



A comparative study of turbulent stratified shear layers: effect of density gradient distribution

Hieu T. Pham¹ · Sutanu Sarkar¹

Received: 12 August 2021 / Accepted: 2 May 2022
© The Author(s) 2022

Abstract

Direct numerical simulations are performed to compare the evolution of turbulent stratified shear layers with different density gradient profiles at a high Reynolds number. The density profiles include uniform stratification, two-layer hyperbolic tangent profile and a composite of these two profiles. All profiles have the same initial bulk Richardson number ($Ri_{b,0}$); however, the minimum gradient Richardson number and the distribution of density gradient across the shear layer are varied among the cases. The objective of the study is to provide a comparative analysis of the evolution of the shear layers in term of shear layer growth, turbulent kinetic energy as well as the mixing efficiency and its parameterization. The evolution of the shear layers in all cases shows the development of Kelvin–Helmholtz billows, the transition to turbulence by secondary instabilities followed by the decay of turbulence. Comparison among the cases reveals that the amount of turbulent mixing varies with the density gradient distribution inside the shear layer. The minimum gradient Richardson number and the initial bulk Richardson number do not correlate well with the integrated TKE production, dissipation and buoyancy flux. The bulk mixing efficiency for fixed $Ri_{b,0}$ is found to be highest in the case with two-layer density profile and lowest in the case with uniform stratification. However, the parameterizations of the flux coefficient based on buoyancy Reynolds number and the ratio of Ozmidov and Ellison scales show similar scaling in all cases.

Keywords Stratified turbulence · Stratified shear layer · Kelvin–Helmholtz instabilities

1 Introduction

Shear-driven mixing is ubiquitous in many environmental flows and it plays a major role in the transport of mass, momentum, heat and scalar in environmental systems. In the ocean, turbulent mixing by shear is responsible for the transport of heat from the surface to the ocean interior. Therefore, it is important to understand the mixing processes in order to quantify the ocean heat budget. Ocean observations often show mixing events which involve complex shear and density stratification profiles [7, 26, 27]; however, due to the

✉ Hieu T. Pham
h8pham@ucsd.edu

¹ University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0411, USA

need for applicability over a wide range of flow conditions, the parameterizations of mixing are often deduced from simplified flow models. It is often assumed that the mixing parameterizations are generic and deviations from the simplified flow conditions do not significantly influence the results of mixing models. In the present study, we explore the sensitivity of mixing parameterizations to variabilities in the flow conditions.

In the past few decades, the direction numerical simulation (DNS) method has become an important tool to explore the parameterization of shear mixing. Nearly all DNS studies make use of one of the following three canonical models for shear-driven turbulence: (1) homogenous shear with uniform stratification [5, 8, 17, 24]; (2) hyperbolic tangent velocity and density profiles [2, 13, 25]; and (3) hyperbolic-tangent velocity profile with uniform stratification [4, 28]. For example, Shih et al. [24] use DNS of homogenous shear with uniform stratification to propose a mixing model that depends on buoyancy Reynolds number. Mashayek et al. [13, 15] use DNS of Kelvin–Helmholtz (K–H) shear instabilities with the hyperbolic tangent velocity and density profiles to propose a different mixing model which also depends on the buoyancy Reynolds number. Recently, VanDine et al. [28] perform DNS of K–H instabilities with hyperbolic tangent velocity profile and uniform stratification to show that the evolution of the shear layer and the amount of turbulent mixing can be significantly different from the results of Mashayek et al. [15]. It is noted that all of these studies use one specific type of velocity and density profile to parameterize the mixing in the complex shear flows of the natural environment. Furthermore, since the evolution of KH shear instabilities at high Reynolds number is sensitive to the numerical setup [9], it is necessary to perform a comparative study to explore whether differences in the background flow conditions, specifically the stratification profiles, can affect the results of mixing parameterizations. In the present study, we aim to explore how the stratification distribution inside shear layers affect the evolution of turbulent mixing. The simulations in the present study include a uniform stratification, a hyperbolic tangent density profile and a composite case of these two profiles.

Previous studies have pointed out the dynamical differences in the evolution of shear layers between the hyperbolic tangent density profile and the continuous stratification. Pham et al. [20] and Pham and Sarkar [19] compare the evolution of the KH instability between these two cases and show that internal waves can be excited when the ambient has sufficiently strong stratification. However, these simulations are performed at low Reynolds numbers so the evolution of the shear layer does not exhibit the growth of secondary instabilities as pointed out by Mashayek and Peltier [11, 12]. Another notable difference between the two density profiles is the development of a transition layer as the shear layers become turbulent. VanDine et al. [28] denote the transition layers with elevated shear and stratification at the edges of the shear layer in their simulations of shear layer with uniform stratification. Significant turbulent dissipation and mixing are found in the transition layers. It is unclear whether transition layers can develop in the hyperbolic tangent density profile. Nearly all studies using the hyperbolic tangent density profile take the momentum thickness to delineate the shear layer boundaries [14, 25]. In this study, we investigate the development of the transition layer in turbulent shear layers with various density gradient distributions.

A popular mixing model for shear-driven mixing is from Osborn [18]. In the model, the turbulent thermal diffusivity (K_ρ) is parameterized as $\Gamma\epsilon/N^2$ where Γ is the flux coefficient, ϵ is the dissipation rate and N^2 is the background stratification. Osborn [18] suggests $\Gamma = 0.2$ to be the upper-bound value to be used in the parameterization. The model is widely used to estimate thermal diffusivity from ocean measurements. However, DNS of stratified shear layers with various density and velocity profiles all point to considerable variability

in the mixing efficiency (E) which is related to the flux coefficient $\Gamma = E/(1-E)$ [13, 21, 24, 28]. Furthermore, Γ is suggested to vary as a function of buoyancy Reynolds number, Prandtl number, Richardson number and Froude number. Notably, the mixing parameterizations suggested by these previous studies show significant differences from one another. For example, Shih et al. [24] suggest three regimes of mixing while Mashayek et al. [15] suggest only two regimes of mixing although both parameterizations are based on buoyancy Reynolds number. VanDine et al. [28] demonstrate that Γ can have a significantly larger value than what is suggested in Mashayek et al. [15] at the same buoyancy Reynolds number. The difference among these studies can be due to the different stratification profiles being used in the simulations. Here, we aim to further explore the effect of stratification profiles on the turbulent mixing and its parameterization.

In the present study, we perform DNS of stratified shear layers at a high Reynolds number with three density gradient distributions: a uniform stratification; a two-layer hyperbolic tangent profile and finally a composite profile with both the hyperbolic tangent density profile at the center of the shear layer and uniform stratification in the ambient. The cases have the same initial non-dimensional parameters like the Reynolds number, Prandtl number and bulk Richardson number. Comparative analysis of the simulations will shed light on how the variation in density profiles can influence the physics of turbulent mixing despite having the same initial bulk control parameters. This work is organized as follows. In Sect. 2, we describe the model setup and discuss the relevant non-dimensional parameters with particular attention on the differences in the gradient Richardson profile. Results from linear stability analysis and evolution of the shear layers in the DNS are compared in Sect. 3. Section 4 focus on the quantification of turbulent mixing and its parameterization. We provide discussions and conclusions in Sect. 5 to address the questions raised in this section.

2 Model formulation

In this section, we introduce the model problem of a shear layer with different types of stratification and the numerical implementation of the DNS method employed. The initial profiles of velocity, density and its gradient are presented in detail with particular attention to the important non-dimensional parameters such as the bulk Richardson number and the profiles of gradient Richardson number. It is our objective to explore how turbulent mixing differs with different types of density stratification.

2.1 Distribution of stratification in a shear layer

We consider a shear layer between two streams flowing in opposite directions:

$$\langle u^* \rangle(z^*, t = 0) = -\frac{\Delta U^*}{2} \tanh\left(\frac{z^*}{0.5\delta^*}\right), \quad (1)$$

where ΔU^* denotes the velocity difference across the shear layer, δ^* denotes the half-thickness and the superscript $*$ denotes a dimensional quantity. When the shear layer is stratified with the two streams having constant density values which differ by $\Delta\rho^*$, the density has a hyperbolic tangent profile similar to the velocity profile:

$$\langle \rho^* \rangle(z^*, t = 0) = \rho_0^* - \frac{\Delta \rho^*}{2} \tanh\left(\frac{z^*}{0.5\delta^*}\right), \quad (2)$$

where ρ_0^* is a reference density value and $\Delta \rho^*$ is the density difference across the shear layer ($-1 \leq z/\delta^* \leq 1$). In contrast, when the shear layer has uniform stratification, the density profile takes the following form:

$$\langle \rho^* \rangle(z^*, t = 0) = \rho_0^* - \frac{\Delta \rho^* z^*}{2\delta^*}, \quad (3)$$

Here, the density gradient is chosen such that density difference across the shear layer is the same between the two-layer density profile and uniform stratification. In order to contrast the effects of the two stratification profiles above, we construct a third profile by combining the two profiles with the following expression:

$$\langle \rho^* \rangle(z^*, t = 0) = \rho_0^* - \frac{\Delta \rho^*}{2} \left[\frac{Rz^*}{\delta^*} + (1 - R) \tanh\left(\frac{z^*}{0.5\delta^*}\right) \right]. \quad (4)$$

Here, the density profile consists of two terms: a linear and a hyperbolic tangent term that are bridged together through the parameter R . The parameter R with values ranging between 0 and 1 is used to vary the density gradient distribution across the shear layers. The density has a two-layer hyperbolic tangent profile when $R = 0$, and it has uniform gradient when $R = 1$. For values of R between the two extremes, the density profile is a composite of both profiles.

