

Nonlinear evolution of linear optimal perturbations of strongly stratified shear layers

A. K. Kaminski^{1,2†}, C. P. Caulfield^{3,1}
and J. R. Taylor¹

¹Department of Applied Mathematics and Theoretical Physics, University of Cambridge,
Wilberforce Road, Cambridge CB3 0WA, UK

²College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, 104 CEOAS
Administration Building, Corvallis, OR 97331, USA

³BP Institute, University of Cambridge, Madingley Rise, Madingley Road, Cambridge CB3
0EZ, UK

(Received ?; revised ?; accepted ?. - To be entered by editorial office)

The Miles-Howard theorem states that a necessary condition for normal-mode instability in parallel, inviscid, steady stratified shear flows is that the minimum gradient Richardson number, $Ri_{g,min}$ is less than $1/4$ somewhere in the flow. However, the non-normality of the Navier-Stokes and buoyancy equations may allow for substantial perturbation energy growth at finite times. We calculate numerically the linear optimal perturbations which maximize the perturbation energy gain for a stably-stratified shear layer consisting of a hyperbolic tangent velocity distribution with characteristic velocity U_0^* and a uniform stratification with constant buoyancy frequency N_0^* . We vary the bulk Richardson number $Ri_b = N_0^{*2} h^{*2} / U_0^*$ (corresponding to $Ri_{g,min}$) between 0.20 and 0.50 and the Reynolds numbers $Re = U_0^* h^* / \nu^*$ between 1000 and 8000, with the Prandtl number held fixed at $Pr = 1$. We find the transient growth of non-normal perturbations may be sufficient to trigger strongly nonlinear effects and breakdown into small-scale structures, thereby leading to enhanced dissipation and non-trivial modification of the background flow even in flows where $Ri_{g,min} > 1/4$. We show that the effects of nonlinearity are more significant for flows with higher Re , lower Ri_b , and higher initial perturbation amplitude E_0 . Enhanced kinetic energy dissipation is observed for higher- Re and lower- Ri_b flows, and the mixing efficiency, quantified here by $\varepsilon_p / (\varepsilon_p + \varepsilon_k)$ where ε_p is the dissipation rate of density variance and ε_k is the dissipation rate of kinetic energy, is found to be approximately 0.35 for the most strongly nonlinear cases.

Key words:

1. Introduction

Vertical mixing in the stably-stratified ocean influences the fluxes of heat, carbon, nutrients, and other important tracers, and couples with the global overturning circulation (Wunsch & Ferrari 2004; Kuhlbrodt *et al.* 2007). Turbulence and mixing in the stably-stratified ocean interior has traditionally been associated with instability of the shear associated with internal waves generated via the action of tides interacting with seafloor bathymetry and winds blowing over the sea surface. Under certain circumstances, such

† Email address for correspondence: kaminska@coas.oregonstate.edu

stratified shear flows may be susceptible to a Kelvin-Helmholtz-type instability, leading to a burst of enhanced mixing and dissipation (Garrett 2003; Smyth & Moum 2012). Regions of turbulence in the ocean are unsteady in time and inhomogeneous in space, leading to difficulties in measuring and quantifying mixing (Ivey *et al.* 2008).

The Miles-Howard theorem quantifies the stability of parallel inviscid steady shear flows to linear instability via the gradient Richardson number, $Ri_g = N^2/S^2$, which is the ratio between the stabilizing effects of stratification (represented by the buoyancy frequency N) and the destabilizing effects of shear (given by S , the vertical shear). Flows with $Ri_g \geq 1/4$ everywhere are stable to linear normal-mode perturbations in a steady flow; having $Ri_g < 1/4$ somewhere in the flow is a necessary (though not sufficient) condition for the KH instability (Miles 1961; Howard 1961). In flows which do exhibit instabilities such as the canonical ‘Kelvin-Helmholtz’ (KH) instability of inflectional shear layers, typically the initial perturbations grow and saturate, become susceptible to secondary instabilities, and eventually break down into small-scale turbulence and mixing via a ‘zoo’ of secondary instabilities (see e.g. Peltier & Caulfield 2003; Mashayek & Peltier 2012). However, it is very important to appreciate that for such free shear layers, the turbulence which arises is inherently transient. Due to the enhanced dissipation associated with the turbulence, these flows inevitably relaminarise eventually, even in the limit as $Ri_g \rightarrow 0$ unless the shear is reinforced by some external mechanism.

Mixing is an inherently small-scale process, not resolved in large-scale circulation models, and so mixing parameterizations are required. Often, these parameterizations employ a critical Richardson number, below which the effects of turbulence are turned on. For example, a critical value of $1/4$, corresponding to the Miles-Howard stability criterion, is sometimes employed (Mellor & Yamada 1982; Price *et al.* 1986; Kunze *et al.* 1990). However, the most appropriate choice of this critical value, and indeed whether such a critical value exists at all, is unclear (Polzin 1996). In particular, despite the correlation between enhanced dissipation with low- Ri flows, there is observational (Polzin 1996; Mack & Schoeberlein 2004), experimental (Augier *et al.* 2014), and numerical evidence (Riley & deBruynKops 2003) indicating that increased dissipation may also be found at higher Richardson numbers. It should be noted that these higher dissipation values are not necessarily in violation of the Miles-Howard theorem, owing to differences between the idealized flow considered in deriving the theorem and the more complicated flows possible in the ocean and laboratory, as well as limitations in measurements. However, they may also suggest that mechanisms other than the classical normal-mode instabilities may be at play in triggering transition to turbulence and mixing in stably stratified fluids.

More recent work has shown that it is possible for disturbances to grow in stratified flows even if the Miles-Howard stability criterion is met and all normal-mode perturbations are stable (Farrell & Ioannou 1993b; Kaminski *et al.* 2014). Owing to the non-normality of the Navier-Stokes equations, it is possible for perturbations to grow at finite times and then subsequently decay in these ostensibly ‘stable’ flows (Trefethen *et al.* 1993; Farrell & Ioannou 1993a). These ‘optimal perturbations’ (i.e. the perturbations associated with maximum energy growth over some finite time horizon) are able to exploit linear transient growth mechanisms such as the Orr mechanism (Orr 1907) and the lift-up mechanism (Ellingsen & Palm 1975) to amplify perturbation amplitude over finite time horizons. Even in flows which admit a normal-mode instability, the perturbation which grows the most at finite times may not be the predicted normal-mode (Tearle 2004; Arratia *et al.* 2013). Farrell & Ioannou (1993b), in investigating a uniformly-stratified constant shear flow, found that at finite times perturbations were able to grow even for flows stable in the Miles-Howard sense. In fact, though the maximum perturbation

growth decreased with increasing stratification, no particular significance was associated with the critical value of $Ri = 1/4$. More recently, as reported in Kaminski *et al.* (2014), we computed the linear optimal perturbations of a flow with uniform background stratification and hyperbolic tangent shear. We showed that while unstable normal-mode perturbations dominated at long target times, at shorter target times the perturbations leading to maximum growth were not necessarily the normal modes predicted by the Taylor-Goldstein equation. Additionally, $O(10 - 100)$ linear perturbation energy gain could be attained even in flows with no predicted unstable normal modes. As we will show here, higher Reynolds numbers can lead to even larger transient growth.

In addition to linear non-normal growth, it is also possible to find *nonlinear* optimal perturbations for a given base flow. Nonlinear optimal perturbations have been found for a variety of unstratified flows, including plane Couette flow (Rabin *et al.* 2012; Duguet *et al.* 2013), pipe flow (Pringle & Kerswell 2010), and boundary layers (Cherubini *et al.* 2015). Recently, nonlinear optimal perturbations have also been identified for stably stratified plane Couette flow, where the presence of stratification is shown to lead to altered structures compared to their unstratified counterparts (Eaves & Caulfield 2015). In general, the structure of the nonlinear optimal perturbations closely matches that of the linear optimal perturbations at small initial amplitude, and becomes increasingly localized in space as the initial perturbation amplitude is increased, until some critical amplitude at which transition to turbulence is able to occur. Nonlinear optimal perturbations have been shown to unpack and grow via a sequence of essentially linear transient growth mechanisms which communicate via nonlinearity as described in Duguet *et al.* (2013) and Kerswell *et al.* (2014).

In free shear flows at high Reynolds number which are susceptible to primary instabilities of ‘Kelvin-Helmholtz’ type, highly disordered motions with substantially enhanced dissipation are commonly observed to occur due to the development and subsequent break down of secondary instabilities on the primary Kelvin-Helmholtz billow (Caulfield & Peltier 2000). However such a turbulent phase in such flows is inevitably transient, as the enhanced dissipation extracts energy from the background flow, leading to eventual laminarization in the absence of any mechanisms by which the background shear is reinforced. Nevertheless, the onset of such transient turbulence is commonly referred to as ‘transition’ in such flows, and is associated in stratified shear flows at sufficiently high Re with the most dynamically significant irreversible mixing of the density field.

Given the substantial linear perturbation growth predicted for finite times, it is natural to question what the fate of these perturbations would be if allowed to evolve in a fully nonlinear fashion. In particular, we are interested in whether the transient growth mechanisms may allow a transition to (inevitably transient) turbulence even when the flow is sufficiently strongly stratified such that the transition is not mediated by the growth of a linear instability. Since the concept of being ‘strongly stratified’ is associated with Ri_g being sufficiently large everywhere in the flow domain, we shall refer to such a transition as an ‘ Ri -subcritical transition’, relying as it does on the eventual development of a finite amplitude perturbation sufficient to trigger transient turbulent motions. It is important to stress that this is not to be confused with the concept of subcritical transition in bounded unstratified shear flows such as pressure-driven flows in pipes or in plane Poiseuille flow, or plane Couette flow between two parallel relatively moving planes, where the ensuing flow is sustained in a statistically-steady turbulent state through the continued forcing.

Exploring this key issue of Ri -subcritical transition is the central aim of this paper. We find that the degree to which nonlinear effects are important in the evolution of the linear optimal perturbations is increased for higher Re , higher initial perturbation

amplitude E_0 , and lower Ri_b . In some cases, the perturbations can grow sufficiently large to trigger the onset of secondary instabilities and break down into transient turbulent motions, characterised by markedly enhanced dissipation rates despite the base flow being linearly stable to a primary Kelvin-Helmholtz instability, and so we can indeed observe an Ri -subcritical transition to turbulence, in the above-defined sense.

The remainder of the paper is organized as follows. In section 2, we describe the method by which the linear optimal perturbations are computed for a range of Richardson and Reynolds numbers. The resulting linear optimal perturbations for different Re are presented in section 3.1. These perturbations are then given finite initial amplitude and used as initial conditions for fully nonlinear direct numerical simulations. A qualitative description of the computed nonlinear flow evolution is presented in section 3.2, and the evolution of mean quantities and turbulent energetics are presented in sections 3.3 and 3.4, respectively. We discuss our results and draw conclusions in section 4.

2. Formulation

The base flow we consider is a stably-stratified shear flow with velocity $U^*(z^*)$ and buoyancy $B^*(z^*)$ given by

$$U^*(z^*) = U_0^* \tanh\left(\frac{z^*}{h^*}\right), \quad B^*(z^*) = N_0^{*2} z^*, \quad (2.1)$$

in which dimensional quantities are denoted by an asterisk. To nondimensionalize, we scale lengths by the half shear layer width h^* , velocities by U_0^* , times by the advective time scale h^*/U_0^* , and buoyancy by $N_0^{*2} h^*$. This gives the nondimensional base flow

$$U(z) = \tanh(z), \quad B(z) = z. \quad (2.2)$$

Using these scales, three nondimensional parameters which control the flow evolution can be formed, namely the Reynolds number Re , Prandtl number Pr , and bulk Richardson number Ri_b , defined as

$$Re = \frac{U_0^* h^*}{\nu^*}, \quad Pr = \frac{\nu^*}{\kappa^*}, \quad Ri_b = \frac{N_0^{*2} h^{*2}}{U_0^{*2}}. \quad (2.3)$$

For the background flow described by (2.2), the minimum gradient Richardson number occurs at the centre of the shear layer and is equivalent to the value of the bulk Richardson number, i.e. $Ri_{g,\min} = Ri_g(z=0) = Ri_b$. We always choose $Pr = 1$ for computational simplicity (as higher values of Pr require higher resolution of the scalar field).

2.1. Linear optimal perturbations

The behaviour of linear perturbations to the background flow given by (2.2) is described by the Boussinesq Navier-Stokes equations, conservation of buoyancy, and mass continuity linearized about the base flow. In nondimensional form, the perturbation governing equations are given by

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial U_i}{\partial x_j} + U_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + Ri_b b \delta_{i3} + \frac{1}{Re} \frac{\partial^2 u_i}{\partial x_j \partial x_j}, \quad (2.4)$$

$$\frac{\partial b}{\partial t} + u_j \frac{\partial B}{\partial x_j} + U_j \frac{\partial b}{\partial x_j} = \frac{1}{Re Pr} \frac{\partial^2 b}{\partial x_j \partial x_j}, \quad (2.5)$$

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (2.6)$$

where \mathbf{u} , b , and p are the velocity, buoyancy, and pressure perturbations, respectively.

We seek the linear optimal perturbations $\mathbf{u}_0(x, y, z, t)$ and $b_0(x, y, z, t)$ for the background flow (2.2), i.e. the perturbations which grow the most over a finite time horizon. As described in Kaminski *et al.* (2014), the problem may be treated as an optimization problem for total perturbation energy gain,

$$G(T) = \frac{\frac{1}{2}(\langle \mathbf{u}, \mathbf{u} \rangle + Ri_b \langle b, b \rangle)}{\frac{1}{2}(\langle \mathbf{u}_0, \mathbf{u}_0 \rangle + Ri_b \langle b_0, b_0 \rangle)}, \quad (2.7)$$

over the finite time interval 0 to T , in which the angle brackets denote the inner product over the volume, $\langle \mathbf{f}, \mathbf{g} \rangle = 1/V \int_V \mathbf{f} \cdot \mathbf{g} dV$. An augmented Lagrangian can be established which constrains perturbation evolution to follow the governing equations, defined as

$$\begin{aligned} \mathcal{L} = & \frac{\langle \mathbf{u}(T), \mathbf{u}(T) \rangle + Ri_b \langle b(T), b(T) \rangle}{\langle \mathbf{u}_0, \mathbf{u}_0 \rangle + Ri_b \langle b_0, b_0 \rangle} \\ & - [\partial_t u_i + u_j \partial_j U_i + U_j \partial_j u_i + \partial_i p - Re^{-1} \partial_j \partial_j u_i - Ri_b b \delta_{i3}, v_i] \\ & - [\partial_t b + u_j \partial_j B + U_j \partial_j b - (RePr)^{-1} \partial_j \partial_j b, \varphi] - [\partial_i u_i, q] \\ & - \langle u_i(0) - u_{0i}, v_{0i} \rangle - \langle b(0) - b_0, \varphi_0 \rangle. \end{aligned} \quad (2.8)$$

In the above, the square brackets denote the spatio-temporal inner product, $[\mathbf{f}, \mathbf{g}] = \int_0^T \langle \mathbf{f}, \mathbf{g} \rangle dt$. Variations of (2.8) with respect to the forward or ‘‘direct’’ state variables \mathbf{u} , b , and p yield the corresponding adjoint equations,

$$\frac{\partial v_i}{\partial t} = v_j \frac{\partial U_j}{\partial x_i} - \frac{\partial}{\partial x_j} (U_j v_i) + \varphi \frac{\partial B}{\partial x_i} - \frac{\partial q}{\partial x_i} - \frac{1}{Re} \frac{\partial^2 v_i}{\partial x_j \partial x_j}, \quad (2.9)$$

$$\frac{\partial \varphi}{\partial t} = - \frac{\partial}{\partial x_j} (U_j \varphi) - Ri_b v_3 - \frac{1}{RePr} \frac{\partial^2 \varphi}{\partial x_j \partial x_j}, \quad (2.10)$$

$$\frac{\partial v_i}{\partial x_i} = 0, \quad (2.11)$$

which govern the evolution of the adjoint velocity \mathbf{v} , adjoint buoyancy φ , and adjoint pressure q , as well as conditions to be satisfied at times $t = 0$ and $t = T$. The direct or forward equations (2.4)-(2.6) and adjoint equations (2.9)-(2.11) are then used in a direct-adjoint looping (DAL) framework to solve for the linear optimal perturbations (see Kaminski *et al.* (2014) for more details of the algorithmic approach).

