J. Fluid Mech. (2019), *vol.* 862, *pp.* 639–658.
doi:10.1017/jfm.2018.973

Stratified shear instability in a field of pre-existing turbulence

A. K. Kaminski[1,†] **and W. D. Smyth**[1]

[1]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA

(Received 6 April 2018; revised 26 November 2018; accepted 29 November 2018)

Turbulent mixing of heat and momentum in the stably-stratified ocean interior occurs in discrete events driven by vertical variations of the horizontal velocity. Typically, these events have been modelled assuming an initially laminar stratified shear flow which develops wavelike instabilities, becomes fully turbulent, and then relaminarizes into a stable state. However, in the real ocean there is always some level of turbulence left over from previous events. Using direct numerical simulations, we show that the evolution of a stably-stratified shear layer may be significantly modified by pre-existing turbulence. The classical billow structure associated with Kelvin–Helmholtz instability is suppressed and eventually eliminated as the strength of the initial turbulence is increased. A corresponding energetics analysis shows that potential energy changes and dissipation of kinetic energy depend non-monotonically on initial turbulence strength, with the largest effects when initial turbulence is present but insufficient to prevent billow formation. The mixing efficiency decreases with increasing initial turbulence amplitude as the development of the Kelvin–Helmholtz billow, with its large pre-turbulent mixing efficiency, is arrested.

Key words: stratified flows, transition to turbulence, turbulent mixing

1. Introduction

In the stably-stratified ocean interior, turbulent mixing is thought to occur in spatially- and temporally-intermittent events driven by vertical shear of horizontal currents (Ivey, Winters & Koseff 2008). These discrete mixing events are often modelled by the Kelvin–Helmholtz instability (Smyth & Moum 2012). The characteristic 'billow' structure associated with the finite-amplitude manifestation of this instability has been observed in a variety of flow environments, ranging from deep ocean settings (van Haren *et al.* 2014) to estuaries (Holleman, Geyer & Ralston 2016). There is a large body of literature examining the process by which this instability triggers turbulent mixing and its dependence on the choice of initial flow parameters (Klaassen & Peltier 1985; Peltier & Caulfield 2003; Smyth 2003; Mashayek & Peltier 2012*a*,*b*; Salehipour, Peltier & Mashayek 2015).

† Present address: Applied Physics Laboratory, University of Washington, Seattle, WA 98105, USA. Email address for correspondence: alexis.kaminski@gmail.com

Turbulent mixing in stratified fluids is frequently parameterized in terms of a mixing efficiency, η, the ratio of energy used in irreversibly mixing the background flow to that lost to viscous dissipation. The mixing efficiency may then be used to relate the viscous dissipation to the turbulent flux of density as $K_\rho = \Gamma\varepsilon/N^2$, where K_ρ is the turbulent mass diffusivity, ε is the viscous dissipation, and $\Gamma = \eta/(1-\eta)$ is known as the mixing coefficient (Osborn 1980; Salehipour & Peltier 2015; Salehipour *et al.* 2016*b*). This parameterization has been used in both large-scale modelling and the interpretation of observational data. A constant value of $\Gamma = 0.2$ ($\eta = 0.17$) is often assumed (Gregg *et al.* 2018). However, observations, laboratory experiments and numerical simulations have all shown the mixing efficiency to depend on a variety of flow parameters, even for idealized flows such as the Kelvin–Helmholtz instability (Ivey *et al.* 2008; Mashayek, Caulfield & Peltier 2013; Gregg *et al.* 2018). Furthermore, the efficiency varies in time throughout the lifecycle of the instability (Peltier & Caulfield 2003).

Previous work on turbulence triggered by stratified shear instabilities has typically considered laminar initial conditions (see, e.g. the summary of DNS studies in Gregg *et al.* (2018)). However, geophysical flows are rarely laminar; turbulence is expected, for example, as the result of previous mixing events. Recent work by Thorpe, Smyth & Li (2013) and Li, Smyth & Thorpe (2015) has shown that pre-existing turbulence, modelled as enhanced (eddy) viscosity and diffusivity, may stabilize or destabilize the flow depending on the vertical distribution of the turbulence. Additionally, a numerical study by Brucker & Sarkar (2007) examines the evolution of initially turbulent shear layers and shows that strongly turbulent initial conditions prevent the formation of coherent billows in both stratified and unstratified flows.

Motivated both by observations of shear instability in geophysical flows and the aforementioned studies highlighting the potential impact of turbulent initial conditions, we ask how pre-existing turbulence might affect mixing in a stratified shear layer. To address this question, we perform a series of direct numerical simulations of transition to turbulence in shear layers with a range of initial flow parameters. We systematically vary the initial turbulence strength for each case, from laminar initial conditions (as in previous studies of the Kelvin–Helmholtz instability) to strongly turbulent (as in Brucker & Sarkar (2007)). In § 2 we describe the setup for our direct numerical simulations. We then present the results of a set of simulations with various initial perturbation amplitudes in § 3.1, and show that the evolution of the shear layer depends strongly on the perturbation amplitude, in terms of both its qualitative characteristics and the resulting energetics and turbulent mixing. In § 3.2 we describe how these results are affected by the bulk stratification and relative importance of viscosity, and show that the overall qualitative effects of the initial turbulence are robust across the parameter space considered. Finally, in § 4 we discuss the broader implications of our results for our understanding of stratified shear layers in geophysical flows.

2. Simulations

We begin with a stably-stratified parallel shear layer,

$$U^*(z^*) = U_0^* \tanh(z^*/h^*) \quad \text{and} \quad B^*(z^*) = B_0^* \tanh(z^*/h^*), \tag{2.1a,b}$$

where U_0^* and B_0^* are half the velocity and buoyancy difference across the shear layer (where buoyancy is defined as $B^* = -g^*(\rho^* - \rho_0^*)/\rho_0^*$, with reference density ρ_0^*), h^* is the initial shear layer half-width, z^* is the vertical coordinate and the asterisks

denote dimensional quantities. We non-dimensionalize velocities by U_0^*, buoyancy by B_0^*, lengths by h^* and time by the advective timescale h^*/U_0^*. The non-dimensional shear layer is then

$$U(z) = B(z) = \tanh(z). \tag{2.2}$$

The evolution of the velocity $\boldsymbol{u}$ and buoyancy b are described by the Boussinesq Navier-Stokes, conservation of buoyancy and mass continuity equations, given in non-dimensional form as:

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}p + Ri_b b\hat{\boldsymbol{z}} + \frac{1}{Re}\nabla^2\boldsymbol{u}, \tag{2.3}$$

$$\frac{\partial b}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}b = \frac{1}{RePr}\nabla^2 b, \tag{2.4}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0. \tag{2.5}$$

The flow is governed by three non-dimensional parameters, namely the initial Reynolds, Prandtl and initial bulk Richardson numbers,

$$Re = \frac{U_0^* h^*}{\nu^*}, \quad Pr = \frac{\nu^*}{\kappa^*}, \quad Ri_b = \frac{B_0^* h^*}{U_0^{*2}}, \tag{2.6a–c}$$

which describe the relative importance of viscosity compared with inertia, diffusion to viscosity and stratification to shear, respectively. It should be noted that here we define Re based on half the shear layer thickness and velocity difference; care should be taken when comparing with other studies of shear-induced mixing which may use the full shear layer thickness and velocity difference. The initial bulk Richardson number also corresponds to the minimum gradient Richardson number of the undisturbed flow (2.2),

$$Ri_g = \frac{\partial b^*/\partial z^*}{(\partial u^*/\partial z^*)^2} = Ri_b \frac{\partial b/\partial z}{(\partial u/\partial z)^2}, \tag{2.7}$$

which occurs at the centre of the shear layer, $z = 0$. As $Ri_b < 1/4$, instability may be possible in this flow (Howard 1961; Miles 1961).

To model the effects of pre-existing turbulence such as that which might be left over from previous mixing events, we consider an initial perturbation centred on the shear layer. Motivated by observed and predicted spectra for stably-stratified turbulent flows (Brethouwer *et al.* 2007), here we consider initial perturbations with

$$E_{pert} \sim \mathcal{A}^2 k_h^{-5/3} \mathrm{e}^{\mathrm{i}\phi}, \tag{2.8}$$

where $k_h = \sqrt{k_x^2 + k_y^2}$ is a horizontal wavenumber, $\mathcal{A}$ is a characteristic perturbation amplitude and ϕ is a random phase. All three components of the velocity are perturbed, and their incompressible projection is superimposed on the initial base flow (2.2), with a Gaussian profile (with standard deviation 1, or h^* in dimensional form) in the vertical.