Using δ^* , ΔU^* , and $\Delta \rho^*$ as the characteristic length, velocity and density scales, respectively, we use a direct numerical simulation (DNS) method to solve the three-dimensional Navier-Stokes equations with the Boussinesq approximation for the evolution of the non-dimensional velocity components (u , v , w), density deviation ($\tilde{\rho}$) from the initial background profile ($\rho_{bg} = \rho(z, t = 0) - \rho_0$) and dynamic pressure (p) in Cartesian coordinates (x , y , z) as follows:

$$\frac{\partial u_j}{\partial x_j} = 0, \quad (5a)$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial(u_j u_i)}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re} \frac{\partial^2 u_i}{\partial x_j \partial x_j} - Ri_{b,0} \tilde{\rho} g_i, \quad (5b)$$

$$\frac{\partial \tilde{\rho}}{\partial t} + \frac{\partial(u_j \tilde{\rho})}{\partial x_j} = \frac{1}{Re Pr} \left(\frac{\partial^2 \tilde{\rho}}{\partial x_j \partial x_j} + \frac{d^2 \rho_{bg}}{dz^2} \right) - w \frac{d\rho_{bg}}{dz}. \quad (5c)$$

The nondimensional parameters are initial Reynolds number (Re), initial bulk Richardson number ($Ri_{b,0}$), and Prandtl number (Pr) which are defined as follows:

$$Re = \frac{\Delta U^* \delta^*}{\nu^*}, \quad Ri_{b,0} = \frac{\Delta \rho^* g^* \delta^*}{\rho_0^* \Delta U^{*2}}, \quad Pr = \frac{\nu^*}{\kappa^*}. \quad (6)$$

Here, ν^* and κ^* are the kinematic viscosity and thermal diffusivity, respectively. We set $Re = 24,000$, $Ri_{b,0} = 0.16$ and $Pr = 1$ for all simulations in the present study. The Reynolds number is sufficiently large for the secondary instabilities to develop in both the linear and

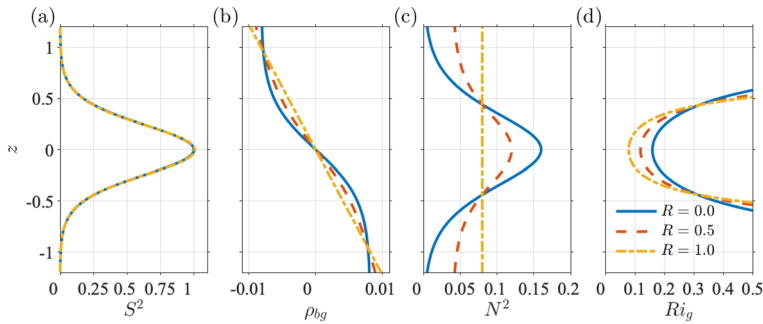


Fig. 1 Initial profiles of stratified shear layers with different density gradient distribution parameter (R): **a** mean squared shear rate (S^2), **b** background density (ρ_{bg}), **c** stratification (N^2), and **d** gradient Richardson number (Ri_g). The density difference across twice of the shear layer thickness measured by $\pm l_u$ is similar among the cases yielding the same initial bulk Richardson number $Ri_{b,0}$. Here, $R = 0$ corresponds to the two-layer density profile and $R = 1$ to uniform stratification

Table 1 Parameters used in the simulations: density gradient distribution parameter (R), initial bulk Richardson number ($Ri_{b,0}$), initial minimum gradient Richardson number Ri_{min} , initial shear layer averaged Richardson number ($\overline{Ri_0}$), growth rate (σ) and wavenumber (k_0) of the fastest growing modes (FGM), number of grid points (N_x , N_y , N_z), grid spacing in the shear layer (Δ), bulk mixing efficiency (representative of the entire evolution) computed over the momentum thickness (E_{lu}) and that computed over the shear layer thickness (E_{sl})

R	$Ri_{b,0}$	Ri_{min}	$\overline{Ri_0}$	σ	k_0	N_x	N_y	N_z	$\Delta (\times 10^{-3})$	E_{lu}	E_{sl}
0	0.16	0.16	0.12	0.081	0.91	2305	1153	1024	6.00	0.36	0.42
0.5	0.16	0.12	0.1	0.105	0.98	2048	1024	1024	6.12	0.37	0.37
1	0.16	0.08	0.08	0.127	1.06	2048	1024	1024	5.79	0.35	0.31

All simulations have the same initial bulk Richardson number $Ri_{b,0} = 0.16$. The domain has the streamwise and spanwise extent of two and one wavelength of the FGM, respectively

hyperbolic tangent density profiles as shown in Mashayek et al. [13] and VanDine et al. [28], respectively. The initial bulk Richardson number of $Ri_{b,0} = 0.16$ allows us to compare our results with the two previous studies.

Three simulations are performed with the density gradient distribution parameter R equal to 0, 0.5 and 1. The two-layer case with $R = 0$ has been considered by Mashayek et al. [13], and we perform the simulation again in order to contrast with other cases. The uniform stratification case $R = 1$ is taken from VanDine et al. [28]. Figure 1 shows the initial profiles of squared shear rate (S^2 where $S = \partial \langle u \rangle / \partial z$), background density (ρ_{bg}), stratification ($N^2 = -g / \rho_0 \partial \langle \rho \rangle / \partial z$) and gradient Richardson number ($Ri_g = N^2 / S^2$) where the angle brackets denote a horizontal average over x - y planes. While the initial bulk Richardson number is the same in the three cases, the minimum gradient Richardson number value (Ri_{min}) and the shear layer averaged Richardson number ($\overline{Ri_0} = N^2 / S^2$ where overbar denotes averaging over $-1 \leq z \leq 1$) are different at $t = 0$ as listed in Table 1. Increasing R values results in the decrease of both Ri_{min} and \overline{Ri} at initial time noting the following relationship: $Ri_{min} = Ri_{b,0}(1 - 0.5R)$. It is the objective of the present study to explore which of the three Richardson numbers is best correlated with the evolution of turbulence in the stratified shear layers.

2.2 Numerical methods

The simulations are performed using the same DNS solver as previously used in Pham et al. [20] and Brucker and Sarkar [1]. Second-order central differences are used for spatial derivatives and a third-order low-storage Runge–Kutta method is used for time advancement in Eq. (5). The dynamic pressure is solved using a multi-grid Poisson solver. In order to set up the computational domain, linear stability analysis (LSA) is performed to identify the fastest growing modes (FGM) in each case. Results from the LSA are discussed in detail in the next section. In all simulations, the domain consists of two wavelengths of the FGM in the streamwise (x) direction and one wavelength in the spanwise (y) direction. The vertical (z) extent spans the region $-16 \leq z \leq 16$. The grid size and grid spacing are listed in Table 1. The grid spacing ($\Delta x = \Delta y = \Delta z = \Delta$) is uniform in the middle region of the domain ($-2.2 \leq z \leq 2.2$ in the $R = 0$ and 0.5 cases and $-2 \leq z \leq 2$ in the $R = 1$ case). The vertical grid spacing is mildly stretched outside this region at a maximum rate of 3%. The fine grid resolution in the middle region is sufficient to resolve at least 2.2 Kolmogorov length scale associated with the turbulence developed inside the shear layer. Periodic boundary conditions in the horizontal directions are used for all independent variables. Free slip conditions are used for the velocity components while the pressure and the density deviation satisfy the homogeneous Neumann condition at the top and bottom boundaries. Sponge layers with a thickness of 5 are implemented at the top and bottom boundaries to prevent the reflection of internal waves.

In order to promote the growth of K–H instabilities, velocity perturbations are added to the three velocity components. The perturbations have a broadband energy spectrum given by $E(k) \propto k^4 \exp[-2(k/k_0)^2]$ where the spectrum is set to peak at the wavenumber k_0 of the FGM as listed in Table 1. The fluctuations have the peak r.m.s. value of $0.1\% \Delta U$ at the center of the shear layer and decay with distance away from the center as $\exp[-(z/\delta)^2]$.

2.3 Statistical analysis

Mean quantities denoted by angle brackets are obtained as x-y horizontal averages and the fluctuations denoted by primes are deviation from the mean. We diagnose the evolution of the turbulent kinetic energy (TKE) to compare the results from different simulations. The TKE budgets are computed using the following evolution equation:

$$\frac{DK}{Dt} = P - \epsilon + B - \frac{\partial T_3}{\partial z}, \quad (7)$$

with turbulent kinetic energy (K), production (P), dissipation (ϵ), buoyancy flux (B), and transport term (T_3) specified as

$$\begin{aligned} K &= \frac{1}{2} (\langle u' \rangle^2 + \langle v' \rangle^2 + \langle w' \rangle^2), \\ P &= -\langle u' w' \rangle \frac{d\langle u \rangle}{dz}, \quad \epsilon = \frac{2}{Re} \langle s'_{ij} s'_{ij} \rangle, \quad s'_{ij} = \frac{1}{2} \left(\frac{\partial u'_i}{\partial x_j} + \frac{\partial u'_j}{\partial x_i} \right), \\ B &= -Ri_{b,0} \langle \rho' w' \rangle, \quad T_3 = \frac{1}{2} \langle w' u'_i u'_i \rangle + \frac{1}{\rho_0} \langle w' p' \rangle - \frac{2}{Re} \langle u'_i s'_{3i} \rangle. \end{aligned} \quad (8)$$

We also compute the scalar dissipation rate χ defined as follows:

$$\chi = \frac{1}{RePr} \left\langle \frac{\partial \rho'}{\partial x_i} \frac{\partial \rho'}{\partial x_i} \right\rangle \quad (9)$$

to quantify the amount of turbulent mixing.

3 Evolution of the stratified shear layer

Previous studies indicate that the stratified shear layer evolves through three stages: (1) the growth of primary K–H instabilities; (2) development of the secondary instabilities that trigger the transition to turbulence; and (3) fully-developed turbulence and its decay until relaminarization of the shear layer. In this section, we explore the role of the density gradient distribution parameter R during each stage of the evolution.

