Baroclinic instability with a simple model for vertical mixing

Matthew N. Crowe* and John R. Taylor†

American Meteorological Society, Boston, Massachusetts

*Matthew N. Crowe, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, UK

†*Corresponding author address:* Current affiliation: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, UK

E-mail: J.R.Taylor@damtp.cam.ac.uk

ABSTRACT

Here, we examine baroclinic instability in the presence of vertical mixing in an idealized setting. Specifically, we employ a simple model for vertical mixing of momentum and buoyancy and expand the buoyancy and vorticity in a series for small Rossby numbers. A flow in a state of subinertial mixed layer (SML) balance (see Young (1994)) exhibits a normal mode linear instability which is studied here using an analytical stability analysis and numerical simulations. The most unstable modes grow by converting potential energy associated with the basic state into kinetic energy of the growing perturbations. However, unlike the inviscid Eady problem, the dominant energy balance is between the buoyancy flux and the energy dissipated by the modeled vertical mixing. Vertical mixing reduces the growth rate of the most unstable modes and changes their orientation with respect to the front. We test our analytical predictions for the angle and growth rate of the most unstable mode using numerical simulations and generally find good agreement. Although the predicted scale of the most unstable mode only matches the simulations for small Rossby numbers, the growth rate and angle agree for a broader range of parameters. A stability analysis of a basic state in SML balance using the inviscid QG equations shows that the angle of the unstable modes is controlled by the orientation of the SML flow, while the stratification associated with an advection/diffusion balance controls the size of growing perturbations for small Ekman numbers and large Rossby numbers. These results imply that mixed layer baroclinic instability can be inhibited by small-scale turbulence when the Ekman number is sufficiently large and might explain the lack of submesoscale eddies in observations and numerical models of the ocean surface mixed layer during summer.

1. Introduction

The ocean surface mixed layer plays a central role in the climate system by mediating transfers of heat, carbon, and other important tracers between the atmosphere and deep ocean and influencing the rate of primary production (Lorbacher et al. 2006; Sverdrup 1953). The mixed layer is subject to intense small-scale turbulence driven by a wide variety of processes including convection, wind stress and breaking waves which lead to the nearly vertically uniform density field that characterizes the mixed layer (Shay and Gregg 1986; Kato and Phillips 1969; Thorpe 2005).

The mixed layer also contains horizontal density gradients (e.g. Rudnick and Ferrari (1999)) in the form of fronts on a wide range of horizontal scales (e.g. Callies and Ferrari (2013)). The available potential energy associated with the horizontal density gradients fuels mixed layer baroclinic instability (or MLI) (Boccaletti et al. 2007; Fox-Kemper et al. 2008) which generates submesoscale eddies while re-stratifying the mixed layer. Although MLI develops in a highly turbulent environment, most previous attempts at a linear stability analysis of MLI have neglected the influence of small-scale turbulence. Our objective in this paper is to examine the influence of vertical mixing on baroclinic instability.

Observations and numerical simulations have reported a strong seasonal cycle in submesoscale activity (Capet et al. 2008; Mensa et al. 2013; Sasaki et al. 2014; Callies et al. 2015; Thompson et al. 2016). Factors that could modulate submesoscale instabilities include the mixed layer depth, horizontal density gradients, and turbulent mixing (e.g. Boccaletti et al. (2007); Bachman and Taylor (2016); Callies and Ferrari (2018)). While the growth rate for MLI does not depend directly on the mixed layer depth (Stone 1966; Fox-Kemper et al. 2008), the potential energy available for release by MLI does (Callies et al. 2015). It remains unclear whether MLI is less energetic and more difficult to detect in the summer, or whether it is arrested entirely. Here, we will show that

vertical mixing can arrest baroclinic instability in the limit of small Rossby numbers and strong mixing. This result might help explain the lack of submesoscale activity in the summer.

Recently Callies et al. (2016) used a two layer Quasi-Geostrophic (QG) model to explore how baroclinic mixed layer instability energizes submesoscale turbulence. Interestingly, their model results are consistent with available observations, despite using the QG limit of small Rossby number to describe structures with a Rossby number in the range of $0.1 - 1$. This suggests that QG dynamics may be useful to qualitatively describe submesoscale processes, although non-QG dynamics are still needed to describe phenomena such as ageostrophic instabilities (e.g. symmetric instability) and submesoscale frontogenesis (Shakespeare and Taylor 2013).

Young (1994) introduced the sub-inertial mixed layer (SML) model using an asymptotic expansion in small Rossby number and a simple parameterization of turbulent mixing to consider the effect of horizontal salinity and temperature gradients on shear and stratification in the mixed layer. For a vertically-sheared flow in thermal wind balance, vertical mixing of momentum leads to an ageostrophic secondary circulation. The secondary circulation acts to restratify the mixed layer, a tendency which is balanced by vertical mixing to leading order. The vertically-sheared cross-front flow associated with the secondary circulation and vertical mixing of temperature work together to spread the front via shear dispersion (Young et al. 1982; Taylor 1953). Shear dispersion acting on fronts was examined in Ferrari and Young (1997) and Crowe and Taylor (2018) for different mixing parameterizations.

Young and Chen (1995) used the SML model to study baroclinic instability associated with horizontal heat and salt gradients. For simplicity only cases of very strong and very weak mixing were considered, with the strong mixing corresponding to a 'slab' mixed layer model with no vertical variation and the weak mixing corresponding to a geostrophically balanced mixed layer. They speculated that the classical Eady model of baroclinic instability (Eady 1949; Vallis 2006)

should be recovered in the limit of weak mixing. However, it was not possible to make this connection explicitly since there is no background vertical stratification in the SML model and the Richardson number is large (and hence stratification is strong) in the limit of small Rossby numbers in the Eady model.

Crowe and Taylor (2018) considered the evolution of an isolated front subject to a depth-dependent turbulent viscosity and diffusivity - a simple vertical mixing parameterization intended to represent the effects of small-scale turbulence. The leading order momentum balance was found to be the so-called 'turbulent thermal wind' (TTW) balance (Gula et al. 2014) between the Coriolis acceleration, the horizontal pressure gradient, and vertical mixing, with the resulting velocity depending linearly on the horizontal buoyancy gradient. As in the SML model, vertically-sheared cross-front flow leads to a re-stratification of the mixed layer, while shear dispersion leads to spreading of the front.

Here, we take a different approach and use the vertical mixing scheme introduced by Young (1994) to consider mixed layer instabilities in the presence of vertical mixing. Unlike Young and Chen (1995) we use a single scalar, buoyancy, which simplifies the analysis for arbitrary mixing intensity. We also include a background vertical stratification to allow direct comparison with the Eady instability and we add horizontal viscous terms to examine the high wavenumber cutoff. While similar to the SML model, our asymptotic approach differs in that the buoyancy and momentum mixing timescales are assumed to be the same order, which leads to a different parameter regime. The parameter regime we use is the same as that considered in Crowe and Taylor (2018), although here the turbulent mixing is represented by relaxation towards the local depth-averaged profile rather than diffusion.

In §2 we describe the governing equations and the asymptotic limit and discuss the differences between our approach and the approach use by Young (1994); Young and Chen (1995). In §3 we

give the asymptotic solution to the governing equations in terms of the background buoyancy field, b_0, and horizontal streamfunction, ψ_0. The governing equations for b_0 and ψ_0 are given in §4 and the instabilities of these equations are considered analytically in §5 and numerically in §6. In §7 we use a quasi-geostrophic model to examine a mechanism that can control the fastest growing mode. Finally in §8 we discuss our results and the limitations of our model.

2. Governing Equations

We start with the 3D non-hydrostatic Boussinesq equations and consider a fluid bounded from above and below by flat, rigid boundaries in a coordinate system rotating about the vertical (z) axis. We invoke a linear equation of state and let the buoyancy, b, denote departures from a background stratification with buoyancy frequency N. We non-dimensionalize the governing equations using the horizontal length scale L, vertical length scale H, buoyancy scale Δb, horizontal velocity scale $U = \Delta b H/(fL)$, vertical velocity scale $W = UH/L = \Delta b H^2/(fL^2)$, pressure scale $P = fUL = \Delta b H$, and timescale $T = L/U = fL^2/(H\Delta b)$. This leads to the non-dimensional parameters defined in Table 1.

We follow Young (1994) and parameterize vertical mixing by adding a forcing term to the RHS of the momentum and buoyancy equations which acts to relax the velocity and buoyancy to the local depth-average. The rates of relaxation for buoyancy and velocity are μ_b and μ_u, respectively. This parameterization is chosen largely for mathematical convenience though it is not conspicuously less realistic than an eddy diffusivity parameterization. A similar analysis could be carried out with the vertical relaxation scheme replaced with a vertical viscosity and diffusivity. Although this complicates the analysis, qualitatively similar results can be obtained (see Appendix C).

With the choices described above, the non-dimensional governing equations are (Charney 1973; Young 1994; Crowe and Taylor 2018):

$$\mathrm{Ro}\frac{Du}{Dt} - v = -\frac{\partial p}{\partial x} + \alpha\left(\overline{u} - u\right), \tag{1a}$$

$$\mathrm{Ro}\frac{Dv}{Dt} + u = -\frac{\partial p}{\partial y} + \alpha\left(\overline{v} - v\right), \tag{1b}$$

$$\mathrm{Ro}\,\varepsilon^2\frac{Dw}{Dt} = -\frac{\partial p}{\partial z} + b, \tag{1c}$$

$$\mathrm{Ro}\frac{Db}{Dt} + \mathrm{Bu}\,w = \frac{\alpha}{\mathrm{Pr}_\alpha}\left(\overline{b} - b\right), \tag{1d}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{1e}$$

where the material derivative is

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla, \tag{2}$$

and

$$\overline{\phi} = \int_{-1/2}^{1/2} \phi\, dz \tag{3}$$

denotes a depth average across the non-dimensional vertical domain $z \in [-1/2, 1/2]$. Note that we might expect the background stratification represented by Bu to be affected by vertical mixing. We instead assume that this stratification is maintained by a process that is not represented here and occurs on a different timescale to the mixing, such as symmetric instability or surface heating, so that the background stratification can be imposed as a constant. Note that a stable stratification will develop in response to mixing of momentum even if $\mathrm{Bu} = 0$. Imposing an additional background stratification is mathematically convenient as it allows for a straightforward comparison with the Eady model in the limit of no vertical mixing.

Our approach differs from Young (1994) where it was assumed that the ratio of the buoyancy mixing timescale to the advection timescale was small compared to one, but large compared to the

Rossby number, such that

$$\frac{1/\mu_b}{T} = \frac{\mathrm{Ro}\,\mathrm{Pr}_\alpha}{\alpha} \gg \mathrm{Ro}. \tag{4}$$

Since no assumptions are made about the size of α, this results in the requirement that Pr_α is large unless α is small. Motivated by simulations and experiments of turbulent mixing in weakly stratified flows, we instead assume that $\mathrm{Pr}_\alpha = O(1)$ (e.g. Schumann and Gerz (1995); Venayagamoorthy and Stretch (2010)). This choice of Pr_α allows us to consider the case where the mixing rates are similar for any value of α, although we are unable to solve the resulting equations to the same order in Ro as Young (1994).

Before proceeding with the analysis, it is useful to relate our nondimensional parameters to physical quantities. We can relate the relaxation (mixing) rates, μ_u and μ_b, to a turbulent eddy turnover time by defining a characteristic turbulent velocity scale, u_*, and a characteristic length scale, l. The parameterized mixing rates, μ_u and μ_b, then scale with

$$\mu_u, \mu_b \sim \frac{u_*}{l}. \tag{5}$$

Therefore, the ratio of the mixing rate to the Coriolis frequency is

$$\alpha \sim \frac{u_*}{lf}. \tag{6}$$

For wind-driven turbulence, the friction velocity provides a characteristic velocity scale such that $u_* = \sqrt{\tau_w/\rho_0}$, where τ_w is the magnitude of the wind stress. In this case the turbulent length scale, l, characterizing the largest turbulent eddies would be the smaller of the mixed layer depth or the Ekman layer depth. On the other hand for convection an appropriate characteristic velocity scale is instead $u_* = w_* = (B_0 l)^{1/3}$ where B_0 is the surface buoyancy flux and l is the mixed layer (or convective layer) depth. Note that the relaxation ratio can be related to the Ekman number,

$\mathrm{E} = \nu/(fl^2)$, by invoking a mixing length argument where the turbulent viscosity, $\nu \sim u_* l$. Hence

$$\alpha \sim \frac{u_*}{lf} \sim \frac{\nu}{fl^2} \sim \mathrm{E}. \tag{7}$$

We can estimate some of the important parameters including the aspect ratio, Rossby number, and relaxation ratio from reported observations of fronts. We have selected three examples using observations reported in Mahadevan et al. (2012); Thompson et al. (2016); Thomas et al. (2013) which correspond to weak, moderate, and strong horizontal density gradients, respectively. Note that the values chosen from Mahadevan et al. (2012) correspond to the north/south density gradient characterizing the North Atlantic as observed during the North Atlantic Bloom Experiment, rather than individual fronts. The estimated parameter values are given in Table 2. Note also that the values are roughly representative of the observations, but the structure of the fronts are complicated and cannot be fully represented with a simple set of parameters. Nevertheless, the relatively weak north/south density gradient observed during the North Atlantic Bloom experiment (Mahadevan et al. 2012) and simulated by Mahadevan et al. (2012) and Taylor (2016) and the fronts reported in Thompson et al. (2016) have relatively small Rossby numbers using our definition. As we will show using comparisons with numerical simulations, aspects of our asymptotic theory are valid at these Rossby numbers. In contrast, the Rossby number associated with the Gulf Stream front is quite large and outside of the range of validity of our asymptotic theory. We note that it is possible to have $\alpha = O(1)$ for both strong and weak fronts and the aspect ratio, ε, is generally small for open ocean fronts.

Note that our definition of Rossby number uses a length scale characteristic of the horizontal density gradient and not necessarily the resulting eddies. As a result, the Rossby number as defined here can be quite small in practice. If we instead define a Rossby number, Ro_s, using the length-scale of a baroclinic eddy, we have $\mathrm{Ro}_s = K\,\mathrm{Ro}$ for nondimensional wavenumber K. In the analysis

that follows, we will show that K can be on the order of 100, and hence $\mathrm{Ro}_s = O(1)$ as typical of a submesoscale eddy. Therefore, even though the values of Ro used in our theory and simulations will be very small, our results are applicable to the formation of submesoscale structures where $\mathrm{Ro}_s \sim 1$.

3. Asymptotic Solution

In this section we will solve Eqns. (1a-1e) using an asymptotic method valid for small Rossby numbers. We begin by assuming that the aspect ratio is small and expand all variables in powers of Ro, e.g. $b = b_0 + \mathrm{Ro} b_1 + \mathrm{Ro}^2 b_2 + \ldots$. We impose no conditions on the relaxation rate, α, and allow it to appear at leading order. We also assume that the stratification is weak with $\mathrm{Bu} = O(\mathrm{Ro})$, hence we write $\mathrm{Bu} = \mathrm{Ro}\mathcal{N}^2$ where

$$\mathcal{N}^2 = N^2 H/\Delta b, \tag{8}$$

is the ratio of the vertical buoyancy difference ($N^2 H$) to the horizontal buoyancy difference (Δb). Here, for the purposes of the asymptotic equations we will assume that $\mathcal{N}^2 = O(1)$, although the result will be valid if $\mathcal{N}^2 \ll 1$. The time derivative is expanded into fast and slow timescales:

$$\frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + \frac{1}{\mathrm{Ro}} \frac{\partial}{\partial \tau}, \tag{9}$$

for fast transient timescale, $\tau = t/\mathrm{Ro}$. The fast timescale, τ, represents the transient evolution from a general initial condition. In order to simplify the analysis we assume that all transients have decayed and hence neglect the τ derivatives. For completeness, the full solution including the transients is given in Appendix A.

a. Order 1 Equations

We now consider separately the $O(1)$ and $O(\mathrm{Ro})$ terms in the governing equations. With the assumption that $\mathrm{Bu} = O(\mathrm{Ro})$, the only term in the buoyancy equation that contributes to $O(1)$ is

the parameterized vertical mixing term. Hence, the $O(1)$ buoyancy balance is

$$\frac{\alpha}{\mathrm{Pr}_\alpha} b_0' = 0, \tag{10}$$

where $(\cdot)'$ denotes a departure from the local depth-average. Eq. 10 implies that b_0 is independent of depth. This is consistent with the limit of strong mixing leading to a well-mixed layer as also found by Young (1994).

Similarly, the leading order balance in the momentum equations is

$$-v_0 = -\frac{\partial p_0}{\partial x} - \alpha u_0', \tag{11a}$$

$$u_0 = -\frac{\partial p_0}{\partial y} - \alpha v_0', \tag{11b}$$

$$0 = -\frac{\partial p_0}{\partial z} + b_0, \tag{11c}$$

$$0 = \frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} + \frac{\partial w_0}{\partial z}. \tag{11d}$$

Eq. 11c can be integrated to give $p_0 = z b_0 + \overline{p}_0$, and the depth-averaged horizontal momentum equations and mass conservation equation reduce to geostrophic balance for the depth-averaged flow:

$$-\overline{v}_0 = -\frac{\partial \overline{p}_0}{\partial x}, \tag{12a}$$

$$\overline{u}_0 = -\frac{\partial \overline{p}_0}{\partial y}, \tag{12b}$$

$$0 = \frac{\partial \overline{u}_0}{\partial x} + \frac{\partial \overline{v}_0}{\partial y}. \tag{12c}$$

Subtracting the depth-averaged horizontal momentum equations from equations 11 gives evolution equations for the horizontal velocity perturbations and vertical velocity

$$\alpha u_0' - v_0' = -z \frac{\partial b_0}{\partial x}, \tag{13a}$$

$$\alpha v_0' + u_0' = -z \frac{\partial b_0}{\partial y}, \tag{13b}$$

$$0 = \frac{\partial u_0'}{\partial x} + \frac{\partial v_0'}{\partial y} + \frac{\partial w_0}{\partial z}. \tag{13c}$$

Equations 13 can be combined to give

$$\mathbf{u}'_{H0} = \gamma[-\alpha\nabla_H b_0 + \mathbf{k}\times\nabla_H b_0]\,z, \tag{14}$$

and

$$w_0 = \frac{\alpha\gamma(4z^2-1)}{8}\nabla_H^2 b_0, \tag{15}$$

where $\gamma = 1/(1+\alpha^2)$. From the depth-averaged mass conservation equation we can write $\overline{\mathbf{u}}_{H0} = -\nabla\times(\psi_0\mathbf{k})$ for streamfunction ψ_0. From equations 12a and 12b we note that $\overline{p}_0 = \psi_0$. Hence

$$\mathbf{u}_{H0} = -\nabla\times(\psi_0\mathbf{k}) + \gamma[-\alpha\nabla_H b_0 + \mathbf{k}\times\nabla_H b_0]\,z. \tag{16}$$

As noted in Young (1994), the horizontal velocity has a non-zero vertical shear at leading order, unlike the buoyancy which is well-mixed at leading order. In the case of $\alpha = 0$, the equation for $\mathbf{u}_{H0}$ reduces to thermal wind balance. For nonzero α, vertical mixing acts to couple the cross-front and along-front flows, leading to a flow with a component in the direction of the buoyancy gradient. For $\alpha < 1$ stronger mixing results in a stronger cross-front shear, while the cross-front shear weakens with stronger mixing for $\alpha > 1$.

b. Order Ro Equations

We now consider the $O(\mathrm{Ro})$ terms in the buoyancy conservation equation. The advection of b_0 by the leading order velocity contributes to $O(\mathrm{Ro})$. Since $b_0 = b_0(x,y,t)$, the $O(\mathrm{Ro})$ buoyancy equation is

$$\frac{\partial b_0}{\partial t} + \mathbf{u}_{H0}\cdot\nabla_H b_0 + \mathcal{N}^2 w_0 = -\frac{\alpha}{\mathrm{Pr}_\alpha} b'_1. \tag{17}$$

Subtracting the depth average gives

$$\frac{\alpha}{\mathrm{Pr}_\alpha} b'_1 = -\mathbf{u}'_{H0}\cdot\nabla_H b_0 - \mathcal{N}^2 w'_0, \tag{18}$$

which describes a balance between advection by the cross-front flow and vertical mixing. Hence the solution for b_1' is

$$b_1' = \mathrm{Pr}_\alpha \gamma \left[z|\nabla_H b_0|^2 - \mathcal{N}^2 \frac{12z^2-1}{24} \nabla_H^2 b_0 \right]. \tag{19}$$

Solving for $\overline{b}_1$ requires the $O(\mathrm{Ro}^2)$ buoyancy equation. Note that even with $\mathcal{N}=0$, there is a stable vertical stratification at this order, consistent with the finding from Tandon and Garrett (1994) that the vertical buoyancy gradient is proportional to the horizontal buoyancy gradient squared.