While previous simulations of instability arising in stratified shear layers have typically considered small-amplitude initial perturbation amplitudes, here we systematically vary the perturbation amplitude by varying $\mathcal{A}$. We consider values that represent a 'laminar' shear layer ($\mathcal{A} = 0.0025$, as a non-zero value is needed to trigger the instability) and shear layers with weak, intermediate and strong initial turbulence ($\mathcal{A} = 0.01$, 0.05 and 0.1, respectively). The simulations are carried

out using DIABLO, which uses a mixed implicit-explicit timestepping routine with pressure projection. The viscous and diffusive terms are stepped forward using the second-order Crank–Nicolson method; other terms employ a third-order Runge–Kutta–Wray algorithm (Taylor 2008). The domain is periodic in the streamwise and spanwise (x and y) directions, while free-slip, impermeable, no-flux boundaries are applied at the vertical (z) extents of the domain. The grid is clustered in the vertical direction around the centre of the shear layer. The domain has a streamwise extent corresponding to one wavelength of the most unstable mode of (2.2) as predicted by the Taylor-Goldstein equation (e.g. Miles 1961), $L_x \sim 14.0$, a spanwise extent of $L_y = 5.0$ to allow for the development of secondary instabilities (Mashayek & Peltier 2012*b*), and $L_z = 25.0$ to minimize boundary effects. Grid sizes are chosen to resolve approximately 2.5 times the Kolmogorov scale, $L_K = (Re^{-3}/\varepsilon)^{1/4}$, after transition to turbulence (where ε is the viscous dissipation of kinetic energy). The simulation resolution may impact the computed values of key mixing quantities, as discussed in Gregg *et al.* (2018); in light of this, we discuss our choice of grid resolution further in appendix A. The choices of Re and Ri_b and corresponding grid sizes for each set of simulations are presented in table 1.

3. Results

Our goal is to explore mixing energetics in the three-dimensional space of initial parameter values Re, Ri_b and $\mathcal{A}$. We begin by focusing on a single mean flow with $Re = 2000$ and $Ri_b = 0.16$ and comparing results with different $\mathcal{A}$. We then repeat the analysis with different values of Re and Ri_b.

3.1. *Effects of initial perturbation amplitude*

In all simulations shown here there is a strong transient adjustment associated with the initial perturbation, with a decrease in kinetic energy, growth of buoyancy fluctuations and enhanced dissipation. This adjustment takes place over the first few non-dimensional time units, on scales comparable to the advective timescale ($t^* \sim h^*/U_0^*$) and buoyancy timescale ($t^* \sim h^*/\sqrt{B_0}^*$), both of which are $O(1)$ quantities when non-dimensionalized.

Profiles of the stratification and the gradient Richardson number of the mean flow (horizontally averaged) at $t = 1$ (i.e. after the flow has started to adjust to the initial perturbation) are shown in figure 1. It is clear that the initial turbulence has led to a more complicated mean flow, particularly for the intermediate and strongly turbulent cases ($\mathcal{A} = 0.05$ and $\mathcal{A} = 0.1$). However, it should be emphasized that in spite of the initial turbulence, the large-scale stratification still varies over the same length scale, and the mean Richardson number still exhibits a minimum on the order of Ri_b. As such, the initial perturbation does not appear to significantly change the bulk stability of the flow.

In addition to the characteristic amplitude of the velocity perturbation, we can further quantify the initial turbulence in terms of a buoyancy Reynolds number, $Re_b = \varepsilon^*/\nu^* N^{*2}$, which describes the turbulence intensity. Characteristic values of Re_b at $t = 1$ corresponding to each individual simulation are listed in table 1. Values vary from $Re_b \lesssim 10$ for the laminar and weak initial turbulence cases ($\mathcal{A} = 0.0025, 0.01$), at which viscous effects are expected to be important (Shih *et al.* 2005) to $Re_b \sim O(10^2)$ for the strong initial turbulence cases, corresponding to highly energetic turbulence. For context, oceanic observations suggest a broad range of $Re_b \sim O(10) - O(10^4)$, with $Re_b \gtrsim 10^2$ during vigorous mixing (Smyth, Moum & Caldwell 2001).

Re	Ri_b	(N_x, N_y, N_z)	L_x	$\mathcal{A}$	Re_b at $t = 1$
1000	0.12	(384, 128, 501)	14.3	0.0025	5.72
				0.01	8.31
				0.05	66.04
				0.1	191.45
1000	0.16	(384, 128, 501)	14.0	0.0025	4.28
				0.01	6.05
				0.05	47.12
				0.1	134.32
2000	0.12	(768, 256, 1153)	14.3	0.0025	5.79
				0.01	9.31
				0.05	84.78
				0.1	224.76
2000	0.16	(512, 192, 751)	14.0	0.0025	4.44
				0.01	8.52
				0.05	89.96
				0.1	224.04
2000	0.20	(384, 128, 501)	13.3	0.0025	3.62
				0.01	8.03
				0.05	103.47
				0.1	268.64
4000	0.16	(1024, 384, 1729)	14.0	0.0025	4.55
				0.01	10.17
				0.05	113.66
				0.1	271.47
4000	0.20	(758, 256, 1153)	13.3	0.0025	3.76
				0.01	10.07
				0.05	119.29
				0.1	265.33

TABLE 1. Flow parameters, grid sizes and initial turbulence strength of direct numerical simulations. All simulations shown have $Pr = 1$, $L_y = 5.0$ and $L_z = 25.0$.

We now present a qualitative overview of the impact of perturbation amplitude on the evolution of the shear instability. Figure 2 shows vertical slices of buoyancy through $y = 0$ at different times for cases with weak, intermediate and strong initial turbulence; the laminar case follows a very similar evolution to the case with weak initial turbulence and is therefore not shown here. For laminar and weakly turbulent initial conditions, the flow follows the usual lifecycle of a classical Kelvin–Helmholtz instability (Peltier & Caulfield 2003; Smyth & Moum 2012): two-dimensional perturbations grow and form finite-amplitude billows (figure 2*b*,*c*), which are susceptible to three-dimensional secondary instabilities leading to a transition to turbulence (figure 2*d*). As the flow considered here is unforced, it eventually relaminarizes after a period of turbulent mixing (figure 2*e*).

As the initial perturbation amplitude is increased, the details of the flow evolution are modified. For the case of intermediate-amplitude initial turbulence ($\mathcal{A} = 0.05$, figure 2*f–j*), a growing perturbation still leads to billow formation, transition to turbulence, and subsequent relaminarization. However, the growing perturbation

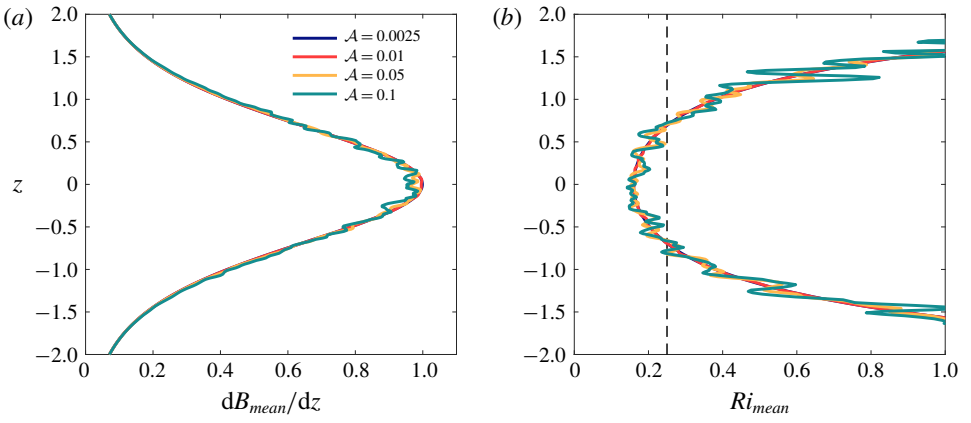

FIGURE 1. Vertical profiles of (*a*) mean buoyancy gradient, dB_{mean}/dz, and (*b*) gradient Richardson number of the mean flow, Ri_{mean} at $t = 1$. $Ri = 1/4$ is denoted by a thin dashed line in panel (*b*). Note that the line corresponding to $\mathcal{A} = 0.0025$ lies almost directly beneath the line for $\mathcal{A} = 0.01$.

