

*Annual Review of Fluid Mechanics*

# Layering, Instabilities, and Mixing in Turbulent Stratified Flows

C.P. Caulfield<sup>1,2</sup>

<sup>1</sup>BP Institute, University of Cambridge, Cambridge CB3 0EZ, United Kingdom;  
email: cpc12@cam.ac.uk

<sup>2</sup>Department of Applied Mathematics and Theoretical Physics, University of Cambridge,  
Cambridge CB3 0WA, United Kingdom

ANNUAL  
REVIEWS **CONNECT**

[www.annualreviews.org](http://www.annualreviews.org)

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Fluid Mech. 2021. 53:113–45

First published as a Review in Advance on  
August 17, 2020

The *Annual Review of Fluid Mechanics* is online at  
[fluid.annualreviews.org](http://fluid.annualreviews.org)

<https://doi.org/10.1146/annurev-fluid-042320-100458>

Copyright © 2021 by Annual Reviews.  
All rights reserved

## Keywords

layering, instabilities, mixing, stratification, turbulence

## Abstract

Understanding how turbulence leads to the enhanced irreversible transport of heat and other scalars such as salt and pollutants in density-stratified fluids is a fundamental and central problem in geophysical and environmental fluid dynamics. This review discusses recent research activity directed at improving community understanding, modeling, and parameterization of the subtle interplay between energy conversion pathways, instabilities, turbulence, external forcing, and irreversible mixing in density-stratified fluids. The conceptual significance of various length scales is highlighted, and in particular, the importance is stressed of overturning or scouring in the formation and maintenance of layered stratifications, i.e., robust density distributions with relatively deep and well-mixed regions separated by relatively thin interfaces of substantially enhanced density gradient.

## 1. INTRODUCTION

The atmosphere and the world's oceans are typically statically stably stratified, and so fluid parcels perturbed vertically feel a restoring (buoyancy) force. The fluid velocity also typically varies in time and space, and the interplay between such velocity variations and the stratification leads to rich and often counter-intuitive dynamical behavior. In particular, modeling the properties of turbulence and the associated irreversible mixing where external large-scale forcing and stratification play central, leading-order roles is a key challenge in geophysical fluid dynamics. Of course, there has been a huge amount of research into various aspects of turbulence in stratified fluids, and a sequence of influential reviews from a variety of viewpoints have shaped our understanding of these flows (e.g., Linden 1979, Fernando 1991, Ivey & Imberger 1991, Riley & Lelong 2000, Peltier & Caulfield 2003, Wunsch & Ferrari 2004, Ivey et al. 2008, Ferrari & Wunsch 2009).

Indeed, much of the recent focus in the fluid dynamical study of the properties of turbulence in stratified fluids has been on improving understanding of heat and scalar transport in the world's oceans. A detailed and insightful review discussing many of the challenges facing this objective has been presented by Gregg et al. (2018). As they note, major and indeed leading-order areas of uncertainty and controversy still remain in both the fluid dynamical and oceanographic literatures. There is not even consensus as to whether the term “stratified turbulence” should be reserved for the description of a specific strongly stratified and anisotropic asymptotic flow regime. It is also not settled whether idealized flows considered in the laboratory and in simulations are particularly relevant to the description of geophysical flows.

Technical challenges abound. The presence of even a weak stratification in a gravitational field inevitably introduces anisotropy to the flow. Furthermore, turbulence in stratified fluids is commonly observed to be spatially and temporally intermittent. Fortunately, due to relatively recent major advances in theory, experimentation, and numerical simulation, there is growing evidence that the relevance gap can be at least partially bridged. The focus of this review is principally on fluid dynamical aspects of these recent advances, and on highlighting the importance of characteristic length scales and velocity shear in the evolution of turbulent flows in stratified fluids.

In Section 2, some of the central mathematical and physical challenges in modeling stratified turbulent flow are presented, various useful quantities are defined, and key simplifying assumptions are spelled out. In Section 3, instability mechanisms by which (sheared) stratified flows undergo the transition to turbulence are discussed, and then in Section 4, the ensuing mixing properties of such inherently freely evolving and transient flows are considered. Section 5 discusses certain aspects of turbulence and mixing in forced stratified flows, focusing in particular on identifying circumstances under which stratified turbulence can or cannot be maintained for extended periods by either wall or even (artificial) body forcing. In all cases, we highlight the emerging importance of layering for the properties of mixing and turbulence, i.e., the development or maintenance of relatively well-mixed deep layers separated by relatively thin interfaces of increasing density gradient. Two qualitatively different kinds of mixing are also highlighted, which may be simplistically classified as either overturning of weak interfaces or scouring of strong interfaces. Finally, conclusions are drawn at the end of the review, posing some key questions for the future.

## 2. MATHEMATICAL AND PHYSICAL CHALLENGES

### 2.1. Assumptions and Equations

The geophysical coordinate system is used throughout this review, so that the gravitational acceleration  $\mathbf{g} = -g\hat{\mathbf{z}}$  points downward in the  $z$ -direction. The velocity field,  $\mathbf{u}(\mathbf{x}, t)$ , is required to be divergence-free,  $\nabla \cdot \mathbf{u} = 0$ , and the fluid density is required to satisfy a linear, single-component

equation of state with respect to some stratifying agent such as heat or salt concentration with molecular diffusivity  $\kappa$ , thus excluding any possibility of double-diffusive effects, which are naturally very important in many circumstances in the ocean (Schmitt 1994) and in astrophysics (Garaud 2018). Density variations are assumed to be sufficiently small [ $|\rho(\mathbf{x}, t) - \rho_r| \ll \rho_r$ , where  $\rho_r$  is a reference density], and the Boussinesq approximation can be made, eliminating many significant complicating effects of thermodynamics in admissible energy conversion pathways, as highlighted by Tailleux (2009).

If the flow is turbulent, a classical Reynolds decomposition of the velocity,  $\mathbf{u} = \langle \mathbf{u} \rangle + \mathbf{u}' \equiv \mathbf{U} + \mathbf{u}'$ , and the density,  $\rho = \langle \rho \rangle + \rho'$  (where the angle brackets here denote formally an ensemble average, which of course is often replaced by a temporal and/or spatial average), leads to the (specific) turbulent kinetic energy equation for a Boussinesq stratified fluid:

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{1}{2} \langle u'_i u'_i \rangle \right) &= - \overbrace{2\nu \langle s'_{ij} s'_{ij} \rangle}^{\mathcal{E}} - \overbrace{\frac{g}{\rho_r} \langle \rho' w' \rangle}^{\mathcal{B}} + \overbrace{\left( -\langle u'_i u'_j \rangle S_{ij} \right)}^{\mathcal{P}} \\ &\quad - \underbrace{\frac{\partial}{\partial x_i} \left[ \langle p' u'_i \rangle + \left\langle (u'_i + U_i) \frac{u'_j u'_j}{2} \right\rangle - \nu \left( \frac{\partial}{\partial x_i} \left\langle \frac{u'_j u'_j}{2} \right\rangle + \frac{\partial}{\partial x_j} \langle u'_i u'_j \rangle \right) \right]}_{\nabla \cdot \mathbf{J}}, \end{aligned} \quad 1.$$

where  $s'_{ij} = (1/2)(\partial u'_i / \partial x_j + \partial u'_j / \partial x_i)$  and  $S_{ij} = (1/2)(\partial U_i / \partial x_j + \partial U_j / \partial x_i)$  are respectively the perturbation and ensemble-averaged deformation tensor,  $p'$  is the perturbation pressure,  $g$  is the acceleration due to gravity,  $\nu$  is the kinematic viscosity, and the Einstein summation convention is used throughout. Here,  $\mathcal{B}$  is the (vertical) turbulent density flux quantifying the generally reversible exchange between the potential energy and kinetic energy reservoirs, and  $\mathcal{E}$  is the turbulent kinetic energy dissipation rate of the turbulent kinetic energy (density),  $\mathcal{K}'$ , which within the Boussinesq approximation with a linear equation of state quantifies the rate of irreversible energy conversion lost into the internal energy reservoir. In general,  $\mathcal{K}'$  may be forced by the turbulence production,  $\mathcal{P}$ , converting kinetic energy in the ensemble-averaged mean flow into the perturbation velocity field, and transported by the divergence of the flux,  $\mathbf{J}$ .

Through analogous consideration of the density advection–diffusion equation (with density diffusivity  $\kappa$ ), the evolution equation for the (specific) potential energy, relative to some reference level (e.g.,  $z = 0$ ), is

$$\frac{\partial}{\partial t} \left\langle \frac{g}{\rho_r} \rho z \right\rangle \equiv \frac{\partial}{\partial t} \langle \mathcal{K}_P \rangle = \frac{g}{\rho_r} \langle \rho' w' \rangle - \nabla \cdot (\langle \mathbf{u} \mathcal{K}_P \rangle) + \kappa \nabla^2 \langle \mathcal{K}_P \rangle - 2\kappa \frac{\partial}{\partial z} \left( \left\langle \frac{g\rho}{\rho_r} \right\rangle \right), \quad 2.$$

which shows the exchange between the potential energy and kinetic energy reservoirs via the density flux term,  $\mathcal{B}$ , using the implicit assumption that  $\langle w \rangle = \langle w' \rangle = 0$ .

A significant alternative formulation to consider arises from the equation for the turbulent buoyancy variance, where the negative of the reduced gravity  $g'$  is termed the “buoyancy” in this context, i.e.,  $b \equiv -g\rho'/\rho_r$ . Multiplying the density advection–diffusion equation by  $g^2 \rho'/\rho_r^2$  and averaging lead to an equation for the evolution of the scaled buoyancy variance:

$$\begin{aligned} \frac{1}{N^2} \frac{\partial}{\partial t} \left\langle \frac{b^2}{2} \right\rangle &= \overbrace{\frac{g}{\rho_r} \langle \rho' w' \rangle}^{\mathcal{B}} - \overbrace{\frac{\kappa}{N^2} \left\langle \frac{\partial b}{\partial x_i} \frac{\partial b}{\partial x_i} \right\rangle}^{\mathcal{X}} - \overbrace{\frac{1}{N^2} \frac{\partial}{\partial x_i} \left[ U_i \left\langle \frac{b^2}{2} \right\rangle + \left\langle \frac{u'_i b^2}{2} \right\rangle - \kappa \left\langle b \frac{\partial b}{\partial x_i} \right\rangle \right]}^{\nabla \cdot \mathbf{J}_\rho}, \\ N^2(z, t) &\equiv -\frac{g}{\rho_r} \frac{\partial}{\partial z} \langle \rho \rangle(z, t), \end{aligned} \quad 3.$$

which defines an appropriately scaled destruction rate of buoyancy variance,  $\chi$ ; transport flux,  $\mathbf{J}_\rho$ ; and buoyancy frequency,  $N$ , which is associated with the ensemble-averaged density and is implicitly assumed to be a function only of  $z$  and possibly  $t$ .

A key insight is thus that, if transport terms have no net effect (e.g., because the system is closed or periodic boundary conditions are imposed), the appropriately scaled buoyancy variance is increased (or decreased) through the action of  $\mathcal{B}$ , and irreversibly decreased at rate  $\chi$ . There is thus a very close connection between the destruction of buoyancy variance and irreversible changes in the potential energy of the flow.

## 2.2. Mixing Versus Stirring

These formulations highlight the inherent connection between energetic pathways and (irreversible) mixing of the density distribution. As discussed by Villiermaux (2019), it is crucially important to distinguish between mixing and stirring. Here, mixing is the irreversible change of the physical properties of fluid parcels, specifically the dynamic scalar density, while stirring is the flow-induced rearrangement of the fluid's density distribution, which is at least in principle reversible.

As discussed in detail by Villiermaux (2019), the ramifications of this distinction are profound. Since it is irreversible, mixing must be associated with diffusive processes, and so it is required that there be a nonzero scalar diffusivity,  $\kappa$ . The destruction rate of density variance  $\chi$  is thus a natural measure of the rate at which mixing is actually taking place, as it captures the inevitable homogenization of the scalar field through mixing. From Equations 2 and 3, and with sweeping assumptions concerning transport terms and averaging, it is then commonplace in geophysical applications to equate mixing with potential energy increase.

However, it is still critical to distinguish between changes in potential energy that are associated with stirring or with mixing, which can be done by use of the particular definition of the concept of available potential energy, as first introduced by Lorenz (1955) and significantly generalized in the influential papers of Winters et al. (1995) and Winters & D'Asaro (1996), who introduced the key connection between mixing and changes in background potential energy,  $\mathcal{K}_{\text{pb}}$ . This is defined as the minimum potential energy of the system, which can be accessed through adiabatic rearrangement of fluid elements, and can be calculated using sorting algorithms (for more discussion, see also Caulfield & Peltier 2000, Peltier & Caulfield 2003) or, to a very good approximation, through the construction of the probability density function (PDF) for the density field (Tseng & Ferziger 2001). The key quantity to be determined is  $z_*(\mathbf{x}, t)$ , i.e., the reference position in the state of background potential energy of the fluid element at  $(\mathbf{x}, t)$  with density  $\rho(\mathbf{x}, t)$ . A beautiful insight of Winters et al. (1995) and Winters & D'Asaro (1996) was that, for the simplest case of a closed system of volume  $V_0$ ,  $E_b$  increases monotonically at a rate  $\Phi_d$ :

$$V_0 \frac{d}{dt} \mathcal{K}_{\text{pb}} = \Phi_d \equiv \kappa \rho_r \int_V \left[ \frac{|\nabla \rho|}{|dz_*/d\rho|^{-1}} \right]^2 N_*^2 dV, \quad N_*^2 = \frac{g}{\rho_r} \left[ \frac{dz_*}{d\rho} \right]^{-1}, \quad 4.$$

where  $V_0$  is the domain volume and  $N_*$  may thus be thought of as the (notional) buoyancy frequency associated with the density field of the background potential energy distribution. Expressed in this way, there is a strong apparent connection to  $\chi$  as defined in Equation 3, with the crucial difference being that  $\chi$  is normalized by the averaged buoyancy frequency, while  $\Phi_d$  is normalized with the sorted buoyancy frequency,  $N_*$ . Indeed, there is also a close connection between the scaled buoyancy variance and the available potential energy,  $\mathcal{K}_{\text{pa}} \equiv \mathcal{K}_{\text{p}} - \mathcal{K}_{\text{pb}}$ , especially in flows with constant background buoyancy gradient.

In an initially statically stable Boussinesq closed system, even in the absence of fluid motions,  $\mathcal{K}_{pb}$  increases due to microscopic diffusive processes at rate  $\Phi_i$  (Winters et al. 1995). Since in the absence of fluid motions, we have  $\mathcal{K}_p \equiv \mathcal{K}_{pb}$ , as argued by Caulfield & Peltier (2000), it is perhaps more natural to consider irreversible conversion into potential energy inherently related to macroscopic fluid motions. This conversion occurs at the strictly non-negative mixing rate,  $\mathcal{M} \equiv \Phi_d - \Phi_i$ . (This subtlety is expected to become less significant as the turbulence becomes more vigorous.)

### 2.3. Eddy Diffusivities and Efficiencies

Although it is clearly too simplistic, the conventional approach for parameterization of stratified mixing has been to develop a classic flux-gradient model for a (vertical) eddy diffusivity of density  $\kappa_\rho$ , parameterizing the mixing properties of small-scale motions in terms of some mean gradient properties of the scalar field. Indeed, even to define this eddy diffusivity, one must make an appropriate choice for the density gradient, and so with respect to stratified fluids, great care must be taken in identifying  $N$  used in the expression

$$\frac{\kappa_\rho}{\kappa} \equiv \frac{\mathcal{B}}{\kappa N^2} = \frac{\nu}{\kappa} \left( \frac{\mathcal{B}}{\mathcal{E}} \right) \frac{\mathcal{E}}{\nu N^2} = Pr \Gamma Re_b. \quad 5.$$

Here,  $Pr$  is the molecular Prandtl number,  $\Gamma \equiv \mathcal{B}/\mathcal{E}$  is the turbulent flux coefficient, and  $Re_b$  is one definition of the buoyancy Reynolds number (Gibson 1980, Gargett et al. 1984), sometimes also called the activity parameter or  $Gn$  (Portwood et al. 2016).

With all of these caveats, the ultimate objective of modeling  $\kappa_\rho$  may thus be reduced to the modeling of the flux coefficient,  $\Gamma$ . In an influential paper, Osborn (1980), relying on available experimental data, postulated  $\Gamma \leq 0.2$  in a statistically steady state, although the inequality is very commonly ignored when estimating mixing from observations of dissipation (Waterhouse et al. 2014). This assumption is on its surface appealing, as it assumes a proportionality between the terms on the right-hand side of Equation 1. However, conceptually, this approach can be criticized, as it relies on a sequence of inferences about the relationship between properties of the turbulent kinetic energy and the quantity of actual interest, namely, the (irreversible) mixing rates. As discussed by Ivey et al. (2018), building on research by Osborn & Cox (1972), it is more natural to focus directly on the buoyancy variance equation, where in steady state and in the absence of transport terms of significance, the buoyancy flux is  $\mathcal{B} \simeq \chi$  when scaled as in Equation 3, and so we have  $\Gamma \simeq \chi/\mathcal{E}$ .