3.1 The development of Kelvin–Helmholtz instability

We perform linear stability analysis to compare the growth rate (σ) of the fastest growing modes (FGM) of K–H instabilities among the cases. The LSA uses a finite-difference method with a grid spacing of $\Delta z = 0.0025$ to solve for the FGM as illustrated in Smyth et al. [27]. The growth rate in the R - $Ri_{b,0}$ parameter space shown in Fig. 2a indicates that the $Ri_{min} < 0.25$ criterion is required, independent of the R value [3, 6]. The growth rate diminishes when $Ri_{b,0}$ and R approach the curve of $Ri_{min} = 0.25$ (shown in magenta). It is important to point out that, for $R = 1$, the shear layer is still unstable to K–H instabilities as long as $Ri_{b,0} < 0.5$. The growth rate varies little with R when $Ri_{b,0}$ is fixed at a small value, i.e. for weakly stratified shear layers. As $Ri_{b,0}$ increases, the growth rate increases with increasing R value. At a fixed value of $Ri_{b,0} = 0.16$, the growth rate in the linear case ($R = 1$) is 50% larger than the value in the hyperbolic tangent case ($R = 0$) as listed in Table 1. The growth rate in the composite case ($R = 0.5$) is closer to the value in the $R = 1$ case indicating the growth rate does not vary linearly with R at a fixed value of $Ri_{b,0}$. Noting that

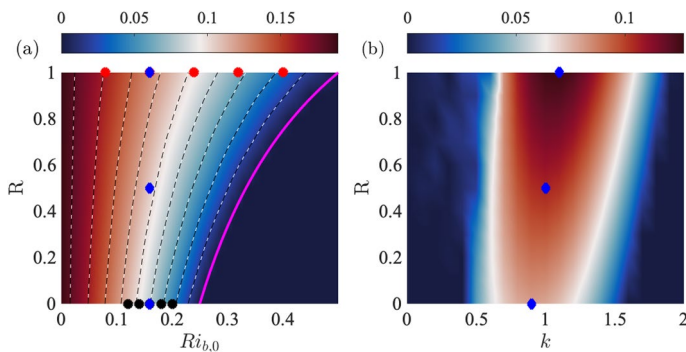


Fig. 2 Variability of the growth rate (σ) of the FGM as a function of: **a** initial bulk Richardson number ($Ri_{b,0}$) and density gradient distribution parameter (R); and **b** wavenumber (k) and density gradient distribution parameter (R) at $Ri_{b,0} = 0.16$. The blue markers denote the parameters used in the present simulations while the black and red markers denote the simulations with $R = 0$ in Mashayek et al. [13] and with $R = 1$ in VanDine et al. [28], respectively. The magenta line in **a** marks $Ri_{min} = 0.25$ while the dashed lines denote curves of constant growth rate

Ri_{min} decreases linearly with increasing R when $Ri_{b,0}$ is fixed, it follows that the growth rate increases as R increases in the present simulations.

Figure 2b reveals that the wavenumber of the FGM increases with increasing R when $Ri_{b,0} = 0.16$, and thus, the wavelength becomes shorter. The wavelength is approximately 17% longer in the $R = 0$ case when compared to the $R = 1$ case. Previous studies have shown that the wavelength becomes shorter when the $Ri_{b,0}$ (and Ri_{min}) increases for either the $R = 0$ or $R = 1$ density profiles [10, 28]. In the present study, we keep $Ri_{b,0}$ to be the same among the cases with different R values and the wavelength becomes shorter with increasing R . Thus, the wavelength of the FGM is correlated better with Ri_{min} rather than $Ri_{b,0}$. Increasing Ri_{min} causes the growth rate to become smaller and the wavelength to become shorter for the K–H modes at different R values. We note that the evolution of K–H instabilities at high Reynolds number is sensitive to the streamwise length (L_x) of the computational domain of the DNS especially when the domain is limited to one or two wavelengths of the most unstable modes. When L_x is not set up in multiples of the most unstable modes, the mode that actually develops in the simulation has a smaller growth rate and thus it can influence the subsequent nonlinear evolution of the shear layer.

3.2 Nonlinear evolution of Kelvin–Helmholtz instabilities and secondary instabilities

In general, the development of K–H instabilities and subsequent secondary instabilities are similar among the simulated cases. The evolution and type of the secondary instabilities (e.g., secondary shear instability, secondary convective instability, secondary core deformation instability, localized core vortex instability, etc.) have been discussed in detail for the $R = 0$ (e.g., Mashayek et al. [13]; Salehipour et al. [21]) and $R = 1$ (e.g., VanDine et al. [28]) cases in many previous studies. For brevity, we refrain from repeating the discussion on the type of secondary instabilities and instead provide Fig. 3 to show the evolution of the shear layer in the $R = 0.5$ case through snapshots of density, spanwise vorticity and turbulent dissipation on the $y = L_y/2$ plane. Secondary instabilities realized by density overturns with enhanced localized vorticity and viscous dissipation develop on the upper eyelid of the left billow in Fig. 3a. The secondary instabilities also develop at the right end of the two braids connecting the two K–H billows. The upper and lower eyelids of the K–H billows are subsequently broken down by the secondary instabilities and the resulting turbulent eddies spread into the cores of the billows as well as in the streamwise direction from the billows into the braid regions. The dissipation is most intense during this period as shown in Fig. 3b and the turbulence field becomes more homogenous throughout the shear layer. After the transition to turbulence, Fig. 3c shows that the shear layer continues to grow vertically by entraining fluid from the ambient despite having weaker vorticity and dissipation relative to just after the transition. The dissipation induced by the eddies at the edges of the shear layer in Fig. 3c takes values comparable to and larger than those at the center of the shear layer.

The shear layer grows differently in the three cases as shown in Fig. 4. Here, we compare the thickness of the shear layer using two measures: the momentum thickness ($I_u = \int_{L_z} (1 - 4\langle u \rangle^2) dz$) and the shear layer thickness (L_{sl}) defined as the region between the edges of the shear layer where the stratification N^2 deviates from its initial background value by 5% of $Ri_{b,0}$. The two measures are shown in Fig. 4a, b, respectively. In all cases, both I_u and L_{sl} exhibit two distinct periods of growth: the first due to the enlargement of K–H billows and the second due to turbulent entrainment. Since the growth rate of the

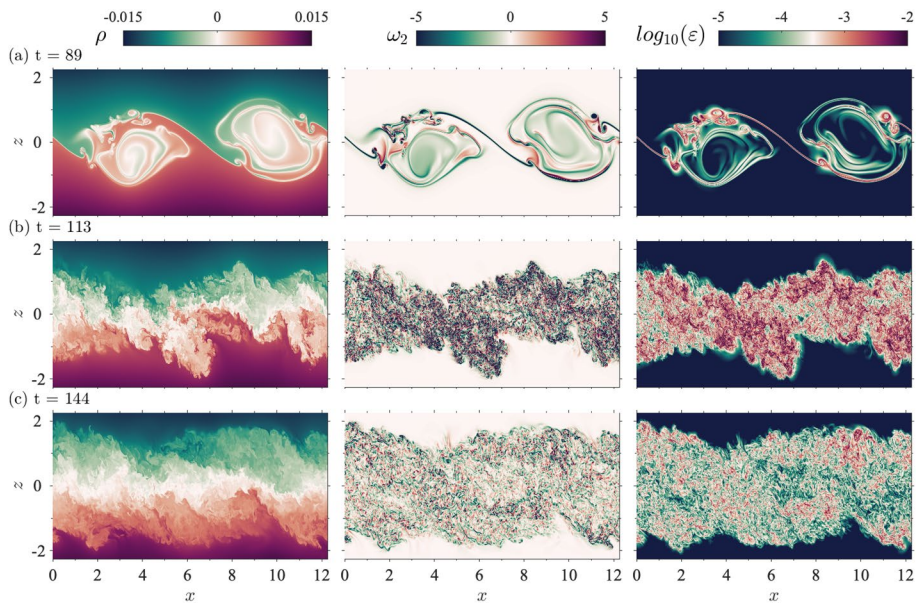


Fig. 3 Evolution of the turbulent shear layer is shown using the density (left column), spanwise vorticity (middle column) and dissipation rate (right column) in the vertical plane at $y = L_y/2$ in the $R = 0.5$ case during **a** the development of primary K–H and secondary instabilities; **b** the transition to 3D turbulence and **c** the decay of turbulence and the relaminarization of the shear layer

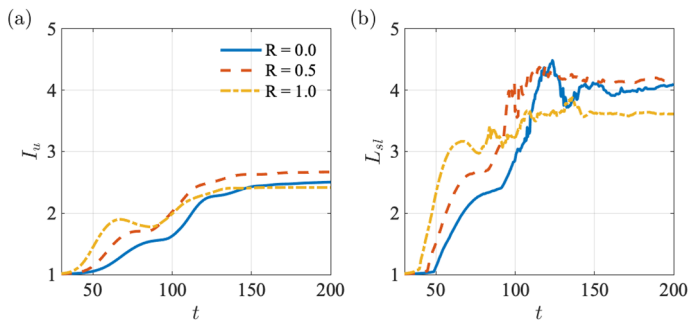


Fig. 4 Evolution of **a** the momentum thickness (I_u) and **b** the shear layer thickness (L_{sl}). The late-time values of I_u are similar among the three cases; however, the L_{sl} is slightly smaller in the $Ri = 1$ case. In all cases, the values of L_{sl} at late time is at least 40% larger than I_u

FGM is largest in the $R = 1$ case, the K–H billows grow first as indicated by the growth of I_u during $40 < t < 65$ in Fig. 4a. Furthermore, the billows grow fastest in this case. The $R = 0$ case with the smallest growth rate has the K–H billows form at a later time and the billows have the smallest size as indicated by the values of I_u at $t = 90$ before the turbulent entrainment period begins. The evolution of L_{sl} in Fig. 4b also shows a similar trend. It is noted that, in the present study where $Ri_{b,0}$ is fixed at 0.16 in all cases, the maximum size of the K–H billows is correlated with the growth rate of the FGM and thus the values of Ri_{min} , i.e., the smaller Ri_{min} produces the larger K–H billows in this study. However, the

trend of growth rate reverses during the turbulent entrainment period. The I_u thickness grows fastest in the $R = 0$ case because the turbulence entrains fluid from an unstratified ambient. In contrast, since the ambient stratification is strongest in the $R = 1$ case, the turbulent entrainment rate is smallest. It is interesting that, despite the different growth rates during the two growth periods, the late time values of I_u are not very different among these cases which have the same $Ri_{b,0}$. The evolution of L_{sl} also shows the fastest growth rate in the $R = 0$ case during the turbulent entrainment period. At late time, L_{sl} is approximately equal in the $R = 0$ and 0.5 cases and it is slightly smaller in the $R = 1$ case. Thus, $Ri_{b,0}$ is a more reliable indicator for the final values of both measures of the shear layer thickness when compared to Ri_{min} as the R values vary.

