} t}.
\end{aligned}
\tag{1.8}
$$

More generally, for any integer $N \geq 0$, there exist $c_N > 0$ and $C_N > 0$ such that for any $k \in \mathbb{Z}$, $t > 0$,

$$
\begin{aligned}
\|D_y^N \theta_k(t)\|_{L_y^2} &\leq C_N e^{-c_N \mu^{\frac{1}{3}}|k|^{\frac{2}{3}} t} \Big(\|D_y^N \theta_k^{(0)}\|_{L_y^2} + (\mu^{-1}|k|)^{\frac{N}{3}} \|\theta_k^{(0)}\|_{L_y^2} \Big), \\
\|D_y^N \omega_k(t)\|_{L_y^2} &\leq C_N e^{-c_N \nu^{\frac{1}{3}}|k|^{\frac{2}{3}} t} \Big(\|D_y^N \omega_k^{(0)}\|_{L_y^2} + (\nu\mu)^{-\frac{1}{6}} |k|^{\frac{1}{3}} \|D_y^N \theta_k^{(0)}\|_{L_y^2} \\
&\qquad + (\nu^{-1}|k|)^{\frac{N}{3}} \big(\|\omega_k^{(0)}\|_{L_y^2} + (\nu\mu)^{-\frac{1}{6}} |k|^{\frac{1}{3}} \|\theta_k^{(0)}\|_{L_y^2} \big) \Big).
\end{aligned}
\tag{1.9}
$$

A similar linear stability result for a slightly different domain was obtained in [27], but the proof presented here is different, simpler and more compact. The estimates in Proposition 1.1 can be converted into a more elegant statement that allows a direct comparison with the nonlinear stability results to be presented. We explain and define a few notations. (1.7) clearly reveals the distinction between the zero mode case $k = 0$ and the nonzero modes $k \neq 0$. This triggers the definitions

$$f_0 := (\mathbb{P}_0 f)(y) = \frac{1}{2\pi} \int_{\mathbb{T}} f(x, y) dx, \qquad f_{\neq} := \mathbb{P}_{\neq} f = f - \mathbb{P}_0 f, \tag{1.10}$$

which represents the projection onto 0 frequency and the projection onto non-zero frequencies. In the process of deriving (1.7), we made the change of variable $\eta = \xi + tk$, which naturally invites the definition of the time-dependent elliptic operator, for $t \geq 0$,

$$\Lambda_t^2 = 1 - \partial_x^2 - (\partial_y + t\partial_x)^2, \tag{1.11}$$

or, in terms of its symbol, $\Lambda_t^2(k, \xi) = 1 + k^2 + (\xi + tk)^2$. In general, we denote Λ_t^b with $b \in \mathbb{R}$ to be the Fourier multiplier with symbol $\Lambda_t^b(k, \xi) = \left(1 + k^2 + (\xi + tk)^2\right)^{\frac{b}{2}}$. It is easy to check that the operator Λ_t^b commutes with the differential operator with variable coefficients $\partial_t + y\partial_x$, namely

$$\Lambda_t^b \left(\partial_t + y\partial_x\right) = \left(\partial_t + y\partial_x\right) \Lambda_t^b. \tag{1.12}$$

Therefore, applying Λ_t^b allows us to obtain the derivative estimates without destroying the structure of the linearized equation (1.5). Furthermore, Λ_t^b shares similarities with the standard fractional Laplacian operators. For example, for any $b > 0$,

$$\|\Lambda_t^b(fg)\|_{L^2} \leq \|f\|_{L^\infty} \|\Lambda_t^b g\|_{L^2} + \|g\|_{L^\infty} \|\Lambda_t^b f\|_{L^2}$$

and, for $b > 1$,

$$\|f(t)\|_{L^\infty(\Omega)} \leq C \|\widehat{f}(t)\|_{L^1(\Omega)} \leq C \|\Lambda_t^b f(t)\|_{L^2(\Omega)}.$$

To precisely state the second linear stability result, we define the horizontal fractional derivative by

$$\widehat{|D_x|^\gamma f}(k, \xi) = |k|^\gamma \widehat{f}(k, \xi).$$

The linear stability result in Proposition 1.1 can be converted into an estimate in the physical space.

Proposition 1.2. *Let (ω, θ) be the solution to (1.5) with initial data $(\omega^{(0)}, \theta^{(0)})$. Then there exists $C > 0$ such that for $b \in \mathbb{R}$,*

$$\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}} \|D_y \Lambda_t^b \omega\|_{L_t^2(L^2)} + \sigma \nu^{\frac{1}{2}} \|D_x \Lambda_t^b \omega\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}} \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L_t^2(L^2)}$$

$$+ (\nu\mu)^{-\frac{1}{6}} \left(\||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^\infty(L^2)} + \mu^{\frac{1}{2}} \|D_y |D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)} + \sigma \mu^{\frac{1}{2}} \||D_x|^{\frac{4}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)} \right.$$

$$\left. + \mu^{\frac{1}{6}} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)} \right)$$

$$\leq C \left(\|\omega^{(0)}\|_{H^b} + (\nu\mu)^{-\frac{1}{6}} \||D_x|^{\frac{1}{3}} \theta^{(0)}\|_{H^b} \right).$$

We assume $\nu = \mu$ for simplicity from now on. The main focus of this paper is actually the nonlinear stability. We are able to establish the stability and large-time behavior for both the full dissipation case and the case with only vertical dissipation. Certainly the proof for the vertical dissipation case also works for the full dissipation. Both results are presented here for a direct comparison. When the dissipation is degenerate, more strict assumptions have to be made on the initial data. The stability result for the fully dissipative Boussinesq equation is stated in the following theorem.

Theorem 1.3. *Assume $b > 1$, $\beta \geq \frac{1}{2}$, $\delta \geq \beta + \frac{1}{3}$, $\alpha \geq \delta - \beta + \frac{2}{3}$ and that the initial data $(\omega^{(0)}, \theta^{(0)})$ satisfies*

$$\|\omega^{(0)}\|_{H^b} \leq \varepsilon \nu^\beta, \qquad \|\theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\alpha, \qquad \||D_x|^{\frac{1}{3}}\theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\delta,$$

for some sufficiently small $\varepsilon > 0$. Then the solution (ω, θ) to (1.3) satisfies that

$$\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla \Lambda_t^b \omega\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\beta,$$

$$\|\Lambda_t^b \theta\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla \Lambda_t^b \theta\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\alpha$$

and

$$\||D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla |D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{2}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\delta.$$

In the case when there is only vertical dissipation, the stability and large-time behavior result is stated as follows.

Theorem 1.4. *Let $b > \frac{4}{3}$, $\beta \geq \frac{2}{3}$, $\delta \geq \beta + \frac{1}{3}$, $\alpha \geq \delta - \beta + \frac{2}{3}$. Assume that*

$$\|\omega^{(0)}\|_{H^b} \leq \varepsilon \nu^\beta, \qquad \|\theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\alpha, \qquad \||D_x|^{\frac{1}{3}}\theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\delta$$

for some sufficiently small $\varepsilon > 0$. Then the solution to the system (1.4) with initial data $(\omega^{(0)}, \theta^{(0)})$ satisfies

$$\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|D_y \Lambda_t^b \omega\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\beta,$$

$$\|\Lambda_t^b \theta\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|D_y \Lambda_t^b \theta\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\alpha$$

and

$$\||D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|D_y |D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{2}{3}}\Lambda_t^b \theta\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon \nu^\delta.$$

Special consequences of Theorem 1.3 and Theorem 1.4 are the nonlinear stability for the 2D Navier-Stokes equation with full dissipation or with only vertical dissipation. When $\theta \equiv 0$, the system (1.3) reduces to the 2D Navier-Stokes vorticity equation with full dissipation. The stability problem of the 2D Couette flow or more general shear flows near the Couette flow has previously been investigated on the 2D Navier-Stokes equations with full dissipation, we refer to the references [6, 23, 24]. In particular, we recover the threshold index estimate $\beta \geq \frac{1}{2}$ with data in H^b, $b > 1$ established firstly in [6]. Moreover, we also obtain the 1/3-horizontal regularity, which is better than the enhanced dissipation quantity $\|\Lambda_t^b \omega_{\neq}\|_{L^2(L^2)}$ obtained in [6, Theorem 1.1]. This is due to the different choice of the multipliers that are used

in the proof. On the other hand, since the stability result for the 2D Navier-Stokes equation with only vertical dissipation is completely new, we state it as a corollary. When $\theta \equiv 0$, the system (1.4) reduces to the 2D Navier-Stokes vorticity equation with only vertical dissipation,

$$\begin{cases} \partial_t \omega + y\partial_x \omega + (\mathbf{u} \cdot \nabla)\omega = \nu \partial_{yy}\omega, \\ \mathbf{u} = -\nabla^\perp(-\Delta)^{-1}\omega. \end{cases} \tag{1.13}$$

Theorem 1.4 yields the following stability result for (1.13).

Corollary 1.5. *Let $b > \frac{4}{3}$ and $\beta \geq \frac{2}{3}$. Assume the initial vorticity $\omega^{(0)}$ satisfies*

$$\|\omega^{(0)}\|_{H^b} \leq \varepsilon \nu^\beta$$

for some suitable small number $\varepsilon > 0$. Then the corresponding solution ω to (1.13) satisfies

$$\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)} + \nu^{\frac{1}{2}}\|D_y \Lambda_t^b \omega\|_{L_t^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L_t^2(L^2)}$$
$$+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)} \leq C\varepsilon\nu^\beta.$$

Remark 1.6. *Various works [3, 6, 23, 24] for Couette flow are done with the change of coordinates $X = x - yt, Y = y$. Our results and proofs can be translated into the coordinates (X, Y) as well. In this paper, we choose to work in the standard physical variables, since we would like to highlight the role played by the non self-adjoint operator $y\partial_x - \nu\partial_{yy}$. The multiplier $\mathcal{M}_1$ defined in (1.14) which allows us to prove the optimal enhanced dissipation estimates, is constructed according to the first-order bracket structure of the linear operator.*

Remark 1.7. *As observed in [19], we can prove an exponential converge of the non-zero modes for both the full dissipation case and the vertical dissipation case. See Remark 4.1.*

Remark 1.8. *After we posted the first version of this paper, C. Zillinger kindly informed us of his very recent manuscript on a similar topic [34]. He studied the stability on perturbations near both the Couette flow and the hydrostatic balance simultaneously. In particular, he obtained the linear stability of the Couette flow for the 2D Boussinesq system with partial dissipation by using explicit solution formulas. In addition, he also showed the nonlinear stability of the Couette flow for the 2D Boussinesq system with full dissipation by using the multipliers constructed in [6], with weaker enhanced dissipation terms obtained. Compared with [34], this paper focuses on the nonlinear stability for the 2D Boussinesq system with only vertical dissipation, and the gain of 1/3-horizontal derivative is well characterized by carefully choosing the multiplier.*

1.2. **Sketch of the proof.** The proofs of the nonlinear stability results stated in Theorem 1.3 and Theorem 1.4 are not trivial. As aforementioned, due to the presence of the buoyancy force, it is not plausible to establish the desired stability results without taking full advantage of the enhanced dissipation, created by the

combination of $y\partial_x\omega$ with $\partial_{yy}\omega$ in the vorticity equation and of $y\partial_x\theta$ with $\partial_{yy}\theta$ in the temperature equation.

Let us explain how to extract the enhanced dissipation, especially the regularity in the horizontal direction, generated by the non-self-adjoint operator $y\partial_x-\nu\partial_{yy}$. We design a Fourier multiplier operator $\mathcal{M}$ defined as follows. Choose a real-valued, non-decreasing function $\varphi \in C^\infty(\mathbb{R})$ satisfying $0 \le \varphi \le 1$, $\varphi(\tau) = \begin{cases} 0, & \tau \in (-\infty, -2], \\ 1, & \tau \in [2, \infty), \end{cases}$ and $\varphi' = \frac{1}{4}$ on $[-1,1]$. Define the Fourier multiplier $\mathcal{M} = \mathcal{M}(D_x, D_y)$ as $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 + 1$ with symbols $\mathcal{M}_1$ and $\mathcal{M}_2$ given by

$$\mathcal{M}_1(k,\xi) = \varphi\big(\nu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\mathrm{sgn}(k)\xi\big), \ k \ne 0,$$

$$\mathcal{M}_2(k,\xi) = \frac{1}{k^2}\Big(\arctan\frac{\xi}{k} + \frac{\pi}{2}\Big), \ k \ne 0, \tag{1.14}$$

$$\mathcal{M}_1(0,\xi) = \mathcal{M}_2(0,\xi) = 0.$$

Then $\mathcal{M}$ is a self-adjoint Fourier multiplier acting on $L^2(\Omega)$ and verifies that

$$1 \le \mathcal{M} \le 2 + \pi.$$

The construction of the multiplier $\mathcal{M}$ is inspired by the works [13, 14] on the (non-self-adjoint) linearized operator of the 2D Navier-Stokes equation around the Oseen vortices. Before we show the key point of our proof, let us remark the fact that for a self-adjoint operator $A = A^*$ and a skew-adjoint operator $B = -B^*$ on L^2, we have the following identity

$$\begin{aligned} 2\mathrm{Re}\langle Af, Bf\rangle_{L^2} &= \langle Af, Bf\rangle_{L^2} + \langle Bf, Af\rangle_{L^2} \\ &= \langle B^*Af, f\rangle_{L^2} + \langle A^*Bf, f\rangle_{L^2} \\ &= \langle (AB - BA)f, f\rangle_{L^2} = \langle [A,B]f, f\rangle_{L^2}, \end{aligned} \tag{1.15}$$

where $[A,B] := AB - BA$ denotes the commutator between A and B.

