J. Fluid Mech. (2013), *vol.* 736, *pp.* 570–593.
doi:10.1017/jfm.2013.551

Time-dependent, non-monotonic mixing in stratified turbulent shear flows: implications for oceanographic estimates of buoyancy flux

A. Mashayek[1],†, C. P. Caulfield[2,3] and W. R. Peltier[1]

[1]Department of Physics, University of Toronto, Ontario, M5S 1A7, Canada

[2]BP Institute, University of Cambridge, Madingley Road, Cambridge CB3 0EZ, UK

[3]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

(Received 28 March 2013; revised 7 October 2013; accepted 14 October 2013;
first published online 11 November 2013)

We employ direct numerical simulation to investigate the efficiency of diapycnal mixing by shear-induced turbulence in stably stratified free shear layers for flows with bulk Richardson numbers in the range $0.12 \leqslant Ri_0 \leqslant 0.2$ and Reynolds number $Re = 6000$. We show that mixing efficiency depends non-monotonically upon Ri_0, peaking in the range 0.14–0.16, which coincides closely with the range in which both the buoyancy flux and the dissipation rate are maximum. By detailed analyses of the energetics of flow evolution and the underlying dynamics, we show that the existence of high mixing efficiency in the range $0.14 < Ri_0 < 0.16$ is due to the emergence of a large number of small-scale instabilities which do not exist at lower Richardson numbers and are stabilized at high Richardson numbers. As discussed in Mashayek & Peltier (*J. Fluid Mech.*, vol. 725, 2013, pp. 216–261), the existence of such a well-populated 'zoo' of secondary instabilities at intermediate Richardson numbers and the subsequent high mixing efficiency is realized only if the Reynolds number is higher than a critical value which is generally higher than that achievable in laboratory settings, as well as that which was achieved in the majority of previous numerical studies of shear-induced stratified turbulence. We furthermore show that the primary assumptions upon which the widely employed Osborn (*J. Phys. Oceanogr.* vol. 10, 1980, pp. 83–89) formula is based, as well as its counterparts and derivatives, which relate buoyancy flux to dissipation rate through a (constant) flux coefficient (Γ), fail at higher Richardson numbers provided that the Reynolds number is sufficiently high. Specifically, we show that the assumptions of fully developed, stationary, and isotropic turbulence all break down at high Richardson numbers. We show that the breakdown of these assumptions occurs most prominently at Richardson numbers above that corresponding to the maximum mixing efficiency, a fact that highlights the importance of the non-monotonicity of the dependence of mixing efficiency upon Richardson number, which we establish to be characteristic of stratified shear-induced turbulence. At high Ri_0, the lifecycle of the turbulence is composed of a rapidly growing phase followed by a phase of rapid decay. Throughout the lifecycle, there is considerable exchange of energy between the small-scale turbulence and larger coherent structures which survive the various stages of flow evolution. Since shear instability is one of the most prominent mechanisms for turbulent dissipation of energy at scales below hundreds of metres and at various depths of the ocean, our results have important

† Email address for correspondence: amashaye@atmosp.physics.utoronto.ca

implications for the inference of turbulent diffusivities on the basis of microstructure measurements in the oceanic environment.

Key words: shear layer turbulence, stratified flows, turbulent mixing

1. Introduction

Quantification of the diapycnal turbulent flux of tracers and momentum in an environment that is stably stratified in density has come to be intensively investigated over the past several decades in both the atmospheric and oceanographic literature. In the oceanographic context especially, the vertical flux of buoyancy remains a primary focus since such diapycnal mixing is a primary mechanism which facilitates upwelling of abyssal waters to the surface, thereby enabling closure of a meridional overturning circulation (see e.g. Wunsch & Ferrari 2004). A key quantity of interest for the quantification of diapycnal mixing is the percentage of the kinetic energy available to turbulence which contributes irreversibly to the vertical buoyancy flux. This quantity is often referred to as the mixing efficiency, or flux Richardson number, Ri_f, and is generally assumed to be in the range $Ri_f \sim 0.15$–0.2 despite the growing body of literature demonstrating that it varies over a wide range of values depending on the nature of the dynamical processes involved (Caulfield & Peltier 2000; Smyth, Moum & Caldwell 2001; Peltier & Caulfield 2003; Ivey, Winters & Koseff 2008). The most widely used relation (due to Osborn 1980) in practical oceanography to estimate a vertical effective diapycnal diffusivity from local measurements of the rate of dissipation of kinetic energy ($\mathscr{E}$) and the buoyancy frequency profile (N) is

$$\kappa = \frac{R_f}{1 - R_f}\frac{\mathscr{E}}{N^2} = \Gamma\frac{\mathscr{E}}{N^2}, \tag{1.1}$$

which is employed along with the flux coefficient itself (often inaccurately referred to as a mixing efficiency) which is assumed to have the value $\Gamma \approx 0.2$. Central to the application of this model are three key assumptions, which we refer to as follows: assumption 'F', that the vertical diffusivity is dominated by fully developed turbulence; assumption 'S', that the turbulence exhibits a quasi-steady balance between production, dissipation and diapycnal mixing; and assumption 'I', that the turbulence is isotropic. At least in the context under consideration of transition induced by shear instability, the validity of each of these assumptions is strongly undermined by the analyses to follow.

The specific numerical value for Γ proposed by Osborn (1980) was strongly influenced by experimental measurement (Britter 1974; Koop 1976), and there has been a very wide range of experimental studies of mixing efficiency: see the reviews of Linden (1979), Fernando (1991) and Ivey *et al.* (2008), and, for example Ivey & Imberger (1991), Park, Whitehead & Gnanadeskian (1994), Strang & Fernando (2001), Rehmann & Koseff (2004), Prastowo *et al.* (2008), Stretch *et al.* (2010) and Oglethorpe, Caulfield & Woods (2013). The evidence points strongly towards the flux coefficient Γ not being constant, but rather a non-monotonic function of the overall stratification, as quantified by an appropriately chosen 'Richardson number', quantifying the relative significance of potential energy and kinetic energy in the flow. Of course, the choice of this Richardson number depends on the central characteristics of the flow of interest since, as noted by Fernando (1991), the physical mechanisms

of entrainment and mixing across sheared and unsheared density interfaces are qualitatively different.

Such non-monotonicity is of interest since, as originally postulated by Phillips (1972), it may lead to the development of the density staircase structures required to explain at least some observations on layering in nature. The Phillips mechanism was described in terms of a Richardson number which included the influence of both background shear and stratification, yet some subsequent research has suggested that the key aspect of this layer formation mechanism does not rely on the flow being subject to vertical shear but rather that the density, or equivalently the buoyancy flux, is a non-monotonic function of the local density gradient, normalized by some measure of the local turbulent kinetic energy (Posmentier 1977; Linden 1980; Park *et al.* 1994; Holford & Linden 1999; Oglethorpe *et al.* 2013). Nevertheless, as observed by Strang & Fernando (2001), for example, non-monotonic variation of mixing efficiency with Richardson number does occur in stratified shear flows.

However, the basic observation that Γ depends non-monotonically on Richardson number (in the stratified shear layer context) is not the only way in which it varies with the external, larger-scale properties of the flow. There is also increasing evidence that there is a significant dependence of Γ on some appropriate measure of the intensity of the turbulence, typically quantified by values of the 'buoyancy' Reynolds number Re_b, defined as

$$Re_b = \frac{\mathscr{E}}{\nu N^2}. \tag{1.2}$$

See for example Ivey & Imberger (1991), Shih *et al.* (2005), Wells, Cenedese & Caulfield (2010) and Lozovatsky & Fernando (2012). This point is particularly significant, as there is a marked difference between the accessible Reynolds numbers in the laboratory and those which are actually characteristic of nature. To quote Ivey *et al.* (2008), 'Both laboratory and DNS work indicate that at these extremes, when either $\mathscr{E}/\nu N^2 \sim O(1)$ or $\mathscr{E}/\nu N^2 \sim O(10^5)$, the mixing efficiency $R_f \to 0$ and the use of large $R_f \simeq 0.2$ in field situations in these limits cannot be justified. This is not simply a matter of curiosity. There is a fundamental inconsistency between the results from the laboratory and DNS experiments and the inference of diffusivity from microstructure in the field that remains unresolved.'

This inconsistency must at least in part be due to an incomplete understanding of the dominant physical mechanisms of mixing at higher Reynolds number. This issue is particularly pressing for the canonical flow of a stratified, shear-driven, mixing layer, as the classical 'tilted tank' experimental geometry originally used by Osborne Reynolds and subsequently popularized by Thorpe (1968) (see also Thorpe 1973) is typically strongly constrained in Reynolds number, as well as being inevitably contaminated by flow acceleration, initialization difficulties (see for example Patterson *et al.* 2006) and end effects, which make it extremely difficult to make quantitative inferences in the laboratory concerning the mixing actually associated with the transition of Kelvin–Helmholtz (KH) billows to turbulence at sufficiently high Re. As noted by Smyth *et al.* (2001), at lower Re a substantial amount of 'mixing' leading to irreversible increases in the potential energy of the system can occur in the 'preturbulent' phase of flow evolution, before the dissipation rate becomes elevated substantially above laminar values. It is an open question as to whether that picture is relevant at higher Reynolds number, as increased intensity in the turbulence, and hence an elevated dissipation rate $\mathscr{E}$ will affect the irreversible mixing in a non-trivial way. However, very recently, Mashayek & Peltier (2013) (hereafter MP13)

described numerical simulations of the evolution of the stratified mixing layer at substantially higher Reynolds numbers $Re = U_0 d/\nu$ (where $2U_0$ is the total velocity difference across the layer of characteristic total thickness $2d$, and ν is the fluid's kinematic viscosity) than are accessible in the laboratory. Their analyses showed that as $Re \to 10^4$, the mixing efficiency associated with the full turbulent lifecycle of an initial shear instability can approach values as high as 0.45 and that the efficiency associated with the fully developed turbulent phase of the flow approaches $\sim 1/3$.

Intriguingly, and perhaps suggestive of a generic behaviour of stratified mixing events at sufficiently high Reynolds number, this result is consistent with the prediction of Caulfield, Tang & Plasting (2004). They identified a rigorous upper bound as a function of Reynolds number on the long-time average of the vertical buoyancy flux in stratified plane Couette flow between two horizontal plates, separated by a distance d, which maintain the boundary fluid with constant background streamwise velocity difference ΔU and density difference $\Delta\rho$. As $Re \propto \Delta U d/\nu \to \infty$, they found that the notional flow associated with this upper bound on the buoyancy flux had an implied mixing efficiency $Ri_f \to 1/3$ as $Re \to \infty$. Furthermore, Tang, Caulfield & Kerswell (2009) have identified further properties of this notional flow, namely that the upper bound on the long-time average of the buoyancy flux implies a notional flow which has $Ri_0 = 1/6$, where

$$Ri_0 = g\Delta\rho d/(\rho_0 \Delta U^2) \tag{1.3}$$

is a 'bulk' Richardson number with g the acceleration due to gravity and ρ_0 a reference density. However, there is no guarantee that this bound is attained, or indeed that this notional flow actually exists, and this approach reveals nothing about the dynamical behaviour of the turbulence which actually leads to this mixing.