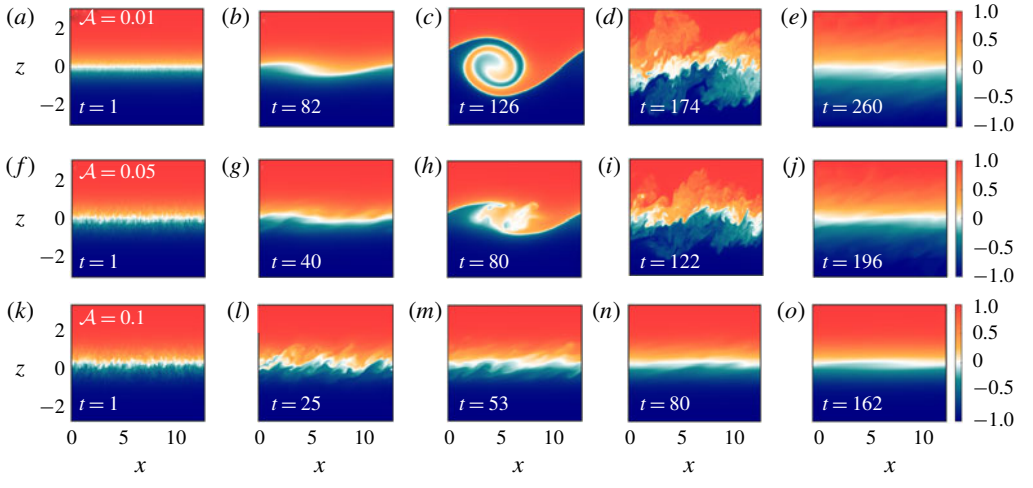

FIGURE 2. Vertical slices of buoyancy through $y = 0$ at various times for (*a–e*) weakly turbulent initial conditions ($\mathcal{A} = 0.01$), (*f–j*) intermediate initial conditions ($\mathcal{A} = 0.05$), and (*k–o*) strongly turbulent initial conditions ($\mathcal{A} = 0.1$). Non-dimensional times corresponding to each snapshot are shown in the bottom left corner of each panel.

is no longer two-dimensional, and the billow becomes turbulent much earlier in its development, without rolling up to the extent seen in the laminar and weakly turbulent cases (compare figure 2*c*,*h*). As the strength of the initial perturbation is further increased, the characteristic billows associated with Kelvin–Helmholtz instability are no longer observed (figure 2*k–o*). Instead, the initial turbulence fills the shear layer and subsequently decays.

Figure 3(*a*) compares the initial buoyancy given in (2.2) with the horizontally-averaged buoyancy B_{mean} after the flow has relaminarized. The vertical extent of the relaminarized layer depends strongly on whether finite-amplitude billows (figure 2*c*,*h*) form and move denser fluid upwards (and *vice versa*). As a result, the final state of

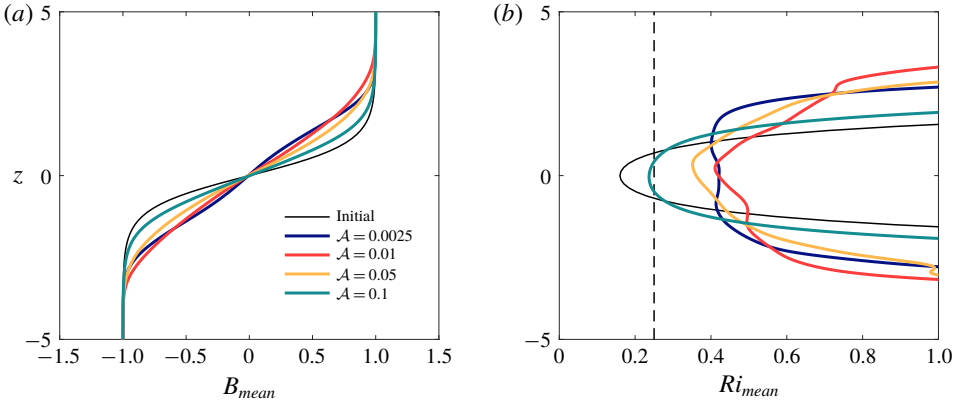

FIGURE 3. Initial (thin black lines) and final (thick coloured lines) vertical profiles of (*a*) mean buoyancy, B_{mean}, and (*b*) gradient Richardson number of the mean flow, Ri_{mean}. $Ri = 1/4$ is denoted by a thin dashed line in panel (*b*).

the flow with strongly turbulent initial conditions (without billow formation) has a much smaller vertical extent than any of the other cases shown. A similar trend is observed for the mean streamwise velocity (not shown here).

In addition to B_{mean}, the gradient Richardson number of the final mean flow, Ri_{mean}, is shown in figure 3(*b*). The minimum Richardson number at the end of the mixing event is well above the critical Miles-Howard value of 1/4 for the laminar, weak and intermediate initial turbulence cases, consistent with experimental results suggesting that the Richardson number after a Kelvin–Helmholtz mixing event should lie in the range 0.32 ± 0.06 (Thorpe 1973). In contrast, strong initial turbulence results in much lower Richardson numbers after relaminarization. The final state is only marginally stable at the centre of the shear layer (green curve).

The effect of varying the initial perturbation amplitude can be studied more quantitatively by considering the evolution of the total kinetic and potential energy of the flow. An equation governing the evolution of the kinetic energy, $\mathcal{K} \equiv \overline{u_j u_j}/2$ (where the overline denotes a volume average) is formed by taking the inner product of (2.3) with $\boldsymbol{u}$ and then averaging over the flow domain, i.e.

$$\frac{\mathrm{d}\mathcal{K}}{\mathrm{d}t} = \mathcal{B} - \varepsilon, \tag{3.1}$$

where $\mathcal{B} \equiv Ri_b \overline{bw}$ is the vertical buoyancy flux and $\varepsilon \equiv \overline{\partial_j u_i \partial_j u_i}/Re$ is the (positive-definite) kinetic energy dissipation. Similarly, an equation governing the total potential energy $\mathcal{P} = -Ri_b \overline{bz}$ is constructed by multiplying (2.4) by $Ri_b z$ and volume-averaging, i.e.

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}t} = -\mathcal{B} + \mathcal{D}_p, \tag{3.2}$$

where $\mathcal{D}_p \equiv Ri_b(b_{top} - b_{bottom})/(RePrL_z)$ denotes the change in potential energy due to diffusion of the mean laminar flow (i.e. the diffusive change in potential energy that would occur even in the absence of any fluid motion). In contrast, the buoyancy flux $\mathcal{B}$ describes the exchange of energy between kinetic and potential forms due to fluid motion, including both reversible stirring and irreversible mixing (Winters *et al.* 1995; Peltier & Caulfield 2003; Davies Wykes, Hughes & Dalziel 2015).

To isolate the irreversible mixing from the total buoyancy flux, we decompose the total potential energy into available and background components, $\mathcal{P} = \mathcal{P}_a + \mathcal{P}_b$. The background potential energy refers to the minimum potential energy state of the system which would be achieved by adiabatically rearranging the buoyancy field into a gravitationally stable configuration,

$$\mathcal{P}_b = -Ri_b\overline{b^\star z}. \tag{3.3}$$

The sorted profile $b^\star(z)$ may be calculated from the full three-dimensional buoyancy field either by sorting or by computing its probability density function, as we do here (Winters *et al.* 1995; Tseng & Ferziger 2001; Davies Wykes *et al.* 2015). The available potential energy $\mathcal{P}_a$ is then the energy that may be used to mix the flow and arises due to horizontal buoyancy gradients or from statically-unstable regions of the flow (Lorenz 1955). The governing equation for potential energy (3.2) can then be split into equations governing the evolution of these two components,

$$\frac{\mathrm{d}\mathcal{P}_a}{\mathrm{d}t} = -\mathcal{B} - \mathcal{M}, \quad \frac{\mathrm{d}\mathcal{P}_b}{\mathrm{d}t} = \mathcal{M} + \mathcal{D}_p. \tag{3.4a,b}$$

It should be noted that Winters *et al.* (1995) refers to the sum of the terms $\mathcal{M} + \mathcal{D}_p$ as Φ_d, describing the total change in background potential energy. However, here we are interested primarily in the irreversible mixing due to the evolution of the shear instability and its transition to turbulence. As such, we follow the convention of Peltier & Caulfield (2003) and distinguish between changes in $\mathcal{P}_b$ due to fluid motion (described by $\mathcal{M}$) and changes due to laminar diffusion of the mean flow (described by $\mathcal{D}_p$).

The evolution of the kinetic and potential energies are shown in figure 4. In the laminar and weakly turbulent cases ($\mathcal{A} = 0.0025$ and 0.01; blue and red curves) there is little change in kinetic or potential energy until $t \sim 100$ when the linearly-growing perturbation reaches finite amplitude. The change in $\mathcal{P}$ is initially associated with the growth of available potential energy arising from the formation of the billow structure. Then, as the secondary instabilities grow and trigger turbulence, the background potential energy begins to increase irreversibly while the available potential energy decays. Finally, as the flow relaminarizes, the energy changes approach near-constant values.

For the case with strongest initial turbulence ($\mathcal{A} = 0.1$; green curves), the energy evolves in a very different manner, as might be expected based on the buoyancy evolution (figure 2*k–o*). The kinetic and potential energies change quickly as the initial turbulence decays and mixes the mean flow (also seen to some extent for the case with intermediate initial turbulence). Changes in available potential energy are small, and changes in total potential energy largely correspond to changes in $\mathcal{P}_b$ from irreversible mixing. After the initial vigorous mixing, the small remaining changes in energy are largely due to the effects of viscosity and diffusion on the relaminarized mean flow, with a plateau in $\mathcal{P}_b - \mathcal{D}_p t$ (figure 4*c*), and no further increases associated with the growth of the Kelvin–Helmholtz instability.