Now taking the inner product of $(y\partial_x - \nu\partial_{yy})\omega$ with $\mathcal{M}\omega$ leads to the quantity

$$R := 2\mathrm{Re}\langle y\partial_x\omega, \mathcal{M}\omega\rangle_{L^2} - 2\nu\mathrm{Re}\langle \partial_{yy}\omega, \mathcal{M}\omega\rangle_{L^2},$$

for which we intend to prove a lower bound. Using the fact that $\mathcal{M}$ is self-adjoint and $y\partial_x$ is skew-adjoint, we have

$$\begin{aligned} 2\mathrm{Re}\langle y\partial_x\omega, \mathcal{M}\omega\rangle_{L^2} &= \langle [\mathcal{M}, y\partial_x]\omega, \omega\rangle_{L^2} \\ &= \sum_k \int_{\mathbb{R}} (k\partial_\xi\mathcal{M})\,|\widehat{\omega}(k,\xi)|^2\,d\xi, \end{aligned}$$

where we have used Plancherel's theorem in the last step. Consequently,

$$R = \sum_k \int_{\mathbb{R}} \big(k\partial_\xi\mathcal{M} + 2\nu\,\mathcal{M}\xi^2\big)\,|\widehat{\omega}(k,\xi)|^2\,d\xi.$$

The Fourier multiplier $\mathcal{M}_1$ is constructed in order to capture the regularity in the horizontal direction: according to the definition of $\mathcal{M}_1$, for any $k \ne 0$ and $\xi \in \mathbb{R}$,

$$k\partial_\xi\mathcal{M}_1(k,\xi) = \nu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'\big(\nu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\mathrm{sgn}(k)\xi\big),$$

which is bounded from below by $\frac{1}{4}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}$ when $|\xi| \leq \nu^{-\frac{1}{3}}|k|^{\frac{1}{3}}$, thanks to the special choice of the function φ. One finds the following important inequality

$$\nu\xi^2 + k\partial_\xi \mathcal{M}_1 \geq \frac{1}{4}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}, \quad \forall \xi \in \mathbb{R}, \ k \in \mathbb{Z}.^1$$

The Fourier multiplier $\mathcal{M}_2$ is designed to control the velocity in the nonlinear term since we have

$$k\partial_\xi \mathcal{M}_2(k, \xi) = \frac{1}{k^2 + \xi^2}.$$

Combining the above estimates, one achieves the lower bound

$$R \geq \nu\|\partial_y \omega\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\omega\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\omega_{\neq}\|_{L^2}^2. \qquad (1.16)$$

(1.16) leads to a control of $\frac{1}{3}$-horizontal derivative of ω and this is the main reason why we can possibly control the buoyancy force, as well as the nonlinear terms. Let us also remark that the exponent $1/3$ on the right hand side of (1.16) is sharp in the sense that there exist $c > 0$ and functions $\omega_\nu \in L^2$ such that the equality $\|(y\partial_x - \nu\partial_{yy})\omega_\nu\|_{L^2}\|\omega_\nu\|_{L^2} = c\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\omega_\nu\|_{L^2}^2$ holds for all $0 < \nu < 1$. This is due to the special first-order bracket structure of the operator $y\partial_x - \nu\partial_{yy}$, see [17], [13] for more details.

Standard Sobolev type energy estimates would not work since they would destroy the combination, see Proposition 1.1. We shall apply the operator Λ_t^b defined below (1.11) which allows differentiate the equations in (1.3) and (1.4) without changing the linear structures of the system, and then apply the multiplier $\mathcal{M}$ to obtain the desired enhanced dissipations for higher-order derivatives.

The buoyancy term in the equation of the vorticity ω takes the form $\partial_x \theta$, which contains full one horizontal derivative. In the process of estimating $\|\Lambda_t^b \omega\|_{L^2}$, the buoyancy term can be bounded by

$$|\langle \partial_x \Lambda_t^b \theta, \mathcal{M}\Lambda_t^b \omega \rangle_{L^2}| \leq \||D_x|^{\frac{2}{3}}\Lambda_t^b \theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L^2},$$

which contains $\frac{2}{3}$-horizontal derivative on θ. Since the enhanced dissipation in the estimate of $\|\Lambda_t^b \theta\|_{L^2}$ contains only $\frac{1}{3}$-horizontal derivative dissipation, we need to estimate $\||D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L^2}$ in order to control the buoyancy term. This explains why we combine the estimates of $\|\Lambda_t^b \omega\|_{L^2}$, $\|\Lambda_t^b \theta\|_{L^2}$ and $\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L^2}$.

Most of the efforts are devoted to obtaining suitable upper bounds on the nonlinear terms. This is a very delicate process especially when there is only vertical dissipation. Let us explain some of the difficulties and our approach in dealing with them when we estimate the nonlinear term $\mathbf{u} \cdot \nabla\theta$. The velocity $\mathbf{u}$ is represented in terms of ω via the Biot-Savart law

$$\mathbf{u} = -\nabla^\perp(-\Delta)^{-1}\omega.$$

---

[1] The choice of the multiplier $\mathcal{M}_1$ is not unique, for example one may use the multiplier $\mathcal{M}_1(k, \xi) = c\arctan(\nu^{\frac{1}{3}}\mathrm{sgn}(k)|k|^{-\frac{1}{3}}\xi) + C$.

To distinguish between the different behaviors of the zeroth mode and the nonzero modes, we split the velocity into two parts according to (1.10)

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_{\neq} = \begin{pmatrix} u_0 \\ 0 \end{pmatrix} + \begin{pmatrix} u_{\neq} \\ v_{\neq} \end{pmatrix} = \begin{pmatrix} u_0 \\ 0 \end{pmatrix} + \begin{pmatrix} \partial_y(-\Delta)^{-1}\omega_{\neq} \\ -\partial_x(-\Delta)^{-1}\omega_{\neq} \end{pmatrix},$$

where $u_0 = \partial_y(-\partial_y^2)^{-1}\omega_0$. Accordingly, $\mathbf{u} \cdot \nabla\theta$ is decomposed into three parts,

$$\mathbf{u} \cdot \nabla\theta = u_0\partial_x\theta + \partial_y(-\Delta)^{-1}\omega_{\neq}\partial_x\theta - \partial_x(-\Delta)^{-1}\omega_{\neq}\partial_y\theta.$$

Due to the lack of dissipation in the horizontal direction, it is impossible to derive suitable bounds for the first two terms in $\mathbf{u} \cdot \nabla\theta$ directly. Our strategy to overcome this difficulty is to estimate the scalar product

$$H := \langle \Lambda_t^b(\mathbf{u} \cdot \nabla\theta), \mathcal{M}\Lambda_t^b\theta \rangle_{L^2}.$$

With the help of the Fourier multiplier $\mathcal{M}$, the frequency space is divided into different subdomains to facilitate cancellations and derivative distributions. Commutator estimates are employed to shift derivatives so that we are able to control the nonlinear terms. Detailed estimates are very technical and left to the proof of Theorem 1.4 in Section 4.

The rest of this paper is divided into three sections. Section 2 proves the linear stability stated in Propositions 1.1 and 1.2. Theorem 1.3 is proved in Section 3 while Section 4 presents the proof of Theorem 1.4.

2. Proofs of Propositions 1.1 and 1.2

This section is devoted to the proofs of the linear stability results stated in Propositions 1.1 and 1.2. These results are valid for both the full dissipation case and the case with only vertical dissipation. Notice that the two equations of the linear system (1.5) are decoupled. As a result, an explicit solution formula to (1.5) is available using Fourier transform as for the model problem (1.6). Indeed, by a direct computation, we have

$$\widehat{\theta}(k, \xi, t) = \widehat{\theta}_0(k, \xi + kt)e^{-\mu\int_0^t(\sigma k^2 + (\xi + kt - k\tau)^2)d\tau},$$

$$\widehat{\omega}(k, \xi, t) = \widehat{\omega}_0(k, \xi + kt)e^{-\nu\int_0^t(\sigma k^2 + (\xi + kt - k\tau)^2)d\tau}$$

$$+ ik\int_0^t e^{-\nu\int_0^{t-s}(\sigma k^2 + (\xi + kt - k\tau)^2)d\tau}\widehat{\theta}(k, \xi, s)ds.$$

Then it is immediate to infer the bounds of Proposition 1.1 and 1.2. Here we present the multiplier method in order to make it more accessible to the nonlinear case. These multiplier operators are constructed to extract the enhanced dissipation from the non-self-adjoint operators $y\partial_x - \nu\partial_{yy}$ and $y\partial_x - \mu\partial_{yy}$. Moreover, the multiplier method is flexible and may be used to study more general models, for which an explicit solution formula might not be available.

We are ready to prove Proposition 1.1.

Proof of Proposition 1.1. By projecting the equations in (1.5) onto each frequency, we obtain the system in the y-variable only,

$$\begin{cases} \partial_t \omega_k + \nu(D_y^2 + \sigma k^2)\omega_k + iky\omega_k = ik\theta_k, \\ \partial_t \theta_k + \mu(D_y^2 + \sigma k^2)\theta_k + iky\theta_k = 0, \\ \omega_k|_{t=0} = \omega_k^{(0)}, \quad \theta_k|_{t=0} = \theta_k^{(0)}, \end{cases} \tag{2.1}$$

where we have used the notation $D = \frac{1}{i}\partial$. We note that $\sigma = 1$ corresponds to the full dissipation case while $\sigma = 0$ to the case with only vertical dissipation. Since ω_k and θ_k may be complex-valued, the the L_y^2-inner product is given by

$$\langle f, g \rangle_{L_y^2} = \int_{\mathbb{R}} f(y)\, \bar{g}(y)\, dy.$$

By taking the L_y^2-inner product of θ_k with the second equation in (2.1), we have

$$\frac{1}{2}\frac{d}{dt}\|\theta_k\|_{L_y^2}^2 + \mu\|D_y\theta_k\|_{L_y^2}^2 + \sigma\mu k^2\|\theta_k\|_{L_y^2}^2 = 0. \tag{2.2}$$

To further the estimates, we define and apply Fourier multiplier operators. If $k > 0$, we define a multiplier M_k by

$$M_k\theta_k := \varphi(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\theta_k,$$

where φ is a real-valued, non-decreasing function, $\varphi \in C^\infty(\mathbb{R})$ satisfying $0 \le \varphi \le 1$ and $\varphi' = \frac{1}{4}$ on $[-1, 1]$. Clearly, M_k is a self-adjoint and non-negative Fourier multiplier operator. We take the L_y^2-inner product of the second equation in (2.1) with $M_k\theta_k$. The following basic identities hold,

$$2\mathrm{Re}\langle \partial_t\theta_k, M_k\theta_k \rangle_{L_y^2} = \frac{d}{dt}\langle M_k\theta_k, \theta_k \rangle_{L_y^2},$$

$$2\mathrm{Re}\langle \mu(D_y^2 + \sigma k^2)\theta_k, M_k\theta_k \rangle_{L_y^2} = \langle 2\mu(D_y^2 + \sigma k^2)M_k\theta_k, \theta_k \rangle_{L_y^2},$$

$$2\mathrm{Re}\langle iky\theta_k, M_k\theta_k \rangle_{L_y^2} = \langle [M_k, iky]\theta_k, \theta_k \rangle_{L_y^2},$$

where in the last equation we have used the fact that M_k is self-adjoint and iky is skew-adjoint. Here the bracket in $[M_k, iky]$ denotes the standard commutator. Noticing that

$$[M_k, iky] = \left[\varphi(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y), iky\right] = \mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y),$$

we obtain

$$\frac{d}{dt}\langle M_k\theta_k, \theta_k \rangle_{L_y^2} + \langle 2\mu(D_y^2 + \sigma k^2)\varphi(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\theta_k, \theta_k \rangle_{L_y^2}$$
$$+ \langle \mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\theta_k, \theta_k \rangle_{L_y^2} = 0.$$

Together with (2.2), this gives

$$\frac{d}{dt}\left(\|\theta_k\|_{L_y^2}^2 + \langle M_k\theta_k, \theta_k \rangle_{L_y^2}\right)$$
$$+ \left\langle \left(2\mu(D_y^2 + \sigma k^2)\left(1 + \varphi(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\right) + \mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\right)\theta_k, \theta_k \right\rangle_{L_y^2} = 0.$$

By the choice of the function φ, there holds

$$\mu(\xi^2 + \sigma k^2)\big(1 + 2\varphi(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\xi)\big) + \mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'(\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\xi) \geq \frac{1}{4}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}$$

for all $k > 0$, $\mu > 0$, $\xi \in \mathbb{R}$. Therefore,

$$\frac{d}{dt}\big(\langle(1 + M_k)\theta_k, \theta_k\rangle_{L_y^2}\big) + \mu\|D_y\theta_k\|_{L_y^2}^2 + \sigma\mu k^2\|\theta_k\|_{L_y^2}^2 + \frac{1}{4}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|\theta_k\|_{L_y^2}^2 \leq 0. \quad (2.3)$$

Integrating in t and using properties of M_k, we obtain the first inequality in (1.8) for $k > 0$. In the case when $k < 0$, we define the multiplier M_k by

$$M_k\theta_k := \varphi(-\mu^{\frac{1}{3}}|k|^{-\frac{1}{3}}D_y)\theta_k,$$

and define $M_0 = 0$, we can deduce the first inequality in (1.8) for $k \leq 0$.

We prove the first inequality in (1.9) by induction. Differentiating the second equation in (2.1) N times with respect to y leads to

$$\partial_t D_y^N \theta_k + \mu(D_y^2 + \sigma k^2)D_y^N \theta_k + iky D_y^N \theta_k + kN D_y^{N-1}\theta_k = 0.$$

Taking the L_y^2-inner product with $(1 + M_k)D_y^N \theta_k$ then gives

$$\frac{d}{dt}\langle(1 + M_k)D_y^N\theta_k, D_y^N\theta_k\rangle_{L_y^2} + \mu\|D_y^{N+1}\theta_k\|_{L_y^2}^2 + \sigma\mu k^2\|D_y^N\theta_k\|_{L_y^2}^2$$

$$+ \frac{1}{4}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|D_y^N\theta_k\|_{L_y^2}^2 \leq -2\mathrm{Re}\langle kN D_y^{N-1}\theta_k, (1 + M_k)D_y^N\theta_k\rangle_{L_y^2}$$

$$\leq \frac{1}{8}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|D_y^N\theta_k\|_{L_y^2}^2 + 32N^2\mu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\|D_y^{N-1}\theta_k\|_{L_y^2}^2.$$

Integrating in t yields

$$\|D_y^N\theta_k(t)\|_{L_y^2}^2 \leq 2\|D_y^N\theta_k^{(0)}\|_{L_y^2}^2 e^{-\frac{1}{16}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}t}$$

$$+ C_N\mu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\int_0^t \|D_y^{N-1}\theta_k(s)\|_{L_y^2}^2 e^{-\frac{1}{16}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}(t-s)}ds.$$

Then the first inequality in (1.9) follows from the induction assumption.

To handle the second inequality in (1.8), we define

$$\overline{M}_k := \varphi(\nu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\mathrm{sgn}(k)D_y) \text{ for } k \neq 0, \qquad \overline{M}_0 = 0$$

and multiply the ω_k equation in (2.1) by ω_k and $\overline{M}_k\omega_k$ to obtain

$$\frac{d}{dt}\langle(1 + \overline{M}_k)\omega_k, \omega_k\rangle_{L_y^2} + \nu\|D_y\omega_k\|_{L_y^2}^2 + \sigma\nu k^2\|\omega_k\|_{L_y^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|\omega_k\|_{L_y^2}^2$$

$$\leq 2\mathrm{Re}\langle ik\theta_k, (1 + \overline{M}_k)\omega_k\rangle_{L_y^2}.$$

Applying Young's inequality to the right-hand side yields

$$\frac{d}{dt}\big(\langle(1 + \overline{M}_k)\omega_k, \omega_k\rangle_{L_y^2}\big) + \nu\|D_y\omega_k\|_{L_y^2}^2 + \sigma\nu k^2\|\omega_k\|_{L_y^2}^2 + \frac{1}{8}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|\omega_k\|_{L_y^2}^2$$

$$\leq 32\nu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\|\theta_k\|_{L_y^2}^2. \qquad (2.4)$$

Due to $\nu \leq L\mu$, we take a small enough constant c so that $c \leq \frac{1}{8}$, $c\nu^{\frac{1}{3}} \leq \frac{1}{32}\mu^{\frac{1}{3}}$. Then we get, by multiplying the above inequality by $e^{c\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}t}$ and using the first inequality in (1.8), that

$$
\begin{aligned}
\frac{d}{dt}\big(\langle(1+\overline{M}_k)\omega_k, \omega_k\rangle_{L_y^2} e^{c\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}t}\big) &\leq 32\nu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\|\theta_k\|_{L_y^2}^2 e^{c\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}t} \\
&\leq C\nu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\|\theta_k^{(0)}\|_{L_y^2} e^{\big(c\nu^{\frac{1}{3}}-\frac{1}{16}\mu^{\frac{1}{3}}\big)|k|^{\frac{2}{3}}t}.
\end{aligned}
$$

Integrating in t, we obtain

$$
\|\omega_k(t)\|_{L_y^2}^2 \leq C\big(\|\omega_k^{(0)}\|_{L_y^2}^2 + (\nu\mu)^{-\frac{1}{3}}|k|^{\frac{2}{3}}\|\theta_k^{(0)}\|_{L_y^2}^2\big)e^{-c\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}t}.
$$

Differentiating the equation of ω_k in (2.1) and using the estimates for θ_k, we can deduce the second inequality in (1.9), under the assumption that $\nu \leq L\mu$. This completes the proof of Proposition 1.1. $\qquad\square$

Proposition 1.2 is a consequence of Proposition 1.1. We recall that the operator Λ_t defined in (1.11) commutes with $\partial_t + y\partial_x$, namely, for any $b \in \mathbb{R}$, there holds (1.12). Therefore it commutes with the linear equation in (1.5).