4. Evolution of the Background Fields

In order to determine the time dependence of the system on the slow timescale, t, we need to determine governing equations for the depth independent functions b_0 and ψ_0. These can be obtained by depth averaging the buoyancy and vertical vorticity equations. The vertical vorticity equation is

$$\mathrm{Ro}\left(\frac{\partial \zeta}{\partial t} + \mathbf{u}\cdot\nabla\zeta - \boldsymbol{\omega}\cdot\nabla w \right) + \nabla_H \cdot \mathbf{u}_H = \alpha(\overline{\zeta} - \zeta), \tag{20}$$

for vertical vorticity $\zeta = \boldsymbol{\omega}\cdot\mathbf{k}$, which can be depth-averaged to give

$$\frac{\partial \overline{\zeta}}{\partial t} + \nabla_H \cdot [\overline{\mathbf{u}_H \zeta - \boldsymbol{\omega}_H w}] = 0, \tag{21}$$

or using depth-averaged and perturbation quantities,

$$\frac{\partial \overline{\zeta}}{\partial t} + \nabla_H \cdot [\overline{\mathbf{u}}_H \overline{\zeta} - \overline{\boldsymbol{\omega}}_H \overline{w} + \overline{\mathbf{u}_H' \zeta'} - \overline{\boldsymbol{\omega}_H' w'}] = 0. \tag{22}$$

Similarly, the depth-averaged buoyancy equation is

$$\frac{\partial \overline{b}}{\partial t} + \overline{\mathbf{u}}_H \cdot \nabla_H \overline{b} + \nabla_H \cdot [\overline{\mathbf{u}_H' b'}] + \mathcal{N}^2 \overline{w} = 0. \tag{23}$$

We now use the leading order solutions for the velocity and buoyancy fields (ψ_0 and b_0) to write the depth-averaged equations in terms of these fields.

a. Buoyancy

Substituting the expansions in Rossby number up to $O(\mathrm{Ro})$ into Eq. 23 gives

$$\frac{\partial b_0}{\partial t} + \overline{\mathbf{u}}_{H0}\cdot\nabla_H b_0 + \mathcal{N}^2\overline{w}_0 + \mathrm{Ro}\left[\frac{\partial \overline{b}_1}{\partial t} + \overline{\mathbf{u}}_{H1}\cdot\nabla_H b_0 + \overline{\mathbf{u}}_{H0}\cdot\nabla_H \overline{b}_1 + \mathcal{N}^2\overline{w}_1 + \nabla_H\cdot[\overline{\mathbf{u}'_{H0}b'_1}]\right] = 0. \tag{24}$$

Using the definition of ψ_0, we can write $\overline{\mathbf{u}}_{H0}\cdot\nabla_H b_0 = J(\psi_0, b_0)$, where J is the Jacobian operator:

$$J(f,g) = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}. \tag{25}$$

We can also write the flux term as

$$\overline{\mathbf{u}'_{H0}b'_1} = \frac{\mathrm{Pr}_\alpha\gamma^2}{12}(-\alpha\nabla_H b_0 + \mathbf{k}\times\nabla_H b_0)|\nabla_H b_0|^2, \tag{26}$$

and hence Eq. 24 can be written as

$$\begin{aligned}\frac{\partial b_0}{\partial t} + J(\psi_0, b_0) + \mathcal{N}^2\overline{w}_0 + \mathrm{Ro}\left[\frac{\partial \overline{b}_1}{\partial t} + \overline{\mathbf{u}}_{H1}\cdot\nabla_H b_0 + \overline{\mathbf{u}}_{H0}\cdot\nabla_H \overline{b}_1 + \mathcal{N}^2\overline{w}_1\right] \\ = \frac{\mathrm{Ro}\,\mathrm{Pr}_\alpha\gamma^2}{12}\nabla_H\cdot\left[(\alpha\nabla_H b_0 - \mathbf{k}\times\nabla_H b_0)|\nabla_H b_0|^2\right].\end{aligned} \tag{27}$$

The limit considered by Young (1994) uses $\mathrm{Pr}_\alpha = \mathcal{P}/\sqrt{\mathrm{Ro}}$ with $\mathcal{P} = O(1)$. For $\mathrm{Ro} \ll 1$ this corresponds to momentum relaxation that is much faster than the buoyancy relaxation. With this choice, and in the absence of background stratification ($N^2 = 0$), the buoyancy evolution equation to order $O(\sqrt{\mathrm{Ro}})$ can be written

$$\frac{\partial b_0}{\partial t} + J(\psi_0, b_0) = \frac{\sqrt{\mathrm{Ro}}\,\mathcal{P}\gamma^2}{12}\nabla_H\cdot\left[(\alpha\nabla_H b_0 - \mathbf{k}\times\nabla_H b_0)|\nabla_H b_0|^2\right]. \tag{28}$$

This result was obtained by Young (1994). The first term in brackets on the right hand side of Eq. 28 is a down-gradient buoyancy flux. The second term is a 'skew' flux directed perpendicular to the buoyancy gradient. The role of the skew flux will be discussed in more detail in §4c.

Here, we take a different approach from Young (1994) and assume that $\mathrm{Pr}_\alpha = O(1)$ while retaining a non-zero background stratification. The $O(1)$ terms in Eq. 27 are then

$$\frac{\partial b_0}{\partial t} + J(\psi_0, b_0) = \frac{\alpha\gamma\mathcal{N}^2}{12}\nabla_H^2 b_0, \tag{29}$$

where we have used $\overline{w}_0 = -\alpha\gamma\left(\nabla_H^2 b_0\right)/12$. With this form for w_0, vertical advection acting on the background stratification ($\mathcal{N}^2 w_0$ in Eq. 27) acts like horizontal diffusion on the leading order buoyancy.

Crowe and Taylor (2018, 2019) studied the evolution of a front in turbulent thermal wind balance. A simple form of this problem can be obtained by considering a two-dimensional front with no y dependence and $\mathcal{N}^2 = 0$. With these assumptions, Eq. 27 reduces to

$$\frac{\partial b_0}{\partial t} = \frac{\mathrm{Ro}\,\mathrm{Pr}_\alpha \gamma^2 \alpha}{12} \frac{\partial}{\partial x}\left(\frac{\partial b_0}{\partial x}\right)^3, \tag{30}$$

as $\overline{b}_1$ and $\overline{u}_{H1}$ can be assumed to be zero by symmetry. This equation describes frontal spreading on the timescale $T = \mathrm{Ro}\,t$ and can be solved with a similarity solution as in Crowe and Taylor (2018).

b. Vorticity

We can formulate a closed system of two equations for the leading order buoyancy, b_0, and the leading order streamfunction, ψ_0, using conservation of vorticity. The leading order vorticity equation is

$$\frac{\partial \overline{\zeta}_0}{\partial t} + \nabla_H \cdot [\overline{\mathbf{u}}_{H0}\overline{\zeta}_0 - \overline{\omega}_{H0}\overline{w}_0 + \overline{\boldsymbol{u}'_{H0}\zeta'_0} - \overline{\boldsymbol{\omega}'_{H0}w'_0}] = 0, \tag{31}$$

where each term can now be written in terms of b_0 and ψ_0. The leading order vertical vorticity is

$$\zeta_0 = \frac{\partial v_0}{\partial x} - \frac{\partial u_0}{\partial y} = \nabla_H^2 \psi_0 + \gamma z \nabla_H^2 b_0. \tag{32}$$

Since b_0 is independent of z, and since z is anti-symmetric about the mid-plane ($z = 0$), the final term does not contribute to the depth-average, which leaves $\overline{\zeta}_0 = \nabla_H^2 \psi_0$. Advection of vorticity by the depth-averaged horizontal velocity can be written as

$$\nabla_H \cdot [\overline{\mathbf{u}}_{H0}\overline{\zeta}_0] = \nabla_H \cdot [\overline{\mathbf{u}}_{H0}\nabla_H^2 \psi_0] = J(\psi_0, \nabla_H^2 \psi_0). \tag{33}$$

The horizontal vorticity is given by

$$\boldsymbol{\omega}_{H0} = \begin{pmatrix} \frac{\partial w_0}{\partial y} - \frac{\partial v_0}{\partial z} \\ \frac{\partial u_0}{\partial z} - \frac{\partial w_0}{\partial x} \end{pmatrix}, \tag{34}$$

and its depth-average is

$$\overline{\boldsymbol{\omega}}_{H0} = \begin{pmatrix} \frac{\partial \overline{w}_0}{\partial y} - \Delta v_0 \\ \Delta u_0 - \frac{\partial \overline{w}_0}{\partial x} \end{pmatrix}, \tag{35}$$

where Δu_0 and Δv_0 are the change in horizontal velocity between the top and bottom boundaries. The second flux term is

$$\nabla_H \cdot [\overline{\boldsymbol{\omega}}_{H0} \overline{w}_0] = \nabla_H \cdot \left[\frac{1}{2} \begin{pmatrix} \frac{\partial \overline{w}_0^2}{\partial y} \\ -\frac{\partial \overline{w}_0^2}{\partial x} \end{pmatrix} + \begin{pmatrix} -\Delta v_0 \, \overline{w}_0 \\ \Delta u_0 \, \overline{w}_0 \end{pmatrix} \right] = \nabla_H \cdot \begin{pmatrix} -\Delta v_0 \, \overline{w}_0 \\ \Delta u_0 \, \overline{w}_0 \end{pmatrix}, \tag{36}$$

or

$$\nabla_H \cdot [\overline{\boldsymbol{\omega}}_{H0} \overline{w}_0] = \frac{\alpha \gamma^2}{12} \nabla_H \cdot \left[\left(\nabla_H b_0 + \alpha \, \mathbf{k} \times \nabla_H b_0 \right) \nabla_H^2 b_0 \right]. \tag{37}$$

using the leading order velocities. The last two flux terms involving departures from the mean vorticity are

$$\nabla_H \cdot [\overline{\boldsymbol{u}'_{H0} \zeta'_0}] = -\frac{\gamma^2}{12} \nabla_H \cdot \left[\left(\alpha \nabla_H b_0 - \mathbf{k} \times \nabla_H b_0 \right) \nabla_H^2 b_0 \right], \tag{38}$$

and

$$\nabla_H \cdot [\overline{\boldsymbol{\omega}'_{H0} w'_0}] = \nabla_H \cdot \begin{pmatrix} \frac{1}{2} \frac{\partial}{\partial y} [\overline{w'^2_0}] - \overline{\frac{\partial v'_0}{\partial z} w'_0} \\ -\frac{1}{2} \frac{\partial}{\partial x} [\overline{w'^2_0}] + \overline{\frac{\partial u'_0}{\partial z} w'_0} \end{pmatrix} = \nabla_H \cdot \begin{pmatrix} -\frac{\partial v'_0}{\partial z} \overline{w'_0} \\ \frac{\partial u'_0}{\partial z} \overline{w'_0} \end{pmatrix} = \mathbf{0}, \tag{39}$$

since u'_0 and v'_0 are linear in z and $\overline{w'} = 0$. The terms in $\overline{w'^2_0}$ can be written as a curl and hence are divergence free.

Combining these results, the vertical vorticity equation can be written

$$\frac{\partial \nabla_H^2 \psi_0}{\partial t} + J(\psi_0, \nabla_H^2 \psi_0) = \frac{\gamma^2}{12} \nabla_H \cdot \left[\left(2\alpha \nabla_H b_0 + (\alpha^2 - 1) \, \mathbf{k} \times \nabla_H b_0 \right) \nabla_H^2 b_0 \right]. \tag{40}$$

The left hand side of Eq. 40 corresponds to advection of vertical vorticity. The first term in brackets on the right hand side corresponds to a cross-front vorticity flux, and the second term in brackets corresponds to an along-front skew flux.

c. The Skew Flux Term

As mentioned earlier, a skew flux term appears in the evolution equation for b_0 (the second term on the right hand side of Eq. 27). This term, $\mathcal{J}_b = -\mathbf{k} \times \nabla_H b_0 |\nabla_H b_0|^2$, represents a flux of buoyancy perpendicular to the buoyancy gradient. As noted above, a skew flux also appears in the vertical vorticity equation (Eq. 40) which we will denote $\mathcal{J}_v = (\mathbf{k} \times \nabla_H b_0)\nabla_H^2 b_0$.

The divergence of the skew flux terms in the buoyancy and vorticity equations can be re-expressed in terms of advection operators. First, note that the divergence of the skew flux terms can be written as

$$\nabla_H \cdot \mathcal{J}_b = -\nabla_H \cdot \left[\mathbf{k} \times \nabla_H b_0 |\nabla_H b_0|^2\right] = -\nabla_H b_0 \cdot \left[-\mathbf{k} \times \nabla_H |\nabla_H b_0|^2\right], \tag{41}$$

and

$$\nabla_H \cdot \mathcal{J}_v = \nabla_H \cdot \left[(\mathbf{k} \times \nabla_H b_0)\nabla_H^2 b_0\right] = -\nabla_H b_0 \cdot \left[(\mathbf{k} \times \nabla_H)\nabla_H^2 b_0\right]. \tag{42}$$

Therefore, the terms in brackets can be written in the form of advection operators with velocities

$$\mathbf{u}_b = -\mathbf{k} \times \nabla_H |\nabla_H b_0|^2 = \nabla_H \times \left[|\nabla_H b_0|^2 \mathbf{k}\right], \tag{43}$$

and

$$\mathbf{u}_v = (\mathbf{k} \times \nabla_H)\nabla_H^2 b_0 = \nabla_H \times \left[-\nabla_H^2 b_0 \mathbf{k}\right]. \tag{44}$$

Therefore $\mathbf{u}_b$ and $\mathbf{u}_v$ can be written in terms of streamfunctions, $\chi_b = -|\nabla_H b_0|^2$ and $\chi_v = \nabla_H^2 b_0$. In the buoyancy equation the skew flux term can be combined with the existing advection term, $J(\psi_0, b_0)$. From the form of the streamfunction, χ_b, we see that the effect of the skew flux term,

$\mathcal{J}_b$, is to advect buoyancy along the contours of $|\nabla_H b_0|^2 = \text{const}$. Equations 27 and 40 can then be written

$$\frac{\partial b_0}{\partial t} + J\left(\psi_0 - \frac{\sqrt{\text{Ro}}\mathcal{P}\gamma^2}{12}|\nabla_H b_0|^2, b_0\right) = \frac{\text{Ro}\,\text{Pr}_\alpha \alpha\gamma^2}{12}\nabla_H \cdot \left[\nabla_H b_0 |\nabla_H b_0|^2\right] + O(\text{Ro}), \tag{45}$$

and

$$\frac{\partial \nabla_H^2 \psi_0}{\partial t} + J(\psi_0, \nabla_H^2 \psi_0) + \frac{(\alpha^2-1)\gamma^2}{12} J(\nabla_H^2 b_0, b_0) = \frac{2\alpha\gamma^2}{12}\nabla_H \cdot \left[\nabla_H b_0 \nabla_H^2 b_0\right]. \tag{46}$$

Note that using Equation 45 with $\text{Pr}_\alpha = \mathcal{P}/\sqrt{\text{Ro}}$ corresponds to the Young (1994) case. In the $\text{Pr}_\alpha = O(1)$ limit that we consider, only the vorticity skew flux term, $\mathcal{J}_v$, enters the equations at leading order and the advection of buoyancy by the buoyancy skew flux term, $\mathcal{J}_b$, is small.

d. Horizontal Diffusion

The system described by Eqns. 1a-1e parameterizes vertical mixing by relaxing the velocity and buoyancy fields towards their local depth average, but the equations do not include any parameterization for horizontal mixing by small-scale turbulence. As will be shown below, the most unstable mode in this system has an infinite horizontal wavenumber or, equivalently, a vanishingly small wavelength. Fortunately, it is relatively straightforward to include a parameterization of horizontal mixing using horizontal Laplacian viscous and diffusive terms with viscosity ν and diffusivity κ. The addition of these terms shifts the most unstable mode to a finite wavenumber. Note that this Laplacian scheme differs from the relaxation parameterization used to represent vertical mixing and is used for mathematical convenience. Appendix C describes a model with Laplacian mixing schemes in the horizontal and vertical directions.

With the addition of parameterized horizontal mixing, the terms $\varepsilon^2 \text{E}\nabla_H^4 \psi_0$ and $\varepsilon^2 \text{E}/\text{Pr}_E \nabla_H^2 b_0$ appear on the right hand sides of the depth-averaged vorticity and buoyancy equations, respectively, where recall that $\varepsilon = H/L$ is the aspect ratio, $\text{E} = \nu/(fH^2)$ is the Ekman number and

$\mathrm{Pr}_E = \nu/\kappa$ is the Prandtl number. In most applications $\varepsilon << 1$ and these terms will be small. However, for very small Ro these terms might contribute significantly. Writing $\mathcal{E} = \varepsilon^2 \mathrm{E}/\mathrm{Ro}$, the resulting equations are

$$\frac{\partial b_0}{\partial t} + J(\psi_0, b_0) = \left[\frac{\alpha\gamma\mathcal{N}^2}{12} + \frac{\mathcal{E}}{\mathrm{Pr}_E}\right]\nabla_H^2 b_0, \tag{47}$$

and

$$\frac{\partial \nabla_H^2 \psi_0}{\partial t} + J(\psi_0, \nabla_H^2\psi_0) - \mathcal{E}\nabla_H^4\psi_0 = \frac{\gamma^2}{12}\nabla_H \cdot \left[\left(2\alpha\nabla_H b_0 + (\alpha^2 - 1)\,\mathbf{k}\times\nabla_H b_0\right)\nabla_H^2 b_0\right], \tag{48}$$

which are asymptotically valid if $\varepsilon^2\mathrm{E} = O(\mathrm{Ro})$. For convenience, we will write the combined buoyancy diffusivity appearing in Eq. 47 as

$$\mathcal{D} = \left[\frac{\alpha\gamma\mathcal{N}^2}{12} + \frac{\mathcal{E}}{\mathrm{Pr}_E}\right]. \tag{49}$$

We note that these equations can be obtained from Young (1994) in the limit of fast buoyancy mixing ($1/\mu_b \ll L/U$). However this result would only strictly be valid for small α based on the analysis in Young (1994) due to the use of different asymptotic limits, while here no constraints have been placed on the size of α.