Another classical measure of mixing is the flux Richardson number,  $Ri_f$ , which is traditionally defined (Turner 1973) as the ratio of the density flux to the turbulence production,  $Ri_f \equiv \mathcal{B}/\mathcal{P}$ . This may be thought of as the taxation rate into the potential energy reservoir that the turbulence production experiences, which in an unstratified flow would be paid in its entirety into the internal energy reservoir through viscous dissipation. The flux Richardson number naturally arises when considering the turbulent Prandtl number because the eddy diffusivity of momentum,  $\kappa_m$ , can be defined in terms of a mean (vertical) shear,  $S = \partial \langle u \rangle / \partial z$  (assuming that it exists):

$$\kappa_m = \frac{-\langle u'w' \rangle}{S} = \frac{\mathcal{P}}{S^2} \rightarrow Pr_\Gamma \equiv \frac{\kappa_m}{\kappa_\rho} \equiv \frac{\langle u'w' \rangle / S}{\mathcal{B} / N^2} = \frac{N^2}{S^2} \frac{\mathcal{P}}{\mathcal{B}} = \frac{Ri_g}{Ri_f}, \quad 6.$$

assuming that the turbulent production is driven by (only) the vertical shear. Here  $Ri_g = N^2/S^2$  is naturally an appropriate definition of a gradient Richardson number.

The implications of this deceptively simple expression are deeply significant, as discussed for example by Venayagamoorthy & Stretch (2010). It is clear that  $Ri_f \leq 1$  by construction. Therefore, sustained turbulence in a strongly stratified flow with associated appropriate large values of

$Ri_g$  inevitably requires  $Pr_T \gg 1$ . Whether it is actually possible to have shear-driven sustained turbulence with truly large values of  $Ri_g$  is not a settled question. Conversely, if a flow is weakly stratified, it seems reasonable to expect that mixing of the scalar would occur on similar scales to the mixing of momentum, hence  $Pr_T \sim 1$ , suggesting  $Ri_f \simeq Ri_g$ .

The entire concept of  $Pr_T$  becomes exceptionally difficult to interpret if any spatiotemporal variability occurs in either  $Ri_g$  or  $Ri_f$ , or indeed if  $Ri_f$  becomes negative, which in itself is not excluded since the buoyancy flux  $\mathcal{B}$  is not sign-definite. Also, it has the implicit assumption that shear-driven turbulence production is the dominant driver for mixing processes to occur. Caulfield & Peltier (2000) and Peltier & Caulfield (2003) argued in favor of defining efficiencies,  $\eta$ , of mixing (or, equivalently, turbulent flux coefficients,  $\Gamma$ ) in terms of the mixing rate,  $\mathcal{M}$ , and the dissipation rate,  $\mathcal{E}$ , either instantaneously or cumulatively over the lifetime of a mixing event, assumed to start at time  $t_0$ :

$$\begin{aligned} \eta_i^*(t) &\equiv \frac{\mathcal{M}}{\mathcal{M} + \mathcal{E}}, & \Gamma_i^*(t) &\equiv \frac{\mathcal{M}}{\mathcal{E}} = \frac{\eta_i^*}{1 - \eta_i^*}, \\ \eta_c^*(t) &\equiv \frac{\int_{t_0}^t \mathcal{M} dt}{\int_{t_0}^t \mathcal{M} dt + \int_{t_0}^t \mathcal{E} dt}, & \Gamma_c^*(t) &\equiv \frac{\eta_c^*}{1 - \eta_c^*}, \end{aligned} \quad 7.$$

where the asterisk (\*) denotes the association with the sorted variable  $z_*$ . Salehipour & Peltier (2015) have demonstrated the preferability of using such definitions and of an analogous quantity,  $Re_b^*$ , which they defined using the sorted buoyancy frequency,  $N_*$ , to parameterize mixing in an evolving shear-driven flow. Comparisons between different definitions for mixing, such as those of Salehipour & Peltier (2015) and Venayagamoorthy & Koseff (2016), are valuable; Gregg et al. (2018) has encouraged community consistency in definitions.

Even for these starred, irreversible quantities, there are three fundamental issues. First, the expressions are inherently nonlocal, as they rely on the determination of the background potential energy field to determine  $\mathcal{M}$ , and so access to the density distribution over the entire volume of the flow domain is necessary to evaluate  $\mathcal{M}$ . As discussed by Davies Wykes et al. (2015), quantitative calculations of the various potential energy reservoirs can depend on the size of the domain considered, thus introducing ambiguities in the evaluation of the efficiency of particular mixing processes, especially when the flow has regions of static instability, leading to buoyancy-driven or convective mixing. Building on the work of Andrews (1981) and Holliday & McIntyre (1981), Scotti & White (2014) have formulated a local measure of available potential energy in an attempt to quantify mixing locally, and this is a promising approach to understand better energetic pathways associated with mixing.

Second, the averaging process underlying the calculation of the dissipation rate  $\mathcal{E}$  must be treated carefully. Particularly in a stratified fluid, mixing is highly spatially and temporally intermittent. Although formally defined in terms of an ensemble average, practical calculations of  $\mathcal{E}$  inevitably rely on averaging over a particular volume and time interval. This inevitably introduces ambiguity and uncertainty in the value of  $\mathcal{E}$  to be identified with a given mixing event, if there is any intermittency, as shown by Portwood et al. (2016).

Indeed, this observation points to the third fundamental issue for stratified mixing parameterizations. In the context of passive scalar mixing, as discussed for example by Sreenivasan (2019), it is accepted that the scalar variance destruction rate most definitely is not in lockstep with the turbulent kinetic energy dissipation rate, and anomalous dynamical behavior is generic. Indeed, Villiermaux (2019) demonstrates the important point that for mixing, history matters in the sense that the prior evolution of the fluid flow leaves an imprint on the scalar field such that subsequent

mixing can be strongly affected. That flow history matters should receive greater attention from the stratified mixing community. There is accumulating evidence that such imprints of the previous flow evolution can play a central role in determining mixing properties, and may go some way to resolving apparent disagreements among different models based around the clearly simplistic eddy diffusivity paradigm.

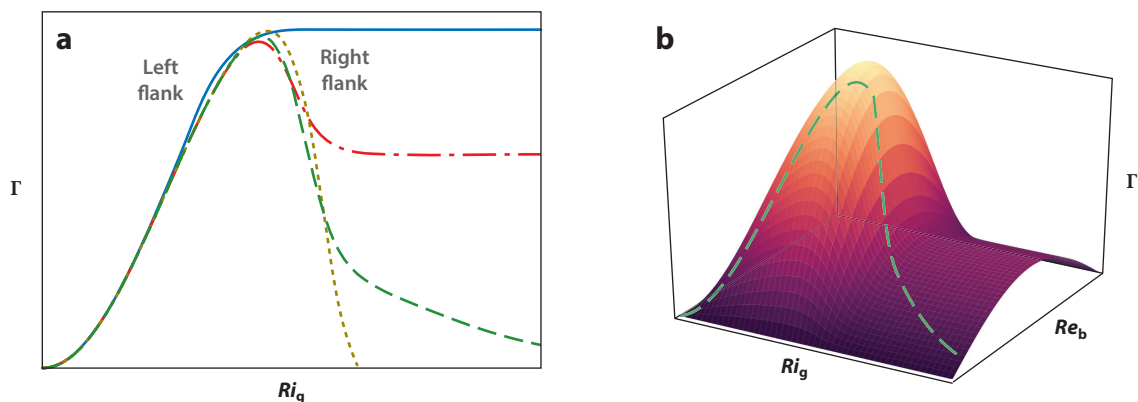
## 2.4. Parameters and Flux Curves

With all of these caveats, the fundamental question for the parameterization of mixing is: What are the functional dependencies of an appropriately defined  $Ri_t$ ,  $\Gamma$ , or  $\eta$  on the flow properties? These may be classified into three groups: properties of the fluid ( $\nu$  and  $\kappa$ ), properties of a background flow ( $N$  and  $S$ ), and properties of the turbulence ( $\mathcal{K}$  and  $\mathcal{E}$ ). Even if it is possible to characterize the flows with unique values of  $\mathcal{E}$ ,  $\mathcal{K}$ ,  $S$ , and  $N$  (all in general functions of space and time), the flow must then be characterized by four nondimensional parameters, which can be chosen to be  $Pr$ ,  $Ri_g$ ,  $Re_b$ , and a horizontal or turbulent Froude number,  $F_h$ :

$$F_h \equiv \frac{\mathcal{E}}{\mathcal{K}N} = \frac{U_T}{L_h N}, \quad L_h \equiv \frac{\mathcal{K}^{3/2}}{\mathcal{E}}, \quad U_T \equiv \mathcal{K}^{1/2}, \quad 8.$$

which implicitly define a velocity scale,  $U_T$ , and a horizontal length scale,  $L_h$ , of the turbulence. Therefore, even accepting all the various sweeping underlying assumptions, the key problem effectively becomes determining, for example,  $\Gamma(Re_b, Ri_g, F_h, Pr)$ . (Of course, it needs to be clear precisely how  $\Gamma$  is defined—for example, whether  $\mathcal{B}$ ,  $\chi$ , or  $\mathcal{M}$  is in the numerator.)

It has been commonplace (see for example Linden 1979, Fernando 1991, Ivey et al. 2008, Wells et al. 2010) to attempt to identify a parametric dependence of  $\Gamma$  on one of these nondimensional parameters, and generic possible pictures are shown in **Figure 1a**. It seems reasonable that  $\Gamma$  must increase from zero as  $Ri_g$  increases on the left flank, as there is no scalar to mix precisely at



**Figure 1**

(a) Four possible flux curves relating the turbulent flux coefficient,  $\Gamma \equiv \mathcal{B}/\mathcal{E}$ , to the gradient Richardson number,  $Ri_g \equiv N^2/S^2$ . The flow is relatively weakly stratified on the left flank, where  $\Gamma$  is expected to increase monotonically with  $Ri_g$ . On the right flank, when the flow is relatively strongly stratified, four possibilities are shown: monotonic saturation of  $\Gamma$  at a finite value (*blue solid line*), decrease of  $\Gamma$  to a finite positive value (*red dash-dotted line*), asymptotic decrease of  $\Gamma$  to zero in the limit  $Ri_g \rightarrow \infty$  (*green dashed line*), and complete switch-off of mixing at a critical value of  $Ri_g$  (*yellow dotted line*). (b) Flux surface showing a plausible hidden correlation between  $Ri_g$  and  $Re_b \equiv \mathcal{E}/(\nu N^2)$  leading to the projected nonmonotonic behavior (*green dashed curve*) when  $\Gamma$  is plotted as a function of  $Ri_g$  alone, as in panel a.

$Ri_g = 0$  and so we have  $\Gamma = 0$ . However, what actually happens at strong stratification is still an open question of some controversy.

Four possibilities are shown. Perhaps the simplest is the (solid blue) curve monotonically increasing toward an asymptotic maximum, assuming that  $\Gamma$  is bounded above by some finite quantity, which pertains in the limit of strong stratification. However, as stratification becomes more significant, for example through increased anisotropy in the turbulent flow with vertical velocities  $w \ll u, v$ , it is plausible that the mixing will become less efficient. Such an argument implies both a global maximum in  $\Gamma$  for some optimal  $Ri_g$  and the existence of a right flank where the flux coefficient decreases with increasing  $Ri_g$ . There are then generically three different possible behaviors in addition to the asymptotic curve described above:  $\Gamma$  may decrease toward a finite nonzero constant (red dash-dotted line) or may tend asymptotically to zero (green dashed line) in the limit  $Ri_g \rightarrow \infty$ . These cases assume implicitly that irreversible turbulent mixing can continue to be maintained at very strong stratification. If that is not true, the mixing will completely switch off at some critical value of  $Ri_g$ , as shown by the yellow dotted line in the figure.

It is still an open question which (if any) of these flux curves are appropriate to underlie mixing parameterizations. Furthermore, such flux curves are a projection of a presumably significantly more complex flux hypersurface representing the full assumed dependence  $\Gamma(Re_b, Ri_g, F_h, Pr)$ , and the true objective should be to determine this hypersurface. That objective is further complicated by the issue of whether it is actually possible to explore this parameter space fully. For example, recall that  $Ri_g$  is defined as  $N^2/S^2$ , where  $S$  is some measure of background shear within a flow. It is at least conceivable that in some turbulent flows the shear  $S$  is proportional to  $U_T/L_h$ , and so  $Ri_g$  is proportional to  $F_h^{-2}$ . Such emergent correlations between assumed independent parameters can make the interpretation of such flux curves, as shown schematically in **Figure 1a**, problematic. A particular issue is whether strongly stratified flows, i.e., those with high values of  $Ri_g$  or small values of  $F_h$ , can also maintain strong turbulence in the sense of high values of  $Re_b$ , or equivalently, whether high values of  $Re_b$  inevitably imply weak stratification. Indeed, there is evidence (see for example Lucas & Caulfield 2017) that certain flows traverse particular curves in parameter space, with functional relationships (such as  $Ri_g \propto Re_b^{-1}$ ) between various parameters, as shown schematically in **Figure 1b**.

## 2.5. Length Scales and Layering

Nondimensional parameters can be thought of as ratios of characteristic timescales or length scales. In a pair of important papers, Mater & Venayagamoorthy (2014a,b) collated data from a wide range of sources and demonstrated that flow dynamics could vary markedly across two-dimensional (2D) projections of the full parameter space. This variation in behavior could be interpreted in terms of the relative size of three timescales: the shear,  $T_s = 1/S$ ; buoyancy,  $T_b = 1/N$ ; and turbulence,  $T_T = \mathcal{E}/K$ , timescales. [Note that we have  $Ri_g = (T_s/T_b)^2$ ,  $F_h = T_b/T_T$ ,  $Re_b = T_b^2/T_K^2$ , and  $Pr = (T_K/T_B)^2$ , where  $T_K = \sqrt{\nu/\mathcal{E}}$  and  $T_B = \sqrt{\kappa/\mathcal{E}}$  are the Kolmogorov and Batchelor timescales, respectively.] Mater & Venayagamoorthy identified three qualitatively different regimes: an inertia-dominated regime with  $T_T \lesssim T_s$ ,  $T_b$ , a shear-dominated regime with  $T_s \lesssim T_T$ ,  $T_b$ , and a buoyancy-dominated regime with  $T_b \lesssim T_s, T_T$ .

However, there are several physical reasons why it is arguably more natural to interpret the various parameters as ratios of length scales. For example, as originally discussed by Phillips (1972) (see also Park et al. 1994 for both a clear discussion and experimental observation), the existence of a right flank with respect to some measure of stratification in a flux curve (as shown in **Figure 1**) implies antidiffusive behavior in the sense that fluctuations in a uniform density gradient would



be expected to be amplified. This behavior should lead inevitably to a layer-interface structure with relatively well-mixed, relatively deep layers (with associated flux coefficient on the left flank) separated by relatively thin interfaces of higher gradient (with matching flux coefficient on the right flank). Although in this simple conceptual model, the flux coefficient should be thought of as associated with a turbulent flux, as argued by Balmforth et al. (1998), if stratification suppresses turbulent motions, eventually molecular diffusive fluxes must become significant, thus leading to regularization of the interfaces at some maximum density gradient.

Indeed, there is an accumulating weight of observations that suggest that strongly stratified flows have a tendency to form and maintain such a quasi-horizontal layer-interface structure. Particularly in sufficiently large flow domains, these layers take the form of pancakes in the sense that they have a finite horizontal extent, which is still much larger than their vertical extent (Waite & Bartello 2004, Praud et al. 2005, Basak & Sarkar 2006, Brethouwer et al. 2007, Almalkie & de Bruyn Kops 2012, Kimura & Herring 2012, Bartello & Tobias 2013, de Bruyn Kops 2015). Such spatial structure calls into question again how credible it is to use single values of the various nondimensional parameters, and such inherently layered structures are another way in which the history of a given flow's evolution may play a critical role in the irreversible mixing that occurs.

Billant & Chomaz (2001) identified a particular regime with  $F_h \ll 1$  and  $Re_T \equiv U_T L_h / \nu \gg 1$ , such that we have  $Re_T F_h^2 \gg 1$ . (Where the inertial scaling  $\mathcal{E} \propto U_T^3 / L_h$  is assumed, this distinguished limit corresponds to  $Re_b \gg 1$ .) With  $Pr \gtrsim 1$ , typical of water and gases, the governing equations become self-similar with respect to  $zN/U_T$ . This self-similarity suggests a characteristic vertical scale,  $L_v \equiv U_T / N \ll L_h$ , for the pancakes such that the vertical Froude number  $F_v \equiv U_T / (NL_v)$  remains  $\mathcal{O}(1)$  even in this distinguished limit.