Proof of Proposition 1.2. For any $b \in \mathbb{R}$, we apply Λ_t^b to the equations in (1.5). Since Λ_t^b commutes the equations in (1.5), the upper bounds in Proposition 1.1 and the estimates in the proof of Proposition 1.1 remain valid if we replace ω and θ by $\Lambda_t^b\omega$ and $\Lambda_t^b\theta$, respectively, in Proposition 1.1. Similarly, since any horizontal derivatives also commute with the linear equations in (1.5), $|D_x|^{\frac{1}{3}}\Lambda_t^b\theta$ enjoys similar estimates as those for θ. In particular, similar to (2.3), one has

$$
\begin{aligned}
\frac{d}{dt}\big(\langle(1+M_k)|D_x|^{\frac{1}{3}}(\Lambda_t^b\theta)_k, |D_x|^{\frac{1}{3}}(\Lambda_t^b\theta)_k\rangle_{L_y^2}\big) &+ \mu\|D_y|D_x|^{\frac{1}{3}}(\Lambda_t^b\theta)_k\|_{L_y^2}^2 \\
+ \sigma\mu k^2\||D_x|^{\frac{1}{3}}(\Lambda_t^b\theta)_k\|_{L_y^2}^2 &+ \frac{1}{4}\mu^{\frac{1}{3}}|k|^{\frac{2}{3}}\||D_x|^{\frac{1}{3}}(\Lambda_t^b\theta)_k\|_{L_y^2}^2 \leq 0.
\end{aligned}
$$

Integrating in t and summing up the resulting inequalities for $k \in \mathbb{Z}$, we find

$$
\begin{aligned}
\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^\infty(L^2)}^2 &+ \mu\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \sigma\mu\||D_x|^{\frac{4}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \\
&+ \frac{1}{4}\mu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \leq 2\||D_x|^{\frac{1}{3}}\theta^{(0)}\|_{H^b}^2.
\end{aligned} \tag{2.5}
$$

While similar to (2.4), we have

$$
\begin{aligned}
\frac{d}{dt}\big(\langle(1+\overline{M}_k)(\Lambda_t^b\omega)_k, (\Lambda_t^b\omega)_k\rangle_{L_y^2}\big) &+ \nu\|D_y(\Lambda_t^b\omega)_k\|_{L_y^2}^2 + \sigma\nu k^2\|(\Lambda_t^b\omega)_k\|_{L_y^2}^2 \\
+ \frac{1}{8}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}}\|(\Lambda_t^b\omega)_k\|_{L_y^2}^2 &\leq 32\nu^{-\frac{1}{3}}|k|^{\frac{4}{3}}\|(\Lambda_t^b\theta)_k\|_{L_y^2}^2.
\end{aligned}
$$

Integrating in t and summing up the resulting inequalities for $k \in \mathbb{Z}$, we obtain

$$
\begin{aligned}
\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}^2 &+ \nu\|D_y\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \sigma\nu\||D_x|\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 \\
&+ \frac{1}{8}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 \leq 2\|\omega^{(0)}\|_{H^b}^2 + 32\nu^{-\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \quad (2.6) \\
&\qquad\qquad\qquad\qquad\qquad \leq 2\|\omega^{(0)}\|_{H^b}^2 + C(\nu\mu)^{-\frac{1}{3}}\||D_x|^{\frac{1}{3}}\theta^{(0)}\|_{H^b}^2,
\end{aligned}
$$

where in the last step, we used (2.5).

By summing up $(\nu\mu)^{-\frac{1}{3}}\times$(2.5) and (2.6), we obtain the desired estimates in Proposition 1.2. We remark that the coefficient $(\nu\mu)^{-\frac{1}{6}}$ in the estimate of Proposition 1.2 helps unify the bound in terms of the initial data. This completes the proof of Proposition 1.2. $\qquad\square$

3. Proof of Theorem 1.3

This section presents the proof of Theorem 1.3 stating the nonlinear stability for (1.3). The framework is the bootstrap argument, which consists of two main steps. The first step is to establish the *a priori* bounds while the second is to apply and complete the bootstrap argument by using the *a priori* bounds. Main efforts are devoted to obtaining suitable *a priori* bounds. As described in the introduction, one component in achieving the bounds is to extract the enhanced dissipation by constructing and applying suitable Fourier multipliers. Another one is to bound the nonlinear terms suitably. To do so, we separate the horizontal zeroth mode from the non-zeroth modes to distinguish their different behaviors. We make use of sharp commutator estimates.

To help prepare for the proof, we recall several notations and basic facts. We make extensive use of the operator Λ_t^b defined below (1.11). The basic properties stated in the following lemma will be used frequently.

Lemma 3.1. *The operator Λ_t^b defined below (1.11) satisfies the following properties*

(1) *For any $b \in \mathbb{R}$, Λ_t^b commutes with $\partial_t + y\partial_x$, namely*

$$\Lambda_t^b \left(\partial_t + y\partial_x \right) = \left(\partial_t + y\partial_x \right) \Lambda_t^b.$$

(2) *For any $b > 0$,*

$$\|\Lambda_t^b(fg)\|_{L^2} \leq \|f\|_{L^\infty}\|\Lambda_t^b g\|_{L^2} + \|g\|_{L^\infty}\|\Lambda_t^b f\|_{L^2}.$$

Moreover, for $b > 1$, we have

$$\|f(t)\|_{L^\infty} \leq C\|\widehat{f}(t)\|_{L^1} \leq C\|\Lambda_t^b f(t)\|_{L^2}.$$

and consequently,

$$\|\Lambda_t^b(fg)\|_{L^2} \leq C \, \|\Lambda_t^b f\|_{L^2} \, \|\Lambda_t^b g\|_{L^2}.$$

Remark 3.2. *It is easy to observe that for any s non-negative and $b > 1$, there holds*

$$\||D_x|^s \Lambda_t^b(fg)\|_{L^2} \leq C \left(\||D_x|^s \Lambda_t^b f\|_{L^2} \|\Lambda_t^b g\|_{L^2} + \|\Lambda_t^b f\|_{L^2} \||D_x|^s \Lambda_t^b g\|_{L^2} \right). \qquad (3.1)$$

Indeed, by applying Plancherel's theorem, we have

$$\||D_x|^s\Lambda_t^b(fg)\|_{L^2}^2 = \sum_{k\in\mathbb{Z}}|k|^{2s}\int_{\mathbb{R}}\big(\Lambda_t^b(k,\xi)\big)^2\Big(\sum_{\ell\in\mathbb{Z}}\int_{\mathbb{R}}\widehat{f}(k-\ell,\xi-\eta)\widehat{g}(\ell,\eta)\,d\eta\Big)^2\,d\xi$$

$$\leq C_s\sum_{k\in\mathbb{Z}}\int_{\mathbb{R}}\big(\Lambda_t^b(k,\xi)\big)^2\Big(\sum_{\ell\in\mathbb{Z}}\big(|k-\ell|^s+|\ell|^s\big)$$

$$\times\int_{\mathbb{R}}\widehat{f}(k-\ell,\xi-\eta)\widehat{g}(\ell,\eta)\,d\eta\Big)^2\,d\xi$$

$$\leq C_s\big(\|\Lambda_t^b(g|D_x|^sf)\|_{L^2}^2+\|\Lambda_t^b(f|D_x|^sg)\|_{L^2}^2\big),$$

from which and Lemma 3.1, we deduce (3.1).

Recall that we now assume that $\mu=\nu$. The Fourier multiplier operator $\mathcal{M}$ employed here is defined in (1.14). We also recall the projectors onto the horizontal zeroth mode and the non-zeroth modes defined in (1.10).

Proof of Theorem 1.3. Applying Λ_t^b to (1.3) and invoking the properties of Λ_t^b in Lemma 3.1, we have

$$\begin{cases} \partial_t\Lambda_t^b\omega + y\partial_x\Lambda_t^b\omega - \nu\Delta\Lambda_t^b\omega + \Lambda_t^b\big((\mathbf{u}\cdot\nabla)\omega\big) = \partial_x\Lambda_t^b\theta, \\ \partial_t\Lambda_t^b\theta + y\partial_x\Lambda_t^b\theta - \nu\Delta\Lambda_t^b\theta + \Lambda_t^b\big((\mathbf{u}\cdot\nabla)\theta\big) = 0. \end{cases} \tag{3.2}$$

We then multiply the equations above by $\overline{\mathcal{M}\Lambda_t^b\omega}$ and $\overline{\mathcal{M}\Lambda_t^b\theta}$, respectively, and integrate over $\mathbb{T}\times\mathbb{R}$. The combination $y\partial_x-\nu\Delta$ creates the enhanced dissipation. As we explained in the introduction, we do not need the full Laplacian dissipation and the vertical dissipation is sufficient. By (1.15), we have

$$2\mathrm{Re}\langle y\partial_x f,\mathcal{M}f\rangle_{L^2} = \langle[\mathcal{M},y\partial_x]f,f\rangle_{L^2} = \langle(k\partial_\xi\mathcal{M})(D)f,f\rangle_{L^2}$$

since $\mathcal{M}$ is self-adjoint and $y\partial_x$ is skew-adjoint. Invoking the equality above, we have

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b\theta\|_{L^2}^2 + 2\nu\|\nabla\sqrt{\mathcal{M}}\Lambda_t^b\theta\|_{L^2}^2 + \langle(k\partial_\xi\mathcal{M})(D)\Lambda_t^b\theta,\Lambda_t^b\theta\rangle_{L^2}$$
$$+ 2\mathrm{Re}\langle\Lambda_t^b\big(\mathbf{u}\cdot\nabla\theta\big),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2} = 0. \tag{3.3}$$

Similarly,

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b\omega\|_{L^2}^2 + 2\nu\|\nabla\sqrt{\mathcal{M}}\Lambda_t^b\omega\|_{L^2}^2 + \langle(k\partial_\xi\mathcal{M})(D)\Lambda_t^b\omega,\Lambda_t^b\omega\rangle_{L^2}$$
$$+ 2\mathrm{Re}\langle\Lambda_t^b\big(\mathbf{u}\cdot\nabla\omega\big),\mathcal{M}\Lambda_t^b\omega\rangle_{L^2} = 2\mathrm{Re}\langle\partial_x\Lambda_t^b\theta,\mathcal{M}\Lambda_t^b\omega\rangle_{L^2}. \tag{3.4}$$

Taking L^2 inner product of the θ equation in (3.2) with $\mathcal{M}|D_x|^{\frac{2}{3}}\Lambda_t^b\theta$ gives

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + 2\nu\|\nabla\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2$$
$$+ \langle|D_x|^{\frac{2}{3}}(k\partial_\xi\mathcal{M})(D)\Lambda_t^b\theta,\Lambda_t^b\theta\rangle_{L^2} + 2\mathrm{Re}\langle\Lambda_t^b\big(\mathbf{u}\cdot\nabla\theta\big),|D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b\theta\rangle_{L^2} = 0, \tag{3.5}$$

According to the definition (1.14) of $\mathcal{M}$, we have

$$k\partial_\xi\mathcal{M}(k,\xi) = \nu^{\frac{1}{3}}|k|^{\frac{2}{3}}\varphi'\big(\nu^{\frac{1}{3}}|k|^{-\frac{1}{3}}\mathrm{sgn}(k)\xi\big) + \frac{1}{k^2+\xi^2} \tag{3.6}$$

for $k \neq 0$, $\xi \in \mathbb{R}$. This implies that, for $k \neq 0$, $\xi \in \mathbb{R}$,

$$2\nu(\xi^2 + k^2)\mathcal{M}(k, \xi) + k\partial_\xi \mathcal{M}(k, \xi) \geq \nu(\xi^2 + k^2) + \frac{1}{4}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}} + \frac{1}{\xi^2 + k^2}.$$

Therefore,

$$2\nu\|\nabla\sqrt{\mathcal{M}}f\|_{L^2}^2 + \langle(k\partial_\xi \mathcal{M})(D)f, f\rangle_{L^2}$$
$$\geq \nu\|\nabla f\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}f\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}f_{\neq}\|_{L^2}^2,$$

where $f_{\neq}$ is defined by (1.10). (3.4), (3.3) and (3.5) then becomes

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b\omega\|_{L^2}^2 + \nu\|\nabla\Lambda_t^b\omega\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}^2$$
$$\leq 2\text{Re}\underbrace{\langle\partial_x\Lambda_t^b\theta, \mathcal{M}\Lambda_t^b\omega\rangle_{L^2}}_{=I_1} - 2\text{Re}\underbrace{\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\omega), \mathcal{M}\Lambda_t^b\omega\rangle_{L^2}}_{=I_2}, \tag{3.7}$$

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b\theta\|_{L^2}^2 + \nu\|\nabla\Lambda_t^b\theta\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2$$
$$\leq -2\text{Re}\underbrace{\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\theta), \mathcal{M}\Lambda_t^b\theta\rangle_{L^2}}_{=I_3} \tag{3.8}$$

and

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \nu\|\nabla|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2$$
$$+ \|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 \leq -2\text{Re}\underbrace{\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\theta), |D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}}_{=I_4}. \tag{3.9}$$

Using the L^2-boundedness of $\mathcal{M}$, we have

$$|I_1| = |\langle\partial_x\Lambda_t^b\theta, \mathcal{M}\Lambda_t^b\omega\rangle_{L^2}| \leq \||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}. \tag{3.10}$$

- Estimates for I_2 and I_3.
  Recall that the velocity field $\mathbf{u}$ is given by the Biot-Savart law

$$\mathbf{u} = -\nabla^{\perp}(-\Delta)^{-1}\omega = \begin{pmatrix} \partial_y(-\Delta)^{-1}\omega \\ -\partial_x(-\Delta)^{-1}\omega \end{pmatrix} =: \begin{pmatrix} u \\ v \end{pmatrix}.$$