The case with intermediate-amplitude initial turbulence ($\mathcal{A} = 0.05$; yellow curves) shares features of both the weak and strong initial turbulence cases. There is an initial change in kinetic and potential energy owing to the turbulent initial conditions. Then, as the billow begins to form (figure 2*h*), available potential energy rises, as for the classical Kelvin–Helmholtz instability. However, the flow becomes turbulent before the billow forms completely, and correspondingly the growth of $\mathcal{P}_a$ is arrested as $\mathcal{P}_b$

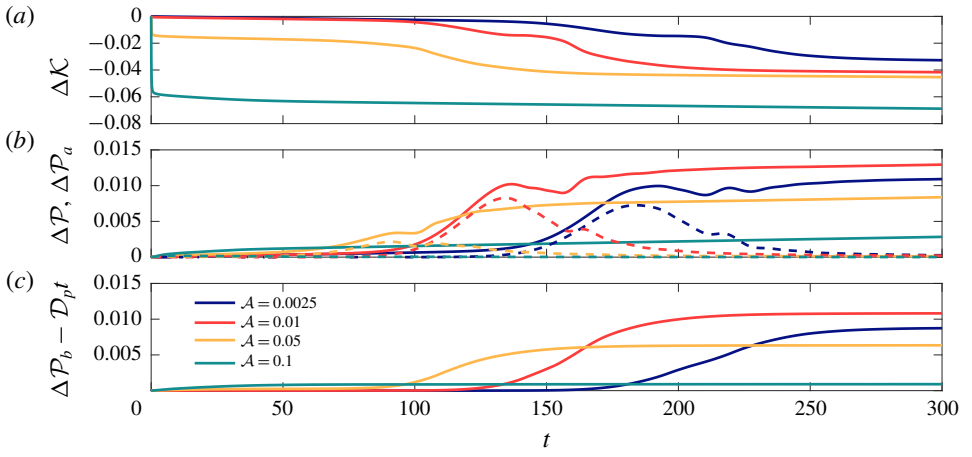

FIGURE 4. Time evolution of changes in (*a*) kinetic energy $\mathcal{K}$, (*b*) total potential energy $\mathcal{P}$ (solid) and available potential energy $\mathcal{P}_a$ (dashed), and (*c*) background potential energy associated with fluid motion, $\mathcal{P}_b - \mathcal{D}_p t$.

begins to rise due to turbulent mixing. Finally, as in the other three cases, the energies plateau as the turbulence decays and the flow relaminarizes. Overall potential energy changes depend non-monotonically on the perturbation amplitude: larger changes in total and background potential energy are observed for $\mathcal{A} = 0.01$ when compared with $\mathcal{A} = 0.0025$ and $\mathcal{A} = 0.05$.

Using the above description of the flow energetics, the mixing efficiency can be defined instantaneously as

$$\eta_i = \frac{\mathcal{M}}{\mathcal{M} + \varepsilon}. \tag{3.5}$$

In the laminar and weak turbulence cases, the evolution of $\mathcal{M}$, ε and η_i (blue and red curves, figure 5*a*,*b*) is similar to that found in previous studies of mixing by Kelvin–Helmholtz instability (Smyth *et al.* 2001; Peltier & Caulfield 2003; Mashayek *et al.* 2013). There is an initial phase during which the irreversible mixing rate is larger than the dissipation, with very high mixing efficiency, $\eta_i \sim 0.6$. This occurs shortly after the peak in available potential energy (figure 4*b*), and is associated with the rollup of the shear instability into the classical billow structure. Then, as the flow becomes turbulent, the mixing efficiency drops to a value $\eta_i \sim 0.3$ as ε peaks. As the flow relaminarizes, $\mathcal{M} \to 0$ while ε remains non-zero owing to the continued viscous dissipation of the mean flow, and so η_i decreases further.

In cases with increased initial perturbation amplitude, the mixing efficiency evolves differently. As figure 2 shows, the billow structure breaks down into turbulence much earlier in the case of intermediate initial turbulence, without rolling up to the same extent as in the classical Kelvin–Helmholtz case. Correspondingly, the peak in η_i associated with laminar rollup of the billow is absent, and the instantaneous irreversible mixing rate $\mathcal{M}$ is always less than the dissipation ε (yellow curves, figure 5*a*,*b*). The maximum instantaneous mixing efficiency in this case is only approximately 0.35, as in the later turbulent mixing stage of the lower initial amplitude cases. As the initial perturbation amplitude is further increased and the billow structure no longer arises (figure 2*k–o*), the mixing efficiency only achieves

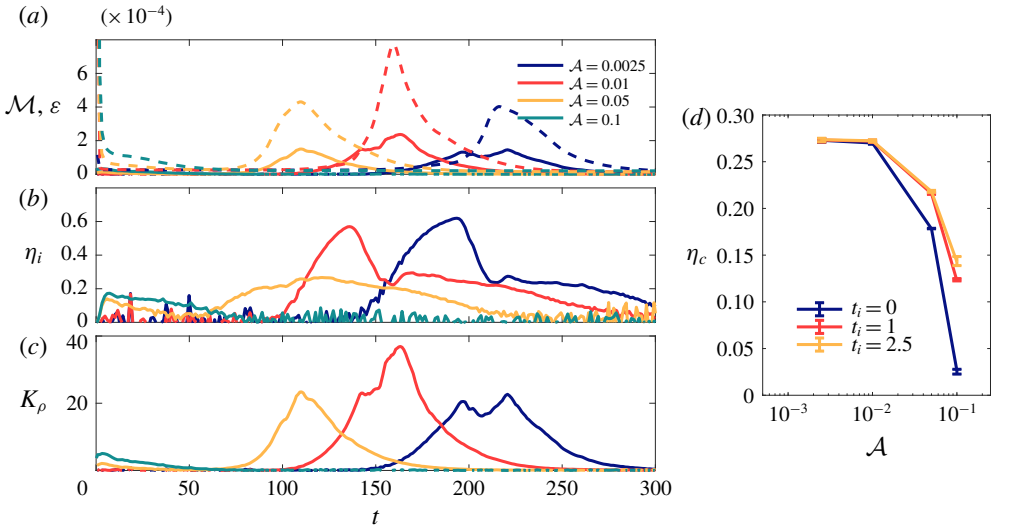

FIGURE 5. (*a*) Time variation of mixing rate, $\mathcal{M}$ (solid), and dissipation, ε (dashed). (*b*) Time variation of instantaneous mixing efficiency, η_i. (*c*) Time variation of turbulent diffusivity, K_ρ. (*d*) Cumulative mixing efficiencies η_c calculated over an entire mixing event for three definitions of the initial time $t = t_i$. The error bars correspond to mixing efficiencies calculated with final times $t_f = \mathcal{D}_p/2$ and $t_f = 2\mathcal{D}_p$.

values of approximately 0.2 during the initial turbulent mixing phase before decaying at later times.

As with the overall changes in potential energy shown in figure 4(*b*,*c*), the peak value of the irreversible mixing rate and the dissipation depend non-monotonically on the initial perturbation amplitude $\mathcal{A}$, with maximum values occurring for $\mathcal{A} = 0.01$. Indeed, the maximum values of both $\mathcal{M}$ and ε are very similar for the laminar ($\mathcal{A} = 0.0025$) and intermediate initial turbulence ($\mathcal{A} = 0.05$) cases, despite the marked differences in both instantaneous and cumulative mixing efficiencies between these two cases. This further emphasizes the role of the highly-efficient pre-turbulent phase associated with billow rollup in determining the overall efficiency of a given mixing event.

As mentioned above, the viscous dissipation can be related to a turbulent density flux via a turbulent diffusivity, K_ρ, expressed in non-dimensional form as (Salehipour & Peltier 2015)

$$K_\rho = \Gamma PrRe_b = \frac{\eta}{1-\eta} PrRe_b. \tag{3.6}$$

All cases show non-zero values of K_ρ at similar times to the changes in $\mathcal{P}_b$ (figure 4*c*). For strongly turbulent initial conditions ($\mathcal{A} = 0.1$, green line in figure 5*c*), there is a small increase in K_ρ at early times and then subsequent decay. For the other three cases, much larger peak values of K_ρ are observed. In all three cases there is a peak in K_ρ associated with peak values of ε (and peak values of Re_b, not shown), with the highest values associated with $\mathcal{A} = 0.01$. However, there is an additional peak for the two cases in which billow formation was not impeded ($\mathcal{A} = 0.0025$ and 0.01). This peak is associated with the onset of turbulence after the rolling up of the billows and the associated peak in mixing efficiency. While the viscous dissipation is still

relatively low at this point in the flow evolution, the highly efficient motions lead to a correspondingly high turbulent buoyancy flux. As a result, K_ρ remains elevated for a much longer period of time in these two cases relative to the case with intermediate initial turbulence ($\mathcal{A} = 0.05$), leading to larger changes overall in the background potential energy.