There is mounting evidence of such characteristic layer depths in a range of flows with a variety of forcing mechanisms, including the zigzag instability of vertical vortices in a stratified fluid (Billant & Chomaz 2000a,b; Deloncle et al. 2008; Waite & Smolarkiewicz 2008; Augier & Billant 2011; Lucas et al. 2017), stirring by vertical grids or rods (Park et al. 1994, Holford & Linden 1999, Praud et al. 2005, Thorpe 2016), stratified Taylor–Couette flow in the annular region between two concentric cylinders (Oglethorpe et al. 2013), and numerical simulations with a variety of forcing mechanisms and geometries (Riley & de Bruyn Kops 2003, Waite & Bartello 2004, Lindborg 2006, Brethouwer et al. 2007, Almalkie & de Bruyn Kops 2012, Kimura & Herring 2012, Bartello & Tobias 2013, de Bruyn Kops 2015, Portwood et al. 2016, Lucas et al. 2019).

Lindborg (2006) considered the regime with the same distinguished limits as Billant & Chomaz (2001) and demonstrated that it still allowed a (profoundly anisotropic) net-forward cascade of energy in strongly stratified turbulence. In particular, the properties of the turbulence must be different above and below the Ozmidov length scale,  $L_O$ , which may be defined consistently with the above development as

$$L_O \equiv \left( \frac{\mathcal{E}}{N^3} \right)^{1/2} \rightarrow Re_b = \left( \frac{L_O}{L_K} \right)^{3/4}, \quad L_K \equiv \left( \frac{\nu^3}{\mathcal{E}} \right)^{1/4}, \quad 9.$$

where  $L_K$  is the Kolmogorov microscale. For vertical scales smaller than  $L_O$ , the flow is at most weakly affected by the density stratification. Interpretation of the meaning of  $L_O$  must be done with care, however, particularly because  $L_O$  is a hybrid parameter comparing an intrinsic property of the turbulence that quantifies locally the intensity of the eddying energy cascade to a background property of the flow that quantifies the strength of the stratification. In particular, implicit in its definition is the idea that the stratification extends over a vertical distance at least as large as the region of strong turbulence.

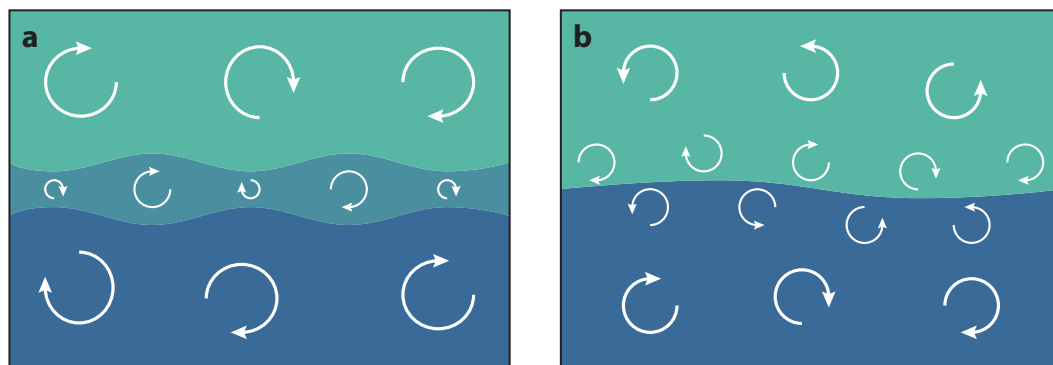
Nevertheless, with this caveat, the buoyancy Reynolds number  $Re_b$  can then be interpreted as a ratio of length scales. Therefore, for there to be absolutely any possibility of a classical inertial

range of isotropic turbulence at smaller scales,  $Re_B$  must be sufficiently large. Furthermore, not only can the horizontal Froude number  $F_h$  be interpreted as the ratio  $F_h = L_v/L_h$  but also, using the definition of  $L_O$  with the inertial (equality) scaling,  $\mathcal{E} \equiv U_T^3/L_h$ , one can define it as  $F_h = (L_O/L_v)^2$ , and so to be in such a turbulent strongly stratified regime, it is required that the length scales are in the asymptotic ordering  $L_K \ll L_O \ll L_v \ll L_h$ .

As argued by Lindborg (2006) and established by further subsequent numerical simulations (Brethouwer et al. 2007, Bartello & Tobias 2013) and reinterpretations of observations (Riley & Lindborg 2008, Falder et al. 2016), in this regime a net-forward (to small scales) energy cascade can be observed, with the horizontal spectra of kinetic and potential energy exhibiting a  $k_h^{-5/3}$  power law range, where  $k_h$  is an appropriate horizontal wavenumber. This regime has been termed the stratified turbulence regime or the strongly stratified turbulence regime (Zhou & Diamessis 2019). However, in the geophysical literature, “stratified turbulence” is often used in a much broader sense to describe any turbulent motions in a stratified fluid, and so following Falder et al. (2016), this regime is referred to here as the layered anisotropic stratified turbulence (LAST) regime.

As discussed by Brethouwer et al. (2007), the oceans and the atmosphere are expected to be in this regime, with very small values of  $F_h \lesssim 10^{-3}$  and very large values of  $Re_T \gtrsim 10^8$ , and so it is of great interest to investigate the properties of mixing within this inherently layered regime. Unfortunately, computing numerical simulations or indeed conducting laboratory simulations in this regime is enormously challenging, as discussed by Bartello & Tobias (2013), due not least to the wide range of (anisotropic) scales that need to be modeled, a challenge that is only exacerbated by attempting to consider  $Pr \sim \mathcal{O}(10)$ , characteristic of thermally stratified water. The existence of a true inertial range requires a range of scales  $L_i$  such that both  $L_i \ll L_O$  and  $L_K \ll L_i$  obtain simultaneously. It is not clear how such a flow can be accessed, maintaining vigorous turbulence in very strong local stratification.

As already noted above, there is a natural tendency, not least in the LAST regime, for stratified fluids to layer and so exhibit substantial vertical variation. Building on the work of Kato & Phillips (1969), Woods et al. (2010) argued that there is a qualitative difference between the mixing expected in relatively weak stratification and in relatively strong stratification. When the stratification is weak, the turbulence is sufficiently strong to overturn density interfaces, as shown in **Figure 2a**. Mixing will tend to smooth out strong gradients, and so it is plausible to describe such mixing with eddy diffusivities. Conversely, as shown in **Figure 2b**, mixing in strong stratification is characterized by turbulent motions that scour the density interfaces, leading to a qualitatively



**Figure 2**

(a) Overturning mixing in weak stratification. (b) Scouring mixing in strong stratification. Figure adapted with permission from Woods et al. (2010); copyright 2010 Cambridge University Press.

different mixing process, which is in a real sense antidiffusive. As originally argued by Crapper & Linden (1974), the properties of such sharp interfaces are affected by values of the Péclet number,  $Pe = U_c L_c / \kappa$ , where  $U_c$  and  $L_c$  are characteristic velocity and length scales, respectively, and Taylor & Zhou (2017) introduced a criterion for the development of such layering in terms of the spatial distribution of an appropriate eddy diffusivity.

This distinction between overturning and scouring has points of similarity with the concepts of internal as opposed to external mixing introduced by Turner (1973), although the equivalence is not perfect, since scouring dynamics within a layered stratification can still occur even though the dominant turbulence production process is inherently local to the flow dynamics. However, it still needs to be established how often such scouring dynamics or indeed layer-interface structures can actually be realized. Although in principle all the possible parameters are independent, it is not clear whether the entire multidimensional parameter space is accessible by real flows. This issue is deceptively complicated when there is both shear and turbulence within the flow, and indeed it is very unclear whether a shear flow, prone to the onset of turbulence, can ever be strongly stratified in the specific sense of having large values of minimum gradient Richardson number. Therefore, a crucially important question to be understood is how stratified (shear) flows can actually become turbulent.

### 3. TURBULENCE ONSET IN STRATIFIED SHEAR FLOWS

There are two natural classes of stratified shear flows: when the background velocity gradient and density gradient are oriented in the same direction (vertical shear flow) and when they are orthogonal (horizontal shear flow).

#### 3.1. Vertical Shear Kelvin–Helmholtz Instabilities

Finite-depth shear layers with an appropriate inflection point are prone to the so-called Kelvin–Helmholtz instability (KHI). Perhaps the most canonical choice considered is the hyperbolic tangent shear layer with background velocity  $\mathbf{u}_b$  and density  $\rho_b$ :

$$\begin{aligned} \mathbf{u}_b &= U(z)\hat{\mathbf{x}}, & U(z) &= U_0 \tanh(z/d_0), \\ \rho_b &= \rho_r - \rho_0 \tanh(z/\delta_0), & Re_0 &= \frac{U_0 d_0}{\nu}, \end{aligned} \quad 10.$$

where  $2d_0$  is the initial total depth of the shear layer (an interface or strip of vorticity), with total velocity jump  $2U_0$ ;  $2\delta_0$  is the initial total depth of the density interface, across which there is a total density jump of  $2\rho_0$ ; and  $\rho_r$  is a reference density. This instability rolls up the initial strip of vorticity into a periodic train of elliptical vortices called billows connected by braids, with a characteristic wavelength on the order of  $\sim 15d_0$  and vertical extent of  $\sim 5d_0$ . Such billow trains can be observed experimentally in tilting tanks (e.g., Thorpe 1973, Caulfield et al. 1996, Patterson et al. 2006).

Unsurprisingly, stratification has a (monotonically) stabilizing effect on the growth rate of the KHI, and the key nondimensional quantity to understand the stability properties of such flows is the gradient Richardson number,

$$Ri_g(z) \equiv \frac{N^2}{S^2} = \frac{g\rho_0 d}{\rho_r U_0^2} \frac{d}{\delta} \frac{\text{sech}^2(z/\delta)}{\text{sech}^4(z/d)} = Ri_0 R \frac{\text{sech}^2(z/\delta)}{\text{sech}^4(z/d)}, \quad 11.$$

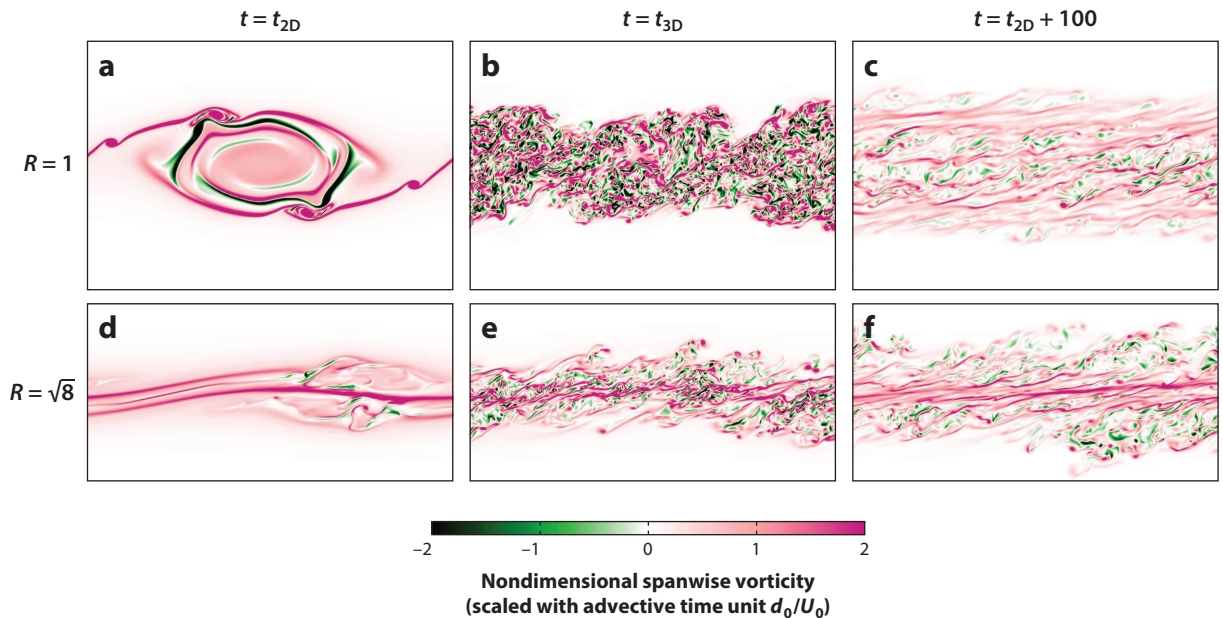
which defines the length scale ratio  $R$  and the bulk Richardson number  $Ri_0$ , which for the simplest case of  $R = 1$  corresponds also to the minimum gradient Richardson number,  $Ri_m \equiv \min_z(Ri_g)$ ,

at the midpoint of the shear layer. As observed by Smyth et al. (1988), this property of minimum Richardson number is important for application of the classical result due to Miles (1961) and Howard (1961). The Miles–Howard theorem, which is strictly applicable only to steady, inviscid Boussinesq flows where the background velocity and density distributions depend only on the vertical coordinate, states that  $Ri_m < 1/4$  is a necessary condition for instability.

However, in the presence of unsteadiness, associated for example with ambient turbulence, the criterion is not formally applicable. It is an open question whether the stability of a mean flow, which typically does not apply for any substantial period of time, has any relevance to the subsequent evolution of a stratified turbulent flow. Observations in the equatorial undercurrent (Smyth & Moum 2013) suggest that the PDF for measured values of  $Ri_g$  peaks around the critical value predicted by linear theory. Thorpe & Liu (2009) postulated that real flows adjust toward a value of marginal stability fundamentally by going through a cycle. Some external forcing mechanism will tend to strengthen a shear until it is sufficiently strong and instabilities onset, triggering turbulence, which then decelerates the shear until the flow is no longer unstable, until the external forcing once again nudges the flow back into instability. Such a cycle is also strongly reminiscent of self-organized criticality, and there is indeed some numerical and observational evidence that stratified shear flows can indeed exhibit such behavior (Salehipour et al. 2018, Smyth et al. 2019) in that stratified shear flows appear unable to access strong stratification, but rather have mean profiles with typical values of  $Ri_g \simeq 0.25$  over much of the shear layer. Indeed, certain aspects of such an adjustment toward criticality were foreshadowed by Turner (1973), who presented physical arguments suggesting that flows adjust to a kind of equilibrium between inertia and buoyancy. Although the effects of both ambient turbulence (Kaminski & Smyth 2019) and transient growth (Kaminski et al. 2014, 2017) must be treated with care, and although there is some evidence that (finite amplitude, yet still quite weak) billow states can exist for  $Ri_m > 1/4$  (Howland et al. 2018, Parker et al. 2019), it does not seem that  $Ri_m$  can be significantly greater than  $1/4$  for significant finite-amplitude billows, although it is conceivable that this observation may depend on  $Pr$ .

The essentially inviscid and unstratified KHI is known to undergo a transition to turbulence through an array of secondary instabilities, provided an appropriate flow Reynolds number is sufficiently high,  $\mathcal{O}(10^4)$ , for the so-called mixing transition to occur (Dimotakis 2005), and there has been a very wide range of numerical simulations of KHI. A first tranche of simulations, carried out at sufficiently high resolution to satisfy the necessary condition that the primary instabilities can be prone to (generically 3D) secondary instabilities, typically had periodic horizontal boundary conditions, chosen to allow one or two billows, with Reynolds numbers  $Re_0 \sim 1,000$ – $2,000$ ,  $R \sim 1$ , and  $Pr \sim 1$  (Caulfield & Peltier 1994, 2000; Palmer et al. 1994, 1996; Smyth & Moum 2000; Peltier & Caulfield 2003). Since the vertical extent of the billow is  $\mathcal{O}(5d_0)$ , it can be argued that an appropriate Reynolds number for this flow is about five times larger than  $Re_0$ , (just) sufficiently large for it to be plausible that the mixing transition can occur.