According to (1.10), $\mathbf{u}$ can be decomposed into $\mathbf{u}_0$ and $\mathbf{u}_{\neq}$,

$$\mathbf{u}_0 = \mathbb{P}_0\mathbf{u} = \begin{pmatrix} u_0 \\ 0 \end{pmatrix}, \quad \text{with } u_0 = \partial_y(-\partial_y^2)^{-1}\omega_0,$$

$$\mathbf{u}_{\neq} = \mathbb{P}_{\neq}\mathbf{u} = -\nabla^{\perp}(-\Delta)^{-1}\omega_{\neq} = \begin{pmatrix} \partial_y(-\Delta)^{-1}\omega_{\neq} \\ -\partial_x(-\Delta)^{-1}\omega_{\neq} \end{pmatrix} = \begin{pmatrix} u_{\neq} \\ v_{\neq} \end{pmatrix}. \tag{3.11}$$

Therefore we can write

$$I_2 = \langle\Lambda_t^b(\mathbf{u}\cdot\nabla\omega), \mathcal{M}\Lambda_t^b\omega\rangle_{L^2} = I_{21} + I_{22},$$

with

$$I_{21} = \langle\Lambda_t^b(\mathbf{u}_{\neq}\cdot\nabla\omega), \mathcal{M}\Lambda_t^b\omega\rangle_{L^2}, \qquad I_{22} = \langle\Lambda_t^b(\mathbf{u}_0\cdot\nabla\omega), \mathcal{M}\Lambda_t^b\omega\rangle_{L^2}.$$

Using the boundedness of $\mathcal{M}$ and Lemma 3.1, we have for $b > 1$,

$$|I_{21}| \leq \|\Lambda_t^b(\mathbf{u}_{\neq} \cdot \nabla \omega)\|_{L^2} \|\Lambda_t^b \omega\|_{L^2} \leq C \|\Lambda_t^b \mathbf{u}_{\neq}\|_{L^2} \|\nabla \Lambda_t^b \omega\|_{L^2} \|\Lambda_t^b \omega\|_{L^2}.$$

By (3.11),

$$\|\Lambda_t^b \mathbf{u}_{\neq}\|_{L^2} \leq \|\nabla^\perp (-\Delta)^{-1} \Lambda_t^b \omega_{\neq}\|_{L^2} \leq \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \omega_{\neq}\|_{L^2},$$

Therefore, for $b > 1$,

$$|I_{21}| \leq C \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \omega_{\neq}\|_{L^2} \|\nabla \Lambda_t^b \omega\|_{L^2} \|\Lambda_t^b \omega\|_{L^2}.$$

The key point is to bound I_{22}. To simplify the notation, we write $\mathcal{M}_t^b = \sqrt{\mathcal{M}} \Lambda_t^b$ or

$$\mathcal{M}_t^b(k, \xi) := \sqrt{\mathcal{M}(k, \xi)} \, \Lambda_t^b(k, \xi) = \sqrt{\mathcal{M}(k, \xi)} \left(1 + k^2 + (\xi + kt)^2\right)^{b/2}. \qquad (3.12)$$

It follows from (3.11) that $\mathbf{u}_0 \cdot \nabla \omega = u_0 \partial_x \omega = u_0 \partial_x \omega_{\neq}$ since ω_0 is independent of x. Therefore,

$$I_{22} = \langle \Lambda_t^b(\mathbf{u}_0 \cdot \nabla \omega), \mathcal{M} \Lambda_t^b \omega \rangle_{L^2} = \langle \mathcal{M}_t^b(u_0 \partial_x \omega_{\neq}), \mathcal{M}_t^b \omega \rangle_{L^2},$$

Due to the cancellations

$$\langle \mathcal{M}_t^b(u_0 \partial_x \omega_{\neq}), \mathcal{M}_t^b \omega_0 \rangle_{L^2} = 0,$$
$$\langle u_0 \partial_x (\mathcal{M}_t^b \omega_{\neq}), \mathcal{M}_t^b \omega_{\neq} \rangle_{L^2} = 0,$$

we have

$$I_{22} = \langle \mathcal{M}_t^b(u_0 \partial_x \omega_{\neq}), \mathcal{M}_t^b \omega_{\neq} \rangle_{L^2}$$
$$= \langle \mathcal{M}_t^b(u_0 \partial_x \omega_{\neq}) - u_0 \partial_x (\mathcal{M}_t^b \omega_{\neq}), \mathcal{M}_t^b \omega_{\neq} \rangle_{L^2}.$$

By Plancherel's theorem,

$$I_{22} = \sum_{k \neq 0} \iint \left(\mathcal{M}_t^b(k, \xi) - \mathcal{M}_t^b(k, \xi - \eta)\right) \widehat{u}(0, \eta) ik \widehat{\omega}(k, \xi - \eta) \mathcal{M}_t^b(k, \xi) \overline{\widehat{\omega}(k, \xi)} d\xi d\eta$$

$$= -\sum_{k \neq 0} \iint \left(\mathcal{M}_t^b(k, \xi) - \mathcal{M}_t^b(k, \xi - \eta)\right) \frac{1}{\eta} \widehat{\omega}(0, \eta) k \widehat{\omega}(k, \xi - \eta) \mathcal{M}_t^b(k, \xi) \overline{\widehat{\omega}(k, \xi)} d\xi d\eta.$$

By Taylor's formula,

$$\left|\mathcal{M}_t^b(k, \xi) - \mathcal{M}_t^b(k, \xi - \eta)\right| \leq \int_0^1 |\partial_\xi \mathcal{M}_t^b(k, \xi - s\eta)| |\eta| ds.$$

Using the explicit expression of $\mathcal{M}_t^b$ we deduce that

$$|\partial_\xi \mathcal{M}_t^b(k, \xi)| \leq C \left(\nu^{\frac{1}{3}} |k|^{-\frac{1}{3}} + \frac{1}{|k|}\right) \left(1 + k^2 + (\xi + kt)^2\right)^{\frac{b}{2}}. \qquad (3.13)$$

Therefore,

$$|I_{22}| \leq \sum_{k \neq 0} C(\nu^{\frac{1}{3}} |k|^{\frac{2}{3}} + 1) \iint \left(\left(1 + k^2 + (\xi + kt)^2\right)^{\frac{b}{2}} + \left(1 + k^2 + (\xi - \eta + kt)^2\right)^{\frac{b}{2}}\right)$$

$$\times |\widehat{\omega}(0, \eta)| |\widehat{\omega}(k, \xi - \eta)| \mathcal{M}_t^b(k, \xi) |\widehat{\omega}(k, \xi)| d\xi d\eta$$

$$\leq C \nu^{\frac{1}{3}} \|\Lambda_t^b \omega_0\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L^2}^2 + C \|\Lambda_t^b \omega_0\|_{L^2} \|\Lambda_t^b \omega_{\neq}\|_{L^2}^2.$$

Consequently,

$$
\begin{aligned}
|I_2| \leq & C \, \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\nabla\Lambda_t^b\omega\|_{L^2}\|\Lambda_t^b\omega\|_{L^2} \\
& + C\,\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}^2 + C\,\|\Lambda_t^b\omega_0\|_{L^2}\|\Lambda_t^b\omega_{\neq}\|_{L^2}^2 \\
\leq & C \, \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\nabla\Lambda_t^b\omega\|_{L^2}\|\Lambda_t^b\omega\|_{L^2} \\
& + C\,\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}^2 \\
& + C\,\|\Lambda_t^b\omega_0\|_{L^2}\,\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\,\|\nabla\Lambda_t^b\omega\|_{L^2}.
\end{aligned}
\tag{3.14}
$$

I_3 can be bounded similarly as I_2. We write I_3 as $I_3 = I_{31} + I_{32}$ with

$$
I_{31} = \langle \Lambda_t^b(\mathbf{u}_{\neq}\cdot\nabla\theta), \mathcal{M}\Lambda_t^b\theta\rangle_{L^2}, \qquad I_{32} = \langle \Lambda_t^b(\mathbf{u}_0\cdot\nabla\theta), \mathcal{M}\Lambda_t^b\theta\rangle_{L^2}
$$

and obtain the following bound

$$
\begin{aligned}
|I_3| \leq & \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\nabla\Lambda_t^b\theta\|_{L^2}\|\Lambda_t^b\theta\|_{L^2} \\
& + C\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + C\|\Lambda_t^b\omega_0\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}^2 \\
\leq & \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\nabla\Lambda_t^b\theta\|_{L^2}\|\Lambda_t^b\theta\|_{L^2} \\
& + C\,\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 \\
& + C\,\|\Lambda_t^b\omega_0\|_{L^2}\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L^2}\,\|\nabla\Lambda_t^b\theta\|_{L^2}.
\end{aligned}
\tag{3.15}
$$

- Estimates for I_4.
  We first decompose I_4 as $I_4 = I_{41} + I_{42}$ with

$$
I_{41} = \langle \Lambda_t^b(\mathbf{u}_{\neq}\cdot\nabla\theta), |D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}, \qquad I_{42} = \langle \Lambda_t^b(\mathbf{u}_0\cdot\nabla\theta), |D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}.
$$

The estimates for I_{42} are the same as those for I_{22},

$$
|I_{42}| \leq C\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2 + C\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2.
$$

For I_{41}, we have

$$
|I_{41}| \leq \||D_x|^{\frac{1}{3}}\Lambda_t^b(\mathbf{u}_{\neq}\cdot\nabla\theta)\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}.
$$

Furthermore, it follows from (3.1) that

$$
\||D_x|^{\frac{1}{3}}\Lambda_t^b(\mathbf{u}_{\neq}\cdot\nabla\theta)\|_{L^2} \leq \|\Lambda_t^b\mathbf{u}_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\nabla\theta\|_{L^2} + \||D_x|^{\frac{1}{3}}\Lambda_t^b\mathbf{u}_{\neq}\|_{L^2}\|\Lambda_t^b\nabla\theta\|_{L^2}
$$

and

$$
\|\Lambda_t^b\mathbf{u}_{\neq}\|_{L^2} \leq \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}, \qquad \||D_x|^{\frac{1}{3}}\Lambda_t^b\mathbf{u}_{\neq}\|_{L^2} \leq \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}.
$$

Therefore, we deduce that

$$
\begin{aligned}
|I_4| \leq & C\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2 + C\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 \\
& + C\,\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\nabla\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2} \\
& + C\,\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}\|\Lambda_t^b\nabla\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}.
\end{aligned}
\tag{3.16}
$$

- The closing of the energy estimate

Inserting the upper bounds (3.10), (3.14), (3.15) and (3.16) in (3.7), (3.8) and (3.9) and integrating in time, we obtain

$$
\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)}^2 + \nu\|\nabla\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \frac{1}{8}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}^2
$$
$$
\leq 2\|\Lambda_0^b\omega^{(0)}\|_{L^2}^2 + 8\nu^{-\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + C_1\nu^{\frac{1}{3}}\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2
$$
$$
+ C_1\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2 L^2}\|\nabla\Lambda_t^b\omega\|_{L_t^2(L^2)}\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}, \tag{3.17}
$$

and

$$
\|\Lambda_t^b \theta\|_{L_t^\infty(L^2)}^2 + \nu\|\nabla\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2
$$
$$
\leq 2\|\Lambda_0^b\theta^{(0)}\|_{L^2}^2 + C_2\,\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}\|\nabla\Lambda_t^b\theta\|_{L_t^2(L^2)}\|\Lambda_t^b\theta\|_{L_t^\infty(L^2)}
$$
$$
+ C_2\nu^{\frac{1}{3}}\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \tag{3.18}
$$
$$
+ C_2\,\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\,\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}\,\|\nabla\Lambda_t^b\theta\|_{L_t^2(L^2)},
$$

and

$$
\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^\infty(L^2)}^2 + \nu\|\nabla|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2
$$
$$
+ \|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2
$$
$$
\leq 2\||D_x|^{\frac{1}{3}}\Lambda_0^b\theta^{(0)}\|_{L^2}^2 + C_3\nu^{\frac{1}{3}}\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \tag{3.19}
$$
$$
+ C_3\,\|\Lambda_t^b\omega_0\|_{L_t^\infty(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2
$$
$$
+ C_3\,\|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}\,\||D_x|^{\frac{1}{3}}\Lambda_t^b\nabla\theta\|_{L_t^2(L^2)}\,\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^\infty(L^2)}
$$
$$
+ C_3\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}\,\|\Lambda_t^b\nabla\theta\|_{L_t^2(L^2)}\,\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^\infty(L^2)}.
$$

The *a priori* bounds in (3.17), (3.18) and (3.19) allow us to prove Theorem 1.3 through the bootstrap argument. We recall the assumptions on the initial data $(\omega^{(0)}, \theta^{(0)})$,

$$
\|\omega^{(0)}\|_{H^b} \leq \varepsilon\nu^\beta, \qquad \|\theta^{(0)}\|_{H^b} \leq \varepsilon\nu^\alpha, \qquad \||D_x|^{\frac{1}{3}}\theta^{(0)}\|_{H^b} \leq \varepsilon\nu^\delta, \tag{3.20}
$$

where $\varepsilon > 0$ is sufficiently small and

$$
\beta \geq \frac{1}{2}, \quad \delta \geq \beta + \frac{1}{3}, \quad \alpha \geq \delta - \beta + \frac{2}{3}. \tag{3.21}
$$

To apply the bootstrap argument, we make the ansatz that, for $T \leq \infty$, the solution of (1.3) obeys

$$
\|\Lambda_t^b\omega\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla\Lambda_t^b\omega\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_T^2(L^2)}
$$
$$
+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_T^2(L^2)} \leq C\varepsilon\nu^\beta, \tag{3.22}
$$
$$
\|\Lambda_t^b\theta\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla\Lambda_t^b\theta\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_T^2(L^2)}
$$
$$
+ \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L_T^2(L^2)} \leq C\varepsilon\nu^\alpha, \tag{3.23}
$$
$$
\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}}\|\nabla|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_T^2(L^2)}
$$

$$+\|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L_T^2(L^2)} \leq \widetilde{C}\varepsilon\nu^{\delta}. \quad (3.24)$$

We then show that (3.22), (3.23) and (3.24) actually hold with C replaced by $C/2$ and $\widetilde{C}$ by $\widetilde{C}/2$. In fact, if we insert the initial condition (3.20) and the ansatz (3.22), (3.23) and (3.24) in (3.17), (3.18) and (3.19), we find

$$\|\Lambda_t^b\omega\|_{L_t^{\infty}(L^2)}^2 + \nu\|\nabla\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \frac{1}{8}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}^2$$
$$\leq 2\varepsilon^2\nu^{2\beta} + 8\widetilde{C}^2\varepsilon^2\nu^{2\delta-\frac{2}{3}} + C_1 C^3\varepsilon^3(\nu^{3\beta-\frac{1}{3}} + \nu^{3\beta-\frac{1}{2}}),$$

$$\|\Lambda_t^b\theta\|_{L_t^{\infty}(L^2)}^2 + \nu\|\nabla\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2$$
$$\leq 2\varepsilon^2\nu^{2\alpha} + C_2 C^3\varepsilon^3(3\nu^{\beta+2\alpha-\frac{1}{2}} + \nu^{\beta+2\alpha-\frac{1}{3}}),$$

$$\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^{\infty}(L^2)}^2 + \nu\|\nabla|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2$$
$$+\|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2$$
$$\leq 2\varepsilon^2\nu^{2\delta} + C_3 C\widetilde{C}\varepsilon^3(2\widetilde{C}\nu^{\beta+2\delta} + C\nu^{\beta+2\alpha-\frac{1}{3}} + 2C\nu^{\beta+\alpha+\delta-\frac{2}{3}}).$$

If we invoke (3.21) and choose

$$\widetilde{C} \geq 8, \quad C \geq 32\widetilde{C}, \quad \varepsilon = \min\left(\frac{1}{128C_1C}, \frac{1}{128C_2C}, \frac{\widetilde{C}}{64C_3C}\right),$$

then the inequalities (3.22)-(3.23) hold with C replaced by $C/2$ and (3.24) holds with $\widetilde{C}$ replaced by $\widetilde{C}/2$. This completes the proof of Theorem 1.3. $\qquad\square$

4. PROOF OF THEOREM 1.4

This section proves the nonlinear stability result stated in Theorem 1.4. We recall that the Boussinesq system concerned here has only vertical dissipation, namely

$$\begin{cases} \partial_t\omega + y\partial_x\omega + (\mathbf{u}\cdot\nabla)\omega = \nu\partial_{yy}\omega + \partial_x\theta, \\ \partial_t\theta + y\partial_x\theta + (\mathbf{u}\cdot\nabla)\theta = \nu\partial_{yy}\theta, \\ \mathbf{u} = -\nabla^{\perp}(-\Delta)^{-1}\omega, \\ \omega(x,0) = \omega^{(0)}, \quad \theta(x,0) = \theta^{(0)}. \end{cases} \quad (4.1)$$

The proof is much more involved than the full dissipation case. The framework is still the bootstrap argument, but it is now much more difficult to prove the desired *a priori* bounds due to the lack of horizontal dissipation. The Fourier multiplier operator is the same as that is designed for the full dissipation case, but the nonlinear terms are now difficult to control. Various techniques are combined to achieve suitable upper bounds. The quantities are decomposed into horizontal zeroth mode and the non-zeroth modes to distinguish their different behaviors. Commutator estimates are employed to shift derivatives. In addition, the frequency space is divided into different subdomains to facilitate cancellations and derivative distribution.

