Diapycnal mixing of passive tracers by Kelvin–Helmholtz instabilities

Jared Penney[1],†, Yves Morel[1], Peter Haynes[2], Francis Auclair[3] and Cyril Nguyen[3]

[1]LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse 31400, France

[2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK

[3]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse 31400, France

This paper considers the diapycnal transport of passive tracers during a Kelvin–Helmholtz mixing event. Numerical simulations of a traditional Kelvin–Helmholtz (KH) configuration of a stratified shear flow are extended to include layers of passive tracer at different locations relative to the shear layer. The evolution of the tracers during the simulation is followed and is analysed using different theoretical approaches. One is to consider the evolution via the distribution in density–tracer space which clearly reveals how the tracers are redistributed across isopycnals by the mixing driven by the growing and saturating KH billow. The shape of the distribution places constraints on the redistribution of the tracer and, for this problem of symmetrically stratified shear, it is shown that the distribution typically tends to a compact form, with significant regions that are nearly linear. The redistribution across isopycnals is also considered via a diffusion equation for the tracer relative to coordinates based on the geometry of density surfaces. The equation is a generalisation of an equation previously derived for transport of density in these coordinates and includes an extra eddy term that arises because there is variation of the tracer along density surfaces. Under certain circumstances and at later stages of the flow, the eddy term can be neglected, and the evolution of the mean tracer profile can be adequately represented using a simple diffusion equation where diffusivity is defined as the effective diffusivity of density, scaled by the molecular diffusion of the tracer.

Key words: stratified turbulence

1. Introduction

On the large scale, both the atmosphere and ocean have stable density stratifications, and processes by which fluid properties are mixed in the vertical direction are crucial to both the circulation and to the distribution of chemical and biological species. Shear between interacting masses of fluid is an integral component of the transfer of energy from the largest scales of flow down to the smallest dissipation and mixing scales through the formation of instabilities and turbulence. Away from boundaries vertical mixing is likely to occur through intermittent events triggered by dynamical shear instabilities, such

† Email address for correspondence: jared.penney@legos.obs-mip.fr

as Kelvin–Helmholtz (KH) instabilities (e.g. Smyth & Moum 2012). The evolution of the flow in simple configurations of KH instabilities has been studied in detail, both in the laboratory (Thorpe 1973; Caulfield, Yoshida & Peltier 1996; Patterson *et al.* 2006) and through numerical simulation (Klaassen & Peltier 1985, 1989; Caulfield & Peltier 1994; Scinocca 1995; Alexakis 2009; Carpenter, Balmforth & Lawrence 2010; Mashayek & Peltier 2012*a*,*b*, 2013). Shear instabilities have also been observed and recorded in nature. For example, in the atmosphere, clear air turbulence represents a specific hazard for aircraft, and is thought to be largely driven by shear instabilities (Browning & Watkins 1970; Fritts & Rastogi 1985), while 'roll-up patterns' in clouds are known to be caused by KH instabilities (Fritts & Rastogi (1985) and the references therein). Measurements indicative of KH instabilities have been recorded near regions of shear along oceanic thermoclines (Woods 1968; Marmorino 1987), during the downwelling tidal phase near seamounts (van Haren & Gostiaux 2010) and the surface mixed layer has been observed to deepen due to shear instabilities along its base (Lincoln, Rippeth & Simpson 2016). KH instabilities have also been observed along interfaces in estuaries (Geyer & Smith 1987; Geyer *et al.* 2010), which can be chemically and nutritionally rich due to material in the river outflows.

Density is a dynamically active tracer in the sense it plays an active role in driving the flow. However, processes such as KH instabilities are important in the vertical mixing of passive tracers, which have no direct effect on the flow, as is the case with certain low concentration or neutrally buoyant species (Warhaft 2000; Canuto, Cheng & Howard 2011), but which are important for other reasons. These include, in the ocean, nutrients and microscopic biological species (e.g. Vaquer-Sunyer & Duarte 2008; Brierley & Kingsford 2009) and, in the atmosphere, chemical species that are radiatively active or affect human health (e.g. Seinfeld & Pandis 1998; Yang *et al.* 2015). In the ocean, studies of the effect of KH instabilities have been motivated by the likely importance of mixing processes on the overall density stratification and hence the large-scale circulation (e.g. Wunsch & Ferrari 2004), and by analogy they are also potentially important to the large-scale distributions of biological species. In the atmosphere, the importance of turbulent mixing on the vertical transport of density and chemical species across isentropic surfaces is uncertain because such transport may also occur through radiative heating and cooling (e.g. Sparling *et al.* 1997). Nonetheless, the intermittent turbulence is likely to control the rate of mixing between air masses of different origin and hence the horizontal and vertical scales of variation of chemical species, particularly in regions such as near the tropopause and in the midlatitude 'surf zone' of the winter stratosphere where there is strong chemical inhomogeneity. Given the likely importance of intermittent turbulence mixing events on the large-scale distribution of atmospheric and oceanic tracers, understanding the details of these events and their effect on tracers is thus of major importance. This motivates the investigation, via numerical simulation, of the mixing of passive tracers in KH instability reported in this paper.

In most atmospheric and oceanic models, the details of intermittent diapycnal mixing events are not simulated directly. Such mixing is typically regarded as a subgrid-scale process that must be parameterised, often through a turbulent or eddy diffusivity (see, for example, Seinfeld & Pandis (1998, chap. 18) for a summary of functional forms of eddy diffusivities in atmospheric models, or Griffies (2004, chap. 7) for a discussion of parameterised phenomena in ocean models). The magnitude of the required eddy diffusivity has been estimated in various ways, for example, with ocean tracer release experiments to estimate diffusivity based on spreading across isopycnals (e.g. Ledwell, Watson & Law 1993), and extracting estimates from atmospheric radar data (e.g. Fukao *et al.* 1994). Parameterisations have been drastically improved and are able to reproduce

the dynamics of the mixed layer from high frequency to seasonal time scales. However, in most developments and existing parameterisations, turbulent diffusivities of passive and active tracers are assumed for simplicity to be equal (or proportional). Whether or not this assumption is justified is not yet clear. One aspect of this uncertainty is the effect of different molecular diffusivities between passive and active species or indeed between different passive species. This certainly needs to be taken into account when considering turbulent mixing in the ocean, where diffusivities of heat and salt, both of which may contribute to density, differ by a factor of approximately 100. The effect of differing molecular diffusivities on vertical transport in KH mixing events has been considered by Smyth, Nash & Moum (2005), who, for practical reasons, set the diffusivity of heat to be approximately 7 times that of salt. In addition, even if the molecular diffusivities of the passive tracer and the density are the same, the assumption of equal turbulent diffusivities may be an oversimplification if they have different large-scale sources and sinks. The geometry of the two may then be different at small scales and the resulting differences in molecular diffusive transport may potentially lead to differences in turbulent diapycnal transport at macroscales. There has been some investigation of this topic using numerical simulation (Nagata & Komori 2001) and some discussion of the potential importance in the atmospheric boundary layer (Li, Bou-Zeid & De Bruin 2012). Further work is needed to evaluate under which circumstances turbulent diffusivities are equal or similar for all tracers, and to assess the implications for representation of diapycnal fluxes of different tracers in oceanic and atmospheric models.

The research reported in this paper examines the vertical mixing of passive tracers by KH instabilities, focusing in particular on whether the extent to which what is known regarding the mixing of density can also be applied to passive tracers, and what other factors need to be taken into account. Numerical simulations of KH instabilities in a standard flow configuration, including a set of passive tracers, are presented (§§ 2 and 3), and analysed using various techniques (§§ 4 and 5). One technique is to use a density–tracer scatter plots to examine the relative distribution of density and tracer and to relate this to the mixing. The shape of the scatter plot places constraints on the redistribution of the tracer and, in particular, it is shown that, for the flow configuration considered, the relationship between the density and the tracer is often in part piecewise linear or close to piecewise linear in the end state. Another technique used in previous studies is to exploit a tracer-based coordinate system in which the effect of mixing can be represented completely by an effective diffusivity. The question addressed here is whether the effective diffusivity for density also usefully represents the mixing of other tracers, which is tested by varying the initial distribution of tracer (§ 7).

2. Numerical model and flow configuration

2.1. *Governing equations and numerical model*

The governing equations are non-dimensionalised using typical scales for the study of KH instabilities, with tildes denoting the dimensional forms of distance $\boldsymbol{x} = (x, y, z)$, time, t, velocity, $\boldsymbol{u} = (u, v, w)$, density, ρ, tracer concentration, ϕ, and pressure, p,

$$x = \tilde{x}/h, \quad y = \tilde{y}/h, \quad z = (\tilde{z} - z_0)/h, \quad t = \tilde{t}U_0/h, \quad \boldsymbol{u} = \tilde{\boldsymbol{u}}/U_0,$$

$$\rho = (\tilde{\rho} - \rho_0)/\Delta\rho, \quad \phi = \tilde{\phi}/\Delta\phi, \quad p = \tilde{p}/\rho_0 U_0^2.$$

The dimensional parameters used here are the value defining the width of the pycnocline and the shear layer, h, the midpoint of the stratification and shear layer, z_0, half of the change in velocity across the two layers, U_0, the density at the midpoint of the density

distribution, ρ_0, half of the change in density across the two layers, $\Delta\rho$, and the maximum initial tracer concentration, $\Delta\phi$. Using these dimensionless variables, the continuity equation and equations of conservation of momentum, density and passive tracer can be written in their incompressible and Boussinesq forms as

$$\nabla \cdot \boldsymbol{u} = 0, \tag{2.1a}$$

$$\partial\boldsymbol{u}/\partial t + \boldsymbol{u} \cdot \nabla\boldsymbol{u} = -\nabla p - Ri_0\rho\hat{\boldsymbol{k}} + Re_0^{-1}\nabla^2\boldsymbol{u}, \tag{2.1b}$$

$$\partial\rho/\partial t + \boldsymbol{u} \cdot \nabla\rho = (Re_0 Pr)^{-1}\nabla^2\rho, \tag{2.1c}$$

$$\partial\phi/\partial t + \boldsymbol{u} \cdot \nabla\phi = (Re_0 Sc)^{-1}\nabla^2\phi, \tag{2.1d}$$

with $\hat{\boldsymbol{k}}$ denoting the unit vector of the vertical axis. The dimensionless parameters included in these equations are the initial Reynolds number, $Re_0 = U_0 h/\nu$, the Prandtl number, $Pr = \nu/\kappa_\rho$, and the Schmidt number, $Sc = \nu/\kappa_\phi$, where ν is the kinematic viscosity, and κ_ρ and κ_ϕ are the prescribed molecular diffusivity constants for density and the passive tracer, respectively. Ri_0 is the minimum initial Richardson number, which follows the standard definition for the gradient Richardson number,

$$Ri = -\frac{g}{\rho_0}\frac{\dfrac{\partial\tilde{\rho}}{\partial\tilde{z}}}{\left(\dfrac{\partial\tilde{u}}{\partial\tilde{z}}\right)^2}, \tag{2.2}$$

and can be expressed as

$$Ri_0 \approx \frac{g\Delta\rho h}{\rho_0 U_0^2}, \tag{2.3}$$

based on the initial conditions presented in the following section. The configuration parameters are chosen so that the necessary criterion for stratified shear instability, $Ri_0 < 1/4$, is satisfied. For the sake of completeness, relevant dimensional parameters prescribed in the simulations presented in this paper are given in table 2 in appendix A.

The numerical simulations presented in this paper were performed using the non-hydrostatic, non-Boussinesq version of the Coastal and Regional Ocean Community model (CROCO). This model was adapted from the Regional Ocean Modeling System (ROMS, Shchepetkin & McWilliams 2005) to include non-hydrostatic and compressible effects (Auclair *et al.* 2018). While numerical simulations of KH instabilities are often considered in a periodic domain with free-slip rigid lid conditions for the upper and lower boundaries (e.g. Mashayek & Peltier 2012*b*, 2013), the implementation presented here utilises a free-surface upper boundary, and a flat, solid, bottom boundary, with periodic lateral boundary conditions in the *x*- and *y*-directions (streamwise and spanwise directions, respectively). The existence of a free surface and compressibility adds two dynamical processes (surface and acoustic waves) compared to more traditional studies of KH instabilities in incompressible, unbounded or rigid lid flows. It has been verified that, with the chosen configurations where the instability develops far from the vertical boundaries, the impact of these additional processes on the results is negligible when compared to the traditional configuration (see appendix A for further details regarding the implementation of CROCO and the discussion in the concluding section). However, in certain circumstances, surface and acoustic waves may play a role in modifying the turbulent cascade (see, for example, Shete & Guha 2018).

2.2. *Initial conditions*

The initial conditions (ICs) for the simulations presented in this paper follow those from previous numerical studies of Kelvin–Helmholtz instabilities (e.g. Salehipour, Peltier & Mashayek 2015). In their dimensionless forms, the extents of the domain in the streamwise, spanwise and vertical directions are given by L_x, L_y and L_z, respectively. Both two-dimensional (2-D) and three-dimensional (3-D) configurations are examined, with periodic boundary conditions in the horizontal directions, a flat, free-slip, rigid bottom and a free surface in the vertical direction. The initial density distribution is defined as two constant-density layers separated by a strongly stratified pycnocline, with a weak stable background stratification superimposed,

$$\rho\,(\boldsymbol{x}, 0) = -\beta z - \tanh(z). \tag{2.4}$$

The linear background term βz is a minor modification to the standard configuration for the purpose of defining the scatter plots in density–tracer space discussed in § 4.1 over the entire vertical domain in physical space (i.e. each density value initially corresponds to a unique value of height). Note that the definition of the initial Richardson number (2.3) ignores the weak linear background term of the initial density profile, since it was confirmed to have a negligible influence on the development of the instability when compared to the initial stratification without the weak linear background ($\beta \approx 3.5 \times 10^{-3} \ll 1$ in this study).

The initial velocity field is given by

$$\left.\begin{aligned} u\,(\boldsymbol{x}, t = 0) &= U\,(z) + u'(\boldsymbol{x}), \\ v\,(\boldsymbol{x}, t = 0) &= v'(\boldsymbol{x}), \\ w\,(\boldsymbol{x}, t = 0) &= w'(\boldsymbol{x}). \end{aligned}\right\} \tag{2.5}$$

Here, $U(z)$ is the initial background flow providing the shear and is defined by a hyperbolic tangent profile, with the upper layer moving leftward, and the lower layer moving rightward,

$$U(z) = -\tanh(z). \tag{2.6}$$

and u', v' and w' are small amplitude perturbations, required to kickstart the instability, defined as

$$u'(x, y, z) = \epsilon f'(z) \sin\left(\frac{2\pi n_x}{L_x}x\right)\left(1 + \epsilon_{3\text{-D}} \sin\left(\frac{2\pi n_y}{L_y}y\right)\right),$$

$$v'(x, y, z) = 0,$$

$$w'(x, y, z) = -\epsilon f(z)\frac{2\pi n_x}{L_x}\cos\left(\frac{2\pi n_x}{L_x}x\right)\left(1 + \epsilon_{3\text{-D}} \sin\left(\frac{2\pi n_y}{L_y}y\right)\right). \tag{2.7}$$

This choice of functional form for the perturbation ensures that the initial velocity field is non-divergent, which is important in the context of this numerical implementation. The function

$$f(z) = 1 - \tanh^2\left(\frac{z}{\alpha}\right), \tag{2.8}$$

ensures that the initial perturbation is localised within the region of the shear, with $\alpha = 4.29$; $\epsilon = 0.01$ is a dimensionless parameter that sets the magnitude of the perturbation in the streamwise direction, while n_x and n_y set the wavelengths of the initial perturbation in
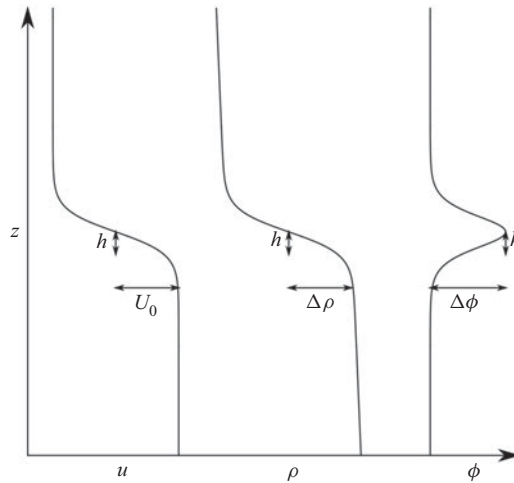
FIGURE 1. (Left to right) Initial vertical profiles of the velocity, density, and passive tracer fields for the described simulations with relevant dimensional parameters.

the x- and y-directions, respectively; ϵ_{3D} sets the magnitude of the spanwise perturbations, and for 3-D configurations, $\epsilon_{3D} = 0.2$. For 2-D configurations, both $n_y = 0$ and $\epsilon_{3D} = 0$.

In addition to the dynamical fields, passive tracer fields are defined by an initial profile of the form

$$\phi(\boldsymbol{x}, 0) = a\mathrm{sech}^2(az - b), \tag{2.9}$$

where a is a dimensionless parameter that determines the width of the tracer distribution, while ensuring the total amount of tracer remains constant between simulations, and b indicates the offset of the tracer from the pycnocline. The tracer is thus initially distributed in a layer, with its maximum initial value located at $z = b/a$. Both the vertical position and the width of the tracer layer will be varied in simulations presented later in this paper, but for the different dynamical configurations discussed first in § 2.3, the tracer layer is collocated with the midpoints of the stratification and the shear ($b = 0$), and the maximum initial tracer value is 1 ($a = 1$). Sketches of the initial vertical profiles of the shear, density and passive tracer fields are depicted in figure 1, with associated dimensional parameters.