3.3 Fully-developed stratified turbulence and the relaminarization of the shear layer

We observe significant differences among the cases as the shear layer becomes turbulent specifically in terms of the vertical distribution of shear, stratification and gradient Richardson number. Figure 5 illustrates how the evolution of N^2 profiles differs in cases with different R

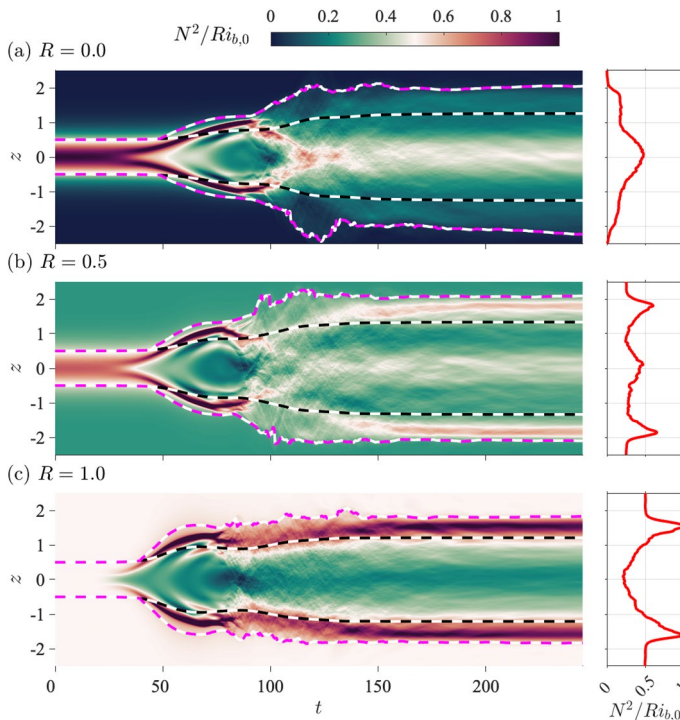


Fig. 5 Comparison of the normalized stratification ($N^2/Ri_{b,0}$) among the simulated cases: **a** $R = 0$; **b** $R = 0.5$; and **c** $R = 1$. The presence of stratification in the ambient induces an overshoot in the stratification in the transition layer in the two latter cases but not in the $R = 0$ case. After the transition to turbulence, a layer of enhanced stratification reappears in the center of the shear layer in the $R = 0$ and 0.5 cases but not in the $R = 1$ case. The right panels show N^2 profiles during the turbulence decay period at $t = 200$. In this figure and hereafter, black and magenta dashed lines mark the shear layer boundaries as indicated by the momentum thickness ($\pm I_u/2$) and the shear layer thickness (L_{sl}), respectively

values. The most significant difference is the absence of transition layer in the $R = 0$ case. For $R = 1$ density profiles, VanDine et al. [28] note the development of a transition layer (TL), which is defined as a thin layer in which the local N^2 is significantly enhanced when compared to the values inside the shear layer. There is a TL at each edge of the shear layer where N^2 is as large as 2.2 times the background N^2 value in the ambient at $Ri_{b,0} = 0.08$ in VanDine et al. [28]. In the present study, the TL only develop in the cases with ambient stratification ($R = 0.5$ and 1) as depicted in Fig. 5b, c. The stratification in the $R = 0$ case is enhanced at the edges of the shear layer when compared to the unstratified ambient; however, unlike Lewin and Caulfield [10], the local N^2 is not larger than the values inside the shear layer. Thus, a distinctive transition layer does not form in the two-layer stratification. We note that Watanabe and Nagata [29] also observe the peak N^2 to occur near the center of the shear layer at late time rather than in the TL in their DNS of a shear layer with $R = 0$ as in the present study. They suggest that the absence of the TL in the $R = 0$ case can be caused by many factors such as the amplitude and spectral content of the broadband initial velocity perturbations as well as the domain size. Regardless to whether the TL form, it is important to delineate the boundaries of the shear layer. The upper and lower black dashed lines in Fig. 5 mark the momentum thickness of the shear layer ($\pm I_u/2$) while the upper and lower magenta dashed lines mark the edge of the shear layer (L_{sl}). It is evident that $\pm I_u/2$ does not capture the TL in the $R = 0.5$ and 1 cases. There is also significant enhancement of stratification outside $\pm I_u/2$ in the $R = 0$ case. Therefore, unlike I_u , L_{sl} is able to capture reliably the edges of the shear layer outside which the stratification is indeed equal to the ambient stratification.

The evolution of the squared shear profiles (S^2) also shows different layering structures among the cases as shown in Fig. 6. The enhanced shear is well correlated with the enhanced N^2 (Fig. 5) indicating that the transport of momentum and density by turbulent eddies occur together. The enhanced shear is most notable in the TL of the $R = 1$ case in which the S^2 profile shows a double peak structure at $t = 200$ after the transition to turbulence.

Unlike the profiles of N^2 and S^2 in which the layered structures are most profound in the $R = 0$ and 1 cases, the evolution of the gradient Richard number (Ri_g) in Fig. 7 shows the strongest layering in the $R = 0.5$ case. During the time period between $t = 140$ and 200, the values of Ri_g span a wide range from 0.25 to values larger than 1.25 across the shear layer. The central region of the shear layer has low values between 0.25 and 0.5 due the compensation of enhanced stratification by enhanced shear. Immediately above and below this low- Ri_g region, Ri_g increases rapidly up to values larger than 1.25. The large values are due to the notably small values of S^2 in these regions as shown in Fig. 6b. Moving outward toward the edges of the shear layer marked by $\pm I_u/2$, Ri_g decreases due to the enhanced shear in the transition layer. Among the three cases, the $R = 1$ case shows the least vertical variability of Ri_g after the transition to turbulence. At late time ($t = 200$), the Ri_g profile in the $Ri = 1$ case is more homogenous across the shear layer when compared to the other cases. Furthermore, the $R = 0$ case shows smaller values of Ri_g at late time on average across the shear layer.

4 Turbulence and the parameterization of mixing

In the previous section, we have illustrated similarities as well as structural differences among the cases with different R values but same $Ri_{b,0}$. In particular, we find the late-time values of the momentum thickness and shear layer thickness to be similar while the profiles of the N^2 , S^2 and Ri_g show considerable differences after the transition-to-turbulence

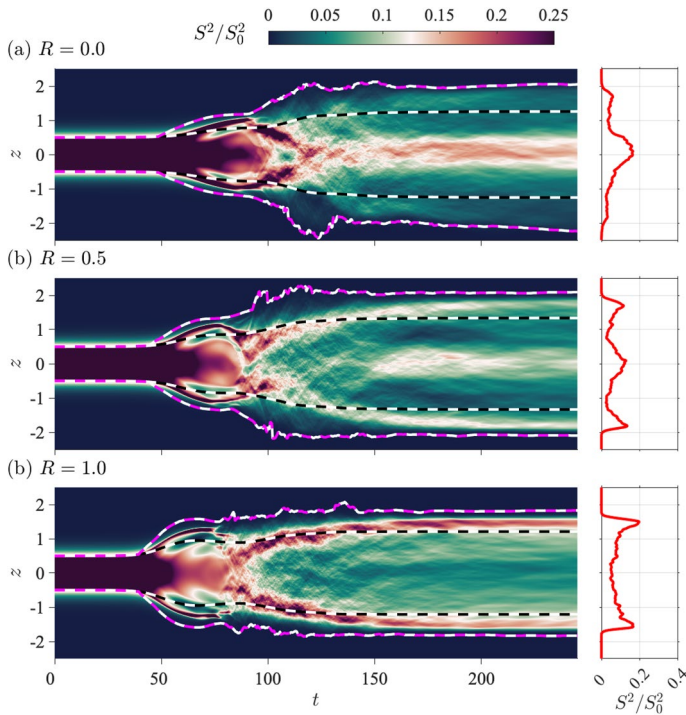


Fig. 6 Comparison of the normalized squared shear (S^2/S_0^2) among the simulated cases: **a** $R = 0$; **b** $R = 0.5$; and **c** $R = 1$. A layer of enhanced shear develops at the center of the shear layer in the $R = 0$ and 0.5 cases after the transition to turbulence but it does not appear in the $R = 1$ case. The shear enhancement near the transition layer becomes more prominent as the external stratification increases (i.e., increasing R value). The right panels show S^2 profile during the turbulence decay period at $t = 200$

period. In this section we diagnose the TKE budgets and the rate of mixing and also explore whether there is any difference in regard to mixing parameterization.

4.1 TKE budgets

Figure 8 compares the terms in the integrated TKE budgets. The integration is carried over the region $-5 \leq z \leq 5$. The integrated production shows double peaks in all three cases. The first peak corresponds to the development of the K–H billows while the second peak corresponds to turbulent entrainment after the shear layer transitions from K–H billows to turbulence. The first production peak (largest in the $R = 1$ case) is correlated with the growth rate of the FGM of the K–H instabilities and the growth rate of I_u during this period. In contrast, the second peak shows the highest value in the $R = 0$ case and the peak values decrease with increasing R . The smaller peak value during the turbulent entrainment period is due to the presence of the external stratification as R increases. Between the two peaks, the production in the $R = 1$ case shows a short period with negative values. The negative production is the result of counter-gradient momentum flux during the contraction period as has been discussed in detail by VanDine et al. [28]. The integrated buoyancy flux shown in Fig. 8b also indicates double negative peaks with a positive peak in-between for all three cases similar to

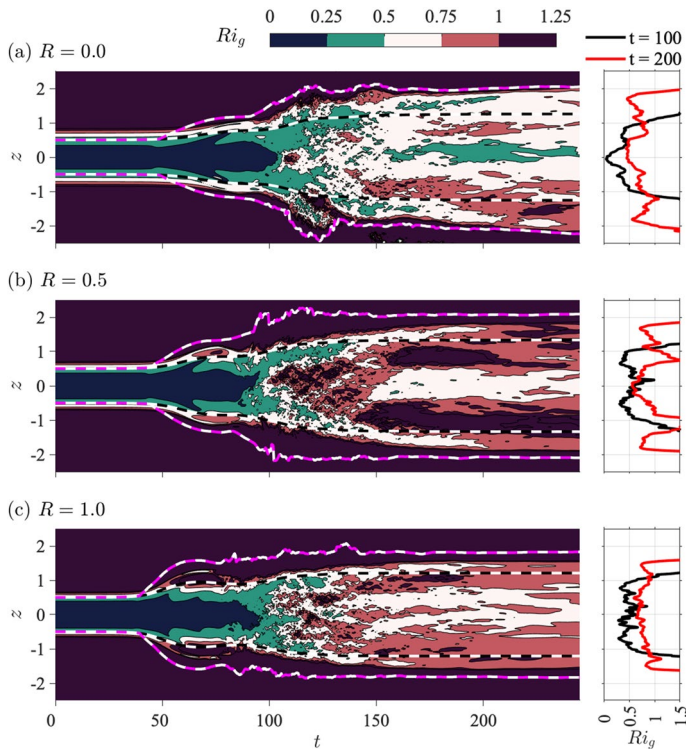


Fig. 7 Comparison of the gradient Richardson number (Ri_g) among the simulated cases: **a** $R = 0$; **b** $R = 0.5$; and **c** $R = 1$. After the transition to turbulence, the three cases show multiple layers with alternating small and large values of Ri_g . The variability across the shear layer is most profound in the $Ri = 0.5$ case while the Ri_g in the $R = 1$ case is more homogenous

the integrated production but with opposite sign. The negative values of the first peak increase in magnitude as the growth rate of the FGM and the value of R increase. For $R = 1$, the transfer from TKE to turbulent potential energy (TPE), i.e. $B < 0$, is strong in the pre-turbulent period and weaker in the latter turbulent period. The opposite trend is noted for $R = 0$.