It is also of interest to describe an individual mixing event in terms of an overall mixing efficiency. A cumulative analogue to (3.5) is defined by integrating $\mathcal{M}$ and ε over the duration of a mixing event (from some initial time t_i to some final time t_f) as

$$\eta_c = \frac{\displaystyle\int_{t_i}^{t_f} \mathcal{M}\,\mathrm{d}t}{\displaystyle\int_{t_i}^{t_f} \mathcal{M}\,\mathrm{d}t + \int_{t_i}^{t_f} \varepsilon\,\mathrm{d}t}. \tag{3.7}$$

As previously discussed, and as figures 4 and 5(*a*) show, there is a strong initial transient adjustment for the cases with intermediate and strong initial turbulence. We find that the computed values of η_c are sensitive to initial integration times t_i within this adjustment period until approximately $t_i \sim 1$; the computed values of η_c do not change significantly for $1 \lesssim t_i \lesssim 10$. We define t_f as the time when $\mathcal{M} = \mathcal{D}_p$ after the mixing event, i.e. the time at which changes in $\mathcal{P}_b$ due to turbulent mixing become comparable to changes due to laminar diffusion in the absence of fluid motion. Note, however, that the computed mixing efficiencies in figure 5(*d*) are not strongly sensitive to the choice of t_f, as long as t_f is after the shear layer begins to relaminarize.

The cases with laminar and weakly turbulent initial conditions (i.e. the cases that follow a typical Kelvin–Helmholtz lifecycle) yield similar cumulative mixing efficiencies, $\eta_c \approx 0.27$, consistent with previous results for simulations of the Kelvin–Helmholtz instability (Peltier & Caulfield 2003; Mashayek *et al.* 2013). As the initial perturbation amplitude is increased and the large peak in η_i associated with laminar rollup is lost, η_c decreases to values of approximately 0.18 for $t_i = 0$ in the intermediate initial turbulence case and 0.025 in the strong initial turbulence case. This decrease in η_c is consistent with the evolution of $\mathcal{M}$, ε and η_i for these amplitudes: irreversible mixing is essentially restricted to times when the flow is turbulent, without the highly efficient pre-turbulent phase associated with billow structure and its large available potential energy. If the initial transient adjustment is excluded ($t_i \geqslant 1$), η_c increases to values of approximately 0.22 and 0.13 for $\mathcal{A} = 0.05$ and $\mathcal{A} = 0.1$, respectively. This increase in η_c is due to the fact that, at very early times, the buoyancy field has not yet adjusted to the initial perturbation, and so the flow is dissipating kinetic energy without mixing buoyancy.

3.2. *Re and Ri_b effects*

Thus far, we have focused on a single choice of the mean flow parameters Re and Ri_b, showing that turbulent initial conditions can impact the evolution of a stratified shear layer. To examine the effect of the mean flow parameters, we now consider simulations at a variety of initial Re and Ri_b for the same four values of $\mathcal{A}$ as considered in the previous section. While the initial perturbation varies slightly depending on the domain size and grid resolution (and, therefore, the corresponding values of k_h in (2.8)), the kinetic energy of the initial perturbation and the turbulence intensity at early times are similar for the same value of $\mathcal{A}$ across the range of simulations presented in table 1.

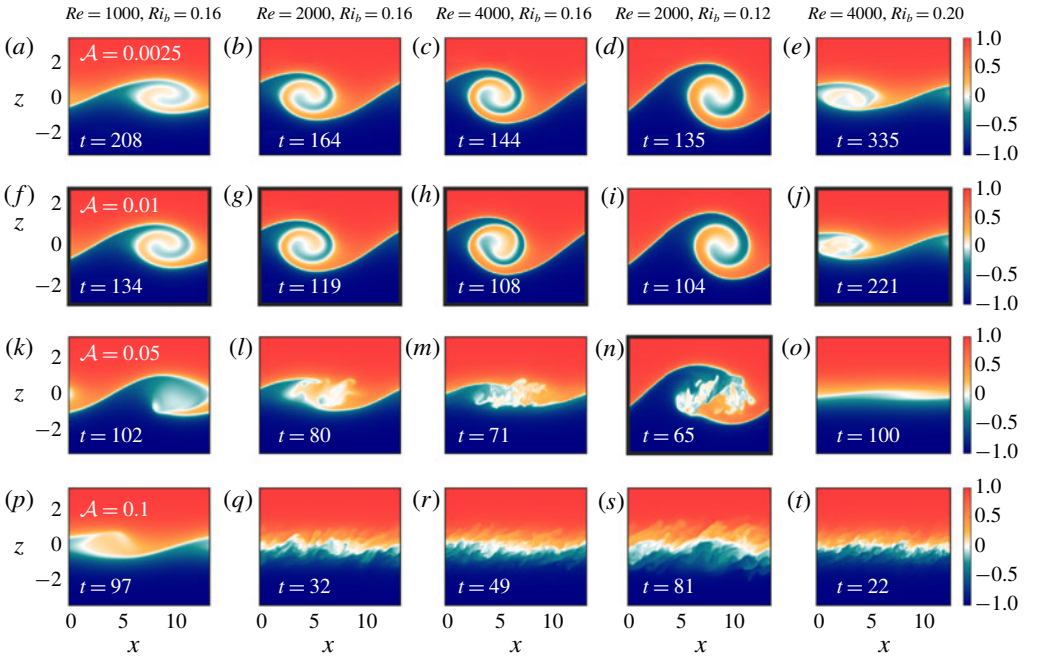

FIGURE 6. Vertical slices of buoyancy through $y = 0$ at various times corresponding to local maxima in $\mathrm{d}\mathcal{P}_a/\mathrm{d}t$. (*a–e*) Laminar initial conditions, $\mathcal{A} = 0.0025$. (*f–j*) Weakly turbulent initial conditions, $\mathcal{A} = 0.01$. (*k–o*) Intermediate turbulence initial conditions, $\mathcal{A} = 0.05$. (*p–t*) Strongly turbulent initial conditions, $\mathcal{A} = 0.1$. Each column corresponds to a different set of simulations with the specified values of Re and Ri_b. Non-dimensional times corresponding to each snapshot are shown in the bottom left corner of each panel. Panels with thick black borders correspond to the value of $\mathcal{A}$ which gives the maximum change in $\mathcal{P}_b$ for a given (Re, Ri_b).

Billow growth due to shear instability is associated with an increase in the available potential energy, $\mathcal{P}_a$. As such, we look for local maxima in the growth rate of available potential energy, $\mathrm{d}\mathcal{P}_a/\mathrm{d}t$, as a means of identifying billows in our suite of simulations. In figure 6 we show the billow structure corresponding to a range of different values of Re and Ri_b and all four values of $\mathcal{A}$, identified using this metric. Note that the specific times corresponding to each snapshot in figure 6 differ depending on both the flow parameters and the initial perturbations. Finite-amplitude billows typically form at later non-dimensional times for higher values of Ri_b (owing to the slower growth rate of the underlying shear instability) and earlier for higher $\mathcal{A}$ (owing to the larger amount of energy initially present in the unstable mode). However, despite the different times highlighted in figure 6, for the cases with clear billow structures the instability follows a similar evolution to that shown in figure 2: the billow becomes unstable to secondary instabilities, the shear becomes fully turbulent, and after some time the turbulence decays and the flow relaminarizes.

Several trends become apparent upon consideration of figure 6. As the initial turbulence amplitude is increased, the development of the shear instability is accordingly modified: the billow structure becomes susceptible to secondary instabilities before fully rolling up, and is prevented from forming for strongly turbulent initial conditions (as shown in the bottom two rows of figure 6). This effect is consistent across the range of Re and Ri_b investigated here.

However, figure 6 also reveals additional effects related to the initial mean flow parameters. The initial Reynolds number of the shear layer appears to affect the sensitivity of the billow structure to modification by pre-existing turbulence. For example, with lower Re and weak initial turbulence (figure 6*f*,*g*,*i*), the resulting billow appears similar to the corresponding structure with laminar initial conditions (figure 6*a*,*b*,*d*). However, as Re is increased, the billow structure is modified; for example, in comparing figures 6(*c*) and 6(*h*), it is clear that the weak initial turbulence has led to asymmetry in the fully developed billow before transition to turbulence. Similarly, Re effects are apparent for the partially developed billows in the case of intermediate initial turbulence. As figure 6(*k*,*l*,*m*) show, when the Reynolds number is increased for the same initial Ri_b, the height of the overturn before breaking down into turbulence is decreased (figure 6*k*,*l*,*m*).

Figure 6 also suggests a strong dependence on the initial bulk Richardson number, Ri_b. The underlying shear instability is less sensitive to the initial turbulence at lower Ri_b, with the billow able to grow to its full vertical height even in the presence of intermediate initial turbulence (figure 6*n*). At the other extreme, for higher Ri_b, intermediate initial turbulence is sufficiently strong to almost prevent billow formation altogether (figure 6*o*).