Faced with these unresolved issues in the literature, this paper has four central aims. Firstly, we want to gain insight into the actual physical mechanisms by which mixing occurs in sufficiently high Re stratified mixing layers to allow for the proliferation of the 'zoo' of secondary instabilities, discussed at length in MP13. Specifically, we extend the analyses of MP13 herein in order to evaluate mixing efficiency for Richardson numbers in the range $0.12 \leqslant Ri_0 \leqslant 0.2$. This range is also relevant in the oceanographic context in which shear instability is often observed to emerge locally once the gradient Richardson number falls below the critical value of $1/4$ (Howard 1961; Miles 1961), due, for example, to transient increase in the background shear. Our second, closely related, aim is to determine whether the existence of different physical mixing mechanisms that become accessible to the flow as the initial Richardson number is varied is the ultimate cause of the non-monotonic dependence of mixing efficiency upon this parameter. Given that the mixing efficiency will be found to be a non-monotonic function of the Richardson number, our third aim is to establish whether the associated maximum does occur near $Ri_0 = 0.16$. Our final aim is to use this understanding to critique the classic model for Γ of Osborn (1980), assessing whether for stratified mixing layers the three central assumptions 'F', 'S' and 'I' are ever appropriate descriptions of the turbulence and attendant irreversible mixing. To achieve these aims, the rest of this paper is organized as follows. After defining various quantities and describing our numerical model in § 2, in § 3 we analyse the properties of our simulations, focusing in particular on the varying phenomenology of mixing as Ri_0 increases. Finally, in § 4 we discuss the significance of our results for the modelling of mixing in stratified mixing layers, confirming that the properties of the turbulence bear little resemblance to those assumed in the classical Osborn (1980) model.

Re	Ri_0	Pr	L_x	L_y	L_z	N_x	N_y	N_z
6000	0.12	1	14.28	3 and 6	30	1280	256 and 512	1216
6000	0.14	1	14.28	3 and 6	30	1280	256 and 512	1216
6000	0.16	1	14.28	3	30	1280	256	1216
6000	0.18	1	14.28	3	30	1280	256	1216
6000	0.20	1	14.28	3	30	1280	256	1216

TABLE 1. Details of three-dimensional numerical experiments.

2. Theoretical preliminaries

We consider temporal evolution of a horizontally periodic stably stratified shear layer. Closely following the formulation and analyses of MP13, the initial background profiles of the velocity and density fields are assumed to be

$$\bar{U}(z) = U_0 \tanh\left(\frac{z}{d}\right), \quad \bar{\rho}(z) = \rho_0 - \Delta\rho \tanh\left(\frac{z}{d}\right), \tag{2.1}$$

where U_0 is reference velocity and where x, y and z denote the streamwise, spanwise and vertical directions respectively. The dimensionless equations of motion, continuity and energy conservation in the Boussinesq approximation are

$$\frac{\mathrm{D}u_i}{\mathrm{D}t} = -\frac{\partial p}{\partial x_i} - Ri_0\rho\delta_{i3} + \frac{1}{Re}\frac{\partial^2 u_i}{\partial x_j^2}, \tag{2.2}$$

$$\frac{\partial u_i}{\partial x_i} = 0, \tag{2.3}$$

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = \frac{1}{Re\ Pr}\frac{\partial^2 \rho}{\partial x_j^2}, \tag{2.4}$$

where $i, j = 1, 2, 3$. We numerically integrate these governing equations using the same methodology as that described in MP13. In short, a pseudo-spectral direct numerical simulation (DNS) methodology is employed which ensures conservation of mass, momentum and energy. The computational domain is assumed to be periodic in both horizontal directions, and thus is treated using a spectral representation, while a second-order finite volume discretization is employed in the vertical. For all of the cases to be discussed herein we set the Reynolds number to 6000, which is sufficiently high to ensure emergence of all the high Reynolds number secondary instabilities introduced in Mashayek & Peltier (2012*a*,*b*) as well as the consequent prohibition of the vortex pairing mechanism as described in MP13. We may therefore, without loss of generality, set the streamwise scale of our numerical domain to a single wavelength ($L_x = 14.28$) of the most unstable mode of linear shear instability (e.g. Hazel 1972). We also set the Prandtl number $Pr = \nu/\kappa = 1$, where κ is the thermal diffusivity. Table 1 provides the required information concerning the three-dimensional numerical experiments to be analysed in the subsequent sections. The case with $Ri_0 = 0.12$ is similar to one of the cases discussed in MP13. For cases with $Ri_0 = 0.12$ and 0.14, simulations were carried out for two spanwise domain extents to ensure that the spanwise domain was sufficiently broad to house all of the fastest growing modes of secondary instability. In all cases the vertical extent was set to $L_z = 30$, with no normal flow and no tangential stress boundary conditions imposed, so that the presence of these boundaries would have no discernible effect on flow evolution.

For our simulations, the effective (peak) buoyancy Reynolds number Re_b as defined in (1.2) (with buoyancy frequency defined as $N = \sqrt{-(g/\rho_0)\ \partial_z \rho}$) is in the range $100 \lesssim Re_b \lesssim 300$, which partially overlaps with the range in which energetic stratified turbulence is known to exist in the ocean (Smyth & Moum 2000).

To facilitate detailed analysis of the flow field, we calculate a hierarchy of velocity fields through spatial averaging, namely

$$\bar{U}(z,t) = \langle u \rangle_{xy}, \tag{2.5}$$

$$(u,v,w)(\boldsymbol{x},t) = (\bar{U} + u_{2D} + u_{3D}, v_{3D}, w_{2D} + w_{3D}), \tag{2.6}$$

$$(u_{2D}, 0, w_{2D})(x,z,t) = \langle (u - \bar{U}, v, w) \rangle_y, \tag{2.7}$$

$$(u_{3D}, v_{3D}, w_{3D}(x,y,z,t) = (u - \bar{U} - u_{2D}, v, w - w_{2D}), \tag{2.8}$$

where $\langle . \rangle_p$ denotes averaging in the p direction. Physically, the non-parallel two-dimensional velocity field $(u_{2D}, 0, w_{2D})$ is associated initially with the primary KH billow. By substituting these expansions into the momentum equation, and by calculating the inner product of the resulting equation with the total velocity vector, we obtain an equation for kinetic energy, which can be decomposed into mean and turbulent components. The derivation is straightforward and follows the standard procedure of derivation of the turbulent energy balances. For further detail see Klaassen & Peltier (1985). The resulting evolution equation for the total kinetic energy is

$$\sigma = \frac{1}{2\mathscr{K}} \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{K} = -\mathscr{H} - \mathscr{D}, \tag{2.9}$$

where σ is the time rate of change of the total kinetic energy $\mathscr{K}$. The buoyancy flux, $\mathscr{H}$, and the positive definite viscous dissipation rate, $\mathscr{D}$ (effectively the non-dimensional, spatially averaged representation of the dissipation rate $\mathscr{E}$), are defined as

$$\mathscr{H} = \frac{Ri_0}{2\mathscr{K}} \langle \rho w \rangle_{xyz}, \quad \mathscr{D} = \frac{1}{2\mathscr{K} Re} \langle 2E_{ij}E_{ij} \rangle_{xyz} = \frac{1}{2\mathscr{K}} \langle \mathscr{E} \rangle_{xyz}, \tag{2.10}$$

where $\langle . \rangle_{x_i}$ denotes averaging in the direction x_i and $\boldsymbol{E}$ is the strain rate tensor for the horizontally averaged velocity defined as

$$E_{ij} = \frac{1}{2} \left(\frac{\partial \bar{U}_i}{\partial x_j} + \frac{\partial \bar{U}_j}{\partial x_i} \right) \quad \text{for } i = 1, 2. \tag{2.11}$$

The evolution equation for the kinetic energy associated with the three-dimensional perturbation velocity field is found to have the form (see Appendix for derivation):

$$\sigma_{3D} = \frac{1}{2\mathscr{K}_{3D}} \frac{\mathrm{d}}{\mathrm{d}t} \mathscr{K}_{3D} = \mathscr{R}_{3D} + \mathscr{S}h_{3D} + \mathscr{A} - \mathscr{H}_{3D} - \mathscr{D}_{3D}, \tag{2.12}$$

where $\mathscr{K}_{3D} = 1/2(u_{3D}^2 + v_{3D}^2 + w_{3D}^2)$. Here $\mathscr{R}_{3D}$ represents extraction of energy from the background mean shear by means of Reynolds stresses and is defined as

$$\mathscr{R}_{3D} = -\frac{1}{2\mathscr{K}_{3D}} \left\langle u_{3D} w_{3D} \frac{\partial \bar{U}}{\partial z} \right\rangle_{xyz}, \tag{2.13}$$

$\mathscr{S}h_{3D}$ represents shear extraction of energy from the background large-scale inherently two-dimensional structures such as the KH billow (and so would vanish in an isotropic

flow) and is given by

$$\mathscr{S}h_{3D} = -\frac{1}{2\mathscr{K}_{3D}}\left\langle u_{3D}w_{3D}\left(\frac{\partial w_{2D}}{\partial x}+\frac{\partial u_{2D}}{\partial z}\right)\right\rangle_{xyz}, \tag{2.14}$$

$\mathscr{A}$ is a further measure of anisotropy, representing the stretching deformation associated with the three-dimensional perturbation velocity, and is defined as

$$\mathscr{A} = -\frac{1}{4\mathscr{K}_{3D}}\left\langle (u_{3D}^2 - w_{3D}^2)\left(\frac{\partial u_{2D}}{\partial x}-\frac{\partial w_{2D}}{\partial z}\right)\right\rangle_{xyz}, \tag{2.15}$$

and $\mathscr{H}_{3D}$ and $\mathscr{D}_{3D}$ are the buoyancy flux and viscous dissipation terms associated with three-dimensional perturbations and are defined, respectively, by

$$\mathscr{H}_{3D} = \frac{Ri_0}{2\mathscr{K}_{3D}}\langle \rho_{3D}w_{3D}\rangle_{xyz}, \tag{2.16}$$

$$\mathscr{D}_{3D} = \frac{1}{2Re\mathscr{K}_{3D}}\langle 2e_{ij}e_{ij}\rangle_{xyz}, \tag{2.17}$$

where $\boldsymbol{e}$ is the fluctuation strain rate tensor defined as

$$e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right). \tag{2.18}$$

To quantify diapycnal mixing, a cumulative mixing efficiency (η_c) is defined following Caulfield & Peltier (2000) as

$$\eta_c = \frac{\displaystyle\int_{t_1}^{t_f}\mathscr{M}(t')\,\mathrm{d}t'}{\displaystyle\int_{t_1}^{t_f}\mathscr{M}(t')\,\mathrm{d}t' - \int_{t_1}^{t_f}\mathscr{D}(t')\,\mathrm{d}t'}, \tag{2.19}$$

where $\mathscr{M}$ represents irreversible diapycnal mixing and t_f marks the end of one lifecycle of the shear instability. $\mathscr{M}$ is determined from the time rate of change in the background potential energy, where the background potential energy is calculated using a sorting methodology similar to that employed by Winters *et al.* (1995), Caulfield & Peltier (2000) and Peltier & Caulfield (2003). For $t_1 = 0$, (2.19) gives an efficiency for the whole lifecycle, whereas if t_1 is set to the time of onset of fully developed turbulence, (2.19) results in an efficiency corresponding to the fully turbulent phase of the flow. As already noted in the Introduction, 'preturbulent' mixing, when $\mathscr{D}$ is small, may be very 'efficient' due to the denominator in expressions such as (2.19) being relatively small, but of central interest to us is in which specific period during the evolution of the flow does the dominant contribution to mixing occur, and what is the mixing efficiency that is achieved during this period.