The flow domain used to compute the linear optimal perturbations has nondimensional size $(L_x, L_y, L_z) = (9.44, 15.0, 60.0)$, with a double hyperbolic-tangent shear layer velocity profile and uniform stable background stratification. This corresponds to the base flow described in (2.2), with an additional shear layer of the opposite sign added above in order to impose periodicity for computational convenience (see also equation (2.2) of Kaminski *et al.* (2014)). Accordingly, the vertical extent for this calculation is twice that used in the fully nonlinear simulations described below. The two shear layers are sufficiently separated that they may be considered independent. The horizontal extent is chosen to be one wavelength of the most unstable mode predicted by the Taylor-Goldstein equation for $Ri_b = 0.20$. We choose this extent as our base choice for two reasons. First, consistently with the previous related studies of Tearle (2004); Arratia *et al.* (2013) and Kaminski *et al.* (2014), for sufficiently long target times in a flow with a linear instability, the optimal perturbation has the same streamwise extent as the most unstable normal mode. Second, our focus is to determine whether the nonlinear time-dependent evolution arising from linear optimal perturbations can trigger (transient) transition. Therefore, it seems natural for ease of comparison to choose the same streamwise extent as Ri_b

is varied. Additional optimal perturbations are computed with larger horizontal and spanwise extents as shown in appendix A, in order to ascertain the importance of domain size on the computed gain. Particular attention is paid to the optimal perturbations computed for the case with $Ri_b = 0.30$ and the larger horizontal extent $(L_x, L_y, L_z) = (12.6, 15.0, 60.0)$ based on two wavelengths of the perturbation structure identified to have optimal linear gain at this value of Ri_b as described in Kaminski *et al.* (2014).

2.2. Fully nonlinear simulations

Direct numerical simulations are carried out by solving the full nonlinear Boussinesq Navier-Stokes and buoyancy conservation equations. The temporal discretization employs a third-order mixed Runge-Kutta-Wray/Crank-Nicolson scheme, in which the viscous terms are treated semi-implicitly and the advective terms are treated explicitly (as described in Taylor (2008)). The horizontal directions are periodic and treated pseudospectrally. In the vertical direction, a second-order finite-difference discretization is used, and the grid points are clustered around the shear layer in the centre of the domain. Sponge layers are used at the top and bottom of the domain to mimic open boundaries.

Linear optimal perturbations are computed as described above, and the structure from the bottom shear layer is subsequently interpolated onto the clustered grid. These perturbations are then given a finite initial amplitude, added onto the background flow (2.2) and used as an initial condition for the direct numerical simulations. Noise with one-tenth the initial perturbation amplitude is also added in order to allow for the possible development of secondary instabilities, and is centred on the shear layer with a Gaussian profile in z . The noise is white up to a cutoff wavenumber corresponding to the maximum wavenumber of the smallest grid in table 1, so as to use the same noise spectrum across all simulations.

Table 1 gives the details of the simulations considered throughout this paper. The Prandtl number is set to $Pr = 1$ for all simulations, while Re and Ri_b are varied. The initial perturbation amplitude E_0 (nondimensionalized by U_0^{*2}/h^{*2}) ranges from 5×10^{-6} to 5×10^{-5} . The grid spacing is determined by the minimum Kolmogorov scale η_{\min} such that $\Delta x_{\min} \leq 2.1\eta_{\min}$, as suggested by Pope (2000). The domain size is set to $(L_x, L_y, L_z) = (9.44, 15.0, 30.0)$ (except for the case $Ri_b = 0.30$, in which a domain of size $(L_x, L_y, L_z) = (12.6, 15.0, 30.0)$ is also considered).

3. Results

3.1. Linear optimal perturbations

Linear optimal perturbations are computed for $Re = 1000 - 8000$, $Ri_b = 0.20, 0.30, 0.35, 0.40, 0.50$, and $Pr = 1$, with a target time of $T = 15$, as shown in figure 1. We choose to focus on this particular value of T as it is to the target time for which maximum transient growth occurs as shown in Kaminski *et al.* (2014). Crucially and conveniently, at this target time, the optimal perturbations are two-dimensional with a horizontal structure corresponding to one or two wavelengths of the domain chosen here. The perturbations take the form of rolls tilted against the background shear which subsequently exploit the Orr mechanism for transient growth as they are tilted upwards over the time interval $(0, T)$ (Orr 1907). Some of the perturbation energy may be seen in wave-like structures outside of the shear layer for more strongly stratified background flows. The total perturbation energy gain decreases with increasing Ri_b , as seen by Farrell & Ioannou (1993a) and Kaminski *et al.* (2014), though $O(10^1 - 10^2)$ energy growth is observed even for the flows with no normal-mode instability. In addition, the optimal perturbation gain increases with increasing Re ; for $Ri_b = 0.40$, the optimal gain at

Re	Ri_b	L_x	E_0	(N_x, N_y, N_z)
1000	0.20	9.44	5E-6, 2E-5, 5E-5	(192, 192, 301), (256, 256, 401), (256, 256, 401)
1000	0.30	9.44	5E-6, 2E-5, 5E-5	(128, 128, 201), (128, 128, 201), (128, 128, 201)
1000	0.30	12.6	5E-6, $\boxed{2E-5}^a$, 5E-5	(128, 128, 201), (192, 192, 301), (192, 192, 301)
1000	0.35	9.44	5E-6, $\boxed{2E-5}$, 5E-5	(128, 128, 201), (192, 192, 301), (192, 192, 301)
1000	0.40	9.44	5E-6, $\boxed{2E-5}^b$, 5E-5	(128, 128, 201), (128, 128, 201), (192, 192, 301)
1000	0.50	9.44	5E-6, $\boxed{2E-5}$, 5E-5	(128, 128, 201), (128, 128, 201), (192, 192, 301)
2000	0.20	9.44	5E-6, 2E-5, 5E-5	(256, 256, 401), (384, 384, 601), (512, 512, 801)
2000	0.30	9.44	5E-6, 2E-5, 5E-5	(192, 192, 301), (192, 192, 301), (256, 256, 401)
2000	0.30	12.6	5E-6, $\boxed{2E-5}^c$, 5E-5	(256, 256, 401), (256, 256, 401), (384, 384, 601)
2000	0.35	9.44	5E-6, $\boxed{2E-5}$, 5E-5	(192, 192, 301), (256, 256, 401), (384, 384, 601)
2000	0.40	9.44	5E-6, $\boxed{2E-5}^d$, 5E-5	(192, 192, 301), (256, 256, 401), (256, 256, 401)
2000	0.50	9.44	5E-6, $\boxed{2E-5}$, 5E-5	(192, 192, 301), (256, 256, 401), (256, 256, 401)
4000	0.30	9.44	5E-6, 2E-5	(384, 384, 601), (384, 384, 601)
4000	0.30	12.6	5E-6, $\boxed{2E-5}^e$	(384, 384, 601), (384, 384, 601)
4000	0.40	9.44	5E-6, $\boxed{2E-5}^f$	(384, 384, 601), (384, 384, 601)
6000	0.40	9.44	2E-6, 5E-6, 2E-5	(384, 384, 601), (256, 384, 601), (384, 512, 801)
8000	0.40	9.44	5E-6, 2E-5	(384, 512, 801), (384, 600, 801)

TABLE 1. Parameters and grid sizes for direct numerical simulations. Boxed cases denote the specific cases considered in figures 4-8, and superscripts a-g correspond to the individual panels a-g in figures 5-8.

$T = 15$ more than doubles between $Re = 1000$ and $Re = 4000$. This is consistent with the optimal gains computed for unstratified shear flows, in which maximum gain increased markedly with Re (Arratia *et al.* 2013).

Figure 1 also shows that the maximum gain is sensitive to the domain size; the gains corresponding to $Ri_b = 0.30$ in the larger domain ($L_x = 12.6$) are higher than those for the standard domain ($L_x = 9.44$). This increase in gain is due to the quantization of the perturbation structure – only discrete wavelengths may fit into the domain, and so the optimal structure (across all streamwise wavenumbers) may be suppressed by the imposed periodicity (associated with the most unstable linear normal mode at $Ri_b = 0.2$). Thus, it is important to note that there may be structures in different-sized domains that would give somewhat larger (essentially linear) perturbation energy gain compared to the perturbations computed here, though once again we are principally interested in the ensuing inherently nonlinear dynamics. Appendix A shows the result of varying the horizontal and spanwise domain sizes in more detail. However, the fact remains that the transient growth mechanism leads to at least the $O(10^1 - 10^2)$ energy gain shown in figure 1, with the possibility of higher growth attainable for differently-sized domains.

The structure of the linear optimal perturbations for $Ri_b = 0.40$ and $Re = 1000$ and 4000 is shown in figure 2. The same colour scale is used for both Re values for comparison purposes. Both perturbations have the same initial energy, though the energy and corresponding velocity amplitudes themselves are arbitrary as these are linear solutions. While the perturbation in both cases takes the form of rolls tilted against the background shear described above, the lower- Re case shows evidence of wavelike structures outside of the shear layer as described in Kaminski *et al.* (2014). For the higher- Re case, the rolls have a smaller horizontal scale and the wavelike structures are absent. Increasing the

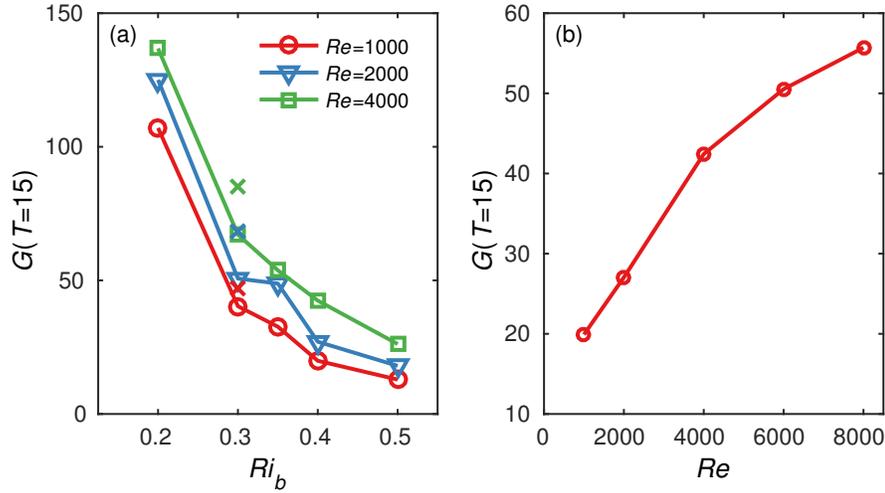


FIGURE 1. Optimal linear gain for a target time of $T = 15$ for (a) varying Ri_b and (b) varying Re with $Ri_b = 0.40$. The \times symbols in (a) denote the gains associated with the optimal perturbations with a larger value of L_x .

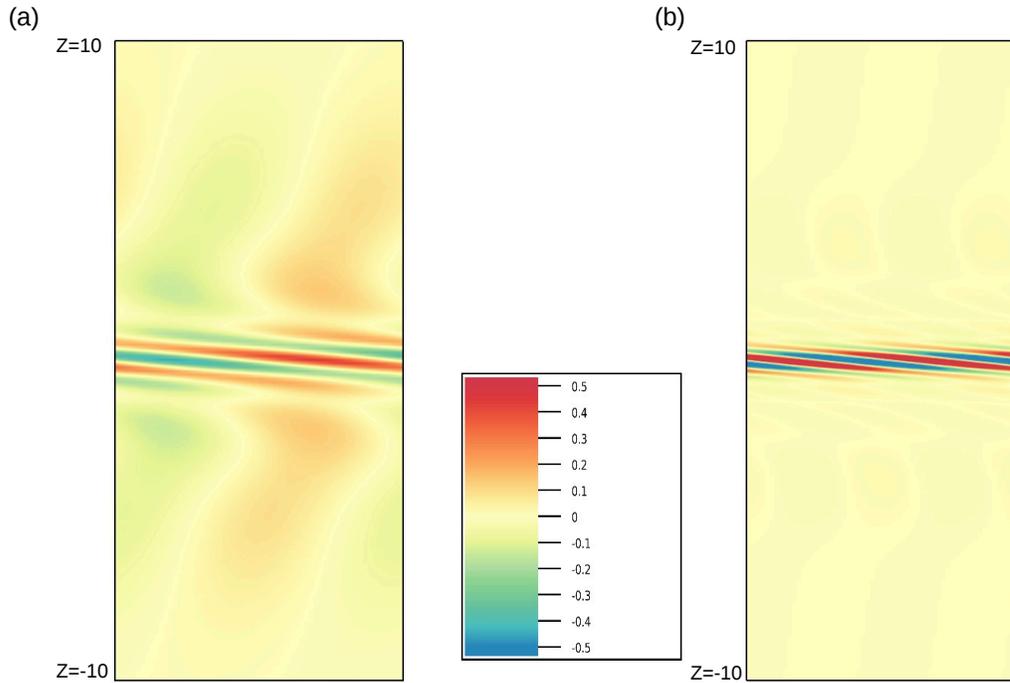


FIGURE 2. Vertical velocity of linear optimal perturbations for $Ri_b = 0.40$, $T = 15$, and (a) $Re = 1000$ and (b) $Re = 4000$. The same colour scheme is used for both and the values in (b) saturate the colour bar.

Reynolds number changes the structure of the linear optimal perturbation in addition to the maximum gain attained at the target time.