Figure 7(*a*) shows the overall change in potential energy associated with turbulent irreversible mixing (i.e. $\Delta\mathcal{P}_b - \mathcal{D}_p t$) as a function of the buoyancy Reynolds number at $t = 1$. As observed in figure 4(*b*), there is a non-monotonic variation of total change in $\mathcal{P}_b$ with Re_b. The specific amplitude at which the change in background potential energy peaks depends on Ri_b: while the cases with $Ri_b = 0.16$ (all Re) and the cases with $Ri_b = 0.20$ and $Re = 4000$ exhibit maximum mixing for weak initial turbulence ($\mathcal{A} = 0.01$), the cases with $Ri_b = 0.12$ instead show a clear peak for intermediate initial conditions ($\mathcal{A} = 0.05$). This can be understood by considering again the billow structures in figure 6 – in particular, those cases indicated by thick black borders which correspond to the large changes in $\mathcal{P}_b$. The cases with the most mixing correspond to cases in which there is some initial turbulence present, but not so much as to prevent the billow from growing to its maximum vertical extent. That is, the overturn associated with the underlying instability is able to reach its maximum amplitude, but then transitions to turbulence before completely rolling up. This is especially noticeable in figure 6(*n*). As initial turbulence strength is further increased and the billow structure is not able to form completely (e.g. figure 6*k*,*l*,*m*), the total change in background potential energy correspondingly begins to decrease. Finally, very small changes in background potential energy are observed for strong initial turbulence across the range of Re and Ri_b considered here. The primary exception to this trend is the set of simulations with $Ri_b = 0.20$ and $Re = 2000$ (yellow squares), in which diffusion of the mean flow plays a significant role before the instability can reach finite amplitude for $\mathcal{A} = 0.0025$ and $\mathcal{A} = 0.01$; for these cases, very little mixing occurs due to a lack of vigorous turbulence.

As figure 3(*b*) shows, the stability of the final mean flow (quantified by its Richardson number) also depends on the initial turbulence amplitude. In figure 7(*b*) we show the final value of Ri_{mean} at the centre of the shear layer after the mixing event as a function of total change in $\mathcal{P}_b$. Unsurprisingly, there is a clear relationship between the total amount of irreversible mixing that occurs and the stability of the final flow, with large values of $\Delta\mathcal{P}_b - \mathcal{D}_p t_f$ leading to large values of centreline Ri_{mean}. This trend is observed across the range of simulations presented here, suggesting that the degree to which a billow is able to form determines the degree to which the mean flow is stabilized.

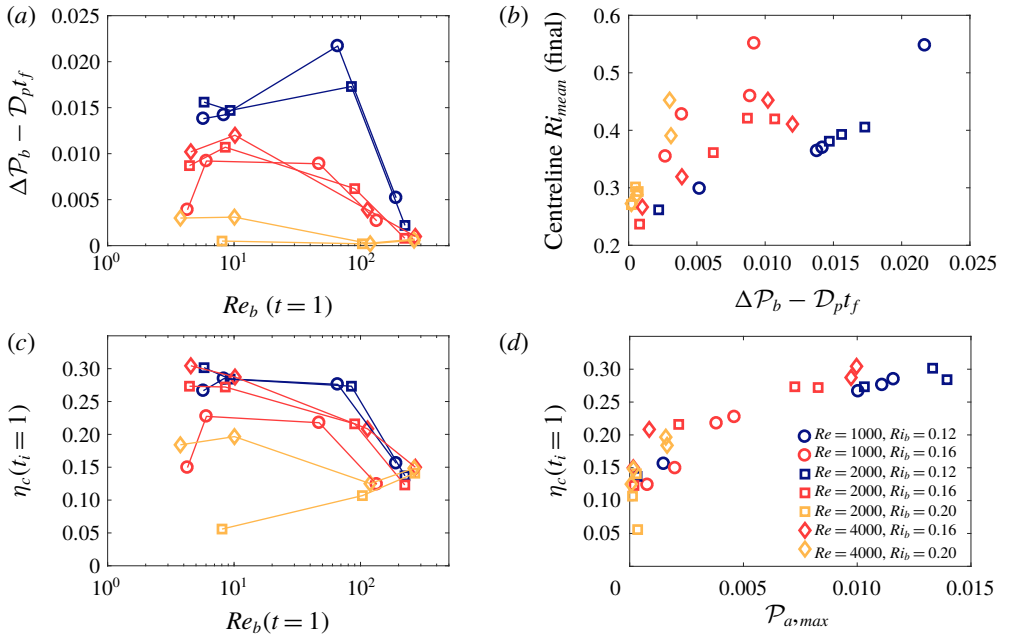

FIGURE 7. (*a*) Overall potential energy change due to irreversible mixing, $\mathcal{P}_b - D_p t$, as a function of Re_b at $t=1$. (*b*) Final centreline Ri_{mean} as a function of $\mathcal{P}_b - D_p t$. (*c*) Cumulative mixing efficiency, η_c, calculated with $t_i=1$ and t_f corresponding to $\mathcal{M}=\mathcal{D}_p$ as a function of Re_b at $t=1$ for a range of different Re and Ri_b. (*d*) Cumulative mixing efficiency, η_c, as a function of the maximum available potential energy during a simulation, $\mathcal{P}_{a,max}$.

Our analysis of the cases with $Re=2000$ and $Ri_b=0.16$ in the previous subsection showed a strong dependence of the cumulative mixing efficiency on the initial perturbation amplitude, with values dropping from $\eta_c \approx 0.27$ to ~ 0.13 as the initial turbulence amplitude is increased (figure 5*d*). This trend is observed for the majority of the cases considered here: individual mixing events are typically more efficient for lower values of $\mathcal{A}$ (figure 7*c*). There is, however, a clear dependence of the mixing efficiency on Ri_b, with much higher cumulative mixing efficiencies (~ 0.3) observed for cases with $Ri_b=0.12$ and lower efficiencies (~ 0.2) for cases with $Ri_b=0.20$. As with the total change in background potential energy, the main exceptions to this trend are cases in which turbulence is weak and, therefore, the diffusion of the background flow plays an important role, namely, the cases with $Re=2000$ and $Ri_b=0.20$, as well as the simulation with $Re=1000$, $Ri_b=0.16$ and $\mathcal{A}=0.002$. The largest overall change in potential energy is not necessarily associated with more efficient mixing (compare figure 7*a*,*c*).

In the previous subsection we showed that a key factor in the efficiency of each mixing event is the degree to which the billow structure, with its associated pre-turbulent peak in instantaneous mixing efficiency, is able to form. Given that the available potential energy $\mathcal{P}_a$ increases with the size of the billow structure (see figure 4*b*), we can use the maximum value of $\mathcal{P}_a$ as a measure of billow development. Figure 7(*d*) shows the cumulative mixing efficiency for each of our simulations as a function of $\mathcal{P}_{a,max}$. There is a clear trend in the efficiency of our simulated mixing events across all values of Re, Ri_b and $\mathcal{A}$ considered here. The cases with strong

initial turbulence in which there is little billow growth (and, accordingly, $\mathcal{P}_{a,max}$ is small) have lower mixing efficiencies of the order of $\eta_c \sim 0.1$–0.15. When larger overturns are able to form and $\mathcal{P}_{a,max}$ increases, the mixing efficiency also increases before eventually reaching a plateau at values of $\eta_c \sim 0.3$, consistent with values from previous studies of the Kelvin–Helmholtz instability in laminar initial flows (Peltier & Caulfield 2003; Mashayek *et al.* 2013; Mashayek & Peltier 2013). This trend even includes the previously identified outlier cases from figure 7(*a*,*c*) in which diffusion of the background flow played a key role and led to smaller maximum billow sizes (figure 6*a*).

Altogether, figures 6 and 7 demonstrate that initial turbulence significantly affects the development of a stratified shear layer, both qualitatively and quantitatively, across a range of Re and Ri_b. These effects are consistent with the results presented in § 3.1, in which initial turbulence may impede the development of the classical Kelvin–Helmholtz billow and, thus, alter the details of the turbulent mixing. For similar initial turbulence amplitudes, these effects appear more pronounced for higher Re and higher Ri_b, suggesting a relationship between the rate of decay of the initial turbulence (which is slower for higher Re) and the growth rate of the underlying instability (which increases for lower Ri_b).

4. Discussion and conclusions

We have shown that the evolution of a dynamically-unstable shear layer may be significantly modified in the presence of pre-existing turbulence, such as may be left behind after a previous mixing event. For small perturbation amplitudes, the classical Kelvin–Helmholtz instability is observed, and the energy evolution and irreversible mixing evolve in a manner similar to that found in previous studies of laminar stratified shear layers. However, as the perturbation amplitude is increased, the development of the canonical billow structure is first suppressed (with transition to turbulence occurring before the billow fully wraps up) and then is eliminated altogether. The latter finding is consistent with the results of Brucker & Sarkar (2007). The total change to the background potential energy (a measure of the total amount of mixing) depends non-monotonically on the initial perturbation amplitude, reaching a maximum when initial turbulence is present but insufficient to prevent the billow from reaching its maximum vertical extent. The instantaneous and cumulative mixing efficiencies depend strongly on whether a Kelvin–Helmholtz billow forms, and decrease as the initial turbulence amplitude is increased. The effects of initial turbulence on the shear layer development are observed for a range of Reynolds and initial bulk Richardson numbers. Modifications of the billow structure associated with the underlying Kelvin–Helmholtz instability are seen for lower initial perturbation amplitudes at higher Re and higher Ri_b.