At this point we can revisit the assumptions behind the Osborn (1980) method of calculating an effective diffusivity through neglect of various terms in (2.12). These three previously mentioned assumptions consist of assumption '*F*' of fully developed turbulence, assumption '*S*' of stationary turbulence and assumption '*I*' of isotropic turbulence. Beginning from (2.12), assumption '*S*' corresponds to the neglect of the time derivative on the left-hand side while assumption '*I*' requires the elimination of the explicitly anisotropic terms $\mathscr{S}h_{3D}$ and $\mathscr{A}$, resulting in a three-term balance between shear production ($\mathscr{R}_{3D}$), turbulent buoyancy flux ($\mathscr{H}_{3D}$) and dissipation ($\mathscr{D}_{3D}$)

in the form

$$\mathscr{R}_{3D} = \mathscr{H}_{3D} + \mathscr{D}_{3D}. \tag{2.20}$$

By defining a turbulent flux Richardson number as

$$Ri_f = \mathscr{H}_{3D}/\mathscr{R}_{3D}, \tag{2.21}$$

and by relating the vertical turbulent buoyancy flux $\langle \rho w_{3D} \rangle$ to a mean background gradient $\mathrm{d}\bar{\rho}/\mathrm{d}z$ through the introduction of an effective turbulent diffusivity, rearranging (2.20) leads directly to (1.1). As one of the primary aims of this paper we seek to investigate the validity of the assumptions leading to elimination of the additional terms in (2.12) in the range of bulk Richardson numbers which is believed to be most relevant to the understanding of energetically turbulent oceanic shear zones.

3. Analyses

Figure 1 illustrates the turbulence transition phase of flow evolution for each of the simulations considered herein by providing contour plots of both density and rate of dissipation of kinetic energy. The contour plots have been made for the mid-plane in the spanwise direction. Although these figures are instructive for determining the form and localization of the structures leading to dissipation and mixing, it is also useful to consider visualizations of the vorticity field, due to the key role of vorticity dynamics in the finite amplitude form of the various instabilities, as well as the ensuing turbulent mixing. Therefore, we also plot iso-surfaces of vorticity (both streamwise and spanwise) for the same cases in figure 2, and interpret the flow evolution by consideration of both sets of figures together. Indeed, in the supplementary section available at http://dx.doi.org/10.1017/jfm.2013.551 which accompanies this paper we have provided a series of animations which the interested reader may use to gain further insight into the shear-induced turbulent flows that are the subject of study in this paper. Taken in combination, these figures and the supplementary animations illustrate clearly a central challenge of interpretation of stratified mixing layers at sufficiently high Reynolds numbers. As the bulk Richardson number increases, the amplitude of the primary KH billow decreases, with several complicated consequences. The types of secondary instabilities which develop on the evolving billow clearly change, and it is by no means obvious whether the mixing is dominated by the effects of the primary larger-scale overturning, by secondary instabilities which develop as the primary billow grows, or by secondary (or indeed tertiary) instabilities 'catalysed' by the non-parallel billow when it is essentially at its saturated amplitude, as was argued to be the case for the low Re, low Ri_0 flows considered by Caulfield & Peltier (2000).

Focusing on the first two cases with $Ri_0 = 0.14$ and 0.16, the evolution of the flow is similar to that of the case with $Ri_0 = 0.12$ discussed at length in MP13, in that following saturation of the primary KH billow, a number of secondary instabilities emerge and facilitate turbulent breakdown of the billow. The most prominent of these secondary instabilities are: (i) the secondary convective instability of the cores (SCI), illustrated by purple and green streamwise vortex tubes in figure 2; (ii) the secondary shear instability (SSI) which forms on the braid (the strained vorticity layer connecting adjacent vortices in a train of KH billows), and is visualized by the grey iso-surface of spanwise vorticity in figure 2; (iii) the stagnation point instability (SPI) which is also visualized by the grey spanwise vorticity surfaces as the largest vortex on the braid at the leftmost edge of the domain in figure 2(*a*). (For more information on the SCI refer to Klaassen & Peltier (1985) and for the other instabilities refer to

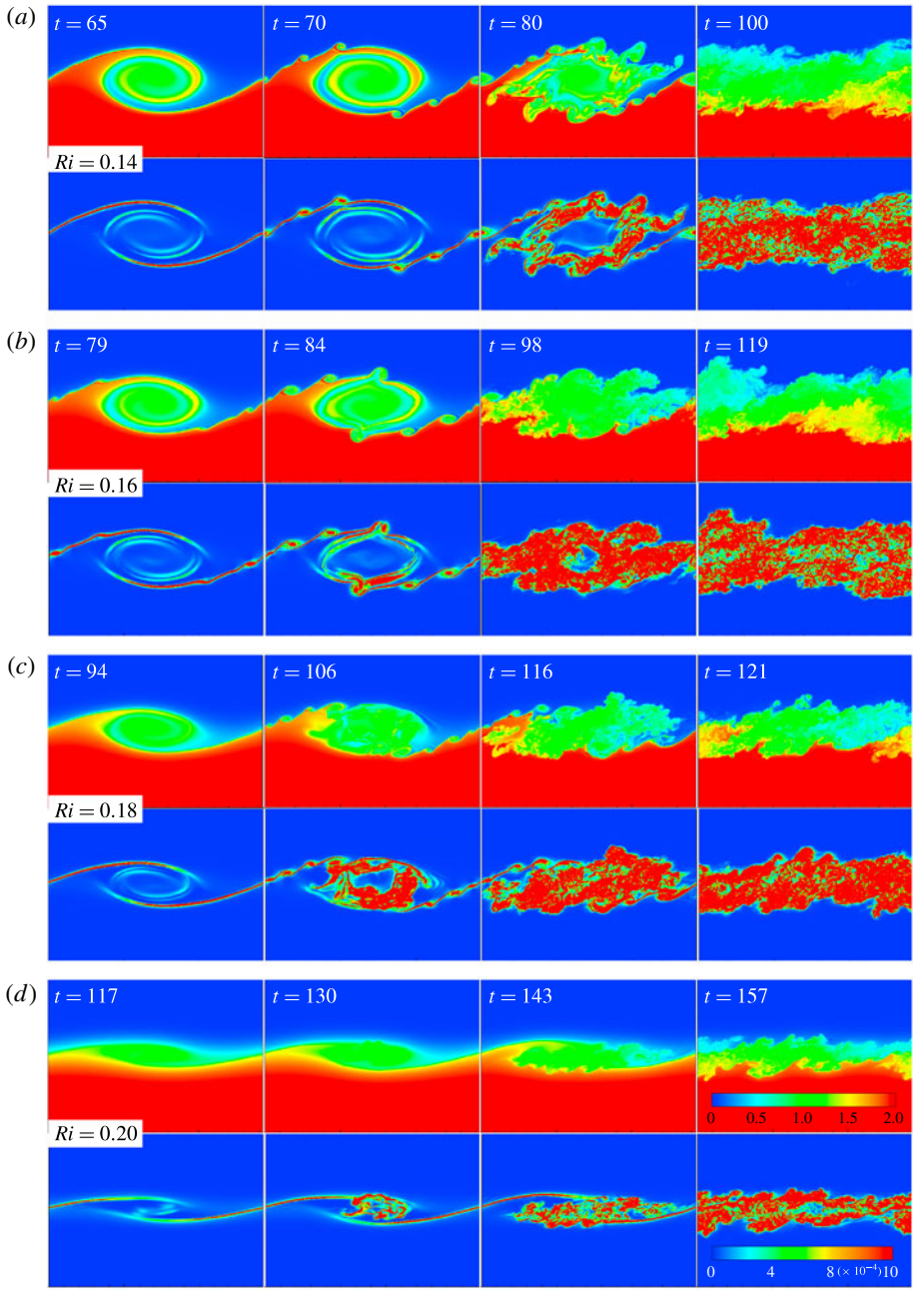

FIGURE 1. Time evolution of flow during the transition phase. For each Ri_0 value, the upper panels show contours of density and the lower panels show contours of the rate of dissipation of kinetic energy. The colour maps for both contours are linear. (*a*) $Ri = 0.14$; (*b*) $Ri = 0.16$; (*c*) $Ri = 0.18$; (*d*) $Ri = 0.20$.