3.2. Nonlinear evolution for $Re = 4000$, $Ri_b = 0.40$, and $E_0 = 2 \times 10^{-5}$

We now shift our focus to the fully nonlinear results. For the nonlinear simulations, we consider the evolution of the mean quantities, given by the horizontally-averaged fields $\bar{\mathbf{u}}(z, t)$ and $\bar{b}(z, t)$, and the fluctuations away from the mean, $\mathbf{u}'(x, y, z, t) = (u', v', w')$ and $b'(x, y, z, t)$. The horizontal average, denoted by an overbar, is defined as

$$\bar{f} = \frac{1}{L_x} \frac{1}{L_y} \int_0^{L_y} \int_0^{L_x} f \, dx \, dy. \quad (3.1)$$

If the perturbations evolve entirely linearly, we expect the mean terms to correspond to the background expressions \mathbf{U} and B , defined in (2.2), and the evolution of \mathbf{u}' and b' to correspond to the linear optimal perturbations \mathbf{u} and b described above. Conversely, as nonlinear effects become important, $\bar{\mathbf{u}}$, \bar{b} , \mathbf{u}' , and b' are expected to deviate from their linear counterparts.

In order to gain a qualitative understanding of how nonlinearity may affect the evolution of the optimal perturbations, figure 3 shows an isosurface of total buoyancy, $\bar{b}(z, t) + b'(x, y, z, t)$, at the centre of the shear layer as well as the vertical component of vorticity, $\omega_z = \partial v' / \partial x - \partial u' / \partial y$, for $Re = 4000$, $Ri_b = 0.40$, and $E_0 = 2 \times 10^{-5}$ at four times. At early times (figure 3(a)), the additive noise obscures the small-amplitude two-dimensional rolls of the optimal perturbation. The rolls are tilted upwards by the mean shear, and if sufficient growth is attained they develop into ‘billows’ (figure 3(b)). The billow structure then becomes susceptible to secondary instabilities as highlighted by the oblique structures in vorticity in figure 3(c), and the flow breaks down into small scales which then dissipate and mix the mean flow (figure 3(d)). The billows and subsequent break down into turbulence are reminiscent of the nonlinear evolution of the classical Kelvin-Helmholtz instability: formation of two-dimensional billows (e.g. Thorpe 1973) followed by subsequent onset of secondary instabilities leading to (inherently transient) turbulence onset (e.g. Mashayek & Peltier 2012). However, it should be emphasized that the development of the billows observed in figure 3 arise not due to a classical normal-mode instability as the base flow is linearly stable, but rather through a transient Orr-type growth triggering the development of inherently nonlinear, yet still recognisably billow-like structures.

3.3. Evolution of mean quantities

We now consider the total perturbation energy evolution,

$$E(t) = \frac{\langle \mathbf{u}'(t), \mathbf{u}'(t) \rangle + Ri_b \langle b'(t), b'(t) \rangle}{2}, \quad (3.2)$$

with respect to time, where the angle brackets denote the same inner product as in (2.7). Figure 4 shows the fully nonlinear perturbation energy evolution as well as corresponding energy evolution for a perturbation with the same initial amplitude evolving in an entirely linear manner (as in the computation of the linear optimal perturbations, described above). The simulations shown in figure 4 have $Ri_b = 0.30$, $L_x = 12.6$, corresponding to perturbations which experience large gain on a base flow which is relatively close to the Miles-Howard criterion, and $Ri_b = 0.40$, $L_x = 9.44$, corresponding to perturbations growing on a more strongly-stratified base flow. The initial perturbation amplitude is $E_0 = 2 \times 10^{-5}$, and a range of Reynolds numbers, $Re = 1000, 2000$, and 4000 are shown. These parameters are chosen in order to highlight a range of different behaviours across the various simulations described in table 1.

The higher gains predicted by figure 1 can be seen in the energy evolution shown in figure 4. In general, the nonlinear flow evolution begins with a decrease in perturbation

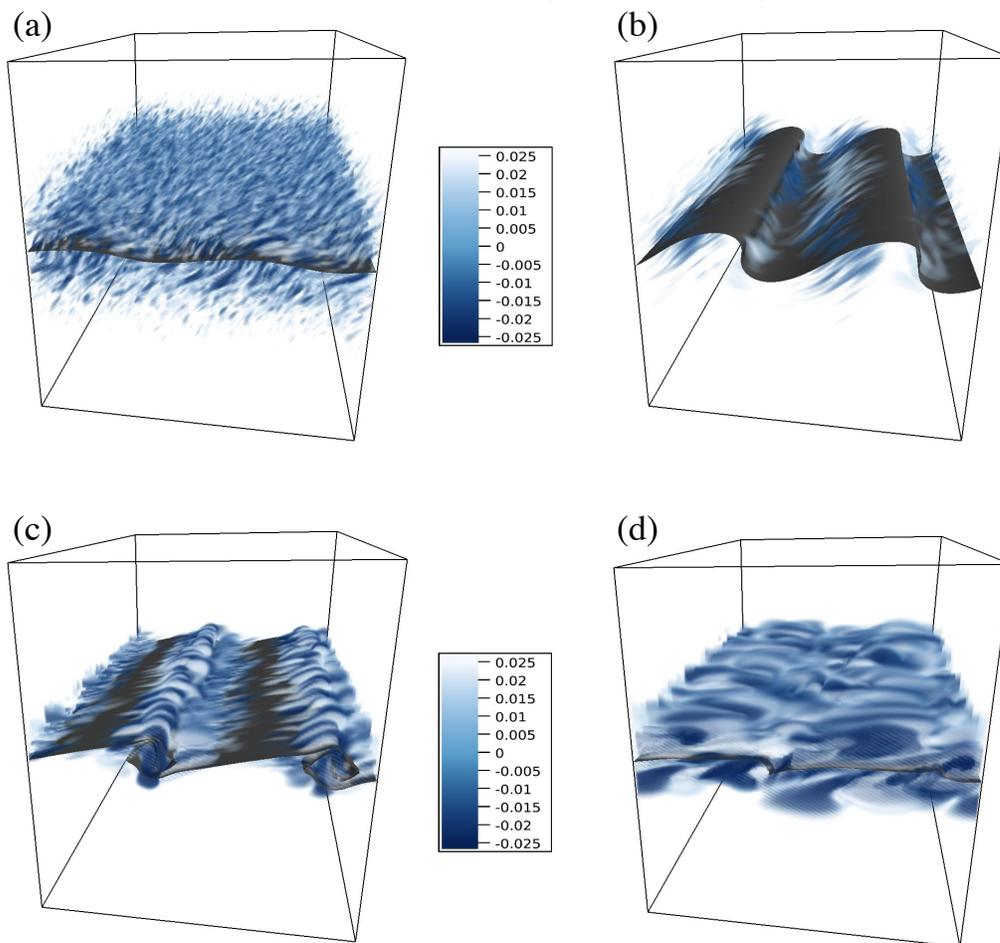


FIGURE 3. Evolution of total buoyancy $\bar{b}(z, t) + b'(x, y, z, t)$ through centre of shear layer (isosurface) and vorticity $\omega_z = \partial v' / \partial x - \partial u' / \partial y$ (colour) for $Re = 4000$, $Ri_b = 0.40$, and $E_0 = 2 \times 10^{-5}$. Snapshots of the flow are shown at (a) $t = 6.0$, (b) $t = 12.2$, (c) $t = 24.4$, and (d) $t = 36.7$. The region shown is $z = -2.5$ to $z = 2.5$, and the aspect ratio has been stretched by a factor of two in the vertical to better show the detail of the evolving flow.

amplitude as the additive noise decays. This is followed by a period of linear perturbation energy growth, in which the perturbations extract energy from the background shear flow via the stratified Orr mechanism, as described by Kaminski *et al.* (2014). Eventually, nonlinear effects become important and the nonlinear perturbation energy saturates, reaching a lower peak amplitude than the purely linear energy evolution. The energy decays at later times in both the linear and nonlinear cases.

In general, the cases with lower initial amplitude E_0 , lower Re , and higher Ri_b more closely follow the linear flow evolution of the computed optimal perturbations. Conversely, cases with higher E_0 , higher Re , and lower Ri_b generally show a more dramatic deviation from the linear case, with nonlinear saturation happening earlier in the flow evolution. The comparison between the linear and nonlinear energy evolution indicates when nonlinear effects start to become important, both in terms of initial amplitude and time. It should be emphasized that even in cases in which nonlinear saturation is significant (e.g. figure 4(c)), this saturation only occurs after an extended period of quasi-linear per-

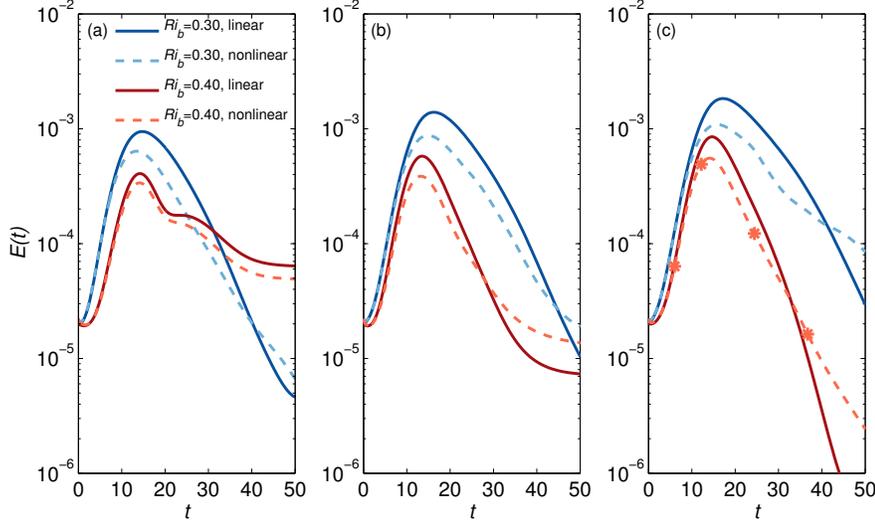


FIGURE 4. Time evolution of total perturbation energy $E(t)$ as defined in (3.2) for (a) $Re = 1000$, (b) $Re = 2000$, and (c) $Re = 4000$. The times corresponding to the flow snapshots shown in figure 3 are indicated by asterisks in (c).

turbation growth, indicating that it is the linear growth mechanism leading to the growth in perturbation energy, rather than some nonlinear response to a large perturbation.

We can examine further the evolution of the optimal perturbation and the subsequent secondary instabilities by considering different components of the perturbation energy. For instance, we can consider the relative contributions of two-dimensional (spanwise-averaged) motions, associated with the optimal perturbations, and of inherently three-dimensional motions, associated with the observed secondary instabilities and subsequent small-scale features which arise. By writing the fluctuation velocity and buoyancy as

$$\mathbf{u}' = \mathbf{u}'_{2D} + \mathbf{u}'_{3D} \quad \text{and} \quad b' = b'_{2D} + b'_{3D} \quad (3.3)$$

where $\mathbf{u}'_{2D} = 1/L_y \int_0^{L_y} \mathbf{u}' dy$ and $b'_{2D} = 1/L_y \int_0^{L_y} b' dy$, we can then define

$$E_{2D} = \frac{\langle \mathbf{u}'_{2D}, \mathbf{u}'_{2D} \rangle + Ri_b \langle b'_{2D}, b'_{2D} \rangle}{2} \quad \text{and} \quad E_{3D} = \frac{\langle \mathbf{u}'_{3D}, \mathbf{u}'_{3D} \rangle + Ri_b \langle b'_{3D}, b'_{3D} \rangle}{2} \quad (3.4)$$

as the energy contributions from two- and three-dimensional motions, respectively. Figure 5(a) and (b) show the time evolution of E_{2D} and E_{3D} for the cases with $Ri_b = 0.40$, $E_0 = 2 \times 10^{-5}$, and $Re = 2000$ to 8000 . In the case of $Re = 2000$, almost all of the perturbation energy is contained in the two-dimensional motions. However, for higher values of Re , there is a clear onset of three-dimensional motion beginning at times around $t \sim 20$, when the secondary instabilities themselves onset (as seen qualitatively in figure 3(b)).

The observed three-dimensionalization of the flow is further supported by decomposing the fluctuation energy budget into terms corresponding to the streamwise, spanwise, and vertical velocities, as well as the buoyancy, i.e.

$$E = \frac{\langle \mathbf{u}', \mathbf{u}' \rangle + Ri_b \langle b', b' \rangle}{2} = \frac{\langle u', u' \rangle}{2} + \frac{\langle v', v' \rangle}{2} + \frac{\langle w', w' \rangle}{2} + \frac{Ri_b \langle b', b' \rangle}{2}. \quad (3.5)$$

Panels (c)-(f) of figure 5 show the contribution of each of these components to the overall energy evolution. It is clear that as E_{3D} increases, so does the component owing

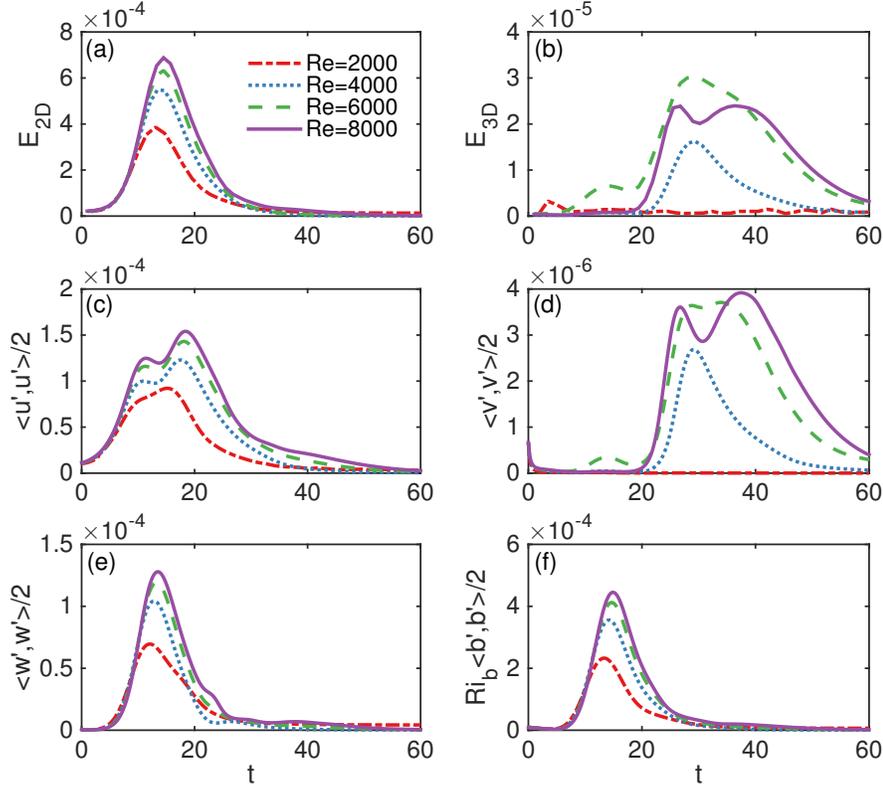


FIGURE 5. Perturbation energy by component for $Ri_b = 0.40$, $E_0 = 2 \times 10^{-5}$, and varying Re . (a) Two-dimensional component of energy, $E_{2D}(t)$. (b) Three-dimensional component of energy, $E_{3D}(t)$. (c) Streamwise velocity component, $\langle u', u' \rangle / 2$. (d) Spanwise velocity component, $\langle v', v' \rangle / 2$. (e) Vertical velocity component, $\langle w', w' \rangle / 2$. (f) Potential energy component, $Ri_b \langle b', b' \rangle / 2$.

to the spanwise fluctuation velocity for $Re \geq 4000$ – the flow is becoming increasingly three-dimensional both in spatial structure and in velocity component. This three-dimensionalization of the flow has been identified as an important step in the transition process for stratified shear flows (Caulfield & Peltier 2000).