2.3. *Description of dynamical configurations*

The primary experiment presented in this section compares simulations of three different dynamical configurations, all with identical initial background fields and prescribed parameters. The relevant dimensionless parameters used in all simulations are $Ri_0 = 0.1158$, $Re = 2000$, $Pr = 1$ and $Sc = 1$. The non-dimensional wavelength of the fastest-growing mode as predicted by the Taylor–Goldstein equations for these given initial background velocity and density profiles is $\lambda_{KH} \approx 14.3$. As such, the length of the domain is the streamwise direction was chosen as $L_x = \lambda_{KH}$, so that a single KH billow develops, or $L_x = 2\lambda_{KH}$, so that two KH billows develop, possibly leading to pairing in 2-D configurations. The parameters N_x, N_y and N_z are the number of grid points in the streamwise, spanwise and vertical directions, respectively. Combined with the length of the domain in each direction, this leads to a consistent dimensionless grid spacing in the x-, y- and z-directions of $\Delta x = \Delta y = \Delta z = 0.0559$. Δt is the dimensionless barotropic time step. Table 1 presents the parameters that vary between each dynamical configuration.

Configuration	L_x	L_y	L_z	n_x	n_y	N_x	N_y	N_z	Δt
I	14.3	—	28.6	1	0	256	—	512	3.125×10^{-4}
II	28.6	—	28.6	2	0	512	—	512	3.125×10^{-4}
III	14.3	5.3625	28.6	1	4	256	96	512	1×10^{-3}

TABLE 1. Relevant computational parameters of the three dynamical configurations presented in this paper.

Additional parameters relevant to the non-Boussinesq components of the model are listed in appendix A. The three dynamical configurations are defined as follows:

(i) Configuration I is 2-D, with $L_x = \lambda_{KH}$ and $n_x = 1$, $n_y = 0$ (so that $v = 0$ at all times). The initial perturbation corresponds to the most unstable wavelength and only one billow develops in this simulation.

(ii) Configuration II is also 2-D, except that $L_x = 2 \lambda_{KH} = L_z$ and $n_x = 2$. The length of the domain and wavelength of the perturbation allow for the formation of two billows, which eventually give way to pairing and additional mixing.

(iii) Configuration III is 3-D, with spanwise parameters $L_y = 0.375 \lambda_{KH}$, $n_y = 4$, while the streamwise and vertical parameters are the same as configuration I ($L_x = \lambda_{KH}$, $n_x = 1$). This allows for mixing by spanwise secondary instabilities that lead to the breakdown of the primary KH billow.

Due to the inclusion of the third dimension and resulting spanwise instabilities, configuration III provides a more realistic simulation than the other two cases. Configurations I and II, while less physically realistic, provide additional pathways to turbulence and mixing, despite showing nearly identical initial behaviour. Comparing the three configurations therefore permits the evaluation of the impact of the route to turbulence on tracer mixing.

3. Diapycnal mixing of a passive tracer by Kelvin–Helmholtz billows: comparison of dynamical configurations

This section presents 2-D and 3-D numerical simulations of stratified turbulence developing from shear-induced Kelvin–Helmholtz billows. Each simulation undergoes the same initial 2-D growth as predicted by linear theory, with later stages leading to different forms of turbulence and mixing.

3.1. *Evolution of the different dynamical configurations*

In order to follow the time evolution of the three dynamical configurations, the mean background and perturbation kinetic energy (KE) are defined in the standard way (e.g. Mashayek & Peltier 2013). This requires definitions for the mean background and perturbation velocity fields. The mean background velocity field is a function of z only, and is given as

$$\overline{u}(z) = \frac{1}{L_x L_y} \iint u \, dx \, dy. \tag{3.1}$$

The 2-D velocity perturbation is a function of x and z, defined as

$$\boldsymbol{u}_{\text{2-D}} = \frac{1}{L_y} \int (\boldsymbol{u} - \overline{\boldsymbol{u}}(z)) \, \mathrm{d}y. \tag{3.2}$$

Finally, the 3-D velocity perturbation can be expressed using the background velocity and the 2-D velocity perturbation,

$$\boldsymbol{u}_{\text{3-D}} = \boldsymbol{u} - \boldsymbol{u}_{\text{2-D}} - \overline{\boldsymbol{u}}(z). \tag{3.3}$$

Following these definitions, the background KE, KE due to 2-D perturbations, and KE due to the 3-D perturbations are defined as

$$\overline{\mathcal{K}} = \frac{1}{L_z} \int \frac{\overline{\boldsymbol{u}}^2}{2} \, \mathrm{d}z, \tag{3.4}$$

$$\mathcal{K}_{\text{2-D}} = \frac{1}{L_x L_z} \iint \frac{\boldsymbol{u}_{\text{2-D}}^2}{2} \, \mathrm{d}x \, \mathrm{d}z, \tag{3.5}$$

$$\mathcal{K}_{\text{3-D}} = \frac{1}{L_x L_y L_z} \iiint \frac{\boldsymbol{u}_{\text{3-D}}^2}{2} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z, \tag{3.6}$$

respectively, noting that the mean KE can be expressed as the sum of these three values,

$$\mathcal{K} = \frac{1}{V} \int \frac{\boldsymbol{u}^2}{2} \, \mathrm{d}V = \overline{\mathcal{K}} + \mathcal{K}_{\text{2-D}} + \mathcal{K}_{\text{3-D}}. \tag{3.7}$$

The general evolution of the dynamical configurations is tracked in figure 2, which plots their mean KE as functions of time. Note that although each of the simulations are performed on domains with different volumes, the initial mean KE is identical for all three cases. This provides a straightforward depiction of the divergent behaviour between the three cases. The 2-D and 3-D perturbation KE will be discussed further in § 4.2 when discussing tracer mixing.

Figures 3 and 4 present the evolution of the density and passive tracer, respectively, for the three dynamical configurations. Each row corresponds to a specific time depicted on the KE time series of figure 2, which indicates an important stage in KH development, or points at which the development of the configurations diverge. The distinct phases enumerated in figure 2 are as follows:

(i) The initial period of linear growth, as predicted by inviscid Taylor–Goldstein theory. This phase is nearly identical for all three simulations, although some weak spanwise effects are visible in the 3-D simulation. Each simulation exhibits stirring of the low and high-density regions by 2-D KH billows with a wavelength close to that of the fastest-growing mode predicted by Taylor–Goldstein theory.

(ii) The branching point between 2-D and 3-D simulations due to the onset of secondary instabilities. All three configurations exhibit secondary shear instabilities along the tilted pycnocline, and secondary convective instabilities within the vortices, while spanwise instabilities begin to emerge in the 3-D configuration. The density field develops alternating layers of high and low density, while the passive tracer takes the shape of ellipses with maximum values centred within the billows, gradually decreasing to no tracer at the billow exterior. A thin strand of low tracer
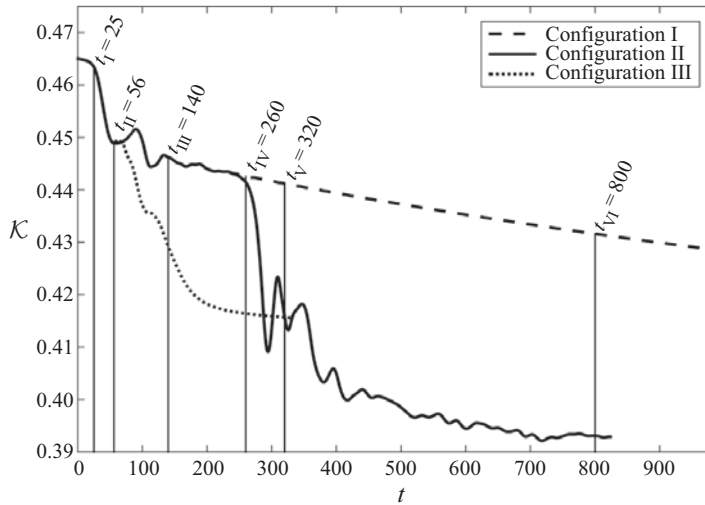
FIGURE 2. Time evolution of the mean KE $\mathcal{K}$ for configuration I (dashed line), configuration II (solid line) and configuration III (dotted line). The labelled times correspond to the rows of cross-sections plotted in figures 3 and 4.

    concentration occurs along the braids, while even lower concentrations of tracer are visible between the horizontal edges of the billows.

(iii) First onset of small-scale features, resulting to greater mixing in all simulations. In the 2-D configurations, this leads to the density within the billows becoming nearly uniform. In the 3-D configuration, this leads to a wider pycnocline that is no longer unstable to shear instabilities. The 2-D cases keep their tracer maxima focused at the centre of the billows, although some tracer is pulled from the high concentration regions towards the outsides of the billows. Meanwhile, the density field experiences greater mixing due to the appearance of the small-scale features, and becomes nearly homogeneous within the billows. The secondary instabilities in configuration III mix the high tracer concentration region at the centre of the billow, redistributing the tracer throughout the rest of the pycnocline.

(iv) The onset of pairing in configuration II, which results in much greater mixing further into the low- and high-density layers, while mixing the concentrated regions of tracer within the two billows with the regions external to the billows without tracer. The single billow of configuration I maintains a nearly constant shape and size.

(v) Second onset of small scales due to enhanced stretching from the pairing in configuration II. Configuration III has reached an essentially steady state.

(vi) The nearly steady final phase of the 2-D configurations. The single billow of configuration I continuously rolls and experiences slow diffusion of density at the edges of the vortex. The pairing of configuration II induces stretching, causing new small-scale instabilities to develop, and enhancing the homogenisation of tracer fields within the billows. This leads to formation of a single larger billow that eventually remains nearly steady.

4. Density–tracer scatter plots

    This section examines density–tracer scatter plots as a diagnostic technique to describe the diapycnal fluxes of a passive tracer.
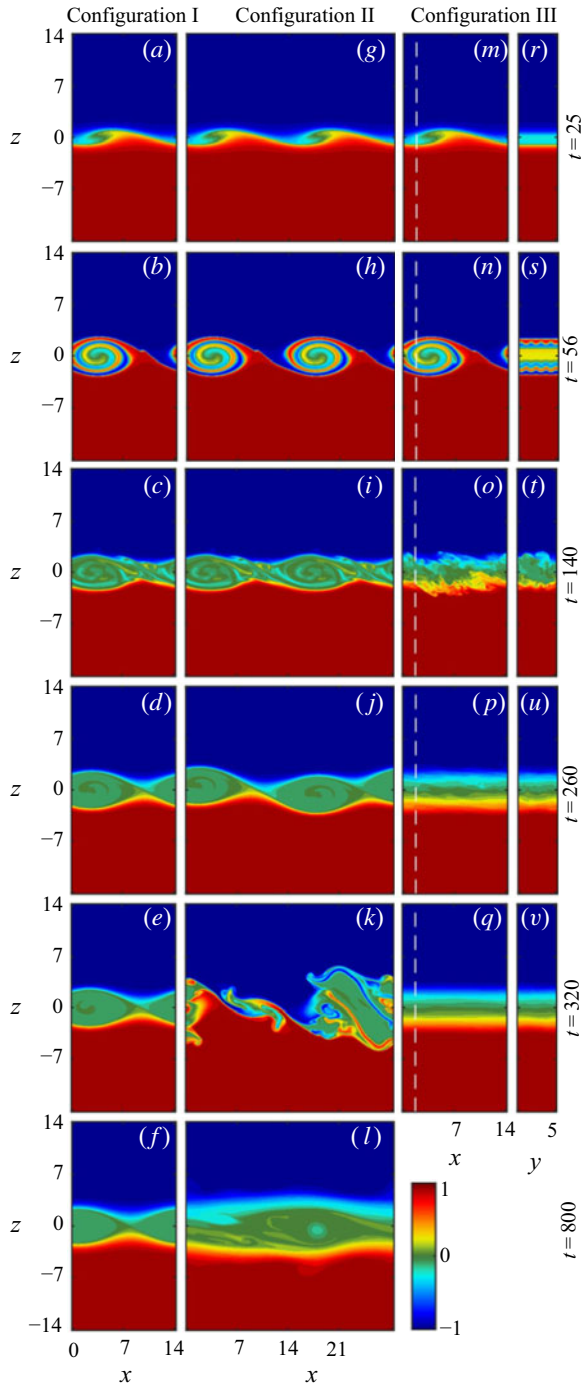
FIGURE 3. Evolution of slices of the density field for configurations I (*a–f*), II (*g–l*) and III (in the streamwise (*m–q*) and spanwise (*r–v*) directions) at the times indicated in figure 2. The dashed lines in the images of the third column indicate the position of the spanwise slices depicted in the fourth column, while the streamwise slices of the third column correspond to the right edge of the spanwise slices.
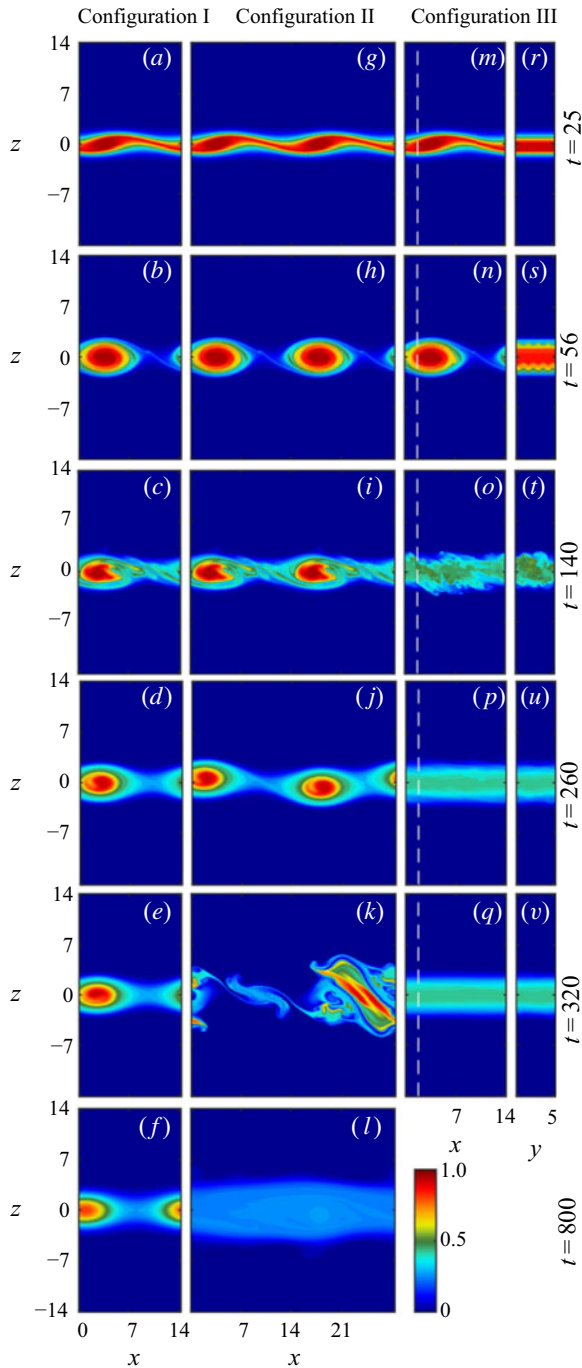
FIGURE 4. Evolution of slices of the passive tracer field for configurations I (*a–f*), II (*g–l*) and III (in the streamwise (*m–q*) and spanwise (*r–v*) directions) at the times indicated in figure 2. The dashed lines in the images of the third column indicate the position of the spanwise slices depicted in the fourth column, while the streamwise slices of the third column correspond to the right edge of the spanwise slices.
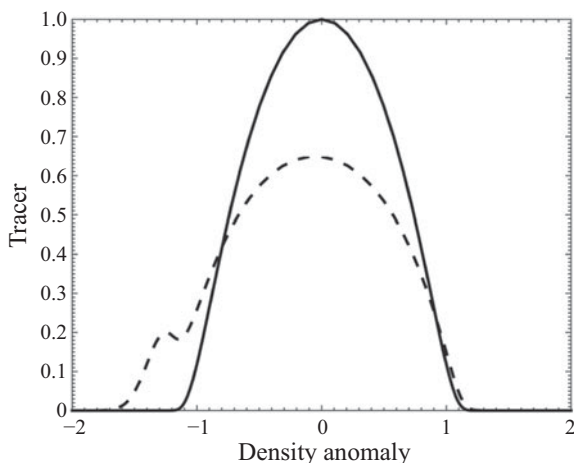
FIGURE 5. Sketch of the initial scatter plot of tracer as a function of density associated with the profiles given by (2.4) and (2.9) where the pycnocline and tracer layer are collocated (solid line). Hypothetical evolution of the scatter plot after a mixing event showing diapycnal fluxes of tracer (dashed line). In the new profile, mixing has led to a flux of tracer towards lower densities.

4.1. *Weighted density–tracer scatter plots*

For definiteness, mixing is here defined as an irreversible process during which the density or tracer concentration of a given fluid parcel is modified. In the equations (2.1*c*) and (2.1*d*), it is allowed by the inclusion of an explicit diffusion term. Following other papers on this topic, the term diabatic is used to denote an irreversible (and correspondingly adiabatic to denote a reversible) process applying to density. Stirring is the process of geometric deformation of fluid elements that leads to diffusive mixing, but before diffusion acts. It is in principle reversible, because it does not change the density or tracer concentration within fluid elements. To characterise the diapycnal mixing of tracers during a dynamical event, scatter plots which show the characteristics of fluid parcels in terms of their location in a 2-D space in which the coordinates are density (ρ) and tracer concentration (ϕ) are employed. In such plots, a fluid parcel will remain at the same position in density–tracer ($\rho\phi$) space whatever its displacement or geometrical deformation in physical space, provided that there is no irreversible mixing due to the action of diffusion. Any change in the distribution of fluid parcels in $\rho\phi$ space is therefore indicative that irreversible mixing has occurred. Similar diagrams are often used to characterise mixing in geophysical flows, such as temperature–salinity diagrams used to qualify the mixing of large water masses in the ocean (e.g. Tomczak 1981; Teramoto 1993); tracer–salinity estuarine mixing curves used to identify whether an estuary may act as a source or sink of a given tracer (e.g. Loder & Reichard 1981; Officer & Lynch 1981); or tracer–tracer diagrams used to examine compact relationships between different atmospheric tracers (e.g. Tilmes *et al.* 2006; Plumb 2007). Scatter plots are thus convenient for the purpose of the analysis presented here as they permit the straightforward identification of diapycnal tracer transport which must be indicated by the generation of new points in density–tracer space. For example, if the density–tracer distribution is initially given by the solid black line in figure 5, then the evolution to a distribution given by the dashed line is indicative of diapycnal tracer flux.