MP13.) Comparing the $Ri_0 = 0.14$ and 0.16 cases in figures 1 and 2 to the results for $Ri_0 = 0.18$ and 0.2 shows that the increase in the bulk Richardson number acts so as to suppress the braid-localized instabilities (SSI and SPI), leaving the SCI as the primary

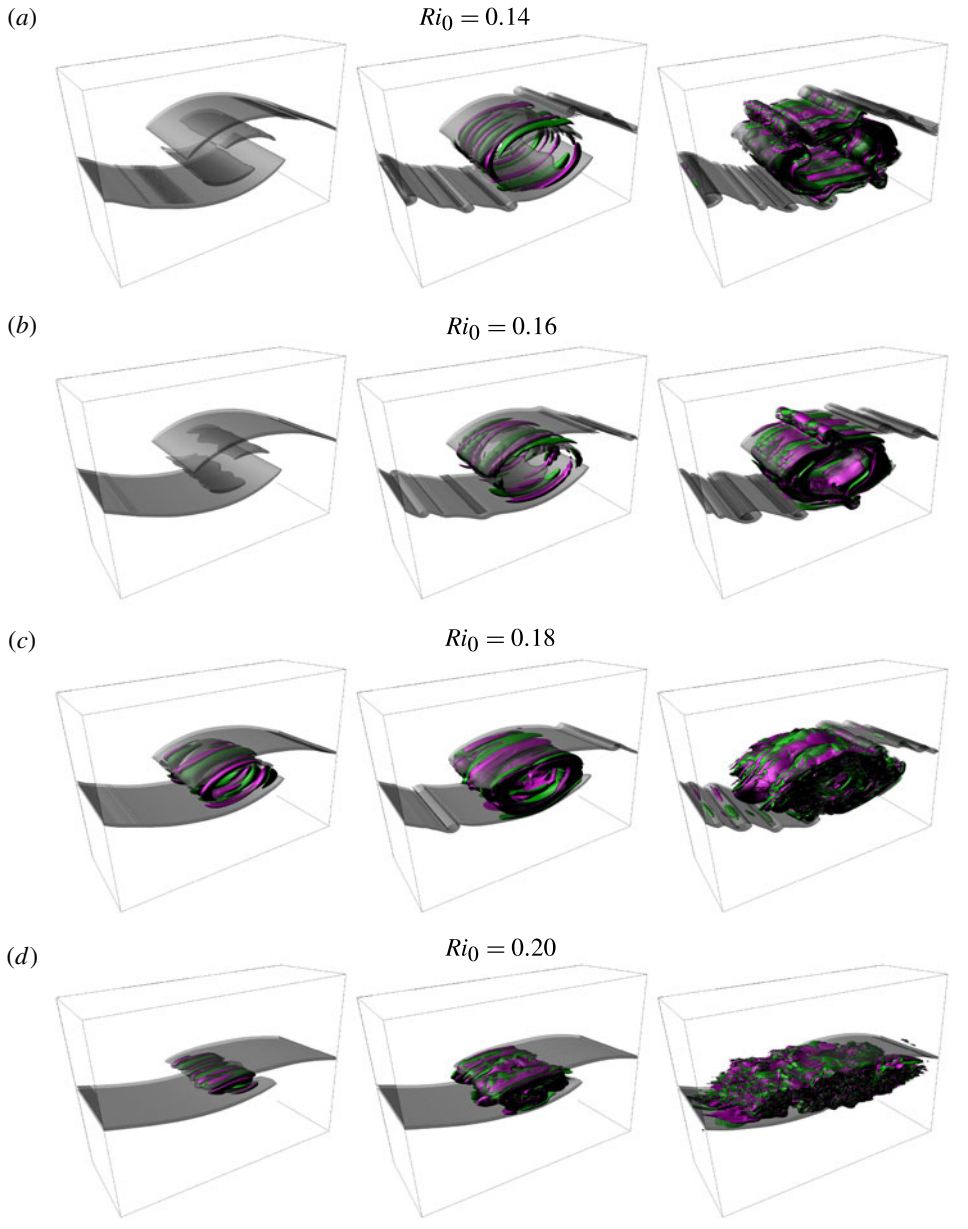

FIGURE 2. Time evolution of flow during the transition phase. The plots show contours of vorticity: grey surfaces represent spanwise vorticity (i.e. in the direction normal to the plane of the original two-dimensional Kelvin–Helmholtz billow) and the green and purple surfaces represent positive and negative streamwise (i.e. in the same direction as the background shear) vorticity, respectively. The structures on the braid (shown in grey) represent the secondary shear instability (SSI) and stagnation point instability (SPI), while the streamwise vortices of purple and green colours in the vortex cores represent the secondary convective instability (SCI). For a description of these instabilities, as well as the mathematical formulation of the various vorticity components plotted in this figure, see MP13. (*a*) $Ri_0 = 0.14$; (*b*) $Ri_0 = 0.16$; (*c*) $Ri_0 = 0.18$; (*d*) $Ri_0 = 0.20$.

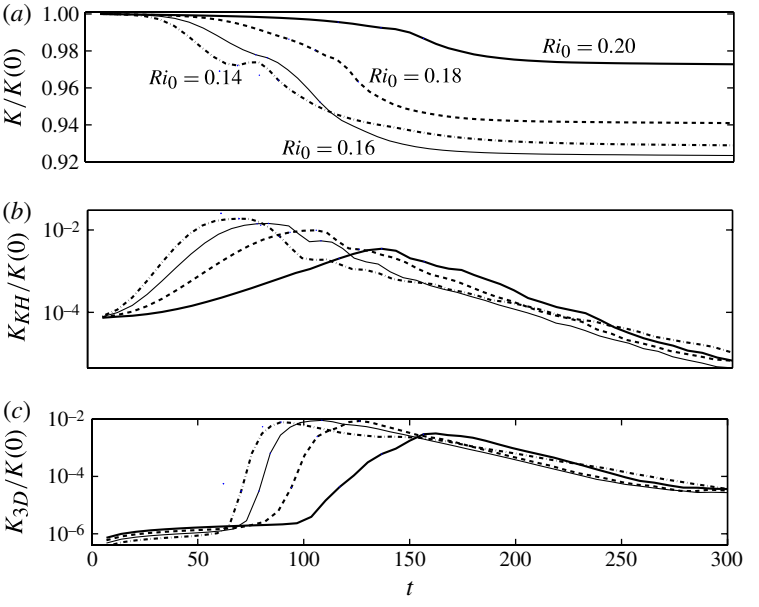

FIGURE 3. Time evolution of the total kinetic energy of the flow (*a*), kinetic energy associated with the roll-up of the two-dimensional KH billow (*b*) and kinetic energy associated with the three-dimensional perturbation field (*c*). All quantities are scaled by the total kinetic energy in the system at time $t = 0$.

mechanism responsible for transition. Mashayek & Peltier (2012*b*) investigated the influence of bulk Richardson number on the evolution of each of the individual secondary instabilities and demonstrated that their growth rates peak at intermediate values of Ri_0. While at low Ri_0 the SSI and SPI do not emerge, essentially due to the stratification being too weak to generate sufficiently intense baroclinic vorticity, at high Ri_0 (0.2 or higher) they do not appear due to the increased stabilizing influence of high Richardson number on the inevitable vertical velocity perturbations associated with these instabilities. It is only at some intermediate value of Ri_0 that the 'zoo' of secondary instabilities is most densely populated, and this we believe is at the heart of the non-monotonicity of the dependence of mixing efficiency upon Richardson number. We explore this further below when we discuss quantitatively the flow energetics, as it is not immediately clear that enhanced turbulent dissipation in and of itself will lead to more, or indeed more efficient, mixing.

As a first step in the quantitative consideration of the flow energetics, figure 3 illustrates the time evolution of the total kinetic energy (figure 3*a*), the kinetic energy associated with the two-dimensional primary KH billow (figure 3*b*) and the kinetic energy of the three-dimensional perturbation field (figure 3*c*) for each of the different flows. As the flow evolves, the energy required for the roll-up of the KH billow and for growth of the three-dimensional perturbation field is provided by the background kinetic energy. After turbulence transition, $\mathscr{K}_{KH}$ and K_{3D} decay and the flow relaminarizes. The loss in the total kinetic energy after relaminarization amounts to a net rise in the background potential energy, which we identify as involving irreversible diapycnal mixing. Inspecting the magnitude and the timing of the peaks of the $\mathscr{K}_{KH}$ curves (hereafter referred to as t^S_{2D}) it is clear that increasing Richardson

number monotonically decreases the saturated magnitude of the primary billows, and monotonically increases the saturation time t^S_{2D}. Furthermore, the timing of the peaks of the $\mathscr{K}_{3D}$ curves (hereafter referred to as t^S_{3D}) also increases monotonically with Ri_0. From this point on, we will refer to the time period prior to t^S_{2D} as the 'preturbulent' phase of flow evolution, to the period between t^S_{2D} and t^S_{3D} as the 'transition' phase, and to the period beyond t^S_{3D} as the 'fully developed turbulent' phase, or simply as the turbulent phase.

All these observations are consistent with the intuitive concept that stratification, appropriately normalized by the flow kinetic energy and hence quantified by the bulk Richardson number Ri_0, may be thought of as a stabilizing effect on flow instability if we may assume that the magnitude of the shear is fixed. However, more careful consideration of the plots shows that the effect of increasing Richardson number is actually much more subtle and complex. The times of onset of growth of the three-dimensional perturbations (hereafter referred to as t^O_{3D}) shows that while at low Ri_0 perturbations begin to grow after saturation of the KH billow (i.e. $t^O_{3D} > t^S_{2D}$, as was discussed in MP13), as Ri_0 increases to 0.14 and beyond, secondary modes actually begin to grow earlier with respect to t_{2D} (noting that while the time of onset of a secondary instability is sensitive to the initial conditions employed in a simulation, as long as similar initial conditions are applied to cases with different Ri_0, the above argument holds). Therefore, at lower Ri_0, the vortex cores have enough time to evolve fully, and hence 'catalyse', and thereafter host a large number of secondary instabilities, whereas at higher Ri_0, perturbations grow early in flow evolution and before the cores build the potential to host multiple instabilities. The results of our new numerical simulations taken together with those discussed in MP13 (which together span the range $0.04 < Ri_0 < 0.2$) reveal that diapycnal mixing is most efficient at $Ri_0 \approx 0.16$, at least for a Prandtl number of unity. The two sets of results together show that at this optimal value of Ri_0, the largest number of secondary instabilities emerge on the primary KH billow and facilitate efficient mixing. The range $0.1 < Ri_0 < 0.18$ seems to be the richest in terms of the number of secondary instabilities facilitating mixing. Figure 3 reveals that the largest transfer of energy from $\mathscr{K}$ and $\mathscr{K}_{KH}$ into $\mathscr{K}_{3D}$ occurs in this range, which might be suggestive of the most efficient mixing occurring at this Richardson number. At $Ri_0 = 0.16$, the primary billow is sufficiently large and stratified to sustain vigorous mixing induced by secondary instability, while it is not too strongly stratified to delay the emergence and limit the vigour of secondary instabilities. Therefore, at $Ri_0 = 0.16$ the secondary instabilities emerge precisely at the time that the available potential energy (APE) is at its peak, and use this APE to mix the density field most efficiently. $Ri_0 = 0.16$ is associated with the maximum value of $\mathscr{K}_{3D}$. Perhaps more significantly, the ultimate value of the total kinetic energy $\mathscr{K}$ is actually minimum for the flow with this Ri_0, showing the largest loss of kinetic energy to both viscous dissipation and diapycnal mixing, which is suggestive at least of the most intense turbulent activity for this flow.