Figure 6 shows the evolution of the mean squared buoyancy frequency, defined as

$$\frac{N^{*2}}{N_0^{*2}} = \frac{\partial \bar{b}}{\partial z}, \quad (3.6)$$

in the vicinity of the shear layer at a series of times over the course of the flow evolution. The cases shown correspond to the same nonlinear simulations as those shown in figure 4. For $Re = 1000$ and $Ri_b = 0.40$ (figure 6(b)), there is very little modification of the background stratification, consistent with the energy evolution shown in figure 4(a): the linear and nonlinear simulations evolve very similarly, indicating that the perturbation evolution in this case is primarily linear. Thus, little modification of the mean flow would be expected in this case, as observed.

For all other cases shown in figure 6, the evolution of the mean stratification shows the signature of localized mixing. A layer of decreased stratification is surrounded by layers of

enhanced stratification. The layered structure is similar to that resulting from classical Kelvin-Helmholtz instability (Klaassen & Peltier 1985; Smyth *et al.* 2001; Salehipour *et al.* 2016). The layer structure is strongest at $t = 15 - 25$ (at which time the flow has noticeably deviated from the linear evolution in figure 4), after which further flow evolution and diffusion act to remove the layer structure and return the mean flow to its initial uniform stratification. For $Re = 1000$ and $Ri_b = 0.30$ (figure 6(a)) and for $Re = 2000$ and $Ri_b = 0.40$ (figure 6(d)), the mean stratification evolves smoothly and is relatively symmetric in the vertical at later times. For the higher- Re flows shown in figure 6(c,e,f), not only is the layering much more pronounced, but the structure of the mean stratification becomes much more complicated at later times as a result of the more vigorous secondary instabilities seen in these flows.

Next, we examine the mean gradient Richardson number of the flow, defined as

$$Ri_g = \frac{\overline{\partial b / \partial z}}{(\overline{\partial \bar{u} / \partial z})^2}, \quad (3.7)$$

where the overbar denotes the horizontal average used in (3.1). While the local gradient Richardson number, N^2/S^2 , varies spatially, this mean quantity gives an overall picture of the stability of the shear layer as the initial perturbation evolves. Figure 7 shows the evolution of the mean gradient Richardson number around the centre of the shear layer for the same simulations as considered in figures 4 and 6.

In the $Re = 1000$, $Ri_b = 0.40$ case, where very little modification of the background stratification was observed, the mean gradient Richardson number gradually increases as the mean flow diffuses outwards. In all other cases the evolution of the mean gradient Richardson number is more complicated. This is particularly true for the higher- Re cases in which more vigorous secondary instabilities lead to Ri_g profiles of increasing complexity at later times. Of note are the local minima observed at the centres of the shear layers. These minima roughly coincide with the time of strongest layering in figure 6, $t \sim 15 - 20$, and in most cases fall well below the classical Miles-Howard “critical” value of $1/4$ – even in the case $Ri_b = 0.40$ and $Re = 4000$ (figure 7(f)). That is, the nonlinear evolution of the non-normal modes is able to drive the flow towards a mean state which may be unstable in a normal-mode sense, despite starting from an initial background state which is linearly stable.

3.4. Turbulent energetics and mixing

In the previous subsection, we observed that linear transient growth may lead to a large increase in the perturbation energy. Subsequently, departure from the linear evolution may be observed, and the resulting nonlinear effects modify the mean density and Ri_g profiles. In this section, we will examine the details of the turbulent energetics and mixing in greater detail.

We construct energy budgets for the perturbation kinetic energy $k = u'_j u'_j / 2$ and potential energy $\phi = Ri_b b'^2 / 2$ from the equations governing conservation of momentum and buoyancy. Applying the same average as in (3.1), the horizontally-averaged perturbation kinetic and potential energy budgets are given by

$$\frac{\partial \bar{k}}{\partial t} = - \underbrace{\frac{\overline{\partial p' w'}}{\partial z}}_{\text{pressure transp.}} - \underbrace{\frac{\overline{\partial k w'}}{\partial z}}_{\text{turbulent transp.}} + \underbrace{Ri_b \overline{b' w'}}_{\text{buoyancy flux}}$$

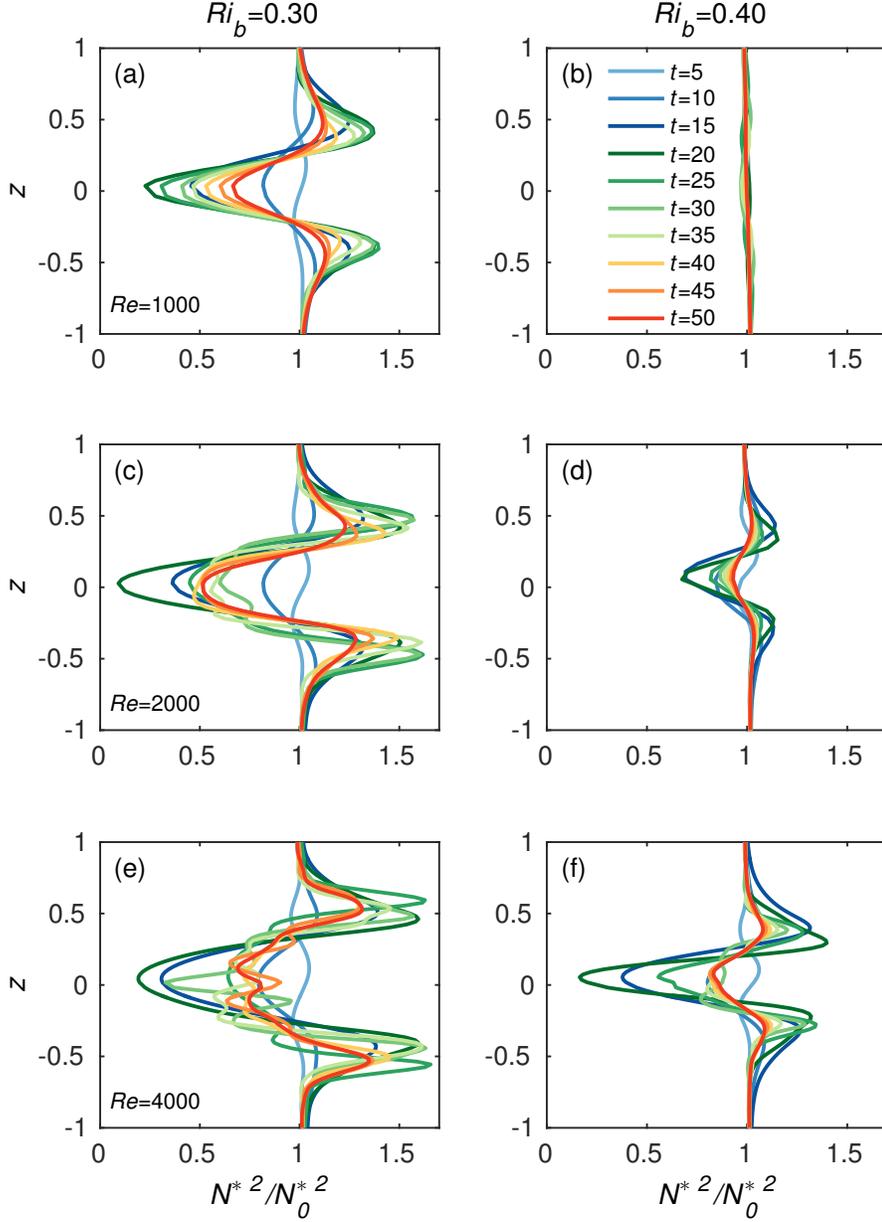


FIGURE 6. Time evolution of scaled mean buoyancy frequency, $N^{*2}/N_0^{*2} = \partial\bar{b}/\partial z$, in the vertical region centred on the shear layer. The left column corresponds to $Ri_b = 0.30$, and the right column corresponds to $Ri_b = 0.40$. (a,b) $Re = 1000$. (c,d) $Re = 2000$. (e,f) $Re = 4000$. Note that $E_0 = 2 \times 10^{-5}$ for all cases shown here.

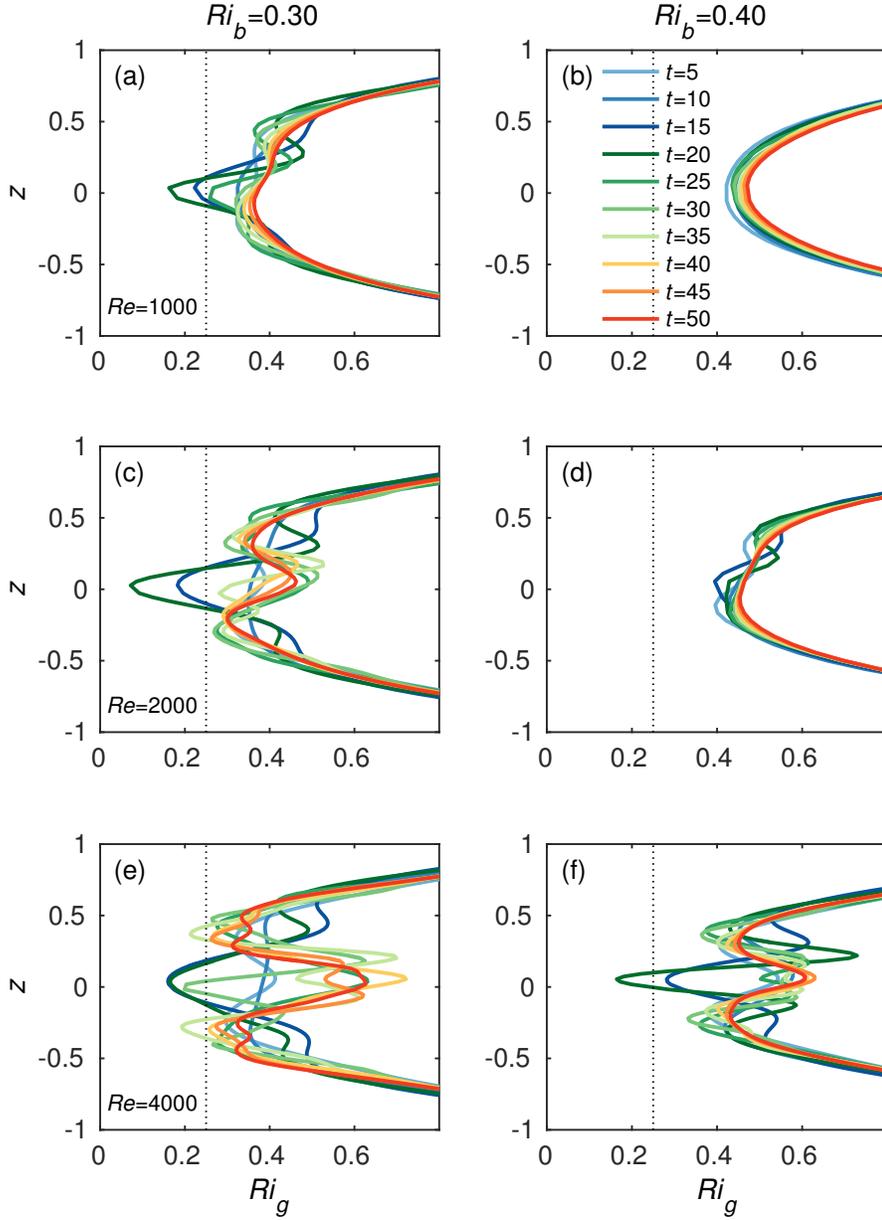


FIGURE 7. Evolution of mean gradient Richardson number (defined by (3.7)), in region around shear layer. The left column corresponds to $Ri_b = 0.30$, and the right column corresponds to $Ri_b = 0.40$. The dotted line denotes $Ri_g = 0.25$. (a,b) $Re = 1000$. (c,d) $Re = 2000$. (e,f) $Re = 4000$. Note that $E_0 = 2 \times 10^{-5}$ for all cases shown here.

$$- \underbrace{\overline{u'w'} \frac{\partial \bar{u}}{\partial z}}_{\text{shear prod.}} + \underbrace{\frac{1}{Re} \frac{\partial^2 \bar{k}}{\partial z^2}}_{\text{viscous transp.}} - \underbrace{\varepsilon_k}_{\text{KE diss.}}, \quad (3.8)$$

and

$$\frac{\partial \bar{\phi}}{\partial t} = - \underbrace{\frac{\partial \overline{\phi w'}}{\partial z}}_{\text{turbulent transp.}} - \underbrace{Ri_b \overline{b'w'} \frac{\partial \bar{b}}{\partial z}}_{\text{buoyancy flux}} + \underbrace{\frac{1}{RePr} \frac{\partial^2 \bar{\phi}}{\partial z^2}}_{\text{diffusive transp.}} - \underbrace{\varepsilon_p}_{\text{PE diss.}}, \quad (3.9)$$

respectively. The strictly positive terms ε_k and ε_p represent sinks of perturbation kinetic energy via viscous dissipation and perturbation potential energy via diffusion, respectively and are defined as

$$\varepsilon_k = \frac{1}{Re} \frac{\overline{\partial u'_i}}{\partial x_j} \frac{\overline{\partial u'_i}}{\partial x_j} \quad \text{and} \quad \varepsilon_p = \frac{Ri_b}{RePr} \frac{\overline{\partial b'}}{\partial x_j} \frac{\overline{\partial b'}}{\partial x_j}. \quad (3.10)$$

The shear production term, $-\overline{u'w'} \partial \bar{u} / \partial z$, describes the interaction between the perturbation kinetic energy and the mean shear, while the vertical buoyancy flux terms $Ri_b \overline{b'w'}$ and $Ri_b \overline{b'w'} \partial \bar{b} / \partial z$ describe the conversion between kinetic and potential energy as well as interaction with the mean stratification. The remaining terms describe turbulent transport of kinetic and potential energy within the domain ($\partial \overline{k w'} / \partial z$ and $\partial \overline{\phi w'} / \partial z$), viscous and diffusive transport ($1/Re \partial^2 \bar{k} / \partial z^2$ and $Ri_b / RePr \partial^2 \bar{\phi} / \partial z^2$), and pressure transport ($\partial \overline{p'w'} / \partial z$).