5. Instabilities of the Depth-Averaged Equations

Equations 47 and 48 are a closed system of equations for the leading order buoyancy and vorticity. In this section, we will analyze the stability of these equations to small amplitude disturbances. For simplicity, we will consider perturbations about a basic state where buoyancy is a linear function of x, i.e. $b_0 = Bx$ for a constant B, and where the vertical vorticity is zero. Introducing normal mode perturbations of the form $\exp[i(kx+ly)+\sigma t]$, the total buoyancy and vorticity can be written using an eigenmode decomposition,

$$(b_0, \psi_0) = \left(\delta A \exp[i(kx+ly)+\sigma t] + Bx, \delta C \exp[i(kx+ly)+\sigma t]\right), \tag{50}$$

for wavevector (k,l), growth rate σ and small parameter δ. The vector (A,C) is the eigenvector of the resulting linear system.

To leading order in δ, the linearized buoyancy and vorticity equations can be written

$$\sigma A - ilBC = -(k^2+l^2)\mathcal{D}A, \tag{51}$$

and

$$-\sigma(k^2+l^2)C = \frac{\gamma^2}{12}\left[-2\alpha ik(k^2+l^2) - (\alpha^2-1)il(k^2+l^2)\right]BA + (k^2+l^2)^2\mathcal{E}C, \tag{52}$$

or in the form of a single matrix equation,

$$\begin{pmatrix} \sigma + (k^2+l^2)\mathcal{D} & -lB \\ \frac{\gamma^2}{12}\left[2\alpha k + (\alpha^2-1)l\right]B & \sigma + (k^2+l^2)\mathcal{E} \end{pmatrix} \begin{pmatrix} A \\ iC \end{pmatrix} = \mathbf{0}. \tag{53}$$

For this equation to be valid for some non-zero vector (A, iC), the determinant of this matrix must vanish. Therefore

$$[\sigma + (k^2+l^2)\mathcal{D}][\sigma + (k^2+l^2)\mathcal{E}] + \frac{\gamma^2 B^2}{12}\left[2\alpha kl + (\alpha^2-1)l^2\right] = 0. \tag{54}$$

The solution to this equation for the growth rate, σ, is

$$\sigma_\pm = -\frac{\mathcal{D}+\mathcal{E}}{2}(k^2+l^2) \pm \sqrt{\left[\frac{\mathcal{D}-\mathcal{E}}{2}\right]^2(k^2+l^2)^2 - \mathcal{B}^2\left[2\alpha kl + (\alpha^2-1)l^2\right]}, \tag{55}$$

where $\mathcal{B}^2 = \gamma^2 B^2/12$ is a re-scaled buoyancy gradient and the parameters

$$\mathcal{D} = \frac{\alpha\gamma \mathrm{Bu}}{12\,\mathrm{Ro}} + \frac{\mathcal{E}}{\mathrm{Pr}_E}, \quad \mathcal{E} = \frac{\varepsilon^2 E}{\mathrm{Ro}}, \tag{56}$$

can be written in terms of the non-dimensional numbers defined in Table 1. The growth rate has a maximum at a finite wavevector (k,l). Note that in the case where $\mathcal{D} = \mathcal{E}$ (e.g. with $\mathcal{N}^2 = 0$ and $\mathrm{Pr}_E = 1$), the growth rate simplifies to

$$\sigma_\pm = -\mathcal{E}(k^2+l^2) \pm \mathcal{B}\sqrt{-\left[2\alpha kl + (\alpha^2-1)l^2\right]}. \tag{57}$$

To find the maximum growth rate in the more general case, it is useful to define a rotated wavevector, $\mathbf{l}' = \mathbf{R}_\alpha \mathbf{l}$, where $\mathbf{l} = (k, l)$ and the rotation matrix,

$$\mathbf{R}_\alpha = \frac{1}{\sqrt{1+\alpha^2}} \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix}, \tag{58}$$

is orthogonal with determinant 1 corresponding to a rotation by $\theta = -\arctan\alpha$. The growth rate in rotated coordinates becomes

$$\sigma_\pm = -\frac{\mathcal{D}+\mathcal{E}}{2}(k'^2 + l'^2) \pm \sqrt{\left[\frac{\mathcal{D}-\mathcal{E}}{2}\right]^2 (k'^2+l'^2)^2 + \mathcal{B}^2\left[l'^2 - \alpha^2 k'^2\right]}. \tag{59}$$

For a fixed wavelength, the growth rate in Eq. 59 is maximum for $(k', l') = (0, \pm K)$, corresponding to waves aligned at an angle of $\arctan\alpha$ to the down-front (y) direction. We note that the fastest growing modes therefore have a wavevector aligned with the horizontal velocity at the top and bottom boundaries. The same is true for the the classical Eady instability where the wavenumber of the fastest growing modes is in the down-front direction (Eady 1949; Vallis 2006), but here the cross-front flow changes the orientation of the growing modes with respect to the front. The maximum growth rate over all directions as a function of the wavenumber is therefore

$$\sigma_{max}(K) = -\frac{\mathcal{D}+\mathcal{E}}{2}K^2 + \sqrt{\left[\frac{\mathcal{D}-\mathcal{E}}{2}\right]^2 K^4 + \mathcal{B}^2 K^2}. \tag{60}$$

Maximizing over K, the most unstable mode has a growth rate

$$\max_K[\sigma_{max}] = \frac{\mathcal{B}^2}{(\sqrt{\mathcal{D}}+\sqrt{\mathcal{E}})^2}, \tag{61}$$

which reduces to $\max_K[\sigma_{max}] = \mathcal{B}^2/(4\mathcal{D})$ in the case $\mathcal{D} = \mathcal{E}$. Therefore the most unstable wavenumber, K_{max} satisfies

$$K_{max}^2 = \frac{-2\mathcal{D}\mathcal{E} + \sqrt{\mathcal{D}\mathcal{E}(\mathcal{D}+\mathcal{E})^2}}{\mathcal{D}\mathcal{E}(\mathcal{D}-\mathcal{E})^2}\mathcal{B}^2, \tag{62}$$

which reduces to

$$K_{max}^2 = \frac{\mathcal{B}^2}{4\mathcal{D}^2}, \tag{63}$$

in the case $\mathcal{D} = \mathcal{E}$.

As noted above, the horizontal viscous/diffusion terms are necessary to produce a finite wavenumber maximum since $K^2_{max} \to \infty$ as $\mathcal{E} \to 0$. Therefore, the system without horizontal viscosity and diffusion appears to produce an 'ultraviolet catastrophe'. However, the maximum growth rate in this case does remain bounded since σ_{max} asymptotes to $\mathcal{B}^2/\mathcal{D}$ for large K^2.

The case of $\mathcal{D} = 0$ and $\mathcal{E} \neq 0$ corresponds to no stratification and an infinite Prandtl number. By symmetry in $\mathcal{E}$ and $\mathcal{D}$ this case is the same as the $\mathcal{E} = 0$ case though with different eigenvectors. When both $\mathcal{E}$ and $\mathcal{D}$ are zero, corresponding to no stratification and no horizontal diffusion, we have

$$\sigma_{max}(K) = \mathcal{B}K, \tag{64}$$

so the growth rate is unbounded and waves with infinite wavenumber will grow infinitely quickly.

The case of $\alpha = 0$ corresponds to the small wavenumber (long wave) limit of the classical Eady problem, where the growth rate is

$$\sigma_{\text{Eady}} = \frac{Bl}{\mu}\left[\left(\coth\frac{\mu}{2} - \frac{\mu}{2}\right)\left(\frac{\mu}{2} - \tanh\frac{\mu}{2}\right)\right]^{1/2}, \tag{65}$$

for scaled wavenumber $\mu^2 = \text{Bu}\, l^2$ (Vallis 2006). Since we consider $\text{Bu} = O(\text{Ro})$, the relevant limit is the small μ limit in which case σ_{Eady} reduces to

$$\sigma_{\text{Eady}} \sim \frac{Bl}{\sqrt{12}} = \mathcal{B}K, \tag{66}$$

consistent with Eq. 64. Note that this result is independent of the background buoyancy gradient represented by Bu. This result is also consistent with the small K limit of ageostrophic baroclinic instability considered by Stone (1966).

As noted earlier, the direction of the most unstable modes described by Eq. 59 corresponds to $k' = 0$. In non-rotated coordinates this corresponds to $k = -\alpha l$, where k is the wavenumber in the cross-front direction and l is the wavenumber in the along-front direction. In contrast, the most

unstable modes in the inviscid Eady problem have $k = 0$ and hence correspond to the limit of $\alpha \to 0$.

Instead, the modes perpendicular to the most unstable modes have $l = \alpha k$ in non-rotated coordinates. For these modes, the coefficient multiplying $\mathcal{B}^2$ inside the square root in Eq. 59 is negative. If $\mathcal{E} = \mathcal{D}$, $\sigma_\pm$ is purely imaginary for these modes, corresponding to traveling waves with a constant amplitude. Note, however, that the neglected higher order terms could add a real part to this growth rate and hence cause these perturbations to grow, while adding horizontal friction and diffusion will cause them to slowly decay with faster decay at higher wavenumbers.

Figure 1 shows the real and imaginary parts of $\sigma_\pm$ in the case of no horizontal friction and diffusion or background stratification ($\mathcal{D} = \mathcal{E} = 0$). Only the σ_+ branch produces growing modes with the fastest growth occurring for large K along the line $k = -\alpha l$. Figure 2 shows the real and imaginary parts of $\sigma_\pm$ with $\mathcal{E} = \mathcal{D} = 2.5 \times 10^{-3}$, corresponding to large horizontal friction and diffusivity or small Rossby number. A maximum in the growth rate can be seen on the plot of $Re[\sigma_+]$ for $K_{max} = 92.4$ along the line $k = -\alpha l$.

We anticipate that the 'ultraviolet catastrophe' in the system without horizontal mixing will be cured by finite Rossby number effects. Equations 45 and 46 contain terms that are $O(\mathrm{Ro})$ which were neglected in equations 47 and 48. These terms involve an extra power of the horizontal wavenumber magnitude, K, compared to the leading order terms. Therefore, the neglect of these terms is asymptotically valid when $K \ll O(1/\mathrm{Ro})$. For sufficiently large K the neglected $O(\mathrm{Ro})$ terms will become important and modify the growth rate, possibly resulting in a maximum growth rate at a lower wavenumber than predicted in Eq. 62 when the Rossby number is not infinitesimally small. This will be discussed further in §7.

6. Numerical Simulations

To test the theory described above, we have conducted a series of fully nonlinear numerical simulations using the code DIABLO. The code solves the incompressible non-hydrostatic Boussinesq equations. Time stepping is performed with a combination of explicit third-order Runge-Kutta and implicit Crank Nicolson schemes while finite differences are used for derivatives in the vertical direction and discrete Fourier transforms, using the pseudo-spectral method for non-linear terms, are used for derivatives in the horizontal direction (Taylor 2008).

The simulations solve the non-dimensional equations where x, y, and z are normalized by the size of the computational domain such that the non-dimensional domain size is $L_x = L_y = L_z = 1$. The boundary conditions in the vertical direction are no stress, no buoyancy flux and no vertical velocity on the top and bottom surfaces. Periodic boundary conditions are applied to the velocity in both horizontal directions (see below for buoyancy).

For numerical stability, viscous terms of the form

$$D_{(u,v)} = \mathrm{E}\left(\frac{\partial^2}{\partial z^2} + \varepsilon^2\nabla_H^2\right)(u,v), \tag{67}$$

$$D_w = \varepsilon^2\mathrm{E}\left(\frac{\partial^2}{\partial z^2} + \varepsilon^2\nabla_H^2\right)w, \tag{68}$$

and

$$D_b = \frac{\mathrm{E}}{\mathrm{Pr}_E}\left(\frac{\partial^2}{\partial z^2} + \varepsilon^2\nabla_H^2\right)b, \tag{69}$$

are added to the horizontal momentum, vertical momentum and buoyancy equations respectively. The simulations use a small Ekman number, E, and a small aspect ratio, ε, such that the dominant vertical mixing process is the relaxation to the depth-average. The simulations are initialized with the solution given in Appendix B with vertical diffusion and relaxation. For the small Ekman numbers considered here this solution exhibits thin boundary layers where vertical diffusion is important. Outside of these boundary layers, the velocity and buoyancy fields correspond to the

solution given in §3. The initial velocity field is set to the leading order solution while the initial buoyancy field is prescribed to be a linear horizontal background gradient plus the resulting $O(\mathrm{Ro})$ correction.

Periodic boundary conditions are inconsistent with the initial conditions for buoyancy which have a constant horizontal buoyancy gradient. To overcome this, we decompose the total buoyancy into a background term with a constant buoyancy gradient and departures from this gradient, i.e.

$$b = Bx + b_p, \tag{70}$$

where B is constant. This form is inserted into the buoyancy equation, and periodic boundary conditions are applied to b_p. This has the effect of fixing the change in buoyancy across the domain in the x direction. A similar approach has been used in a number of previous studies (e.g. Taylor and Ferrari (2011); Taylor (2016)).

Small amplitude normal mode perturbations are then added the buoyancy and streamfunction of the form

$$b_0' = \mathcal{R}\left[\sum_{(k,l)} A_{kl} \exp[i(kx + ly + \phi_{kl})]\right], \tag{71}$$

and

$$\psi_0' = \mathcal{R}\left[\sum_{(k,l)} C_{kl} \exp[i(kx + ly + \phi_{kl})]\right], \tag{72}$$

where ϕ_{kl} is a random phase, $\mathcal{R}[f]$ denotes the real part of f and $(k,l) = 2\pi(n_k, n_l)$ for $n_{k,l} = 1, 2, \ldots$ and $n_k^2 + n_l^2 < N_{max}^2$ describing a disc in phase space of radius $2\pi N_{max}$.

Note that the leading order depth-dependent velocity depends on b_0 and hence the velocity perturbation can be found from b_0', while perturbations to the depth-independent velocity are introduced through ψ_0'. Similarly, the leading order depth-dependent buoyancy, b_1, depends on b_0 and therefore perturbations to b_1 are introduced through b_0'. In the simulations, we set the amplitudes $|A_{kl}| = |C_{kl}| = 10^{-12}$ which ensure an interval of linear perturbation growth, while the phase

difference between A_{kl} and C_{kl} is randomized. We use a background buoyancy gradient of $B = 2$ such that $b = \pm 1$ at $x = \pm 0.5$.

There are several non-dimensional parameters in the system described here. For simplicity, the numerical simulations are conducted for fixed Burger number, Prandtl number, aspect ratio, and Ekman numbers, with $\mathrm{Bu} = 0$, $\mathrm{Pr}_\alpha = \mathrm{Pr}_E = 1$, $\varepsilon = 0.05$ and $\mathrm{E} = 10^{-4}$. The Rossby number, Ro, and relaxation ratio, α, are varied over the set of values $\mathrm{Ro} \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ and $\alpha \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. Each simulation is run until growing modes develop and transition to a nonlinear state.

a. Description

First, we compare the linear instabilities captured by the numerical simulations with the predictions from the theory outlined above. We find that for sufficiently small Rossby numbers, the predicted angles of the instability and growth rates closely match the analytical predictions. This is perhaps not surprising since the theory is developed in the limit of asymptotically small Rossby number. However, by comparing the simulations and theory, we can quantify how large the Rossby number can be before the analytical theory breaks down.

Figure 3 illustrates the development and nonlinear breakdown of the unstable modes from a simulation with $\mathrm{Ro} = 10^{-3}$ and $\alpha = 0.4$. Here, the depth-averaged buoyancy field is plotted, where the background buoyancy gradient, B, has been removed. For reference, the unperturbed basic state is $\overline{b} = Bx$, which would have vertical buoyancy contours in this figure. At a relatively early time ($t = 0.314$, upper right panel), growing perturbations develop with a distinctive angle with respect to the buoyancy gradient. Note that the fastest growing modes occur on a larger scale compared to the initial perturbations, suggesting a scale-selective process. By $t = 0.384$ (lower left panel) the flow transitions to a nonlinear regime and the growing perturbations roll up into

coherent vortices. These vortices then merge resulting in an energy cascade to larger scales (see lower right panel).

Figure 4 shows two simulations with different values of α during the period when the perturbations are linear and the growth is exponential. The theoretical prediction for the direction of the fastest growing modes, $k = -\alpha l$, is plotted as a black dashed line. As predicted, the wave crests of the most unstable modes are nearly perpendicular to the predicted wavenumber vector.

Figure 5 shows the buoyancy perturbation from four simulations with different values of the Rossby number. In all cases, $\alpha = 0.4$, and hence the predicted angle of the most unstable modes is the same. The wavelength of the most unstable modes changes with Ro, but interestingly the dependence is not monotonic. For the range of Ro tested, the shortest waves are observed for $\mathrm{Ro} = 10^{-3}$. For $\mathrm{Ro} = 10^{-4} - 10^{-2}$ the direction of the wavevector is independent of Ro and closely matches the theoretical prediction.

In the case with the largest Rossby number, $\mathrm{Ro} = 0.1$, the fastest growing mode does not fit in the domain, and instead a quantized mode with $(k,l) = (0, 2\pi)$ appears. There also appear to be growing perturbations at an angle nearly perpendicular to the analytical prediction of $k = -\alpha l$. These modes might be an indication of symmetric instability modified by vertical mixing, although this is not captured by our theory and we do not focus on it here.

b. Energetics

To describe the dynamics of the unstable modes, it is useful to diagnose the perturbation energy budgets. To start, we define the horizontal domain average to be

$$\langle f \rangle = \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} f \, dx \, dy, \tag{73}$$

and let $\tilde{f} = f - \langle f \rangle$ denote the departure from the horizontal average. The nondimensional energy equation can be derived from the governing equations and written in conservative form as

$$\mathrm{Ro}\,\frac{\partial \mathcal{K}}{\partial t} + \nabla \cdot [(\mathrm{Ro}\,\mathcal{K} + p)\mathbf{u}] - b\,w = -\alpha \left(uu' + vv' \right), \tag{74}$$

for kinetic energy density

$$\mathcal{K} = \frac{1}{2}\left(u^2 + v^2 + \varepsilon^2 w^2 \right), \tag{75}$$

assuming that the diffusive terms are small. We now consider the perturbation kinetic energy,

$$e = \frac{1}{2}\left\langle \tilde{u}^2 + \tilde{v}^2 + \varepsilon^2 \tilde{w}^2 \right\rangle, \tag{76}$$

and using Eq. 74 and the horizontally averaged governing equations, the perturbation energy budget can be written

$$\begin{aligned} \mathrm{Ro}\left[\frac{\partial e}{\partial t} + \langle w \rangle \frac{\partial e}{\partial z} + \underbrace{\langle \tilde{u}\tilde{w} \rangle \frac{\partial \langle u \rangle}{\partial z} + \langle \tilde{v}\tilde{w} \rangle \frac{\partial \langle v \rangle}{\partial z}}_{\mathcal{S}} + \underbrace{\frac{1}{2}\frac{\partial}{\partial z}\left\langle [\tilde{u}^2 + \tilde{v}^2 + \varepsilon^2 \tilde{w}^2]\tilde{w} \right\rangle}_{\mathcal{T}} \right] = \\ \underbrace{-\frac{\partial}{\partial z}\langle \tilde{p}\tilde{w} \rangle}_{\mathcal{P}} + \underbrace{\langle \tilde{b}\tilde{w} \rangle}_{\mathcal{B}} \underbrace{-\alpha \left\langle \tilde{u}'\tilde{u} + \tilde{v}'\tilde{v} \right\rangle}_{\mathcal{R}}. \end{aligned} \tag{77}$$

The terms in Eq. 77 can be interpreted as $\mathcal{S}$: production of perturbation kinetic energy by the mean shear, $\mathcal{T}$: turbulent transport, $\mathcal{P}$: pressure transport, $\mathcal{B}$: buoyancy flux, and $\mathcal{R}$: dissipation by the parameterized vertical mixing. From mass conservation $\partial \langle w \rangle / \partial z = 0$ and hence $\langle w \rangle = \text{const.}$ and using the vertical boundary conditions we have that $\langle w \rangle = 0$. We can now vertically average Eq. 77 to remove the transport terms. The resulting equation for the domain averaged perturbation kinetic energy is

$$\mathrm{Ro}\,\frac{\partial \overline{e}}{\partial t} = \overline{\mathcal{S}} + \overline{\mathcal{F}} + \overline{\mathcal{R}}. \tag{78}$$

The depth-averaged dissipation associated with the vertical relaxation term is given by

$$\overline{\mathcal{R}} = -\alpha \left[\overline{\langle \tilde{u}'^2 \rangle} + \overline{\langle \tilde{v}'^2 \rangle} \right] = -\alpha \left[\overline{\langle \tilde{u}^2 \rangle} + \overline{\langle \tilde{v}^2 \rangle} - \left\langle \overline{\tilde{u}}^2 \right\rangle - \left\langle \overline{\tilde{v}}^2 \right\rangle \right], \tag{79}$$

which is negative by the Cauchy-Schwarz inequality. We now calculate the four terms in Eq. 78 using our numerical data and consider the energy balance in order to determine the energy source and mechanism for the instability.