Proof of Theorem 1.4. Applying the operator Λ_t^b to (4.1) and making use of the fact that Λ_t^b commutes with $\partial_t + y\partial_x$, we obtain

$$\begin{cases} \partial_t \Lambda_t^b \omega + y\partial_x \Lambda_t^b \omega - \nu\partial_y^2 \Lambda_t^b \omega + \Lambda_t^b\big((\mathbf{u}\cdot\nabla)\omega\big) = \partial_x \Lambda_t^b \theta, \\ \partial_t \Lambda_t^b \theta + y\partial_x \Lambda_t^b \theta - \nu\partial_y^2 \Lambda_t^b \theta + \Lambda_t^b\big((\mathbf{u}\cdot\nabla)\theta\big) = 0. \end{cases} \tag{4.2}$$

We then take the scalar product of the equations with $\mathcal{M}\Lambda_t^b \omega$ and $\mathcal{M}\Lambda_t^b \theta$, respectively, where the Fourier multiplier $\mathcal{M}$ is defined in (1.14). Using (1.15), due to the fact that $\mathcal{M}$ is self-adjoint and $y\partial_x$ is skew-adjoint,

$$2\mathrm{Re}\langle y\partial_x f, \mathcal{M}f\rangle_{L^2} = \langle [\mathcal{M}, y\partial_x]f, f\rangle_{L^2} = \langle (k\partial_\xi \mathcal{M})(D)f, f\rangle_{L^2}.$$

Invoking this equality, we have

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b \omega\|_{L^2}^2 + 2\nu\|D_y\sqrt{\mathcal{M}}\Lambda_t^b \omega\|_{L^2}^2 + \langle (k\partial_\xi \mathcal{M})(D)\Lambda_t^b \omega, \Lambda_t^b \omega\rangle_{L^2}$$
$$+ 2\mathrm{Re}\langle \Lambda_t^b\big(\mathbf{u}\cdot\nabla\omega\big), \mathcal{M}\Lambda_t^b \omega\rangle_{L^2} = 2\mathrm{Re}\langle \partial_x \Lambda_t^b \theta, \mathcal{M}\Lambda_t^b \omega\rangle_{L^2} \tag{4.3}$$

and

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b \theta\|_{L^2}^2 + 2\nu\|D_y\sqrt{\mathcal{M}}\Lambda_t^b \theta\|_{L^2}^2 + \langle (k\partial_\xi \mathcal{M})(D)\Lambda_t^b \theta, \Lambda_t^b \theta\rangle_{L^2}$$
$$+ 2\mathrm{Re}\langle \Lambda_t^b\big(\mathbf{u}\cdot\nabla\theta\big), \mathcal{M}\Lambda_t^b \theta\rangle_{L^2} = 0. \tag{4.4}$$

Similarly, taking the L^2-inner product of $\mathcal{M}|D_x|^{\frac{2}{3}}\Lambda_t^b \theta$ with the θ equation of (4.2) gives

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L^2}^2 + 2\nu\|D_y\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b \theta\|_{L^2}^2$$
$$+ \langle |D_x|^{\frac{2}{3}}(k\partial_\xi \mathcal{M})(D)\Lambda_t^b \theta, \Lambda_t^b \theta\rangle_{L^2} + 2\mathrm{Re}\langle \Lambda_t^b\big(\mathbf{u}\cdot\nabla\theta\big), |D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b \theta\rangle_{L^2} = 0. \tag{4.5}$$

Using (3.6) and the properties of the function φ, especially $\varphi'(\tau) = \frac{1}{4}$ when $|\tau| \le 1$, we have, for $k \ne 0$, $\xi \in \mathbb{R}$

$$2\nu\xi^2 \mathcal{M}(k,\xi) + k\partial_\xi \mathcal{M}(k,\xi) \ge \nu\xi^2 + \frac{1}{4}\nu^{\frac{1}{3}}|k|^{\frac{2}{3}} + \frac{1}{\xi^2 + k^2}.$$

As a consequence, it comes out

$$2\nu\|D_y\sqrt{\mathcal{M}}f\|_{L^2}^2 + \langle (k\partial_\xi \mathcal{M})(D)f, f\rangle_{L^2}$$
$$\ge \nu\|D_y f\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}f\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}f_{\ne}\|_{L^2}^2, \tag{4.6}$$

where $f_{\ne}$ is given in (1.10). Inserting (4.6) into (4.3), (4.4), (4.5) yields

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b \omega\|_{L^2}^2 + \nu\|D_y\Lambda_t^b \omega\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b \omega\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b \omega_{\ne}\|_{L^2}^2$$
$$\le 2\mathrm{Re}\underbrace{\langle \partial_x \Lambda_t^b \theta, \mathcal{M}\Lambda_t^b \omega\rangle_{L^2}}_{=I_1} - 2\mathrm{Re}\underbrace{\langle \Lambda_t^b\big(\mathbf{u}\cdot\nabla\omega\big), \mathcal{M}\Lambda_t^b \omega\rangle_{L^2}}_{=I_2}, \tag{4.7}$$

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}\Lambda_t^b\theta\|_{L^2}^2 + \nu\|D_y\Lambda_t^b\theta\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2$$
$$\leq -2\mathrm{Re}\underbrace{\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\theta),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}}_{=I_3} \qquad (4.8)$$

and

$$\frac{d}{dt}\|\sqrt{\mathcal{M}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \nu\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2$$
$$+ \|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 \leq -2\mathrm{Re}\underbrace{\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\theta),|D_x|^{\frac{2}{3}}\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}}_{=I_4}. \qquad (4.9)$$

The term I_1 is easy to deal with, using the L^2-boundedness of $\mathcal{M}$, we have

$$|I_1| = |\langle\partial_x\Lambda_t^b\theta,\mathcal{M}\Lambda_t^b\omega\rangle_{L^2}| \leq \||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}$$
$$\leq \frac{1}{16}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}^2 + 8\nu^{-\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2. \qquad (4.10)$$

- Estimates for I_2 and I_3.

  The terms I_2 and I_3 have the same structure so that we only estimate I_3. In view of (3.11), we write

$$\mathbf{u}\cdot\nabla\theta = u_0\partial_x\theta + \partial_y(-\Delta)^{-1}\omega_{\neq}\partial_x\theta - \partial_x(-\Delta)^{-1}\omega_{\neq}\partial_y\theta.$$

Correspondingly, we decompose I_3 as

$$I_3 = \langle\Lambda_t^b(\mathbf{u}\cdot\nabla\theta),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2} = I_{31} + I_{32} + I_{33}, \quad \text{with}$$
$$I_{31} := -\langle\Lambda_t^b(\partial_x(-\Delta)^{-1}\omega_{\neq}\partial_y\theta),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2},$$
$$I_{32} := \langle\Lambda_t^b(u_0\partial_x\theta),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2},$$
$$I_{33} := \langle\Lambda_t^b(\partial_y(-\Delta)^{-1}\omega_{\neq}\partial_x\theta),\mathcal{M}\Lambda_t^b\theta\rangle_{L^2}.$$

For the term I_{31}, we have

$$I_{31} \leq \|\Lambda_t^b(\partial_x(-\Delta)^{-1}\omega_{\neq}\partial_y\theta)\|_{L^2}\|\Lambda_t^b\theta\|_{L^2}$$
$$\leq \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta\|_{L^2}\|\Lambda_t^b\theta\|_{L^2}. \qquad (4.11)$$

The estimates for I_{32} and I_{33} are much more elaborate since we only have $\frac{1}{3}$-derivative enhanced dissipation in the x-direction, which is not enough to control $\partial_x\theta$ directly. By (1.10), we write $\theta = \theta_0 + \theta_{\neq}$. Since θ_0 is independent of x, we have $\partial_x\theta_0 = 0$ and the cancellations

$$\langle\mathcal{M}_t^b(u_0\partial_x\theta_{\neq}),\mathcal{M}_t^b\theta_0\rangle_{L^2} = 0, \qquad \langle u_0\partial_x(\mathcal{M}_t^b\theta_{\neq}),\mathcal{M}_t^b\theta_{\neq}\rangle_{L^2} = 0,$$

for $\mathcal{M}_t^b = \sqrt{\mathcal{M}}\Lambda_t^b$ with symbol $\mathcal{M}_t^b(k,\xi)$ given by (3.12). Therefore,

$$I_{32} = \langle\mathcal{M}_t^b(u_0\partial_x\theta_{\neq}),\mathcal{M}_t^b\theta_{\neq}\rangle_{L^2}$$
$$= \langle\mathcal{M}_t^b(u_0\partial_x\theta_{\neq}) - u_0\partial_x(\mathcal{M}_t^b\theta_{\neq}),\mathcal{M}_t^b\theta_{\neq}\rangle_{L^2}.$$

Using Plancherel's theorem, we have

$$I_{32} = \sum_{k \neq 0} \iint \left(\mathcal{M}_t^b(k,\xi) - \mathcal{M}_t^b(k,\xi-\eta)\right)\widehat{u}(0,\eta)ik\widehat{\theta}_{\neq}(k,\xi-\eta)\mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta$$

$$= -\sum_{k \neq 0} \iint \left(\mathcal{M}_t^b(k,\xi) - \mathcal{M}_t^b(k,\xi-\eta)\right)\frac{1}{\eta}\widehat{\omega}(0,\eta)k\widehat{\theta}_{\neq}(k,\xi-\eta)\mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta,$$

where we used $\widehat{u}(0,\eta) = i\eta^{-1}\widehat{\omega}(0,\eta)$ by (3.11). Therefore, in view of (3.13), we get, by applying Young's convolution inequality, that

$$|I_{32}| \leq \sum_{k \neq 0} C(\nu^{\frac{1}{3}}|k|^{\frac{2}{3}} + 1) \iint \left(\Lambda_t^b(k,\xi-\eta) + \Lambda_t^b(0,\eta)\right)|\widehat{\omega}(0,\eta)|$$

$$\times |\widehat{\theta}_{\neq}(k,\xi-\eta)|\Lambda_t^b(k,\xi)|\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta$$

$$\leq C\big(\nu^{\frac{1}{3}}\|\widehat{\omega}_0\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\|\widehat{|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2} \tag{4.12}$$

$$+ \|\widehat{\omega}_0\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b\omega_0\|_{L^2}\|\widehat{\theta}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\big)$$

$$\leq C\nu^{\frac{1}{3}}\|\Lambda_t^b\omega_0\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}^2 + C\|\Lambda_t^b\omega_0\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}^2.$$

Due to $\text{div } \mathbf{u}_{\neq} = 0$, we have the cancellation

$$\langle \mathbf{u}_{\neq} \cdot \nabla(\mathcal{M}_t^b\theta), \mathcal{M}_t^b\theta\rangle_{L^2} = 0$$

and we can rewrite

$$I_{33} = \underbrace{\langle \mathcal{M}_t^b(u_{\neq}\partial_x\theta) - u_{\neq}\partial_x(\mathcal{M}_t^b\theta), \mathcal{M}_t^b\theta\rangle_{L^2}}_{=:J} - \underbrace{\langle v_{\neq}\partial_y(\mathcal{M}_t^b\theta), \mathcal{M}_t^b\theta\rangle_{L^2}}_{=:J'}.$$

The term J' is easy to control

$$|J'| \leq \|v_{\neq}\|_{L^\infty}\|D_y\mathcal{M}_t^b\theta\|_{L^2}\|\mathcal{M}_t^b\theta\|_{L^2}$$

$$\leq \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta\|_{L^2}\|\Lambda_t^b\theta\|_{L^2}. \tag{4.13}$$

It remains to estimate the term J. Noticing that $\partial_x\theta_0 = \partial_x(\mathcal{M}_t^b\theta_0) = 0$, we can write

$$J = \langle \mathcal{M}_t^b(u_{\neq}\partial_x\theta_{\neq}) - u_{\neq}\partial_x(\mathcal{M}_t^b\theta_{\neq}), \mathcal{M}_t^b\theta\rangle_{L^2} = J_1 + J_2 \text{with}$$

$$J_1 := \langle \mathcal{M}_t^b(u_{\neq}\partial_x\theta_{\neq}) - u_{\neq}\partial_x(\mathcal{M}_t^b\theta_{\neq}), \mathcal{M}_t^b\theta_{\neq}\rangle_{L^2},$$

$$J_2 := \langle \mathcal{M}_t^b(u_{\neq}\partial_x\theta_{\neq}) - u_{\neq}\partial_x(\mathcal{M}_t^b\theta_{\neq}), \mathcal{M}_t^b\theta_0\rangle_{L^2}.$$