However, another mixing transition has been effectively identified by a second tranche of studies at yet larger Reynolds number. For  $Re_0 \gtrsim 4,000$ , a further zoo of secondary instabilities have been identified by Mashayek & Peltier (2012a,b). Perhaps most significantly, these vigorous secondary instabilities suppress the onset of billow merging for flows with sufficiently high Reynolds numbers and Richardson numbers (Mashayek & Peltier 2013), and the mixing properties (discussed below) of such high-Reynolds number flows susceptible to vigorous overturning primary KHI also exhibit nontrivial differences with those at lower  $Re_0$ . Due to this qualitative change in behavior between the two tranches of simulations with different characteristic ranges of  $Re_0$ , a particular point of caution must be expressed regarding whether an ultimate regime of shear-induced stratified mixing has been identified by such simulations. It is at least possible that, once



**Figure 3**

Contours of the nondimensional spanwise vorticity (scaled with advective time units  $d_0/U_0$ ) for simulations with Reynolds number  $Re_0 \equiv U_0 d_0 / \nu = 4,000$ , molecular Prandtl number  $Pr = \nu / \kappa = 8$ , and bulk Richardson number  $Ri_0 = g \rho_0 d_0 / (\rho_r U_0^2) = 0.16$  of (a–c) flows with a length scale ratio of  $R = d_0 / \delta_0 = 1$ , which are prone to the primary Kelvin–Helmholtz instability (KHI), and (d–f) flows with  $R = \sqrt{8}$ , which are prone to the primary Holmboe wave instability (HWI). Contours are shown at nondimensional times: at  $t = t_{2D}$ , when spanwise-averaged turbulent kinetic energy is maximum; at  $t = t_{3D}$ , when three-dimensional perturbations are maximum; and at  $t = t_{2D} + 100$ . Note the overturning by the primary KHI and scouring by the HWI. Figure adapted with permission from Salehipour et al. (2016a); copyright 2016 Cambridge University Press.

again, higher-Reynolds number simulations will reveal different, and inevitably richer, dynamics. Three snapshots of the roll-up and turbulent breakdown of a KHI with  $Re_0 = 4,000$  are shown in **Figure 3a–c**.

### 3.2. Vertical Shear Inherently Stratified Instabilities

Stratified shear flows are also prone to a range of other instabilities that owe their very existence to the presence of a statically stable stratification. At finite amplitude, the appearance of trains of elliptical vortices does not in itself imply that these vortices are associated with KHI, and may well be associated with instabilities with quantitatively different mixing properties. As originally appreciated by Taylor (1931), shear flow instability can often be interpreted in terms of wave interactions (Carpenter et al. 2011, Guha & Lawrence 2014, Smyth & Carpenter 2019). KHI can be interpreted as being due to a resonance between the vorticity or Rayleigh waves (not to be confused with the Rayleigh waves of solid mechanics, and perhaps more appropriately considered as a variant of the Rossby waves of geophysical fluid dynamics), which are localized at either edge of the shear layer. These marginally stable waves propagate locally upstream, and so under certain circumstances the waves localized at each edge of the shear layer can have the same wavenumber and frequency, thus possibly leading to instability growth. Although this formalism is applicable straightforwardly when the density and vorticity distributions are piecewise constant, it carries over to smooth velocity and density distributions (Smyth et al. 1988, Alexakis 2005).

For  $R \gg 1$  when the density interface is much sharper than the shear layer, it is possible for the density interface to support two effectively localized internal gravity waves. One of these waves propagates upstream, while the other propagates downstream, relative to the local fluid. Furthermore, for sufficiently large  $R$ , the gradient Richardson number actually is maximum in the vicinity of the density interface and drops to small values in the far field, irrespective of the value of the bulk Richardson number,  $Ri_0$ , and so the Miles–Howard criterion is irrelevant. Generically, if the density distribution has a layer-interface structure, then strong stratification, in the sense of high values of the bulk Richardson number, does not preclude flow instability since generically there will be somewhere within the flow where the gradient Richardson number is small (see, e.g., Alexakis 2009).

The classic example demonstrating the propensity of layered stratified shear flows to be unstable to high values of  $Ri_0$  arises in the limits of inviscid flow and  $R \rightarrow \infty$ , with a two-layer density distribution having an infinitesimally thin density interface. For all values of  $Ri_0$  there are (generically) two bands of wavenumbers where the flow is unstable to the so-called Holmboe (wave) instability (HWI) (Holmboe 1962, Browand & Winant 1973, Koop & Browand 1979, Smyth et al. 1988), although if the flow is symmetrical about the midpoint of the shear layer, these two bands coincide (Lawrence et al. 1991). In terms of wave interactions, this instability can be interpreted as being due to an interaction between a vorticity wave at the edge of the shear layer and one of the internal waves localized at the density interface (Caulfield 1994, Baines & Mitsudera 1994) propagating upstream or downstream relative to the local velocity at the density interface.

At finite amplitude, this instability has been observed both numerically (Smyth et al. 1988, Smyth & Winters 2003, Carpenter et al. 2010) and experimentally (Browand & Winant 1973, Caulfield et al. 1995, Zhu & Lawrence 1996, Hogg & Ivey 2003, Tedford et al. 2009, Lefauve et al. 2018). Classically, HWI at finite amplitude has been associated with cusped waves propagating along the density interface, with occasional wisps of fluid being ejected from the top of the cusps. These cusped waves can be identified with propagating vortices displaced above or below the density interface; in particular, when the density interface is substantially displaced from the midpoint of the shear layer, such a vortex can have significant qualitative similarity to a KHI billow. Analogously to the KHI, the transition to turbulence is qualitatively different at  $Re \gtrsim 4,000$  (Salehipour et al. 2016a, 2018), as shown in **Figure 3d–f**. A further complication when considering the dynamics of such instabilities is that, since the HWI relies, at least initially, on the existence of sharp density interfaces,  $Pr$  typically needs to be substantially greater than one so that the background density distribution does not diffuse too rapidly.

The HWI still relies on the existence of at least one significant change in the velocity gradient (at which the constituent vorticity wave is localized) and, in a finite-depth shear layer, can be thought of as arising as a bifurcation of the KHI (Hogg & Ivey 2003). Therefore, although it is an inherently stratified instability, it still shares some characteristics with unstratified inflectional instabilities like the KHI. However, as also first appreciated by Taylor (1931), layered stratified shear flows can be unstable to instabilities that have no connection at all to inflectional shear layers. This instability, further considered by Caulfield (1994) and experimentally reported by Caulfield et al. (1995), arises in a sheared flow with at least two density interfaces, with at least one finite-depth layer of intermediate density relative to outer layers. Each density interface supports two localized internal waves propagating upstream and downstream relative to the local fluid velocity. If there is a background shear (not necessarily inflectional), there is once again the possibility of the interaction of Doppler-shifted internal waves localized on each interface, provided they have the same wavenumber and frequency. Such an interaction can lead to a qualitatively different instability.

This instability, which was termed the Taylor–Caulfield instability (TCI) by Carpenter et al. (2010) (see also Churilov 2016), is also manifest at finite amplitude as an array of elliptical

vortices (Caulfield et al. 1995, Lee & Caulfield 2001, Balmforth et al. 2012, Eaves & Caulfield 2017, Ponetti et al. 2018, Eaves & Balmforth 2019) trapped between the bounding density interfaces. As demonstrated by Ponetti et al. (2018), the general concept of instability arising due to interaction of Doppler-shifted interfacial waves within a layered stratified shear flow can be generalized to an arbitrary number of interfaces. Due to the apparent necessity of very sharp density interfaces for the development of this instability, and hence very high values of the Péclet number,  $Pe = RePr$ , most numerical studies of this instability have been restricted to two dimensions. Nevertheless, the general picture seems to point toward the finite-amplitude saturated TCI also triggering a wide range of secondary instabilities, specifically apparently leading to structures highly reminiscent of HWI on the various (sheared) density interfaces. Significantly, the saturated elliptical vortex associated with TCI is once again reminiscent of KHI, although the growth mechanisms and flow signature of these three canonical instabilities (KHI, HWI, and TCI) are qualitatively distinct (Eaves & Balmforth 2019).

### 3.3. Horizontal Shear Instabilities

The finite amplitude form of the canonical (vertical) shear instabilities generically takes the form of (horizontally aligned) elliptical vortices, and it is natural to expect that the rolling up of such vortices will have an energetic cost within a (vertically) stratified fluid. Such a cost clearly does not apply if the fluid is sheared horizontally, and so such shear instabilities should be able to grow straightforwardly and unconstrained by stratification. Basak & Sarkar (2006) demonstrated that for a horizontal hyperbolic tangent shear layer with an initially constant vertical density gradient  $N$ , the primary shear-driven instability still developed, creating an array of elliptical vortices. Crucially, they demonstrated that these vortices were then prone to strong secondary instabilities, which led to the spontaneous development of a layer-interface structure in the vertical density distribution, with characteristic layer depth scaling as  $l_v \sim U/N$ , where  $U$  is a characteristic velocity scale of the ambient shear, highly reminiscent of the LAST regime discussed above.

At high  $Re$ ,  $U/N$  is of course a natural inviscid vertical length scale. Billant & Chomaz (2000a,b) demonstrated that vertically oriented vortices in a stratified fluid (as would be expected to form from instability of a horizontally aligned shear layer) are prone to the zigzag instability, which naturally leads to layering on the scale of  $U/N$ . Subsequent numerical simulations (see, e.g., Deloncle et al. 2008, Waite & Smolarkiewicz 2008) have demonstrated that, at sufficiently high Reynolds number, subsequent turbulent breakdown does indeed lead to characteristic layering on a scale of  $U/N$ . [Although Billant & Chomaz initially considered the interaction of two oppositely signed vertical vortices, subsequently Otheguy et al. (2006) demonstrated that like-signed vortices were also prone to a zigzag instability.] Such a scaling does not appear to rely on a particular shear layer structure, as Lucas et al. (2017) observed similar behavior in a stratified Kolmogorov flow forced by sinusoidal horizontal shear.

This layer-interface structure arises in a variety of circumstances. Experimentally, Oglethorpe et al. (2013) demonstrated that a layer-interface structure also spontaneously appears in vertically stratified Taylor–Couette flow, i.e., the flow in the annular region between two concentric cylinders. Oglethorpe et al. (2013) observed very long-lived and robust well-mixed layers of characteristic depth that scaled with  $U_H/N$ , where  $N$  is the buoyancy frequency of the initial linear stratification and the horizontal velocity scale,  $U_H$ , is given by  $\Omega_I \sqrt{R_I(R_O - R_I)}$ , where  $\Omega_I$  is the angular velocity of the inner cylinder of radius  $R_I$  and  $R_O$  is the radius of the stationary outer cylinder. The initial depth of these layers may be imprinted by some variant of the strato-rotational instability (Molemaker et al. 2001, Le Bars & Le Gal 2007), but the



robustness of relatively sharp interfaces, even in the presence of significant turbulent motions [as first observed by Guyez et al. (2007) in two-layer stratified Taylor–Couette flow], strongly suggests that it is a generically accessible state within stratified turbulent flows. In real flows, such layered states might be prone to the subsequent occurrence of vertical shear, thus motivating further study of mixing due to inherently stratified instabilities that rely on the existence of sharp interfaces, such as HWI and TCI.

Three key questions concerning these flows arise. First, is the layer-interface structure generic in the sense can it arise in a wide range of situations? Second, what are the circumstances in which the layer-interface density structure remains robust? Finally, what sets the vertical scale of the interfaces? Attempts to answer these questions appear to be inherently linked to the mixing properties of such flows and the distinction between relatively weakly stratified interfaces, which are overturned for example through the breakdown of the KHI, and relatively strongly stratified interfaces, which are scoured for example through the breakdown of the HWI.

## 4. TEMPORAL EVOLUTION OF TURBULENCE IN STRATIFIED FLOWS

Turbulent mixing triggered by shear instabilities in stratified flows, if the base flow is not forced, is an inherently transient run-down problem. Conversely, it is possible both experimentally and numerically to consider continually forced flows, where it is at least conceivable that the flow approaches some kind of quasi-steady equilibrium. A natural question to ask is then whether the mixing properties associated with these two situations are equivalent. That is, are the properties of turbulence in a transient flow (with in general time-varying parameter values of  $Ri_g$ ,  $Re_b$ , etc.) defined adequately by the instantaneous values of those parameters, and thus are they closely related to a steady flow with at least approximately constant values of the relevant parameters? It is fair to say that such equivalence relies on a pretty heroic set of assumptions, not least due to the overwhelming evidence that history does indeed matter (crucially) to the irreversible mixing induced by the turbulence.

### 4.1. Mixing Induced by Overturning Kelvin–Helmholtz Instability

Mixing induced by the development and ultimate breakdown of primary KHI at sufficiently high  $Re$  and  $Ri$  is appropriately categorized as overturning, driven locally by a wide range of secondary instabilities (Mashayek & Peltier 2012a,b, 2013). A quantitative measure to describe the extent of overturning, which varies in a time-dependent fashion in such flows, is the so-called Thorpe scale (Thorpe 1977, Smyth & Moum 2000, Odier & Ecke 2017, Mashayek et al. 2017a). In a (1D) typical density profile  $\rho(z, t)$ , generically there may be regions of static instability with  $\partial\rho/\partial z > 0$ . Such a profile can be sorted to be statically stable everywhere, analogously to the construction of the background potential energy,  $\mathcal{K}_{pb}$ . Each individual fluid parcel,  $p(i)$ , has to be displaced a distance  $d_i$  vertically to form this statically stable profile. The Thorpe scale,  $L_T$ , can then be defined as the root mean square (RMS) of these displacement lengths,  $d_i$ .

In numerical simulations, an analogous 3D Thorpe scale,  $L_T^{3D}$ , can be constructed as the RMS value of all the displacements across an entire computational volume required to construct the background density distribution,  $\rho_b(z)$ , associated with  $\mathcal{K}_{pb}$ . As discussed in detail by Mashayek et al. (2017a), following Smyth & Moum (2000) and Smyth et al. (2001), there can be differences between these two length scales, and so the use of (either of) these quantities in parameterizations must be treated with caution. Indeed, there is a wide range of uncertainty in using this purely geometrical construct to estimate mixing, which is associated with, for example, inherent biases (see Mater et al. 2015, Scotti 2015) or sampling issues in inferring typical values



from a relatively sparse set of vertical profiles (Itsweire et al. 1993, Taylor et al. 2019). As discussed by Mater et al. (2013), under certain circumstances, however,  $L_T$  can be related directly to the so-called Ellison scale,  $L_E = \langle (\rho')^2 \rangle^{1/2} / |\partial \bar{\rho} / \partial z|$ , where angle brackets denote ensemble averaging and  $\rho'$  is the (turbulent) fluctuation away from the mean density profile,  $\bar{\rho}$ .

With these caveats, there is still a real attraction to using estimates of the Thorpe scale to quantify mixing, as estimations of  $L_T$  can be done purely from knowledge of a profile of the density field, as can be relatively straightforwardly measured by observational oceanographers, whereas estimates of dissipation rates such as  $\chi$  and  $\mathcal{E}$  are extremely challenging to obtain, not least due to the need to measure accurately spatial gradients of fluctuating quantities. [Even  $L_E$  requires accurate measurement of fluctuation quantities, as discussed for example by Ivey et al. (2018).] Dillon (1982) hypothesized that a transient turbulent mixing event could be characterized by a particular time evolution of the ratio  $R_{OT}(t) = L_O/L_T$ , and he argued that the event's age could be inferred from the value of  $R_{OT}$ . It seems reasonable that vigorous turbulence, and hence increasing  $\mathcal{E}$ , will tend to mix a stratified fluid, and hence to decrease  $L_T$  while still having a large value of  $L_O$ . As a turbulent event ages,  $R_{OT}$  should thus increase, allowing a particular value of  $R_{OT}$  to be identified with a particular stage of a particular mixing event.

Furthermore, if a typical value of  $R_{OT}$  can be identified [Dillon (1982) suggested  $R_{OT} \sim 0.8$ ] then measurements of  $L_T$  can be used to infer values of  $\mathcal{E}$  and hence (assuming  $\Gamma$  is also known) to estimate values of  $\kappa_\rho$ , the quantity of interest. Though commonly used in observational practice by necessity (Waterhouse et al. 2014, Gregg et al. 2018), this approach has a large amount of uncertainty, not least in the choice of the most appropriate estimates of the various background quantities such as  $N$  (Arthur et al. 2017). Smyth & Moum (2000) and Smyth et al. (2001) considered the evolution of  $R_{OT}$  for evolving KHI-dominated flows (with  $Re_0 \sim 500$ –1,250, as defined in Equation 10) and indeed demonstrated that such flows exhibit  $R_{OT}$  increasing with time. A similar analysis for much more vigorously turbulent flows with  $Re_0 \gtrsim 4,000$  by Mashayek et al. (2017a) also observed  $R_{OT}$  increasing with time and found no evidence of a specific typical value of  $R_{OT}$ .