Figure 6 shows the four terms in Eq. 78 for four different values of (Ro, α). Clear regions of exponential growth (with constant slope on the semi-log plot) develop in each case. When $\alpha = 0$, corresponding to the classical Eady model, we can see that the dominant energy balance is between the time rate of change in kinetic energy and the buoyancy flux, representing the transformation of perturbation potential energy into perturbation kinetic energy and indicative of baroclinic instability. For nonzero α, the dominant balance is between the buoyancy flux and the dissipation associated with the vertical relaxation term, with the residual corresponding to the time rate of change of kinetic energy. Therefore, in the presence of vertical mixing, the instability is driven by a transfer of potential energy from the buoyancy field consistent with baroclinic instability, although most of the energy extracted from the potential energy reservoir is dissipated through the vertical mixing (relaxation) term. We note that the balance between $\overline{\mathcal{F}}$ and $\overline{\mathcal{R}}$ is closer for smaller Ro which is consistent with the asymptotic theory. Once the instability reaches the nonlinear phase, the neglected viscous dissipation term becomes significant due to the appearance of small scale vortices.

c. Growth Rate

In this section, we diagnose the growth rate of the unstable perturbations from the numerical simulations and compare these with the prediction from the analytical theory. We define the growth rate of perturbations captured in the numerical simulations by

$$\sigma_N = \frac{1}{2\overline{e}} \frac{d\overline{e}}{dt}. \tag{80}$$

This can then be compared with the theoretical growth rate, σ, given by the eigenmode decomposition in Eq. 50. We now define $\sigma_{avg}(t_1,t_2)$ to be the average of σ_N in the time interval $[t_1,t_2]$ and $\sigma_{rms}(t_1,t_2)$ to be the RMS deviation from this average. Specifically,

$$\sigma_{avg}(t_1,t_2) = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} \sigma_N(t)\,dt, \tag{81}$$

and

$$\sigma_{rms}(t_1,t_1) = \left[\frac{1}{t_2-t_1}\int_{t_1}^{t_2} \left[\sigma_N(t)-\sigma_{avg}(t_1,t_2)\right]^2 dt\right]^{1/2}. \tag{82}$$

We define the interval of exponential growth to be the largest time interval in which the ratio of σ_{rms} to σ_{avg} is below a specified tolerance, i.e. $\sigma_{rms}/\sigma_{avg} < \delta$. The value of the growth rate is then taken to be σ_{avg} within the region of exponential growth. We use a tolerance of $\delta = 0.01$ and do not define a growth rate if the region of exponential growth is small or σ_N is strongly oscillatory. We also use 2D discrete Fourier transforms to determine the wavevector of the fastest growing modes in each simulation.

As an illustration of this procedure, Figure 7 shows σ_N diagnosed from four simulations with different values of Ro and α. For large Ro, large oscillations in σ_N prevent us from accurately diagnosing the growth rate for $\alpha > 0.6$. Figure 8 shows the growth rate and wavenumber of the fastest growing modes diagnosed in this way for each simulation. We exclude results for large Ro and α where we are unable to accurately diagnose the growth rate. For $\mathrm{Ro} = 0.1$ the dominant mode is $(k,l) = (0,2\pi)$ which is likely not the fastest growing mode due to the restrictions of the domain size. For small Rossby number, the wavenumber of the fastest growing mode depends on α while for $\mathrm{Ro} \geq 10^{-3}$ it is independent of α. This is an indication that there are different processes controlling the most unstable modes for small and large Ro.

Figure 9 shows the 2D Fourier transform of the depth-averaged buoyancy perturbation for several values of Ro and α. When viscous effects are included the wavenumber associated with the

most unstable mode is given by 63 and can be written as

$$K_{max} = \frac{\mathrm{Ro}\, B}{\sqrt{48}(1+\alpha^2)\varepsilon^2 \mathrm{E}}, \tag{83}$$

by taking $\mathrm{Pr}_E = 1$ and $\mathrm{Bu} = 0$. The dependence of K_{max} on α matches the simulations for $\mathrm{Ro} = 10^{-4}$ (see Figure 8). The circles on Figure 9 have radius given by Eq. 83 and we can see that the numerical results match the predictions of fastest growing wavenumber for $\mathrm{Ro} = 10^{-4}$. However, for larger Ro the fastest growing wavenumber is significantly smaller than the theoretical prediction. It appears that there is a second, α independent effect which controls the fastest growing modes and is not captured by the theory. This will be examined further in §7. Note from Eq. 83 that K_{max} depends on the aspect ratio, ε, when viscous effects set the scale of the most unstable mode. However, as seen in Figure 8, the scale of the most unstable mode for $\mathrm{Ro} > 10^{-3}$ appears to be independent of viscosity (and independent of the aspect ratio).

Along the direction $k = -\alpha l$, the growth rate is given by Eq. 60. For $\mathrm{Bu} = 0$ and $\mathrm{Pr}_E = 1$, the growth rate of the most unstable mode given in Eq. 60 can be written

$$\sigma_{max} = \frac{BK}{\sqrt{12}(1+\alpha^2)} - \frac{\varepsilon^2 \mathrm{E} K^2}{\mathrm{Ro}}. \tag{84}$$

Figure 10 shows a comparison between the growth rates predicted by Eq. 84 (left panel) and the growth rates diagnosed from the numerical simulations (right panel) where the wavenumber corresponding to the most unstable mode as diagnosed in the numerical simulations is used to set K in Eq. 84. There is very good agreement between the growth rates from the theory and simulations across a wide range of Rossby numbers and relaxation ratios. Interestingly, the growth rates match reasonably well even in cases where the most unstable wavenumber in the theory (Eq. 83) doesn't match the most unstable wavenumber diagnosed in the simulations (e.g. the cases in the bottom panels of Figure 9).

7. QG analysis of a stratified basic state

The numerical simulations described above indicate that the wavenumber of the most unstable mode is set by a process other than viscosity for the larger values of Ro. In this section we use the quasi-geostrophic (QG) equations to examine the stability of a depth-dependent basic state associated with vertical mixing of momentum and buoyancy. Specifically the velocity and buoyancy of the basic state will be given by Eqns. 16 and 19. Importantly, here the stratification of the basic state is non-zero and is the result of a balance between cross-front advection and vertical mixing. Since the stratification in Eq. 19 appears at $O(\mathrm{Ro})$, it did not appear in the basic state analyzed in §5. Here, we also assume that departures from the basic state are not directly affected by vertical or horizontal mixing. This allows us to isolate the influence of vertical mixing on the background flow from its influence on the growing perturbations.

The total velocity and buoyancy fields can be written as

$$(u,v,w,b) = \left(U+\hat{u}, V+\hat{v}, \hat{w}, Bx+N^2z+\hat{b}\right), \tag{85}$$

where capital letters denote the basic state and $\hat{\cdot}$ denotes a perturbation to the basic state. The nondimensional QG equation can be written

$$\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{1}{\mathrm{Ro}}\frac{\partial}{\partial z}\left(\frac{1}{N^2}\frac{\partial}{\partial z}\right)\right]\hat{\psi}=0, \tag{86}$$

where the streamfunction satisfies $\hat{u}=-\partial\hat{\psi}/\partial y$ and $\hat{v}=\partial\hat{\psi}/\partial x$. Applying the boundary condition $w=0$ at $z=\pm 1/2$ to the buoyancy equation gives

$$\left[\frac{\partial^2}{\partial t\partial z}+U\frac{\partial^2}{\partial x\partial z}+V\frac{\partial^2}{\partial y\partial z}-B\frac{\partial}{\partial y}\right]\hat{\psi}=0, \tag{87}$$

where the nondimensional buoyancy perturbation is $\hat{b}=\partial\hat{\psi}/\partial z$ using the QG approximation. From Eqns. 16 and 19 we now write

$$(U,V,N^2) = \left(-\alpha\gamma Bz,\ \gamma Bz,\ \mathrm{Ro}\,\mathrm{Pr}_\alpha\gamma B^2\right), \tag{88}$$

and note that N^2 describes the stratification that results from the balance between cross-front advection and vertical mixing. Eq. 86 has solutions of the form

$$\hat{\psi} = [A\sinh\kappa z + C\cosh\kappa z]\, e^{ikx+ily+\sigma t}, \tag{89}$$

for $\kappa = \sqrt{\mathrm{Ro}N^2(k^2+l^2)}$ and following Vallis (2006) we Eq. 87 to determine a linear system for (A,C). The requirement that the determinant of this system vanishes determines the growth rate, which can be written

$$\sigma^2 = \frac{B^2}{\kappa^2}\left[\gamma(l-\alpha k)\frac{\kappa}{2} - l\tanh\frac{\kappa}{2}\right]\left[l\coth\frac{\kappa}{2} - \gamma(l-\alpha k)\frac{\kappa}{2}\right]. \tag{90}$$

We note that this result reduces to the classical Eady result (Eady 1949) for $\alpha = 0$. Working in our rotated coordinate system (k',l'), we can show that σ is maximal for $k = -\alpha l$ where

$$\sigma^2 = \frac{\gamma B^2}{\mathrm{Ro}N^2}\left[\frac{\kappa}{2} - \tanh\frac{\kappa}{2}\right]\left[\coth\frac{\kappa}{2} - \frac{\kappa}{2}\right]. \tag{91}$$

Therefore, following Eady (1949) and Vallis (2006), we have maximum growth rate

$$\sigma_{max} = \frac{0.31B}{\sqrt{\mathrm{Ro}\,(1+\alpha^2)}N}, \tag{92}$$

for most unstable wavenumber

$$K_{max} = \frac{1.6}{\sqrt{\mathrm{Ro}}N}. \tag{93}$$

Using N^2 from Eq. 88 this result becomes

$$\sigma_{max} = \frac{0.31}{\mathrm{Ro}\sqrt{\mathrm{Pr}_\alpha}}, \tag{94}$$

and

$$K_{max} = \frac{1.6\sqrt{1+\alpha^2}}{\mathrm{Ro}\sqrt{\mathrm{Pr}_\alpha}B}. \tag{95}$$

Therefore the most unstable mode is set by the interaction of edge waves, moderated by the stratification that develops in response to vertical mixing of momentum. Since this stratification is an

$O(\mathrm{Ro})$ term in the buoyancy equation, it does not appear in the leading order evolution equations for b_0 or ψ_0 (equations 47 and 48), and hence its influence on the unstable modes is not captured by our asymptotic model.

The QG predictions for σ_{max} and K_{max} are shown in Figure 11 as functions of α and Ro. Comparing these results with Figure 8 we find that they provide reasonably accurate predictions for the growth rate and wavenumber for $\mathrm{Ro} > 10^{-3}$ where the scale is not set by horizontal diffusion and the wavenumber of the most unstable mode decreases with increasing Ro. However, the growth rate from the QG analysis (Eq. 94) is less accurate than the prediction from the asymptotic theory (Eq. 84) when compared with the numerical simulations. For example, the growth rate in Eq. 94 is independent of α, while the prediction in 84 and the growth rate diagnosed from the simulations decrease with increasing α. This suggests that vertical mixing acts to damp the perturbations and reduces their growth rate. Nevertheless, the estimate from Eq. 94 still provides a reasonable approximation to the growth rate.

We expect the mixing-induced stratification to limit the size of the most unstable modes when it would give a smaller value of K_{max} than horizontal diffusion. Using Eq. 83, this occurs when

$$\frac{\mathrm{Ro}\, B}{\sqrt{48}(1+\alpha^2)\varepsilon^2 \mathrm{E}} > \frac{1.6\sqrt{1+\alpha^2}}{\mathrm{Ro}\sqrt{\mathrm{Pr}_\alpha} B}. \tag{96}$$

Therefore the mixing-induced stratification will be important when

$$\mathrm{Ro} > \frac{3.3(1+\alpha^2)^{3/4}\varepsilon\, \mathrm{E}^{1/2}}{\mathrm{Pr}_\alpha^{1/4} B}. \tag{97}$$

For the parameters used in our numerical simulations (specifically $\varepsilon = 0.05$ and $\mathrm{E} = 10^{-4}$), this condition is satisfied for $\mathrm{Ro} \gtrsim 10^{-3}$, consistent with our observations that the fastest growing mode is not set by horizontal diffusion for this parameter range. We note that using a turbulent Ekman number scaling of $\mathrm{E} \sim u^*/fH$ for turbulent velocity u^* and mixed layer depth H can give values of E on the order of $10^{-2} - 1$. Therefore in a highly turbulent mixed layer, a diffusive cutoff may

be possible for Rossby numbers up to about Ro ~ 0.1. Mathematically, this can help us explain the apparent inaccuracy of our original prediction for the fastest growing mode. Since a very small value of ε^2E was used in our simulations, the horizontal mixing terms were smaller than any terms describing the $O(\mathrm{Ro})$ stratification, N^2, even for small values of Ro. Our theory assumes that these horizontal mixing terms are dominant and hence we have discrepancies for much smaller values of Ro than might be anticipated. If we were to instead use a much larger value of E for our numerical simulations, representative of a turbulent Ekman number, we would find agreement over a much wider range of Ro as described by Eq. 97. It is also worth re-emphasizing that the lengthscale used to define Ro is the horizontal domain size in the simulations. Since the most unstable mode is typically much smaller than the domain size (see Fig. 3), the scale-dependent Rossby number associated with the size of the most unstable mode will be much larger than Ro.

The angle of the most unstable mode from the QG analysis agrees with the theory in §5 and the simulations. Therefore, we can conclude that the orientation of the most unstable modes are primarily set by the background flow and is not strongly influenced by the effects of vertical mixing acting directly on the perturbations.

For small K the growth rate from the QG analysis (Eq. 90) becomes

$$\sigma \sim \frac{BK}{\sqrt{12(1+\alpha^2)}\mathrm{Bu}}, \tag{98}$$

for $\mathrm{Bu} = \mathrm{Ro}N^2$. This does not have the same α dependence as Eq. 64, although we note that both expressions reduce to the classical Eady case for $\alpha = 0$. This discrepancy is likely because the QG approach does not consider the action of vertical mixing on the perturbations.

The analysis of the instability using the QG equations also provides insight into the relative accuracy of growth rate from the analytical theory. In the absence of horizontal mixing, the analytical theory predicted that the growth rate is a linearly proportional to the horizontal wavenumber (see

Eq. 64). As shown in Eq. 98, the growth rate in the QG analysis also increases linearly with K for small values of K, while stratification decouples the Eady edge waves and suppresses the growth rate for large K. However, the maximum growth rate in the QG analysis is relatively close to the value that would be obtained by using the wavenumber of the fastest growing mode in Eq. 98, which has the same form as the theory in §5. We expect that the true growth rate for the problem admits both a viscous cutoff and a decoupled edge wave cutoff and reduces to the analytical result for small K. If the maximum growth rate in the case of decoupled edge waves is close to the linear, small K region (as is the case in the QG model) then the analytical theory would well describe the growth rate even though it does not capture the cutoff mechanism. This may explain why our growth rate predictions in Figure 10 closely match the numerical simulations.

Note that the Richardson number of this system can be shown to be $\mathrm{Ri} = \mathrm{Pr}_\alpha$ hence it would be more accurate to use the ageostrophic analysis of Stone (1966). This analysis can be performed using the background state in Eq. 88, although it is much more complicated than the QG analysis. Including non-QG effects reduces the growth rate of the most unstable mode (σ_{max}) by a factor of $\sqrt{1+\mathrm{Ri}}$ and it somewhat reduces the wavenumber of the most unstable mode (K_{max}). Importantly, the dependence of σ_{max} and K_{max} on Ro and α are unchanged by the inclusion of non-QG effects, and hence we use the QG equations here for simplicity.

Recall from figure 3 that modes with $l = \alpha k$ appeared in the simulation with $Ro = 0.1$ which were perpendicular to the anticipated most unstable mode. Setting $l = \alpha k$ in the QG analysis gives

$$\sigma^2 = -\frac{B^2 l^2}{\kappa^2}. \tag{99}$$

These modes are stable and correspond to travelling waves. Therefore, the perpendicular modes observed in Figure 5 do not appear to arise through QG dynamics, and are likely associated with finite Ro effects which we have not considered here.

Finally, we note that a similar QG analysis could be carried out for the TTW system with vertical mixing parameterized using a Laplacian viscosity and diffusivity, as described in Appendix C. However, in this case Eq. 86 would have to be solved numerically since N^2 depends on z.

8. Conclusions and Discussion

Here, we examined baroclinic instability in the presence of vertical mixing, where mixing is parameterized using a simple relaxation towards the local depth average. A theory was developed which is valid in the limit of small Rossby number, but arbitrary mixing rates. In the limit of no mixing we recover the long wave limit of baroclinic instability in the Eady model. Vertical mixing reduces the growth rate and tilts the unstable modes such that they are aligned with the horizontal velocity, with the angle determined by the relaxation timescale.

In the absence of horizontal mixing and a turbulent Prandtl number of 1, the growth rate associated with the fastest growing modes (from Eq. 61 with $B = 1$) is

$$\sigma = \frac{\mathrm{Ro}}{\alpha\,(1+\alpha^2)\,\mathrm{Bu}}, \tag{100}$$

where σ is nondimensionalised by $1/T = HM^2/(fL)$, H is the mixed layer depth, L is a characteristic horizontal length scale, M^2 is the horizontal buoyancy gradient, and f is the Coriolis parameter. The nondimensional parameters in Eq. 100 are the Rossby number, $\mathrm{Ro} = M^2H/(f^2L)$, the Burger number, $\mathrm{Bu} = N^2H^2/(f^2L^2)$, where N is the buoyancy frequency associated with a stable background stratification, and the mixing ratio, $\alpha = \mu/f$, where μ is the vertical mixing rate. Note that the horizontal length scale, L, characterizes the width of the front and not necessarily the size of the unstable modes. Indeed, Figure 8 shows that the non-dimensional wavenumber of the most unstable modes is $K >> 1$ and therefore the scale-dependent Rossby number associated with the growing perturbations will be significantly larger than Ro.