The evolution of the horizontally-averaged turbulent dissipation, ε_k , is shown as a function of time t and the vertical coordinate z in figure 8. The simulations shown correspond to those from figure 4. After a short initial period in which the added noise perturbation decays ($t \sim 0 - 5$), the dissipation is fairly low and slightly offset from the centre of the shear layer, corresponding to the edges of the rolls as they are tilted upward by the mean shear. As the rolls lead to the billow structure and smaller-scale secondary instabilities onset, a period of significantly enhanced dissipation is observed in most cases (corresponding to those which exhibited strong nonlinear effects in figures 6 and 7). This enhanced dissipation is confined to a narrow region around the centre of the shear layer. The observed dissipation decreases with increasing Ri_b , and increases with increasing Re , as might be expected based on the increasingly complex structure observed for e.g. figures 6(e,f), indicative of vigorous small-scale structures that would lead to higher dissipation. Also in accordance with figures 6 and 7, in the case of $Ri_b = 0.40$ and $Re = 1000$ (figure 8(b)), ε_k remains low as the flow remains quasi-linear throughout the evolution of the perturbations.

In order to assess the relative importance of the various terms in the perturbation energy budgets throughout the domain, we can integrate equations (3.8) and (3.9) in a control volume defined by the vertical extents z_1 and z_2 . The kinetic and potential energy budgets then become

$$\frac{\partial}{\partial t} \int_{z_1}^{z_2} \bar{k} dz = \left[-\overline{k w'} - \overline{p' w'} - \frac{1}{Re} \frac{\partial \bar{k}}{\partial z} \right]_{z_1}^{z_2} + \int_{z_1}^{z_2} \left(Ri_b \overline{b'w'} - \overline{u'w'} \frac{\partial \bar{u}}{\partial z} - \varepsilon_k \right) dz \quad (3.11)$$

and

$$\frac{\partial}{\partial t} \int_{z_1}^{z_2} \bar{\phi} dz = \left[-\overline{\phi w'} - \frac{1}{RePr} \frac{\partial \bar{\phi}}{\partial z} \right]_{z_1}^{z_2} + \int_{z_1}^{z_2} \left(-Ri_b \overline{b'w'} \frac{\partial \bar{b}}{\partial z} - \varepsilon_p \right) dz \quad (3.12)$$

respectively. For a control volume with vertical extents outside of the shear layer, we

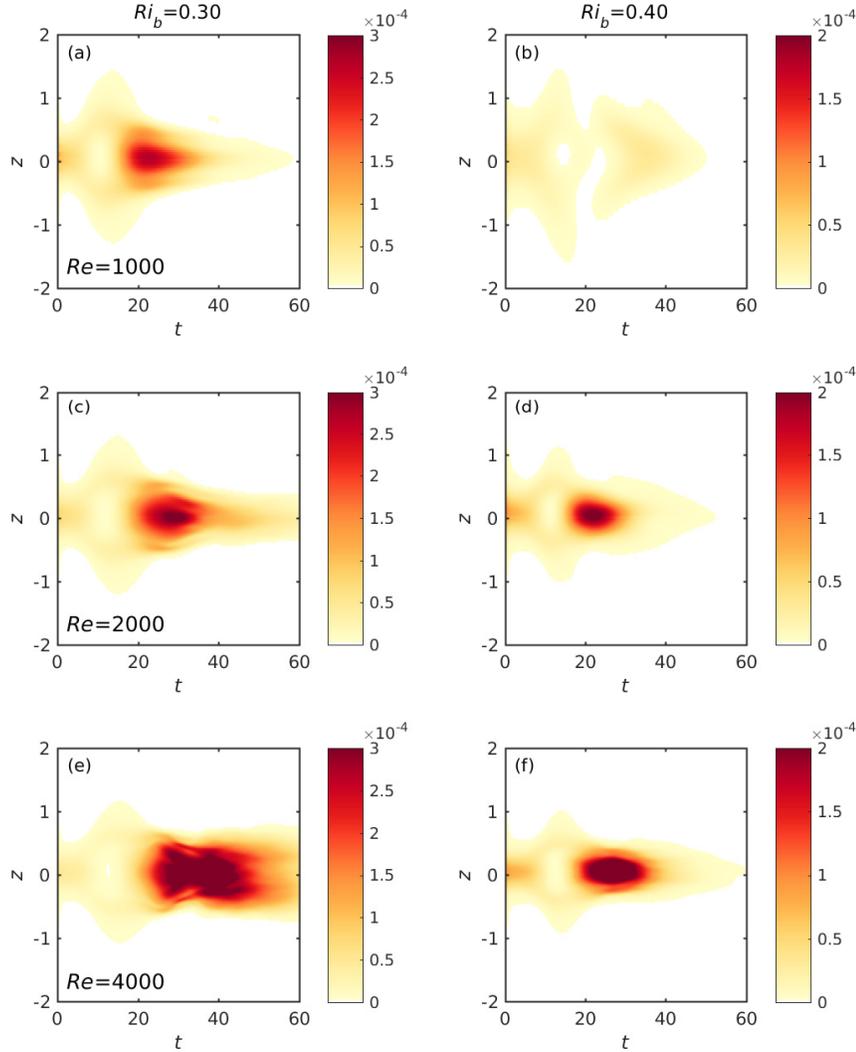


FIGURE 8. Planewise-averaged turbulent dissipation, $\varepsilon_k(z, t)$, for the same simulations as in figures 6 and 7.

might expect that the the shear production, buoyancy flux terms, and sinks owing to dissipation are important, as in the purely linear case (Kaminski *et al.* 2014), and that the energy flux associated with the viscous and advective transport terms would be small. The pressure transport term may be important if waves are generated during the flow evolution.

In order to assess the relative importance of the two sinks ε_k and ε_p , we consider a control volume with boundaries at $z_1 = -6$ and $z_2 = 6$. As figure 8 shows, these vertical extents are well separated from the region of high dissipation occurring near the centre of the shear layer. The results presented here are relatively insensitive to the vertical extent

of the control volume provided the boundaries are away from the shear layer and evolving perturbation. Figure 9 shows the evolution of integrated dissipation in the control volume as a function of time for the $E_0 = 2 \times 10^{-5}$, as in the previous figures. In general, as Re increases, the total dissipation of both kinetic and potential energy in the control volume reaches a higher maximum value. For $Ri_b = 0.30$ and $Re = 1000$ and 2000 (figure 9(a,c)), the peak values in dissipation of kinetic and potential energy occur at approximately the same time, and have roughly the same magnitude. However, at the higher Reynolds number $Re = 4000$, shown in figure 9(e), the kinetic energy dissipation reaches a larger maximum value and remains elevated for a much longer period of time than the potential energy dissipation, suggesting that ‘mixing’ of potential energy and dissipation of kinetic energy are not necessarily well time-correlated (see e.g. Mashayek *et al.* 2013). A similar trend can be seen when comparing the dissipation rates for $Ri_b = 0.40$ (figure 9(b,d,e)): as Re increases, the kinetic energy dissipation increases relative to the potential energy dissipation. In the specific case of $Ri_b = 0.40$ and $Re = 1000$, for which the flow evolution is quasi-linear (as shown in figures 6 and 7), the amount of dissipation in the control volume is quite low in comparison to the other cases shown in figure 9. In all cases, the peak values of ε_k and ε_p occur after nonlinear effects have become important in the flow, as discussed in the previous section.

In addition to dissipation of kinetic and potential energy via viscosity and diffusion, energy may also leave the control volume through the effect of internal waves. Given the internal waves observed for the linear scenario in Kaminski *et al.* (2014), it is natural to question how significant the transport of energy by waves is when compared with the dissipative terms. We quantify the flux of energy through the edges of the control volume by the pressure transport $\overline{p'w'}|_{z_1}^{z_2}$. Away from the shear layer, where the flow is stratified and quasi-linear, we expect the effects of internal waves to dominate the pressure transport. We plot $\overline{p'w'}|_{z_1}^{z_2}$ as a function of time in figure 9 in order to quantify the importance of the flux via waves relative to the dissipative sink terms.

For $Ri_b = 0.30$ (figure 9(a,c,e)), $\overline{p'w'}|_{z_1}^{z_2}$ is significantly lower than that of the sinks $\int_{z_1}^{z_2} \varepsilon_k dz$ and $\int_{z_1}^{z_2} \varepsilon_p dz$, and arises from the localized evolution of the perturbation in the shear layer rather than from internal waves propagating away from the centre of the domain. However, when $Ri_b = 0.40$, the role of internal waves is potentially very different. While the dissipative terms are still the dominant terms when $Ri_b = 0.40$ and $Re = 4000$ (figure 9(f)), similarly to the cases with $Ri_b = 0.30$, for $Re = 2000$ (figure 9(d)) the magnitude of the wave flux term is comparable to and in fact slightly larger than the peak values in the dissipative terms, and for $Re = 1000$ (the nearly linear case discussed above, (figure 9(b)) the flux of energy via internal waves is much larger than the dissipation of both kinetic and potential energy. In general, there is a shift from an internal wave-dominated regime to one in which turbulent dissipation is dominant as Re increases and Ri_b decreases.

As seen in figure 5, for higher- Re flows there is a clear three-dimensionalization of the flow as secondary instabilities set in on the primarily two-dimensional optimal perturbation. In order to assess the energetics corresponding to these three-dimensional structures, we construct spectral energy budgets from the nonlinear perturbation governing equations (as with the spatial energy budgets (3.8) and (3.9), above). We first Fourier transform the governing equations in the x and y directions, denoting transformed variables with a caret. Then, by multiplying the governing equations by the complex conjugates of velocity and buoyancy (denoted by an asterisk), we obtain the spectral form of the

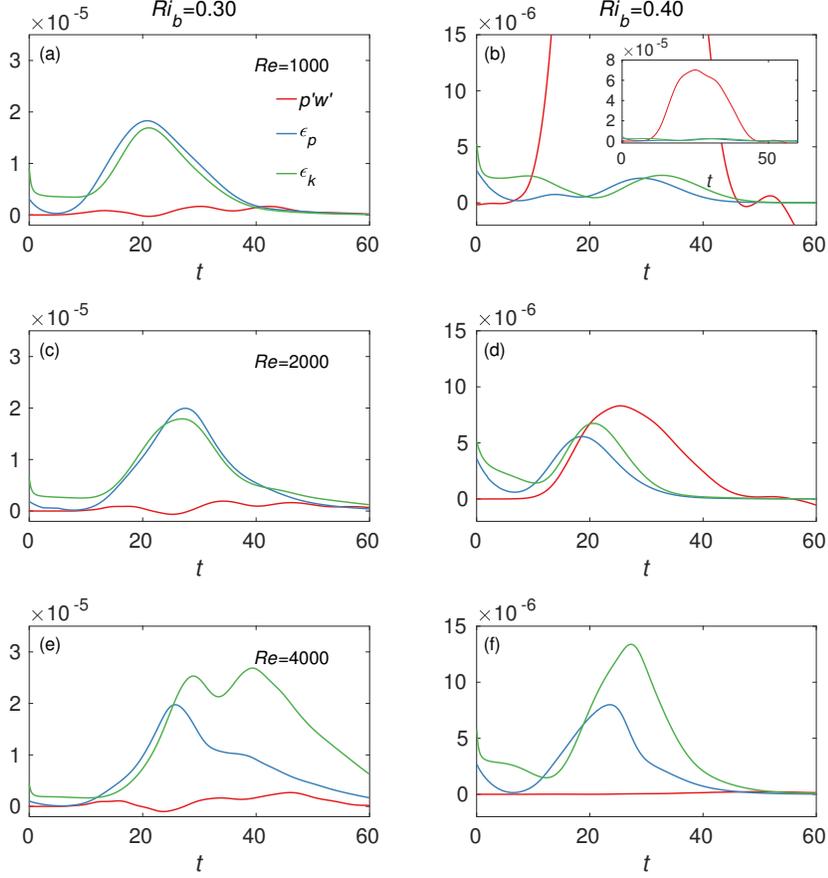


FIGURE 9. Evolution of $\int_{z_1}^{z_2} \epsilon_k dz$, $\int_{z_1}^{z_2} \epsilon_p dz$, and $\overline{p'w'}|_{z_1}^{z_2}$, where z_1 and z_2 correspond to the control volume shown in figure 8, as a function of time. The simulations are the same as those in figures 6 and 7.

kinetic energy budget,

$$\frac{\partial}{\partial t} \Re \left[\frac{\hat{\mathbf{u}}'^* \cdot \hat{\mathbf{u}}'}{2} \right] = \Re \left[-\hat{\mathbf{u}}'^* \cdot (\hat{\mathbf{u}}' \cdot \widehat{\nabla} \mathbf{u}') - \hat{u}'^* \hat{w}' \frac{\partial \bar{u}}{\partial z} - \hat{\mathbf{u}}'^* \cdot \widehat{\nabla} p + Ri_b \hat{w}'^* \hat{b}' + \frac{\hat{\mathbf{u}}'^* \cdot \widehat{\nabla}^2 \mathbf{u}'}{Re} \right], \quad (3.13)$$

and the spectral form of the potential energy budget,

$$\frac{\partial}{\partial t} \Re \left[\frac{Ri_b \hat{b}'^* \cdot \hat{b}'}{2} \right] = \Re \left[-Ri_b \hat{b}'^* (\hat{\mathbf{u}}' \cdot \widehat{\nabla} \mathbf{b}') - Ri_b \hat{b}'^* \hat{w}' \frac{\partial \bar{b}}{\partial z} + \frac{Ri_b \hat{b}'^* \cdot \widehat{\nabla}^2 \mathbf{b}'}{Re Pr} \right]. \quad (3.14)$$

Equations (3.13) and (3.14) allow us to examine the evolution of specific lengthscales in the flow. We define

$$\mathcal{P}_{low} = \iint_{|\mathbf{k}| < k_c} \left(-\hat{u}'^* \hat{w}' \frac{\partial \bar{u}}{\partial z} \right) dk_x dk_y, \quad \mathcal{P}_{high} = \iint_{|\mathbf{k}| > k_c} \left(-\hat{u}'^* \hat{w}' \frac{\partial \bar{u}}{\partial z} \right) dk_x dk_y \quad (3.15)$$

as the low- and high-wavenumber contributions to the shear production,

$$\mathcal{B}_{low} = \iint_{|\mathbf{k}| < k_c} \left(Ri_b \hat{w}'^* \hat{b}' \right) dk_x dk_y, \quad \mathcal{B}_{high} = \iint_{|\mathbf{k}| > k_c} \left(Ri_b \hat{w}'^* \hat{b}' \right) dk_x dk_y \quad (3.16)$$

as the low- and high-wavenumber contributions to the buoyancy flux, and

$$\mathcal{T}_{low} = \iint_{|\mathbf{k}| < k_c} \left(-\hat{\mathbf{u}}'^* \cdot (\widehat{\mathbf{u}' \cdot \nabla \mathbf{u}'}) \right) dk_x dk_y, \quad \mathcal{T}_{high} = \iint_{|\mathbf{k}| > k_c} \left(-\hat{\mathbf{u}}'^* \cdot (\widehat{\mathbf{u}' \cdot \nabla \mathbf{u}'}) \right) dk_x dk_y \quad (3.17)$$

as the nonlinear transfer terms for low- and high-wavenumber motions, respectively. In the above definitions, $\mathbf{k} = (k_x, k_y)$ and k_c is a cutoff wavenumber. We plot these terms in figure 10 for $Ri_b = 0.40$, $E_0 = 2 \times 10^{-5}$, and $Re = 4000, 6000$, and 8000 at the approximate times when growth rates of E_{2D} and E_{3D} are largest (i.e. $t \sim 7$ for the top row and $t \sim 25$ for the bottom row). The cutoff wavenumber is chosen to be $k_c = 2$.