Analysis of a wide range of oceanic observations by Ijichi & Hibiya (2018) demonstrated a dependence of estimates of the flux coefficient on  $\Gamma \propto R_{OT}^{-4/3}$ . As they argued, this scaling arises naturally from a set of length-scale-based assumptions. Effectively, these all rely on a central concept that the observed turbulence itself is at most weakly affected by the stratification and does not in itself require consideration of the age of a particular turbulent event. First, the classical inertial scaling for the dissipation rate relies on the assumption that  $\mathcal{E}$  can be expressed in terms of the kinetic energy and some characteristic length scale,  $\mathcal{E} \sim U^3/L \sim K^{3/2}/L$ . Second, it is assumed that this length scale may be interpreted as a mixing length, connecting the eddy diffusivity for momentum to the characteristic velocity scales and hence the turbulent kinetic energy,  $\kappa_m \sim LK^{1/2}$ . Third, if the flow is relatively weakly stratified, it is reasonable to assume that momentum and the (density) scalar are mixed in qualitatively the same way, so that we have  $Pr_T \sim 1$  and  $\kappa_\rho \sim LK^{1/2} \sim L^{4/3}\mathcal{E}^{1/3}$ , using the classical inertial scaling. Finally, using the definitions for  $\kappa_\rho$  and  $L_O$  in Equations 5 and 9, respectively, we can then write  $\Gamma$  as

$$\Gamma \equiv \frac{\kappa_\rho N^2}{\mathcal{E}} \sim \frac{N^2 L^{4/3}}{\mathcal{E}^{2/3}} = \left( \frac{L_O}{L} \right)^{-4/3} \sim R_{OT}^{-4/3}, \quad 12.$$

under the further (strong) assumption that  $L_T \sim L$ . The evidence that this scaling is applicable to a wide range of oceanic observations is at least suggestive that the observed mixing events may be, at least in some sense, characterized as weakly stratified. Unfortunately, of course, the observational measurements cannot identify the specific character of the flow structures leading to this mixing

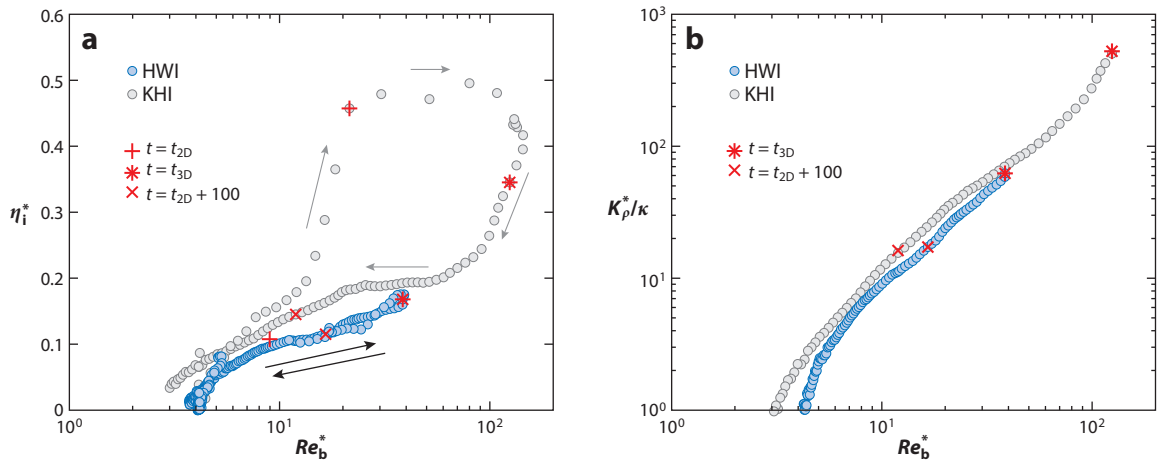
in great detail. In particular, it cannot be assumed that the mixing is associated with classical shear instabilities.

However, at least for KHI-dominated flows at sufficiently high  $Re_0$ , there is clear temporal evolution of  $R_{OT}$ , which must be interpreted differently, as there is no justification to connect the rapidly varying  $L_T$  to any characteristic value of a mixing length. The primary billow roll-up does lead to a rapid increase or flare in any reasonable estimate of  $L_T$ . However, as the flow is not turbulent at this stage of its evolution,  $\mathcal{E} \simeq 0$  and so  $L_O$  remains small. Once the billow saturates, the wide range of secondary instabilities occur and trigger identifiable turbulent motions. This has three interconnected and not fully understood effects. First,  $L_T$  relatively rapidly decays. Second, turbulence onset rapidly increases  $L_O$ . The turbulent mixing event is inherently transient, with an inevitable eventual decay. Although  $L_O$  does increase and then decrease as the turbulence decays, the rate of decay of  $L_O$  is much slower than for  $L_T$ :  $L_O$  slowly burns down. The increase in  $R_{OT}$  with time should thus be interpreted as being essentially due to the slower decay of its numerator compared to its denominator during a KHI-induced mixing event. The third, purely empirically observed effect is that  $L_O$  reaches its maximum value almost precisely coincidentally with the instant when  $L_T \simeq L_O$  (Mashayek et al. 2017a). The dissipation rate and  $Re_B$  are maximum then as well, since in these flows appropriate estimates of the buoyancy frequency,  $N$ , remain close to constant.

**4.1.1. Properties and parameterization of Kelvin–Helmholtz instability-induced mixing events.** Mashayek et al. (2017a) found that the instant when  $L_T \simeq L_O$  also corresponds to the time when the instantaneous mixing efficiency,  $\eta_i^*$ , reaches its maximum value, a mixing property originally hypothesized by Ivey & Imberger (1991). Loosely, energy is being injected (at the overturning scale,  $L_T$ ) exactly at the largest possible scale,  $L_O$ , of a potentially quasi-isotropic inertial range. It is becoming increasingly accepted that the efficiency of such KHI-induced mixing is relatively large (see also Mashayek & Peltier 2013, Mashayek et al. 2013, Salehipour et al. 2018, Salehipour & Peltier 2015) in the sense that instantaneous values of  $\eta_i^* \sim 0.4$ – $0.5$  (as defined in Equation 7) are observed at the optimal instant when  $L_O \sim L_T$ , as shown in **Figure 4a**, with cumulative values of  $\eta_c^* \sim 0.3$ – $0.4$ . Generating such data from simulations is quite challenging, as each individual simulation (with different initial values of  $Ri_0$ ,  $Re$ , and  $Pr$ , for example) requires both significant spatial resolution and temporal evolution. Furthermore, such flows generally exhibit nonmonotonic variability in their mixing properties with respect to external parameters. Just to take one example, Mashayek et al. (2013) demonstrated that the mixing appeared to be most efficient for a mixing layer with  $Ri_0 = 0.16$ : Although the initial growth rate and saturated amplitude of the primary billow decreases monotonically with minimum initial Richardson number, several of the secondary instabilities grow more strongly as  $Ri_0$  increases. There are thus optimal (what might be termed Goldilocks) conditions when the mixing is most efficient.

Time-evolving properties of the flow can be calculated, with associated values of  $Re_b$  or the minimum gradient Richardson number,  $Ri_m$ , determined from (close to instantaneous) spatially averaged properties of the flow. These snapshots of both parameters and mixing properties can then be used in an attempt to populate a parameterization of mixing. Mashayek et al. (2013) showed unsurprisingly that the underlying assumptions of steady-state energetically balanced models such as those proposed by Osborn (1980) are not satisfied. In particular, the subsequent mixing always retains an imprint of the primary overturning triggered by the primary billow, demonstrating again that history does indeed matter.

Such data are definitely relevant to real flows, but they are associated with vigorous, shear-driven, overturning-dominated mixing events. Mashayek et al. (2017b) demonstrated, consistently with observational data, that KHI-induced mixing exhibits a nonmonotonic dependence on  $Re_b$ ,



**Figure 4**

(a) Variation of the irreversible mixing efficiency  $\eta_i^* \equiv \mathcal{M}/(\mathcal{M} + \mathcal{E})$  with the buoyancy Reynolds number  $Re_b^* = \mathcal{E}/(\nu N_*^2)$ , defined using the buoyancy frequency  $N_*$  associated with the sorted density profile, for the simulation prone to Kelvin-Helmholtz instability (KHI) in **Figure 3a–c** (gray circles) and the simulation prone to Holmboe wave instability (HWI) in **Figure 3d–f** (blue circles). The direction of time evolution is indicated by arrows, and the flare of mixing efficiency to relatively large values for the KHI flow is apparent. (b) Variation of the scaled irreversible diapycnal diffusivity  $\kappa_\rho^*/\kappa \equiv \mathcal{M}/(\kappa N_*^2)$  with  $Re_b^*$  for the same simulations for  $t \geq t_{3D}$ . Figure adapted with permission from Salehipour et al. (2016a); copyright 2016 Cambridge University Press.

such that the (irreversible) flux coefficient is

$$\Gamma_c^*(Re_b) = \frac{2\Gamma^m \left( \frac{Re_b}{Re_b^m} \right)^{1/2}}{1 + \left( \frac{Re_b}{Re_b^m} \right)}, \quad 13.$$

where  $Re_b^m \sim \mathcal{O}(100)$  is the value of  $Re_b$  at which  $\Gamma_c^*(Re_b) = \Gamma^m \sim 0.5$  is maximum.

This decrease at large  $Re_b$  is consistent with observations (Monismith et al. 2018) and the influential sheared and forced numerical simulations of Shih et al. (2005), as discussed in detail by Ivey et al. (2008). However, caution must be shown in interpretation for at least two reasons. First,  $Re_b$  is a ratio, and so large values can be associated with large values of  $\mathcal{E}$  or small values of  $N$ , however defined. Indeed, for KHI-induced mixing, it would never be appropriate to characterize the flow as being strongly stratified. Therefore, the observed right flank of decreasing values of  $\Gamma \sim Re_b^{-1/2}$  with larger  $Re_b$  is at best a signal that the efficiency of mixing apparently decreases for sufficiently intense turbulence. (The eddy diffusivity, and so the actual amount of transport, still increases with  $Re_b$  since  $\kappa_\rho$  is proportional to  $\Gamma Re_b$ .)

Second, it is quite difficult to disentangle the dependence on other variables. Salehipour et al. (2016b) presented a mixing parameterization for the (irreversible) mixing efficiency in terms of both  $Re_b$  and  $Ri_g$ , although it is difficult to avoid correlation between these two parameters in KHI-induced turbulent flows, with smaller values of  $Ri_g$  typically being associated with larger values of  $Re_b$ . Therefore, in particular, a decrease in mixing efficiency with increasing  $Re_b$ , consistent with Shih et al. (2005), and hence an apparent right-flank behavior with respect to  $Re_b$ , may be principally due to left-flank behavior of increasing mixing efficiency with increasing Richardson number.

In short, there are four fundamental criticisms of using a time-evolving flow prone to KHI as a test bed for generating data for the construction of a generic mixing parameterization. First, the flow is varying so rapidly that it is hard to argue that the mixing properties at any particular instant

are not strongly influenced by the previous flow history. Second, the initial imprint of the large overturning of the primary billow remains significant throughout the flow evolution. Third, the flow is inherently shear driven, and it is naturally of interest to investigate whether flows forced by an alternative mechanism have similar (or indeed completely different) mixing properties. Finally, the flows are inevitably in a regime that should be characterized as weakly stratified, and so cannot access flow regimes that could reasonably be characterized as strongly stratified.

## 4.2. Mixing Induced by Scouring Holmboe Wave Instability

A natural way to address at least some of these criticisms is to consider the mixing properties induced by the breakdown of HWI. As demonstrated by Salehipour et al. (2016a), analogously to the situation with KHI, there is a qualitative change in the mixing properties associated with the breakdown of HWI when the flow is at a high value of  $Re_0 \gtrsim 4,000$  compared to previous, more moderate values of  $Re_0 \sim \mathcal{O}(500)$  (Smyth & Winters 2003), due principally to  $Re_b$  transiently reaching a sufficiently large value. However, there are certain qualitative differences in the properties of the ensuing turbulence, as explored further by Salehipour et al. (2018). As shown in **Figure 3d–f**, the initially sharp density interface embedded within the shear layer remains, and the ensuing mixing following the breakdown of the primary HWI clearly scours rather than overturns the density interface. Furthermore, as shown in **Figure 4**, the variation of an appropriate measure of the irreversible mixing efficiency with  $Re_b$  is less than for a flow driven by a primary KHI, and in particular there is no evidence of the flare of large optimal values of  $\mathcal{E}$  associated with the KHI overturning.

Indeed, the HWI-induced turbulence appears to burn for longer than the KHI-induced turbulence, generally approaching a quasi-equilibrium, which importantly is apparently largely independent of the initial conditions in terms of the initial value of the ratio of shear layer and density interface depth,  $R$ , or the bulk Richardson number. As shown by Salehipour et al. (2018), the cumulative irreversible flux coefficient for a wide range of initial conditions tends very closely toward  $\Gamma_c^* \simeq 0.2$ , just as hypothesized by Osborn (1980). By considering the horizontal averages of the streamwise velocity and density distribution, and then considering (pointwise in space and time) the implied values of the gradient Richardson number  $Ri_g$  in such a turbulent flow, Salehipour et al. (2018) found that the PDF of  $Ri_g$  is strongly peaked around the Miles–Howard critical value of  $1/4$ . This apparent self-organized criticality is highly suggestive of the hypothesis that  $\Gamma \simeq 0.2$  is indeed characteristic of scouring shear-driven mixing that is approximately in steady state and is also not contaminated by the imprint of an initial vigorous overturning. Furthermore, such mixing would appear not to be characteristic of right-flank behavior, as shown in **Figure 1**, as the effective turbulent Prandtl number (as defined in Equation 6) is  $Pr_T \sim \mathcal{O}(1)$ . Therefore, even for such flows where the bulk Richardson number can have relatively large values, the flow adjusts so that significant turbulent mixing occurs in regions where the stratification is sufficiently weak for  $Pr_T \sim 1$  in a layered flow.

## 5. MIXING IN FORCED FLOWS

There is an argument that a natural approach characterizing irreversible stratified mixing should be based around consideration of continuously forced flows, in which the partitioning (termed the taxation rate above) between turbulent dissipation and irreversible mixing can be determined.

### 5.1. Boundary Forcing: Stratified Plane Couette Flow

An attractive boundary-forced flow geometry is stratified plane Couette flow, i.e., the flow between two plates a distance  $2b$  apart, which are maintained at different velocities,  $\pm U_0$ , with  $Re = U_0 b / \nu$ .

Such flows (when unstratified) are linearly stable for all  $Re$ , and so the transition to turbulence must be inherently of finite amplitude and nonlinear. If the fluid beside the boundaries is maintained at constant different densities,  $\rho_r \mp \rho_0$  [modeling, for example, constant temperature boundaries with a linear equation of state such that  $\rho = \rho_r[1 - \alpha_v \theta]$ , where  $\alpha_v$  is the thermal coefficient of expansion], and if gravity is chosen to point in the wall-normal direction, it is also possible to define a natural bulk Richardson number,  $Ri_0 \equiv g\rho_0 b / \rho_r U_0^2$ . A particular attraction of such a flow (Deusebio et al. 2015, Scotti & White 2016, Zhou et al. 2017a) is that, by construction with such boundary conditions, there is both a friction velocity,  $u_\tau$ , defined in terms of the wall shear stress,  $\tau_w$ , emergent from the structure of the horizontally averaged velocity field, and a constant vertical heat flux,  $q_w$ , through the system, emergent from the structure of the temperature/density field induced by the flow. The constant vertical heat flux ensures that the density flux  $\mathcal{B} > 0$  is sign-definite, thereby avoiding subtleties concerning the use of  $\mathcal{M}$  or  $\chi$  in the definition of  $\Gamma$ .

These quantities then allow the definition of the Obukhov length scale,  $L$ , which in turn can be scaled with the near-wall viscous length scale,  $\delta_v = \nu/u_\tau$ :

$$\begin{aligned} u_\tau^2 &\equiv \frac{\tau_w}{\rho_r} = \nu \left| \frac{\partial \bar{u}}{\partial z} \right|_{z=\pm b}, & q_w &\equiv \kappa \left| \frac{\partial \bar{\theta}}{\partial z} \right|_{z=\pm b}, \\ L &\equiv \frac{u_\tau^3}{k_m g \alpha_v q_w}, & L^+ &\equiv \frac{L}{\delta_v}, \end{aligned} \quad 14.$$

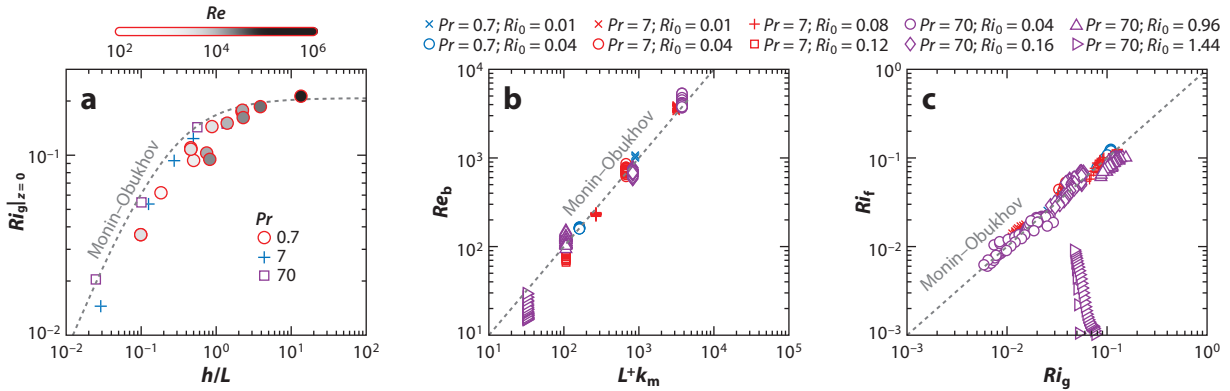
where an overline denotes a horizontal average and  $k_m$  is the von Karman constant for momentum,  $k_m \simeq 0.41$ . Considering the flow energetics, the moving walls via the shear stress are forcing the flow, generically leading to turbulence production. The length scale  $L$  quantifies the relative strength of this turbulent production to the power required to maintain the (wall-normal) flux of buoyancy. As observed by Flores & Riley (2010), for  $L^+ \lesssim 200$ , stratified turbulence cannot be sustained and the flow becomes spatiotemporally intermittent in a fundamentally different way from unstratified intermittency at transitional  $Re$  (Duguet et al. 2010), as it is the stabilizing effects of buoyancy that suppress the wall-normal transfer of momentum, rather than viscous effects, that lead to the turbulence being switched off.