Deeper understanding of the role of the underlying dynamics on the mixing process requires a detailed analysis of the flow energetics, through consideration of the time-evolution of both perturbations and mean flow. Of particular interest will be whether there is evidence of the 'lag' postulated by Barenblatt *et al.* (1993) to regularize flows where the amount of mixing varies non-monotonically with stratification, perhaps equivalent in the case considered here to a non-monotonic variation with Ri_0. To facilitate this, in figure 4 we plot the various contributions to the growth rate of both the total and the turbulent kinetic energy reservoirs defined in (2.9) and (2.12). Moreover, to facilitate further interpretation of the results of figure 4, in figure 5 we

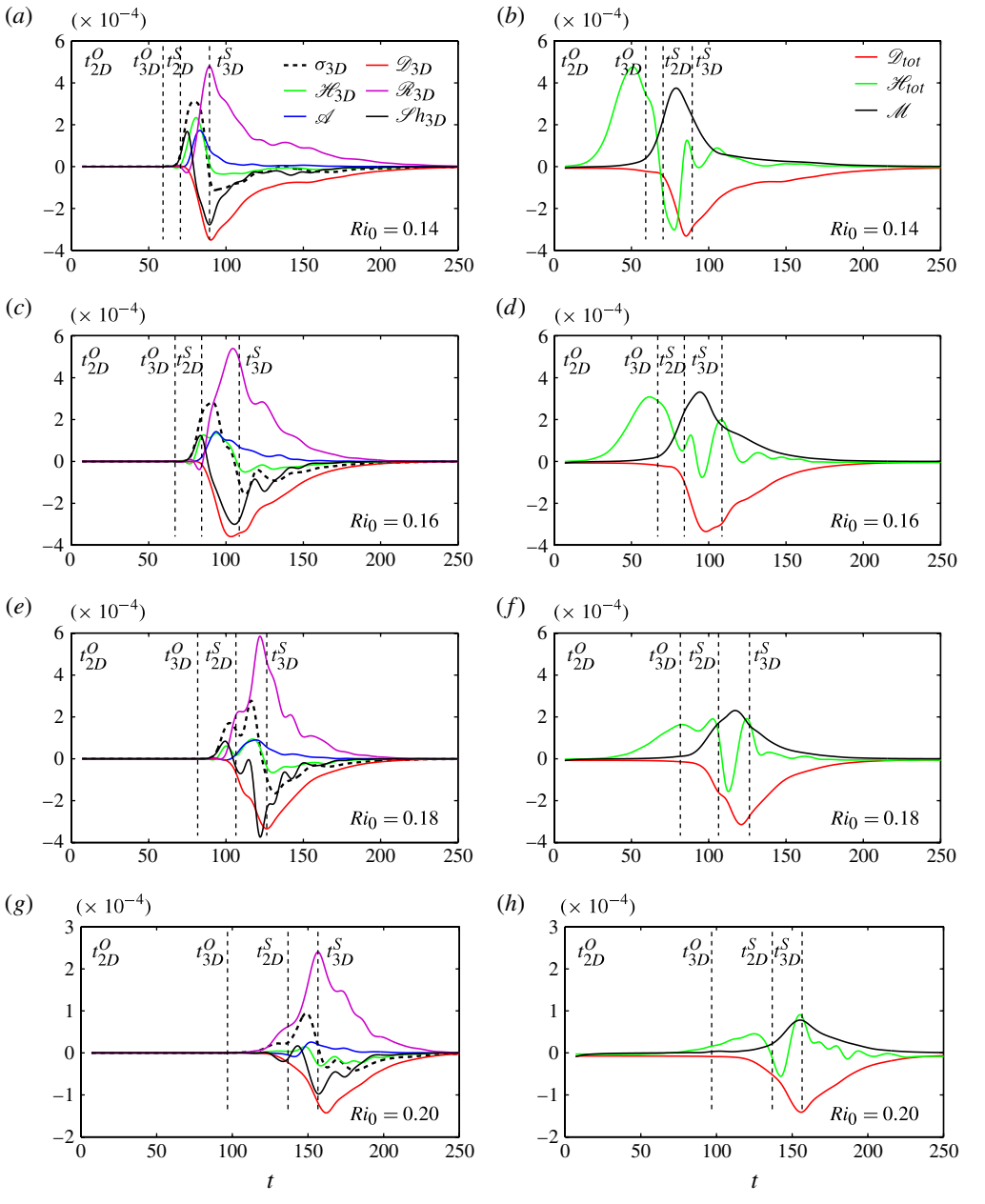

FIGURE 4. (*a,c,e,g*) Time evolution of various contributors to growth of the turbulence as defined in (2.12). (*b,d,f,h*) Time evolution of contributors to the total kinetic energy as defined in (2.9) as well as the total mixing. t^O_{2D} and t^S_{2D} mark the onset and saturation of growth of the primary KH billow while t^O_{3D} and t^S_{3D} mark onset and saturation of growth of the three-dimensional perturbation field. (Note that the range of the vertical axis in the last row has been halved.) To facilitate comparison of various terms in (2.9) and (2.12) in the same figure, all terms in the two relations have been normalized by the total kinetic energy. (*a,b*) $Ri_0 = 0.14$; (*c,d*) $Ri_0 = 0.16$; (*e,f*) $Ri_0 = 0.18$; (*g,h*) $Ri_0 = 0.20$.

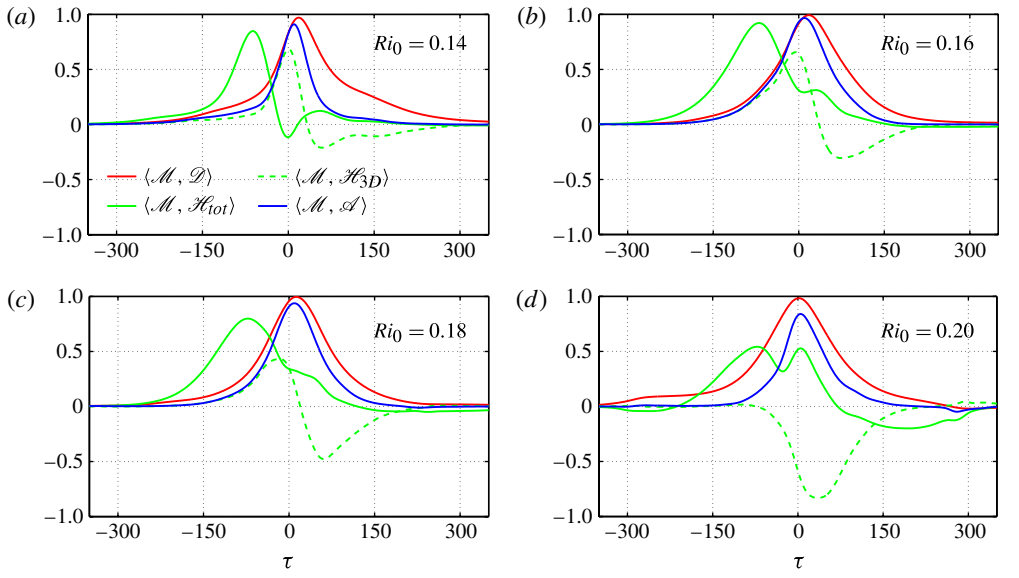

FIGURE 5. Normalized cross correlation functions (defined in (3.1) between total diapycnal mixing, $\mathscr{M}$, and various other quantities defined in (2.9)–(2.12), plotted as functions of the correlation time lag τ. (*a*) $Ri_0 = 0.14$; (*b*) $Ri_0 = 0.16$; (*c*) $Ri_0 = 0.18$; (*d*) $Ri_0 = 0.20$.

plot normalized two-point cross-correlation functions of various of these contributions, in order to provide information concerning both the relative timing of various processes, but also, crucially, their duration. Mathematically, the cross-correlation of functions $f(t)$ and $g(t)$ is simply defined as

$$(f \star g)(t) = \int_{-\infty}^{\infty} f^*(\tau) g(t+\tau)\, \mathrm{d}\tau, \tag{3.1}$$

where f^* denotes the complex conjugate of f. The correlation coefficients plotted in figure 5 are normalized in such a way that the autocorrelation of each of f and g are identically unity at zero lag. In the discussion below, we interpret figures 4 and 5, together.

For $Ri_0 = 0.14$, prior to saturation of the KH billow, three-dimensional instabilities begin growing at t^O_{3D} (see figure 4*a*) and they saturate at t^S_{3D}. Prior to t^O_{3D}, a large increase in the buoyancy flux (see figure 4*b*) leads to the accumulation of APE in the billows. During the time period between t^O_{3D} and t^S_{3D}, this APE is transformed into kinetic energy associated with the rise of a number of three-dimensional instabilities which facilitate significant diapycnal mixing. For this case ($Ri_0 = 0.14$), that rise in three-dimensionality (see σ_{3D} in figure 4*a*) is coincident with a large negative buoyancy flux (see the minimum of the green curve in figure 4*b*) which is due to convective instability of the vortex cores. This large exchange of APE back into kinetic energy also leads to the local maximum in total kinetic energy apparent in figure 3(*a*), which only occurs for this relatively low Ri_0 simulation. This mixing through convection driven by APE conversion is known to be a characteristic of stratified KH mixing at low to intermediate Ri_0 (Klaassen & Peltier 1985; Caulfield & Peltier 2000, MP13). The rise in the magnitude of the $\mathscr{S}h_{3D}$ curve in figure 4(*a*) is a manifestation of

the secondary 'braid' instabilities such as the secondary shear instability (SSI) and stagnation point instability (SPI) discussed above.

At higher Ri_0, the strong growth of these braid modes (which was predicted by Mashayek & Peltier 2012*a*,*b*) dramatically changes the nature of the diapycnal mixing. To demonstrate this, we compare figure 4(*c*,*d*) with figure 4(*a*,*b*). The primary difference between $Ri_0 = 0.14$ and $Ri_0 = 0.16$ is that for the latter value, mixing does not correlate negatively with buoyancy flux in the period $t_{3D}^O < t < t_{3D}^S$ and the time-averaged contribution of $\mathscr{S}h_{3D}$ is larger. This can be observed more clearly in figure 5, which shows that for $Ri_0 = 0.14$ there is a negative correlation between mixing and buoyancy flux at zero lag while that negative peak (which is diagnostic of the central role of APE conversion in mixing in low Re and low to medium Ri_0 flows) disappears for the $Ri_0 \geqslant 0.16$ simulations. The correlation between $\mathscr{A}$ and mixing also increases from $Ri_0 = 0.14$ to $Ri_0 = 0.16$, showing that the energy exchange between the SSI and SPI and the background structured vortices becomes the dominant contributions to mixing along with the SCI.