By Plancherel's theorem,

$$J_1 = \sum_{k,l} \iint \left(\mathcal{M}_t^b(k,\xi) - \mathcal{M}_t^b(k-l,\xi-\eta)\right)$$

$$\times \widehat{u}_{\neq}(l,\eta) \cdot i(k-l)\widehat{\theta}_{\neq}(k-l,\xi-\eta) \cdot \mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta$$

$$= -\sum_{\substack{k \neq 0, l \neq 0 \\ k-l \neq 0}} \iint \left(\mathcal{M}_t^b(k,\xi) - \mathcal{M}_t^b(k-l,\xi-\eta)\right)$$

$$\times \frac{\eta(k-l)}{l^2+\eta^2}\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\cdot\mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta,$$

where in the last equality we used $\widehat{u}_{\neq}(l,\eta)=i\eta(l^2+\eta^2)^{-1}\widehat{\omega}_{\neq}(l,\eta)$ by (3.11). In order to estimate J_1, the idea is to use Taylor's formula for $\mathcal{M}_t^b(k,\xi)-\mathcal{M}_t^b(k-l,\xi-\eta)$ as in the estimates of I_{32}. However, $\mathcal{M}(k,\xi)$ and $\mathcal{M}_t^b(k,\xi)$ are not smooth at $k=0$. We then have to divide into four different cases:

$$A_1=\{k>0,k-l>0\}, \quad A_2=\{k<0,k-l<0\},$$
$$A_3=\{k>0,k-l<0\}, \quad A_4=\{k<0,k-l>0\} \tag{4.14}$$

and denote by

$$J_{1i} := -\sum_{(k,l)\in A_i}\iint \left(\mathcal{M}_t^b(k,\xi)-\mathcal{M}_t^b(k-l,\xi-\eta)\right)$$

$$\times\frac{\eta(k-l)}{l^2+\eta^2}\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\cdot\mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta.$$

We first estimate J_{11} and J_{12}. When $k>0,k-l>0$, we use Taylor's formula,

$$|\mathcal{M}_t^b(k,\xi)-\mathcal{M}_t^b(k-l,\xi-\eta)| \leq \int_0^1|\partial_\xi\mathcal{M}_t^b(k-sl,\xi-s\eta)||\eta|ds$$
$$+\int_0^1|\partial_k\mathcal{M}_t^b(k-sl,\xi-s\eta)||l|ds.$$

A direct computation gives

$$|\partial_k\Lambda_t^b(k,\xi)|\leq C\Lambda_t^{b-2}(k,\xi)(|k|+|\xi+kt||t|), \qquad |t|\leq\frac{1}{|k|}\left(|\xi|+\Lambda_t(k,\xi)\right),$$

which implies

$$|\partial_k\mathcal{M}_t^b(k,\xi)|\leq\left(\frac{1}{k}+\frac{|\xi|}{k^2}\right)\Lambda_t^b(k,\xi) \quad \text{for } k>0.$$

Together with (3.13), we obtain

$$|\mathcal{M}_t^b(k,\xi)-\mathcal{M}_t^b(k-l,\xi-\eta)|$$
$$\leq\int_0^1\left(\frac{\nu^{\frac{1}{3}}|\eta|}{(k-sl)^{\frac{1}{3}}}+\frac{|\eta|+|l|}{k-sl}+\frac{|\xi-s\eta||l|}{(k-sl)^2}\right)\Lambda_t^b(k-sl,\xi-s\eta)ds$$
$$\lesssim\left(\frac{\nu^{\frac{1}{3}}|\eta|}{\min(k-l,k)^{\frac{1}{3}}}+\frac{|\eta|+|l|}{\min(k-l,k)}+\frac{(|\xi|+|\xi-\eta|)|l|}{(k-l)k}\right)\left(\Lambda_t^b(k-l,\xi-\eta)+\Lambda_t^b(l,\eta)\right).$$

Here and in all that follows, $a\lesssim b$ always means that there is a uniform constant C so that $a\leq Cb$. Therefore, by the convolution inequality,

$$|J_{11}^{(1)}|:=\Big|\sum_{\substack{(k,l)\in A_1\\l>0}}\iint\left(\mathcal{M}_t^b(k,\xi)-\mathcal{M}_t^b(k-l,\xi-\eta)\right)$$

$$\times\frac{\eta(k-l)}{l^2+\eta^2}\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\cdot\mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta\Big|$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l>0}} \iint \Big(\frac{\nu^{\frac{1}{3}}|\eta|}{(k-l)^{\frac{1}{3}}} + \frac{|\eta|+|l|}{k-l} + \frac{(|\xi|+|\xi-\eta|)|l|}{(k-l)k}\Big)\big(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta)\big)$$

$$\times \frac{|\eta|(k-l)}{l^2+\eta^2}|\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\Lambda_t^b(k,\xi)\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l>0}} \iint \big(\nu^{\frac{1}{3}}(k-l)^{\frac{2}{3}} + 1 + \frac{|\xi|+|\xi-\eta|}{(l^2+\eta^2)^{\frac{1}{2}}}\big)\big(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta)\big)$$

$$\times |\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\Lambda_t^b(k,\xi)\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta,$$

from which, we infer

$$\begin{aligned}
|J_{11}^{(1)}| \quad &\lesssim \quad \|\widehat{\omega}_{\neq}\|_{L^1}\big(\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b\theta_{\neq}\|_{L^2}^2\big) \\
&\quad + \|\widehat{(-\Delta)^{-\frac{1}{2}}\omega_{\neq}}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2} \\
&\quad + \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\nu^{\frac{1}{3}}\||\widehat{D_x|^{\frac{1}{3}}\theta}_{\neq}\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2} + \|\widehat{\theta}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\quad + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\|\widehat{\theta}_{\neq}\|_{L^1}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2} + \|\widehat{D_y\theta}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\lesssim \quad \|\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2},
\end{aligned}$$

where we have used that $(k-l)^{\frac{2}{3}} \leq (k-l)^{\frac{1}{3}}k^{\frac{1}{3}}$ for $k>0, k-l>0, l>0$.

On the other hand, when $k \geq 1, l < 0$, we have the inequalities

$$\frac{k-l}{k^{\frac{1}{3}}} \leq \min\big((k-l)^{\frac{1}{3}}k^{\frac{1}{3}} + (k-l)^{\frac{1}{3}}|l|^{\frac{2}{3}}, 2(k-l)^{\frac{2}{3}}|l|^{\frac{1}{3}}\big),$$

$$\frac{k-l}{k} \leq 2\min\big((k-l)^{\frac{1}{3}}|l|^{\frac{2}{3}}, (k-l)^{\frac{2}{3}}|l|^{\frac{1}{3}}\big).$$

$J_{11}^{(2)}$ can be estimated as follows,

$$|J_{11}^{(2)}| := \Big| \sum_{\substack{(k,l)\in A_1 \\ l<0}} \iint \big(\mathcal{M}_t^b(k,\xi) - \mathcal{M}_t^b(k-l,\xi-\eta)\big)$$

$$\times \frac{\eta(k-l)}{l^2+\eta^2}\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta) \cdot \mathcal{M}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta\Big|$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l<0}} \iint \Big(\frac{\nu^{\frac{1}{3}}|\eta|}{k^{\frac{1}{3}}} + \frac{|\eta|+|l|}{k} + \frac{(|\xi|+|\xi-\eta|)|l|}{(k-l)k}\Big)\big(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta)\big)$$

$$\times \frac{|\eta|(k-l)}{l^2+\eta^2}|\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\Lambda_t^b(k,\xi)\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l<0}} \iint \Big(\big(\nu^{\frac{1}{3}}(k-l)^{\frac{1}{3}}k^{\frac{1}{3}} + |\xi|+|\xi-\eta|\big)\big(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta)\big)$$

$$+ (k-l)^{\frac{1}{3}}|l|^{\frac{2}{3}}\Lambda_t^b(k-l,\xi-\eta) + (k-l)^{\frac{2}{3}}|l|^{\frac{1}{3}}\Lambda_t^b(l,\eta)\Big)$$

$$\times |\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\Lambda_t^b(k,\xi)\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta,$$

which ensures that

$$
\begin{aligned}
|J_{11}^{(2)}| &\lesssim \nu^{\frac{1}{3}}\|\widehat{\omega}_{\neq}\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \||\widehat{|D_x|^{\frac{2}{3}}\omega_{\neq}}\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2} \\
&\quad + \|\widehat{\omega}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2} + \nu^{\frac{1}{3}}\|\Lambda_t^b\omega_{\neq}\|_{L^2}\||\widehat{D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2} \\
&\quad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||\widehat{D_x|^{\frac{2}{3}}\theta_{\neq}}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2} \\
&\quad + \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\|\widehat{\theta}_{\neq}\|_{L^1}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2} + \|\widehat{D_y\theta_{\neq}}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\quad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2},
\end{aligned}
$$

where we have used

$$
\||\widehat{D_x|^{\frac{2}{3}}\omega_{\neq}}\|_{L^1} \le \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}, \qquad \text{provided that } b > \frac{4}{3}.
$$

Combining the bounds for $J_{11}^{(1)}$ and $J_{11}^{(2)}$ yields

$$
\begin{aligned}
|J_{11}| &\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\quad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}.
\end{aligned}
$$

The term J_{12} can be treated in the same way.

To estimate J_{13} and J_{14}, we notice that, when $k > 0, k - l < 0$ or $k < 0, k - l > 0$, we have $|k - l| < |l|$ and thus

$$
\begin{aligned}
|J_{13} + J_{14}| &\le \sum_{(k,l)\in A_3\cup A_4} \iint \big(\Lambda_t^b(k - l, \xi - \eta) + \Lambda_t^b(l, \eta)\big)\frac{|\eta||k - l|}{l^2 + \eta^2}|\widehat{\omega}_{\neq}(l, \eta)| \\
&\qquad\qquad\qquad \times |\widehat{\theta}_{\neq}(k - l, \xi - \eta)\Lambda_t^b(k, \xi)\widehat{\theta}_{\neq}(k, \xi)|d\xi d\eta \\
&\lesssim \|\widehat{\omega}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\widehat{\theta}_{\neq}\|_{L^1}\|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2} \\
&\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}^2.
\end{aligned}
$$

This finishes the estimate for J_1,

$$
\begin{aligned}
|J_1| &\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta_{\neq}\|_{L^2}\big) \\
&\quad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}.
\end{aligned}
$$

To estimate J_2, we observe that

$$
\Lambda_t^b(0, \xi) \le C\big(\Lambda_t^b(l, \eta) + \Lambda_t^b(-l, \xi - \eta)\big).
$$

We thus apply Plancherel's theorem and (3.11) to write

$$
\begin{aligned}
|J_2| &= \Big|\sum_{l\neq 0} \iint \big(\mathcal{M}_t^b(0, \xi) - \mathcal{M}_t^b(-l, \xi - \eta)\big)\widehat{u}_{\neq}(l, \eta)i(-l)\widehat{\theta}_{\neq}(-l, \xi - \eta)\overline{\mathcal{M}_t^b\widehat{\theta}(0, \xi)}d\xi d\eta\Big| \\
&\lesssim \sum_{l\neq 0} \iint \big(\Lambda_t^b(-l, \xi - \eta) + \Lambda_t^b(l, \eta)\big)\frac{|l||\eta|}{l^2 + \eta^2}|\widehat{\omega}_{\neq}(l, \eta)\widehat{\theta}_{\neq}(-l, \xi - \eta)\Lambda_t^b\widehat{\theta}(0, \xi)|d\xi d\eta \\
&\lesssim \|\widehat{\omega}_{\neq}\|_{L^1}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_0\|_{L^2} + \|\widehat{\theta}_{\neq}\|_{L^1}\|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\Lambda_t^b\theta_0\|_{L^2} \\
&\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\Lambda_t^b\theta_{\neq}\|_{L^2}\|\Lambda_t^b\theta_0\|_{L^2}.
\end{aligned}
$$

Combining the bounds for J_1 and J_2, we obtain

$$|J| \lesssim \|\Lambda_t^b \omega_{\neq}\|_{L^2} \big(\||D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b \theta_{\neq}\|_{L^2} \|D_y \Lambda_t^b \theta_{\neq}\|_{L^2} + \|\Lambda_t^b \theta_{\neq}\|_{L^2} \|\Lambda_t^b \theta_0\|_{L^2}\big)$$
$$+ \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega_{\neq}\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2} \|\Lambda_t^b \theta_{\neq}\|_{L^2}.$$

Together with (4.13), we finish the estimates for I_{33}:

$$|I_{33}| \lesssim \|\Lambda_t^b \omega_{\neq}\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2}^2 + \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2} \|\Lambda_t^b \theta_{\neq}\|_{L^2}$$
$$+ \|\Lambda_t^b \omega_{\neq}\|_{L^2} \|\Lambda_t^b \theta\|_{L^2} \big(\|D_y \Lambda_t^b \theta\|_{L^2} + \|\Lambda_t^b \theta_{\neq}\|_{L^2}\big). \tag{4.15}$$

It follows from (4.11), (4.12) and (4.15) that

$$|I_3| \lesssim \|\Lambda_t^b \omega\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2}^2 + \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2} \|\Lambda_t^b \theta_{\neq}\|_{L^2}$$
$$+ \|\Lambda_t^b \omega_{\neq}\|_{L^2} \|\Lambda_t^b \theta\|_{L^2} \big(\|D_y \Lambda_t^b \theta\|_{L^2} + \|\Lambda_t^b \theta_{\neq}\|_{L^2}\big). \tag{4.16}$$

Similarly, the upper bound for I_2 is given by

$$|I_2| \lesssim \|\Lambda_t^b \omega\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L^2}^2 + \|\Lambda_t^b \omega\|_{L^2} \|\Lambda_t^b \omega_{\neq}\|_{L^2} \|D_y \Lambda_t^b \omega\|_{L^2}. \tag{4.17}$$

- Estimates for I_4.