   The geometry of density–tracer scatter plots does not by itself uniquely define density and tracer distributions in physical space. Further information is provided by associating

each point in density–tracer space with a 'weight', which is a discrete formulation of the density–tracer probability function presented in appendix D of Plumb (2007), which quantifies the amount of fluid in a given density–tracer bin. The procedure to calculate the weight is outlined as follows. The density and tracer domains are subdivided into N_ρ and N_ϕ individual bins, respectively, with sizes

$$\delta\rho = \frac{\rho_{Max} - \rho_{Min}}{N_\rho}, \tag{4.1}$$

$$\delta\phi = \frac{\phi_{Max} - \phi_{Min}}{N_\phi}, \tag{4.2}$$

where the subscripts *Max* and *Min* indicate the maximum and minimum values of the density and the tracer. The centre of a given bin is defined by the point (ρ_i, ϕ_j), where

$$\rho_i = \rho_{Min} + \frac{2i - 1}{2}\delta\rho, \quad i = 1, 2, \ldots, N_\rho, \tag{4.3}$$

$$\phi_j = \phi_{Min} + \frac{2j - 1}{2}\delta\phi, \quad j = 1, 2, \ldots, N_\phi. \tag{4.4}$$

The weight corresponding to a given bin with centre (ρ_i, ϕ_j), $W_{ij}(t)$, is calculated as

$$\begin{aligned} W_{ij}(t) &= W\left(\rho_i, \phi_j, t\right) \\ &= \frac{1}{V}\sum I_{ij}(\rho, \phi, t)\Delta V, \end{aligned} \tag{4.5}$$

where $\Delta V = \Delta x \times \Delta y \times \Delta z$, and

$$I_{ij}(\rho, \phi, t) = \begin{cases} 1, & \left(\rho\left(\boldsymbol{x}, t\right) - \rho_i, \phi\left(\boldsymbol{x}, t\right) - \phi_j\right) \in \left[-\frac{1}{2}\delta\rho, \frac{1}{2}\delta\rho\right) \times \left[-\frac{1}{2}\delta\phi, \frac{1}{2}\delta\phi\right) \\ 0, & \text{otherwise,} \end{cases} \tag{4.6}$$

and its value is represented in colour on the scatter plot. Note that the total weight is conserved (i.e. the total is always 1), and the integral of the weight multiplied by the tracer concentration or density is also conserved in the absence of sources and sinks.

Since graphical limitations can make it difficult to discern when a scatter plot has converged to a compact relationship, it is useful to define a diagnostic that acts as a measure of the scatter,

$$R(t) = \frac{\int \left(\phi(\boldsymbol{x}, t) - \phi_*(z, t)\right)^2 \mathrm{d}V}{\int \left(\phi(\boldsymbol{x}, t) - \overline{\phi}\right)^2 \mathrm{d}V}. \tag{4.7}$$

This value will be referred to as the scatter variance. It relates the total variance of the tracer from the isopycnal mean ϕ_* (as defined by (C 11) in appendix C) to the total variance of tracer from the mean over the whole domain $\overline{\phi}$. As such, it presents a time evolution of the scatter plots, with larger values of R indicating greater relative variance over a given density bin, and smaller values indicating a convergence towards functional, compact relationships between ρ and ϕ.

4.2. *Scatter plot evolution for the different dynamical configurations*

Figure 6 provides the evolution of the weighted density–tracer scatter plots for each of the dynamical configurations, with each of the rows corresponding to those in figures 3 and 4. The weight of each density–tracer bin is represented by the colours indicated by the colour bar, with white indicating no fluid occupies that region of density–tracer space. Prior to $t = 25$, all scatter plots remain close to their initial shape, reflecting that the evolution is mostly adiabatic and advective. Most of the fluid remains located at the edges of the scatter plots (near $\rho = \pm 1$), which is indicative of the near-constant high- and low-density layers below and above the pycnocline, while the rest of the curve indicates the pycnocline itself, which initially occupies a relatively small region of the domain. Starting around $t = 25$ (figure 6*a,g,m*), the scatter plots begin to spread just below the top of the initial curve, corresponding to the initial roll-up of the pycnocline and slow diffusion of ρ and ϕ near the pycnocline. The greater spreading of the scatter plots at $t = 56$ relates to the irreversible mixing of both density and passive tracer which starts when the roll-up of the billow has significantly stretched the interface between the regions of high and low density. While the scatter plots of the 2-D configurations appear identical, the slight deviation in shape of the 3-D configuration scatter plot is due to the effect of the spanwise instabilities. The development of these secondary instabilities is when the evolutions of the 2-D and 3-D configurations diverge. The spanwise instabilities lead to the rapid breakdown of the 3-D billows, and thus rapid homogenisation of the density and passive tracer. By $t = 140$, the scatter plot reduces to an almost piecewise linear tent-like shape with minor spreading along the edges, and slightly more spreading at the peak around $\rho = 0$ (figure 6*o*). As the density and tracer further homogenise, this tent-like scatter plot becomes more compact, with the top around $\rho = 0$ becoming more rounded, and branches on either side remaining nearly linear, as shown in figure 6(q). The convergence of the scatter plots towards more compact curves indicates that the range of tracer concentrations for a given density value has decreased. The 2-D billows continue to rotate uninhibited, which mixes and homogenises the density within the billow, while trapping local maxima of the passive tracer within the core of the billows. These tracer maxima localised in regions of relatively constant density are visible in the scatter plot as narrow vertical protrusions centred at $\rho = 0$, as in figure 6(c,i). As the density field homogenises, these vertical protrusions converge to more compact vertical lines, as in figure 6(d,j), indicating a wide range of tracer values in a region of nearly constant density. However, while configuration I equilibrates and maintains this shape, configuration II experiences new instabilities and undergoes a new turbulent phase associated with billow pairing. This generates large meanders protruding deep into the upper and lower layers of the fluid, which involves mixing over a wide density range, indicated by the increased scattering in figure 6(k). A new single billow is formed that eventually stabilises. The scatter plot evolves to a new shape where the central branch has vanished.

The scatter variance is plotted with different components of the perturbation KE in figure 7, in order to compare the scatter to the dynamical evolution of the instabilities. Note that until $t \approx 50$, the evolution of the 2-D KE is essentially identical for all three configurations, while it remains the same for the 2-D configurations until $t \approx 200$. Each configuration shows an increase in scatter variance that occurs just after the initial increase in 2-D KE, with both quantities depicting similar rates of increase. The first peak in the plots of scatter variance correspond to the period during which secondary instabilities have started to develop. For configuration I, the scatter variance undergoes a rapid decrease as the scatter plot begins to converge to its three branch shape, with a slight increase around $t = 150$. After this, the decrease in scatter variance is quite slow, as is the decrease in 2-D
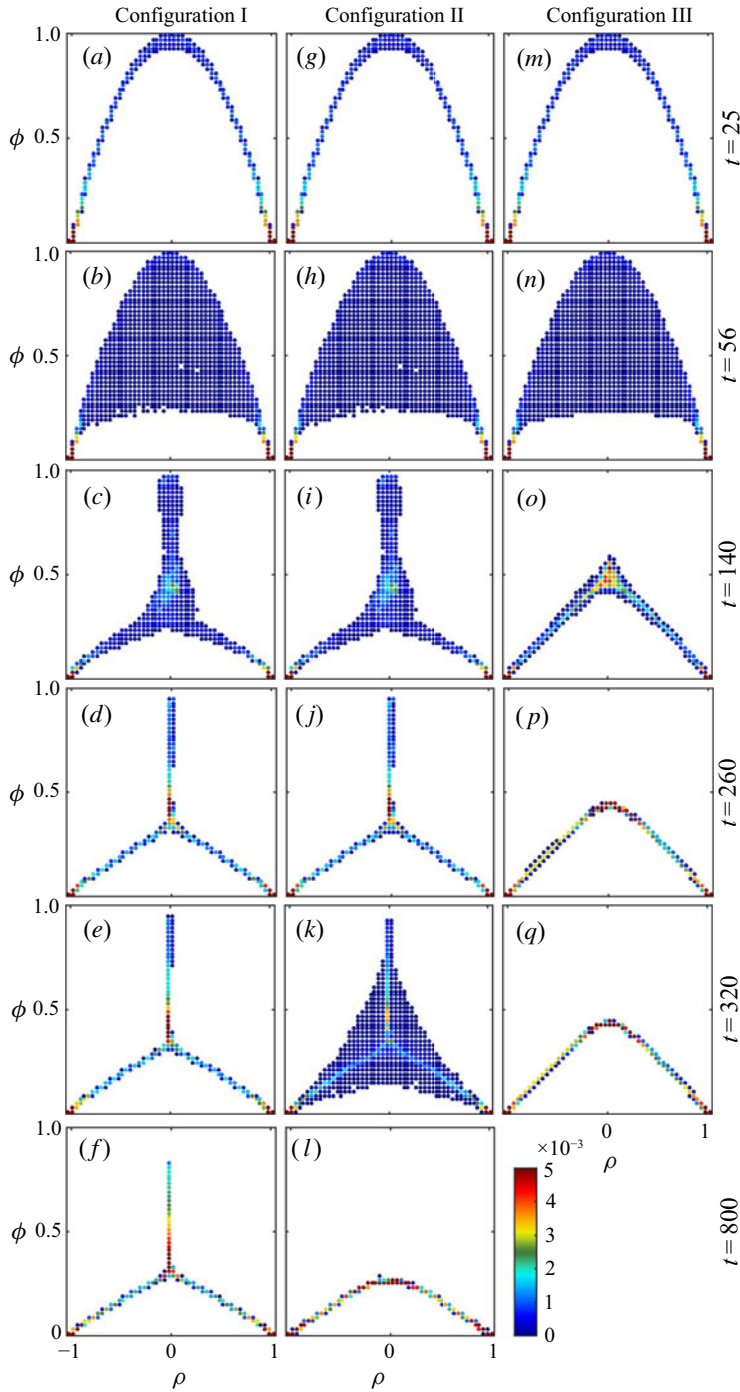
FIGURE 6. Weighted density–tracer scatter plots for configurations I (*a–f*), II (*g–l*) and III (*m–q*). Blue, yellow and red indicate low, intermediate and high values of the weight, respectively.
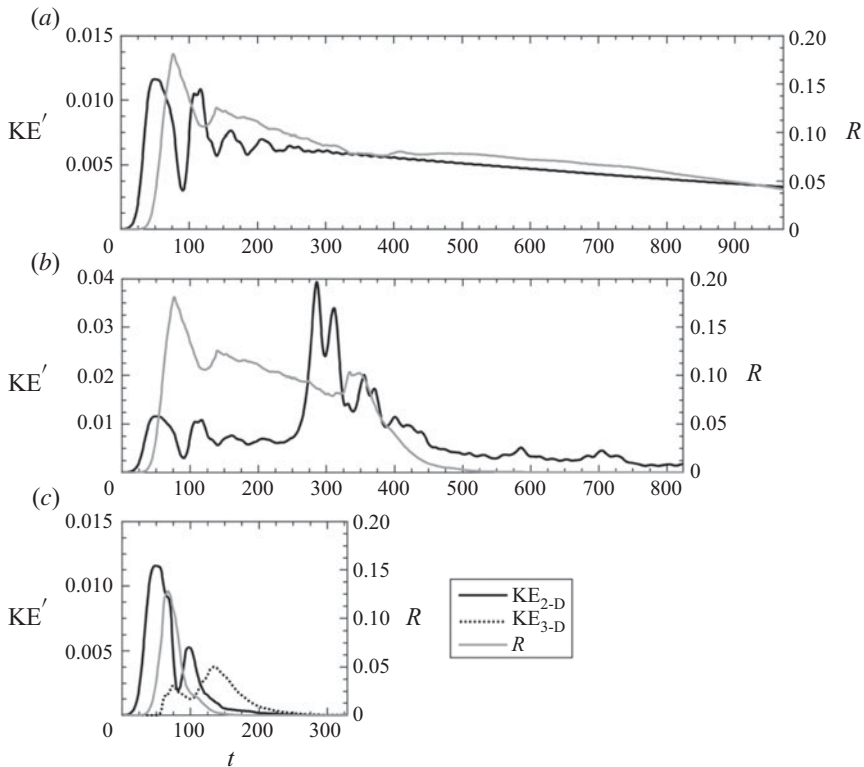
FIGURE 7. Perturbation KE (black lines), and scatter variance (grey lines) for (*a*) configuration I, (*b*) configuration II and (*c*) configuration III. Configurations I and II depict only the 2-D kinetic energy, while configuration III depicts both the 2-D KE (solid line) and 3-D KE (dotted line). The values for the kinetic energy are depicted on the left-hand axes, while the values for the scatter variance are depicted on the right-hand axes.

KE, as the flow has reached a stable state. Configuration II observes similar behaviour in the evolution of its scatter variance, but sees a small rapid increase after the onset of pairing (depicted by the rapid increase in 2-D KE around $t = 300$), corresponding to the spreading of the scatter plot into away from its three branch structure to the filled triangle structure, as depicted in figure 6(*j,k*), respectively. The scatter variance then undergoes a relatively rapid decrease to near zero as the scatter plot reaches its stable compact shape. The scatter variance of configuration III experiences the same rapid increase as the other cases after the initial billow roll-up, but rapidly decreases after the organised 3-D secondary instabilities develop (the 2-D KE sees a rapid decrease during the development of the spanwise instabilities, which appears to relate to a brief breakdown of the billow). The rate of decrease of the scatter variance slows from around $t = 90$ to 130, following a rapid increase in the 2-D KE (related to a brief reformation of the coherent billow which occurs prior to the complete breakdown of the primary instability and further development of turbulence). It quickly reaches zero around $t = 150$ as the scatter plot begins to reach its compact tent-like shape.

4.3. *Convergence principle for scatter plots*

Because density and the passive tracer are both governed by advection–diffusion equations without sources or sinks, there exists an important constraint on the evolution of the
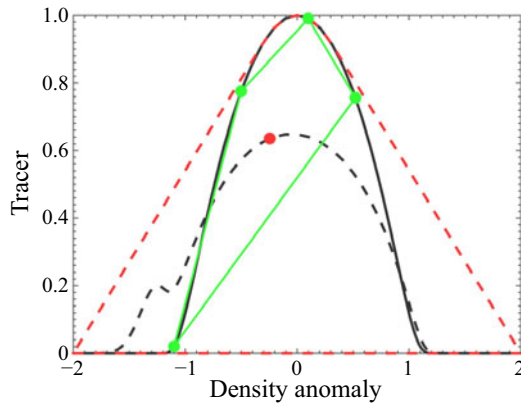
FIGURE 8. Diagram illustrating the convex envelope constraint on the evolution of a typical density–tracer scatter plot as described in §4.3.

density–tracer scatter plot (see Plumb (2007) and references therein. Lauritzen & Thuburn (2012) used this constraint to determine if the mixing in numerical models is physical). In a general sense, scatter plot evolution can be understood as follows. As previously mentioned, stirring will not modify the scatter plot, but bring fluid parcels from different regions closer together (for example, the four parcels represented by the green points in figure 8). Provided the molecular diffusivities of the density and tracer are equal, mixing will homogenise the density–tracer characteristics of parcels within a cell whose size is determined by the diffusivity coefficient. The resulting density–tracer characteristics of this cell are then the averaged values of the initial parcels, weighted by their volumetric ratio (e.g. the red point in figure 8). This implies that the scatter plot after stirring and mixing will be contained within the convex envelope of the initial distribution (region within the red dashed curve in figure 8). Certain properties can be inferred from this constraint:

(*a*) Whatever the route to turbulence and mixing, the tracer concentration in a given density range remains within the interval determined by the initial convex envelope. This limits the extrema of the tracer concentration in a given density range to the values set by the initial convex envelope.

(*b*) During mixing, the scatter plot evolves continuously. Thus, at any given time, a scatter plot must lie within the convex envelope of every preceding scatter plot. In addition, extreme values of tracer or density are eroded by mixing. As a result, the convex envelope reduces with time and may eventually converge to a more compact scatter plot.

(*c*) The convex envelope of a straight line is simply the same straight line. If fluid parcels along a given straight line mix, they will remain confined to that line.

Note that if the density and tracer have different diffusivities then these convex envelope constraints can be broken, although whether or not this effect is important in practice remains to be determined.

The scatter plot evolution depicted in figure 6 follows each of the constraints listed above. The large-scale dynamical mixing due to the KH billows (both the initial billows and pairing) reduces the overall size of the convex envelopes and thus the scatter plots, and the maximum possible value of the tracer concentration in specific density ranges.

After a certain amount of time, the maximum tracer value slowly decreases, showing that mixing at this point is no longer dynamically active between the layers, but acts at smaller, localised scales between adjacent points in density–tracer space. This localised mixing appears responsible for the formation of the nearly linear regions of the converged scatter plots on either side of $\rho = 0$ because the previous arguments can be applied to restricted portions of the scatter plots provided that mixing acts locally in $\rho\phi$ space or at reduced scales.

5. Background density and isopycnal mean tracer profile evolution

This section presents the second approach used to describe and quantify the evolution of the tracer field and contrast it with the evolution of the density. It is based on the evolution of mean density and tracer profiles obtained by an adiabatic rearrangement of the density field.