For the higher Ri_0 cases illustrated in figures 4 and 5, the same trend holds. Mixing is facilitated more and more by small-scale instabilities distinct from the SCI. In fact, a comparison between all panels in figure 5 clearly illustrates this change in behaviour. For all cases, the buoyancy flux is initially dominated by the roll-up of the KH billow in the early stages of flow evolution, leading to the large positive correlation at negative lag for each flow, a contribution which diminishes in magnitude as Ri_0 increases due to the suppression in the vertical extent of the vortex cores (see figure 1). The decrease in the vertical extent of the vortex cores also leads to a marked reduction in the vertical extent of the statically unstable layers within the billow in which convective instability is realized at lower values of Ri_0. Because the effective Rayleigh number of these unstable regions scales as the cube of layer thickness, the increased static stability (more accurately the increasing proximity of the initial Richardson number to the critical value for the onset of inviscid instability) of the initially parallel flow leads to suppression of the shear aligned convective mode of secondary instability (Klaassen & Peltier 1985; Mashayek & Peltier 2012*b*). Instead, for the flows with $Ri_0 \geqslant 0.16$, there is a positive correlation between mixing and the total buoyancy flux at zero lag. This can be understood physically since the smaller-scale overturning for the braid-centred instabilities is located in a region which is statically stable, and hence the overturning corresponds to positive buoyancy flux, unlike that in the region surrounding the periphery of the primary billow core for the flow with $Ri_0 = 0.14$, where the fluid is locally statically unstable, and so overturnings correspond to negative buoyancy flux. For intermediate Ri_0 flows, these overturnings lead to mixing which is characterized also by a positive correlation with the three-dimensional buoyancy flux at zero lag, showing that the three-dimensional overturnings associated with the secondary instabilities are still energetic as the mixing is occurring. However, there is eventually a negative correlation between mixing and three-dimensional buoyancy flux as the flow restratifies, and for the most strongly stratified flow with $Ri_0 = 0.2$ this restratification is actually contemporaneous with the peak in irreversible mixing, leading to a negative correlation between $\mathscr{H}_{3D}$ and $\mathscr{M}$.

In the fully turbulent phase, at $Ri_0 = 0.14$, buoyancy flux correlates negatively with mixing (see figure 5) due to convective instability, and the peak of mixing activity (due to convection) precedes proliferation of secondary shear and strain instabilities. In the fully turbulent phase (beyond t_{3D}^S), small-scale overturns due to shear instabilities (which are microcosms of the original KH billow) lead to an increase in the total buoyancy flux. We shall refer to these structures as *small-scale coherent structures*

(SSCS). The existence of these structures relies inherently on the accessibility of a wide range of energy-containing length scales, and hence sufficiently high Re.

As Ri_0 increases to and beyond $Ri_0 = 0.16$, the negative correlation between mixing and total flux (which was shown in figure 4*b* and discussed above) disappears and the lag between the emergence of shear and braid instabilities and mixing decreases, so that for the two highest values of Ri_0 in the figure, the lag tends towards zero. This implies that mixing becomes increasingly influenced by instabilities other than the classic convective instability (SCI) of the core, and that the more intense turbulence at $Ri_0 = 0.16$ is associated with the abundance of braid-centred secondary instabilities which give rise to an increase in small-scale overturns manifested by a negative turbulent buoyancy flux. This can be further observed by comparing the $\mathscr{A}$ curves in figure 4(*a,c,e,g*) which quantify the anisotropic stretching of the small-scale turbulence by the background coherent structures. While in the early stages of flow evolution (i.e. prior to t^S_{2D}) $\mathscr{A}$ is due to the primary KH billow (and is negligible), for the post-transition period it represents stretching of turbulent structures by the localized SSCS discussed above. As already noted, for $Ri_0 = 0.2$, the vertical turbulent flux has a strong negative correlation with mixing. This mixing is primarily due to structures which are sufficiently large to facilitate numerous localized overturnings. As is apparent in figure 2 for $Ri_0 = 0.2$, the turbulent flux can be inferred to be highly anisotropic, stronger in the horizontal than in the vertical. This is a very different picture from that envisioned by Osborn (1980), which is based on assumption '*I*', i.e. that isotropic turbulence is responsible for producing the (turbulent) vertical buoyancy flux. It is therefore clear that, even if the turbulence were stationary in this regime of high Ri_0, which it is not for the Reynolds number of 6000, application of (1.1) (or variants thereof) to the quantification of diapycnal mixing associated with shear-induced turbulence under initial Richardson number conditions that are only slightly smaller than the Miles–Howard critical value would be highly inappropriate, since the vertical buoyancy flux is not maintained by isotropic turbulence.

Another important observation follows by comparing the various rows of figure 4. At lower Ri_0, the flow evolution can be divided into three primary periods: an initial period of turbulent growth, an intermediate period in which energetic turbulence is maintained and may be approximated as stationary, and a final decaying period (similar to the categorization of Shih *et al.* 2005). This categorization is not specific to shear-layer turbulence and is applicable to other turbulence events such as breaking of an internal wave or evolution of a stratified turbulent wake (Spedding 1997). In general, it is only the intermediate period which may be characterizable by the quasi-steady balance associated with assumption '*S*' used in the derivation of (2.20) and (1.1). It was shown in MP13 that this intermediate period makes up a larger fraction of the whole lifecycle of turbulence with increase in the Reynolds number at a fixed moderate degree of initial inviscid instability for $Ri_0 = 0.12$. Our present analyses, however, show that with increasing Richardson number towards the Miles–Howard critical value, the intermediate period shrinks, and by $Ri_0 = 0.2$ the flow evolution includes only growth and decay periods. Therefore, it is clear that increase in bulk Richardson number further limits the applicability of the Osborn formula by violating assumption '*S*' of stationary turbulence, and indeed, since the flow is continually evolving, it seems unlikely that assumption '*F*', that the turbulence is fully developed, is appropriate either. Especially for values of Ri_0 only slightly less than the critical value of 1/4 for the onset of Kelvin–Helmholtz instability, the Osborn formula is not appropriate as a parametrization of diapycnal mixing by the ensuing turbulent breakdown of KH instability in a stratified mixing layer. For such idealized flows,

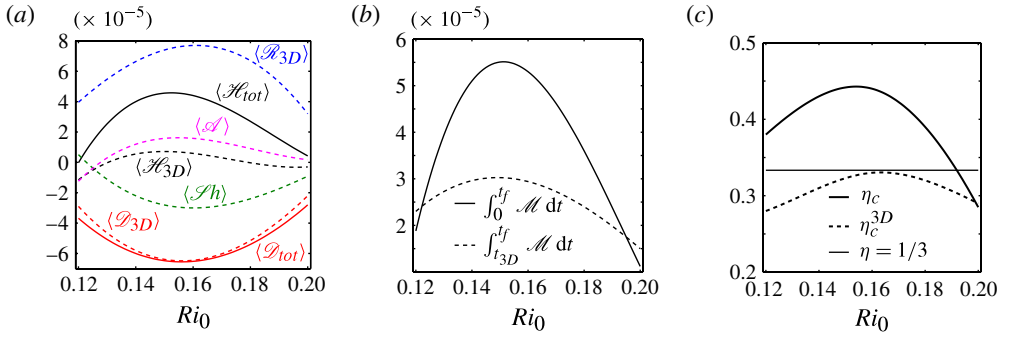

FIGURE 6. (Colour online) Time-averaged quantities of the parameters plotted in figure 4 as a function of the Richardson number. Brackets represent time averaging. Parameters with subscript $3D$ have been integrated only over the post-transition phase of the flow whereas those with the subscript *tot* have been integrated over the whole lifecycle. The thin horizontal line in (*c*) shows $Ri_f = 1/3$, consistent with the theoretical predictions of Caulfield *et al.* (2004) and Tang *et al.* (2009). Note that all curves are obtained from cubic spline fits to data points obtained from simulations for $Ri_0 = 0.12$, 0.14, 0.16.0.18 and 0.2.

our analyses above demonstrate that the underlying central assumptions of the Osborn model do not apply, due specifically to the fact that the turbulence is both non-stationary and anisotropic. Furthermore, the observed mixing in this situation does not exhibit the 'general relationship between the dissipation rate and the buoyancy flux due to the small-scale turbulent velocity fluctuations' which was derived by Osborn in his influential paper, and thus caution should be exercised in the use of that idealized model to the evidently vastly more complex mixing events which occur in the world's oceans.

4. Discussion

As discussed in the previous section, it seems clear, at least qualitatively, that the key assumptions (that the turbulence is fully developed, isotropic and stationary) of the Osborn model are not appropriate to the stratified mixing layer at sufficiently high Re and Ri_0. Furthermore, we have demonstrated that the mixing is, at least at intermediate Ri_0, dominated by the 'zoo' of secondary instabilities introduced in Mashayek & Peltier (2012*a*,*b*). To address the remaining central aims of this paper, i.e. whether the mixing efficiency is a non-monotonic function of the bulk Richardson number Ri_0, and if so, at which value of Ri_0 is the mixing most efficient, we now turn our attention to quantitative consideration of the mixing efficiency of the various flows. Figure 6 presents time averages of the quantities plotted in figure 4 as a function of Ri_0. Figure 6(*a*) shows that the buoyancy flux is not entirely due to the post-transition turbulent flux, but that it is dominated by the larger-scale stirring due to the primary KH billow (the contribution of which to total flux diminishes with increase in Ri_0) and by localized stirring due to small-scale instabilities growing inside the mixing layer. Indeed, the turbulent flux may even have a negative contribution as explained in the previous section. The maximum buoyancy flux occurs in the range $0.14 < Ri_0 < 0.16$ and coincides with the range of maximal total mixing, as shown in figure 6(*b*). The high flux in this range is undoubtedly due to the existence of a large number of secondary instabilities which do not exist at either lower or higher Ri_0. This leaves an intermediate range of Ri_0 in which the flow is richest in

terms of the small-scale instabilities responsible for a rapid transition to turbulence. Viscous dissipation is also maximal in this range because of the strength of the three-dimensional perturbation field enriched by the high secondary instability activity, implying that both the numerator and denominator of the mixing efficiency η_c as defined in (2.19) are maximal for intermediate Ri_0. It actually appears that the mixing efficiency attains its maximum between $Ri_0 = 0.14$ and $Ri_0 = 0.16$, which is similar to the range in which buoyancy flux is maximized.

This exhibits an interesting though perhaps fortuitous agreement with the predictions of Tang *et al.* (2009) for stratified plane Couette flow, which of course is a completely different flow from the one considered here. We speculate that there are two generic characteristics of 'optimal' mixing which the two flows might share, which are indeed effectively the physically motivated arguments presented by Linden (1979). Firstly, it appears that, at least at sufficiently high Re, 'efficient' mixing (i.e. mixing with large η_c) is directly associated with flows with high mixing (i.e. flows with large values of the buoyancy flux in the numerator of η_c). Secondly, while at relatively low bulk Richardson number the thickness of the turbulent layer can grow large, turbulent diapycnal mixing is not highly efficient (disregarding the preturbulent laminar phase of the flow, which is efficient due to negligible dissipation) as the stratification can rapidly be homogenized. Conversely, at values of the Richardson number closer to the Miles–Howard threshold, vertical flux, and the existence of instabilities which would facilitate that flux, are suppressed. There is an intermediate range in which the mixing layer is thick enough to host a large number of secondary (and tertiary) instabilities which owe their existence to the presence of stratification, for example through the baroclinic generation of vorticity. Thus, the intermediate regime is most effective in maintaining a large vertical flux through efficient mixing. Since for both these classes of flows $Ri_0 \geqslant 1/4$ leads to complete suppression of instability, it is thus plausible that flows with $Ri_0 \simeq 1/6$ lie in this intermediate optimal regime.