  As in the estimates of I_3, we decompose the term I_4 as

$$I_4 = I_{41} + I_{42} + I_{43} \quad \text{with} \quad I_{41} := \langle \Lambda_t^b(v_{\neq} \partial_y \theta), |D_x|^{\frac{2}{3}} \mathcal{M} \Lambda_t^b \theta \rangle_{L^2},$$
$$I_{42} := \langle \Lambda_t^b(u_0 \partial_x \theta), |D_x|^{\frac{2}{3}} \mathcal{M} \Lambda_t^b \theta \rangle_{L^2},$$
$$I_{43} := \langle \Lambda_t^b(u_{\neq} \partial_x \theta), |D_x|^{\frac{2}{3}} \mathcal{M} \Lambda_t^b \theta \rangle_{L^2}.$$

By (3.1), one has

$$I_{41} \lesssim \||D_x|^{\frac{1}{3}} \Lambda_t^b(v_{\neq} \partial_y \theta)\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2}$$
$$\lesssim \big(\||D_x|^{\frac{1}{3}} \Lambda_t^b v_{\neq}\|_{L^2} \|D_y \Lambda_t^b \theta\|_{L^2} + \|\Lambda_t^b v_{\neq}\|_{L^2} \||D_x|^{\frac{1}{3}} D_y \Lambda_t^b \theta\|_{L^2}\big) \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2}. \tag{4.18}$$

Setting $\mathcal{N}_t^b(k, \xi) := |k|^{\frac{1}{3}} \mathcal{M}_t^b(k, \xi)$ and $\mathcal{N}_t^b$ the corresponding Fourier multiplier, we can write

$$I_{42} = \langle \mathcal{N}_t^b(u_0 \partial_x \theta_{\neq}) - u_0 \partial_x \mathcal{N}_t^b \theta_{\neq}, \mathcal{N}_t^b \theta_{\neq} \rangle_{L^2}.$$

The estimates for I_{42} are similar to those for I_{32},

$$|I_{42}| \lesssim \nu^{\frac{1}{3}} \|\Lambda_t^b \omega_0\|_{L^2} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2}^2 + \|\Lambda_t^b \omega_0\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2}^2. \tag{4.19}$$

In order to estimate the term I_{43}, we decompose it as

$$I_{43} = \langle \mathcal{N}_t^b(u_{\neq} \partial_x \theta), \mathcal{N}_t^b \theta \rangle_{L^2} = K + K' \quad \text{with}$$
$$K = \langle \mathcal{N}_t^b(u_{\neq} \partial_x \theta) - u_{\neq} \partial_x(\mathcal{N}_t^b \theta), \mathcal{N}_t^b \theta \rangle_{L^2},$$
$$K' = -\langle v_{\neq} \partial_y(\mathcal{N}_t^b \theta), \mathcal{N}_t^b \theta \rangle_{L^2}.$$

The term K' can be bounded easily,

$$|K'| \leq \|v_{\neq}\|_{L^\infty} \|\partial_y \mathcal{N}_t^b \theta\|_{L^2} \|\mathcal{N}_t^b \theta\|_{L^2}$$
$$\lesssim \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \omega_{\neq}\|_{L^2} \|D_y |D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L^2}. \tag{4.20}$$

For the term K, due to $\partial_x \theta_0 = \partial_x \mathcal{N}_t^b \theta_0 = 0$,

$$K = \langle \mathcal{N}_t^b(u_{\neq} \partial_x \theta_{\neq}) - u_{\neq} \partial_x (\mathcal{N}_t^b \theta_{\neq}), \mathcal{N}_t^b \theta_{\neq} \rangle_{L^2}.$$

By Plancherel's theorem and (3.11),

$$K = -\sum_{k,l} \iint \left(\mathcal{N}_t^b(k,\xi) - \mathcal{N}_t^b(k-l,\xi-\eta) \right) \frac{\eta(k-l)}{l^2+\eta^2} \widehat{\omega}_{\neq}(l,\eta)$$

$$\times \widehat{\theta}_{\neq}(k-l,\xi-\eta) \mathcal{N}_t^b(k,\xi) \overline{\widehat{\theta}_{\neq}(k,\xi)} d\xi d\eta$$

$$= K_1 + K_2 + K_3 + K_4,$$

where, for $i = 1,2,3,4$,

$$K_i := -\sum_{(k,l)\in A_i} \iint \left(\mathcal{N}_t^b(k,\xi) - \mathcal{N}_t^b(k-l,\xi-\eta) \right) \frac{\eta(k-l)}{l^2+\eta^2} \widehat{\omega}_{\neq}(l,\eta)$$

$$\times \widehat{\theta}_{\neq}(k-l,\xi-\eta) \mathcal{N}_t^b(k,\xi) \overline{\widehat{\theta}_{\neq}(k,\xi)} d\xi d\eta$$

with A_i being defined in (4.14). For any $k \neq 0$,

$$|\partial_\xi \mathcal{N}_t^b(k,\xi)| \lesssim (\nu^{\frac{1}{3}} + |k|^{-\frac{2}{3}}) \Lambda_t^b(k,\xi),$$

$$|\partial_k \mathcal{N}_t^b(k,\xi)| \lesssim (|k|^{-\frac{2}{3}} + |k|^{-\frac{5}{3}}|\xi|) \Lambda_t^b(k,\xi).$$

When $k > 0, k - l > 0$, using Taylor's formula, we have

$$|\mathcal{N}_t^b(k,\xi) - \mathcal{N}_t^b(k-l,\xi-\eta)|$$

$$\lesssim \left(\nu^{\frac{1}{3}}|\eta| + \frac{|l| + |\xi-\eta| + |\xi|}{\min(k-l,k)^{\frac{2}{3}}} \right) \left(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta) \right).$$

Therefore, by the convolution inequality,

$$|K_1^{(1)}| = \left| \sum_{\substack{(k,l)\in A_1 \\ l>0}} \iint \left(\mathcal{N}_t^b(k,\xi) - \mathcal{N}_t^b(k-l,\xi-\eta) \right) \frac{\eta(k-l)}{l^2+\eta^2} \widehat{\omega}_{\neq}(l,\eta) \right.$$

$$\left. \times \widehat{\theta}_{\neq}(k-l,\xi-\eta) \mathcal{N}_t^b(k,\xi) \overline{\widehat{\theta}_{\neq}(k,\xi)} d\xi d\eta \right|$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l>0}} \iint \left(\nu^{\frac{1}{3}}|\eta| + \frac{|l| + |\xi-\eta| + |\xi|}{(k-l)^{\frac{2}{3}}} \right) \left(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta) \right)$$

$$\times \frac{\eta(k-l)}{l^2+\eta^2} |\widehat{\omega}_{\neq}(l,\eta) \widehat{\theta}_{\neq}(k-l,\xi-\eta) \mathcal{N}_t^b \widehat{\theta}_{\neq}(k,\xi)| d\xi d\eta$$

$$\lesssim \sum_{\substack{(k,l)\in A_1 \\ l>0}} \iint \left(\nu^{\frac{1}{3}}(k-l) + (k-l)^{\frac{1}{3}} + (|\xi-\eta| + |\xi|)(k-l)^{\frac{1}{3}} \right)$$

$$\times \left(\Lambda_t^b(k-l,\xi-\eta) + \Lambda_t^b(l,\eta) \right) |\widehat{\omega}_{\neq}(l,\eta) \widehat{\theta}_{\neq}(k-l,\xi-\eta) \mathcal{N}_t^b \widehat{\theta}_{\neq}(k,\xi)| d\xi d\eta,$$

from which, we deduce that

$$|K_1^{(1)}| \lesssim \|\widehat{\omega}_{\neq}\|_{L^1} \left(\nu^{\frac{1}{3}} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2} \||D_x|^{\frac{1}{3}} \mathcal{N}_t^b \theta_{\neq}\|_{L^2} + \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2} \|\mathcal{N}_t^b \theta_{\neq}\|_{L^2} \right.$$

$$\left. + \||D_y|D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2} \|\mathcal{N}_t^b \theta_{\neq}\|_{L^2} + \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L^2} \|D_y \mathcal{N}_t^b \theta_{\neq}\|_{L^2} \right)$$

$$+\|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\nu^{\frac{1}{3}}\|\widehat{|D_x|^{\frac{2}{3}}\theta_{\neq}}\|_{L^1}\||D_x|^{\frac{1}{3}}\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\|\widehat{|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}$$

$$+\|\widehat{D_y|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\|\widehat{|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\|D_y\mathcal{N}_t^b\theta_{\neq}\|_{L^2}\big)$$

$$\lesssim\ \ \|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2+\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\big),$$

where we used $k-l\leq(k-l)^{\frac{2}{3}}k^{\frac{1}{3}}$ for $k>0,k-l>0,l>0$.

On the other hand, using the fact that, for $k>0,l<0$,

$$k-l\leq(k-l)^{\frac{2}{3}}(k^{\frac{1}{3}}+|l|^{\frac{1}{3}}),\quad\frac{k-l}{k^{\frac{2}{3}}}\leq2(k-l)^{\frac{2}{3}}|l|^{\frac{1}{3}},\quad\frac{k-l}{k^{\frac{2}{3}}|l|}\leq2(k-l)^{\frac{1}{3}},$$

we have

$$|K_1^{(2)}|\ =\ \Big|\sum_{\substack{(k,l)\in A_1\\ l<0}}\iint\big(\mathcal{N}_t^b(k,\xi)-\mathcal{N}_t^b(k-l,\xi-\eta)\big)\frac{\eta(k-l)}{l^2+\eta^2}\widehat{\omega}_{\neq}(l,\eta)$$

$$\times\widehat{\theta}_{\neq}(k-l,\xi-\eta)\mathcal{N}_t^b(k,\xi)\overline{\widehat{\theta}_{\neq}(k,\xi)}d\xi d\eta\Big|$$

$$\lesssim\ \sum_{\substack{(k,l)\in A_1\\ l>0}}\iint\Big(\nu^{\frac{1}{3}}|\eta|+\frac{|l|+|\xi-\eta|+|\xi|}{k^{\frac{2}{3}}}\Big)\big(\Lambda_t^b(k-l,\xi-\eta)+\Lambda_t^b(l,\eta)\big)$$

$$\times\frac{\eta(k-l)}{l^2+\eta^2}|\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\mathcal{N}_t^b\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta$$

$$\lesssim\ \sum_{\substack{(k,l)\in A_1\\ l>0}}\iint\big(\nu^{\frac{1}{3}}(k-l)^{\frac{2}{3}}k^{\frac{1}{3}}+(k-l)^{\frac{2}{3}}|l|^{\frac{1}{3}}+(|\xi-\eta|+|\xi|)(k-l)^{\frac{1}{3}}\big)$$

$$\times\big(\Lambda_t^b(k-l,\xi-\eta)+\Lambda_t^b(l,\eta)\big)|\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta}_{\neq}(k-l,\xi-\eta)\mathcal{N}_t^b\widehat{\theta}_{\neq}(k,\xi)|d\xi d\eta.$$

As a result, it comes out

$$|K_1^{(2)}|\lesssim\nu^{\frac{1}{3}}\big(\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\widehat{\omega}_{\neq}\|_{L^1}\||D_x|^{\frac{1}{3}}\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\widehat{|D_x|^{\frac{2}{3}}\theta_{\neq}}\|_{L^1}\||D_x|^{\frac{1}{3}}\mathcal{N}_t^b\theta_{\neq}\|_{L^2}\big)$$

$$+\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\widehat{|D_x|^{\frac{1}{3}}\omega_{\neq}}\|_{L^1}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|\widehat{|D_x|^{\frac{2}{3}}\theta_{\neq}}\|_{L^1}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}$$

$$+\|\widehat{\omega}_{\neq}\|_{L^1}\big(\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|D_y\mathcal{N}_t^b\theta_{\neq}\|_{L^2}\big)$$

$$+\|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\|\widehat{D_y|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}+\|\widehat{|D_x|^{\frac{1}{3}}\theta_{\neq}}\|_{L^1}\|D_y\mathcal{N}_t^b\theta_{\neq}\|_{L^2}\big)$$

$$\lesssim\|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2+\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\big)$$

$$+\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}.$$

This completes the estimates for K_1,

$$K_1\lesssim\|\Lambda_t^b\omega_{\neq}\|_{L^2}\big(\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2+\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\big)$$

$$+\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}.$$

We can estimate the term K_2 in the same way.

To estimate K_3 and K_4, we notice that when $k > 0, k - l < 0$ or $k < 0, k - l > 0$, $|k - l| < |l|$. Therefore,

$$
\begin{aligned}
|K_3 + K_4| &\leq \sum_{(k,l)\in A_3\cup A_4} \iint \left(\mathcal{N}_t^b(k,\xi) + \mathcal{N}_t^b(k-l,\xi-\eta)\right) \\
&\qquad \times |\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta_{\neq}}(k-l,\xi-\eta)\mathcal{N}_t^b\widehat{\theta_{\neq}}(k,\xi)|d\xi d\eta \\
&\lesssim \sum_{(k,l)\in A_3\cup A_4} \iint \left((|k|^{\frac{1}{3}} + |k-l|^{\frac{1}{3}})\Lambda_t^b(k-l,\xi-\eta) + |k|^{\frac{1}{3}}\Lambda_t^b(\ell,\eta)\right) \\
&\qquad \times |\widehat{\omega}_{\neq}(l,\eta)\widehat{\theta_{\neq}}(k-l,\xi-\eta)\mathcal{N}_t^b\widehat{\theta_{\neq}}(k,\xi)|d\xi d\eta,
\end{aligned}
$$

so that we have

$$
\begin{aligned}
|K_3 + K_4| &\lesssim \|\widehat{\omega}_{\neq}\|_{L^1}\left(\|\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\mathcal{N}_t^b\theta_{\neq}\|_{L^2} + \||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\|\mathcal{N}_t^b\theta_{\neq}\|_{L^2}\right) \\
&\qquad + \|\Lambda_t^b\omega_{\neq}\|_{L^2}\|\widehat{\theta}_{\neq}\|_{L^1}\||D_x|^{\frac{1}{3}}\mathcal{N}_t^b\theta_{\neq}\|_{L^2} \\
&\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\left(\|\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2} + \||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2\right).
\end{aligned}
$$

Summarizing the estimates, we achieve that

$$
\begin{aligned}
|K| &\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\left(\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}^2 + \|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\right) \\
&\qquad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L^2}.
\end{aligned}
$$

Together with (4.20), we obtain

$$
\begin{aligned}
|I_{43}| &\lesssim \|\Lambda_t^b\omega_{\neq}\|_{L^2}\left(\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}\right) \\
&\qquad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}.
\end{aligned}
\tag{4.21}
$$

Then by (4.18), (4.19) and (4.21), we finish the estimates for I_4,

$$
\begin{aligned}
|I_4| &\lesssim \|\Lambda_t^b\omega\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}^2 + \|\Lambda_t^b\omega_{\neq}\|_{L^2}\|D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2} \\
&\qquad + \||D_x|^{\frac{1}{3}}(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}\|D_y\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2} \\
&\qquad + \||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L^2}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L^2}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L^2}.
\end{aligned}
\tag{4.22}
$$

- <u>The closing of the energy estimate</u>

Integrating (4.7), (4.8) and (4.9) in time and making use of the upper bounds in (4.10), (4.16), (4.17) and (4.22), we obtain, for $b > \frac{4}{3}$,

$$
\begin{aligned}
&\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}^2 + \nu\|D_y\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \frac{1}{8}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}^2 \\
&\leq 2\|\Lambda_0^b\omega^{(0)}\|_{L^2}^2 + 8\nu^{-\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + C_1\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}^2 \\
&\qquad + C_1\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\|\Lambda_t^b\omega_{\neq}\|_{L_t^2(L^2)}\|D_y\Lambda_t^b\omega\|_{L_t^2(L^2)},
\end{aligned}
\tag{4.23}
$$

and

$$
\begin{aligned}
&\|\Lambda_t^b\theta\|_{L_t^\infty(L^2)}^2 + \nu\|D_y\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2 \\
&\leq 2\|\Lambda_0^b\theta^{(0)}\|_{L^2}^2 + C_2\|\Lambda_t^b\omega\|_{L_t^\infty(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 \\
&\qquad + C_2\||D_x|^{\frac{1}{3}}\Lambda_t^b\omega\|_{L_t^2(L^2)}\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}\|\Lambda_t^b\theta_{\neq}\|_{L_t^\infty(L^2)}
\end{aligned}
\tag{4.24}
$$

$$+ C_2 \|\Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)} \|\Lambda_t^b \theta\|_{L_t^\infty(L^2)} \big(\|D_y \Lambda_t^b \theta\|_{L_t^2(L^2)} + \|\Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)} \big)$$

and

$$
\begin{aligned}
&\||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^\infty(L^2)}^2 + \nu \|D_y |D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)}^2 + \frac{1}{4} \nu^{\frac{1}{3}} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)}^2 \\
&\quad + \|(-\Delta)^{-\frac{1}{2}} |D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)}^2 \\
&\leq 2 \||D_x|^{\frac{1}{3}} \Lambda_0^b \theta^{(0)}\|_{L^2}^2 + C_3 \|\Lambda_t^b \omega\|_{L_t^\infty(L^2)} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)}^2 \quad\quad (4.25) \\
&\quad + C_3 \|\Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)} \big(\|D_y |D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)} + \|D_y \Lambda_t^b \theta\|_{L_t^2(L^2)} \big) \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^\infty(L^2)} \\
&\quad + C_3 \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L_t^2(L^2)} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^\infty(L^2)}.
\end{aligned}
$$