The theoretical growth rate in Eq. 100 decreases with decreasing Ro (e.g. for weak horizontal buoyancy gradients) and decreases with increasing vertical mixing rate. In the absence of a background stratification (Bu = 0) the growth rate is unbounded. However, when a horizontal Laplacian viscosity and diffusivity is included to parameterize horizontal mixing, the growth rate is bounded and equal to

$$\sigma = \frac{\mathrm{Ro}}{12\,(1+\alpha^2)^2\left[\sqrt{\varepsilon^2\mathrm{E}}+\sqrt{\varepsilon^2\mathrm{E}+\frac{\alpha\,\mathrm{Bu}}{12(1+\alpha^2)}}\right]^2}, \tag{101}$$

where $\mathrm{E} = \nu/(fH^2)$ is the Ekman number, ν is the horizontal viscosity (equal to the diffusivity since the Prandtl number is assumed to be 1) and H is the mixed layer depth.

The theory developed here is valid for asymptotically small Rossby numbers (although the vertical mixing rate can be large). To test the range of validity of the theory, we conducted a series of numerical simulations. The growth rate and wavenumbers predicted by the theory match those diagnosed from the simulations very closely for small Rossby numbers. The predicted growth rate matches the simulations for Rossby numbers up to $O(0.1)$. However for $\mathrm{Ro} > O(10^{-3})$, the most unstable modes in the simulations are significantly larger than those predicted from the theory. This implies that in this range of Ro, the neglected higher order terms become important and provide a scale selecting mechanism.

To investigate this further, we used the quasi-geostrophic (QG) equations to analyze the stability of a depth-dependent basic state. Here the density of the basic state was set through a balance between cross-front advection and vertical mixing. Since the stratification that results from this balance appears at $O(\mathrm{Ro})$ it was not included in the theory described earlier. We also neglected the direct influence of vertical mixing on the perturbations when applying the QG equations. Results from the QG analysis show that the horizontal orientation of the fastest growing modes is largely inherited from the orientation of the background flow. The QG analysis also shows that the

stratification that develops from the cross-front flow can decouple the Eady edge waves, thereby providing a high wavenumber cutoff. Although this effect was not included in the theory presented in §5, the growth rate predicted from our theory agrees well with the growth rate diagnosed from the numerical simulations, even in parts of parameter space where the most unstable wavenumber is not set by a viscous cutoff.

Motivated by this, we can combine the predicted growth rate from our theory with the high wavenumber cutoff from the QG analysis. To put the results in the context of typical ocean conditions, it is useful to normalize the growth rate by f and write it as a function of $|\nabla b|/f^2$, which has the effect of eliminating the dependence of the growth rate on the aspect ratio. In the case with $\mathrm{Bu} = 0$, $\mathrm{Pr}_E = 1$, and a non-dimensional horizontal buoyancy gradient $B = 1$, the growth rate given in Eq. 84 can be written

$$\frac{\sigma}{f} = \frac{K\mathrm{Ro}}{\sqrt{12}(1+\alpha^2)} - \frac{f^4}{|\nabla b|^2}\mathrm{E}K^2\mathrm{Ro}^2. \tag{102}$$

When the size of the most unstable mode is limited by horizontal mixing,

$$K\mathrm{Ro} = \frac{|\nabla b|^2}{\sqrt{48}f^4(1+\alpha^2)\mathrm{E}}, \tag{103}$$

and when it is limited by the influence of mixing-induced stratification on the interaction between Eady edge waves,

$$K\mathrm{Ro} = 1.6\sqrt{1+\alpha^2}. \tag{104}$$

When vertical and horizontal mixing are described using the same characteristic turbulent velocity u_* and length scale, l, we have $\alpha \sim \mathrm{E} \sim u_*/(fl)$ (see Eq. 7).

Figure 12 shows the growth rate prediction from Eq. 102 with $\mathrm{E} = \alpha$, and $K\mathrm{Ro}$ set by the minimum of Eqns. 103 and 104. The dashed line separates regions where the most unstable mode is controlled by horizontal mixing through Eq. 103 (the region below the line) and mixing-induced stratification through Eq. 104 (the region above the line). The symbols show typical parameters

corresponding to winter and summer conditions based on the observations reported in Thompson et al. (2016). In the winter when submesoscale activity was clearly observed, the size of the most unstable mode is limited by mixing-induced stratification and the growth rate from Eq. 102 is close to the inviscid prediction from Stone (1966). However, for parameters more typical of conditions during summer, Figure 12 suggests that mixing associated with small-scale turbulence can limit the size and dramatically reduce the growth rate of the unstable modes. This result might help explain the apparent absence of submesoscale activity in the summer months.

For the inviscid Eady problem (here with $\alpha = 0$), the dominant source of energy for the growing baroclinic modes is the buoyancy flux. Here, we find that for relatively large mixing rates, the dominant energy balance is between the buoyancy flux and the dissipation associated with vertical mixing, with the small residual corresponding to the kinetic energy growth. Therefore the instability is driven by a transfer of potential energy to the growing perturbations, consistent with baroclinic instability, although now most of the energy is dissipated by vertical mixing, which acts to reduce the growth rate of the unstable modes.

For larger values of α and Ro, the numerical simulations show evidence of small scale modes with $l = \alpha k$ that are perpendicular to the predicted direction. These might be associated with the skew flux term in the leading order buoyancy equation (Eq. 27) which cannot be neglected for large Ro and acts to destabilize modes with $l = \alpha k$. The energy budget suggests that these modes have a different energy source involving both the buoyancy flux and the shear production. These modes have not been studied in detail here since our theory is not valid for this range of parameters.

As noted above, the growth rate of the most unstable mode predicted from our theory matches the numerical simulations up to a Rossby number of about 0.1. This range includes many open ocean fronts. For example, based on a year-long timeseries from the OSMOSIS campaign, Thompson et al. (2016) found that the strongest fronts observed had $|\nabla b| \sim 10^{-7}\mathrm{s}^{-2}$. For mixed layer depths

in the range $20-200$m and a horizontal scale between 20km, this corresponds to a Rossby number ranging from $0.01-0.1$. Stronger and/or sharper fronts such as the Gulf Stream (Thomas et al. 2013) are likely to be strongly influenced by the relatively large Rossby numbers characterizing these fronts, and our results might not be applicable.

The turbulent thermal wind (TTW) model considered in Wenegrat and McPhaden (2016); McWilliams (2016); Crowe and Taylor (2018) used a large turbulent Ekman number instead of relaxation to represent mixed layer turbulence. As shown in Appendix C, the TTW model is also susceptible to the instability described here, and we expect that any turbulence parametrization in which the leading order velocity is linear in the buoyancy gradient will exhibit the same instability.

We have approached the stability problem by seeking analytical solutions to the asymptotic equations which are valid for small Rossby numbers. Another approach would be to solve the linearized equations numerically, without making any assumptions about the size of the Rossby number. This could be viewed as an extension to Stone (1970) and Stamper and Taylor (2017) with the addition of vertical mixing. This would permit non-geostrophic processes such as symmetric instability which are not included in the limit of small Ro.

Here, we have assumed that the relaxation ratio, α is constant which effectively prescribes the vertical mixing rate. This allows us to isolate and study the influence of vertical mixing on mixed layer instabilities, but the assumption of constant α does not allow the instabilities to modify the vertical mixing rate. Previous studies (e.g. Taylor and Ferrari 2011; Taylor 2016) have found that the stable stratification induced by baroclinic and symmetric instability significantly reduces the rate of vertical mixing. We speculate that a reduction in α would enhance the growth rate of the unstable modes, providing a positive feedback mechanism. This hypothesis could be tested in future work.

Acknowledgments. The authors are grateful to two anonymous referees for constructive comments. The authors would also like to thank Jörn Callies, Raffaele Ferrari, Leif Thomas, and Jacob Wenegrat for stimulating and helpful conversations. We are particularly grateful to Bill Young for very helpful discussions and for comments on a draft of the paper.

APPENDIX A

Transient Solution

In §3 we calculated the asymptotic solution for $\mathbf{u}_0$, b_0 and b_1' for the long time evolution. Here we include the transient evolution on the timescale τ for an initial flow with arbitrary vertical structure (e.g. thermal wind flow). Again the leading order buoyancy is assumed to be depth independent but we allow the buoyancy deviation, b_1', to have arbitrary initial vertical structure. This setup allows us to initialize the flow in thermal wind balance with a depth independent buoyancy, the transient evolution causes the front to slump over and the velocity to develop a cross-front component.

a. Order 1 Equations

The leading order buoyancy balance is

$$\frac{\partial b_0}{\partial \tau} + \frac{\alpha}{\mathrm{Pr}_\alpha} b_0' = 0, \tag{A1}$$

so we take solution with b_0 to be depth independent and hence independent of τ, therefore $b_0 = b_0(x,y,t)$.

The leading order velocity balance is given by

$$\frac{\partial u_0}{\partial \tau} - v_0 = -\frac{\partial p_0}{\partial x} - \alpha u_0', \tag{A2a}$$

$$\frac{\partial v_0}{\partial \tau} + u_0 = -\frac{\partial p_0}{\partial y} - \alpha v_0', \tag{A2b}$$

$$0 = -\frac{\partial p_0}{\partial z} + b_0, \tag{A2c}$$

$$\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} + \frac{\partial w_0}{\partial z} = 0, \tag{A2d}$$

hence the pressure can be written as

$$p_0 = z b_0 + \overline{p}_0, \tag{A3}$$

and the horizontal momentum equations and mass conservation equation can be depth averaged to give

$$\frac{\partial \overline{u}_0}{\partial \tau} - \overline{v}_0 = -\frac{\partial \overline{p}_0}{\partial x}, \tag{A4a}$$

$$\frac{\partial \overline{v}_0}{\partial \tau} + \overline{u}_0 = -\frac{\partial \overline{p}_0}{\partial y}, \tag{A4b}$$

$$\frac{\partial \overline{u}_0}{\partial x} + \frac{\partial \overline{v}_0}{\partial y} = 0. \tag{A4c}$$

Subtracting the depth-averaged horizontal momentum equations from equations A2 gives evolution equations for the horizontal velocity perturbations and vertical velocity

$$\left[\frac{\partial}{\partial \tau} + \alpha\right] u_0' - v_0' = -z\frac{\partial b_0}{\partial x}, \tag{A5a}$$

$$\left[\frac{\partial}{\partial \tau} + \alpha\right] v_0' + u_0' = -z\frac{\partial b_0}{\partial y}, \tag{A5b}$$

$$\frac{\partial u_0'}{\partial x} + \frac{\partial v_0'}{\partial y} + \frac{\partial w_0}{\partial z} = 0. \tag{A5c}$$

Equations A5 can be solved to get

$$\mathbf{u}_{H0}' = A_1(z,\tau)\,\nabla_H b_0 + A_2(z,\tau)\,\mathbf{k}\times\nabla_H b_0, \tag{A6}$$

and

$$w_0 = A_3(z,\tau)\nabla_H^2 b_0. \tag{A7}$$

where

$$A_1 = \frac{-\alpha z}{1+\alpha^2} + \left[A_1^0(z) + \frac{\alpha z}{1+\alpha^2}\right] e^{-\alpha\tau}\cos\tau + \left[A_2^0(z) - \frac{z}{1+\alpha^2}\right] e^{-\alpha\tau}\sin\tau, \tag{A8a}$$

$$A_2 = \frac{z}{1+\alpha^2} + \left[A_2^0(z) - \frac{z}{1+\alpha^2}\right] e^{-\alpha\tau}\cos\tau - \left[A_1^0(z) + \frac{\alpha z}{1+\alpha^2}\right] e^{-\alpha\tau}\sin\tau, \tag{A8b}$$

$$A_3 = \frac{\alpha(z^2-\frac{1}{4})}{2(1+\alpha^2)} - e^{-\alpha\tau}\cos\tau\int_{-1/2}^{z} A_1^0(z') + \frac{\alpha z'}{1+\alpha^2}\,dz' - e^{-\alpha\tau}\sin\tau\int_{-1/2}^{z} A_2^0(z') - \frac{z'}{1+\alpha^2}\,dz', \tag{A8c}$$

where (A_1^0, A_2^0) describes the initial horizontal flow. Once the transients have decayed the balanced solutions are

$$\mathbf{u}'_{H0} = \gamma[-\alpha\nabla_H b_0 + \mathbf{k}\times\nabla_H b_0]\,z, \tag{A9}$$

and

$$w_0 = \frac{\alpha\gamma(4z^2-1)}{8}\nabla_H^2 b_0, \tag{A10}$$

for $\gamma = 1/(1+\alpha^2)$. From the depth-averaged mass conservation equation we can write

$$\overline{\mathbf{u}}_{H0} = -\nabla\times(\psi_0\mathbf{k}), \tag{A11}$$

for streamfunction $\psi_0 = \overline{p}_0$. Hence

$$\mathbf{u}_{H0} = -\nabla\times(\psi_0\mathbf{k}) + A_1(z,\tau)\,\nabla_H b_0 + A_2(z,\tau)\,\mathbf{k}\times\nabla_H b_0. \tag{A12}$$

b. Order Ro Equations

The $O(\mathrm{Ro})$ buoyancy equation is

$$\frac{\partial b_1}{\partial\tau} + \frac{\partial b_0}{\partial t} + \mathbf{u}_{H0}\cdot\nabla_H b_0 + \mathcal{N}^2 w_0 = -\frac{\alpha}{\mathrm{Pr}_\alpha} b_1', \tag{A13}$$

and subtracting the depth average gives

$$\left[\frac{\partial}{\partial\tau}+\frac{\alpha}{\mathrm{Pr}_\alpha}\right]b_1' = -\mathbf{u}_{H0}'\cdot\nabla_H b_0 - \mathcal{N}^2 w_0'. \tag{A14}$$

When $\mathcal{N}^2 = 0$ the general solution is given by

$$b_1' = A_4(z,\tau)|\nabla_H b_0|^2, \tag{A15}$$

where

$$A_4 = \frac{\mathrm{Pr}_\alpha z}{1+\alpha^2} + \left[A_4^0(z) - \frac{\mathrm{Pr}_\alpha z}{1+\alpha^2} + \frac{\left[A_1^0(z)+\frac{\alpha z}{1+\alpha^2}\right]\alpha\left(1-\frac{1}{\mathrm{Pr}_\alpha}\right)+\left[A_2^0(z)-\frac{z}{1+\alpha^2}\right]}{1+\alpha^2\left(1-\frac{1}{\mathrm{Pr}_\alpha}\right)^2}\right]e^{-\frac{\alpha}{\mathrm{Pr}_\alpha}\tau}$$
$$+\left[\frac{\left[A_1^0(z)+\frac{\alpha z}{1+\alpha^2}\right]\left(\sin\tau-\alpha\left(1-\frac{1}{\mathrm{Pr}_\alpha}\right)\cos\tau\right)-\left[A_2^0(z)-\frac{z}{1+\alpha^2}\right]\left(\cos\tau+\alpha\left(1-\frac{1}{\mathrm{Pr}_\alpha}\right)\sin\tau\right)}{1+\alpha^2\left(1-\frac{1}{\mathrm{Pr}_\alpha}\right)^2}\right]e^{-\alpha\tau}, \tag{A16}$$

for initial vertical structure described by $A_4^0(z)$. The general steady state solution is given by

$$b_1' = \mathrm{Pr}_\alpha\gamma\left[z|\nabla_H b_0|^2 - \mathcal{N}^2\frac{12z^2-1}{24}\nabla_H^2 b_0\right], \tag{A17}$$

and calculating $\overline{b}_1$ requires the $O(\mathrm{Ro}^2)$ buoyancy equation.

APPENDIX B

Analytic Solution with Relaxation and Diffusion

If we include vertical diffusion in the leading order velocity balance by taking $\mathrm{E} = O(1)$, we can obtain the solution

$$\mathbf{u}_{H0}' = -\sqrt{\mathrm{E}}\left[B_1\nabla_H b_0 + B_2\,\mathbf{k}\times\nabla_H b_0\right], \tag{B1}$$

where

$$B_1 = \alpha\zeta/(1+\alpha^2) + iC_1\sinh[\sqrt{\alpha+i}\zeta] - iC_2\sinh[\sqrt{\alpha-i}\zeta], \tag{B2a}$$

$$B_2 = -\zeta/(1+\alpha^2) + C_1\sinh[\sqrt{\alpha+i}\zeta] + C_2\sinh[\sqrt{\alpha-i}\zeta], \tag{B2b}$$

and

$$\zeta = z/\sqrt{\mathrm{E}}. \tag{B3}$$

Using boundary conditions of no vertical shear on the top and bottom surfaces gives that

$$\begin{bmatrix} i\sqrt{\alpha+i}\cosh\left[\sqrt{\frac{\alpha+i}{4\mathrm{E}}}\right] & -i\sqrt{\alpha-i}\cosh\left[\sqrt{\frac{\alpha-i}{4\mathrm{E}}}\right] \\ \sqrt{\alpha+i}\cosh\left[\sqrt{\frac{\alpha+i}{4\mathrm{E}}}\right] & \sqrt{\alpha-i}\cosh\left[\sqrt{\frac{\alpha-i}{4\mathrm{E}}}\right] \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \frac{1}{1+\alpha^2}\begin{bmatrix} -\alpha \\ 1 \end{bmatrix}, \tag{B4}$$

which can be inverted to obtain solution

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \frac{1}{2(1+\alpha^2)}\begin{bmatrix} \frac{1+\alpha i}{\sqrt{\alpha+i}\cosh\zeta_{\alpha+}} \\ \frac{1-\alpha i}{\sqrt{\alpha-i}\cosh\zeta_{\alpha-}} \end{bmatrix}, \tag{B5}$$

for

$$\zeta_{\alpha\pm} = \sqrt{\frac{\alpha\pm i}{4\mathrm{E}}} = \sqrt{\alpha\pm i}\,\zeta_0, \tag{B6}$$

and $\sqrt{*}$ denoting the principle value of the square root with branch cut taken along the line $z \in -\mathbb{R}_0^+$. The leading order vertical velocity can be obtained by integrating the mass conservation equation as before which gives solution

$$w_0 = \mathrm{E}\left[\frac{\alpha^2-1}{(\alpha^2+1)^2} + \frac{\alpha(\zeta^2-\zeta_0^2)}{2(1+\alpha^2)} + \frac{iC_1}{\sqrt{\alpha+i}}\cosh[\sqrt{\alpha+i}\zeta] - \frac{iC_2}{\sqrt{\alpha-i}}\cosh[\sqrt{\alpha-i}\zeta]\right]\nabla_H^2 b_0, \tag{B7}$$

for $\zeta_0 = 1/2\sqrt{\mathrm{E}}$.

We can now use this leading order solution for the velocity to calculate the $O(\mathrm{Ro})$ solution for the buoyancy perturbation, the governing equation is

$$\mathbf{u}'_{H0}\cdot\nabla_H b_0 = \frac{\mathrm{E}}{\mathrm{Pr}}\frac{\partial^2 b'_1}{\partial z^2} - \frac{\alpha}{\mathrm{Pr}}b'_1, \tag{B8}$$

hence

$$\left[\frac{\partial^2}{\partial\zeta^2} - \alpha\right] b'_1 = -\sqrt{\mathrm{E}}\,\mathrm{Pr}\,B_1\,|\nabla_H b_0|^2, \tag{B9}$$

which has solution

$$b'_1 = -\sqrt{\mathrm{E}}\,\mathrm{Pr}\,B_2\,|\nabla_H b_0|^2. \tag{B10}$$

For small E this solution reduces to the solution given above in the region away from the boundaries. We use this solution to initialize the numerical simulations so that the initial state matches the no stress boundary conditions and hence will not produce inertial waves while adjusting to a balanced state.