5.1. *Background density and effective diffusivity*

The traditional approach to representing transport and mixing of tracers in turbulent flows is via a diffusive formulation, i.e. a turbulent diffusivity is sought that represents the effect of the turbulent flow on the tracer. One of the fundamental limitations of this approach is that the diffusive representation of random walks, and by extension, of the effect on tracers of quasi-random flows, describes evolution that is the result of many small random steps. This condition cannot be justified for flows that are spatially inhomogeneous, such as the KH instability considered here, where the typical fluid particle displacement by a single eddy is comparable to the length scale on which the properties of the flow change. One way of overcoming this is to a move to a tracer-based coordinate system. This is the approach in the effective-diffusivity formalism devised by Nakamura (1996) and Winters & D'Asaro (1996). This formalism applies to systems where tracers are advected and diffused. Contours (in two dimensions) or surfaces (in three dimensions) of tracer concentration are used to define coordinate surfaces or contours, but the latter are labelled not by the value of the tracer concentration but the area (in two dimensions) or volume (in three dimensions) enclosed by the surface. If the tracer has some kind of geometric organisation, then the coordinate system and the variation of tracer concentration within that system represent that organisation. For the KH instability and for other flows in density-stratified fluids, the natural tracer is the density and the corresponding tracer-based coordinate, z_*, represents a vertical coordinate. By construction, the density is a function of one space variable, z_*, and time t alone, written as $\rho_*(z_*, t)$, which may be shown to satisfy the diffusion equation

$$\partial\rho_*/\partial t = \partial/\partial z_* \left(K_\rho \partial\rho_*/\partial z_* \right), \tag{5.1}$$

where K_ρ is the dimensional effective diapycnal diffusivity of density, defined as

$$\frac{K_\rho(z_*, t)}{\kappa_\rho} = \left\langle |\nabla\rho|^2 \right\rangle_{z_*} (\partial\rho_*/\partial z_*)^{-2}, \tag{5.2}$$

where κ_ρ is the molecular diffusivity of density, and $\langle \cdot \rangle_{z_*}$ denotes an appropriately defined average over a z_* surface (as defined by (C 11) in appendix C). Note that the rearranged (or background) density field $\rho_*(z_*, t)$ may be calculated from the 3-D simulation by rearrangement to construct a monotonic profile. That is, fluid elements, each with a specified infinitesimal volume, are ordered by their density, giving density as a function

of cumulative volume (i.e. the volume of fluid with density less than a given value), and then the volume is converted to a vertical coordinate z_* by dividing by the horizontal area of the fluid domain. The coordinate z_* is then a decreasing function of ρ_*. The value of z_* for a given ρ_* is therefore proportional to the integral, down to ρ_* of the weight defined previously in § 4.1. (Note that, typically, (5.1) is written in terms of ρ and z_* (for example, Smyth *et al.* 2005), with $\frac{\partial \rho}{\partial z_*}$ implicitly referring to the adiabatic rearrangement of the density field. For clarity, ρ_* will be used here for density when it is being regarded as a function of z_*, and ρ will be used when it is being regarded as a function of the Cartesian coordinates (x, y, z), such as in the numerator of the right-hand side of (5.2).)

The left-hand column of figure 9 shows the time evolution of the background density profiles sorted from the simulations of the different dynamical configurations (solid lines), and the profiles calculated from (5.1), using the effective diffusivity calculated from those three simulations. The right-hand column plots the final profiles based on the rearrangement of the simulation results and the diffusion equation. The area around the pycnocline is magnified to show the area of interest in better detail. As suggested by the evolution of the cross-sections in figure 3, the profiles of all three simulations undergo similar evolution until the roll-up of the primary KH billows, at which point the pycnocline of configuration III undergoes greater spreading than the 2-D configurations due to the mixing from secondary instabilities. The profiles for the 2-D configurations continue to evolve identically until the onset of pairing, at which point the pycnocline of the configuration II profile widens significantly. After mixing, for both 2-D configurations, the edges of the pycnocline widen slightly until the end of the simulation. The final state of configuration I is an approximately three-layer profile, with a centre layer near $\rho_* = 0$, and rapid changes to the upper and lower layers. Configuration II has a near constant $\rho_* = 0$ layer of similar width to configuration I, with less steep changes in profile towards the upper and lower layers. The final profile of configuration III differs by not having a constant density middle layer, instead showing the density continuously decrease with height.

Equation (5.1) describes the transport which occurs solely through molecular diffusion of ρ, and hence ρ_*, across z_* surfaces (there is no advective component). As discussed by Nakamura (1996) and Winters & D'Asaro (1996) (and in subsequent papers that exploit this formalism, such as Smyth *et al.* (2005), Salehipour & Peltier (2015), Zhou *et al.* (2017)), the effective diffusivity K_ρ is determined by the geometry of the full 3-D density field. The dimensionless factor $\langle |\nabla \rho|^2 \rangle_{z_*} (\partial \rho_* / \partial z_*)^{-2}$ has minimum value 1 when the ρ surfaces are planes, and increases as the ρ surfaces become more complex. Thus whilst there is no advective transport required in (5.1), the indirect effect of advection is to deform the surfaces of ρ and hence to increase the effective diffusivity.

The approach first followed here is to investigate whether (5.1) provides a quantitatively useful expression of the effect of transport and mixing on density and whether it can be extended to other tracers. Whilst (5.1) follows exactly from the partial differential equation for advection–diffusion, with a specified value of molecular diffusivity κ_ρ without any approximation, the KH instability simulations rely on a numerical implementation of this partial differential equation, and it is first important to establish whether or not this implementation (5.1) remains quantitatively accurate. It is straightforward to solve the one-dimensional diffusion equation (5.1) numerically by providing the history of $K_\rho(z_*, t)$ calculated using the density field from the numerical simulations, and to compare this predicted evolution of $\rho_*(z_*, t)$ with that given by the numerical simulation itself. In appendix B it is shown that applying this approach using the constant value of κ_ρ specified for the numerical simulation under-predicts the mixing of density in the evolution of the KH instability. The explanation is that the numerical approximations
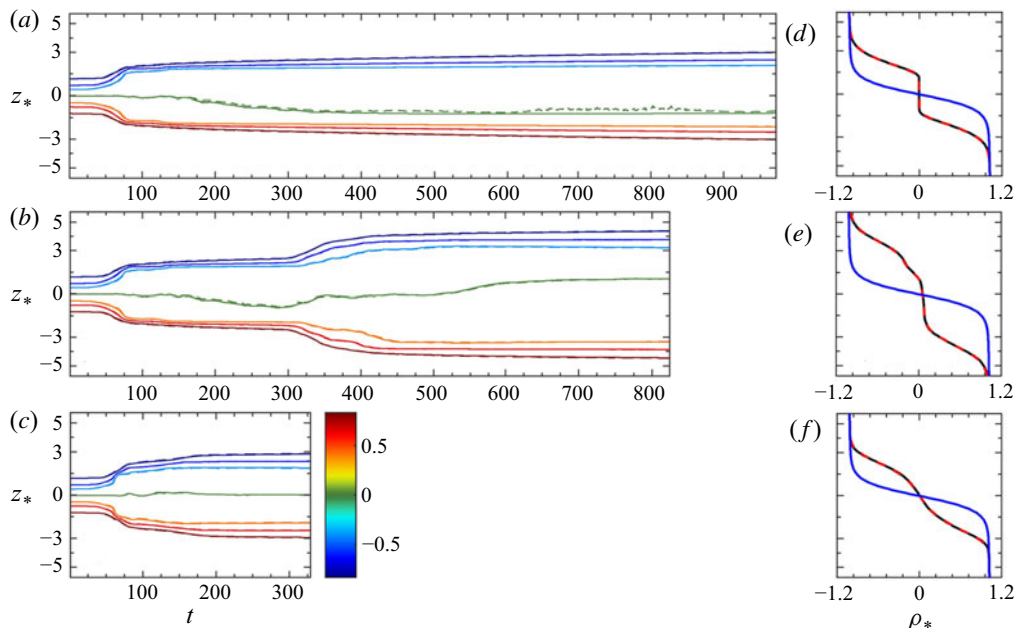
FIGURE 9. Time evolution of the background density profile from the simulations (solid lines) and as calculated from the diffusion equation (5.1) (dashed lines) for (*a*) configuration I, (*b*) configuration II and (*c*) configuration III. The corresponding final density profiles are presented in the right-hand column (*d–f*), with the background profile given by the solid black line, and the profile from the diffusion equation given by the solid red line. The initial profile is depicted by the blue line. The solid black and dashed red lines overlap almost perfectly.

to each of the derivative terms of the governing equations, which have been designed with certain properties (for example, controlling spurious oscillations near discontinuities Shu (1999)), in effect augment the specified molecular diffusivity. An estimate of the extra diffusivity provided by the numerical schemes is provided by considering the ratio between the time rate of change of the density variance and the variance of the density gradient (see (B 3)). It may be concluded that by this measure, in the simulation depicted (configuration III), the additional time-varying 'numerical' diffusivity reaches up to 100 % of the specified molecular diffusivity during the initial billow roll-up. The numerical diffusivity deduced on this basis can be added to κ_ρ to give a time-dependent 'net diffusivity' $\kappa_\rho^{Net}(t)$, which helps better represent the behaviour of the higher-order numerical terms in the simulation by a simple advection–diffusion equation. Fortunately, the derivation of the effective diffusivity formulation in appendix C permits the diffusivity provided in the advection–diffusion equation to vary with time without modifying the end result. By adjusting a single quantity, κ_ρ, improved agreement is seen in the predicted ρ_* as a function of z_*. Figure 9 indicates that both the computed and rearranged profiles are in very good agreement for the duration of the simulations, especially towards the edges of the pycnocline. The greatest disparity appears for the $\rho_* = 0$ isopycnal (green lines) in the configuration I (figure 9*a*), but remains modest. Note that tracer profiles presented in subsequent sections will also be calculated using net tracer diffusivities.

For reference, the corresponding time evolution of the dimensionless K_ρ (i.e. the right-hand side of (5.2)) for configuration III is presented in figure 10. As has been
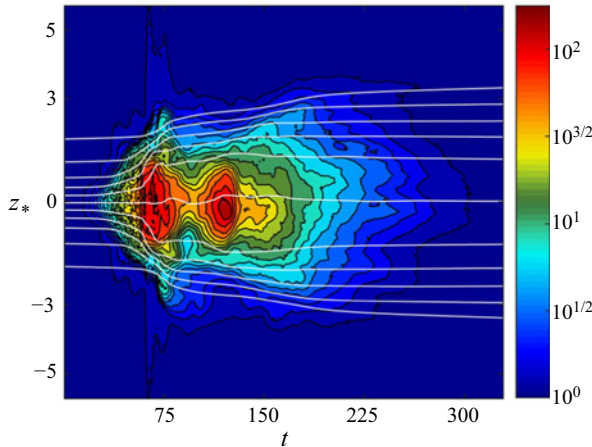
FIGURE 10. Example of dimensionless effective diffusivity (i.e. the right-hand side of (5.2)) computed from the sorted density profile for configuration III. The colour indicates the dimensionless values of effective diffusivity, while the white lines indicate different values of the background density profile.

demonstrated in previous papers, and in applications to different flows (e.g. Nakamura 1996; Winters & D'Asaro 1996; Shuckburgh & Haynes 2003), the distortion of ρ-surfaces leads to a substantial enhancement of K_ρ relative to κ_ρ, and in the case shown, a factor of several hundred at certain stages of the flow evolution. Note also that the difference between $\kappa_\rho^{\mathrm{Net}}(t)$ and κ_ρ is not very significant in this enhancement but, again, that it is significant in giving precise quantitative agreement between the evolution of $\rho_*(z_*, t)$ predicted by (5.1) and the evolution according to the full numerical simulation.

5.2. *Effective tracer diffusivity*

Once ρ_* is determined, a tracer in the flow can be sorted relative to this density profile by calculating tracer means over isopycnal surfaces,

$$\phi_* (z_*, t) = \langle \phi (\mathbf{x}, t) \rangle_{z_*}, \tag{5.3}$$

as defined by (C 11). Following the process described in appendix C, a complete governing equation for this isopycnal mean tracer profile can be derived. A preliminary step during the derivation provides a formulation that holds for any ϕ and ρ,

$$(\langle \phi \rangle_{z_*})_t = \langle \phi_t \rangle_{z_*} + \frac{\partial}{\partial z_*} \left(\left(\langle \rho_t \rangle_{z_*} \langle \phi \rangle_{z_*} - \langle \phi \rho_t \rangle_{z_*} \right) \left(\frac{\partial \langle \rho \rangle_{z_*}}{\partial z_*} \right)^{-1} \right), \tag{5.4}$$

provided the time derivatives ϕ_t and ρ_t are supplied via specific evolution equations. The final step in the derivation supplies these time derivatives via the advection–diffusion equations (2.1c) and (2.1d) to give

$$\partial \phi_* / \partial t = \partial / \partial z_* \left(\frac{\kappa_\phi}{\kappa_\rho} K_\rho \partial \phi_* / \partial z_* \right) + \partial / \partial z_* \left((\partial \rho_* / \partial z_*)^{-1} \left\langle \kappa_\phi \nabla \phi' \cdot \nabla \rho - \kappa_\rho \phi' \nabla^2 \rho \right\rangle_{z_*} \right), \tag{5.5}$$

where ϕ and ρ have diffusivities κ_ϕ and κ_ρ, respectively. Here, $\phi' = \phi - \phi_*$ represents the perturbation from the mean value of tracer for a given density. The density at each

point corresponds to a specific z_*, and thus $\phi_*(z_*, t)$, which is subtracted from the value of the tracer concentration to give ϕ'. An intermediate formulation could be derived by supplying ρ_t via an advection–diffusion equation, but allowing for a more general form of the equation for ϕ. This equation, and the special limit cases in which either one or both tracers are non-diffusive, are discussed in greater detail in appendix C.3.

Focusing on (5.5), the first term on the right-hand side is straightforward – it simply captures the molecular diffusion of ϕ_* taking account of the geometry of the ρ-surfaces. The second term on the right-hand side takes account of the variation of ϕ over ρ-surfaces, represented by the quantity ϕ'. Estimating this 'eddy term' therefore presents a closure problem, since information about ϕ' is not available from ϕ_*. There are two distinct contributions to this second term, both requiring non-zero diffusivity. The first, proportional to $\kappa_\phi \nabla \phi' \cdot \nabla \rho$ is associated with a diffusive flux of ϕ across ρ-surfaces. The second, proportional to $-\kappa_\rho \phi' \nabla^2 \rho$, is associated with the motion of ρ-surfaces relative to the fluid. Rather than considering the closure problem further, the importance of the tracer eddy term will be investigated by considering a 'virtual tracer' $\phi_v(z_*, t)$ which satisfies (5.5) with the tracer eddy term neglected,

$$\partial \phi_v / \partial t = \partial / \partial z_* \left(\frac{\kappa_\phi}{\kappa_\rho} K_\rho \partial \phi_v / \partial z_* \right), \tag{5.6}$$

where initially $\phi_v(z_*, t = 0) = \phi_*(z_*, t = 0)$. Figure 11 shows the time evolution of ϕ_* (solid lines) as calculated from the different dynamical configurations using (5.3), and ϕ_v as calculated from (5.6) (dashed lines), for the tracer layer collocated with the pycnocline. Profiles at specific times corresponding to the cross-sections of figure 4 are presented in figure 12. During the evolution, there exist strong differences between both profiles for all three configurations, starting around $t = 50$, immediately after the first mixing event. However, with time, both configurations II and III show an increase in agreement. By the end of these simulations (figure 11e,f), there are relatively minor differences towards the centre of the domain and along the edges of the profiles, where the virtual tracer slightly underestimates and overestimates the tracer concentration, respectively. In contrast, configuration I continues to show marked disagreement between the isopycnal mean profiles from simulation and the virtual profiles, with ϕ_v being wider than ϕ_*, and significantly underestimating the amount of tracer at the centre of the domain. This does indicate that, provided there are adequate mixing events, ϕ_v and ϕ_* can eventually reach good agreement, suggesting that the mean tracer profile resulting from mixing can be predicted using a simple diffusive equation with diffusivity given by the density effective diffusivity. This is not a trivial result as it requires that, during the evolution, the cumulative effect of the eddy term vanishes. The fact that the eddy term vanishes at the end of the simulation may not be enough.

5.3. *Relative contribution of the eddy term*

In this section, the eddy term is examined in greater detail to assess its relative importance in tracer profile evolution. First note that (5.5) can rewritten in a way that omits the calculation of ϕ' or its gradient, facilitating the numerical computation of the eddy term,

$$\partial \phi_* / \partial t = \partial / \partial z_* \left(\frac{\kappa_\phi}{\kappa_\rho} K_\rho \partial \phi_* / \partial z_* \right) + \partial / \partial z_* \Big((\partial \rho_* / \partial z_*)^{-1} \big(\kappa_\phi \langle \nabla \phi \cdot \nabla \rho \rangle_{z_*}$$
$$- \kappa_\phi \partial \phi_* / \partial \rho_* \langle \nabla \rho \cdot \nabla \rho \rangle_{z_*} - \kappa_\rho \langle \phi \nabla^2 \rho \rangle_{z_*} + \kappa_\rho \phi_* \langle \nabla^2 \rho \rangle_{z_*} \big) \Big). \tag{5.7}$$
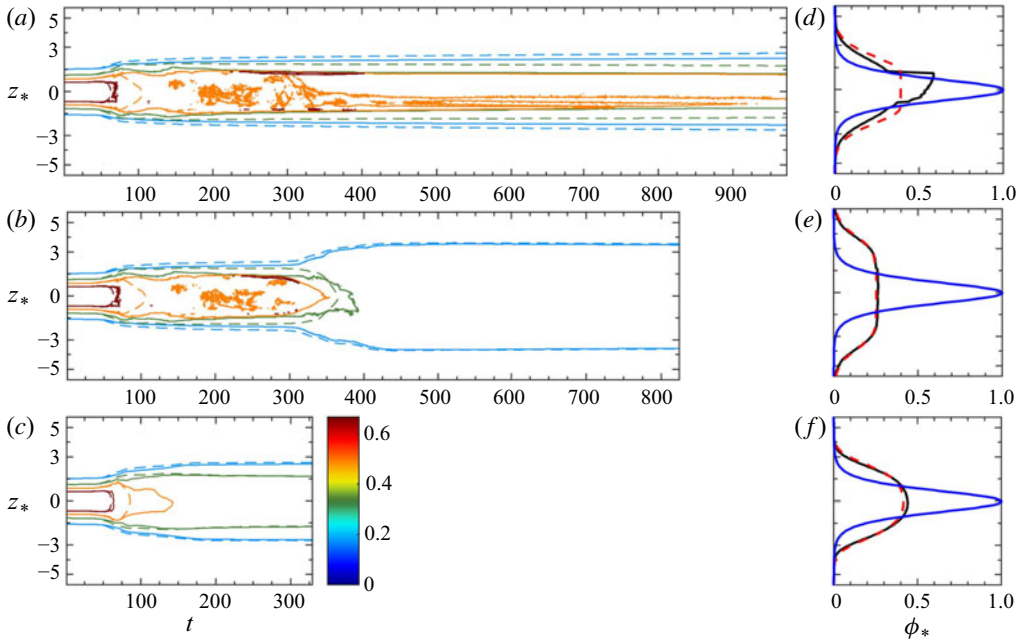
FIGURE 11. Time evolution of the virtual tracer profiles (dashed lines) and the isopycnal mean tracer profiles from simulations (solid lines) for (*a*) configuration I, (*b*) configuration II and (*c*) configuration III. The corresponding final tracer profiles are presented in the right-hand column (*d–f*), with the isopycnal mean profile given by the solid black line, and the virtual profile given by the solid red line. The initial profile is given by the blue line.