A further attraction of this flow is that it is possible to construct a Monin–Obukhov (M–O) similarity theory for various key properties of the flow in terms of  $L$ , which agrees very well with the results of direct numerical simulations. In particular, when both  $L^+$  and  $\zeta$  (the distance from the wall scaled with  $L$ ) are sufficiently large, the mean density profile, the mean velocity profile across the interior of the flow, and the associated mixing lengths  $l_m$  and  $l_s$  for momentum and scalar mixing, respectively, relating perturbation quantities to mean gradients, should all be self-similar functions of  $\zeta$  alone.

The application of this similarity theory has several significant implications, as discussed in more detail by Deusebio et al. (2015), Scotti & White (2016) and Zhou et al. (2017a). Crucially, the spatial variation of gradient Richardson number can be expressed in terms of  $\zeta$ :

$$\begin{aligned} Ri_g(\zeta) &= \frac{k_m}{k_s} \frac{\zeta^{-1} + \beta_s}{(\zeta^{-1} + \beta_m)^2}, \\ Ri_g|_{z=0} &= \frac{k_m}{k_s} \frac{L/b + \beta_s}{(L/b + \beta_m)^2} \rightarrow \frac{k_m \beta_s}{k_s \beta_m^2} \simeq 0.21 \quad \text{as } b/L \rightarrow \infty, \end{aligned} \quad 15.$$

where  $k_s \simeq 0.48$  is the von Karman constant for the scalar profiles and  $\beta_m = 4.8$  and  $\beta_s = 5.6$  are other empirical constants. The length scale  $L$  must be sufficiently large so that the turbulence does not become intermittent. Increasing  $Re$  (i.e., increasing  $b$ ) will still constrain the Richardson



**Figure 5**

Verification of Monin–Obukhov (M–O) scalings for direct numerical simulations of stratified plane Couette flow. Different symbols denote different combinations of molecular Prandtl number,  $Pr \equiv \nu/\kappa$ , and bulk Richardson number,  $Ri_0 \equiv g\rho_0 b/(\rho_\tau U_0^2)$ . In panel *a*, shading denotes various Reynolds numbers,  $Re \equiv U_0 b/\nu$ , for the simulations with  $Pr = 0.7$  (open circles). All other simulations have  $Re = 4,250$ . Dashed lines denote the predicted M–O scalings: Equation 15 for panel *a*, and equality of  $x$  and  $y$  coordinates for panels *b* and *c*. Shown is the variation of (a) the mid-channel gradient Richardson number,  $Ri_g|_{z=0}$ , with the channel half-depth scaled by the Obukhov length scale,  $b/L$ ; (b) buoyancy Reynolds number,  $Re_b \equiv \mathcal{E}/(\nu N^2)$  (using pointwise estimates in the interior of the channel), with the product of the Obukhov length in wall units and von Karman's constant,  $L^+k_m$ ; and (c) local flux Richardson number,  $Ri_f \equiv \mathcal{B}/\mathcal{P}$ , with gradient Richardson number,  $Ri_g = N^2/S^2$ , both estimated pointwise in the interior of the channel. Figure adapted with permission from Zhou et al. (2017a); copyright 2017 Cambridge University Press.

number at the midpoint of the channel within such a flow to remain small,  $\sim 0.21$ , irrespective of the value of the imposed  $Ri_0$ , as shown in **Figure 5a**.

Furthermore, as argued by Scotti & White (2016), an explicit expression can be derived for  $Re_b = L^+k_m$ , which also agrees very well with numerical simulations, as shown in **Figure 5b**. They also called into question (at least for such flows) whether the buoyancy Reynolds number is actually an appropriate parameter to use to describe the mixing efficiency. Indeed, for such flows, both  $F_h$  and the flux Richardson number,  $Ri_f$ , prove to be closely related to characteristic values of  $Ri_g$  ( $F_h^{-2} \sim Ri_g$  and  $Ri_f \sim Ri_g$ ), and so the turbulent Prandtl number is (close to) one. This is a result that, as demonstrated by Zhou et al. (2017a), is largely independent of the molecular Prandtl number,  $Pr$ , as shown in **Figure 5c** [the anomalous data are effectively at small  $Re_b = \mathcal{O}(10)$ ].

Therefore, in such a steadily forced flow, there is an emergent mixing efficiency that is largely consistent [though slightly larger given  $\Gamma = Ri_f/(1 - Ri_f)$ ] with the hypothesized bounded flux coefficient of Osborn (1980). [Scotti & White's (2016) observations were in close agreement, but they had not accessed the asymptotically possible midplane gradient Richardson number considered by Zhou et al. (2017a).] Additionally, as discussed in detail by Zhou et al. (2017a), there is no evidence of strongly stratified right-flank behavior:  $Ri_f$  increases monotonically (linearly with  $Ri_g$ ), until the flow can no longer support turbulence, toward a value largely consistent with the assumptions of Osborn (1980), and also largely irrespective of the particular imposed values of  $Ri_0$ ,  $Re$ , or  $Pr$ . Equivalently, this particular wall-forced flow never enters a low-turbulent Froude number regime (in fact,  $F_h \gtrsim 2$ ) and so  $\Gamma$  is proportional to  $F_h^{-2}$ , consistent with Ivey & Imberger (1991) and Wells et al. (2010), as well as the scaling analyses (discussed below) associated with body-forced stratified turbulence.

The onset of intermittency in these flows does not appear to be directly related to the Miles–Howard criterion for at least two reasons. First, as demonstrated by Zhou et al. (2017a), even if stratified plane Couette flow is initialized with an initially two-layer density distribution such that

scouring rather than overturning mixing occurs with a sharp and robust interface within the flow, the ensuing mixing still appears to have the character of weak stratification in that the measures of the mixing efficiency scale with  $F_h^{-2}$ .

Second, it is possible to consider a flow where the mean density gradient and velocity gradient are orthogonal, and sufficiently strongly stratified flows still appear unable to sustain turbulence. Such a flow (as discussed by Lucas et al. 2019) may be thought of as having two vertical bounding plates moving relative to each other at velocities  $\pm U_0$  with (vertically) stratified fluid between the two plates, with initially constant buoyancy frequency  $N_0$ . Analogously to stratified Taylor–Couette flow, spontaneous layers form in this flow with a characteristic depth proportional to  $U_0/N_0$ , provided these layers are sufficiently deep, which in some circumstances are, analogously to stratified Taylor–Couette flow, imprinted by a linear instability mechanism identified by Facchini et al. (2018). However, as the (initial) stratification is increased, ultimately the layers become so thin that the stratification suppresses the near-wall generation and maintenance mechanisms for the turbulence, and the flow becomes intermittent, once again apparently precluding sustained turbulence when the flow is (in any reasonable sense) strongly stratified.

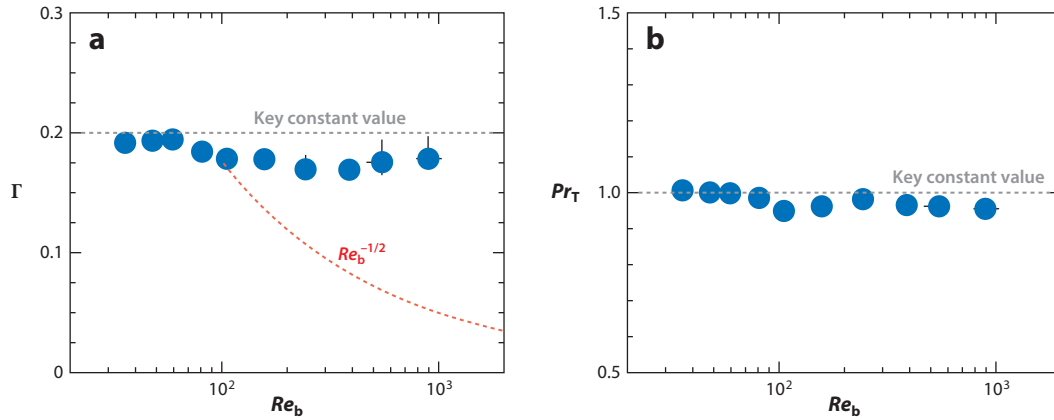
## 5.2. Body-Forced Turbulent Flows

It is straightforward in a numerical simulation to add artificial body forces to inject energy into the flow in an attempt to sustain turbulence, as it seems difficult to boundary-force a strongly stratified turbulent flow. Although it is of course possible to consider a time-dependent body-forced flow [see, for example, the turbulent wake simulations of Zhou & Diamessis (2019)], here our focus is on quasi-steady forced flows.

**5.2.1. Forced stratified shear flows.** The influential simulations of Shih et al. (2005), which exhibited an  $Re_b^{-1/2}$  decrease in mixing efficiency for energetic values of  $Re_b > 100$ , forced the flow to have both a uniform mean shear and density gradient. Chung & Matheou (2012) took a different approach, throttling the mean shear so that the turbulent production stayed constant. They found that with such a control mechanism, the flow evolved into a state once again well described by M–O similarity theory, with  $Ri \sim 0.15$  and (again)  $Pr_T \simeq 1$ .

However, fixing the turbulence production to a particular value makes it somewhat difficult to interpret the dependence of the various flow properties on, in particular, the buoyancy Reynolds number,  $Re_b$ . An alternative, though closely related, three-stage approach was implemented by Portwood et al. (2019) (see also Taylor et al. 2016) for a flow with a fixed uniform (vertical) shear and a fixed background density gradient. First, the viscosity within the simulation is set. Second, a target turbulent kinetic energy  $\mathcal{K}'_t$  is chosen. Third, the value of gravity (and hence the characteristic value of the coupling gradient Richardson number,  $Ri_g$ ) is adjusted using a second-order mass-spring-damper control system to converge the turbulent kinetic energy toward the target value. If  $\mathcal{K}'$  is increasing,  $Ri$  is increased, while if  $\mathcal{K}'$  is decreasing,  $Ri$  is decreased, until  $\mathcal{K}' \rightarrow \mathcal{K}'_t$ .

Crucially, several key quantities are emergent for a range of  $\mathcal{K}'_t$ . The dissipation rates of both kinetic energy,  $\mathcal{E}$ , and buoyancy variance,  $\chi$ , vary naturally with  $\mathcal{K}'_t$ , but their relative size, and so the mixing taxation rate, is not predetermined, and neither is the final value of  $Ri$ . As  $\mathcal{K}'_t$  is varied, it is reasonable to expect that  $Re_b$  will vary, but the specific numerical value is also emergent. For  $30 < Re_b < 900$ , the emergent value of  $Ri \simeq 0.15$ – $0.16$  is close to constant. For this range of  $Re_b$ , Portwood et al. (2019) found (yet again)  $Pr_T \simeq 1$  and the appropriate emergent flux coefficient,  $\Gamma = \chi/\mathcal{E} \simeq 0.2$ , as shown in **Figure 6**, which is close to the upper bound proposed by Osborn (1980), perhaps unsurprisingly since this flow more closely satisfies the underlying assumptions. There is no evidence of the  $Re_b^{-1/2}$  scaling. The requirement that the turbulence remains sustained



**Figure 6**

Variation with the buoyancy Reynolds number,  $Re_b \equiv \mathcal{E}/(\nu N^2)$ , of (a) the turbulent flux coefficient,  $\Gamma \equiv \mathcal{B}/\mathcal{E}$ , and (b) the turbulent Prandtl number,  $Pr_T \equiv \kappa_m/\kappa_\rho$ , for constant vertical shear, constant buoyancy frequency simulations using the control method described by Portwood et al. (2019). Thin black lines mark estimates of variability arising from the control method. The scalings  $\Gamma = 0.2$  and  $\Gamma \propto Re_b^{-1/2}$  for  $Re_b \gtrsim 100$  are shown with dashed lines in panel a, and the scaling  $Pr_T = 1$  is shown with a dashed line in panel b. Figure adapted with permission from Portwood et al. (2019); copyright 2019 American Physical Society.

appears to force the flow evolution to be weakly stratified in that the mixing of the scalar is closely coupled to the mixing of momentum. Furthermore, these calculations are consistent with the argument of Scotti & White (2016) that the hybrid parameter,  $Re_b$ , is not the most appropriate parameter to describe turbulent mixing. The control algorithm leads to effectively constant values of both  $Ri_g$  and (for this flow at least) the turbulent Froude number,  $F_h \simeq 5$ . For small  $Ri_g$ , there is certainly evidence of  $Ri_f \simeq Ri_g$ , although more strongly stratified sheared flows do not seem to be able to support sustained turbulence.

**5.2.2. Body-forced stratified turbulence: dependence on  $F_h$ .** In an important paper, Maffioli et al. (2016) presented convincing arguments that the most appropriate measure to parameterize mixing generically is the (turbulent) Froude number, and that it is plausible that there are at least two regimes. They identified these regimes through numerical simulations with artificial body forces designed to maintain (unsheared) stratified turbulence. In particular, for simulations with low Froude numbers,  $F_h < 0.2$ , they found that it was necessary to force only (vertical) vortical modes with relatively small horizontal wavenumbers, as isotropic forcing of such flows leads to nonstationary flows with substantial shear. This curious phenomenon is unexplained, and it would be of interest to investigate further.

The high-Froude number regime once again exhibited a scaling  $\Gamma \propto F_h^{-2}$ , consistent with the results for weak stratification in shear flows. Building upon this, Garanaik & Venayagamoorthy (2019) presented scaling arguments leading to this behavior, arguing that at high  $F_h$  the key timescale for  $\chi$  (and hence the mixing) is the (unmodified) turbulence timescale,  $\mathcal{K}/\mathcal{E}$ . Maffioli et al. (2016) also observed that, with this particular forcing, turbulence could indeed be maintained to relatively low values of  $F_h \sim 0.02$ , although for these flows inevitably  $Re_b$  dropped to smaller values of  $\mathcal{O}(10)$  in their simulations. They observed nonmonotonic variation of  $\Gamma$ , with a peak value  $\Gamma \simeq 0.5$  for  $F_h \simeq 0.3$ , before an apparent asymptotic value of  $\Gamma \simeq 0.33$ , suggestive of independence of  $\Gamma \sim F_h^0$  as  $F_h \rightarrow 0$ .

Garanaik & Venayagamoorthy (2019) argued that such a constant asymptotic value is consistent with the assumption that the dominant timescale for both  $\chi$  and  $\mathcal{E}$  is the buoyancy timescale,



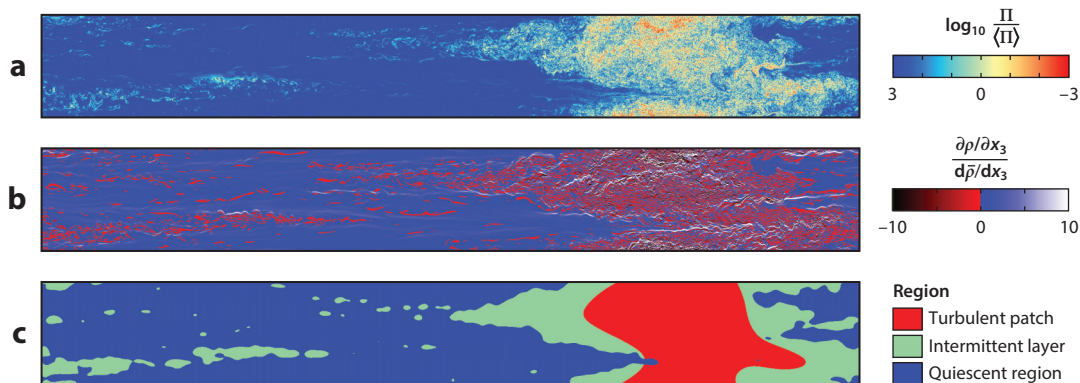
$1/N$ , which thus cancels out in the ratio of  $\Gamma$ . Indeed, they argued in favor of a hybrid intermediate regime where the timescale for  $\chi$  is the buoyancy timescale, while the timescale for  $\mathcal{E}$  is the turbulence timescale, thus leading to a prediction of  $\Gamma \propto F_h^{-1}$  for intermediate values of the Froude number. Their collated, significantly scattered data are not inconsistent with this intermediate scaling. [They also presented arguments to infer mixing properties from various powers of the ratio  $L_E/L_O$ , where the Ellison scale,  $L_E$ , is effectively being used as a proxy for the Thorpe scale,  $L_T$ , analogously to the arguments presented above from Ijichi & Hibiya (2018).]