Overturns associated with shear instability are frequently observed in the atmosphere and oceans. However, the characteristic rolled up structure typically found in numerical simulations of the Kelvin–Helmholtz instability (figure 2*c*) is not often seen (e.g. van Haren *et al.* 2014; Holleman *et al.* 2016). The associated mixing efficiencies are much lower than the values of $\sim 1/3$ found in previous numerical studies, suggesting that the pre-turbulent rollup associated with the formation of the billow may be less important in true geophysical flows (Holleman *et al.* 2016). In their recent review Gregg *et al.* (2018) suggested several possibilities to explain this mismatch between observations and simulations, including differences in parameter values (including Re_b and Pr), non-Boussinesq effects and differences between the

mechanisms driving mixing in highly-idealized simulations versus field measurements. Consistent with this final point, our results (especially our intermediate-amplitude case with its partial billow formation) suggest that pre-existing turbulence may play an important role. In particular, the question of how important the highly efficient pre-turbulent rollup is in shear instabilities may have implications for the parameterization of turbulent mixing events in oceanographic flows (Mashayek *et al.* 2017*a*).

Recent work by Mashayek, Caulfield & Peltier (2017*b*) examining the length scales characterizing turbulent mixing arising from a Kelvin–Helmholtz instability suggests that stratified turbulence is most efficient when distinct overturns exist. Their results suggest that this state of maximum efficiency corresponds to times when the Thorpe length (quantifying the size of the overturns) and the Ozmidov length (quantifying the size at which background stratification is important) are comparable. They argue that these overturns, associated with secondary instabilities, are able to feed off the stored available potential energy in the Kelvin–Helmholtz billow structure and, thereby, efficiently drive the turbulent cascade of energy to small scales. Our results, summarized in figure 7(*d*), are broadly consistent with this conjecture, insofar as we observe less efficient mixing as $\mathcal{A}$ is increased and the primary Kelvin–Helmholtz billow either becomes turbulent before rolling up and storing $\mathcal{P}_a$ ($\mathcal{A}=0.05$) or fails to form at all ($\mathcal{A}=0.1$). Future work will consider how pre-existing turbulence modifies the evolution of these turbulent length scales.

While we have considered a range of Re and Ri_b here, we have kept the Prandtl number fixed at $Pr=1$ for all of the simulations presented above. Previous studies have shown that the transition to turbulence and mixing observed for the Kelvin–Helmholtz instability depends on the value of Pr (e.g. Salehipour *et al.* 2015), and it is likely that the evolution of initially turbulent shear layers will also depend on the Prandtl number.

In this work we consider only a single wavelength of the underlying shear instability, thus, precluding the emergence of the subharmonic pairing instability (significant for lower Re and Ri_b flows). Other secondary instabilities have been shown to emerge earlier in the flow at higher Re and Ri_b (more relevant for the geophysical flows motivating this work), breaking down the primary billow and preventing pairing (Mashayek & Peltier 2012*a*, 2013). We might expect that pre-existing turbulence will likewise impact the emergence of the pairing instability as a result of preventing full billow formation. Future research should focus specifically on how pre-existing turbulence interacts with the pairing mechanism.

Additionally, here we have considered a single vertical profile of initial turbulence, in which the turbulence is concentrated in the centre of the shear layer. However, the distribution of the turbulence within the domain may also play a significant role in the evolution of the shear layer. Thorpe *et al.* (2013) and Li *et al.* (2015) showed that turbulence can act to either stabilize or destabilize a flow, depending on its distribution and the flow parameters. Furthermore, recent numerical studies have shown that external turbulence can modify the evolution of other stratified flows in a variety of ways, including sharpening an existing shear layer at high Pr (Zhou *et al.* 2017) and increasing the rate of decay of a stratified wake (Pal & Sarkar 2015). As such, understanding the impact of different distributions of pre-existing turbulence on the fully nonlinear evolution of stratified shear flows such as the one we have considered here is an important avenue of future research.

Finally, it should be noted that while Kelvin–Helmholtz-type instabilities are a common model for mixing in the ocean interior (Smyth & Moum 2012), they are

not the only mechanism by which mixing may occur in the ocean. Mixing may be associated with a variety of mechanisms, including convective breaking of internal waves (Thorpe 2018), turbulent hydraulic jumps (Thorpe *et al.* 2018) and other types of shear instabilities (Salehipour, Caulfield & Peltier 2016*a*). It is not well known how pre-existing turbulence might modify the mixing characteristics of these different instabilities. Understanding if and how turbulent initial conditions affect the underlying dynamics of these other types of mixing events may therefore also have important ramifications for future parameterizations of mixing in the ocean.

Acknowledgements

This work was funded by the NSF grant OCE1537173. High-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) was provided by NCAR's Computational and Information Systems Laboratory, sponsored by the NSF. We acknowledge the helpful comments of our referees, which added significantly to this work.

Appendix. Simulation resolution

As mentioned in § 2, the grid sizes for our simulations are chosen with the aim of resolving $2.5L_K$. However, as discussed in a recent review of mixing efficiency (Gregg *et al.* 2018), inadequate numerical resolution may impact the computed efficiency. In light of this, it is worth revisiting our choice of grid spacing.

We can quantify the amount of energy at various length scales by considering horizontal spectra of kinetic energy, $\hat{\mathcal{K}}_h$ (computed by taking the Fourier transform of $\boldsymbol{u}\cdot\boldsymbol{u}$ in the x and y directions and then averaging in the z direction). Figure 8(*a–c*) show $\hat{\mathcal{K}}_h$ at several times corresponding to the panels in figure 2. The amount of energy contained at the highest wavenumbers is quite small, suggesting that the flow is well resolved in our simulations. Interestingly, figure 8(*a–c*) is also consistent with the qualitative evolution of the flow shown in figure 2, with elevated energy at small scales when the shear layer is fully turbulent and more energy at the lowest wavenumbers during billow growth.

In figure 8(*d*) we compare the minimum Kolmogorov scale in the simulations from § 3.1 (calculated using the maximum value of horizontally-averaged dissipation at each point in time) to our grid spacing. For our chosen grid spacing, we expect $2.5L_K \geqslant \Delta x$. It is clear that, apart from the very early times in the simulation, this condition is achieved. The early transient for the cases with stronger initial turbulence is unavoidable, given that our initial perturbation includes energy at the smallest scales of the flow; however, by $t \sim 1$ even the case with $\mathcal{A} = 0.1$ is well resolved.

However, the question remains as to whether a resolution based on $2.5L_K$ is sufficient. To examine this, we linearly interpolate the full three-dimensional velocity and buoyancy fields onto a grid with coarser resolution, $(N_x, N_y, N_z) = (384, 128, 501)$, and a grid with finer resolution, $(N_x, N_y, N_z) = (768, 256, 1153)$. We consider the case with $Re = 2000$, $Ri_b = 0.16$ and $\mathcal{A} = 0.01$, as this is the case with the highest peak dissipation (figure 5*a*). We restart these additional simulations at $t \sim 132$, when the primary billow structure has formed and secondary instabilities are beginning to grow.

There is a brief initial transient in the computed mixing efficiency in figure 8(*e*) for the higher-resolution simulation, related to the use of linear interpolation (which affects the probability distribution function of the buoyancy field and,