We observe that, consistent with the results of Kaminski *et al.* (2014), the initial two-dimensional energy growth is largely due to shear production, with buoyancy flux acting to convert kinetic energy into potential energy. At later times when the three-dimensional motions are growing, shear production is acting to decrease energy, while buoyancy flux is leading to conversion of potential energy into kinetic energy, consistent with the development of convectively unstable regions of the flow (see figure 3). In addition, there is a net transfer of energy from larger to smaller scales via the nonlinear transfer term. The down-scale energy transfer is seen at earlier times for the case with $Re = 8000$. Altogether, the spectral energetics suggest that the secondary instabilities and subsequent three-dimensionalization arise as a result of the nonlinear development of the linear optimal perturbation itself, rather than due to linear instability arising on the nonlinearly-modified base flow shown in figures 6 and 7. In fact, at the time when the mean Ri_g first drops below $1/4$, the shear production term is negative for both the low and high wavenumber components of the flow.

Mixing efficiency is an important metric characterizing the proportion of energy expended by a fluid in mixing the density field to the total energy lost to mixing and dissipation (Peltier & Caulfield 2003). One way to consider this is by comparing the dissipation of potential energy to the total dissipation (see e.g. Brucker & Sarkar 2007). Defining the time- and domain-integrated (perturbation) dissipations $\mathcal{E}_k = \int_0^{t_f} \int_0^{L_z} \varepsilon_k dz dt$ and $\mathcal{E}_p = \int_0^{t_f} \int_0^{L_z} \varepsilon_p dz dt$, where t_f is the end time of the simulation, we can quantify the relative importance of potential energy dissipation to total dissipation via a ‘‘mixing efficiency’’, which we define as

$$\eta_\varepsilon = \frac{\mathcal{E}_p}{\mathcal{E}_k + \mathcal{E}_p}. \quad (3.18)$$

In the above, t_f denotes the end time of the simulation, where the end of the simulation occurs when the fluctuation energy as well as the kinetic and potential energy dissipation terms are small, and the transient turbulence has relaminarized, as is characteristic of such free shear flows. In all cases, $d\mathcal{E}_k/dt < 0.005\mathcal{E}_k$ and $d\mathcal{E}_p/dt < 0.005\mathcal{E}_p$ by time t_f .

Figure 11 shows this mixing efficiency calculated for a range of Re , Ri_b , and initial amplitude E_0 . It also shows the mixing efficiency as calculated for a purely linear flow evolution; it should be noted that as the linear quantities are ratios of very small numbers associated with quantities, the ratio itself may be relatively high despite the magnitudes of the perturbation dissipation rates being quite low. The linear mixing efficiencies in figure 11 show that, for $Ri_b \leq 0.35$, $\eta_\varepsilon \sim 0.5 - 0.6$, and $\eta_\varepsilon \sim 0.4$ for $Ri_b > 0.35$; these values are not strongly dependent on Re . In general, across the range of Ri_b considered

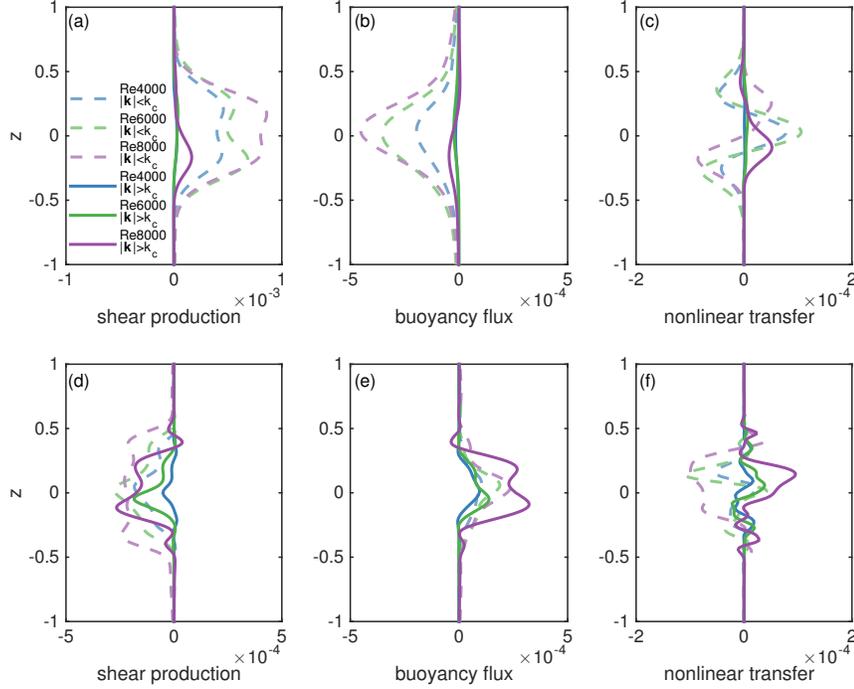


FIGURE 10. High- and low-wavenumber contributions to perturbation energy growth when growth of E_{2D} (top row) and E_{3D} (bottom row) are maximum, for $Ri_b = 0.40$, $E_0 = 2 \times 10^{-5}$, and $Re = 4000, 6000$, and 8000 . (a,d) Shear production terms, \mathcal{P}_{low} and \mathcal{P}_{high} . (b,e) Buoyancy flux terms, \mathcal{B}_{low} and \mathcal{B}_{high} . (c,f) Nonlinear transfer terms, \mathcal{T}_{low} and \mathcal{T}_{high} . For the cases plotted here, the cutoff wavenumber is chosen to be $k_c = 2$.

here, it is expected that the kinetic and potential energy dissipations will be of similar magnitudes in the case of purely linear flow evolution.

For the direct numerical simulations, the mixing efficiency deviates from the linear values as nonlinear effects become increasingly important. For the lower- Re cases (figure 11(a,b)), the curves collapse to the linear value for $Ri_b \geq 0.40$. For $Ri_b < 0.40$, the mixing efficiency approaches the linear value as Ri_b increases and E_0 decreases; conversely, for high E_0 and low Ri_b , nonlinear effects become much more important and mixing efficiency becomes substantially lower than the linear value. The exception to this trend is for $Ri_b = 0.30$ in the smaller domain ($L_x = 9.44$) at $Re = 1000$ and 2000 , for which the mixing efficiency increases slightly with E_0 due to an increase in \mathcal{E}_p relative to \mathcal{E}_k . We find that \mathcal{E}_p reaches a maximum ($d\mathcal{E}_p/dt \approx 0$) around the time at which the turbulent kinetic energy reaches its maximum value. While there is no evidence of transition to turbulence for these Reynolds numbers, as E_0 is increased the perturbation evolution does lead to some local modification of the background flow, strengthening buoyancy gradients and consequently leading to the observed increase in \mathcal{E}_p . For $Ri_b = 0.20$, where the initial condition gives rise to the Kelvin-Helmholtz instability and becomes nonlinear for all initial perturbation amplitudes considered here, the resulting mixing efficiency is also significantly lower ($\eta_\varepsilon \sim 0.30$) than the value predicted by the purely linear flow evolution. Similar results are seen at $Re = 4000$, with a decrease in mixing efficiency seen even for $Ri_b = 0.40$ and $E_0 = 2 \times 10^{-5}$, where nonlinear effects are significant, as

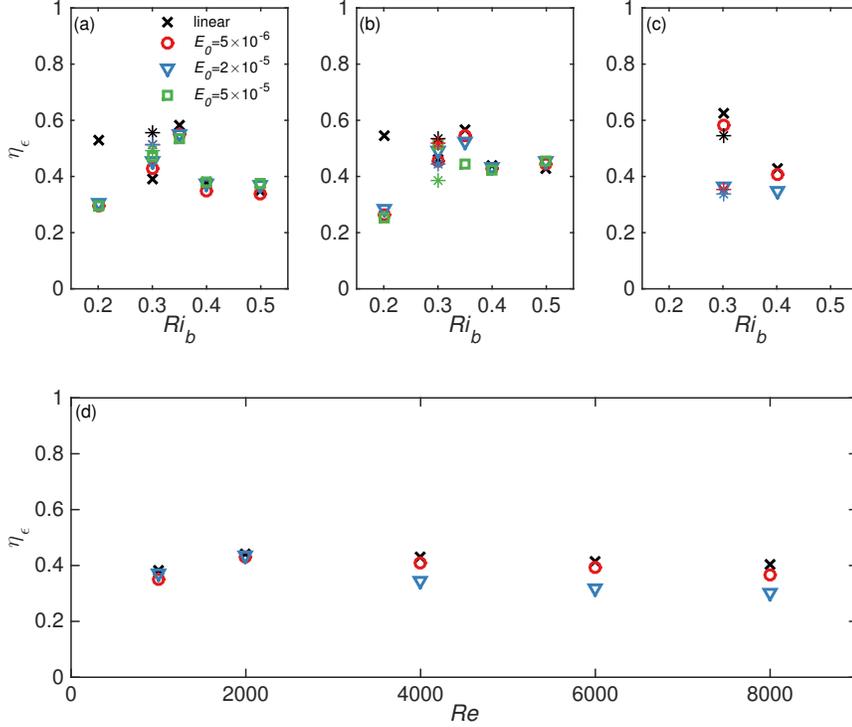


FIGURE 11. Mixing efficiency η_ϵ , as defined in (3.18), as a function of Ri_b for different initial amplitudes E_0 and (a) $Re = 1000$, (b) $Re = 2000$, and (c) $Re = 4000$. Open symbols correspond to simulations with $L_x = 9.44$; asterisks denote mixing efficiencies from simulations with $L_x = 12.6$ and $Ri_b = 0.30$. (d) Mixing efficiency for $Ri_b = 0.40$ as a function of Re .

shown in figures 3, 6 and 7 – the mixing efficiency in this case drops to approximately $\eta_\epsilon \sim 0.35$, as does the mixing efficiency calculated for $Ri_b = 0.30$ and $Re = 4000$ for both domain sizes L_x . At higher values of Re , the mixing efficiency tends towards a value of approximately 0.30, and the nonlinear values begin to deviate more from the linear values at lower initial amplitudes (figure 11(d)). While lower than the mixing efficiency in the linear cases, these values of the mixing efficiency are still quite large when compared to the commonly used value of $1/6$, but consistent with Mashayek *et al.* (2013) and the more recent results of Maffioli *et al.* (2016) and Venayagamoorthy & Koseff (2016).

For all of the flows considered, the region of elevated kinetic energy dissipation is localized in the vertical around the centre of the shear layer, as seen in figure 8. The decrease in mixing efficiency seen in figure 11 is associated with flows that experience significant nonlinearity and therefore an elevated kinetic energy dissipation relative to potential energy dissipation as seen in figure 9. In these flows, figure 6 shows that the centre of the shear layer becomes less stratified than the surrounding fluid as the perturbations evolve. While more kinetic energy is being dissipated by the strongly nonlinear flows, comparing figures 6 and 8 shows that the bulk of the dissipation is occurring in a weakly stratified region of fluid, and so the resulting mixing is less efficient than if it were occurring in a more strongly-stratified flow.

4. Discussion and conclusions

We have computed the linear optimal perturbations – that is, the perturbations that experience the largest linear energy growth – of a stably-stratified shear flow for a target time $T = 15$ advective time units. The linear optimal perturbations have been computed for various bulk Richardson numbers ranging from $Ri_b = 0.20$ to $Ri_b = 0.50$ and for Reynolds numbers from $Re = 1000$ up to $Re = 8000$ in some cases. We then used these optimal perturbations as an initial condition for direct numerical simulation with varying initial amplitude in order to examine what, if any, nonlinear effects may occur during the evolution of the flow.

We find that substantial perturbation energy gains are possible, and the maximum linear gain at the target time increases with increasing Re and decreasing Ri_b . Additionally, the structure of the optimal perturbations is sensitive to Re , Ri_b , and domain size. However, in all cases considered here $O(10^1 - 10^2)$ energy gain is achieved in the purely linear case.

The effects of varying Re , Ri_b , and initial amplitude E_0 are also apparent when the perturbations are allowed to evolve nonlinearly. In general, cases with higher E_0 and Re and lower Ri_b are prone to stronger nonlinear effects, including modification of the mean flow, formation of billows, vigorous secondary instabilities, three-dimensionalization of the flow and break down into smaller-scale structures, leading to enhanced yet transient turbulent dissipation. In some cases, the perturbation evolution may undergo a regime shift with increasing nonlinearity, from a case in which the perturbations generate strong internal waves propagating away from the shear layer to one in which dissipation dominates the energetics. In addition, flows which are susceptible to strongly nonlinear effects exhibit less efficient mixing as a result of higher viscous dissipation.

Here, we have focused primarily on the evolution of quasi-two-dimensional linear optimal perturbations for a target time of $T = 15$, which attain close to maximum transient growth for the hyperbolic-tangent stratified shear layer (Kaminski *et al.* 2014). As we show in appendix B, despite lower perturbation energy gains the inherently three-dimensional perturbations corresponding to a target time of $T = 5$ may also lead to saturation and modification of the mean flow. Additionally, substantial transient growth may occur over a range of different perturbation wavelengths (see, for instance, the computed gains for differently-sized domains in appendix A, or the results for unstratified flows in Arratia *et al.* (2013)) – the nonlinear evolution described above is not restricted only to a certain wavelength or target time. Altogether, it appears that non-normal perturbation growth may lead to significant nonlinear effects for a variety of different initial conditions, even in flows stable to normal-mode perturbations. Further nonlinear simulations of optimal perturbations with a variety of initial wavelengths and target times would therefore be of great interest, in particular to examine any correlation between the length scales of the initial perturbations with those observed during the transition process.

It is natural to question what would happen for flows with even higher Re and Pr than those considered here, which would be more relevant for the geophysical flows motivating this study. As figure 1(b) shows, the maximum perturbation gain increases with increasing Re , and as we have discussed above the onset of secondary instabilities and mixing of the mean flow are also strongly affected by increasing Re (as shown, for example, in figure 11).