APPENDIX C

Instability in the viscous TTW model

In this section, we analyze the stability of the the Turbulent Thermal Wind (TTW) model used in Crowe and Taylor (2018) where vertical mixing is parameterized by a Laplacian viscosity and diffusivity. We also include a background streamfunction, ψ_0. The Steady state solution is

$$\mathbf{u}_H = -\nabla \times (\psi_0 \mathbf{k}) - \sqrt{\mathrm{E}}\left(K_0'' \nabla_H b_0 + K_0 \mathbf{k} \times \nabla_H b_0\right) + O(\mathrm{Ro}), \tag{C1}$$

$$w = \mathrm{E} K_0' \nabla_H^2 b_0 + O(\mathrm{Ro}), \tag{C2}$$

and

$$b = b_0 - \mathrm{Ro}\,\mathrm{Pr}\sqrt{\mathrm{E}} K_0 |\nabla_H b_0|^2 + O(\mathrm{Ro}^2), \tag{C3}$$

where K_0 and its derivatives are given in Crowe and Taylor (2018). Note that this model does not include a stratification so $\mathrm{Bu} = 0$ and we are using an order 1 Ekman number to describe the effects of turbulence. Here we have used a constant vertical turbulent viscosity and diffusivity profile, this is just for convenience and the resulting equations will be similar for arbitrary vertical profiles.

Using the depth-averaged buoyancy and vorticity equations, equations 23 and 40, and including horizontal diffusion we can write the governing equations for the background buoyancy and streamfunction as

$$\frac{\partial b_0}{\partial t} + J(\psi_0, b_0) = \frac{\varepsilon^2 \mathrm{E}}{\mathrm{Ro}\,\mathrm{Pr}} \nabla_H^2 b_0, \tag{C4}$$

and

$$\frac{\partial \nabla_H^2 \psi_0}{\partial t} + J(\psi_0, \nabla_H^2 \psi_0) - \frac{\varepsilon^2 \mathrm{E}}{\mathrm{Ro}} \nabla_H^4 \psi_0 = \\ \mathrm{E} \nabla_H \cdot \left[\left(2\overline{K_0'^2} \nabla_H b_0 + \left[\overline{K_0''^2} - \overline{K_0^2} \right] \mathbf{k} \times \nabla_H b_0 \right) \nabla_H^2 b_0 \right]. \tag{C5}$$

Note that the right hand side of Eq. C4 can be related to the diabatic PV flux (Thomas 2005; Wenegrat et al. 2018). These equations are of the same form as the governing equations for the relaxation model considered above hence we expect instabilities with the corresponding growth rate

$$\sigma_\pm = -\frac{\varepsilon^2 \mathrm{E}}{2\mathrm{Ro}} \left[1 + \frac{1}{\mathrm{Pr}} \right] (k^2 + l^2) \pm \\ \sqrt{\left[\frac{1}{\mathrm{Pr}} - 1 \right]^2 \frac{\varepsilon^4 \mathrm{E}^2}{4\mathrm{Ro}^2} (k^2 + l^2)^2 - \mathrm{E} B^2 \left[2\overline{K_0'^2} kl + (\overline{K_0''^2} - \overline{K_0^2}) l^2 \right]}, \tag{C6}$$

for frontal gradient B and horizontal wavevector (k, l). In the case $\mathrm{Pr} = 1$ this simplifies to

$$\sigma_\pm = -\frac{\varepsilon^2 \mathrm{E}}{\mathrm{Ro}} (k^2 + l^2) \pm \sqrt{-\mathrm{E} B^2 \left[2\overline{K_0'^2} kl + (\overline{K_0''^2} - \overline{K_0^2}) l^2 \right]}. \tag{C7}$$

The last term in square brackets in equations C6 and C7 is a symmetric quadratic form so can be diagonalized by an orthogonal transformation. Therefore the fastest growing modes for a given wavenumber will be tilted with angle dependent only on a function of the Ekman number. The fastest growing mode for a given wavenumber, $K = \sqrt{k^2 + l^2}$, is

$$\sigma_\pm = -\frac{\varepsilon^2 \mathrm{E}}{2\mathrm{Ro}} \left[1 + \frac{1}{\mathrm{Pr}} \right] K^2 + \sqrt{\left[\frac{1}{\mathrm{Pr}} - 1 \right]^2 \frac{\varepsilon^4 \mathrm{E}^2}{4\mathrm{Ro}^2} K^4 + \lambda \mathrm{E} B^2 K^2}, \tag{C8}$$

for eigenvalue

$$\lambda = -\frac{1}{2} \left[\overline{K_0''^2} - \overline{K_0^2} - \sqrt{4\overline{K_0'^2}^2 + \left(\overline{K_0''^2} - \overline{K_0^2} \right)^2} \right], \tag{C9}$$

with angle

$$\theta = \arctan \left[\frac{\overline{K_0''^2} - \overline{K_0^2} + \sqrt{4\overline{K_0'^2}^2 + \left(\overline{K_0''^2} - \overline{K_0^2} \right)^2}}{2\overline{K_0'^2}^2} \right], \tag{C10}$$

from the down-front direction.

In the case $\mathrm{Pr} = 1$ the maximum growth rate reduces to

$$\sigma_{max} = \sqrt{\lambda \mathrm{E}} BK - \frac{\varepsilon^2 \mathrm{E} K^2}{\mathrm{Ro}}, \tag{C11}$$

which is the analogous result to Eq. 84 and has fastest growing mode

$$K_{max} = \frac{\sqrt{\lambda}\, B \mathrm{Ro}}{2\varepsilon^2 \sqrt{\mathrm{E}}}, \tag{C12}$$

with corresponding growth rate

$$\sigma(K_{max}) = \frac{\lambda B^2 \mathrm{Ro}}{4\varepsilon^2}. \tag{C13}$$

Figure C1 shows the formation of baroclinic instability for $\alpha = 0$, $\mathrm{E} = 0.1$ and $\mathrm{Ro} = 0.01$. We can see that the evolution and structure of the instability is similar to the case of the relaxation parametrisation with modes tilted by the angled TTW flow.

References

Bachman, S. D., and J. R. Taylor, 2016: Numerical simulations of the equilibrium between eddy-induced restratification and vertical mixing. *Journal of Physical Oceanography*, **46 (3)**, 919–935.

Boccaletti, G., R. Ferrari, and B. Fox-Kemper, 2007: Mixed layer instabilities and restratification. *Journal of Physical Oceanography*, **37 (9)**, 2228–2250.

Callies, J., and R. Ferrari, 2013: Interpreting energy and tracer spectra of upper-ocean turbulence in the submesoscale range (1–200 km). *Journal of Physical Oceanography*, **43 (11)**, 2456–2474.

Callies, J., and R. Ferrari, 2018: Baroclinic instability in the presence of convection. *Journal of Physical Oceanography*, **48 (1)**, 45–60.

Callies, J., R. Ferrari, J. M. Klymak, and J. Gula, 2015: Seasonality in submesoscale turbulence. *Nature communications*, **6**, 6862.

Callies, J., G. Flierl, R. Ferrari, and B. Fox-Kemper, 2016: The role of mixed-layer instabilities in submesoscale turbulence. *J. Fluid Mech.*, **788**, 5–41.

Capet, X., E. Campos, and A. Paiva, 2008: Submesoscale activity over the argentinian shelf. *Geophysical Research Letters*, **35 (15)**.

Charney, J. G., 1973: *Planetary Fluid Dynamics*, chap. Symmetric Circulations in Idealized Models, 128–141. D. Reidel Publishing Company.

Crowe, M. N., and J. R. Taylor, 2018: The evolution of a front in turbulent thermal wind balance, part 1: Theory. *J. Fluid Mech.*, **850**, 179–211.

Crowe, M. N., and J. R. Taylor, 2019: The evolution of a front in turbulent thermal wind balance, part 2: Numerical simulations. *J. Fluid Mech.*, **in press**.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1**, 33–52.

Ferrari, R., and W. R. Young, 1997: On the development of thermohaline correlations as a result of nonlinear diffusive parameterizations. *J. Mar. Res.*, **55**, 1069–1101.

Fox-Kemper, B., R. Ferrari, and R. Hallberg, 2008: Parameterization of mixed layer eddies. part i: Theory and diagnosis. *Journal of Physical Oceanography*, **38 (6)**, 1145–1165.

Gula, J., M. J. Molemaker, and J. C. McWilliams, 2014: Submesoscale cold filaments in the gulf stream. *J. Phys. Oceanogr.*, **44**, 2617–2643.

Kato, H., and O. M. Phillips, 1969: On the penetration of a turbulent layer into stratified fluid. *J. Fluid Mech.*, **37**, 643–655.

Lorbacher, K., D. Dommenget, P. Niiler, and A. Köhl, 2006: Ocean mixed layer depth: A subsurface proxy of ocean-atmosphere variability. *Journal of Geophysical Research: Oceans*, **111 (C7)**.

Mahadevan, A., E. D'asaro, C. Lee, and M. J. Perry, 2012: Eddy-driven stratification initiates north atlantic spring phytoplankton blooms. *Science*, **337 (6090)**, 54–58.

McWilliams, J. C., 2016: Submesoscale surface fronts and filaments: Secondary circulation, buoyancy flux, and frontogenesis. *unpublished*.

Mensa, J. A., Z. Garraffo, A. Griffa, T. M. Özgökmen, A. Haza, and M. Veneziani, 2013: Seasonality of the submesoscale dynamics in the gulf stream region. *Ocean Dynamics*, **63 (8)**, 923–941.

Rudnick, D. L., and R. Ferrari, 1999: Compensation of horizontal temperature and salinity gradients in the ocean mixed layer. *Science*, **283 (5401)**, 526–529.

Sasaki, H., P. Klein, B. Qiu, and Y. Sasai, 2014: Impact of oceanic-scale interactions on the seasonal modulation of ocean dynamics by the atmosphere. *Nature communications*, **5**, ncomms6636.

Schumann, U., and T. Gerz, 1995: Turbulent mixing in stably stratified shear flows. *Journal of Applied Meteorology*, **34 (1)**, 33–48.

Shakespeare, C. J., and J. Taylor, 2013: A generalized mathematical model of geostrophic adjustment and frontogenesis: uniform potential vorticity. *J. Fluid Mech.*, **736**, 366–413.

Shay, T. J., and M. C. Gregg, 1986: Convectively driven turbulent mixing in the upper ocean. *J. Phys. Oceanogr.*, **16**, 1777–1798.

914 Stamper, M. A., and J. R. Taylor, 2017: The transition from symmetric to baroclinic instability in
915 the eady model. *Ocean Dynamics*, **67 (1)**, 65–80.

916 Stone, P. H., 1966: On non-geostrophic baroclinic stability. *J. Atmos. Sci.*, **23**, 390–400.

917 Stone, P. H., 1970: On non-geostrophic baroclinic stability: Part ii. *Journal of the Atmospheric*
918 *Sciences*, **27 (5)**, 721–726.

919 Sverdrup, H., 1953: On vernal blooming of phytoplankton. *Conseil Exp. Mer*, **18**, 287–295.

920 Tandon, A., and C. Garrett, 1994: Mixed layer restratification due to a horizontal density gradient.
921 *J. Phys. Oceanogr.*, **24 (6)**, 1419–1424.

922 Taylor, G. I., 1953: Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc.*
923 *Roy. Soc. A.*, **219**, 186–203.

924 Taylor, J. R., 2008: Numerical simulations of the stratified oceanic bottom boundary layer. Ph.D.
925 thesis, University of California, San Diego.

926 Taylor, J. R., 2016: The influence of submesoscale restratification and convection on light-limited
927 phytoplankton growth. *Geophys. Res. Lett.*, 2016GL069106.

928 Taylor, J. R., and R. Ferrari, 2011: Numerical simulations of the competition between wind-driven
929 mixing and surface heating in triggering spring phytoplankton blooms. *Limnol. Oceanogr.*,
930 **56 (6)**, 2293–2307.

931 Thomas, L. N., 2005: Destruction of potential vorticity by winds. *J. Phys. Oceanogr.*, **35**, 2457–
932 2466.

933 Thomas, L. N., J. R. Taylor, R. Ferrari, and T. M. Joyce, 2013: Symmetric instability in the gulf
934 stream. *Deep Sea Research Part II: Topical Studies in Oceanography*, **91**, 96–110.

Thompson, A. F., A. Lazar, C. Buckingham, A. C. Naveira Garabato, G. M. Damerell, and K. J. Heywood, 2016: Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders. *J. Phys. Oceanogr.*, **46 (4)**, 1285–1307.

Thorpe, S. A., 2005: *The turbulent ocean*. Cambridge University Press.

Vallis, G. K., 2006: *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press.

Venayagamoorthy, S. K., and D. D. Stretch, 2010: On the turbulent prandtl number in homogeneous stably stratified turbulence. *Journal of Fluid Mechanics*, **644**, 359–369.

Wenegrat, J., L. N. Thomas, J. Gula, and J. C. McWilliams, 2018: Effects of the submesoscale on the potential vorticity budget of ocean mode waters. *J. Phys. Oceanogr.*, **48**, 2141–2165.

Wenegrat, J. O., and M. J. McPhaden, 2016: A simple analytical model of the diurnal ekman layer. *J. Phys. Oceanogr.*, **46 (9)**, 2877–2894.

Young, W. R., 1994: The subinertial mixed layer approximation. *J. Phys. Oceanogr.*, **24**, 1812–1826.

Young, W. R., and L. Chen, 1995: Baroclinic instability and thermohaline gradient alignment in the mixed layer. *J. Phys. Oceanogr.*, **25**, 3172–3185.

Young, W. R., P. B. Rhines, and C. J. R. Garrett, 1982: Shear-flow dispersion, internal waves and horizontal mixing in the ocean. *J. Phys. Oceanogr.*, **12**, 515–527.

LIST OF TABLES

Parameter	Symbol	Definition
Rossby Number	Ro	$\varepsilon \Delta b / f^2 L$
Relaxation Ratio	α	μ_u / f
Prandtl Number (α)	Pr_α	μ_u / μ_b
Aspect Ratio	ε	H/L
Burger Number	Bu	$N^2 H^2 / f^2 L^2$

TABLE 1: Definitions of the dimensionless parameters and their values for buoyancy difference, Δb, Coriolis parameter, f, background stratification, N^2, horizontal lengthscale, L, vertical lengthscale, H, and momentum and buoyancy relaxation rates, μ_u and μ_b.

Parameter	Symbol	Mahadevan et al. (N. Atl., spring)	Thompson et al. (N. Atl., summer)	Thompson et al. (N. Atl., winter)	Thomas et al. (Gulf Stream, winter)
Mixed layer depth	$H\,(\mathrm{m})$	300	20	200	100
Coriolis parameter	$f\,(\mathrm{s}^{-1})$	1.3×10^{-4}	1.1×10^{-4}	1.1×10^{-4}	9.0×10^{-5}
Horiz. buoyancy grad.	$\lvert\nabla b\rvert\,(\mathrm{s}^{-2})$	7×10^{-9}	10^{-8}	10^{-7}	10^{-7}
Horizontal scale	$L\,(\mathrm{km})$	300	5	15	10
Turbulent velocity	$u_*\,(\mathrm{m\,s}^{-1})$	10^{-2}	10^{-2}	2×10^{-2}	2×10^{-2}
Aspect ratio	ε	10^{-3}	4×10^{-3}	10^{-2}	10^{-2}
Rossby number	$\mathrm{Ro}=\frac{\lvert\nabla b\rvert}{f^2}\varepsilon$	10^{-3}	4×10^{-3}	0.1	0.5
Relaxation ratio	$\alpha=\frac{u_*}{fH}$	0.3	5	1	1

TABLE 2: Estimates of physical scales and nondimensional parameters for three open ocean fronts as estimated based on observations reported in Mahadevan et al. (2012); Thompson et al. (2016); Thomas et al. (2013).

LIST OF FIGURES

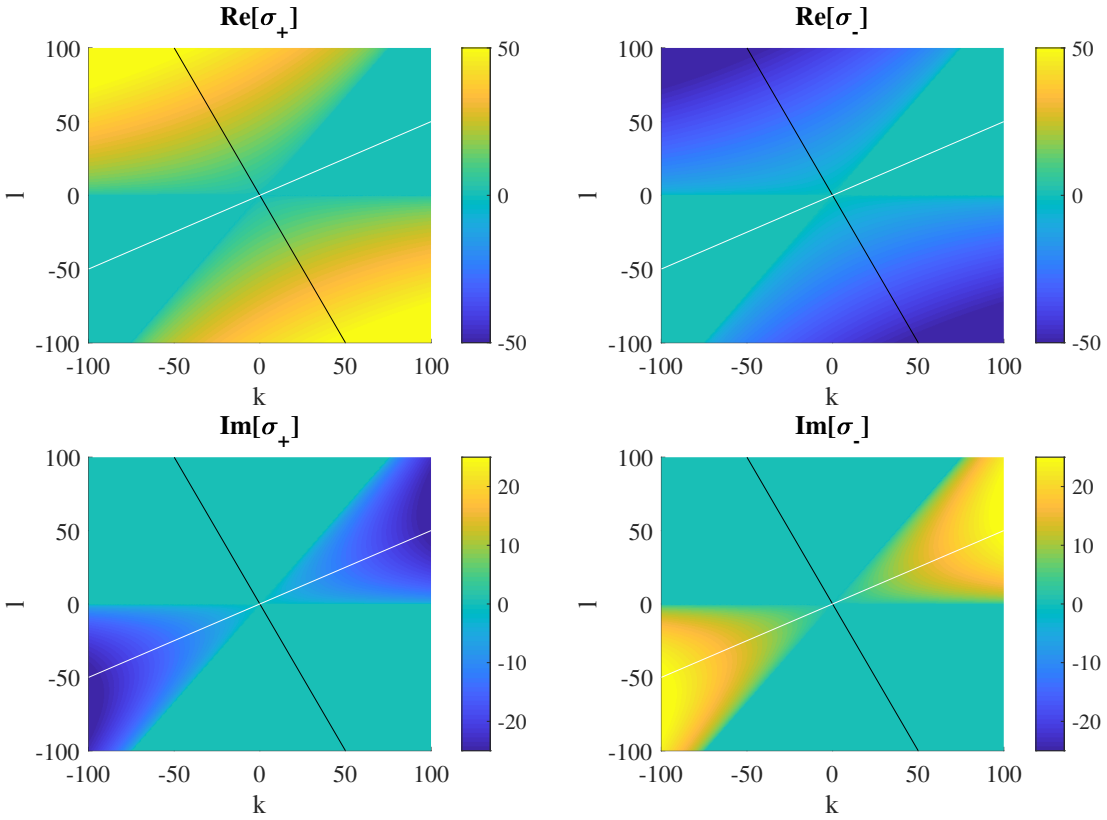

FIG. 1: Real and imaginary parts of the growth rate $\sigma_{\pm}$, predicted from the theory for $\mathcal{E} = \mathcal{D} = 0$ and $\mathcal{B}^2 = 0.213$ corresponding to $\alpha = 0.5$ and $B = 2$. The black lines are $k = -\alpha l$ and the white lines are $l = \alpha k$. Note that the top and bottom rows have different color bars.