It is fair to say that it is still an open question of great interest whether there is nonmonotonicity in the dependence of mixing efficiency with  $F_h$ , and so the key, fundamental fluid dynamical questions concerning the parameterization of mixing are still open. Indeed, it is not clear what the specific numerical value of  $\Gamma$  might be expected to be as  $F_h \rightarrow 0$ ; for example Garanaik & Venayagamoorthy (2019) argued that a broader data set suggests  $\Gamma \rightarrow 0.5$  in the limit  $F_h \rightarrow 0$ . This particular constant value is actually consistent with Khani's (2018) large eddy simulations of forced stratified turbulence, when  $Pr_T \simeq 1$ , remembering that  $\Gamma/(1 + \Gamma)$  corresponds to Khani's definition of mixing efficiency. Indeed, even larger implied values of  $\Gamma$ , although still asymptotically independent of the strength of the stratification, occur when a sufficiently strong density interface is scoured by impinging vortex rings, as considered by Olsthoorn & Dalziel (2018). It should certainly be investigated further what influence the particular form of forcing might have on the observed mixing. In particular, it is of interest whether the mixing might best be categorized as shear driven or convective in character, since convective overturning (for example, associated with the breaking of an internal wave) generically leads to substantially more efficient irreversible mixing.

**5.2.3. Layering and local dynamics.** A final and very important point is a recurring theme: Layering of strongly stratified turbulent flows appears to occur spontaneously, and so identifying flows as strongly stratified must be treated with great caution. There is evidence that such body-forced flows spontaneously organize into quiescent, locally more strongly stratified regions and vigorously turbulent, more weakly stratified regions, calling into question the classification of flows in terms of single, volume-averaged quantities.

Howland et al. (2020) qualitatively demonstrated such a spontaneous layering by conducting stratified turbulent simulations with a range of different forcings. Furthermore, Portwood et al. (2016) developed a robust method to identify vigorous turbulent patches (associated with local density overturnings), intermittent layers, and quiescent regions within a forced stratified flow with imposed constant  $N$ . A typical example is shown in **Figure 7**. This automated method can be applied to simulation data from flows with different overall  $F_h$  and different  $Re_b$ . These parameters are typically defined in terms of a volume-averaged dissipation rate: for the two most extreme cases,  $F_h = 0.0743$  and  $Re_b = 218$  for a weakly stratified simulation (F1), while  $F_h = 0.0152$  and  $Re_b = 13.4$  for a strongly stratified simulation (F3), possibly just satisfying the conditions for being in the LAST regime. F1 is identified as being 96% patch. The absolute contributions from within those patches to the dissipation rates across the whole volume are over 99% for both  $\mathcal{E}$  and  $\chi$ . Unsurprisingly, the conditional value of  $Re_b = 240$  within the patch is similar to the volume average of 218.

The picture is markedly different for the strongly stratified simulation F3 (a slice of which is shown in **Figure 7**). For this simulation, only 4% of the flow is vigorously turbulent. However the conditional value of  $Re_b = 177$  within those patches (i.e., using the average value of the local dissipation rate in parts of the flow identified as patches) is quite similar to the equivalent value for the (virtually all-turbulent) simulation F1. Furthermore, these small patches contribute 56% of the total dissipation rate and 66% of the total buoyancy variance destruction rate. Therefore, the truly turbulent regions in the strongly stratified flow are very similar to those in the weakly



**Figure 7**

A vertical slice through the strongly stratified simulation F3 illustrating (a) the logarithm of the potential enstrophy normalized by its domain average; (b) the pointwise vertical density gradient, normalized by the background density gradient; and (c) the results of the automated flow classification method developed by Portwood et al. (2016). The coloring in panel *b* highlights the convectively unstable fluid elements in red. The coloring in panel *c* corresponds to the region classification, where red indicates a turbulent patch region, green an intermittent layer region, and blue a quiescent flow region. Figure adapted with permission from Portwood et al. (2016); copyright 2016 Cambridge University Press.

stratified flow, and the fundamental difference is that the strongly stratified flow is just substantially more spatiotemporally intermittent and, once again, layered. This is yet more evidence that it is actually extremely difficult for a (in some sense quasi-steady) flow to be both strongly stratified and vigorously turbulent. It is clearly of great interest either to identify (and measure the properties of) such a flow or, on theoretical grounds, to establish the impossibility of sufficiently vigorous turbulence of surviving for a long period in a strongly stratified environment without the formation of a layer-interface structure. As discussed in this review, such layering does appear to arise in a wide range of flows with a variety of forcing mechanisms, and the characteristics of such flows are clearly deserving of greater consideration and study in the future.

## SUMMARY POINTS

1. Although there has been a large range of deeply insightful research contributions to our understanding of transition, turbulence, and irreversible mixing in stratified fluids, it still remains extremely difficult to say anything generic about mixing.
2. History matters in describing stratified turbulent mixing, with marked differences between turbulence triggered by shear instabilities and turbulence forced either at a boundary or by artificial body forces.
3. It is important to consider structures rather than statistics when analyzing the turbulence and ensuing irreversible mixing in stratified fluids.
4. Mixing associated with overturning instabilities, such as the classic Kelvin–Helmholtz instability (KHI), is qualitatively different from scouring instabilities, such as the Holmboe wave instability (HWI).
5. Evidence is accumulating that a very wide range of instability and turbulence mechanisms and mixing processes can lead generically to a layer-interface structure. Under certain circumstances (Taylor & Zhou 2017), such a structure can be robust.

6. Weakly stratified sustained turbulent flows often have  $Pr_T \simeq 1$ , and thus  $Ri_f \sim Ri_g \lesssim 0.2$ , largely consistent with the classical model of Osborn (1980).
7. There is evidence suggesting that strongly stratified turbulence should be thought of as patches of vigorous turbulence in local regions of relatively weak stratification embedded in relatively quiescent regions of significantly stronger stratification.

## FUTURE ISSUES

1. Is it even reasonable to attempt to develop parametric models for such important quantities as mixing efficiency (however defined) in terms of natural nondimensional parameters, such as Richardson numbers, buoyancy Reynolds numbers, turbulent Froude numbers, and the Prandtl number?
2. Does nonmonotonicity of dependence on a parameter ever actually occur, and in particular, is right-flank behavior real, where the mixing efficiency decreases as (for example)  $Ri_g$  increases, where all other parameters are properly controlled, without hidden correlations between independent parameters?
3. When combined with other recent techniques from modern data science, what role can automated structure identification (e.g., Portwood et al. 2016) play in improving models for mixing? For example, Salehipour & Peltier (2019) used deep convolutional neural networks to develop an alternative data-driven parameterization of shear-driven mixing in flows susceptible to primary KHI- and HWI-induced mixing, which out-performed previous physics-based parameterizations.
4. Is it worth investigating whether data-driven techniques such as machine learning and artificial intelligence can be applied productively to model stratified mixing (Bolton & Zanna 2019)? Through vastly enhanced computational, experimental, and observational resources available to the research community, there is a proliferation of data available to analyze and from which to extract insight. The challenge is to use the right combination of data-driven and physical-process-based approaches such as those described here to unravel the fascinating mystery of irreversible mixing in stratified fluids.

## DISCLOSURE STATEMENT

The author is not aware of any biases that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

I owe a huge debt to Paul Linden and Dick Peltier. I have been fortunate to work with many inspirational students, postdoctoral fellows, and collaborators. I especially wish to acknowledge both the financial support of EPSRC (Engineering and Physical Sciences Research Council) and the contributions of all those involved in the 2013–2018 Programme Grant EP/K034529/1 entitled Mathematical Underpinnings of Stratified Turbulence. I am grateful to the National Science Foundation for the sustained support of the Summer Study Program in Geophysical Fluid Dynamics at Woods Hole Oceanographic Institution. In particular, I thank all the participants in the

2019 program on Stratified Turbulence and Ocean Mixing Processes, which had a major influence on this review.

## LITERATURE CITED

- Alexakis A. 2005. On Holmboe's instability for smooth shear and density profiles. *Phys. Fluids* 17:084103
- Alexakis A. 2009. Stratified shear flow instabilities at large Richardson numbers. *Phys. Fluids* 21:054108
- Almalkie S, de Bruyn Kops SM. 2012. Kinetic energy dynamics in forced, homogeneous, and axisymmetric stably stratified turbulence. *J. Turbul.* 13:N29
- Andrews DG. 1981. A note on potential-energy density in a stratified compressible fluid. *J. Fluid Mech.* 107:227–36
- Arthur RS, Venayagamoorthy SK, Koseff JR, Fringer OB. 2017. How we compute  $N$  matters to estimates of mixing in stratified flows. *J. Fluid Mech.* 831:R2
- Augier P, Billant P. 2011. Onset of secondary instabilities on the zigzag instability in stratified fluids. *J. Fluid Mech.* 682:120–31
- Baines PG, Mitsudera H. 1994. On the mechanism of shear flow instabilities. *J. Fluid Mech.* 276:327–42
- Balmforth NJ, Llewellyn Smith SG, Young WR. 1998. Dynamics of interfaces and layers in a stratified turbulent fluid. *J. Fluid Mech.* 355:329–58
- Balmforth NJ, Roy A, Caulfield CP. 2012. Dynamics of vorticity defects in stratified shear flow. *J. Fluid Mech.* 694:292–331
- Bartello P, Tobias SM. 2013. Sensitivity of stratified turbulence to the buoyancy Reynolds number. *J. Fluid Mech.* 725:1–22
- Basak S, Sarkar S. 2006. Dynamics of a stratified shear layer with horizontal shear. *J. Fluid Mech.* 568:19–54
- Billant P, Chomaz J-M. 2000a. Experimental evidence for a new instability of a vertical columnar vortex pair in a strongly stratified fluid. *J. Fluid Mech.* 418:167–88
- Billant P, Chomaz J-M. 2000b. Theoretical analysis of the zigzag instability of a vertical columnar vortex pair in a strongly stratified fluid. *J. Fluid Mech.* 419:29–63
- Billant P, Chomaz J-M. 2001. Self-similarity of strongly stratified inviscid flows. *Phys. Fluids* 13:1645–51
- Bolton T, Zanna L. 2019. Applications of deep learning to ocean data inference and subgrid parameterization. *J. Adv. Model. Earth Syst.* 11:376–99
- Brethouwer G, Billant P, Lindborg E, Chomaz J-M. 2007. Scaling analysis and simulation of strongly stratified turbulent flows. *J. Fluid Mech.* 585:343–68
- Browand FK, Winant CD. 1973. Laboratory observations of shear layer instability in a stratified fluid. *Bound.-Layer Meteorol.* 5:67–87
- Carpenter JR, Tedford EW, Heifetz E, Lawrence GA. 2011. Instability in stratified shear flow: review of a physical interpretation based on interacting waves. *Appl. Mech. Rev.* 64:060801
- Carpenter JR, Tedford EW, Rahmani M, Lawrence GA. 2010. Holmboe wave fields in simulation and experiment. *J. Fluid Mech.* 648:205–23
- Caulfield CP. 1994. Multiple linear instability of layered stratified shear flow. *J. Fluid Mech.* 258:255–85
- Caulfield CP, Peltier WR. 1994. Three dimensionalization of the stratified mixing layer. *Phys. Fluids* 6:3803
- Caulfield CP, Peltier WR. 2000. The anatomy of the mixing transition in homogeneous and stratified free shear layers. *J. Fluid Mech.* 317:179–93
- Caulfield CP, Peltier WR, Yoshida S, Ohtani M. 1995. An experimental investigation of the instability of a shear flow with multilayered density stratification. *Phys. Fluids* 7:3028–41
- Caulfield CP, Yoshida S, Peltier WR. 1996. Secondary instability and three-dimensionalization in a laboratory accelerating shear layer with varying density differences. *Dyn. Atmos. Oceans* 23:125–38
- Chung D, Matheou G. 2012. Direct numerical simulation of stationary homogeneous stratified sheared turbulence. *J. Fluid Mech.* 696:434–67
- Churilov SM. 2016. Stability of shear flows with multilayered density stratification and monotonic velocity profiles having no inflection points. *Geophys. Astrophys. Fluid Dyn.* 110:78–108
- Crapper PF, Linden PF. 1974. Structure of turbulent density interfaces. *J. Fluid Mech.* 65:45–63

- Davies Wykes MS, Hughes GO, Dalziel SB. 2000. On the meaning of mixing efficiency for buoyancy-driven mixing in stratified turbulent flows. *J. Fluid Mech.* 781:261–75
- de Bruyn Kops S. 2015. Classical scaling and intermittency in strongly stratified Boussinesq turbulence. *J. Fluid Mech.* 775:436–63
- Deloncle A, Billant P, Chomaz J-M. 2008. Nonlinear evolution of the zigzag instability in stratified fluids: a shortcut on the route to dissipation. *J. Fluid Mech.* 599:229–39
- Deusebio E, Caulfield CP, Taylor JR. 2015. The intermittency boundary in stratified plane Couette flow. *J. Fluid Mech.* 781:298–329
- Dillon TM. 1982. Vertical overturns: a comparison of Thorpe and Ozmidov length scales. *J. Geophys. Res. Oceans* 87:9601–13
- Dimotakis PE. 2005. Turbulent mixing. *Annu. Rev. Fluid Mech.* 37:329–56
- Duguet Y, Schlatter P, Henningson DS. 2010. Formation of turbulent patterns near the onset of transition in plane Couette flow. *J. Fluid Mech.* 650:119–29
- Eaves TS, Balmforth NJ. 2019. Instability of sheared density interfaces. *J. Fluid Mech.* 860:145–71
- Eaves TS, Caulfield CP. 2017. Multiple instability of layered stratified plane Couette flow. *J. Fluid Mech.* 813:250–78
- Facchini G, Favier B, Le Gal P, Wang M, Le Bars M. 2018. The linear instability of the stratified plane Couette flow. *J. Fluid Mech.* 853:205–34
- Falder M, White NJ, Caulfield CP. 2016. Seismic imaging of rapid onset of stratified turbulence in the South Atlantic ocean. *J. Phys. Oceanogr.* 46:1023–44
- Fernando HJS. 1991. Turbulent mixing in stratified fluids. *Annu. Rev. Fluid Mech.* 23:455–93
- Ferrari R, Wunsch C. 2009. Ocean circulation kinetic energy: reservoirs, sources, and sinks. *Annu. Rev. Fluid Mech.* 41:253–82
- Flores O, Riley JJ. 2010. Analysis of turbulence collapse in stably stratified surface layers using direct numerical simulation. *Bound.-Layer Meteorol.* 129:241–59
- Garanaik A, Venayagamoorthy SK. 2019. On the inference of the state of turbulence and mixing efficiency in stably stratified flows. *J. Fluid Mech.* 867:323–33
- Garaud P. 2018. Double-diffusive convection at low Prandtl number. *Annu. Rev. Fluid Mech.* 50:275–98
- Gargett AE, Osborn TR, Nasmith PW. 1984. Local isotropy and the decay of turbulence in a stratified fluid. *J. Fluid Mech.* 144:231–80
- Gibson CH. 1980. Fossil temperature, salinity, and vorticity turbulence in the ocean. In *Marine Turbulence*, ed. J Nihoul, pp. 221–57. Amsterdam: Elsevier
- Gregg MC, D’Asaro EA, Riley JJ, Kunze E. 2018. Mixing efficiency in the ocean. *Annu. Rev. Mar. Sci.* 10:443–73
- Guha A, Lawrence GA. 2014. A wave interaction approach to studying non-modal homogeneous and stratified shear instabilities. *J. Fluid Mech.* 755:336–64
- Guyez E, Flor J-B, Hopfinger EJ. 2007. Turbulent mixing at a stable density interface: the variation of the buoyancy flux–gradient relation. *J. Fluid Mech.* 577:127–36
- Hogg AMcC, Ivey GN. 2003. The Kelvin–Helmholtz to Holmboe instability transition in stratified exchange flows. *J. Fluid Mech.* 477:339–62
- Holford JM, Linden PF. 1999. Turbulent mixing in a stratified fluid. *Dyn. Atmos. Ocean* 30:173–98
- Holliday D, McIntyre ME. 1981. On potential-energy density in an incompressible stratified fluid. *J. Fluid Mech.* 107:221–25
- Holmboe J. 1962. On the behaviour of symmetric waves in stratified shear layers. *Geophys. Publ.* 24:67–113
- Howard LN. 1961. Note on a paper of John W. Miles. *J. Fluid Mech.* 10:509–12
- Howland CJ, Taylor JR, Caulfield CP. 2018. Testing linear marginal stability in stratified shear layers. *J. Fluid Mech.* 839:R4
- Howland CJ, Taylor JR, Caulfield CP. 2020. Mixing in forced stratified turbulence and its dependence on large-scale forcing. *J. Fluid Mech.* 898:A7
- Ijichi T, Hibiya T. 2018. Observed variations in turbulent mixing efficiency in the deep ocean. *J. Phys. Oceanogr.* 48:1815–30
- Itsweire EC, Koseff JR, Briggs DA, Ferziger JH. 1993. Turbulence in stratified shear flows: implications for interpreting shear-induced mixing in the ocean. *J. Phys. Oceanogr.* 23:1508–22