Figure 6(*b*) shows that while the mixing averaged over the whole lifecycle of the flow (solid curve) peaks over the range $0.14 < Ri_0 < 0.16$, the mixing during the post-transition phase (i.e. $t > t^S_{3D}$, dashed curve) is less variable with Ri_0. Therefore, it is the transition period of the flow ($t^0_{3D} < t < t^S_{3D}$) which contributes the most to the maximization of the mixing at $Ri_0 \approx 0.16$. Figure 6(*c*) shows that the efficiency of mixing during the post-transition phase varies in the range 1/4–1/3, with the upper bound corresponding to $Ri_0 = 0.16$. The total mixing efficiency corresponding to the whole lifecycle, however, varies between 0.28 and 0.45, with the peak again at $Ri_0 = 0.16$ and the smallest values at high Ri_0. Both figures 6(*b*) and 6(*c*) show that the lifecycle-averaged (i.e. averaged over the period between $t = 0$ and $t = t_f$, where t_f marks the final time of simulation at which the flow is fully relaminarized) and turbulent phase-averaged (i.e. averaged between t^S_{3D} and t_f) quantities merge at high Ri_0. We wish to emphasize that the existence of an intermediate range of values of the bulk Richardson number Ri_0 in which mixing is most efficient (and thus the validity of the above arguments) is conditional upon the Reynolds number being sufficiently high that the 'zoo' of secondary instabilities is fully populated so as to facilitate enhanced mixing, as was discussed in detail in MP13. These are the mechanisms for this particular flow by which substantially enhanced (and Ri_0-dependent) mixing can occur. Therefore, it is clear that the stratified mixing layer does indeed exhibit non-monotonic variation of mixing efficiency with Ri_0.

Finally, we are now in a position to test quantitatively the validity of the Osborn (1980) method of calculating an effective diapycnal diffusivity (described in (1.1)) by revisiting again the three basic assumptions upon which this method is based:

assumption 'F' of fully developed turbulence; assumption 'S' of stationary turbulence and assumption 'I' of isotropy. In (4.1), we repeat for clarity (from (2.12)) the complete balance of various terms contributing to evolution of the turbulent kinetic energy and in (4.2), we repeat (from (2.20)) the reduced balance considered by Osborn (1980):

$$\text{complete balance} \quad \frac{1}{2\mathscr{K}_{3D}}\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{K}_{3D} = \mathscr{R}_{3D} + \mathscr{S}h_{3D} + \mathscr{A} - \mathscr{H}_{3D} - \mathscr{D}_{3D}, \tag{4.1}$$

$$\text{Osborn (1980) balance} \quad \mathscr{R}_{3D} = \mathscr{H}_{3D} + \mathscr{D}_{3D}. \tag{4.2}$$

It was shown in figure 4 that while the flow evolves into a fully developed turbulent phase beyond t^S_{3D}, turbulent kinetic energy (thick black dashed lines in figure 4*a,c,e,g*) decays afterwards. Although at low to intermediate values of the bulk Richardson number Ri_0 (such as those considered in MP13 and in the case $Ri_0 = 0.14$ here) there exists an extended period of weak decay of the turbulence which one can treat as a 'stationary period', this phase shrinks in time with increase in Ri_0 and in flows with moderately high Ri_0, assumption 'S' of stationary turbulence in (4.2) completely breaks down, at least for the value of the Reynolds number of 6000 and Prandtl number of unity upon which all of the analyses in this paper have been based. Moreover, we showed in figure 4 that there exists a non-negligible shear-driven exchange of energy between the turbulent structures and the background billow upon which turbulence has grown (i.e. $\mathscr{S}h_{3D}$) as well as a non-negligible contribution due to deformation of anisotropic turbulent structures by the background billow (i.e. $\mathscr{A}$). Both $\mathscr{S}h_{3D}$ and $\mathscr{A}$ exist due to anisotropy in the flow and are comparable to (if not larger than) the other terms in the Osborn balance of (4.2). Therefore, assumption 'I' of isotropy also fails. A final and perhaps the most lethal blow to the ability of (4.2) to represent accurately the buoyancy flux in our simulations actually comes not from the failure of the three underlying assumptions mentioned above, but rather from the incorrect modelling argument that the turbulent buoyancy flux accounts for the total flux. It was shown in figure 4 that coherent structures (such as the primary KH billow and the attendant 'zoo' of secondary instabilities) which exist at scales larger than the turbulent fluctuations contribute significantly to the effective flux and that the turbulent flux can in fact contribute negatively to the upward flux in some periods of flow evolution. As pointed out by Ivey *et al.* (2008), the important contribution of overturns at scales larger than that of turbulence to the buoyancy flux often goes unaccounted for in estimation of the flux.

To formally quantify the error (due to the inapplicability of the underlying assumptions) associated with application of the Osborn balance (4.2) to our numerical experiments, in figure 7 we plot the difference between the left- and right-hand sides of the balance, normalized by the total dissipation for all the cases discussed in this study. All the curves in the plot correspond to the fully developed turbulent phase ($t > t^S_{3D}$) of the flows, which is most suitable for direct comparison with (4.2). As is clearly shown in the figure, for the whole period of fully developed turbulence, (4.2) breaks down. At early stages after t^S_{3D}, large errors are associated with the existence of remnants of the several secondary instabilities which facilitated transition to turbulence, and thus lead to a large influence of anisotropy and non-stationarity effects on the failure of balance. At later times shown in the plot, turbulence is decaying and thus is inherently non-stationary. For the mid-range period of each curve, the non-negligible influence of $\mathscr{A}$ and $\mathscr{S}h_{3D}$ contributes to errors comparable to the total dissipation rate. Figure 7 also illustrates the fact that the duration of the turbulent phase of

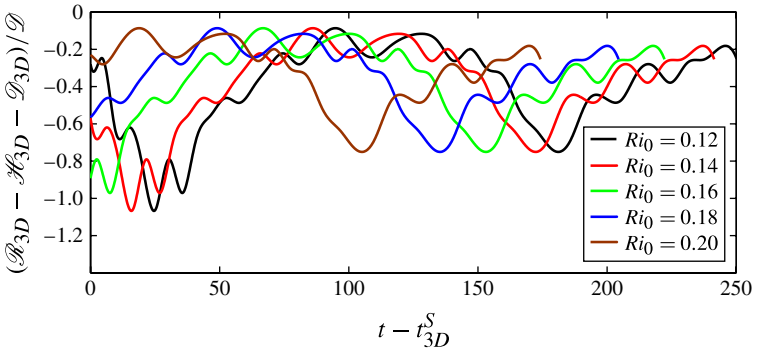

FIGURE 7. $(\mathscr{R}_{3D} - \mathscr{H}_{3D} - \mathscr{D}_{3D})/\mathscr{D}$ plotted versus time for the fully developed turbulent phase of flow ($t > t^S_{3D}$) as a measure of error associated with balance (4.2) due to Osborn (1980).

flow evolution diminishes with increase of the Richardson number, as was previously pointed out in the discussion of figure 4. Thus, figure 7 demonstrates breakdown of the assumptions of 'S' and 'I' in the fully developed turbulent phase of flow evolution. Noting that the balance (4.2) is not even applicable to periods prior to t^S_{3D}, and that the preturbulent phase makes a non-negligible contribution to the time-averaged net buoyancy flux, our analysis questions the applicability of (1.1) to the estimation of an effective turbulent diffusivity in stratified shear layers. Furthermore, based on the above discussion, it is clear that the definition of the flux Richardson number given in (2.21) (which is commonly employed in the oceanographic community), is not an accurate measure of the mixing efficiency and can deviate from the more accurate definition given by (2.19) significantly. This deviation at Richardson numbers which are only slightly below the critical value exists both because the turbulent flux does not account for the total buoyancy flux and because the buoyancy flux is not precisely equivalent to diapycnal mixing due to the existence of stirring at scales larger than the turbulent structures.

We conclude by pointing out once again that the general conclusions of this work are expected not to be unique to the particular framework on which our analyses are based nor to the specific choice of shear instability which we study. The animations of our numerical simulations available as supplementary materials show that once the layer becomes unstable, it becomes rife with many different instabilities and wave breaking events. Therefore, the KH billows examined in this work merely provide a pathway to turbulence and can themselves be considered as sub-structures within larger dynamical systems, just as much smaller KH billows grow on the braid of our primary KH billows. Studies on internal wave breaking often also lead to results suggestive of great variations in mixing properties with changes in the background flow parameters. We expect that the picture of vigorous 'instabilities', which grow preferentially at some intermediate value of bulk Richardson number, leading to a non-monotonic dependence of mixing efficiency upon some appropriate Ri_0, may be generic, provided the Reynolds number of the flow is sufficiently high. Such mixing appears to be inherently transient, with a finite characteristic relaxation time allowing for regularization of the interfaces which must inevitably develop. We aim to report on the dynamics of such layered flows in due course, in particular considering their characteristic depth and drawing connections to the 'stratified turbulence' self-similar scaling predicted by Billant & Chomaz (2001) and observed numerically at sufficiently

high buoyancy Reynolds number, as defined in (1.2) (see for example Riley & de Bruyn Kops 2003 and Brethouwer *et al.* 2007).

Acknowledgements

A.M. acknowledges the award of a David Crighton Fellowship from DAMTP, University of Cambridge, where this research was partially conducted. The simulations were performed on the SciNet facility which is a component of the Compute Canada HPC platform located at the University of Toronto. The research of W.R.P. in Toronto is supported by NSERC Discovery Grant A9627. The authors would like to thank H. Salehipour as well as three anonymous referees for their valuable comments on our manuscript.

Supplementary movies

Supplementary movies are available at http://dx.doi.org/10.1017/jfm.2013.551.

Appendix

In this appendix, we present the derivation of the evolution equation for the kinetic energy associated with the three-dimensional perturbation velocity, (2.12). The following derivation closely follows that of Laprise & Peltier (1989), Klaassen & Peltier (1991), Smyth & Peltier (1991) and Caulfield & Peltier (2000).