For the simulations presented here, it has been verified that calculating the evolution of ϕ_* using (5.7) with an eddy term computed directly from simulation significantly improves the agreement with the isopycnal mean tracer profile computed from the simulation at all times. The profiles obtained from (5.7) closely match the mean profiles from the 2-D or 3-D fields, even for configuration I, where the isopycnal mean tracer profile from simulation does not converge to a shape similar to the purely diffusive virtual profile. Figure 13 presents the time evolution of the profiles of the diffusive term, eddy term, and time derivative of the isopycnal mean tracer, for the 3-D configuration III. Following the initial roll-up of the billow and the during the development of secondary instabilities (from about $t = 50$ to 130), both terms show similar structure of the same magnitude, but of opposite sign. This shows that the eddy term is non-negligible, and provides closure for the accurate computation of ϕ_*. Over this time period, both tend to mostly cancel each other, with the resulting time derivative approximately an order of magnitude smaller than either term. This time derivative shows that ϕ_* undergoes the most rapid change as the secondary instabilities and 3-D structure develop (around $t = 60$ to 80), although there are still slow changes in ϕ_* after $t = 130$. The slow changes indicated by the blue patches about $z_* = 0$ between $t = 130$ and 150 are due to contributions from both terms, while the streaks of red in figure 13(*c*) between $t = 130$ and 200 around $z_* = \pm 2.5$ are due entirely to the diffusive term.

6. Preliminary interpretations of results

For configurations II and III, the reasoning behind the relatively good agreement between the stable state profiles of ϕ_v and ϕ_* is not entirely clear. This section presents
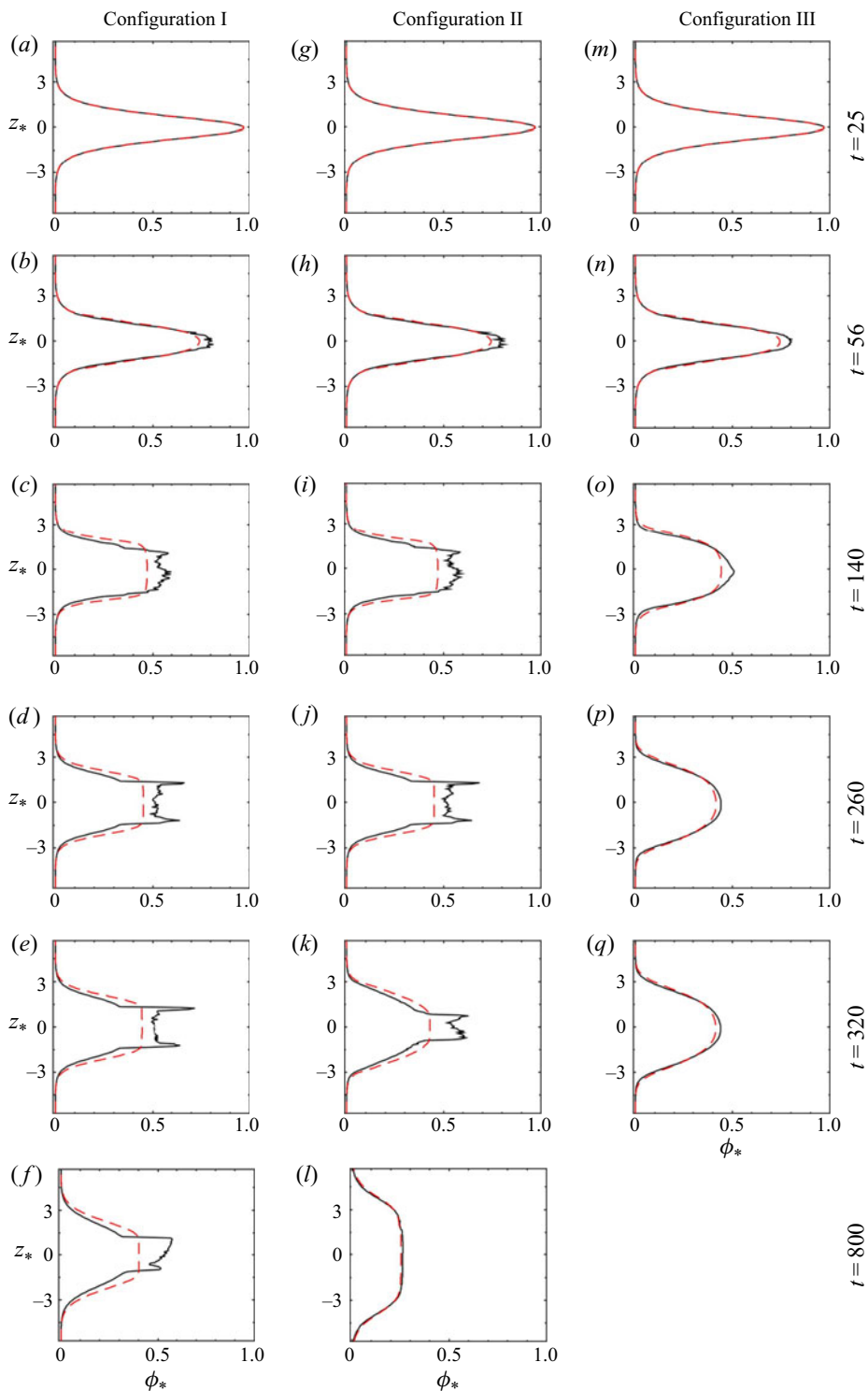
FIGURE 12. Isopycnal mean tracer profiles computed from simulations (solid black lines) and virtual tracer profiles calculated from (5.6) (dashed red lines) for the three dynamical configurations (configuration I (*a–f*), configuration II (*g–l*) and configuration III (*m–q*)).
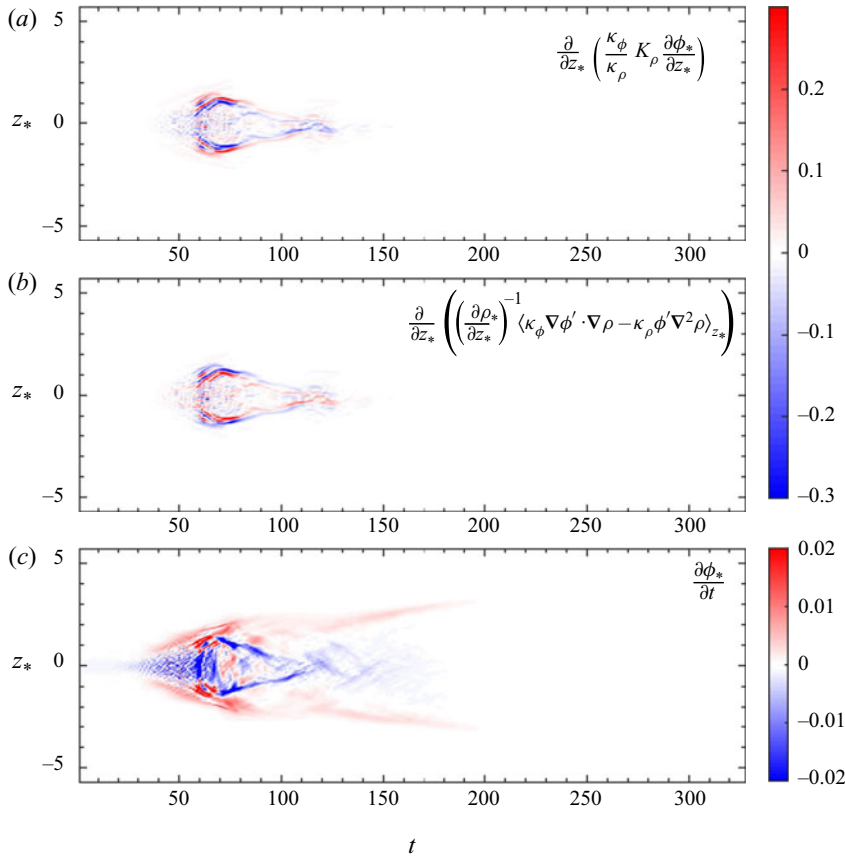
FIGURE 13. Time evolution of the profile of (*a*) the diffusive term, (*b*) the eddy term and (*c*) the time derivative of the isopycnal mean tracer for the 3-D dynamical configuration (III).

possible behaviours and features that may be indicative of a tendency towards good agreement between ϕ_v and ϕ_*, as well as situations that may yield discrepancies between the two.

In § 4.2, it is observed that the scatter plots for the fully mixed configurations II and III (figure 6*l*,*q*) tend to compact forms in which there is little to no variation in tracer concentration for a given value of density. Note that there is a strong relationship between the eddy term of (5.5) and the scatter plots, in that perfect compactness ensures that $\phi' = 0$, and thus the eddy term is null. The behaviour depicted in figure 13 can thus be interpreted as a tendency for the eddy term to generate small-scale perturbations on the mean profile ϕ_* and for diffusive term to compensate this effect, as long as the eddy term remains active. After the strong mixing phase, the eddy term diminishes, reflecting the convergence to a compact relationship. The scatter plots of both configurations tend to shapes with outer regions in which the density–tracer relationship appears linear (as suggested in § 4.3), and a central region where the relation lies on a curve. The highest values of tracer concentration are located in the curved central region, with tracer values decreasing along the piecewise linear regions towards the edges. Consideration of the weights in density–tracer space shows that each of these three regions correspond to a non-negligible volume of the fluid.
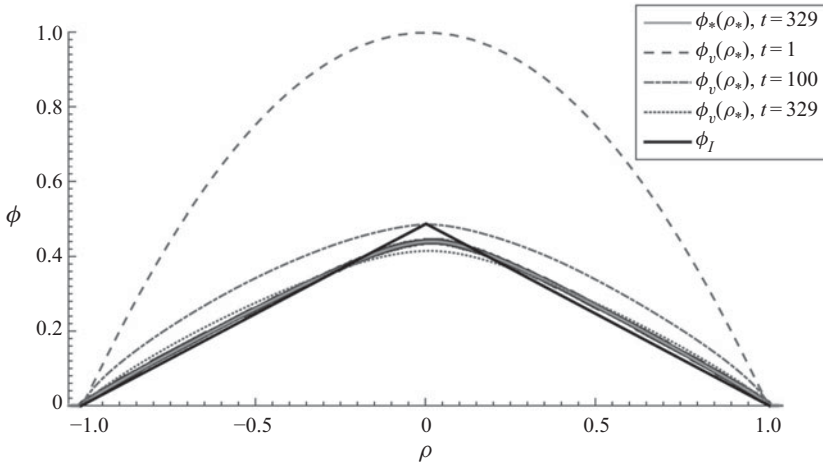
FIGURE 14. Evolution of the virtual tracer profiles (dashed lines) as a function of the background density for the default tracer field of configuration III. The blue and yellow points (inside and outside $\rho = -1$ and 1, respectively) indicate the density–tracer scatter plot at $t = 329$, with the solid grey line indicating the isopycnal mean tracer profile as a function of background density at $t = 329$. The solid black line indicates the ideal tent shape given by (6.3) at $t = 329$.

The density–tracer scatter plot is now compared to the virtual tracer plotted as a function of density, $\phi_v(\rho_*)$ for configuration III. Figure 14 plots the end time ($t = 329$) density–tracer scatter plot (indicated by primarily blue dots between $\rho = -1$ and 1), which has converged to a compact relationship that recreates $\phi_*(\rho_*)$ (solid grey line overlapping the scatter plot). This convergence is to be expected, since ϕ_* represents the mean value of the tracer over a given density surface. $\phi_v(\rho_*)$ is plotted at different times as dashed grey lines, showing that the shape of the virtual tracer curve converges to a shape similar to the compact scatter plot relationship and $\phi_*(\rho_*)$. Both figures 11(f) and 14 indicate that although $\phi_v(\rho_*)$ underestimates the amount of tracer at the central region around $\rho_* = 0$, and overestimates the amount of tracer along the two outer regions, ϕ_* and ϕ_v are in relatively good agreement.

As is the case for the scatter plots (and thus $\phi_*(\rho_*)$), $\phi_v(\rho_*)$ has tendency towards nearly piecewise linear relationships. Since ϕ_v is a one-dimensional profile, it can be rewritten as

$$\phi_v(z_*, t) = F(\rho_*, t). \tag{6.1}$$

Substituting (6.1) into (5.6), a new equation in terms of F can be written as

$$\partial F/\partial t - \frac{\partial^2 F}{\partial \rho_*^2} K_\rho \, (\partial \rho_*/\partial z)^2 = 0. \tag{6.2}$$

This is a type of diffusion equation, under which the curvature of the graph of the function F might be expected to reduce, and indeed the equation allows a possible steady state with $\partial^2 F/\partial \rho_*^2 = 0$, i.e. with F a linear function of ρ_*. This suggests that the steady state form of F might be modelled by a piecewise linear function, which can be used as a framework in which to compare the final stages of $\phi_v(\rho_*)$ and $\phi_*(\rho_*)$.

The mixing event sets the distribution of density weights within a given density range, say $[\rho_L, \rho_H]$, which will be affected by turbulent mixing over the course of the entire event.

Within this range, the total amount of passive tracer is conserved. Therefore, assuming symmetric mixing across the mid-density, a unique ideal piecewise linear tent shape can be entirely determined by the distribution of the density and the total amount of tracer within this range. It can be defined as

$$
\phi_I(\rho_*) = \begin{cases} (\rho_* - \rho_L)\left(\dfrac{2\phi_{Max}}{\rho_H - \rho_L}\right), & \rho_* \in [\rho_L, \rho_M], \\[4mm] -(\rho_* - \rho_H)\left(\dfrac{2\phi_{Max}}{\rho_H - \rho_L}\right), & \rho_* \in [\rho_M, \rho_H], \end{cases} \tag{6.3}
$$

where the maximum tracer concentration value ϕ_{Max} occurs at the mid-density defined as $\rho_M = (\rho_L + \rho_H)/2$, and is given by solving

$$
\phi_{Tot} = \int_{\rho_L}^{\rho_H} \phi_I(\rho)W(\rho)\,\mathrm{d}\rho, \tag{6.4}
$$

where ϕ_{Tot} is the total amount of tracer within the range $[\rho_L, \rho_H]$, and $W(\rho)$ is the weight function of density at the final state, which is determined from the simulation. Figure 14 depicts $\phi_I(\rho_*)$ for configuration III as solid black lines. Though there are important differences that must be considered, the general shape of $\phi_I(\rho_*)$ is close to that of $\phi_*(\rho_*)$. Primarily, the maximum of the tracer concentration is greater for the ideal tent shape. Given the conservation of total amount of tracer in the mixing region, this discrepancy requires that part of the external branches of $\phi_*(\rho_*)$ lie above the linear regions of the ideal tent shape. This indicates that $\phi_*(\rho_*)$ has a slightly convex shape in the branch regions. The piecewise linear profile (6.3), with the constraint on ϕ_{Max} implied by (6.4), is an idealisation that, as can be seen from figure 14, has some skill in predicting the actual final relation between tracer and density. However, the prediction is not exact because there is a central range of values of density where the gradient of the function $F(\rho_*)$ varies smoothly, rather than changing abruptly as would be required if it were. Thus what happens in the central region is of particular importance for determining the shape of the late time profile defined by $F(\rho_*)$ and, in particular, its deviation from a piecewise linear form. Note also that the symmetry assumed in (6.3) is not exact in the simulations discussed here and will not be in general.

The lack of convergence between $\phi_v(z)$ and $\phi_*(z)$ for configuration I can be explained by the fact that, in the 2-D simulation, the scatter plot does not tend to a compact relationship. There is a prominent vertical branch at the mid-density (in the central region) corresponding to incomplete homogenisation of tracer along that density surface. Such a feature is impossible to achieve for the virtual profile. However, configuration I seems anomalous for the dynamical configurations tested here, as configurations II and III (the most physically relevant case) demonstrate convergence towards similar vertical profiles. A possible explanation for this is that at intermediate times (for example, configuration III between $t = 50$ and 130), the forcing provided by the eddy term creates a positive difference between ϕ_* and ϕ_v near the centre, and negative differences along the sides (e.g. figure 12o). Once the scatter plot reaches a more compact shape, the forcing provided by the eddy term diminishes, allowing diffusion to redistribute ϕ_* faster than ϕ_v due to its structure. A consequence is that the effective diffusivity of density provides an excellent approximation for the tracer diffusivity in these flows.

7. Sensitivity to variations in the initial tracer field

The comparison of the three different dynamical configurations only considers a tracer that is centred along and of similar width to the pycnocline. The tracer is fully distributed within the density range affected by turbulence and mixing, and its initial gradient is similar in magnitude and geometry to the density. In order to test whether the effective diffusivity provides an adequate representation of tracer mixing in configurations where this is not the case, this section examines the mixing of tracer layers not initially collocated with or of similar width to the pycnocline in the 3-D dynamical configuration III. The background shear and density profiles are maintained as before, but the width and vertical position of the initial tracer profiles are varied by modifying parameters a and b in (2.9), respectively.