The integrated TKE dissipation (ϵ) and scalar dissipation (χ) in Fig. 8c, d reach their peak values when the shear layer exhibits fully-developed turbulence. The peak values of ϵ and χ are largest in the $R = 0.5$ case while the $R = 1$ case shows the smallest values unlike the integrated production or buoyancy flux. Thus, while the peak value of the production is correlated with the R value, the peak TKE and scalar dissipation are not correlated with R .

We quantify the total amount of turbulent mixing by computing the bulk evolution of the TKE budget terms over the entire simulation. We integrate the TKE budget terms shown in Fig. 8 in time and compare the results of all cases in Fig. 9. In the figure, we also include the results from the vertical integration over the momentum thickness ($\pm l_u/2$), shear layer thickness (L_s), and integration over $-5 \leq z \leq 5$ region. Overall, all terms show a peak value in the $R = 0.5$ case. The differences when R varies from 0 to 1 are approximately 10-15%. The bulk energetic terms are considerably different

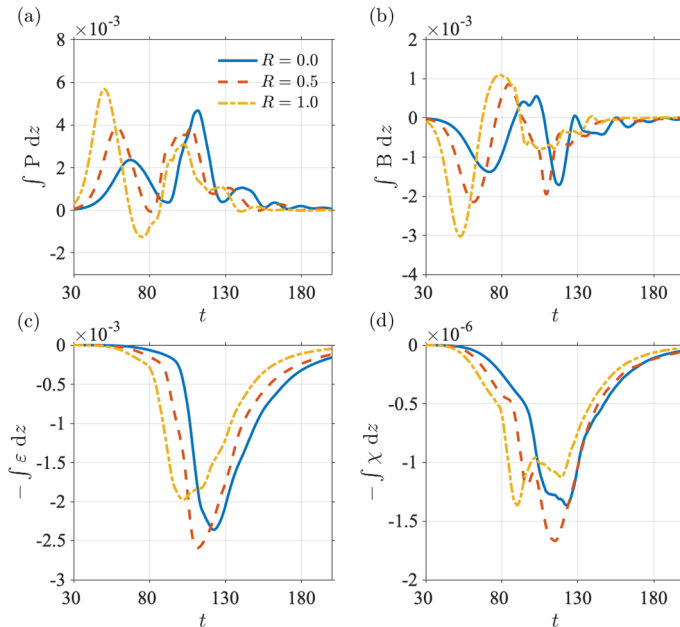


Fig. 8 Evolution of the integrated TKE budget: **a** TKE production (P); **b** buoyancy flux (B); and **c** TKE dissipation (ϵ). The scalar dissipation (χ) is shown in **d**. The integration is taken in the region $-5 \leq z \leq 5$

when comparing results from the integration over $\pm I_u/2$ with the integration over L_{sl} due to a significant amount of turbulence generated outside $\pm I_u/2$. When comparing the results from the integration over L_{sl} to that from the integration over $-5 \leq z \leq 5$, all of the terms are well captured except for the buoyancy flux in the $R = 0.5$ and 1 cases. The difference in the buoyancy flux is due to the presence of turbulence-generated internal waves in the regions outside L_{sl} .

It is important to highlight the significant amount of TKE present inside the TL. Figure 10 shows the accumulation of TKE and scalar budget terms inside the TL normalized by the corresponding amount inside the regions bounded by $\pm I_u/2$. In all cases, the dissipation inside the TL is between 20 to 30% of the amount inside the $\pm I_u/2$ region. Lewin and Caulfield [10] also observed a large amount of dissipation inside the TL for both the $R = 0$ and 1 flow configurations. According to Fig. 10, the accumulation inside the TL varies with R which suggests the turbulence inside the TL is influenced by the local shear and stratification. We recall that the turbulent entrainment also varies with R as indicated by the the growth of L_{sl} in Fig. 4b. Therefore, including dynamics at the edges of the shear layer is important to the quantification of the turbulent budgets independent of the type of density profile (i.e. independent of the precise value of R).

4.2 Mixing efficiency and its parameterization

While the turbulence evolution as indicated by the time evolution of terms in the integrated TKE budgets as well as the profiles of $N^2(z)$ and $S^2(z)$ are different in the simulated cases, it remains in question whether the different initial density profiles can affect the mixing

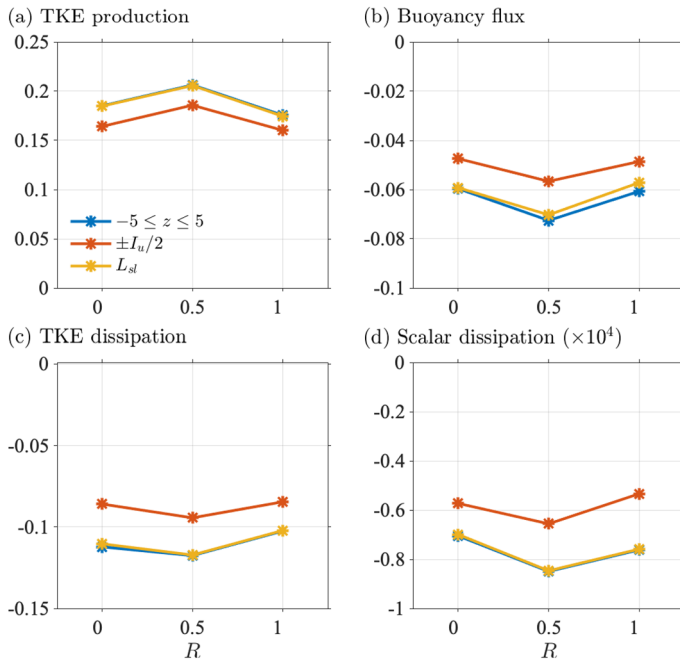


Fig. 9 Comparison of the integrated bulk values of **a** TKE production; **b** buoyancy flux; **c** TKE dissipation and **d** scalar dissipation among the simulated cases. The integration is carried out in space and in time. Results from different domains of spatial integration are shown: in the region $-5 \leq z \leq 5$, across the momentum thickness ($\pm I_u/2$), and across the shear layer thickness (L_{sl})

efficiency and its parameterization. The mixing efficiency is a measure of the rate of mixing wherein the change in background potential energy is compared to the total change in both the background potential energy and kinetic energy. The processes of computing the available and background potential energy require sorting of the 3D density fields. Instead, we follow the method of Scotti and White [23] to quantify the mixing efficiency $\bar{E} = \bar{\epsilon}_\rho / (\bar{\epsilon} + \bar{\epsilon}_\rho)$ where the shear-layer-averaged values of the TKE dissipation and TPE dissipation are defined as follows:

$$\bar{\epsilon} = \frac{1}{L_{sl}} \int_{L_{sl}} \epsilon \, dz \quad \text{and} \quad \bar{\epsilon}_\rho = \frac{g^2}{\rho_0^2 N^2} \frac{1}{L_{sl}} \int_{L_{sl}} \chi \, dz. \quad (10)$$

Here, the overbar indicates vertical averaging over the region marked by L_{sl} and \bar{N}^2 is the stratification averaged over the shear layer. It is noted that Vandine et al. (2021) use the constant initial stratification across the shear layer to compute $\bar{\epsilon}_\rho$ instead of the time-varying \bar{N}^2 for their $R = 1$ simulations. Such an approximation cannot be used in the $R = 0$ and 0.5 cases in the present study since the density gradient is non-uniform and its value changes considerably in time as shown in Fig. 11a. The value of N^2 decreases approximately by a factor of 3 in the $R = 0$ case while the value varies little in the $R = 1$ case. The reason for the little variation of \bar{N}^2 despite having significant turbulence and reduction of N^2 from its initial value in the core of the shear layer in the $R = 1$ case is due to the large values of stratification in the TL. Averaging the stratification over $\pm I_u/2$, which excludes

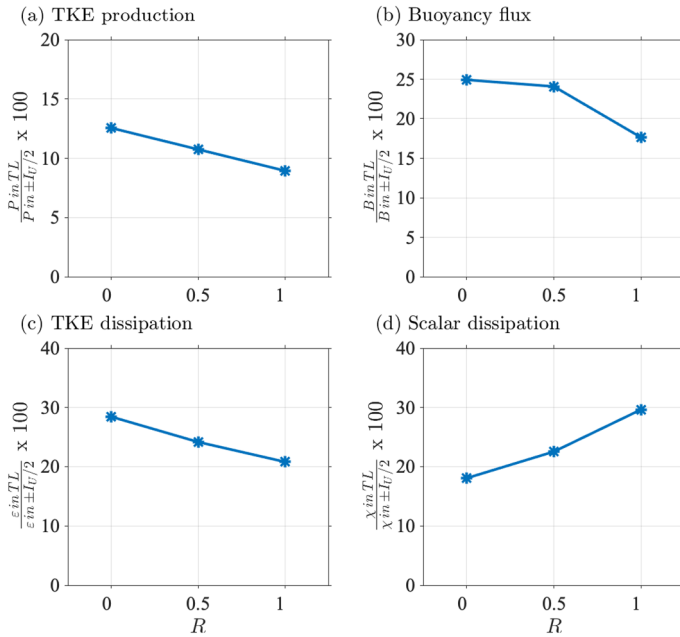


Fig. 10 Accumulation of TKE and scalar budget terms inside the transition layers: **a** production; **b** buoyancy flux; **c** TKE dissipation; and **d** scalar dissipation. The terms are plotted in percentage of the amount accumulated inside the region bounded by $\pm L_u/2$