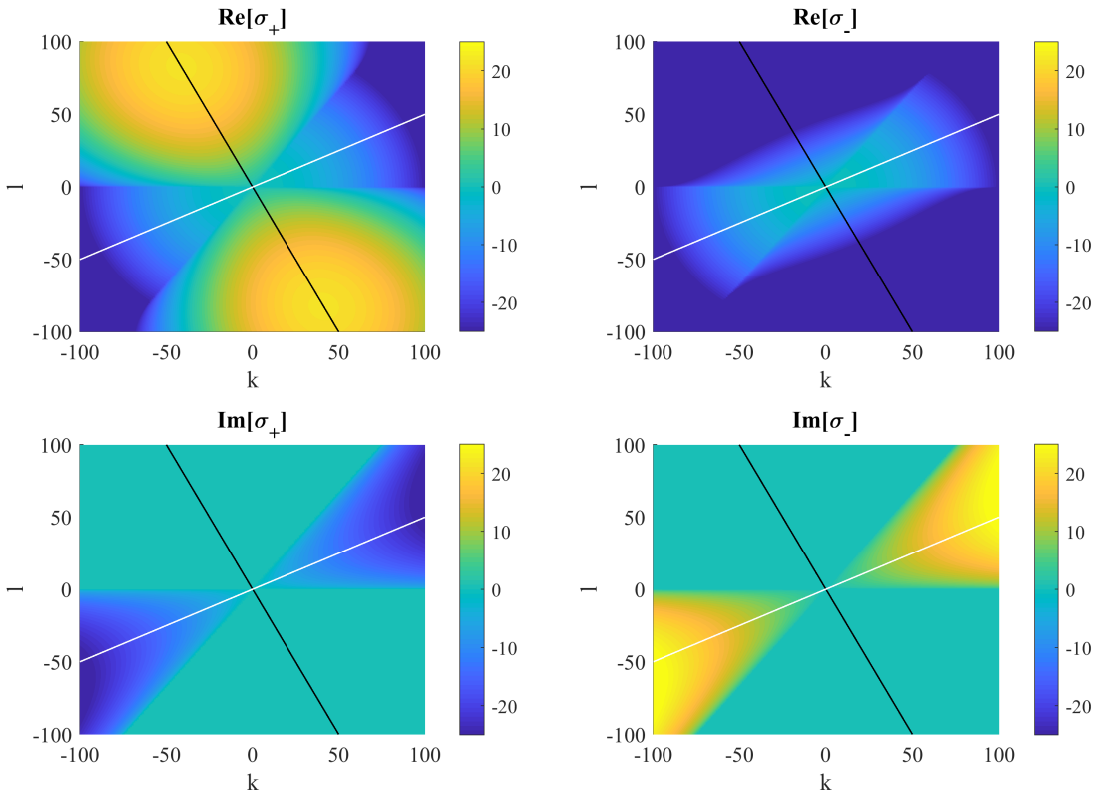

FIG. 2: The real and imaginary parts of the growth rate, $\sigma_{\pm}$, predicted from the theory for $\mathcal{E} = \mathcal{D} = 2.5 \times 10^{-3}$ and $\mathcal{B}^2 = 0.213$ corresponding to $\alpha = 0.5$ and $B = 2$. The black lines are $k = -\alpha l$ and the white lines are $l = \alpha k$. Note that the top and bottom rows have different color bars.

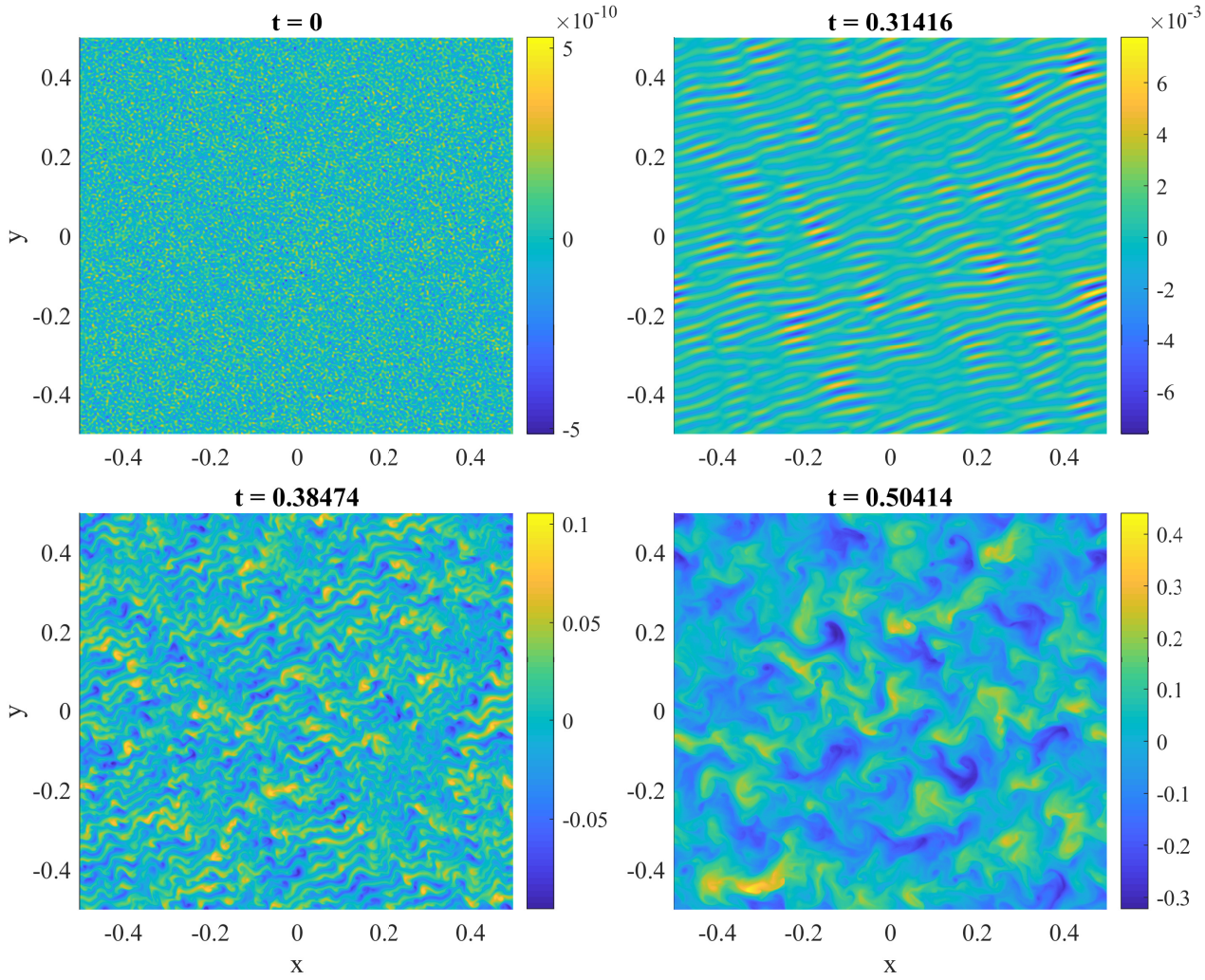

FIG. 3: Depth-averaged buoyancy perturbation, $\overline{b}(x,y) - Bx$ from a nonlinear numerical simulation with $\mathrm{Ro} = 10^{-3}$ and $\alpha = 0.4$ at several times as indicated. The formation of the linear instability and the transition to nonlinear instability can be seen.

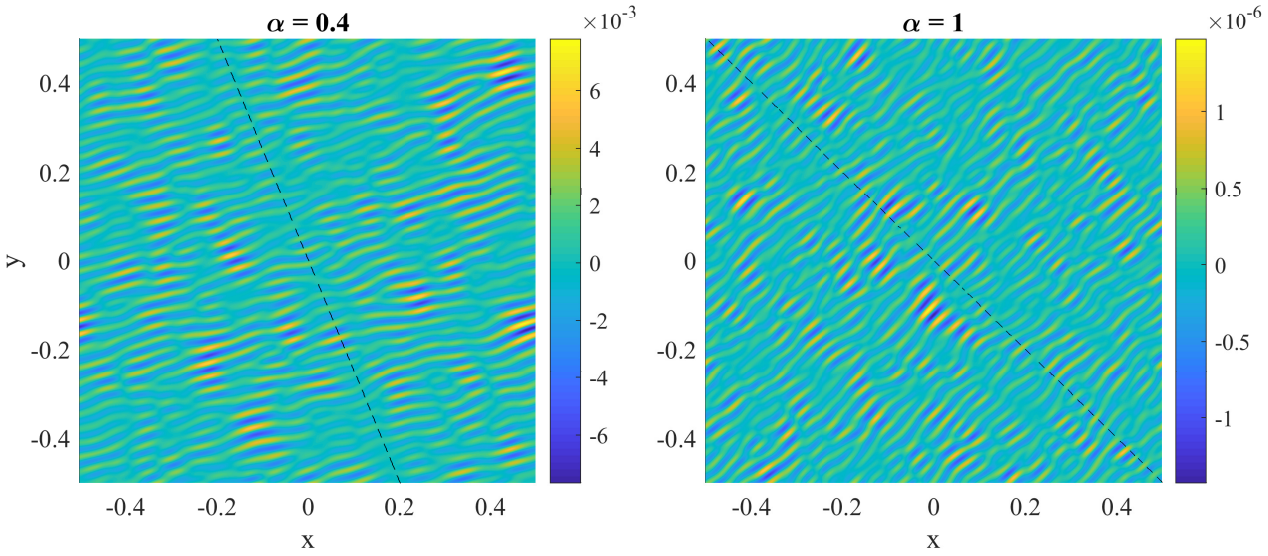

FIG. 4: Depth-averaged buoyancy perturbation, $\overline{b}(x,y) - Bx$ from numerical simulations with $\alpha = 0.4$ and $\alpha = 1$. In both cases $\mathrm{Ro} = 10^{-3}$ and the fields are shown at time $t = 0.314$. The black lines show the predicted wavevector direction, $k = -\alpha l$, which should be perpendicular to lines of constant phase.

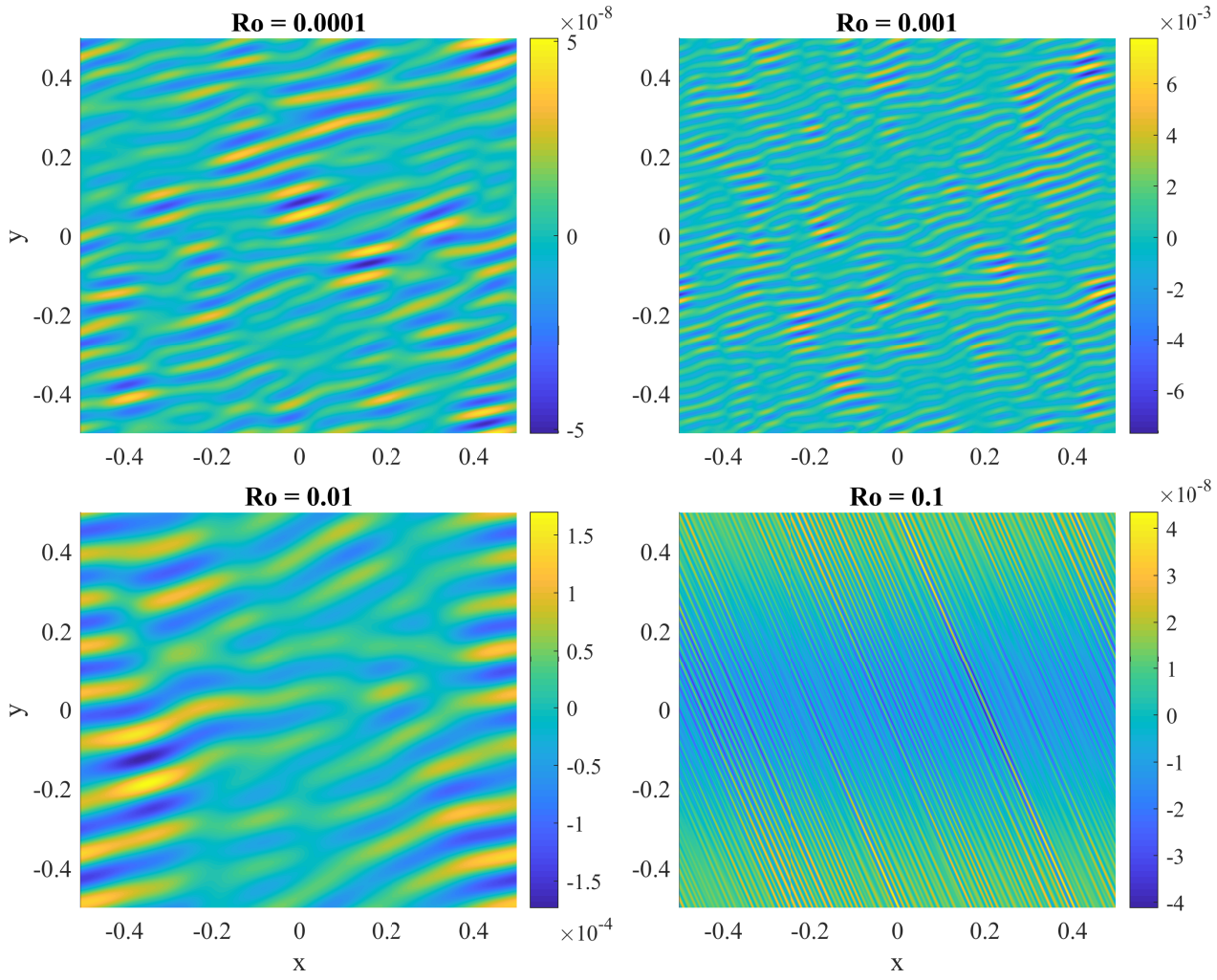

FIG. 5: Depth-averaged buoyancy perturbation, $\overline{b}(x,y) - Bx$ for $\alpha = 0.4$ and several Rossby numbers during the phase of linear perturbation growth in several numerical simulations. For Ro $= 0.1$ growing modes appear which are perpendicular to those predicted by the analytical theory, indicating a breakdown of the theory due to the relatively large Rossby number.

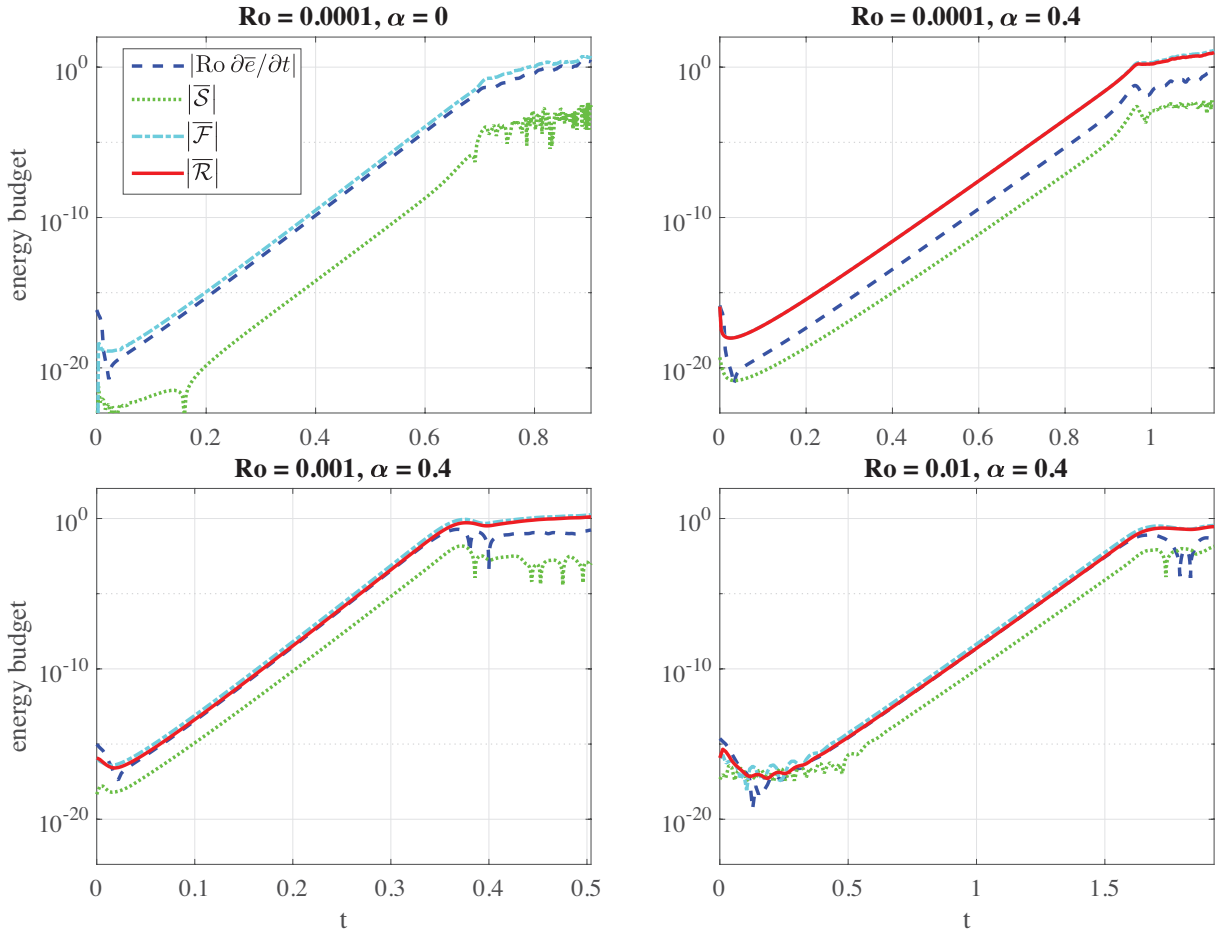

FIG. 6: The magnitude of terms in the volume-averaged energy budget from numerical simulations for several values of Rossby numbers and α. The terms are as given in Eqs. 77 and 78, specifically, the shear production is denoted $\overline{\mathcal{S}}$, buoyancy flux, $\overline{\mathcal{F}}$, and dissipation via vertical mixing (relaxation), $\overline{\mathcal{R}}$. The case of $\alpha = 0$ corresponds to the classical Eady model.