7.1. *Sensitivity to the vertical extent of the tracer*

In § 6, the ideal piecewise linear structure defined by (6.3) depends only the initial values of the tracer at $\rho = \rho_L$ and $\rho = \rho_H$ (taken here to be 0), and the integral quantity of the tracer within the density range $[\rho_L, \rho_H]$ affected by the dynamics of the flow, ϕ_{Tot}. To test how strongly the details of the initial tracer distribution influence the final profile, the width and maximum of the initial tracer profile are varied in (2.9) by modifying a, with $b = 0$ to keep the tracer collocated with the pycnocline. In the reference configuration for the tracer (i.e. the configuration presented in previous sections), $a = 1$. When $a > 1$, the width of the tracer layer is narrower than the that of the reference configuration, while the initial maximum is greater. Therefore, the tracer in these simulations initially resides entirely within the region subject to mixing. The left-hand column of figure 15 presents the evolution of the isopycnal mean tracer profiles from simulation (solid lines) and the virtual profiles (dashed lines) for three passive tracers with initial fields given by (2.9) with $a = 1$, 2, and 5 (top to bottom), all subject to the same initial dynamical profiles as configuration III. The right-hand column depicts the initial and final profiles for the isopycnal mean (solid blue and black lines, respectively) and final virtual profiles (dashed red line), while profiles at specific times are depicted in figure 16.

During the development of the fastest-growing instabilities, and prior to the formation of secondary instabilities around $t = 56$, the virtual and isopycnal mean profiles agree almost perfectly for all three tracers. After the onset of the secondary instabilities, the effective diffusivity overestimates the width of the tracer layers, and the virtual tracer layer is broader than the isopycnal mean tracer layer for all three tracers. At intermediate times (e.g. $t = 86$), the virtual profiles are smoother than the isopycnal mean profiles, which have slowly changing, low concentration values at the edges, with sharp increases leading towards flatter maximums towards the middle of the domain. This trend continues for the duration of the simulation, although the isopycnal mean and virtual profiles show much better agreement by the end. In each case, the virtual profile tends to underestimate the tracer concentration at the middle of the domain, and overestimate the amount at the edges of the layer. Despite the variation in initial profile widths and maximums, the final profiles for each of the tracers are all of similar magnitudes and shapes.

Figure 17 presents the evolution of density–tracer scatter plots for each tracer layer with different initial widths. The tracer axis has been normalised by the maximum value for the wide tracer layer ($a = 1$) at each time step to better compare scatter plots. Additionally, scatter plots are magnified in the vertical direction between the depicted times. During the initial roll-up of the KH billows, as depicted in figure 17(*a,g,m*), there is an increase in scattering along ϕ-space as a increases. By $t = 56$ at the onset of secondary
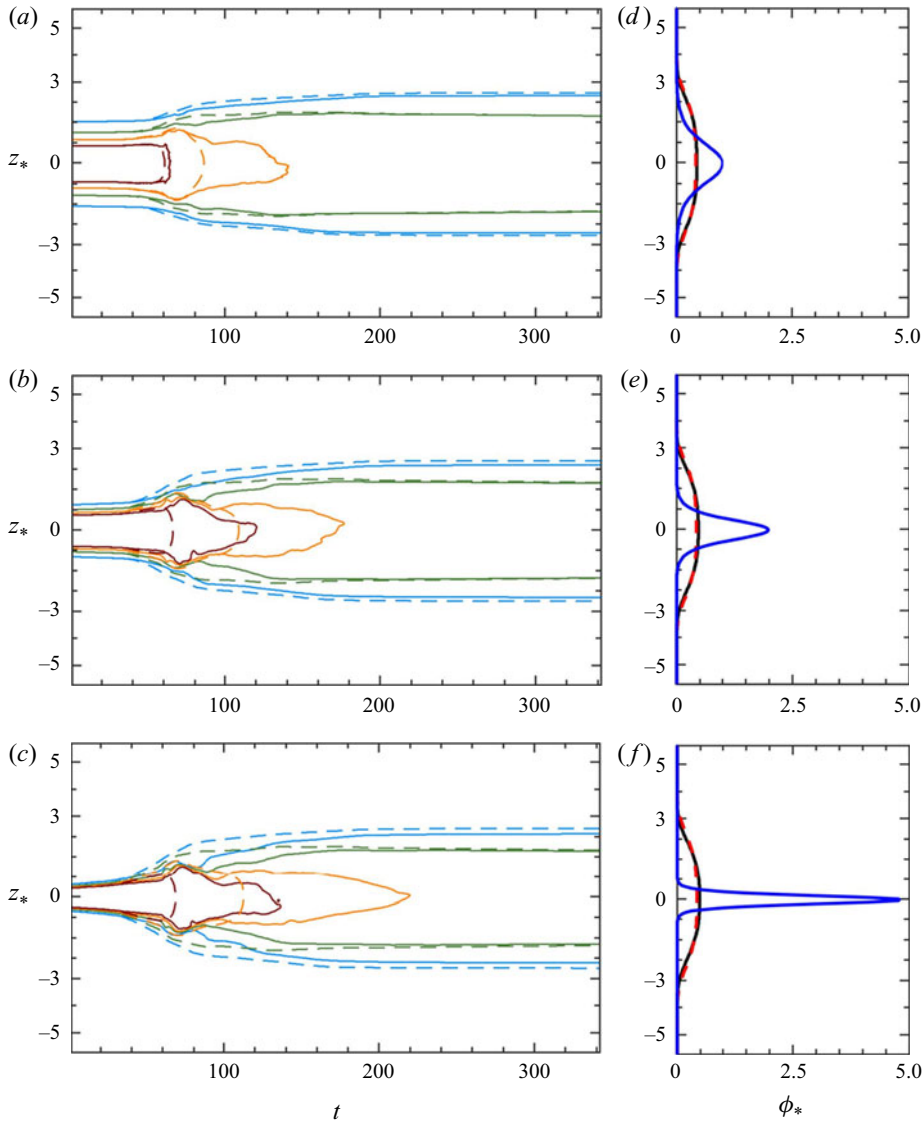
FIGURE 15. (*a–c*) Evolution of the isopycnal mean tracer profiles from simulation (solid lines) and virtual tracer profiles (dashed lines) with time: (*a*) wide ($a = 1$), (*b*) medium ($a = 2$) and (*c*) narrow ($a = 5$). (*d–e*) Corresponding initial tracer profiles (blue), final isopycnal mean tracer profiles (black) and final virtual profiles (red).

instabilities, there is significant scattering induced by the turbulent mixing, filling the area below the initial curves. The maximum concentration of each tracer has been reduced by approximately 20 %, and there are strong discrepancies between all scatter plots. The outer edges of the scatter plots collapse to protrusions mostly centred at $\rho = 0$ ($t = 85$) while maintaining similar maximum tracer values to the previous times. By $t = 140$, the difference in maximum tracer values of each of the scatter plots has greatly reduced. The amount of scatter over a given density bin varies, however, increasing as the initial profile becomes narrower (*a* increasing). By $t = 260$, convergence of the scatter plots to similar shapes becomes apparent, with each displaying similar compact relationships in
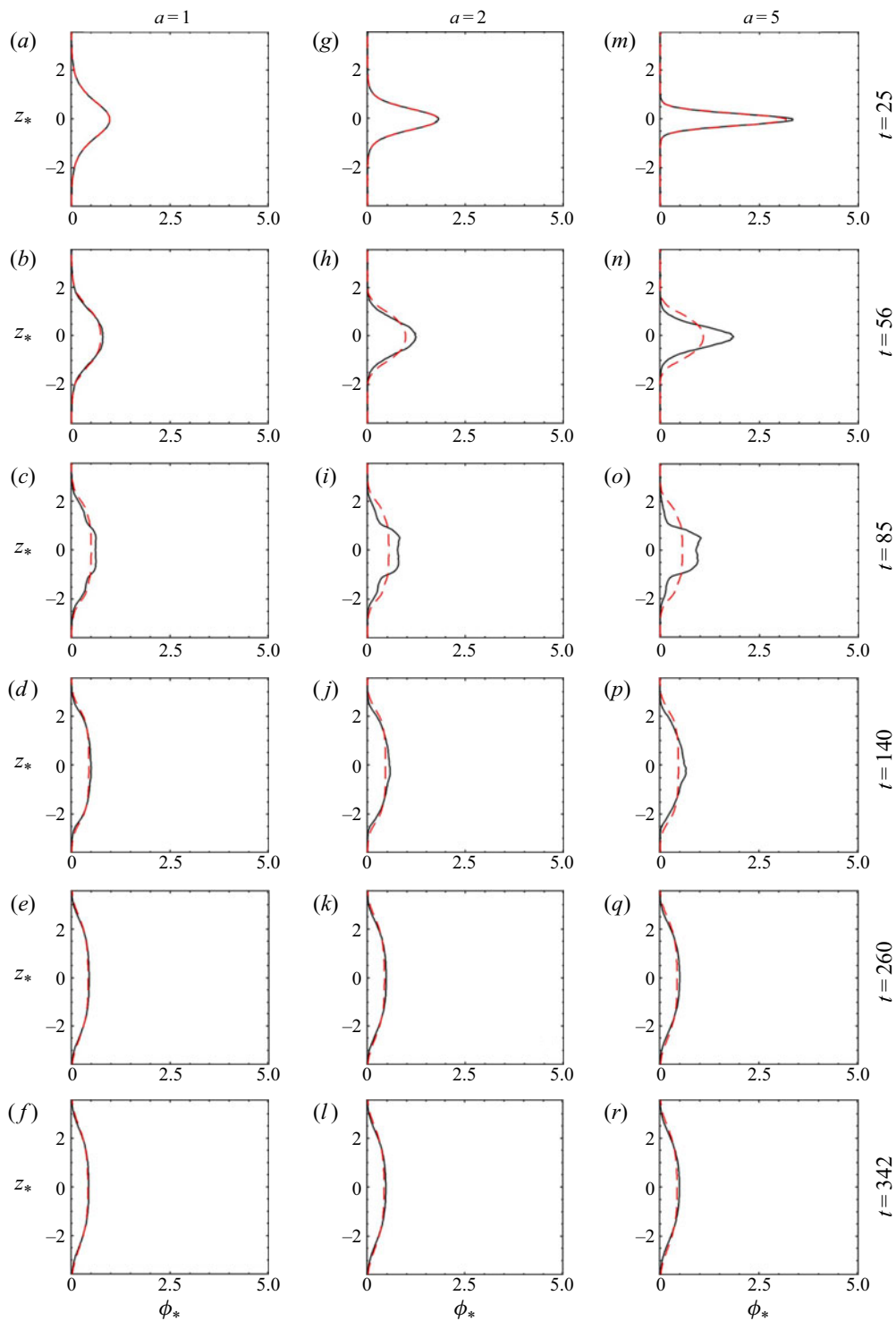
FIGURE 16. Isopycnal mean tracer profiles from simulation (solid black lines) and virtual tracer profiles (dashed red lines) for the variable tracer layer width simulations ((a–f) wide ($a = 1$), (g–l) medium ($a = 2$) and (m–r) narrow ($a = 5$)).

much closer agreement, with minor differences in tracer maximum and concavity. As per figure 10, the effects of turbulent mixing have diminished significantly by this point, with a maximum effective diffusivity of approximately 2.5. Magnifying the scatter plots at the end of the simulation ($t = 342$) shows that the differences in tracer maximum and concavity persist after the turbulent phase has ended, and the tracers are subject only to diffusive mixing. The similarity between these scatter plots at the end of the simulation suggests that for an arbitrary initial tracer distribution, provided it is entirely within the mixing region, and a sufficiently mixed state has been reached, a reasonable estimate of the final tracer distribution might be obtained using the piecewise linear formulation suggested in § 6.

The density–tracer scatter plots with the narrow and medium distributions show important cross-isopycnal fluxes of tracers, with a final distribution extending to a wider density range. The amount of tracer affecting new density ranges can be calculated from the knowledge of the density evolution and the total initial tracer amount. Finally, while all three scatter plots tend to nearly piecewise linear tent shapes, the wide (reference) case ($a = 1$) is slightly convex on either side of the tracer maximum, while the scatter plots for the other two tracers are concave on either side of the maximum, with the narrow case ($a = 5$) being more concave than the medium case ($a = 3$). Additionally, the tracer concentration maximum increases slightly as a increases. This relationship between scatter plot concavity and tracer concentration maximum reflects the observations made about deviations away from the ideal tent shape made in § 6.