To examine further what might happen at higher Re , we can consider a case in which $Re = 6000$, $Ri_b = 0.40$, and $E_0 = 2 \times 10^{-6}$ (a smaller initial amplitude than the cases considered above). For this set of parameters, by $t = 6.0$ the gradient Richardson number

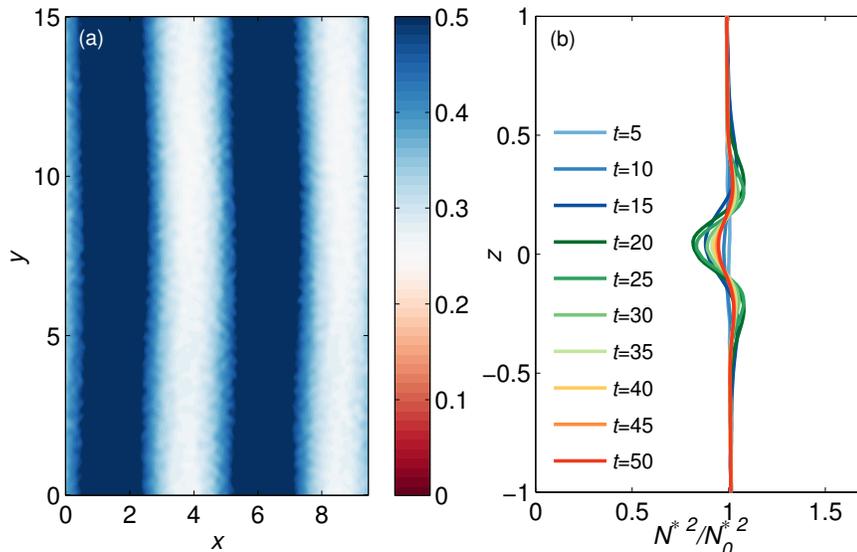


FIGURE 12. Richardson number and buoyancy frequency for $Re = 6000$, $Ri_b = 0.40$, and $E_0 = 2 \times 10^{-6}$. (a) Local gradient Richardson number on a horizontal plane passing through $z = 0$ at $t = 6.0$. Note that the white region corresponds to $Ri_g = 1/4$. (b) Evolution of scaled mean buoyancy frequency, $N^{*2}/N_0^{*2} = \partial b/\partial z$, in region around shear layer.

of the flow is everywhere greater than $1/4$, as seen in figure 12. Unlike the case considered in figure 3 or the higher- Re cases described in figures 5 and 10, this flow does not fully develop into billows or break down into turbulence. However, the perturbation does grow enough to modify the mean flow as seen in figure 12(b). In fact, more modification of the base flow occurs in this case than in the flow shown in figure 6(b), despite the initial perturbation amplitude being a factor of 10 lower in figure 12. This points to transient non-normal perturbation growth as a possible mechanism to trigger turbulence and mixing in stably-stratified geophysical environments.

It should be noted that a similar dependence on Reynolds number has been observed when considering transition to turbulence in unstratified flows. For example, Reddy *et al.* (1998) shows a decrease in the required amplitude for transition with increasing Re due to linear optimal perturbations in Poiseuille and Couette flow. Similarly, Duguet *et al.* (2013) and Cherubini *et al.* (2015) find that the required perturbation amplitude for transition to turbulence triggered by nonlinear optimal perturbations in plane Couette and boundary layer flows, respectively, decreases with increasing Re . Furthermore, Chapman (2002) derives asymptotically the scaling of the decrease in initial perturbation amplitude with various negative powers of Re for subcritical transition to sustained turbulence in plane Couette and plane Poiseuille flow. As such, our observation that nonlinear effects can be observed at lower initial amplitudes for higher- Re stratified flows is consistent with these past results.

In all of our simulations, the Prandtl number has been held fixed at a value of $Pr = 1$, and examining the role of Pr in both the linear and nonlinear behaviour of the optimal perturbations would be of great interest. In particular, recent work examining Kelvin-Helmholtz-unstable flows (Salehipour *et al.* 2015) suggests that the preferred secondary instabilities (and therefore the process by which free stratified shear flows undergo the transition to transient turbulence) as well as the mixing efficiency may depend strongly

on Pr . Examining the nonlinear evolution of linear optimal perturbations at different values of Pr for the more strongly-stratified flows considered here may also reveal a similar dependence on Prandtl number.

An additional point of interest is the emergence of billows as the result of transient perturbation energy growth in flows with $Ri_b > 1/4$. It is interesting to compare the billows seen in figure 3 to those arising from the canonical Kelvin-Helmholtz instability. While the billows seen at $Ri_b = 0.40$ have a higher aspect ratio than those arising in a more weakly-stratified environment, the resemblance is still remarkable, despite a different physical mechanism leading to their formation.

Despite the fact that the predicted growth rate of the normal-mode Kelvin-Helmholtz instability $\sigma_{KH} \rightarrow 0$ as the critical minimum gradient Richardson number of $1/4$ is approached, oceanic observations of the turbulence associated with apparently marginal instability (Smyth & Moum 2013) indicate that the nonlinear billow state has finite amplitude for flows with Ri_g just below the critical value of $1/4$. This is suggestive of an Ri -subcritical transition (as described in the Introduction, with some appropriate Richardson number as the control parameter) where the nonlinear (unstable) billow state commonly referred to as the ‘KH billow’ exists above an as yet unknown threshold $Ri^{-1} > Ri_{KH}^{-1}$. For $Ri^{-1} > 4$, the flow is linearly unstable and approaches the billow state transiently, attaining some maximum amplitude, before secondary instabilities onset. However, for $Ri_{KH}^{-1} < Ri^{-1} < 4$, only sufficiently large amplitude perturbations approach the billow state, (once again attaining transiently some maximum, non-trivial amplitude) while smaller perturbations are attracted back to the linearly stable laminar state.

It is important to reiterate that the billow state itself is an unstable state embedded within an inherently transient flow, in that it is typically prone to a ‘zoo’ of secondary instabilities and the (unforced) flow will inevitably decay back to the laminar state. Therefore, it is not appropriate to assert that this ‘ Ri -subcritical’ transition corresponds formally to a canonical subcritical bifurcation of a dynamical system. Nevertheless, the billow state is an identifiable intermediate state of the transient flow evolution for at least some range of Richardson numbers, and our results support the above-described scenario that the nonlinear billow state persists for Richardson numbers greater than those associated with linear instabilities of the laminar state.

Specifically, we observe that the finite-amplitude billow state may be reached for infinitesimal perturbations via the normal-mode instability when $Ri^{-1} > 4$. For finite perturbations above a certain initial amplitude, non-normal growth allows perturbations to develop and approach the billow state when $Ri_{KH}^{-1} < Ri^{-1} < 4$. We conjecture that the billow state ceases to exist for $Ri^{-1} < Ri_{KH}^{-1}$, in the sense that a flow with such a high Richardson number never approaches a billow-like state at finite amplitude, irrespective of the initial form (or amplitude) of perturbation.

Further examination of the finite-amplitude saturation of perturbations for base flows with $Ri_{g,\min} \sim 1/4$ would help to clarify the details of the (maximum in time) amplitude of the billow state, and in particular to identify the value $Ri_{KH}^{-1} < 4$ (which may be a function of Re and/or Pr) above which this intermediate billow state is approached. It is important to note again that here we are not referring to a transition to a sustained turbulent state, but rather to the unstable finite-amplitude billow state which is then susceptible to secondary instabilities, leading to a transient period of flow evolution with enhanced dissipation rates compared to the laminar state. However, it is still undoubtedly remarkable that despite the lack of linear instability connection to such a state when $Ri^{-1} < 4$, perturbations with a sufficiently large initial amplitude are still able to approach closely to a billow state and become at least ‘Kelvin-Helmholtz-like’ at some point in their transient flow evolution.

It should also be emphasized that larger linear growth does not necessarily imply a more efficient transition to turbulence. For example, in the case of unstratified nonlinear optimal perturbations there may exist quasi-linear optimal perturbations with larger peak energy gains but less efficient transition (Rabin *et al.* 2012). Pralits *et al.* (2015) have shown that weakly nonlinear optimal perturbations may be able to exploit nonlinear modification of the mean flow to allow for enhanced growth and transition to turbulence at amplitudes for which linear optimal perturbations are unable to lead to transition. Even for the linear optimal perturbations considered here there may be other structures with suboptimal linear gains that are nonetheless able to transition to turbulence from lower initial amplitudes. As such, the amplitudes for which the optimal perturbations are able to transition to turbulence should be viewed, in some sense, as an upper bound for the minimum amplitude for transition to turbulence in these strongly stratified flows: there may exist more efficient mechanisms by which lower-amplitude perturbations may be able to transition. In the language of Pringle & Kerswell (2010), we do not expect these linear optimal perturbations to be ‘minimal seeds’ for the transition to turbulence, which must be identified using an inherently nonlinear DAL method (Eaves & Caulfield (2015)). While we would expect, based on our results above, that transition will occur at lower initial amplitudes for increasing Re and decreasing Ri_b , the exact minimal energy needed for transition remains the topic of future work.

Here, we have considered the stability of stratified shear flows inspired by those observed in geophysical settings. However, further modifications can be made to mimic oceanic and atmospheric flows better. Given the strong Re - and Ri_b -dependence observed above, extending this work to more geophysically-relevant parameter regimes is needed before reaching firm conclusions about the importance of the transient growth mechanism in these flows. Additionally, while the base flow considered above is steady and parallel, geophysical flows generally vary in both time and space. Extending this work to more complicated base flows would also be of great interest in applying these results to observations.

Acknowledgements

This manuscript benefited from the constructive comments of the referees. A.K.K. was supported by a NSERC doctoral scholarship. The research activity of C.P.C. and J.R.T. was supported by EPSRC Programme Grant EP/K034529/1, entitled ‘Mathematical Underpinning of Stratified Turbulence’.

Appendix A. Linear perturbation energy gains with different domain sizes

In the above discussion, we have focused primarily on optimal perturbations computed over a domain with $L_x = 9.44$ and $L_y = 15.0$. However, it has been shown in unstratified flows that the optimal gain can depend strongly on the spatial wavelengths of the initial perturbation (see e.g. Andersson *et al.* 1999; Luchini 2000; Arratia *et al.* 2013), and so we consider the optimal gains for different domain sizes here. In figure 13, we show the resulting perturbation energy gains at $T = 15$ for three different values of the horizontal extent L_x and two different values of the spanwise extent L_y for $Re = 1000$ and $Re = 2000$. There is little effect on the maximum linear gain when the spanwise extent is doubled, consistent with the two-dimensional optimal perturbation structure mentioned above.

While varying the horizontal domain size is observed to have an effect on the maximum

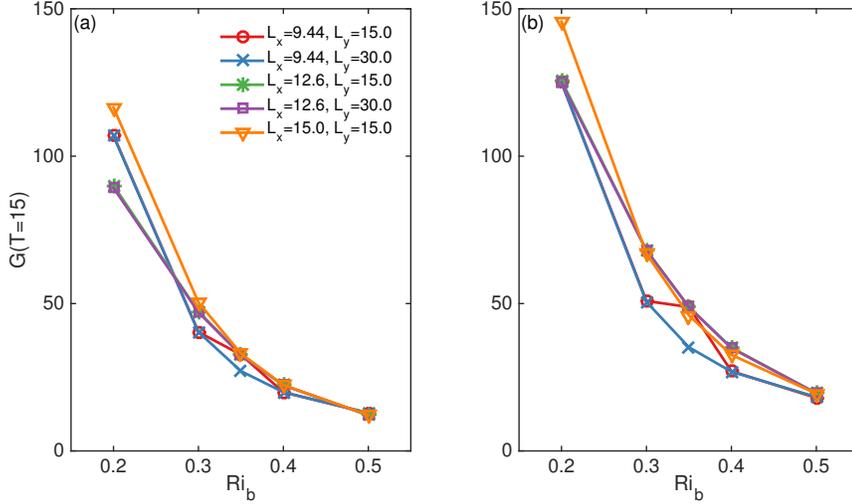


FIGURE 13. Perturbation energy gain for a target time of $T = 15$ and a variety of horizontal and spanwise domain sizes L_x and L_y . (a) $Re = 1000$. (b) $Re = 2000$.

attainable linear growth, generally speaking $G(T)$ remains the same order of magnitude across the sizes considered here. Thus, we would expect growing structures over a variety of wavelengths, rather than restricted to a specific range of length scales.

Appendix B. Nonlinear evolution for linear optimal perturbations with target time $T = 5$

In the above discussion, we have considered optimal perturbations corresponding to a target time of 15 advective units, which are primarily two-dimensional in structure. A set of additional simulations, the details of which are presented in table 2, were run for linear optimal perturbations with a target time $T = 5$, $Ri_b = 0.30$ and 0.40 , and $Re = 1000$, 2000 , and 4000 . The domain size was chosen to be $(L_x, L_y, L_z) = (9.44, 15.0, 30.0)$, and the grid sizes were selected such that $\Delta x_{min} \leq 3\eta$. Only one initial amplitude ($E_0 = 2 \times 10^{-5}$) is considered here. As described in Kaminski *et al.* (2014), the perturbations for this target time take the form of rolls tilted against the background shear flow, as with $T = 15$, but with significant three-dimensional structure. The computed linear perturbation energy gains are approximately $10 - 20$, as shown in figure 14(a), and increase with increasing Reynolds number.

As in figure 4, the nonlinear energy evolution is shown in figure 14(b-d), along with the theoretical linear evolution. As Re is increased, the nonlinear fluctuation energy saturates and deviates from the linear predictions, reaching lower peak amplitudes than the linear case but remaining elevated at later times in the flow evolution. The deviation between the linear and nonlinear cases is more pronounced for $Ri_b = 0.30$ than for $Ri_b = 0.40$. The corresponding kinetic energy dissipation (not shown) reaches peak values comparable to those in figure 8, though the burst of elevated ε_k is short-lived.

Figure 15 shows the mean scaled buoyancy frequency N^{*2}/N_0^{*2} and mean gradient Richardson number for the simulations with $Re = 4000$. As in figure 6, the perturbation evolution leads to modification of the background buoyancy profile. However, the resulting structure of the mean flow shows two minima located above and below the centre of

Re	Ri_b	(N_x, N_y, N_z)
1000	0.30	(128, 128, 201)
1000	0.40	(128, 128, 201)
2000	0.30	(192, 256, 301)
2000	0.40	(128, 192, 301)
4000	0.30	(256, 384, 401)
4000	0.40	(256, 384, 401)

TABLE 2. Parameters for nonlinear simulations of $T = 5$ optimal perturbations with $E_0 = 2 \times 10^{-5}$.