- Ivey GN, Bluteau CE, Jones NL. 2018. Quantifying diapycnal mixing in an energetic ocean. *J. Geophys. Res. Oceans* 123:346–57
- Ivey GN, Imberger J. 1991. On the nature of turbulence in a stratified fluid. Part I: the energetics of mixing. *J. Phys. Oceanogr.* 21:650–58
- Ivey GN, Winters KB, Koseff JR. 2008. Density stratification, turbulence, but how much mixing. *Annu. Rev. Fluid Mech.* 40:169–84
- Kaminski AK, Caulfield CP, Taylor JR. 2014. Transient growth in strongly stratified shear layers. *J. Fluid Mech.* 758:R4
- Kaminski AK, Caulfield CP, Taylor JR. 2017. Nonlinear evolution of linear optimal perturbations of strongly stratified shear layers. *J. Fluid Mech.* 758:R4
- Kaminski AK, Smyth WD. 2019. Stratified shear instability in a field of pre-existing turbulence. *J. Fluid Mech.* 862:639–58
- Kato H, Phillips OM. 1969. On the penetration of a turbulent layer into stratified fluid. *J. Fluid Mech.* 37:643–55
- Khani S. 2018. Mixing efficiency in large-eddy simulations of stratified turbulence. *J. Fluid Mech.* 849:373–94
- Kimura Y, Herring JR. 2012. Energy spectra of stably stratified turbulence. *J. Fluid Mech.* 698:19–50
- Koop CG, Browand FK. 1979. Instability and turbulence in a stratified fluid with shear. *J. Fluid Mech.* 93:135–59
- Lawrence GA, Browand FK, Redekopp LG. 1991. The stability of a sheared density interface. *Phys. Fluids A* 3:2360–70
- Le Bars M, Le Gal P. 2007. Experimental analysis of the stratorotational instability in a cylindrical Couette flow. *Phys. Rev. Lett.* 99:064502
- Lee V, Caulfield CP. 2001. Nonlinear evolution of a layered stratified shear flow. *Dyn. Atmos. Oceans* 34:103–24
- Lefauve A, Partridge JL, Zhou Q, Dalziel SB, Caulfield CP, Linden PF. 2018. The structure and origin of confined Holmboe waves. *J. Fluid Mech.* 848:508–44
- Lindborg E. 2006. The energy cascade in a strongly stratified fluid. *J. Fluid Mech.* 550:207–42
- Linden PF. 1979. Mixing in stratified fluids. *Geophys. Astrophys. Fluid Dyn.* 13:431–64
- Lorenz EN. 1955. Available potential energy and the maintenance of the general circulation. *Tellus* 7:157–67
- Lucas D, Caulfield CP. 2017. Irreversible mixing by unstable periodic orbits in buoyancy dominated stratified turbulence. *J. Fluid Mech.* 832:R1
- Lucas D, Caulfield CP, Kerswell RR. 2017. Layer formation in horizontally forced stratified turbulence: connecting exact coherent structures to linear instabilities. *J. Fluid Mech.* 832:409–37
- Lucas D, Caulfield CP, Kerswell RR. 2019. Layer formation and relaminarisation in plane Couette flow with spanwise stratification. *J. Fluid Mech.* 868:97–118
- Maffioli A, Brethouwer G, Lindborg E. 2016. Mixing efficiency in stratified turbulence. *J. Fluid Mech.* 794:R3
- Mashayek A, Caulfield CP, Peltier WR. 2013. Time-dependent, non-monotonic mixing in stratified shear flows: implications for oceanographic estimates of buoyancy flux? *J. Fluid Mech.* 736:570–93
- Mashayek A, Caulfield CP, Peltier WR. 2017a. Role of overturns in optimal mixing in stratified mixing layers. *J. Fluid Mech.* 826:522–52
- Mashayek A, Peltier WR. 2012a. 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 WR. 2012b. 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 WR. 2013. Shear-induced mixing in geophysical flows: Does the route to turbulence matter to its efficiency? *J. Fluid Mech.* 725:216–61
- Mashayek A, Salehipour H, Bouffard D, Caulfield CP, Ferrari R, et al. 2017b. Efficiency of turbulent mixing in the abyssal ocean circulation. *Geophys. Res. Lett.* 44:6296–306
- Mater BD, Schaad SM, Venayagamoorthy SK. 2015. Relevance of the Thorpe length scale in stably stratified turbulence. *Phys. Fluids* 25:076604
- Mater BD, Venayagamoorthy SK. 2014a. A unifying framework for parameterizing stably stratified shear-flow turbulence. *Phys. Fluids* 26:036601
- Mater BD, Venayagamoorthy SK. 2014b. The quest for an unambiguous parameterization of mixing efficiency in stably stratified geophysical flows. *Geophys. Res. Lett.* 41:4646–53

- Mater BD, Venayagamoorthy SK, St. Laurent L, Moum JN. 2015. Biases in Thorpe-scale estimates of turbulence dissipation. Part I: assessments from large-scale overturns in oceanographic data. *J. Phys. Oceanogr.* 45:2497–521
- Miles JW. 1961. On the stability of heterogeneous shear flows. *J. Fluid Mech.* 10:496–508
- Molemaker MJ, McWilliams JC, Yavneh I. 2001. Instability and equilibration of centrifugally stable stratified Taylor-Couette flow. *Phys. Rev. Lett.* 86:5270–73
- Monismith SG, Koseff JR, White B. 2018. Mixing efficiency in the presence of stratification: When is it constant? *Geophys. Res. Lett.* 45:5627–34
- Odier P, Ecke RE. 2017. Stability, intermittency and universal Thorpe length distribution in a laboratory turbulent stratified shear flow. *J. Fluid Mech.* 815:243–56
- Ogblethorpe RLF, Caulfield CP, Woods AW. 2013. Spontaneous layering in stratified turbulent Taylor-Couette flow. *J. Fluid Mech.* 721:R3
- Olsthoorn J, Dalziel SB. 2018. Vortex-ring-induced stratified mixing: mixing model. *J. Fluid Mech.* 837:129–46
- Osborn TR. 1980. Estimates of the local-rate of vertical diffusion from dissipation measurements. *J. Phys. Oceanogr.* 10:83–89
- Osborn TR, Cox CS. 1972. Oceanic fine structure. *Geophys. Fluid Dyn.* 3:321–45
- Otheguy P, Chomaz J-M, Billant P. 2006. Elliptic and zigzag instabilities on co-rotating vertical vortices in a stratified fluid. *J. Fluid Mech.* 553:253–72
- Palmer TL, Fritts DC, Andreassen O. 1996. Evolution and breakdown of Kelvin-Helmholtz billows in stratified compressible flows. Part II: instability structure, evolution, and energetics. *J. Atmos. Sci.* 53:3192–212
- Palmer TL, Fritts DC, Andreassen O, Lie I. 1994. Three-dimensional evolution of Kelvin-Helmholtz billows in stratified compressible flow. *Geophys. Res. Lett.* 21:2287–90
- Park YG, Whitehead JA, Gnanadeskian A. 1994. Turbulent mixing in stratified fluids: layer formation and energetics. *J. Fluid Mech.* 279:279–311
- Parker JP, Caulfield CP, Kerswell RR. 2019. Kelvin-Helmholtz billows above Richardson number 1/4. *J. Fluid Mech.* 879:R1
- Patterson MD, Caulfield CP, McElwaine JN, Dalziel SB. 2006. Time-dependent mixing in stratified Kelvin-Helmholtz billows: experimental observations *Geophys. Res. Lett.* 33:L15608
- Peltier WR, Caulfield CP. 2003. Mixing efficiency in stratified shear flows. *Annu. Rev. Fluid Mech.* 35:135–67
- Phillips OM. 1972. Turbulence in a strongly stratified fluid—Is it unstable? *Deep Sea Res.* 19:79–81
- Ponetti G, Balmforth NJ, Eaves TS. 2019. Instabilities in a staircase stratified shear flow. *Geophys. Astrophys. Fluid Dyn.* 112:1–19
- Portwood GD, de Bruyn Kops SM, Caulfield CP. 2019. Asymptotic dynamics of high dynamic range stratified turbulence. *Phys. Rev. Lett.* 122:194504
- Portwood GD, de Bruyn Kops SM, Taylor JR, Salehipour H, Caulfield CP. 2016. Robust identification of dynamically distinct regions in stratified turbulence. *J. Fluid Mech.* 807:R2
- Praud O, Fincham AM, Sommeria J. 2005. Decaying grid turbulence in a strongly stratified fluid. *J. Fluid Mech.* 522:1–33
- Riley JJ, de Bruyn Kops SM. 2003. Dynamics of turbulence strongly influenced by buoyancy. *Phys. Fluids* 15:2047–59
- Riley JJ, Lelong MP. 2000. Fluid motions in the presence of strong stable stratification. *Annu. Rev. Fluid Mech.* 32:613–57
- Riley JJ, Lindborg E. 2008. Stratified turbulence: a possible interpretation of some geophysical turbulence measurements. *J. Atmos. Sci.* 65:2416–24
- Salehipour H, Caulfield CP, Peltier WR. 2016a. Turbulent mixing due to the Holmboe wave instability at high Reynolds number. *J. Fluid Mech.* 856:228–56
- Salehipour H, Peltier WR. 2015. Diapycnal diffusivity, turbulent Prandtl number and mixing efficiency in Boussinesq stratified turbulence. *J. Fluid Mech.* 775:464–500
- Salehipour H, Peltier WR. 2019. Deep learning of mixing by two ‘atoms’ of stratified turbulence. *J. Fluid Mech.* 861:R4
- Salehipour H, Peltier WR, Caulfield CP. 2018. Self-organised criticality of turbulence in strongly stratified mixing layers. *J. Fluid Mech.* 856:228–56

- 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
- Salehipour H, Peltier WR, Whalen CB, MacKinnon JA. 2016b. A new characterization of the turbulent diapycnal diffusivities of mass and momentum in the ocean. *Geophys. Res. Lett.* 43:3370–79
- Schmitt RW. 1994. Double diffusion in oceanography. *Annu. Rev. Fluid Mech.* 26:255–85
- Scotti A. 2015. Biases in Thorpe-scale estimates of turbulence dissipation. Part II: energetics arguments and turbulence simulations. *J. Phys. Oceanogr.* 45:2497–521
- Scotti A, White B. 2014. Diagnosing mixing in stratified turbulent flows with a locally defined available potential energy. *J. Fluid Mech.* 740:114–35
- Scotti A, White B. 2016. The mixing efficiency of stratified turbulent boundary layers. *J. Phys. Oceanogr.* 46:3181–91
- 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
- Smyth WD, Carpenter JR. 2019. *Instability in Geophysical Flows*. Cambridge, UK: Cambridge Univ. Press
- Smyth WD, Klaassen GP, Peltier WR. 1988. Finite amplitude Holmboe waves. *Geophys. Astrophys. Fluid Dyn.* 43:181–222
- Smyth WD, Moum JN. 2000. Length scales of turbulence in stably stratified mixing layers. *Phys. Fluids* 12:1327–42
- Smyth WD, Moum JN. 2013. Marginal instability and deep cycle turbulence in the eastern equatorial Pacific Ocean. *Geophys. Res. Lett.* 40:6181–85
- Smyth WD, Moum JN, Caldwell DR. 2001. The efficiency of mixing in turbulent patches: inferences from direct simulations and microstructure observations. *J. Phys. Oceanogr.* 31:1969–92
- Smyth WD, Nash JD, Moum JN. 2019. Self-organized criticality in geophysical turbulence. *Sci. Rep.* 9:3747
- Smyth WD, Winters KB. 2003. Turbulence and mixing in Holmboe waves. *J. Phys. Oceanogr.* 33:694–711
- Sreenivasan KR. 2019. Turbulent mixing: a perspective. *PNAS* 116:18175–83
- Tailleux R. 2009. On the energetics of stratified turbulent mixing, irreversible thermodynamics, Boussinesq models and the ocean heat engine controversy. *J. Fluid Mech.* 638:339–82
- Taylor GI. 1931. Effect of variation in density on the stability of superposed streams of fluid. *Proc. R. Soc. A* 132:499–532
- Taylor JR, de Bruyn Kops SM, Caulfield CP, Linden PF. 2019. Testing the assumptions underlying ocean mixing methodologies using direct numerical simulations. *J. Phys. Oceanogr.* 49:2761–79
- Taylor JR, Deusebio E, Caulfield CP, Kerswell RR. 2016. A new method for isolating turbulent states in transitional stratified plane Couette flow. *J. Fluid Mech.* 808:R1
- Taylor JR, Zhou Q. 2017. A multi-parameter criterion for layer formation in a stratified shear flow using sorted buoyancy coordinates. *J. Fluid Mech.* 823:R5
- Tedford EW, Pieters R, Lawrence GA. 2009. Symmetric Holmboe instabilities in a laboratory exchange flow. *J. Fluid Mech.* 636:137–53
- Thorpe SA. 1973. Experiments on instability and turbulence in a stratified shear flow. *J. Fluid Mech.* 61:731–51
- Thorpe SA. 1977. Turbulence and mixing in a Scottish loch. *Philos. Trans. R. Soc. A* 286:125–81
- Thorpe SA. 2016. Layers and internal waves in uniformly stratified fluids stirred by vertical grids. *J. Fluid Mech.* 793:380–413
- Thorpe SA, Liu Z. 2009. Marginal instability? *J. Phys. Oceanogr.* 39:2373–81
- Tseng Y, Ferziger JH. 2001. Mixing and available potential energy in stratified flows. *Phys. Fluids* 13:1281–93
- Turner JS. 1973. *Buoyancy Effects in Fluids*. Cambridge, UK: Cambridge Univ. Press
- Venayagamoorthy SK, Koseff JR. 2016. On the flux Richardson number in stably stratified turbulence. *J. Fluid Mech.* 644:359–69
- Venayagamoorthy SK, Stretch DD. 2010. On the turbulent Prandtl number in homogeneous stably stratified turbulence. *J. Fluid Mech.* 644:359–69
- Villiermaux E. 2019. Mixing versus stirring. *Annu. Rev. Fluid Mech.* 51:245–73
- Waite ML, Bartello P. 2004. Stratified turbulence dominated by vortical motion. *J. Fluid Mech.* 517:281–308
- Waite ML, Bartello P. 2008. Instability and breakdown of a vertical vortex pair in a strongly stratified fluid. *J. Fluid Mech.* 517:281–308



- Waterhouse AF, MacKinnon JA, Nash JD, Alford MH, Kunze E, et al. 2014. Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *J. Phys. Oceanogr.* 44:1854–72
- Wells M, Cenedese C, Caulfield CP. 2010. The relationship between flux coefficient and entrainment ratio in density currents. *J. Phys. Oceanogr.* 40:2713–27
- Winters KB, D’Asaro EA. 1996. Diascalar flux and the rate of fluid mixing. *J. Fluid Mech.* 317:179–93
- Winters KB, Lombard PN, Riley JJ, D’Asaro EA. 1995. Available potential-energy and mixing in density-stratified fluids. *J. Fluid Mech.* 289:115–28
- Woods AW, Caulfield CP, Landel JR, Kuesters A. 2010. Non-invasive mixing across a density interface in a turbulent Taylor–Couette flow. *J. Fluid Mech.* 663:347–57
- Wunsch C, Ferrari R. 2004. Vertical mixing, energy and the general circulation of the oceans. *Annu. Rev. Fluid Mech.* 36:281–314
- Zhou Q, Diamessis PJ. 2019. Large-scale characteristics of stratified wake turbulence at varying Reynolds number. *Phys. Rev. Fluids* 4:084802
- Zhou Q, Taylor JR, Caulfield CP. 2017a. Self-similar mixing in stratified plane Couette flow for varying Prandtl number. *J. Fluid Mech.* 820:86–120
- Zhou Q, Taylor JR, Caulfield CP, Linden PF. 2017b. Diapycnal mixing in layered stratified plane Couette flow quantified in a tracer-based coordinate. *J. Fluid Mech.* 821:198–229
- Zhu DZ, Lawrence GA. 2001. Holmboe’s instability in exchange flows. *J. Fluid Mech.* 429:391–409