7.2. *Sensitivity to the vertical position of the tracer*

In the results presented in previous sections, all tracer profiles are initially concentrated within the region subject to turbulent mixing. As a result, the amount of tracer initially for densities less than $\rho = \rho_L$ and greater than $\rho = \rho_H$ is null. In this section, the validity of the convergence principle is tested for configurations where the tracer structure is offset above the main turbulent region, so not all the tracer is located within the mixing range. These initial tracer profiles are defined by (2.9), with $a = 1$, and b varied to change the initial vertical position of the tracer. For this simulation, $b = 0$ sets a layer centred along the pycnocline (i.e. the reference tracer configuration), $b = 2.86$ is offset above the pycnocline by 10 % of the vertical extent of the domain and $b = 5.72$ is offset above the pycnocline by 20 % of the vertical extent of the domain. The evolution of tracers subject to these initial conditions for the 3-D configuration III are presented as cross-sections in figure 18. From left to right, the tracers are aligned with the midpoint of the pycnocline ($b = 0$), slightly offset from the pycnocline ($b = 2.86$), and completely offset from the pycnocline ($b = 5.72$). The corresponding density evolution is unchanged, and is depicted in the two right-hand columns of figure 3. The slightly offset tracer initially resides on the edge of the fastest-growing vortex, so that when the KH billow first develops, some of the tracer is entrained into the pycnocline by the vortex, and the tracer layer is significantly distorted. With the onset of small-scale features from the secondary instabilities, the bottom half of this tracer layer is eroded and a thick layer of low tracer concentration develops, extending to the lower edge of the pycnocline. These secondary instabilities affect the interior and upper parts of the high concentration region, but do not significantly mix regions of low and high tracer concentration (e.g. figure 18k), and there remains a layer of relatively high tracer concentration mostly unaffected by mixing. Along the upper edge of this layer, molecular diffusion is the primary source of slow mixing, but this mixing is weak enough to be considered negligible when compared to the strong mixing happening below. The layer of tracer completely offset from the pycnocline is weakly distorted by the
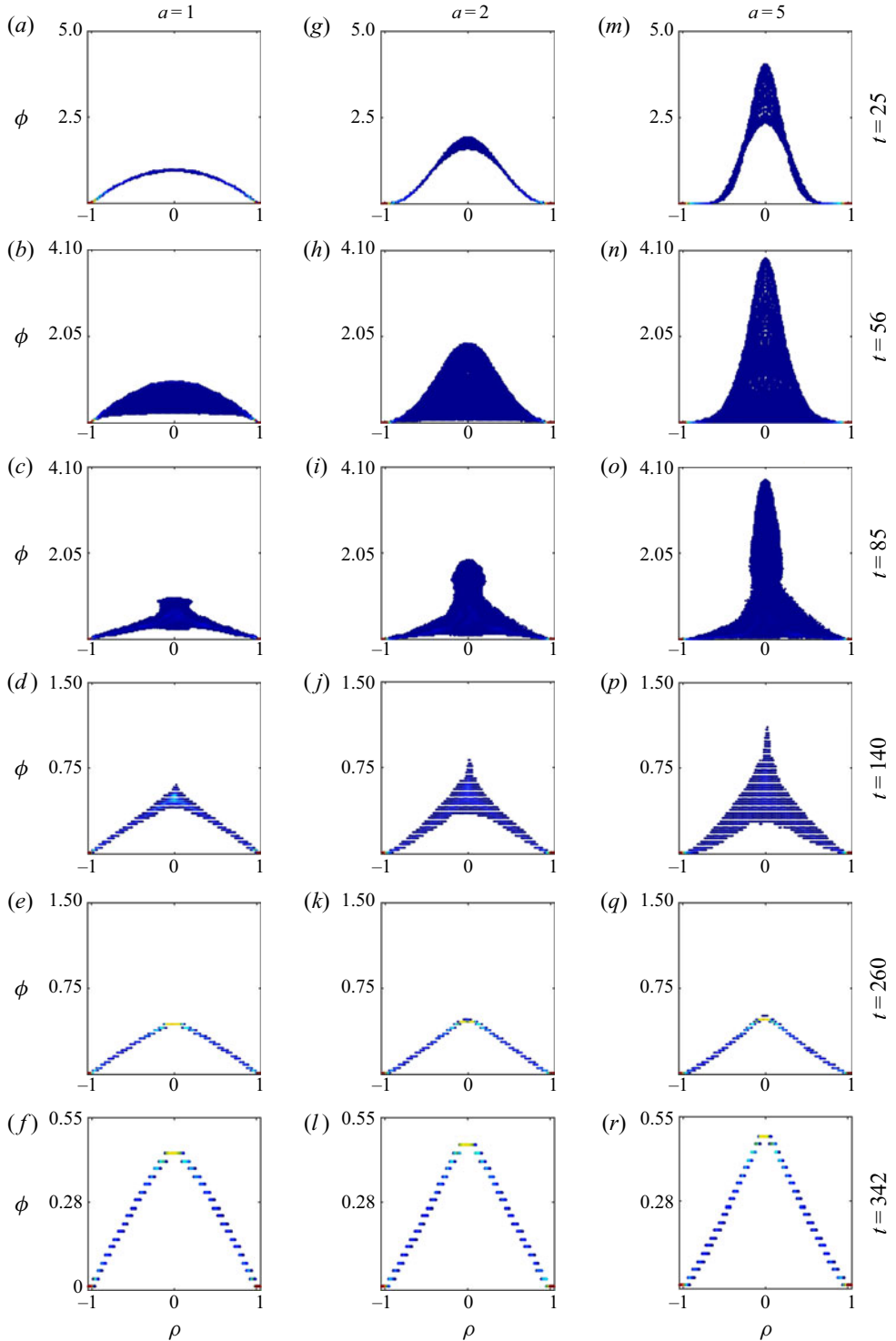
FIGURE 17. Density–tracer scatter plots for tracer layers centred along the pycnocline with varying widths, but identical tracer totals ((a–f) wide ($a = 1$), (g–l) medium ($a = 2$) and (m–r) narrow ($a = 5$)). Note that the tracer axis is adjusted after each time step as the scatter plots collapse.
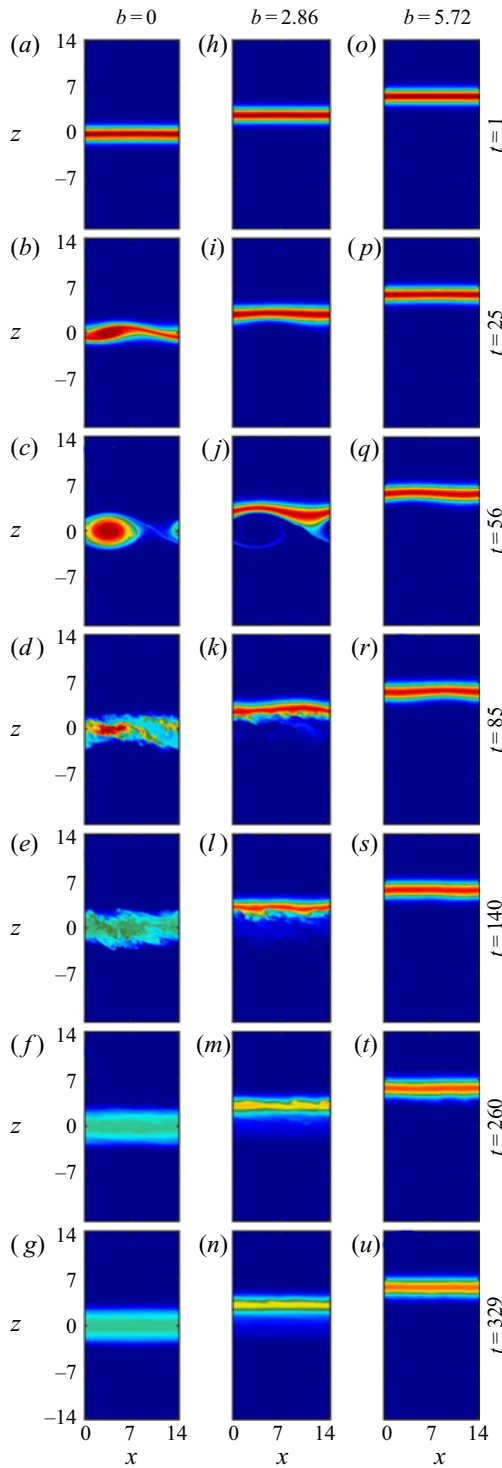
FIGURE 18. Cross-sections of the evolution of tracer layers centred at various depths ((*a–g*) along the pycnocline ($b = 0$), (*h–n*) slight offset ($b = 2.86$) and (*o–u*) completely offset ($b = 5.72$)).
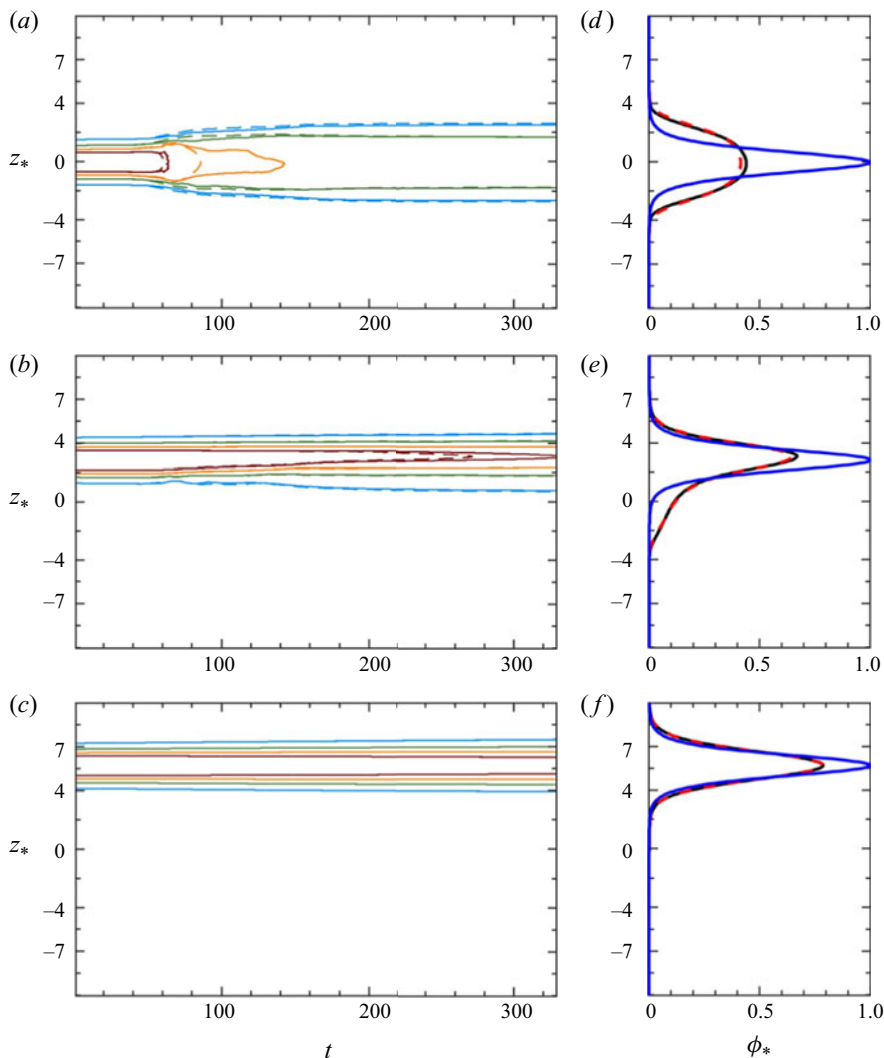
FIGURE 19. (*a–c*) Evolution of the isopycnal mean tracer profiles with time, (*a*) along the pycnocline ($b = 0$), (*b*) slightly offset from the pycnocline ($b = 2.86$) and (*c*) completely offset from the pycnocline ($b = 5.72$), as computed from the simulation (solid lines) and the virtual tracer profiles (dashed lines). (*d–f*) Corresponding initial tracer profiles (blue), final isopycnal mean tracer profile (black) and final virtual profile (red).

outer edge of the billow, but experiences almost no mechanical mixing (some strands of tracer pulled into the secondary instabilities are visible in figure 18*t*), mostly a spreading of tracer due to diffusion. The effect of the localised mechanical mixing away from the offset tracer layers is visible in the isopycnal mean tracer profiles. The complete time evolution of the isopycnal mean and virtual tracer profiles is presented in the left-hand column of figure 19, with the initial and final profiles presented in the right-hand column. Profiles at specific times corresponding to the cross-sections of figure 18 are presented in figure 20.

Compared to the tracers collocated with the pycnocline presented in previous sections, the offset isopycnal mean tracer profiles typically show greater agreement with the virtual
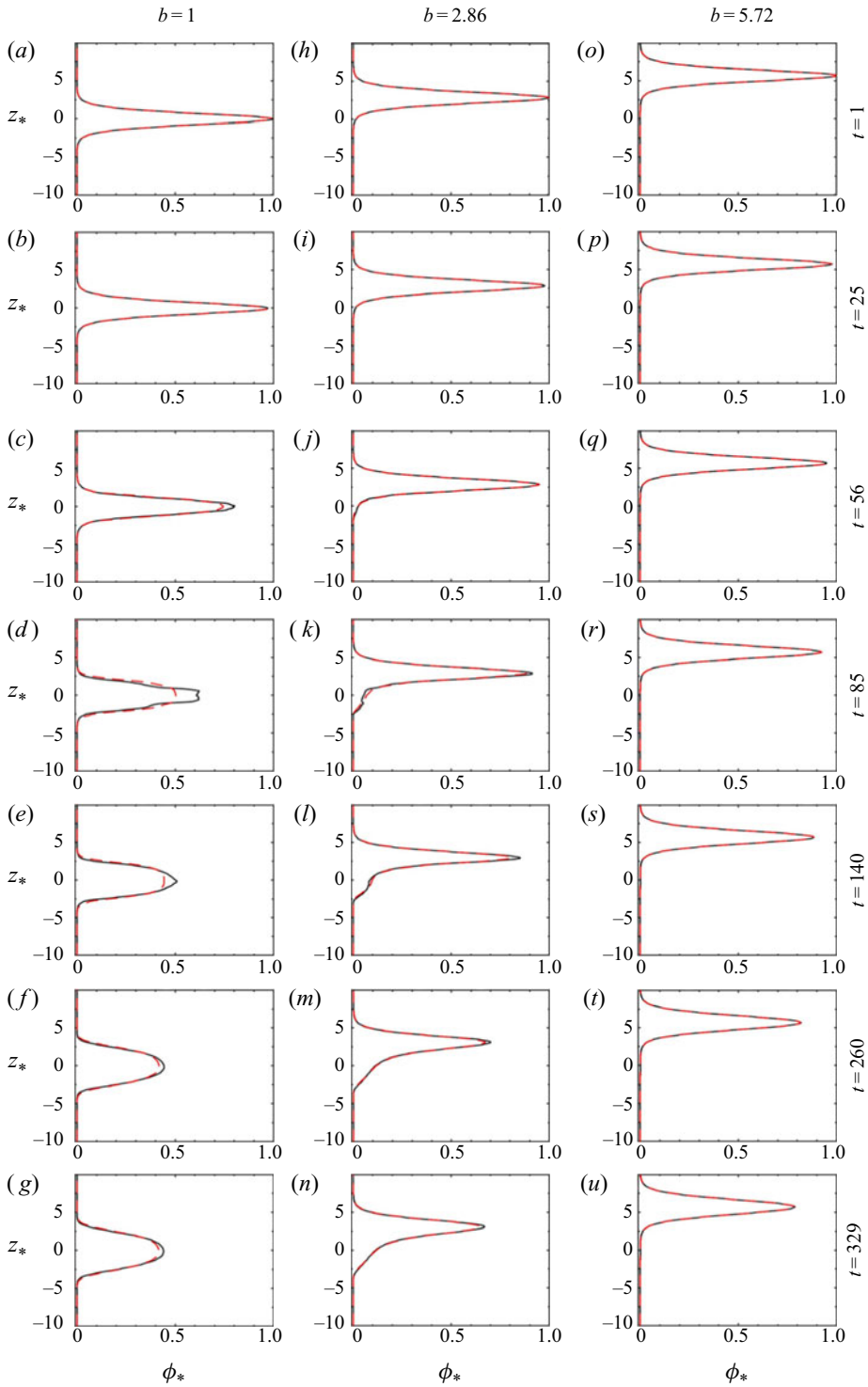
FIGURE 20. Isopycnal mean tracer profiles (solid black lines) and virtual tracer profiles (dashed red lines) for the simulation where the passive tracers are offset from the pycnocline: $(a–g)$ along the pycnocline ($b = 0$), $(h–n)$ slightly offset from the pycnocline ($b = 2.86$) and $(o–u)$ completely offset from the pycnocline ($b = 5.72$).

FIGURE 21. Density–tracer scatter plots for tracer layers centred at various depths: (*a–g*) along the pycnocline (*b* = 0), (*h–n*) slight offset (*b* = 2.86) and (*o–u*) completely offset (*b* = 5.72).

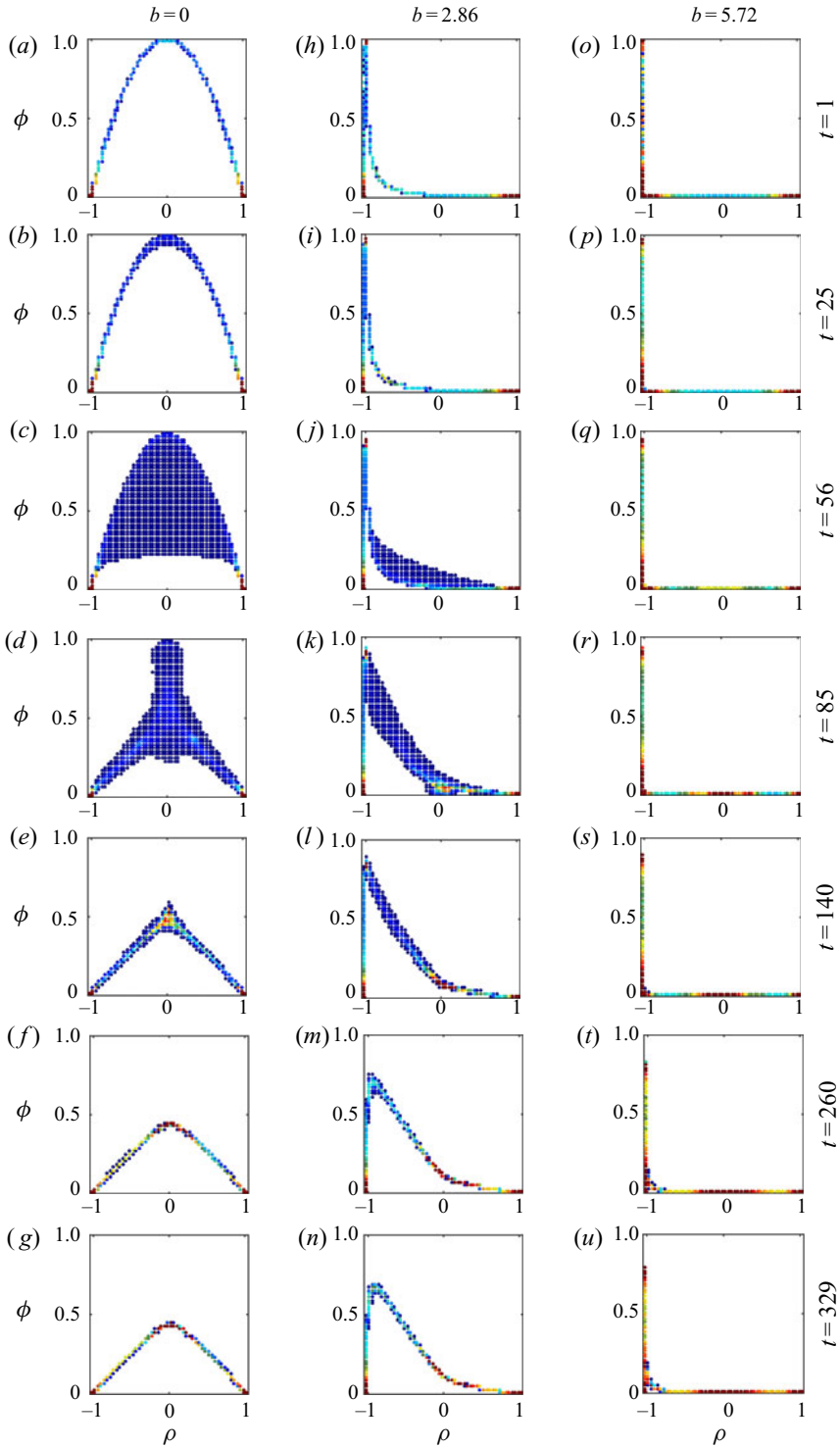profiles throughout the evolution of the flow. The slightly offset tracer shows some disparity at the lower edge of the virtual and isopycnal mean profiles after the onset of the secondary instabilities, and the virtual profile underestimates the maximum value of the profile towards the midpoint of the simulation (e.g. $t = 140$). For the completely offset tracer, the virtual and isopycnal mean profiles appear to agree perfectly for the duration of the simulation. This can be explained by considering that the tracer is located mostly outside the region of strong turbulent mixing and is essentially subject only to slow molecular diffusion. The tracer eddy term is almost null for this tracer distribution, since both ϕ' and the gradient of density are nearly zero over the density levels in that region. The full equation for ϕ_* therefore reduces to the virtual tracer equation in this case.

Figure 21 presents the scatter plots of each of the tracers in figure 18 relative to density. The first thing to note is the offset tracer scatter plots take different initial forms than the centred tracer scatter plot. They are asymmetric, and have different convex envelopes, essentially triangles with corners at the top left, bottom left and bottom right corner of the graphs. The slightly offset tracer ($b = 2.86$) experiences most of its scattering above the concave region of the plot, and tends to a final shape that is nearly piecewise linear, with lines of two different slopes that intersect near $\rho = 0$. The tracer maximum occurs along an isopycnal with a density lower than the mean. The cross-isopycnal flux of tracer is strong in this case, since at the end of the simulation, a significant quantity of the tracer is distributed over a much wider density range. For most of the flow, the completely offset tracer is modestly affected by mixing and does not show much scattering, although some redistribution of the scatter plot is visible at the lower edge of the tracer layer where it is more significantly eroded by mixing. This tendency for the completely offset tracer to maintain a nearly compact relationship is indicative of the fact that the eddy term remains negligible for the entire simulation. As for the simulations with different tracer layer widths, the present offset configurations show that mixing can strongly modify tracer concentration and entrain significant amounts of tracer into new density ranges.

8. Conclusions

This paper uses the classical problem of turbulence arising from Kelvin–Helmholtz instabilities to investigate the mixing of passive tracers. Three different stratified shear configurations providing different routes to turbulence are simulated, and the mixing of different passive tracers layers is examined. Simulations examining the mixing of tracer layers with different widths and vertical positions are performed. Weighted density–tracer scatter plots are proposed as a method through which to analyse diapycnal transport of tracer. When tracer and density diffusivities are equal, a convex envelope constraint can be placed on the evolving scatter plot, which limits on the maximum tracer values achievable in specific density ranges. In the initial stage of the KH instability, when the density and tracer are materially conserved there is no modification of the scatter plots. The stirring effect of the growing KH disturbance eventually enhances diffusive mixing which yields irreversible modifications of the density–tracer scatter plot, such that the plotted points fill a finite area, over the density range affected by the mixing. The area filled then reduces as the tracer and density then relax towards a different compact relation than that set by the initial conditions. Note that a peculiarity of the 2-D case is that there may be a persistent volume of fluid with essentially constant density within which the tracer varies. This does not occur in the more realistic 3-D case. The formulation for the scatter variance as introduced by (4.7) and presented in figure 7 provides an additional diagnostic to compare the evolution of the tracer layer collocated with the pycnocline in each of the three dynamical configurations. Times at which the diagnostic is near zero can also be

used to identify if the relationship between the density and tracer has become compact. This can supplement the information available in the weighted scatter plots, as it can be difficult to ascertain whether a compact density–tracer relationship has been reached based only on graphical analysis.