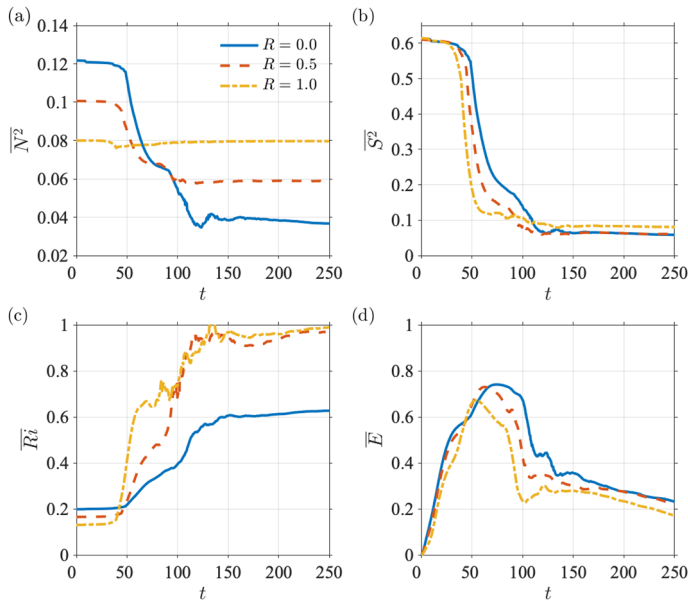


Fig. 11 Evolution of depth-averaged **a** stratification ($\overline{N^2}$); **b** squared shear ($\overline{S^2}$); **c** Richardson number (\overline{Ri}) and **d** mixing efficiency (\overline{E}). The quantities are averaged across the shear layer thickness (L_{sl})

the TL, shows significant reduction in $\overline{N^2}$ (not shown). In the $R = 0$ and 0.5 cases, the stratification at the edges of the shear layer is relatively smaller than at the central region of the shear layer so that the contribution from the transition layer is less significant. Due to the decrease of N^2 over time in the $R = 0$ and 0.5 cases, we will use the time-varying N^2 to compute $\bar{\epsilon}_\rho$ instead of the constant value at the initial time N_0^2 . Lewin and Caulfield [10] show that the use of Eq. (10) with N^2 in the denominator leads to a satisfactory agreement with mixing rate using density sorting.

The shear-layer average of the squared shear rate ($\overline{S^2}$) shown in Fig. 11b exhibits less difference among the cases than $\overline{N^2}$. The three cases show nearly equal values of $\overline{S^2}$ at late time. The different values of $\overline{N^2}$ among the cases result in a wide difference in the late-time value of the shear layer averaged Richardson number ($\overline{Ri} = \overline{N^2}/\overline{S^2}$) as shown in Fig. 11c. The late-time value is approximately equal to 0.6 in the $R = 0$ case while it is nearly equal to 1 in the $R = 0.5$ and 1 cases. The late-time values of \overline{Ri} in the $R = 0$ case agree with the result in Mashayek et al. [13] while the larger values in the $R = 0.5$ and 1 cases fall in line with the results of VanDine et al. [28]. The addition of uniform stratification into the background causes the late-time values of \overline{Ri} to increase. When we average the shear and stratification only over $\pm I_u/2$, the late-time values of \overline{Ri} are approximately equal to 0.85 and 0.8 in the $R = 0.5$ and 1 cases, respectively (not shown).

The mixing efficiency $\bar{E}(t)$ shown in Fig. 11d exhibits some differences among the cases. During the development of the K–H billows, \bar{E} increases similarly to reach a comparable peak value of approximately 0.7 in all cases. After that, \bar{E} decreases rapidly during the transition-to-turbulence period. During the turbulence decay period, \bar{E} gradually decreases toward the value of 0.2. We note that the difference in \bar{E} among the cases is smaller when we perform the integration over $\pm I_u/2$ (not shown).

The values of bulk mixing efficiency (E_{sl} and E_{lu}) are listed in Table 1. Here, E_{sl} is obtained as follows:

$$E_{sl} = \frac{\int \bar{\epsilon}_\rho dt}{\int \bar{\epsilon} + \bar{\epsilon}_\rho dt}, \quad (11)$$

where $\bar{\epsilon}$ and $\bar{\epsilon}_\rho$ are the shear-layer-averaged values as defined in Eq. (10) and the time integration is over the entire evolution of the shear layer. For comparison, we also compute the bulk mixing efficiency (E_{lu}) using the depth averaging over $\pm I_u/2$, i.e. replace L_{sl} with $\pm I_u/2$ and take the overbars in Eqs. (10) and (11) to be the average over $\pm I_u/2$ instead of L_{sl} . The values of E_{lu} are approximately equal to 0.36 in all the cases. However, the values of E_{sl} decrease with increasing R . The $R = 0$ case has the highest value of 0.42 and the $R = 1$ case has the lowest value of 0.31. Increasing R from 0 to 1 reduces the bulk mixing efficiency by nearly 25%.

An important application of mixing efficiency is its wide use to estimate turbulent thermal diffusivity (K_ρ) in oceanic flows. From the mixing efficiency, one can use Osborn model to estimate K_ρ from the relationship: $K_\rho = \Gamma \epsilon / N^2$ where $\Gamma = E/(1-E)$ is the flux coefficient. The typically used value for Γ is 0.2 which is the upper bound value suggested by Osborn [18]. Results from numerous DNS studies indicate that Γ should not be taken as a constant. Shih et al. [24] use DNS of homogenous stratified shear turbulence to suggest that Γ should vary with buoyancy Reynolds number ($Re_b = \bar{\epsilon}/\nu N^2$) in three different regimes: $\Gamma = 2Re_b^{-1/2}$ in an energetic regime ($Re_b > 100$), $\Gamma = 0.2$ in an intermediate regime ($7 \leq Re_b \leq 100$) and a diffusive regime ($Re_b < 7$) in which K_ρ is equal to the molecular value. More recently, Mashayek et al. [15] use DNS of stratified shear layers with $R = 0$ to suggest two regimes: a buoyancy-dominated regime in which $\Gamma \propto Re_b^{-1/2}$ and

a shear-dominated regime where $\Gamma \propto \overline{Re}_b^{-1/2}$ and the two regimes are separated by a transitional value of \overline{Re}_b ranging between 100 and 300. However, VanDine et al. [28] simulate stratified shear layers with $R = 1$ and show that the \overline{Re}_b -dependence is different from what has been suggested by Mashayek et al. [15]. They conclude that \overline{Re}_b -dependence of Γ is more aligned with what has been suggested by Shih et al. [24].

Figure 12a–c show the relationship between Γ and \overline{Re}_b in the three cases with different R values in the present study. We note that the figure only shows the time period after the turbulence in the shear layer has fully developed (i.e., after the integrated dissipation in Fig. 8c has peaked). Results from averaging Γ and \overline{Re}_b over $\pm L_u/2$ and L_{sl} are both included in the figure for comparison. The Γ - \overline{Re}_b relationship using averaging over L_{sl} in the $R = 1$ case exhibits similar trends as in Shih et al. [24] particularly in the following aspects. First, Γ decreases with increasing \overline{Re}_b when \overline{Re}_b exceeds values of approximately 100. Second, Γ remains relatively constant for values of \overline{Re}_b between 10 and 100. Finally, Γ gradually decreases when \overline{Re}_b decreases below 10. In the $R = 0$ and 0.5 cases, the Γ - \overline{Re}_b relationship indicates good agreement with Mashayek et al. [15] for the range of \overline{Re}_b between 100 and 300. However, Γ decreases at a considerably slower rate than what has been suggested by Mashayek et al. [15] when \overline{Re}_b decreases to less than 100. The difference between our result in the $R = 0$ case and what has been suggested by Mashayek et al. [15] is possibly due to the different methods employed to compute Γ . Mashayek et al. [15] perform 3D sorting of the density field to compute the change in background potential energy while we use the approximation of potential energy mixing by $\bar{\epsilon}_\rho$ [23]. Changing the averaging

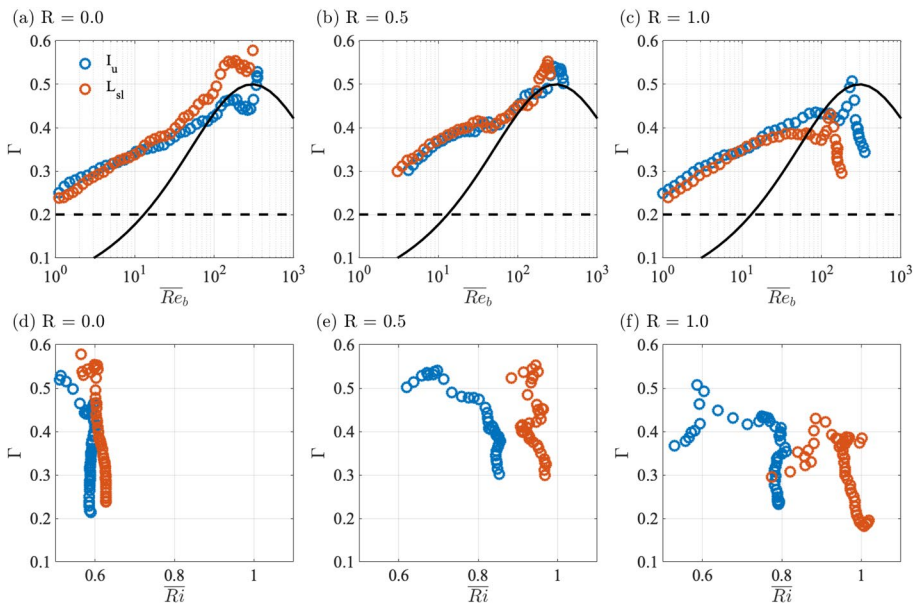


Fig. 12 a–c Dependence of flux coefficient (Γ) on a–c depth-averaged buoyancy Reynolds number (\overline{Re}_b) and d–f depth-averaged Richardson number (\overline{Ri}) during the turbulence decay period in the three simulated cases. Quantities that are averaged across the momentum thickness ($\pm L_u/2$) are shown by blue markers while those averaged across the shear layer thickness (L_{sl}) are shown by red markers. Dashed black lines in panels (a–c) denotes the $\Gamma = 0.2$ parameterization in Osborn [18]. Solid black lines denotes the upper bound of the parameterization suggested by Mashayek et al. [15]

method from L_{sl} to $\pm I_u/2$ does not affect $\Gamma\text{-}\overline{Re_b}$ relationship in the $\overline{Re_b} < 100$ range as shown in Fig. 12a–c. Figure 12d, e show the $\Gamma\text{-}Ri$ relationship during the fully-developed turbulence period. In the $Ri = 0$ case, Γ decreases monotonically at a constant value of $Ri \approx 0.6$, slightly larger than the critical value of 0.5 that Salehipour et al. [22] suggest to shut off turbulent mixing. In the $R = 0.5$ and 1 cases, Γ remains significantly large even when Ri approaches near unity. The simulations in VanDine et al. [28] show the late-time values of Ri to be much larger than unity.