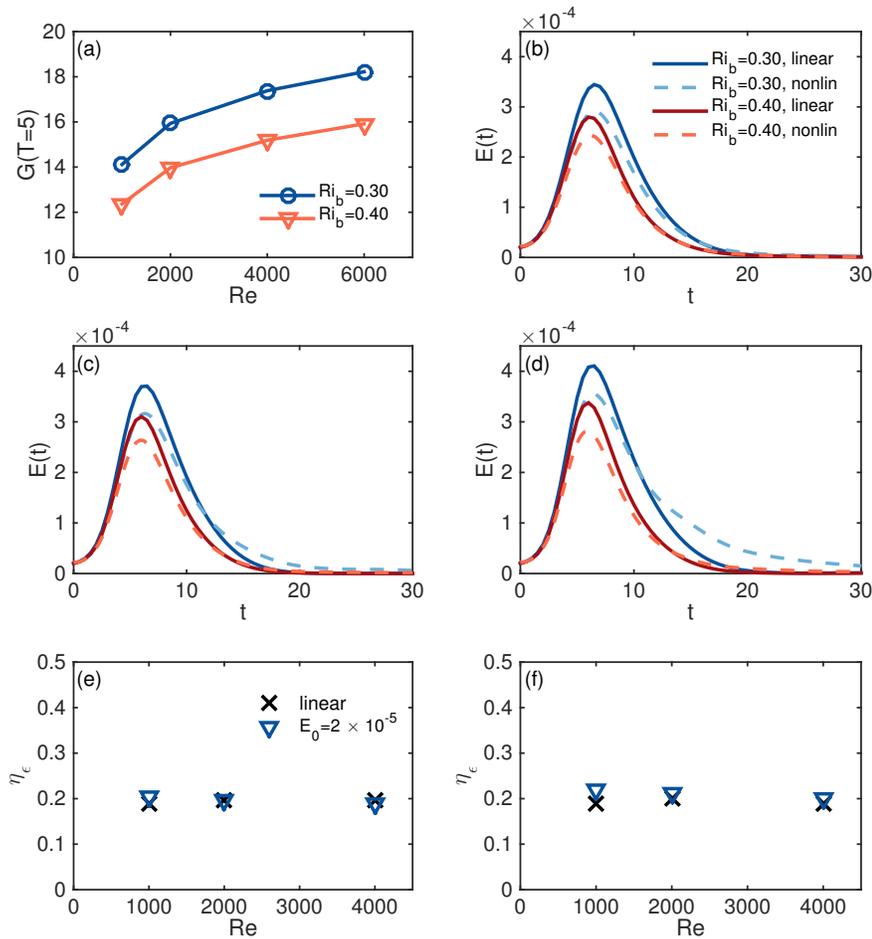


FIGURE 14. (a) Perturbation energy gain for $T = 5$ as a function of Re . (b-d) Linear and nonlinear perturbation energy evolution for $E_0 = 2 \times 10^{-5}$, $Ri_b = 0.30$ and 0.40 , and (b) $Re = 1000$, (c) $Re = 2000$, and (d) $Re = 4000$. (e-f) Computed linear and nonlinear mixing efficiencies, η_ϵ , as a function of Re for (e) $Ri_b = 0.30$ and (f) $Ri_b = 0.40$.

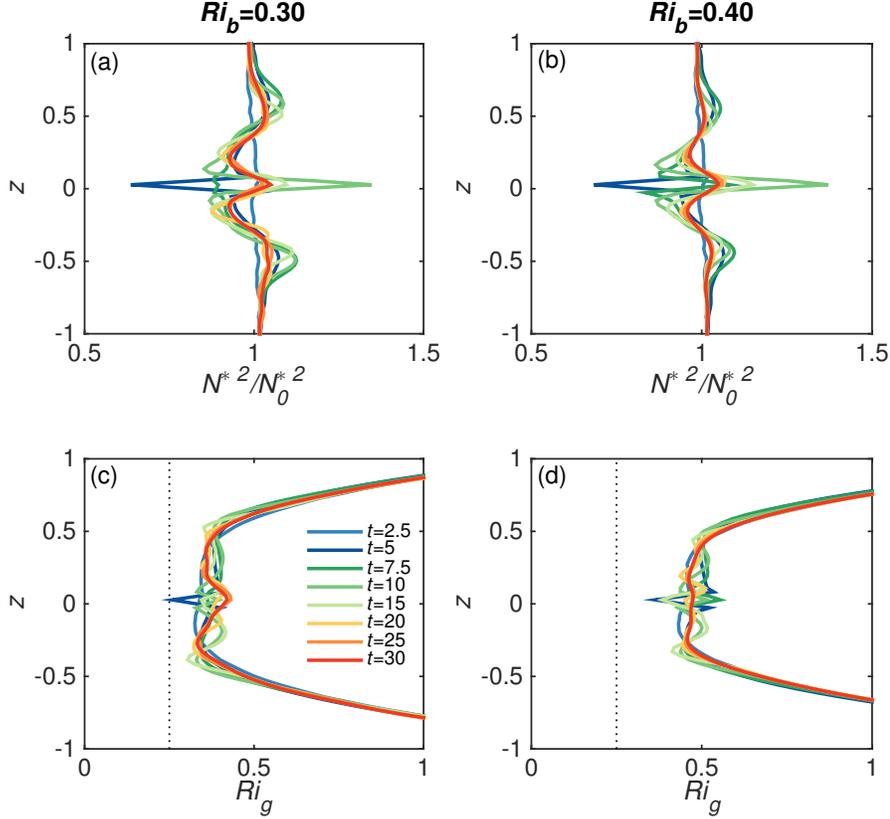


FIGURE 15. (Top row) Time evolution of mean scaled buoyancy frequency, $N^{*2}/N_0^{*2} = \partial\bar{b}/\partial z$, for initial perturbations corresponding to a target time $T = 5$ with $E_0 = 2 \times 10^{-5}$, $Re = 4000$, and (a) $Ri_b = 0.30$ and (b) $Ri_b = 0.40$. (Bottom row) Time evolution of mean gradient Richardson number for the same cases, i.e. (c) $Ri_b = 0.30$ and (d) $Ri_b = 0.40$.

the shear layer rather the layered profiles seen for the $T = 15$ cases. The changes in background stratification are less dramatic than those seen in figure 6, and the corresponding mean Ri_g does not drop below $1/4$.

The mixing efficiency η_ε associated with the evolution of the optimal perturbations with $T = 5$ is shown in figure 14(e-f). Both the linear and nonlinear mixing efficiencies are much lower for the $T = 5$ linear optimal perturbations (with values of $\eta_\varepsilon \sim 0.2$) than for their $T = 15$ counterparts. This decrease in mixing efficiency is due primarily to higher kinetic energy dissipation relative to potential energy dissipation for the three-dimensional $T = 5$ perturbations when compared with their two-dimensional $T = 15$ counterparts. The nonlinear simulations generally have slightly higher mixing efficiencies than the linear predictions, owing to an increase in ε_p around $t = 10$, when the strongest changes in mean $\partial B/\partial z$ are also occurring (figure 15). It would be interesting to see if, at higher initial amplitude or higher Reynolds number, the nonlinear mixing efficiencies of the $T = 5$ optimal perturbations further increased to the $\eta_\varepsilon \sim 0.3$ values seen for $T = 15$ above and by Maffioli *et al.* (2016) and Venayagamoorthy & Koseff (2016).

It is clear from figures 14 and 15 that the inherently three-dimensional linear optimal perturbations corresponding to shorter target times may also show strongly nonlinear

effects at higher Re , even for strongly-stratified base flows. This may point to transient growth as an important process in normal-mode stable flows over a variety of target times and initial conditions.

REFERENCES

- ANDERSSON, P., BERGGREN, M. & HENNINGSON, D. S. 1999 Optimal disturbances and bypass transition in boundary layers. *Phys. Fluids* **11** (1), 134–150.
- ARRATIA, C., CAULFIELD, C. P. & CHOMAZ, J.-M. 2013 Transient perturbation growth in time-dependent mixing layers. *J. Fluid Mech.* **717**, 90–133.
- AUGIER, P., BILLANT, P., NEGRETTI, M. E. & CHOMAZ, J.-M. 2014 Experimental study of stratified turbulence forced with columnar dipoles. *Phys. Fluids* **26** (4), 046603.
- BRUCKER, K. A. & SARKAR, S. 2007 Evolution of an initially turbulent stratified shear layer. *Phys. Fluids* **19**, 105105.
- CAULFIELD, C. P. & PELTIER, W. R. 2000 The anatomy of the mixing transition in homogeneous and stratified free shear layers. *J. Fluid Mech.* **413**, 1–47.
- CHAPMAN, S. J. 2002 Subcritical transition in channel flows. *J. Fluid Mech.* **451**.
- CHERUBINI, S., DE PALMA, P. & ROBINET, J.-C. 2015 Nonlinear optimals in the asymptotic suction boundary layer: Ttransition thresholds and symmetry breaking. *Phys. Fluids* **27**, 034108.
- DUGUET, Y., MONOKROUSOS, A., BRANDT, L. & HENNINGSON, D. S. 2013 Minimal transition thresholds in plane Couette flow. *Phys. Fluids* **25**, 084103.
- EAVES, T. S. & CAULFIELD, C. P. 2015 Disruption of SSP/VWI states by a stable stratification. *J. Fluid Mech.* **784**, 548–564.
- ELLINGSEN, T. & PALM, E. 1975 Stability of linear flow. *Phys. Fluids* **18** (4), 487–488.
- FARRELL, B. F. & IOANNOU, P. J. 1993a Optimal excitation of three-dimensional perturbations in viscous constant shear flow. *Phys. Fluids A* **5** (6), 13901400.
- FARRELL, B. F. & IOANNOU, P. J. 1993b Transient development of perturbations in stratified shear flow. *J. Atmos. Sci.* **50** (14), 2201–2214.
- GARRETT, C. 2003 Mixing with latitude. *Nature* **422**, 477–478.
- HOWARD, L. N. 1961 Note on a paper of John W. Miles. *J. Fluid Mech.* **10**, 509–512.
- IVEY, G. N., WINTERS, K. B. & KOSEFF, J. R. 2008 Density stratification, turbulence, but how much mixing? *Annu. Rev. Fluid Mech.* **40**, 169–184.
- KAMINSKI, A. K., CAULFIELD, C. P. & TAYLOR, J. R. 2014 Transient growth in strongly stratified shear layers. *J. Fluid Mech.* **758**, R4.
- KERSWELL, R. R., PRINGLE, C. C. T. & WILLIS, A. P. 2014 An optimization approach for analyzing nonlinear stability with transition to turbulence in fluids as an exemplar. *Rep. Prog. Phys.* **77**, 085901.
- KLAASSEN, G. P. & PELTIER, W. R. 1985 Evolution of finite amplitude Kelvin-Helmholtz billows in two spatial dimensions. *J. Atmos. Sci.* **42** (12), 1321–1339.
- KUHLBRODT, T., GRIESEL, A., MONTOYA, M., LEVERMANN, A., HOFMANN, M. & RAHMSTORF, S. 2007 On the driving processes of the Atlantic meridional overturning circulation. *Rev. of Geophys.* **45**, RG2001.
- KUNZE, E., WILLIAMS III, A. J. & BRISCOE, M. G. 1990 Observations of shear and vertical stability from a neutrally buoyant float. *J. Geophys. Res.* **95**, 18127–18142.
- LUCHINI, P. 2000 Reynolds-number-independent instability of the boundary layer over a flat surface: optimal perturbations. *J. Fluid Mech.* **404**, 289–309.
- MACK, S. A. & SCHOEBERLEIN, H. C. 2004 Richardson number and ocean mixing: towed chain observations. *J. Phys. Oceanogr.* **34**, 736–754.
- MAFFIOLI, A., BRETHOUWER, G. & LINDBORG, E. 2016 Mixing efficiency in stratified turbulence. *J. Fluid Mech.* **794**, R3.
- MASHAYEK, A., CAULFIELD, C. P. & PELTIER, W. R. 2013 Time-dependent, non-monotonic mixing in stratified turbulent shear flows: implications for oceanographic estimates of buoyancy flux. *J. Fluid Mech.* **736**, 570–593.
- MASHAYEK, A. & PELTIER, W. R. 2012 The ‘zoo’ of secondary instabilities precursory to strat-

- ified shear flow transition. part 1 Shear aligned convection, pairing, and braid instabilities. *J. Fluid Mech.* **708**, 5–44.
- MELLOR, G. L. & YAMADA, T. 1982 Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.* **20** (4), 851–875.
- MILES, J. W. 1961 On the stability of heterogeneous shear flows. *J. Fluid Mech.* **496**, 496–508.
- ORR, W. M'F. 1907 The stability or instability of the steady motions of a perfect liquid and of a viscous liquid. Part I: A perfect liquid. *Proc. R. Irish Acad. A* **27**, 9–68.
- PELTIER, W. R. & CAULFIELD, C. P. 2003 Mixing efficiency in stratified shear flows. *Annu. Rev. Fluid Mech.* **35**, 135–167.
- POLZIN, K. 1996 Statistics of the Richardson number: mixing models and finestructure. *J. Phys. Oceanogr.* **26**, 1409–1425.
- POPE, S. B. 2000 *Turbulent Flows*. Cambridge, UK: Cambridge Univ. Press.
- PRALITS, J. O., BOTTARO, A. & CHERUBINI, S. 2015 Weakly nonlinear optimal perturbations. *J. Fluid Mech.* **785**, 135–151.
- PRICE, J. F., WELLER, R. A. & PINKEL, R. 1986 Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. *J. Geophys. Res.* **91** (C7), 8411–8427.
- PRINGLE, C. C. T. & KERSWELL, R. R. 2010 Using nonlinear transient growth to construct the minimal seed for shear flow turbulence. *Phys. Rev. Lett.* **105**, 154502.
- RABIN, S. M. E., CAULFIELD, C. P. & KERSWELL, R. R. 2012 Triggering turbulence efficiently in plane Couette flow. *J. Fluid Mech.* **712**, 244–272.
- REDDY, S. C., SCHMID, P. J., BAGGETT, J. S. & HENNINGSON, D. S. 1998 On stability of streamwise streaks and transition thresholds in plane channel flows. *J. Fluid Mech.* **365**, 269–303.
- RILEY, J. J. & DEBRUYNKOPS, S. M. 2003 Dynamics of turbulence strongly influenced by buoyancy. *Phys. Fluids* **15** (7), 2047–2059.
- SALEHIPOUR, H., CAULFIELD, C. P. & PELTIER, W. R. 2016 Turbulent mixing due to the Holmboe wave instability at high Reynolds number. *J. Fluid Mech.* **803**, 591–621.
- SALEHIPOUR, H., PELTIER, W. R. & MASHAYEK, A. 2015 Turbulent diapycnal mixing in stratified shear flows: the influence of Prandtl number on mixing efficiency and transition at high Reynolds number. *J. Fluid Mech.* **773**, 178–223.
- SMYTH, W. D. & MOUM, J. N. 2012 Ocean mixing by Kelvin-Helmholtz instability. *Oceanography* **25** (2), 140–149.
- SMYTH, W. D. & MOUM, J. N. 2013 Marginal instability and deep cycle turbulence in the eastern equatorial Pacific Ocean. *Geophys. Res. Lett.* **40**, 6181–6185.
- SMYTH, W. D., MOUM, J. N. & CALDWELL, D. R. 2001 The efficiency of mixing in turbulent patches: Inferences from direct simulations and microstructure observations. *J. Phys. Oceanogr.* **31**, 1969–1992.
- TAYLOR, J. R. 2008 Numerical simulations of the stratified oceanic bottom boundary layer. PhD thesis, University of California, San Diego.
- TEARLE, M. O. 2004 Optimal perturbation analysis of stratified shear flow. PhD thesis, University of Colorado.
- THORPE, S. A. 1973 Experiments on instability and turbulence in a stratified shear flow. *J. Fluid Mech.* **61** (4), 731–751.
- TREFETHEN, L. N., TREFETHEN, A. E., REDDY, S. C. & DRISCOLL, T. A. 1993 Hydrodynamic stability without eigenvalues. *Science* **261**, 578–584.
- VENAYAGAMOORTHY, S. K. & KOSEFF, J. R. 2016 On the flux Richardson number in stably stratified turbulence. *J. Fluid Mech.* **798**, R1.
- WUNSCH, C. & FERRARI, R. 2004 Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.* **36**, 281–314.