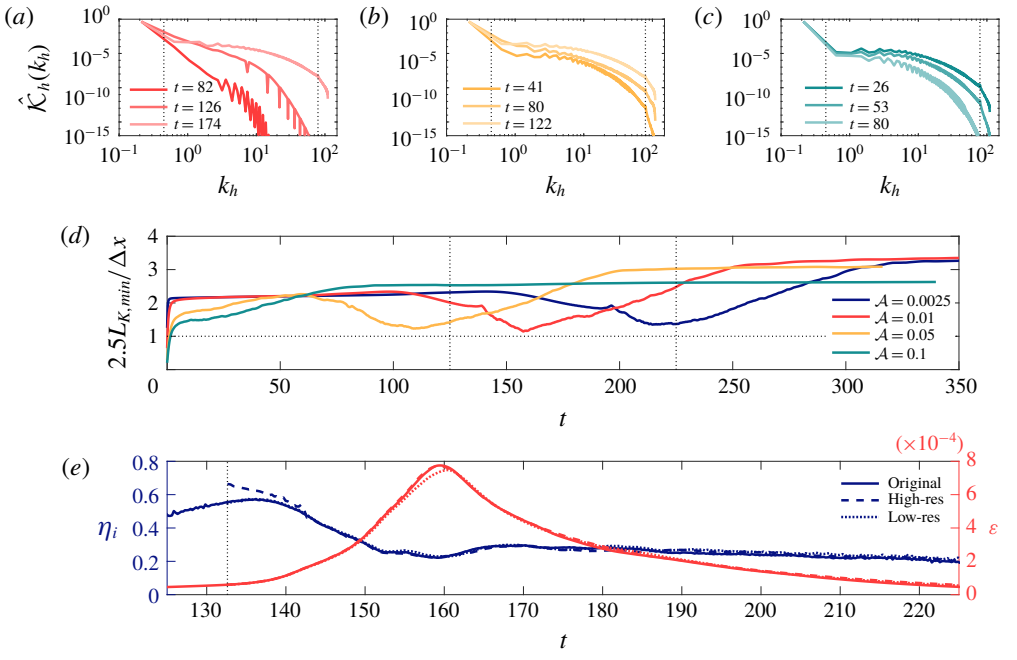

FIGURE 8. (*a–c*) Horizontal spectra of kinetic energy for cases with $Re = 2000$ and $Ri_b = 0.16$ at several times shown in figure 2. Panels (*a*), (*b*) and (*c*) correspond to $\mathcal{A} = 0.01$, 0.05 and 0.1, respectively. Vertical dotted lines denote the Kelvin–Helmholtz wavenumber and the wavenumber associated with the grid spacing. (Note that there is still some energy at higher wavenumbers, corresponding to small three-dimensional scales with high k_x and k_y.) (*d*) $2.5L_{K,min}$ scaled by the grid spacing Δx for the simulations presented in § 3.1. The horizontal dotted line denotes $2.5L_K = \Delta x$ and the vertical dotted lines denote the range of times presented in the bottom panel. (*e*) Instantaneous mixing efficiency (blue lines) and viscous dissipation (red lines) for simulations with three different grid sizes. The simulations correspond to the case with $\mathcal{A} = 0.01$, $Re = 2000$ and $Ri_b = 0.16$, with the higher- and lower-resolution cases restarted from $t \sim 132$ (shown by the vertical dotted line).

consequently, $\mathcal{P}_b$). However, even with this adjustment, there is very little difference at later times in the instantaneous mixing efficiency for either the higher- or lower-resolution cases. In contrast, the viscous dissipation is affected slightly by the grid resolution, with a lower peak value observed for the coarser grid. There is very little difference, however, between ε corresponding to the original simulation and that with the finer grid. The similarity in both η_i and ε between the original and higher-resolution grids therefore shows that the original resolution (determined using $2.5L_K$) is sufficient to quantify turbulent mixing in our simulations.

REFERENCES

BRETHOUWER, G., BILLANT, P., LINDBORG, E. & CHOMAZ, J.-M. 2007 Scaling analysis and simulation of strongly stratified turbulent flows. *J. Fluid Mech.* **585**, 343–368.

BRUCKER, K. A. & SARKAR, S. 2007 Evolution of an initially turbulent stratified shear layer. *Phys. Fluids* **19**, 105105.

DAVIES WYKES, M. S., HUGHES, G. O. & DALZIEL, S. B. 2015 On the meaning of mixing efficiency for buoyancy-driven mixing in stratified turbulent flows. *J. Fluid Mech.* **781**, 261–275.

GREGG, M. C., D'ASARO, E. A., RILEY, J. J. & KUNZE, E. 2018 Mixing efficiency in the ocean. *Annu Rev. Mar. Sci.* **10**, 443–473.

HOLLEMAN, R. C., GEYER, W. R. & RALSTON, D. K. 2016 Stratified turbulence and mixing efficiency in a salt wedge estuary. *J. Phys. Oceanogr.* **46**, 1769–1783.

HOWARD, L. N. 1961 Note on a paper of John W. Miles. *J. Fluid Mech.* **10**, 509–512.

VAN HAREN, H., GOSTIAUX, L., MOROZOV, E. & TARAKANOV, R. 2014 Extremely long Kelvin–Helmholtz billow trains in the Romanche fracture zone. *Geophys. Res. Lett.* **41**, 8445–8451.

IVEY, G. N., WINTERS, K. B. & KOSEFF, J. R. 2008 Density stratification, turbulence, but how much mixing? *Annu. Rev. Fluid Mech.* **40**, 169–184.

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.

LI, L., SMYTH, W. D. & THORPE, S. A. 2015 Destabilization of a stratified shear layer by ambient turbulence. *J. Fluid Mech.* **771**, 1–15.

LORENZ, E. N. 1955 Available potential energy and the maintenance of the general circulation. *Tellus* **7**, 157–167.

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., CAULFIELD, C. P. & PELTIER, W. R. 2017*b* Role of overturns in optimal mixing in stratified mixing layers. *J. Fluid Mech.* **826**, 522–552.

MASHAYEK, A. & PELTIER, W. R. 2012*a* The 'zoo' of secondary instabilities precursory to stratified shear flow transition. Part 1. Shear aligned convection, pairing, and braid instabilities. *J. Fluid Mech.* **708**, 5–44.

MASHAYEK, A. & PELTIER, W. R. 2012*b* The 'zoo' of secondary instabilities precursory to stratified shear flow transition. Part 2. The influence of stratification. *J. Fluid Mech.* **708**, 45–70.

MASHAYEK, A. & PELTIER, W. R. 2013 Shear-induced mixing in geophysical flows: does the route to turbulence matter to its efficiency? *J. Fluid Mech.* **725**, 216–261.

MASHAYEK, A., SALEHIPOUR, H., BOUFFARD, D., CAULFIELD, C. P., FERRARI, R., NIKURASHIN, M., PELTIER, W. R. & SMYTH, W. D. 2017*a* Efficiency of turbulent mixing in the abyssal ocean circulation. *Geophys. Res. Lett.* **44**, 6296–6306.

MILES, J. W. 1961 On the stability of heterogeneous shear flows. *J. Fluid Mech.* **496**, 496–508.

OSBORN, T. R. 1980 Estimates of the local rate of vertical diffusion from dissipation measurements. *J. Phys. Oceanogr.* **10**, 83–89.

PAL, A. & SARKAR, S. 2015 Effect of external turbulence on the evolution of a wake in stratified and unstratified environments. *J. Fluid Mech.* **772**, 361–385.

PELTIER, W. R. & CAULFIELD, C. P. 2003 Mixing efficiency in stratified shear flows. *Annu. Rev. Fluid Mech.* **35**, 135–167.

SALEHIPOUR, H., CAULFIELD, C. P. & PELTIER, W. R. 2016*a* Turbulent mixing due to the Holmboe wave instability at high Reynolds number. *J. Fluid Mech.* **803**, 591–621.

SALEHIPOUR, H. & PELTIER, W. R. 2015 Diapycnal diffusivity, turbulent Prandtl number and mixing efficiency in Boussinesq stratified turbulence. *J. Fluid Mech.* **775**, 464–500.

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.

SALEHIPOUR, H., PELTIER, W. R., WHALEN, C. B. & MACKINNON, J. A. 2016*b* A new characterization of the turbulent diapycnal diffusivities of mass and momentum in the ocean. *Geophys. Res. Lett.* **43**, 3370–3379.

SHIH, L. H., KOSEFF, J. R., IVEY, G. N. & FERZIGER, J. H. 2005 Parameterization of turbulent fluxes and scales using homogeneous sheared stably stratified turbulence simulations. *J. Fluid Mech.* **525**, 193–214.

SMYTH, W. D. 2003 Secondary Kelvin–Helmholtz instability in weakly stratified shear flow. *J. Fluid Mech.* **497**, 67–98.

SMYTH, W. D. & MOUM, J. N. 2012 Ocean mixing by Kelvin–Helmholtz instability. *Oceanography* **25** (2), 140–149.

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.

THORPE, S. A. 1973 Experiments on instability and turbulence in a stratified shear flow. *J. Fluid Mech.* **61** (4), 731–751.

THORPE, S. A. 2018 Models of energy loss from internal waves breaking in the ocean. *J. Fluid Mech.* **836**, 72–116.

THORPE, S. A., MALARKEY, J., VOET, G., ALFORD, M. H., GIRTON, J. B. & CARTER, G. S. 2018 Application of a model of internal hydraulic jumps. *J. Fluid Mech.* **834**, 125–148.

THORPE, S. A., SMYTH, W. D. & LI, L. 2013 The effect of small viscosity and diffusivity on the marginal stability of stably stratified shear flows. *J. Fluid Mech.* **731**, 461–476.

TSENG, Y. & FERZIGER, J. H. 2001 Mixing and available potential energy in stratified flows. *Phys. Fluids* **13** (5), 1281–1293.

WINTERS, K. B., LOMBARD, P. N., RILEY, J. J. & D'ASARO, E. A. 1995 Available potential energy and mixing in density-stratified fluids. *J. Fluid Mech.* **289**, 115–128.

ZHOU, Q., TAYLOR, J. R., CAULFIELD, C. P. & LINDEN, P. F. 2017 Diapycnal mixing in layered stratified plane Couette flow quantified in a tracer-based coordinate. *J. Fluid Mech.* **823**, 198–229.