With these *a priori* bounds at our disposal, our final step is to prove Theorem 1.4 via the bootstrap argument. We assume that the initial data $(\omega^{(0)}, \theta^{(0)})$ satisfies

$$\|\omega^{(0)}\|_{H^b} \leq \varepsilon \nu^\beta, \quad \|\theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\alpha, \quad \||D_x|^{\frac{1}{3}} \theta^{(0)}\|_{H^b} \leq \varepsilon \nu^\delta,$$

where $\varepsilon > 0$ is sufficiently small, and β, α, δ are constants satisfying

$$\beta \geq \frac{2}{3}, \qquad \delta \geq \beta + \frac{1}{3}, \qquad \alpha \geq \delta - \beta + \frac{2}{3}. \quad\quad (4.26)$$

The bootstrap argument starts with the ansatz that, for $T \leq \infty$, the solution (ω, θ) of (4.1) satisfies

$$
\begin{aligned}
\|\Lambda_t^b \omega\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}} \|D_y \Lambda_t^b \omega\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}} \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L_T^2(L^2)} & \\
+ \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \omega_{\neq}\|_{L_T^2(L^2)} &\leq C \varepsilon \nu^\beta, \quad\quad (4.27)
\end{aligned}
$$

$$
\begin{aligned}
\|\Lambda_t^b \theta\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}} \|D_y \Lambda_t^b \theta\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_T^2(L^2)} & \\
+ \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \theta_{\neq}\|_{L_T^2(L^2)} &\leq C \varepsilon \nu^\alpha, \quad\quad (4.28)
\end{aligned}
$$

$$
\begin{aligned}
\||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_T^\infty(L^2)} + \nu^{\frac{1}{2}} \|D_y |D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_T^2(L^2)} + \nu^{\frac{1}{6}} \||D_x|^{\frac{2}{3}} \Lambda_t^b \theta\|_{L_T^2(L^2)} & \\
+ \|(-\Delta)^{-\frac{1}{2}} |D_x|^{\frac{1}{3}} \Lambda_t^b \theta_{\neq}\|_{L_T^2(L^2)} &\leq \widetilde{C} \varepsilon \nu^\delta. \quad (4.29)
\end{aligned}
$$

The constants $\varepsilon > 0, C, \widetilde{C} > 0$ are suitably selected and will be specified later. Making use of the bounds in (4.23), (4.24) and (4.25), we show that (4.27), (4.28) and (4.29) actually holds with C replaced by $C/2$ and $\widetilde{C}$ replaced by $\widetilde{C}/2$. The bootstrap argument then implies that $T = +\infty$ and (4.27), (4.28) and (4.29) holds for all time.

In fact, if we substitute the ansatz given by (4.27), (4.28) and (4.29) in the *a priori* estimates in (4.23), (4.24) and (4.25), we find

$$
\begin{aligned}
&\|\Lambda_t^b \omega\|_{L_t^\infty(L^2)}^2 + \nu \|D_y \Lambda_t^b \omega\|_{L_t^2(L^2)}^2 + \frac{1}{8} \nu^{\frac{1}{3}} \||D_x|^{\frac{1}{3}} \Lambda_t^b \omega\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \omega_{\neq}\|_{L_t^2(L^2)}^2 \\
&\quad\quad \leq 2 \varepsilon^2 \nu^{2\beta} + 8 \widetilde{C}^2 \varepsilon^2 \nu^{2\delta - \frac{2}{3}} + C_1 C^3 \varepsilon^3 (\nu^{3\beta - \frac{1}{3}} + \nu^{3\beta - \frac{2}{3}}),
\end{aligned}
$$

$$
\begin{aligned}
&\|\Lambda_t^b \theta\|_{L_t^\infty(L^2)}^2 + \nu \|D_y \Lambda_t^b \theta\|_{L_t^2(L^2)}^2 + \frac{1}{4} \nu^{\frac{1}{3}} \||D_x|^{\frac{1}{3}} \Lambda_t^b \theta\|_{L_t^2(L^2)}^2 + \|(-\Delta)^{-\frac{1}{2}} \Lambda_t^b \theta_{\neq}\|_{L_t^2(L^2)}^2 \\
&\quad\quad \leq 2 \varepsilon^2 \nu^{2\alpha} + C_2 C^3 \varepsilon^3 (3 \nu^{\beta + 2\alpha - \frac{1}{3}} + \nu^{\beta + 2\alpha - \frac{2}{3}}),
\end{aligned}
$$

$$\||D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^\infty(L^2)}^2 + \nu\||D_y|D_x|^{\frac{1}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2 + \frac{1}{4}\nu^{\frac{1}{3}}\||D_x|^{\frac{2}{3}}\Lambda_t^b\theta\|_{L_t^2(L^2)}^2$$

$$+\|(-\Delta)^{-\frac{1}{2}}|D_x|^{\frac{1}{3}}\Lambda_t^b\theta_{\neq}\|_{L_t^2(L^2)}^2$$

$$\leq 2\varepsilon^2\nu^{2\delta} + C_3C\widetilde{C}\varepsilon^3(2\widetilde{C}\nu^{\beta+2\delta-\frac{1}{3}} + \widetilde{C}\nu^{2\delta+\beta-\frac{2}{3}} + C\nu^{\beta+\alpha+\delta-\frac{2}{3}}).$$

If we recall (4.26) and choose

$$\widetilde{C} \geq 8, \quad C \geq 32\widetilde{C}, \quad \varepsilon = \min\big(\frac{1}{128C_1C}, \frac{1}{128C_2C}, \frac{\widetilde{C}}{64C_3C}\big),$$

then (4.27-4.28) hold with C replaced by $C/2$ and (4.29) holds with $\widetilde{C}$ replaced by $\widetilde{C}/2$. This completes the proof of Theorem 1.4. $\qquad\square$

Remark 4.1. As observed in [19], once we have the bounds for the enhanced dissipation term, we can deduce an explicit exponential converge of the non-zero modes for both systems (1.3) and (1.4). Let us take the vertical dissipation case (1.4) for example. First apply the projection $\mathbb{P}_{\neq}$ to the system (4.2) and then take the scalar product of the equations with $\mathcal{M}\Lambda_t^b\omega_{\neq}$ and $\mathcal{M}\Lambda_t^b\theta_{\neq}$, respectively. Multiplying the energy inequality by a weight $e^{2\gamma\nu^{\frac{1}{3}}t}$, we obtain

$$\frac{d}{dt}\|e^{\gamma\nu^{\frac{1}{3}}t}\sqrt{\mathcal{M}}\Lambda_t^b\omega_{\neq}\|_{L^2}^2 + \nu\|e^{\gamma\nu^{\frac{1}{3}}t}D_y\Lambda_t^b\omega_{\neq}\|_{L^2}^2 + \|e^{\gamma\nu^{\frac{1}{3}}t}(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2}^2$$

$$+ \frac{1}{4}\nu^{\frac{1}{3}}\|e^{\gamma\nu^{\frac{1}{3}}t}|D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2}^2 - 2\gamma\nu^{\frac{1}{3}}\|e^{\gamma\nu^{\frac{1}{3}}t}\sqrt{\mathcal{M}}\Lambda_t^b\omega_{\neq}\|_{L^2}^2$$

$$\leq 2e^{2\gamma\nu^{\frac{1}{3}}t}\text{Re}\langle\partial_x\Lambda_t^b\theta_{\neq}, \mathcal{M}\Lambda_t^b\omega_{\neq}\rangle_{L^2} - 2e^{2\gamma\nu^{\frac{1}{3}}t}\text{Re}\langle\Lambda_t^b(\mathbf{u}\cdot\nabla\omega)_{\neq}, \mathcal{M}\Lambda_t^b\omega_{\neq}\rangle_{L^2},$$

and similar energy inequalities for $\theta_{\neq}$. By choosing $\gamma > 0$ small enough and using estimates for the nonlinear terms, we can deduce

$$\|e^{\gamma\nu^{\frac{1}{3}}t}\Lambda_t^b\omega_{\neq}\|_{L^\infty(L^2)} + \nu^{\frac{1}{2}}\|e^{\gamma\nu^{\frac{1}{3}}t}D_y\Lambda_t^b\omega_{\neq}\|_{L^2(L^2)} + \nu^{\frac{1}{6}}\|e^{\gamma\nu^{\frac{1}{3}}t}|D_x|^{\frac{1}{3}}\Lambda_t^b\omega_{\neq}\|_{L^2(L^2)}$$

$$+ \|e^{\gamma\nu^{\frac{1}{3}}t}(-\Delta)^{-\frac{1}{2}}\Lambda_t^b\omega_{\neq}\|_{L^2(L^2)} \lesssim \nu^\beta,$$

and similar bounds for $\theta_{\neq}$.

Acknowledgments

References

[1] J. Bedrossian and M. Coti Zelati, Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows, *Arch. Ration. Mech. Anal.* **224** (2017), 1161-1204.

[2] J. Bedrossian, M. Coti Zelati and V. Vicol, Vortex axisymmetrization, inviscid damping, and vorticity depletion in the linearized 2D Euler equations, *Ann. PDE* **5** (2019), Art. 4, 192 pp.

[3] J. Bedrossian, P. Germain and N. Masmoudi, On the stability threshold for the 3D Couette flow in Sobolev regularity, *Ann. Math.* **185** (2017), 541-608.

[4] J. Bedrossian, P. Germain and N. Masmoudi, Stability of the Couette flow at high Reynolds numbers in two dimensions and three dimensions, *Bull. Amer. Math. Soc. (N.S.)* **56** (2019), 373-414.

[5] J. Bedrossian, N. Masmoudi and V. Vicol, Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the two dimensional Couette flow, *Arch. Ration. Mech. Anal.* **219** (2016), 1087-1159.

[6] J. Bedrossian, V. Vicol and F. Wang, The Sobolev stability threshold for 2D shear flows near Couette, *J. Nonlinear Sci.* **28** (2018), 2051-2075.

[7] A.J. Bernoff and J.F. Lingevitch, Rapid relaxation of an axisymmetric vortex, *Phys. Fluids* **6** (1994), 3717-3723.

[8] L. Brandolese and M.E. Schonbek, Large time decay and growth for solutions of a viscous Boussinesq system, *Trans. Amer. Math. Soc.* **364** (2012), 5057-5090.

[9] A. Castro, D. Córdoba and D. Lear, On the asymptotic stability of stratified solutions for the 2D Boussinesq equations with a velocity damping term, *Math. Models Methods Appl. Sci.* **29** (2019), 1227-1277.

[10] C. Cao and J. Wu, Global regularity for the 2D anisotropic Boussinesq equations with vertical dissipation, *Arch. Ration. Mech. Anal.* **208** (2013), 985-1004.

[11] P. Constantin and C. Doering, Heat transfer in convective turbulence, *Nonlinearity* **9** (1996), 1049-1060.

[12] P. Constantin, A. Kiselev, L. Ryzhik and A. Zlatos, Diffusion and mixing in fluid flow, *Ann. Math.* **168**(2) (2008), 643-674.

[13] W. Deng, Resolvent estimates for a two-dimensional non-self-adjoint operator, *Commun. Pure Appl. Anal.* **12** (2013), 547-596.

[14] W. Deng, Pseudospectrum for Oseen vortices operators, *Int. Math Res. Not.* **2013**, 1935-1999.

[15] C. Doering and J. Gibbon, *Applied analysis of the Navier-Stokes equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1995.

[16] C. R. Doering, J. Wu, K. Zhao and X. Zheng, Long time behavior of the two-dimensional Boussinesq equations without buoyancy diffusion, *Physica D* **376/377** (2018), 144-159.

[17] L. Hörmander, Hypoelliptic second order differential equations. *Acta Math.* **119** (1967), 147-171.

[18] M. Latini and A.J. Bernoff, Transient anomalous diffusion in Poiseuille flow, *J. Fluid Mech.* **441** (2001), 399-411.

[19] K. Liss, On the Sobolev stability threshold of 3D Couette flow in a uniform magnetic field. *Comm. Math. Phys.* **377** (2020), no. 2, 859-908.

[20] A. Majda, *Introduction to PDEs and Waves for the Atmosphere and Ocean*, Courant Lecture Notes **9**, Courant Institute of Mathematical Sciences and American Mathematical Society, 2003.

[21] A. Majda and A. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge University Press, 2002.

[22] A. Majda and M. Grote, Model dynamics and vertical collapse in decaying strongly stratified flows, *Phys. Fluids* **9** (1997), 2932-2940.

[23] N. Masmoudi and W. Zhao, Enhanced dissipation for the 2D Couette flow in critical space, *Comm. Partial Differential Equations*, **45** (2020), 1682-1701.

[24] N. Masmoudi and W. Zhao, Stability threshold of the 2D Couette flow in Sobolev spaces, arXiv:1908.11042.

[25] J. Pedlosky, *Geophysical fluid dynamics*, Springer, New York, 1987.

[26] P.B. Rhines and W.R. Young, How rapidly is a passive scalar mixed within closed streamlines?, *J. Fluid Mech.* **133** (1983), 133-145.

[27] L. Tao and J. Wu, The 2D Boussinesq equations with vertical dissipation and linear stability of shear flows, *J. Differential Equations* **267** (2019), 1731-1747.

[28] L. Tao, J. Wu, K. Zhao and X. Zheng, Stability near hydrostatic equilibrium to the 2D Boussinesq equations without thermal diffusion, *Arch. Ration. Mech. Anal.*, **237** (2020), 585-630.

[29] C. Villani, Hypocoercivity. *Mem. Amer. Math. Soc.* **202** (2009), no. 950, iv+141.

[30] W. Thomson, Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates, *Phil. Mag.* **24** (1887), 188-196.

[31] R. Wan, Global well-posedness for the 2D Boussinesq equations with a velocity damping term, *Discrete Contin. Dyn. Syst.*, **39** (2019), 2709-2730.

[32] D. Wei and Z. Zhang, Threshold for the 3D Couette flow in Sobolev space, arXiv 1803.01359.

[33] D. Wei, Z. Zhang and W. Zhao, Linear inviscid damping for a class of monotone shear flow in Sobolev spaces, *Comm. Pure Appl. Math.* **71** (2018), 617-687.

[34] C. Zillinger, On Enhanced Dissipation for the Boussinesq Equations, *J. Differential Equations* **282** (2021), 407-445.

¹ ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE AND HUA LOO-KENG KEY LABORATORY OF MATHEMATICS, CHINESE ACADEMY OF SCIENCES, BEIJING 100049, CHINA

Email address: `dengwen@amss.ac.cn`

² DEPARTMENT OF MATHEMATICS, OKLAHOMA STATE UNIVERSITY, STILLWATER, OK 74078, UNITED STATES

Email address: `jiahong.wu@okstate.edu`

³ ACADEMY OF MATHEMATICS & SYSTEMS SCIENCE AND HUA LOO-KENG KEY LABORATORY OF MATHEMATICS, THE CHINESE ACADEMY OF SCIENCES, CHINA, AND SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF CHINESE ACADEMY OF SCIENCES, BEIJING 100049, CHINA

Email address: `zp@amss.ac.cn`