We are primarily interested in the initial development of three-dimensional perturbations which grow on the background KH billow. So, we begin by linearizing the governing equations about the spanwise averaged flow to estimate σ_{3d}, the instantaneous growth rate of the three-dimensional perturbation velocity, noting that instantaneously $|\boldsymbol{u}_{3D}| \propto \mathrm{e}^{\sigma_{3D}t}$. Beginning from the Boussinesq governing equations introduced in § 2, we divide the total flow fields into a basic time-evolving two-dimensional state, and three-dimensional perturbation fields. This decomposition may be written explicitly as

$$f(x, y, z, t) = \tilde{f}(x, z, t) + f'(x, y, z, t), \tag{A 1}$$

where f represents any one of the three components of the velocity field or the density or pressure fields. The corresponding fields of the two-dimensional background flow are represented by $\tilde{f}$ and the three-dimensional perturbations to the background fields are represented by f'. For the specific case of three-dimensionalization of KH billows discussed in this paper, the background fields include the original shear and density layers and the over-growing KH billow:

$$\tilde{U} = \bar{U} + u_{2D}, \quad \tilde{W} = 0 + w_{2D}, \quad \tilde{\rho} = \bar{\rho} + \rho_{2D}, \tag{A 2}$$

where u_{2D}, w_{2D} and ρ_{2D} correspond to the KH billow. Upon substitution of (A 1) into the governing equations, and linearizing in the perturbations, the coefficients of the various perturbation terms in the resulting equations comprise the $\tilde{f}$ fields and their derivatives. These coefficients are all independent of the cross-stream y-coordinate and are periodic in x with the same wavelength as the primary KH billow. Hence, we may separate the structures of the perturbation fields as

$$f'(x, y, z, t) = f_{3D}(x, z, t)\mathrm{e}^{\mathrm{i}(bx+dy)}, \tag{A 3}$$

in which $f_{3D}(x, z, t)$ is periodic in x and y, and b and d are the streamwise and spanwise wavenumbers of the three-dimensional perturbations. Equation (A 3) may be

further simplified by noting that the three-dimensional perturbations of interest to us grow to finite amplitude very quickly compared with the evolution of the background two-dimensional flow (Mashayek & Peltier 2013). We can therefore safely ignore the time dependence of the background flow coefficients, and so (A 3) is reduced to the form

$$f'(x, y, z, t) = f_{3D}(x, z)\mathrm{e}^{\mathrm{i}(bx+dy)+\sigma_{3D}t}. \tag{A 4}$$

Upon substitution of the total fields $\tilde{f} + f_{3D}\mathrm{e}^{\mathrm{i}(bx+dy)+\sigma_{3D}t}$ into the governing equations, and after subtraction of the background two-dimensional flow, the resulting linearized equations become

$$\begin{aligned}
\sigma_{3D}u_{3D} &= -\tilde{U}(\partial_x + \mathrm{i}b)u_{3D} - \tilde{W}\partial_z u_{3D} \\
&\quad - (\partial_x + \mathrm{i}b)\tilde{U}u_{3D} - \tilde{U}_z w_{3D} - (\partial_x + \mathrm{i}b)p_{3D} + Re^{-1}\nabla^2 u_{3D}, && \text{(A 5)}\\
\sigma_{3D}v_{3D} &= -\tilde{U}(\partial_x + \mathrm{i}b)v_{3D} - \tilde{W}\partial_z v_{3D} - \mathrm{i}dp_{3D} + Re^{-1}\nabla^2 v_{3D}, && \text{(A 6)}\\
\sigma_{3D}w_{3D} &= -\tilde{U}(\partial_x + \mathrm{i}b)w_{3D} - \tilde{W}\partial_z w_{3D} \\
&\quad - (\partial_x + \mathrm{i}b)\tilde{W}u_{3D} - \tilde{W}_z w_{3D} - \partial_z p_{3D} + Ri\rho_{3D} + Re^{-1}\nabla^2 w_{3D}, && \text{(A 7)}\\
\sigma_{3D}\rho_{3D} &= -\tilde{U}(\partial_x + \mathrm{i}b)\rho_{3D} - \tilde{W}\partial_z \rho_{3D} \\
&\quad - (\partial_x + \mathrm{i}b)\tilde{\rho}u_{3D} - \tilde{\rho}_z w_{3D} + (RePr)^{-1}\nabla^2 \rho_{3D}, && \text{(A 8)}\\
0 &= (\partial_x + \mathrm{i}b)u_{3D} + \mathrm{i}dv_{3D} + \partial_z w_{3D}. && \text{(A 9)}
\end{aligned}$$

We may now construct a budget for the kinetic energy density of the three-dimensional perturbations, $\mathscr{K}_{3D} = 1/2(u_{3D}^2 + v_{3D}^2 + w_{3D}^2)$, by taking

$$\begin{aligned}
&\big[(\text{A 5}) \times u_{3D}^* + (\text{A 5}) \times u_{3D} + (\text{A 6}) \times v_{3D}^* + (\text{A 6}) \times v_{3D} \\
&\quad + (\text{A 7}) \times w_{3D}^* + (\text{A 7}) \times w_{3D}\big]
\end{aligned} \tag{A 10}$$

(where stars represent complex conjugates), and by averaging the resulting equation for $\mathscr{K}$ over the three-dimensional spatial domain to obtain (2.12). Note that in the last step we employed (A 2) and also we assumed that $Pr = 1$. The spatial averaging is done through

$$\langle\rangle_{xyz} = \frac{1}{L_x L_y L_z}\int_0^{L_x}\int_0^{L_y}\int_0^{L_z} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{A 11}$$

where L_x, L_y and L_z are the streamwise, spanwise, and vertical dimensions of the computational domain, respectively.

REFERENCES

BARENBLATT, G. I., BERTSCH, M., PASSO, R. DAL, PROSTOKISHIN, V. M. & UGHI, M. 1993 A mathematical model of turbulent heat and mass transfer in stably stratified shear flow. *J. Fluid Mech.* **253**, 341–358.

BILLANT, P. & CHOMAZ, J. M. 2001 Self-similarity of strongly stratified inviscid flows. *Phys. Fluids* **13**, 1645–1651.

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.

BRITTER, R. E. 1974 An experiment on turbulence in a density-stratified fluid. PhD thesis, Monash University, Victoria, Australia.

CAULFIELD, C. P. & PELTIER, W. R. 2000 Anatomy of the mixing transition in homogeneous and stratified free shear layers. *J. Fluid Mech.* **413**, 1–47.

CAULFIELD, C. P., TANG, W. & PLASTING, S. C. 2004 Reynolds number dependence of an upper bound for the long-time-averaged buoyancy flux in plane stratified Couette flow. *J. Fluid Mech.* **498**, 315–332.

FERNANDO, H. J. S. 1991 Turbulent mixing in stratified fluids. *Annu. Rev. Fluid Mech.* **23**, 455–493.

HAZEL, P. 1972 Numerical studies of the stability of inviscid parallel shear flows. *J. Fluid Mech.* **51**, 39–62.

HOLFORD, J. M. & LINDEN, P. F. 1999 Turbulent mixing in a stratified fluid. *Dyn. Atmos. Oceans* **30**, 173–198.

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

IVEY, G. N. & IMBERGER, J. 1991 On the nature of turbulence in a stratified fluid. Part 1. The energetics of mixing *J. Phys. Oceanogr.* **21**, 650–658.

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 The onset of turbulence in finite amplitude Kelvin–Helmholtz billows. *J. Fluid Mech.* **155**, 1–35.

KLAASSEN, G. P. & PELTIER, W. R. 1991 The influence of stratification on secondary instabilities in free shear layers. *J. Fluid Mech.* **227**, 71–106.

KOOP, C. G. 1976 Instability and turbulence in a stratified shear layer. NASA STI/Recon Technical Report N 77, 16303.

LAPRISE, R. & PELTIER, W. R. 1989 The linear stability of nonlinear mountain waves: implications for the understanding of severe downslope windstorms. *J. Atmos. Sci.* **46** (4), 545–564.

LINDEN, P. F. 1979 Mixing in stratified fluids. *Geophys. Astrophys. Fluid Dyn.* **13**, 3–23.

LINDEN, P. F. 1980 Mixing across a density interface produced by grid turbulence. *J. Fluid Mech.* **100**, 691–703.

LOZOVATSKY, I. D. & FERNANDO, H. J. S. 2012 Mixing efficiency in natural flows. *Phil. Trans. R. Soc.* A **371**, 20120213.

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.

MILES, J. W. 1961 On the stability of heterogeneous shear flows. *J. Fluid Mech.* **10**, 496–508.

OGLETHORPE, R. L. F., CAULFIELD, C. P. & WOODS, A. W. 2013 Spontaneous layering in stratified turbulent Taylor–Couette flow. *J. Fluid Mech.* **721**, R3.

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

PARK, Y. G., WHITEHEAD, J. A. & GNANADESKIAN, A. 1994 Turbulent mixing in stratified fluids: layer formation and energetics. *J. Fluid Mech.* **279**, 279–311.

PATTERSON, M. D., CAULFIELD, C. P., MCELWAINE, J. N. & DALZIEL, S. B. 2006 Time-dependent mixing in stratified Kelvin–Helmholtz billows: experimental observations. *Geophys. Res. Lett.* **33**, L15608.

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

PHILLIPS, O. M. 1972 Turbulence in a strongly stratified fluid: is it unstable? *Deep-Sea Res.* **19**, 79–81.

POSMENTIER, E. S. 1977 The generation of salinity fine structure by vertical diffusion. *J. Phys. Oceanogr.* **7** (2), 298–300.

PRASTOWO, T., GRIFFITHS, R. W., HUGHES, G. O. & HOGG, A. MCC. 2008 Mixing efficiency in controlled exchange flows. *J. Fluid Mech.* **600**, 235–244.

REHMANN, C. R. & KOSEFF, J. R. 2004 Mean potential energy change in stratified grid turbulence. *Dyn. Atmos. Oceans* **37**, 271–294.

RILEY, J. J. & DE BRUYN KOPS, S. M. 2003 Dynamics of turbulence strongly influenced by buoyancy. *Phys. Fluids* **15**, 2047–2059.

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. & MOUM, J. N. 2000 Length scales of turbulence in stably stratified mixing layers. *Phys. Fluids* **12**, 1327–1342.

SMYTH, W. D., MOUM, J. & CALDWELL, D. 2001 The efficiency of mixing in turbulent patches: inferences from direct simulations and microstructure observations. *J. Phys. Oceanogr.* **31**, 1969–1992.

SMYTH, W. D. & PELTIER, W. R. 1991 Instability and transition in finite amplitude Kelvin–Helmholtz and Holmboe waves. *J. Fluid Mech.* **228**, 387–415.

SPEDDING, G. R. 1997 The evolution of initially turbulent bluff-body wakes at high internal Froude number. *J. Fluid Mech.* **337**, 283–301.

STRANG, E. J. & FERNANDO, H. J. S. 2001 Entrainment and mixing in stratified shear flows. *J. Fluid Mech.* **428**, 349–386.

STRETCH, D. D., ROTTMAN, J. W., VENAYAGAMOORTHY, S. K., NOMURA, K. K. & REHMANN, C. R. 2010 Mixing efficiency in decaying stably stratified turbulence. *Dyn. Atmos. Oceans* **49**, 25–36.

TANG, W., CAULFIELD, C. P. & KERSWELL, R. R. 2009 A prediction for the optimal stratification for turbulent mixing. *J. Fluid Mech.* **634**, 487–497.

THORPE, S. A. 1968 A method of producing a shear flow in a stratified fluid. *J. Fluid Mech.* **32**, 693–704.

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

WELLS, M., CENEDESE, C. & CAULFIELD, C. P. 2010 The relationship between flux coefficient and entrainment ratio in density currents. *J. Phys. Oceanogr.* **40**, 2713–2727.

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.

WUNSCH, C. & FERRARI, R. 2004 Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.* **36**, 281–314.