Beside the $\overline{Re_b}$ dependence, Smyth et al. (2001) suggest that Γ can be used to parameterize the thermal diffusivity using the ratio ($R_{OT} = L_O/L_T$) between the Ozmidov scale ($L_O = (\overline{\epsilon}/N^2)^{1/2}$) and the Thorpe scale (L_T). The motivation behind the parameterization is to relate Γ to the relative evolution of two length scales: one representative of the energy-containing scales like that of the K–H billows and one in the inertial subrange of the TKE spectrum like the turbulent eddies during the transition-to-turbulence period. In ocean applications, the Thorpe scale is taken to be the length scale of large-eddy overturns while the Ozmidov scale provides a good estimate of the inertial subrange. From the scaling argument, if we take $K_p \propto \Delta UL_T$, then the Osborn model yields:

$$\Gamma = \frac{K_p \overline{N^2}}{\overline{\epsilon}} \propto \frac{\Delta U}{N^2} \frac{L_T}{L_O^2}, \quad (12)$$

in which Γ varies as a function of the mixing length scale L_O and two energy-containing scales L_T and $\Delta U/N^2$. Now, if we further assume $\overline{\epsilon} \propto \Delta U^3/L_O$ which implies $L_O \propto \Delta U/N^2$, then Eq. (12) suggests:

$$\Gamma \propto (L_O/L_T)^{-1} \propto R_{OT}^{-1}. \quad (13)$$

In contrast, if we assume $\overline{\epsilon} \propto \Delta U^3/L_T$ which leads to $L_O \propto (\Delta U/\overline{N^2})^{3/2} L_T^{-1/2}$, then Eq. (12) indicates:

$$\Gamma \propto (L_O/L_T)^{-4/3} \propto R_{OT}^{-4/3}. \quad (14)$$

Recent analysis using observational data by Ijichi and Hibiya [7] indicates $\Gamma \propto R_{OT}^{-4/3}$ while Mashayek et al. [16] suggest a parameterization scheme that involves both of the R_{OT}^{-1} and $R_{OT}^{-4/3}$ scalings. Using DNS of stratified shear layers at low Reynolds numbers, Smyth et al. [26], however, suggest $\Gamma \propto R_{OT}^{-0.63}$. It is of interest to examine which scaling of Γ (equivalently, $\overline{\epsilon}$) prevails in the present study.

We use the Ellison scale ($L_E = \langle \rho'^2 \rangle^{1/2} / \partial \langle \rho \rangle / \partial z$) in place of the Thorpe scale to test the scaling arguments. Figure 13a–c illustrate the evolution of L_E and L_O in the simulated cases. It is noted that the length scales are averaged over the shear layer and we include both the averaging over $\pm I_u/2$ and L_{sl} in the figure. As the K–H billows develop, L_E increases drastically and then it decreases when the shear layer transitions into turbulence. While the K–H billows grow to different sizes in the cases with different R as indicated by I_u and L_{sl} , the peak value of L_E is similar among the cases for the both methods of averaging. It is interesting that despite the differences in stratification profile, in billow size and in integrated buoyancy flux (e.g., Fig. 8b) among the cases with $R = 0, 0.5$ and 1 during this period, the billows generate a similar value of maximum Ellison scale in these cases. The peak values of L_E are significantly different between the two averaging methods. When the averaging is done over L_{sl} , the peak value of L_E is smaller due to contribution of the enhanced stratification in the TL. The Ozmidov scale

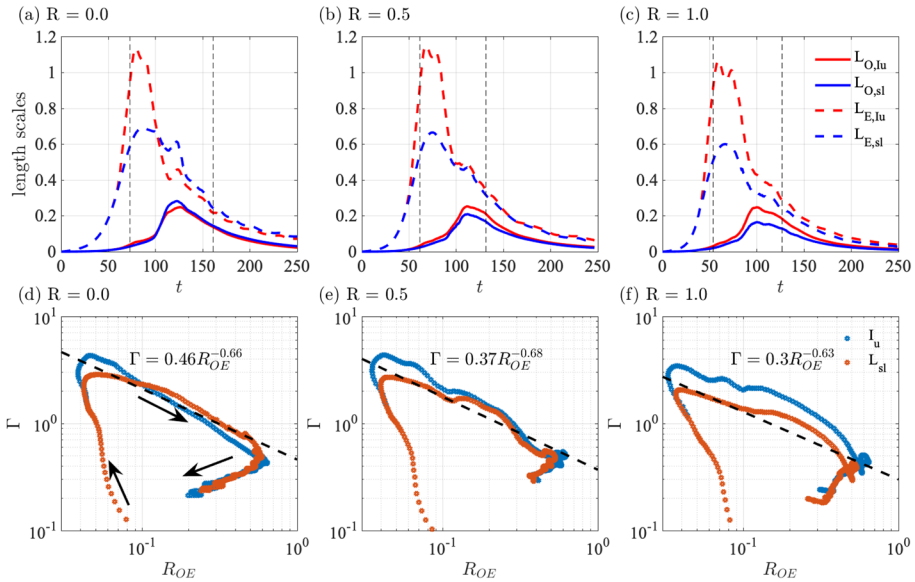


Fig. 13 **a–c** Evolution of Ozmidov and Ellison length scales and **d–f** the parameterization of flux coefficient (Γ) using the R_{OE} ratio between the two length scales. Each panel includes results for both methods of averaging, i.e., across L_{sl} and $\pm I_u/2$. Arrows in panel **d** indicate time progression through three regimes in Γ - R_{OE} relationship. The dashed lines in **d–f** denote a power law fit using the values of Γ and R_{OE} that are averaged over L_{sl} . The vertical dashed lines in panels **a–c** mark the second regime during which R_{OE} increases as Γ decreases

begins to grow at a later time when the Ellison scale has decreased. The peak value of L_O is comparable among the cases when the averaging is over $\pm I_u/2$. When comparing L_O results between the two methods of averaging, the $R = 0$ case shows larger values when averaging over L_{sl} while the other two cases show smaller values. The reason is that there exists significant dissipation but the stratification is relatively weak in the TL in the $R = 0$ case. The enhanced stratification in the transition layer in $R = 0.5$ and 1 cases causes L_O to decrease.

The relationship between Γ and the ratio between the Ozmidov and Ellison length scale (R_{OE}) is shown in Fig. 13d, e. Unlike Fig. 12 in which the Γ - Re_b relationship is only shown during the period of the fully-developed turbulence, the entire evolution of the shear layer is included in Fig. 13. The Γ - R_{OE} relationship in all simulated cases exhibits three evolutionary regimes. During the first regime when the K–H billows develop, R_{OE} decreases to its minimum value as Γ increases drastically. As R_{OE} increases during the second regime (marked by the dashed lines in Fig. 13a–c), Γ decreases monotonically. Finally, R_{OE} and Γ decreases in the third regime at late time. Figure 13d–f show the correlation between Γ and R_{OE} . A power law fit using the values of Γ and R_{OE} averaged across the L_{sl} during the second regime results in $\Gamma = 0.46 R_{OE}^{-0.66}$ in the $R = 0$ case, $\Gamma = 0.37 R_{OE}^{-0.68}$ in the $R = 0.5$ case, and $\Gamma = 0.3 R_{OE}^{-0.63}$ in the $R = 1$ case. When the average is across $\pm I_u/2$, the results are: $\Gamma = 0.32 R_{OE}^{-0.85}$, $\Gamma = 0.36 R_{OE}^{-0.79}$, $\Gamma = 0.42 R_{OE}^{-0.66}$ for $R = 0, 0.5$ and 1 , respectively. Results in Lewin and Caulfield [10] show the scaling of $\Gamma \propto R_{OE}^{-1}$ during the early period of the second regime and the scaling of $\Gamma \propto R_{OE}^{-4/3}$ during the later period of the second regime. The different scaling observed in

the present study perhaps is due to how the K–H billows in Lewin and Caulfield [10] are allowed to develop first in 2D simulations and then extended into 3D simulations when the billows reach its maximum sizes. In the present study, the K–H billows are allowed to develop in 3D simulations since the beginning and the secondary instabilities are observed to develop before the K–H billows grow to its full size. A different transition to turbulence can result in different Γ - R_{OE} relationship.

Furthermore, if we assume the behavior of R_{OE} and R_{OT} is similar during the second regime as suggested by Lewin and Caulfield [10], then the Γ - R_{OE} relationship in the present study agrees best with the scaling of $\Gamma = 0.33R_{OT}^{-0.63}$ found in Smyth et al. [26]. We do not observe the scaling of either $\Gamma \propto R_{OT}^{-1}$ or $R_{OT}^{-4/3}$ as indicated in Ijichi and Hibiya [7], Mashayek et al. [16], and Lewin and Caulfield [10]. It has been hypothesized that the scaling $\Gamma \propto R_{OT}^{-0.63}$ in Smyth et al. [26] is possibly due to the effect of a low Reynolds number of 5000 [7]. Here, we observe the same scaling as in Smyth et al. [26] at $Re = 24,000$.