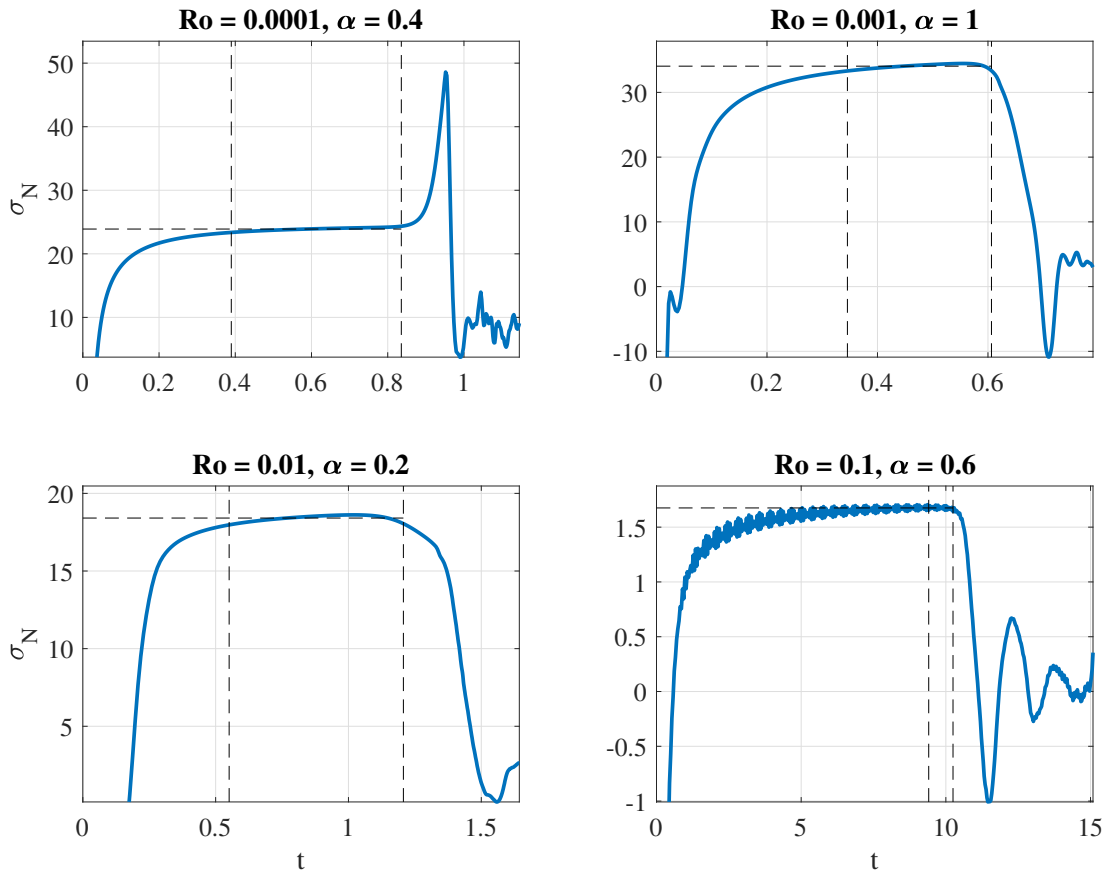

FIG. 7: Perturbation growth rate, σ_N, diagnosed from the numerical simulations for a range of Rossby numbers, Ro and relaxation ratios, α. The dashed lines show the interval of exponential growth and the average value of σ_N within this region.

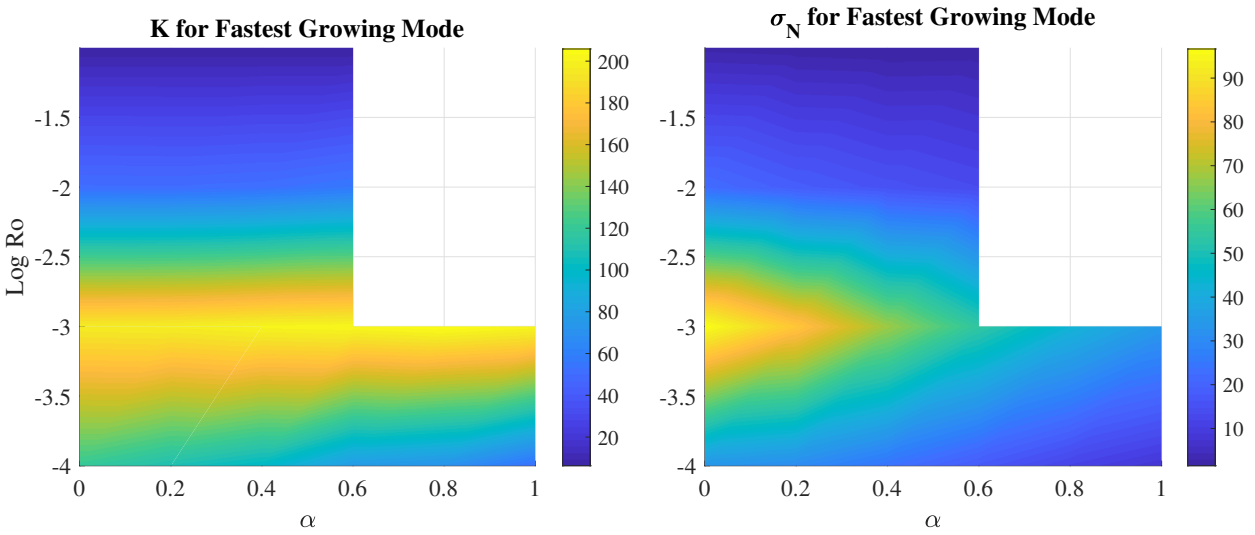

FIG. 8: Perturbation growth rate, σ_N, and the wavenumber, $K = \sqrt{k^2 + l^2}$, for the fastest growing mode inferred from the numerical simulations as functions of Ro and α.

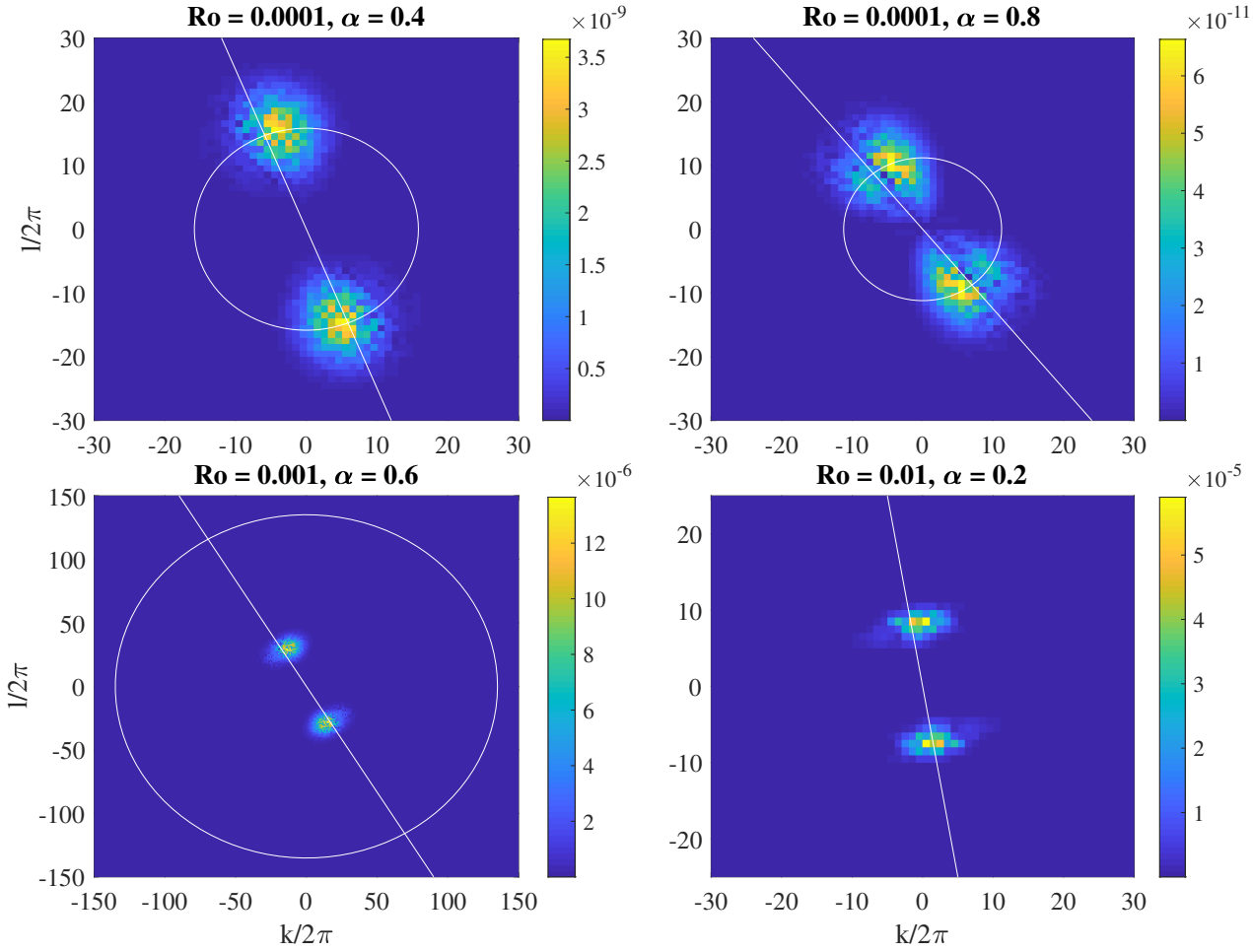

FIG. 9: Amplitudes of the horizontal Fourier coefficients from the depth-averaged buoyancy perturbation, $\overline{b}(x,y) - Bx$, calculated form the numerical simulations for a range of Ro and α. The line $k = -\alpha l$ is shown in white and the maximum wavenumber, K_{max}, is given by the white circle.

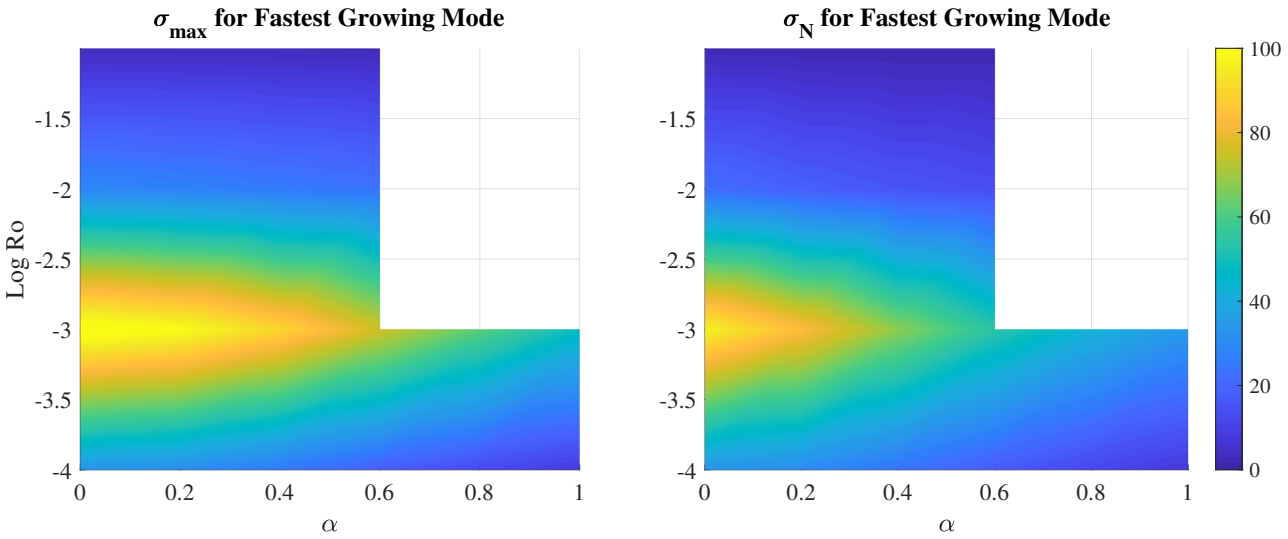

FIG. 10: Perturbation growth rate, σ, diagnosed from the numerical simulations (σ_N) and the maximum growth rate predicted from the theory (σ_{max}).

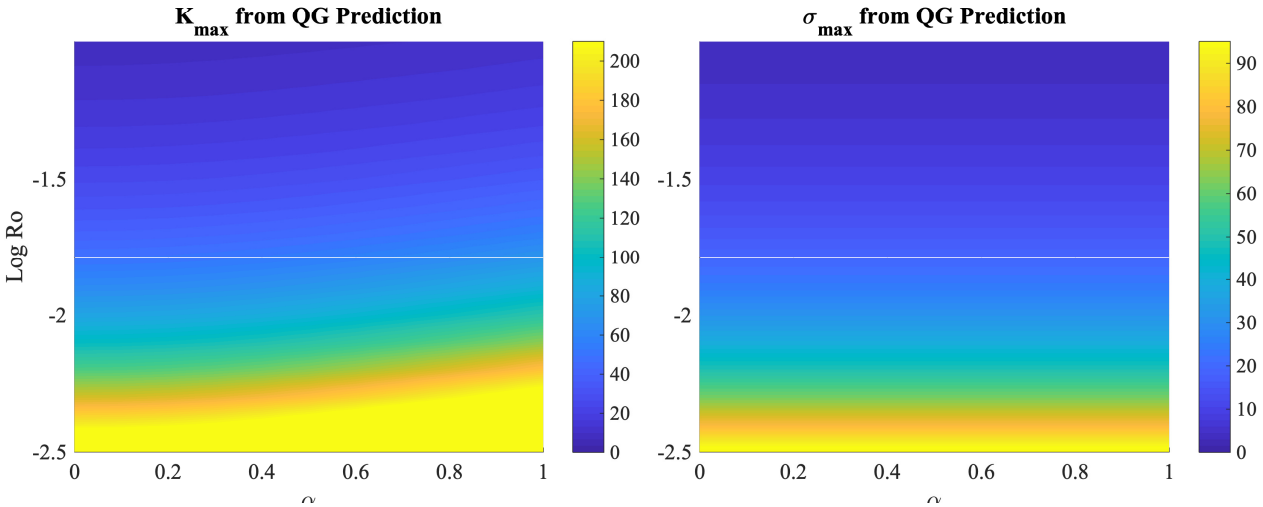

FIG. 11: Wavenumber and growth rate of the most unstable mode from the QG analysis (Eqs. 94 and 95) as functions of Ro and α.

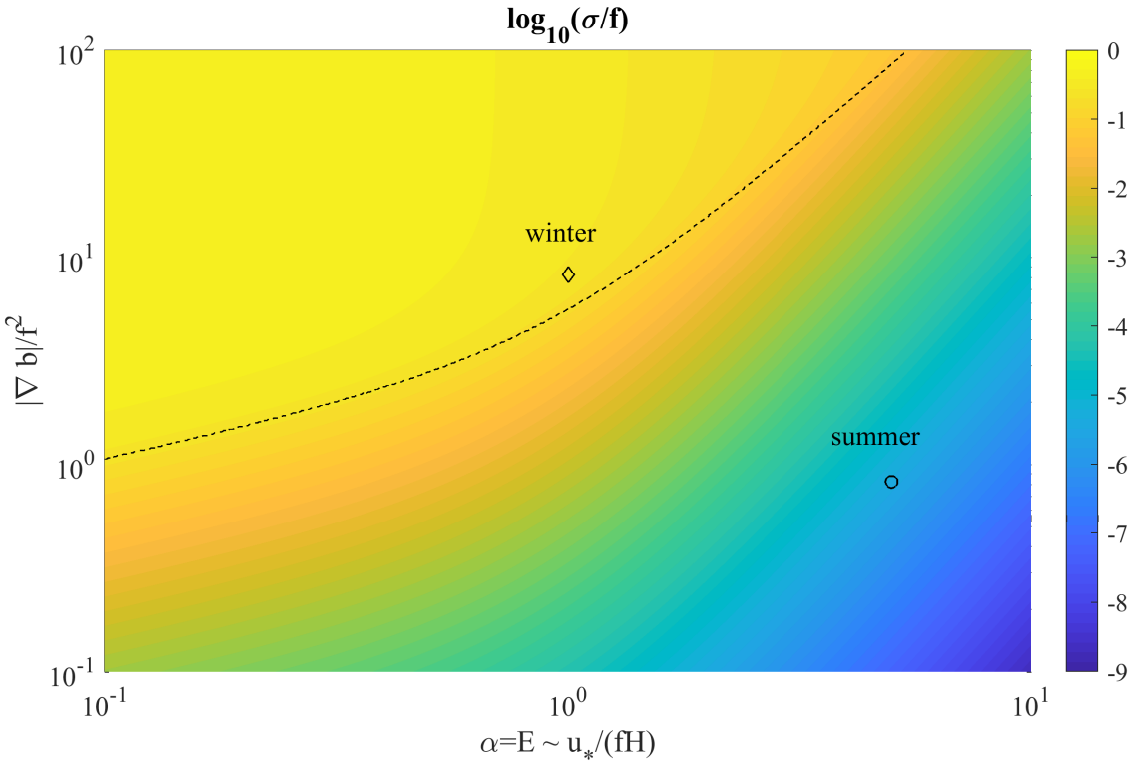

FIG. 12: Predicted growth rate of the most unstable modes from Eq. 102 where the wavenumber is set by the smaller of Eq. 103 and 104. The dashed line separates regions where the wavenumber is set by Eq. 103 (below the line) from regions where the wavenumber is set by Eq. 104 (above the line). The symbols indicate typical parameters from the OSMOSIS survey as reported in Thompson et al. (2016) in the winter and summer (see Table 2 for values).

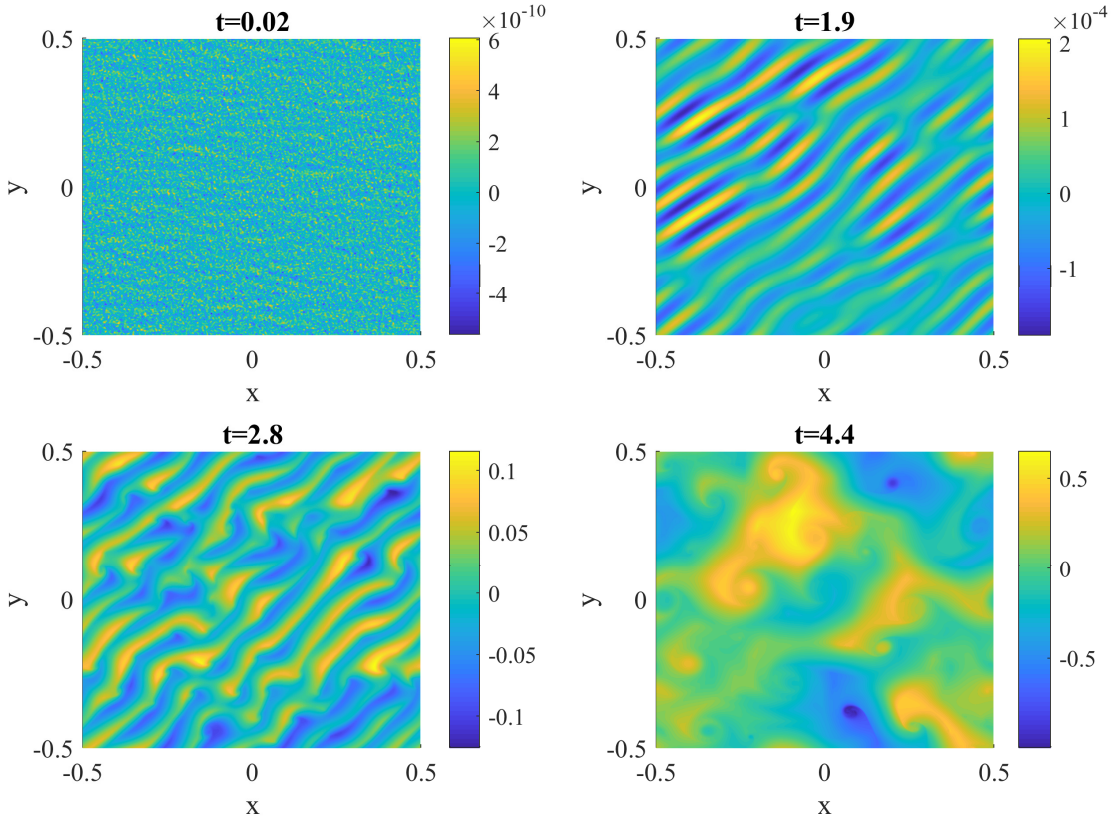

Fig. C1: The formation of baroclinic instability for E = 0.1 and Ro = 0.01. We plot $b_0(x, y, z = 0) - Bx$ as a function of cross-front coordinate x and along-front coordinate y.