5 Discussion and conclusions

In the present study, we have used DNS to investigate the evolution of stratified shear layers with different profiles of density gradient at a high Reynolds number. The shear layer has constant initial bulk Richardson number ($Ri_{b,0}$); however, the density gradient distribution is varied among the three cases by varying the density gradient distribution parameter R . One case has the two-layer density profile ($R = 0$) while another case has uniform density gradient throughout the fluid ($R = 1$). The third case has a composite of the two-layer profile and the linear density gradient profile ($R = 0.5$). The three cases have the same $Ri_{b,0}$ but the minimum gradient Richardson number (Ri_{min}) and the shear-layer-averaged value ($\overline{Ri_0}$) vary among the cases. The value of $Ri_{min} = Ri_{b,0}(1 - 0.5R)$ decreases with increasing R . We are interested in how the evolution of the shear layer correlates with the different Richardson number parameters.

The evolution of the shear layer in all cases shows the development of K–H shear instabilities and billows, the transition to turbulence by secondary instabilities and subsequent decay of turbulence. Results from linear stability analysis indicate the K–H instabilities grow the fastest in the $R = 1$ case and the growth rate decreases with decreasing R value (equivalently increasing Ri_{min}). The evolution of the momentum thickness (I_u) and shear layer thickness (L_{sl}) indicate the shear layer evolves through two distinct periods of significant growth. The first period corresponds to the enlargement of the KH billows while the second period corresponds to the turbulent entrainment of fluid from the external ambient. During the first period, the growth rate of I_u and L_{sl} increases with increasing R values. The maximum size of the K–H billows is observed in the $R = 1$ case. After the first period of growth, secondary instabilities occur on the eyelids of the K–H billows and they trigger the shear layer to transition to turbulence. Unlike the first period, the growth rate during the turbulent entrainment period is slower when R has larger values. The reason for the slower growth rate is due to the external stratification in the ambient when R increases. Overall, the momentum thickness and the shear layer thickness are comparable among the cases at late time suggesting $Ri_{b,0}$ exerts control on these late-time thickness metrics. We note that the shear layer in the present study has a fixed value of $Ri_{b,0} = 0.16$ in all cases. Whether the insensitivity of the momentum thickness and shear layer thickness at late time to R persists at different values of $Ri_{b,0}$ requires future study.

After the K–H billows have developed, transition layers with peak values of shear and stratification form at the edges of the shear layer in the $R = 0.5$ and 1 cases, namely the cases where stratification exists in the ambient. The $R = 0$ case does not exhibit peak value of shear and stratification at the edges where there exists enhanced turbulence. Significant turbulent mixing is found in the TL defined as the region outside $\pm I_u/2$ and bounded by L_{sl} in all cases. The TKE dissipation rate integrated over the TL can be as large as 28% of the amount found inside $\pm I_u/2$. Therefore, it is important to use L_{sl} as physical boundaries of the shear layer instead of $\pm I_u/2$ as was typically used in previous studies.

The bulk mixing efficiency (integrated over space and time) is found to be approximately equal to 0.36 in all cases when only the mixing over $\pm I_u/2$ is taken into consideration. When the mixing in the TL is taken into account, the bulk mixing efficiency has a peak value of 0.42 in the $R = 0$ case which agrees with the result of Mashayek et al. [13]. The mixing efficiency decreases to 0.31 in the $R = 1$ case. While the flux coefficient (Γ) during the period of fully-developed turbulence in the $R = 1$ case shows three regimes of mixing efficiency based on the Re_b parameterization similar to the study of Shih et al. (2002), Γ in the $R = 0$ and $R = 0.5$ cases decreases monotonically with decreasing Re_b when $Re_b < 100$ as suggested by Mashayek et al. [15]. However, we observe a slower decrease with Re_b and thus Γ remains at larger values relative to Mashayek et al. [15]. When we examine the relationship between Γ and the ratio between the Ozmidov and Ellison length scales, we find that $\Gamma \propto R_{OE}^{-0.66}$ similar to the scaling of $\Gamma \propto R_{OT}^{-0.63}$ found in Smyth et al. [26].

Overall, the density gradient distribution parameter R influences the layering structures in N^2 , S^2 and Ri_g inside the shear layer after the transition to turbulence. The TL (a layer with enhanced stratification and shear) develops at the edges of the shear layer when $R \neq 0$. As the shear layer evolves, the shear-layer-averaged stratification (N^2) exhibits a significant reduction when R is small. The change in N^2 becomes smaller when R increases. The shear-layer-averaged Richardson number Ri at late time is larger when $R \neq 0$. Nonetheless, the shear layer thickness (I_u and L_{sl}) at late time is not influenced by the value of R although the details of the shear and stratification profiles are influenced. VanDine et al. [28] investigate the uniform stratification case ($R = 1$) with Ri_{min} varying between 0.04 and 0.2, equivalently $Ri_{b,0}$ between 0.08 and 0.4, and they find the TL develop in all cases. As $Ri_{b,0}$ increases in that study, both I_u and L_{sl} have a smaller value at late time and the net turbulent dissipation is also smaller. All cases in VanDine et al. [28] exhibit TL. The Γ – Re_b relationship discussed here for $Ri_{b,0} = 0.16$ and $R = 1$ are also found for all the different values of $Ri_{b,0}$ in that study.

Acknowledgements The authors thank W. D. Smyth for providing the LSA code. This work was supported by the National Science Foundation (Grant No. OCE-1851390).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Brucker KA, Sarkar S (2010) A comparative study of self-propelled and towed wakes in a stratified fluid. *J Fluid Mech* 652:373–404
2. Caulfield CP, Peltier WR (2000) The anatomy of the mixing transition in homogeneous and stratified free shear layers. *J Fluid Mech* 413:1–47
3. Drazin PG (1958) The stability of a shear layer in an unbounded heterogeneous inviscid fluid. *J Fluid Mech* 4(2):214–224
4. Fritts DC, Baumgarten WK, Werne J, Lund T (2014) Quantifying Kelvin–Helmholtz instability dynamics observed in noctilucent clouds: 2. Modeling and interpretation of observations. *J Geophys Res Atmos* 119:9359–9375
5. Garanaik A, Venayagamoorthy SK (2019) On the inference of the state of turbulence and mixing efficiency in stably stratified flows. *J Fluid Mech* 867:323–333
6. Hazel P (1972) Numerical studies of the stability of inviscid stratified shear flows. *J Fluid Mech* 51:39–61
7. Ijichi T, Hibiya T (2018) Observed variations in turbulent mixing efficiency in the deep ocean. *J Phys Oceanogr* 48:1851–1830. <https://doi.org/10.1175/JPO-D-17-0275.1>
8. Jacobitz FG, Sarkar S, VanAtta CW (1997) Direct numerical simulations of the turbulence evolution in a uniformly sheared and stably stratified flow. *J Fluid Mech* 342:231–261
9. Kaminski AK, Smyth WD (2019) Stratified shear instability in a field of pre-existing turbulence. *J Fluid Mech* 862:639–658
10. Lewin S, Caulfield CP (2021) The influence of far field stratification on shear-induced turbulent mixing. *J Fluid Mech*. <https://doi.org/10.1017/jfm.2021.755>
11. Mashayek A, Peltier WR (2012) 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
12. Mashayek A, Peltier WR (2013) Shear-induced mixing in geophysical flows: does the route to turbulence matter to its efficiency? *J Fluid Mech* 725:216–261
13. Mashayek A, Caulfield CP, Peltier WR (2013) Time-dependent, non-monotonic mixing in stratified turbulent shear flows: implications for oceanographic estimates of buoyancy flux. *J Fluid Mech* 736:570–593
14. Mashayek A, Caulfield C, Peltier W (2017) Role of overturns in optimal mixing in stratified mixing layers. *J Fluid Mech* 826:522–552
15. Mashayek A, Salehipour H, Bouffard CE, Caulfield CP, Ferrari R, Nikurashin M, Peltier WR, Smyth WD (2017) Efficiency of turbulent mixing in the abyssal ocean circulation. *Geophys Res Lett* 44:6296–6303
16. Mashayek A, Caulfield CP, Alford MH (2021) Goldilocks mixing in oceanic shear-induced turbulent overturns. *J Fluid Mech*. <https://doi.org/10.1017/jfm.2021.740>
17. Mater BD, Venayagamoorthy SK (2014) A unifying framework for parameterizing stably stratified shear-flow turbulence. *Phys Fluids* 26(3):036601
18. Osborn T (1980) Estimates of the local rate of vertical diffusion from dissipation measurements. *J Phys Oceanogr* 10:80–89
19. Pham HT, Sarkar S (2010) Transport and mixing of density in a continuously stratified shear layer. *J Turbul* 11:1–23
20. Pham HT, Sarkar S, Brucker KA (2009) Dynamics of a stratified shear layer above a region of uniform stratification. *J Fluid Mech* 630:191–223
21. Salehipour H, Peltier WR, 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
22. Salehipour H, Peltier WR, Whalen CB, MacKinnon JA (2016) A new characterization of the turbulent diapycnal diffusivities of mass and momentum in the ocean. *Geophys Res Lett* 43(7):3370–3379
23. Scotti A, White B (2014) Diagnosing mixing in stratified turbulent flows with a locally defined available potential energy. *J Fluid Mech* 740:114–135
24. Shih LH, Koseff JR, Ivey GN, Ferziger JH (2005) Parameterization of turbulent fluxes and scales using homogeneous sheared stably stratified turbulence simulations. *J Fluid Mech* 525:193–214
25. Smyth WD, Moun JN (2000) Length scales of turbulence in stably stratified mixing layers. *Phys Fluids* 12(6):1327–1342
26. Smyth WD, Moun JN, Caldwell DR (2001) The efficiency of mixing in turbulent patches: inferences from direct simulations and microstructure observations. *J Phys Oceanogr* 31:1969–1992
27. Smyth WD, Moun JN, Nash JD (2011) Narrowband oscillations at the upper equatorial ocean. Part II. Properties of shear instabilities. *J Phys Oceanogr* 41:412–428

28. VanDine A, Pham HT, Sarkar S (2021) Turbulent shear layers in a uniformly stratified background: Dns at high Reynolds number. *J Fluid Mech* 916(A42):1–35
29. Watanabe T, Nagata K (2021) Large-scale characteristics of a stably stratified turbulent shear layer. *J Fluid Mech*. <https://doi.org/10.1017/jfm.2021.773>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.