The use of the effective diffusivity derived from the density field as a model for the mixing of the passive tracer has also been examined. This model does not predict the details of the mixing as the KH billow develops (e.g. see figure 12c,i,o), but in configuration II and configuration III (which is the 3-D case and therefore considered to be the most realistic) does make a good prediction of the cumulative effect of the mixing (e.g. figures 12(l,q), 15 and 19). Note that the disagreement is most apparent during the stages of the mixing when a compact density–tracer relationship has not been established. In the single billow case, configuration I, such a relationship is never established, although this is considered to be rather unrealistic.

The approach taken in this paper is to investigate, for ϕ representing the concentration of a passive tracer, to what extent most of the evolution of ϕ_* is captured by (5.5), with the eddy term ignored. A question beyond the scope of this paper is whether, for cases where the eddy term cannot be ignored, can it be estimated using a closure assumption and whether this can be accomplished whilst retaining the advantages that have been demonstrated for the standard use of a tracer-based coordinate and the corresponding effective diffusivity. Note that Shuckburgh & Haynes (2003) consider whether particle transport can be approximated using an effective diffusivity calculated from a tracer field, i.e. to $\kappa_\phi = 0$, corresponding to Example 2 in appendix C. Some success (at least semi-quantitative) was demonstrated. In terms of the new results given above, this would be equivalent to

$$\frac{\partial}{\partial z_*}\left(-\left\langle\phi'\kappa_\rho\nabla^2\rho\right\rangle_{z_*}(\partial z_*/\partial\rho)\right) \simeq \frac{\partial}{\partial z_*}\left(K_\rho\frac{\partial\left\langle\phi\right\rangle_{z_*}}{\partial z_*}\right). \tag{8.1}$$

It remains to investigate whether the approximate equality above holds (and, if it does, under what circumstances).

This paper has described tracer transport in KH billows when the diffusivity of the tracer is equal to the diffusivity of the density, i.e. the diffusivity ratio $\tau = \kappa_\phi/\kappa_\rho$ is equal to 1. However, as noted in the introduction, in geophysical cases τ is typically much smaller than 1. For example, in an oceanic environment where the density is controlled by temperature and the passive tracer is some dissolved chemical species, $\tau \sim O(10^{-2})$. Reducing τ from 1 may change the description of tracer transport in several ways. In particular it will break the convex envelope constraint placed on the scatter plot. In the hypothetical limit where only density diffuses and the passive tracer is non-diffusive ($\tau = 0$), the displacement of the points on the scatter plot would be strictly horizontal, eventually converging to a vertical line located at the mean density. Similarly, if only the passive tracer were diffusive and the density is non-diffusive ($\tau \to \infty$), the displacement of points would be entirely vertical, leading to a horizontal line at the tracer mean. It therefore remains to be determined what constraints can be placed on scatter plots for tracers with different diffusivities. Reducing τ may also alter how well the virtual tracer represents the isopycnal mean tracer evolution. Preliminary simulations in which the tracer density is reduced moderately relative to that for the density ($\tau \sim O(10^{-1})$) have shown reasonable agreement between the virtual and isopycnal mean tracer profiles in the sense that the virtual tracer still provides useful estimate of the cumulative effect of the mixing event. However, substantial further work would be needed to determine whether this held at smaller values of τ potentially relevant to oceanic mixing.

Unfortunately, modelling systems in which two tracers have significantly different diffusivities can be computationally expensive due to the large number of grid points required to adequately resolve the lower diffusivity tracer (Penney & Stastna 2016), and is beyond the scope of the present study.

It also remains to be determined if these scatter plots provide a useful analysis in the event where the tracer is an active component of the density. For example, if ρ is a function of ϕ only, then a $\rho\phi$-scatter plot would simply depict the functional form of $\rho = \rho(\phi)$. However, if ρ is a function of two tracers, $\rho = \rho(\phi_1, \phi_2)$, it remains to be seen what information could be derived from $\rho\phi_1$ or $\rho\phi_2$ scatter plots. It is also worth considering a situation in which one is concerned with the mixing of a passive tracer when density is a function of two different active tracers, such as for double-diffusive flows (for example in the ocean, where salt and temperature are the active tracers). Because double diffusion allows for the creation of new density anomalies, this immediately presents the possibility of breaking the convex envelope constraint presented here.

Diffusion is a major component of the physics discussed in the present paper. It is the process by which tracer characteristics are eventually modified, although as demonstrated here, there exist constraints on these modifications that lead to a predictable evolution. The knowledge and control of diffusive terms in numerical models is thus a major issue. In this respect, another result of interest discussed in the present study concerns the evaluation of a net diffusivity to be used in place of the prescribed 'explicit' one. Indeed, discretisation and numerical schemes used in models can involve implicit diffusion, accounting for which is shown here to be essential when evaluating (5.5)–(5.7). Note also that implicit diffusivity does not modify the physics of mixing discussed here. For instance, in addition to the $N_x \times N_z = 512^2$ resolution simulation of configuration II presented in this article, additional tests at resolutions of $N_x \times N_z = 256^2$, 384^2 and 768^2 were performed (not shown). In general, the lower the resolution, the larger the implicit numerical diffusivity, and thus, the larger the net diffusivities of the density and tracers. Each of these simulations (including the 768^2 resolution simulation for which the implicit diffusivity is practically negligible), lead to similar convergence of the scatter plots, as well as the virtual and isopycnal mean tracer profiles. This process is thus controlled by diffusivity whatever its origin (prescribed or implicit), not by other properties of the model's numerical schemes. Finally, it is worth mentioning that in oceanic or atmospheric circulation models, horizontal and vertical advection terms are often discretised with different operators involving parameterisations. In these cases, the implicit and explicit diffusion are no longer isotropic, which may violate the constraints and taint the mixing physics discussed in the present paper. A detailed study of such numerical effects is worth investigating.

An important part of the behaviour described here is the evolution towards a compact density–tracer relation during the final stages of the mixing event. This is rather similar to the tracer redistribution along streamlines examined from a theoretical perspective for steady and oscillatory 2-D flows by Rhines & Young (1983). They describe how, from an arbitrary initial condition, passive tracers will tend to homogenise along streamlines (hence resulting in a compact relationship between a tracer and streamfunction, or indeed between any two tracers). This description cannot be easily extended to the configurations studied here, as density is not a passive tracer, there are strong 3-D effects (for configuration III), and convergence is achieved during the chaotic mixing phase (see figure 7) when no clear streamlines can be identified. On the other hand, the density field clearly plays a similar organisation role on the flow as the steady streamfunction in the Rhines & Young (1983) case. On general dynamical grounds it is plausible that motion will be stronger along density surfaces than across them and that passive tracers will therefore

tend to homogenise along density surfaces. A more complete description of this process would predict the time evolution for the homogenisation and its relation to the time decay of the turbulence in the later stages of the KH mixing event.

Acknowledgements

Declaration of interests

The authors report no conflict of interest.

Appendix A. Non-Boussinesq, non-hydrostatic system of equations (CROCO model)

The numerical models used in the study of stratified shear generally consider that volume (rather than mass) is conserved, and that the flow satisfies the Boussinesq assumption. For non-hydrostatic, incompressible processes such as Kelvin–Helmholtz instabilities, the system of Boussinesq equations (2.1) accurately represents the flow. From a numerical point of view, the non-hydrostatic implementation of an incompressible flow degenerates to an elliptic system of equations (with respect to acoustic waves). A 3-D Poisson equation must thus be solved to obtain a pressure 'avatar' for the resulting flow and several types of algorithms (based on, for example, pressure projection (e.g. Subich, Lamb & Stastna 2013) or pressure correction (e.g. Jiang 2019)) can be implemented to simulate non-hydrostatic flows. When dealing with geophysical flows, whether in the atmosphere or in the ocean, the re-introduction of acoustic waves can lead to efficient numerical algorithms for non-hydrostatic flows (see for instance Janjic, Gerrity & Nickovic (2001) for atmospheric flows and Auclair *et al.* (2018) for oceanic flows). Several reasons can explain this at-first-glance paradoxical conclusion: energy transfers are simulated in a presumably consistent way (Tailleux 2009), and the resulting computations are 'local' and they are thus well adapted to massively parallel implementations. In the ocean, the re-introduction of acoustic waves additionally provides a convenient way to simulate non-hydrostatic flows with time-splitting algorithms (Auclair *et al.* 2018).

The numerical simulations presented in this article were performed using the CROCO ocean model. CROCO is a free-surface, non-hydrostatic, non-Boussinesq model. It inherited from ROMS and ROMS-AGRIF (Shchepetkin & McWilliams 2005; Debreu *et al.* 2012) their numerically efficient two-mode leapfrog–third-order Adams–Moulton (LF-AM3) time splitting. The non-hydrostatic implementation of this time splitting is conveniently obtained by calling into question the Boussinesq assumption and re-introducing acoustic waves together. The resulting compressible processes are treated in the LF-AM3 fast mode which cannot remain 2-D (barotropic) and becomes 3-D. The resulting system of equations is given in dimensional form by

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \boldsymbol{u}), \qquad (A\,1a)$$

U_0	1 m s^{-1}
ρ_0	1017 kg m^{-3}
$\Delta\rho$	1.2 kg m^{-3}
h	10 m
z_0	-143 m
ν	$5 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$
κ_ρ	$5 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$
κ_ϕ	$5 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$
g	9.81 m s^{-2}
λ	$100 \text{ m}^2 \text{ s}^{-1}$

TABLE 2. Prescribed dimensional parameters consistent across all simulations presented in this paper.

Configuration	$N_{\Delta t_{Fast}}$	c_s
I	10	500 m s^{-1}
II	10	500 m s^{-1}
III	8	350 m s^{-1}

TABLE 3. Relevant computational parameters of the dynamical configurations presented in this paper.

$$\frac{\partial \rho \boldsymbol{u}}{\partial t} = -\boldsymbol{\nabla} \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) - \boldsymbol{\nabla} p - \rho g \hat{\boldsymbol{k}} + \rho \nu \nabla^2 \boldsymbol{u} + \rho \lambda \boldsymbol{\nabla} \boldsymbol{\nabla} \cdot \boldsymbol{u}, \tag{A 1b}$$

$$\frac{\partial \rho_h}{\partial t} = -\boldsymbol{\nabla} \cdot (\rho_h \boldsymbol{u}) + \kappa_\rho \nabla^2 \rho_h, \tag{A 1c}$$

$$\frac{\partial \rho \phi}{\partial t} = -\boldsymbol{\nabla} \cdot (\rho \phi \boldsymbol{u}) + \rho \kappa_\phi \nabla^2 \phi, \tag{A 1d}$$

$$\rho = \rho_h + \delta\rho, \tag{A 1e}$$

where λ is the bulk viscosity. Dimensional parameters used in all simulations presented in this paper, including those used in non-dimensionalisation, are listed in table 2.

Density components ρ_h and $\delta\rho$ are respectively the hydrostatic, Boussinesq and the non-hydrostatic, non-Boussinesq contributions to density. In the current implementation, the latter compressible density anomaly is linked to the pressure by the first-order relation $\delta\rho = c_s^{-2}\delta p$, with $\delta p = p - p_h$, where p_h is the hydrostatic component of pressure and c_s is the speed of sound. As in the original ROMS-AGRIF hydrostatic time splitting, the free-surface anomaly is prognosticated in CROCO fast mode. The speed of sound and number of fast time steps per slow time step, $N_{\Delta t_{Fast}}$, are given in table 3.

Concerning first the consequences of the explicit modelling of acoustic waves, sensitivity tests have been carried out in the frame of the present study. The propagation speed of acoustic waves has been modified, and no significant consequences have been found on the evolution of KH instabilities. As far as the consequences of the surface-induced processes are concerned, the studied configurations are based on deep pycnocline and shear-stress layers, hence minimising the interactions between the

free-surface and Kelvin–Helmholtz instabilities. Sensitivity studies with a deeper water column (and a deeper pycnocline and shear layer) have shown that the development of the instability is not influenced by the free surface. Only a slight influence in the difference between upper and lower boundary conditions was observed during the pairing and subsequent mixing in the 2-D case with two billows, as the vertical extent of the final billow becomes large. This leads to an asymmetry of the mixing at the edge of the billow, but this effect does not modify the conclusions presented here.

Appendix B. Net diffusivity estimate

Provided the fluid is incompressible, for a tracer governed by an advection–diffusion equation in the form,

$$\partial \rho / \partial t + \boldsymbol{u} \boldsymbol{\nabla} \cdot \rho = \kappa_\rho \nabla^2 \rho, \tag{B 1}$$

one can solve for the diffusivity at a given time by multiplying the equation by the tracer field, then integrating over the entire domain, giving

$$\kappa_\rho = -\frac{\int_V \dfrac{\partial \rho^2}{\partial t} \, \mathrm{d}V}{2 \int_V |\boldsymbol{\nabla} \rho|^2 \, \mathrm{d}V}. \tag{B 2}$$

For an ideal numerical simulation with perfect numerical schemes, the resulting value of κ_ρ would be constant and equal to the prescribed value at all times. In reality, numerical schemes are imperfect, and the resulting effects due to extra terms can act as extra diffusion inherent to the schemes. The resulting 'net diffusivity' will vary with time, and depend on the schemes used, the grid size, the time step and the flow geometry,

$$\kappa_\rho^{Net}(t) = -\frac{\int_V \dfrac{\partial \rho^2}{\partial t} \, \mathrm{d}V}{2 \int_V |\boldsymbol{\nabla} \rho|^2 \, \mathrm{d}V}. \tag{B 3}$$

As an example, the net diffusivities as calculated by (B 3) and normalised by the prescribed molecular value (5×10^{-3} m^2 s^{-1}) are presented as a function of time in figure 22 for the density and passive tracer fields with initially different vertical positions (as described in § 7.2) of configuration III. The net estimates all begin at the prescribed molecular value, with the time series for the density and the $b = 0$ tracer showing similar overall trends: an initial increase of approximately 100 % corresponding to the stretching of the field interfaces by the initial billow formation before decreasing to near the prescribed value as the secondary instabilities collapse the billow. The evolution of the net diffusivity of the middle tracer ($b = 2.86$) is similar to that of the density and $b = 0$ tracer, although it only sees a maximum increase of approximately 70 %. In contrast, the upper layer of tracer ($b = 5.72$) experiences a much smaller increase in net diffusivity of about 10 % when the billow initially distorts the field. This increase away from the maximum is much less significant and shorter lived, as there is little meaningful interaction being the upper tracer and the dynamic mixing process.

For this study, the net diffusivity was calculated in order to account for the effects of numerical processes when calculating profile evolution. Figure 23 depicts the complete time evolution (with figure 24 depicting profiles at individual times) of the background density profile as a calculated from the rearrangement of the density from simulations (solid lines) versus the profiles calculated from (5.1) using the constant prescribed
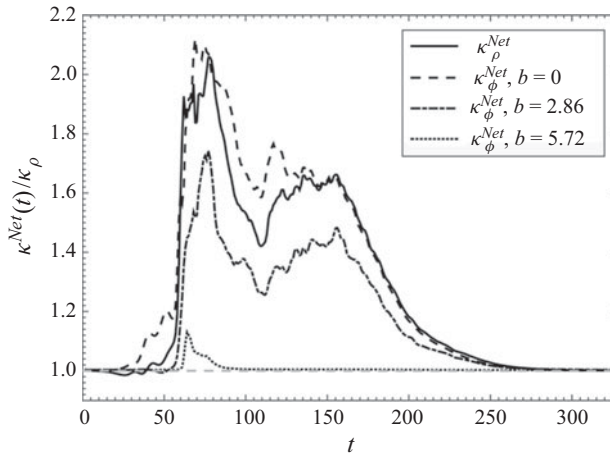
FIGURE 22. Evolution of the estimate of global effective diffusivity for the density (solid line) and passive tracer (dashed lines) fields for each of the offset tracers described in § 7.2 normalised by the prescribed molecular diffusivity.

diffusivity value κ_ρ for the simulation (figure 23a), and the time-dependent net diffusivity $\kappa_\rho^{Net}(t)$ computed from the simulation using (B 3) (figure 23b). Note that the theory used to derive (5.1) and (5.5) as described in appendix C remains unchanged if the diffusivity in the advection–diffusion equation is time dependent, and thus the net diffusivity can be directly applied. As shown in figure 23(a), the isopycnals calculated from the diffusion equation typically lie within the sorted isopycnals, indicating that the constant diffusivity leads to the spreading of the pycnocline being underestimated. Using the time-dependent net diffusivity shows better agreement between the calculated and sorted isopycnals. This is also apparent in the later times shown in figure 24, which also demonstrates that the profiles calculated using the prescribed diffusivity underestimate the steepness of transition at $\rho_* = 0$.

Appendix C. Diffusion equation of one tracer relative to another

C.1. *Starting point*

Consider flow in a 3-D domain with horizontal area A_d and, for definiteness, assume that the domain is bounded in the vertical coordinate z. The density $\rho(\boldsymbol{x}, t)$ is assumed to satisfy the advection diffusion equation

$$\frac{\partial \rho}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho = \kappa_\rho \nabla^2 \rho, \tag{C 1}$$

with the velocity field $\boldsymbol{u}(\boldsymbol{x}, t)$ being non-divergent, and the diffusivity κ_ρ being either constant or a function of time. Note that for convenience, all variables presented in this appendix are given in their dimensional forms. Following Nakamura (1996) and Winters & D'Asaro (1996), the density field may be used to define a new vertical coordinate z_*, as follows. For a specified value ρ_* of the density, define

$$z_*(\rho_*, t) = \frac{1}{A_d} \int_{\rho(\boldsymbol{x}, t) \geq \rho_*} \mathrm{d}V. \tag{C 2}$$
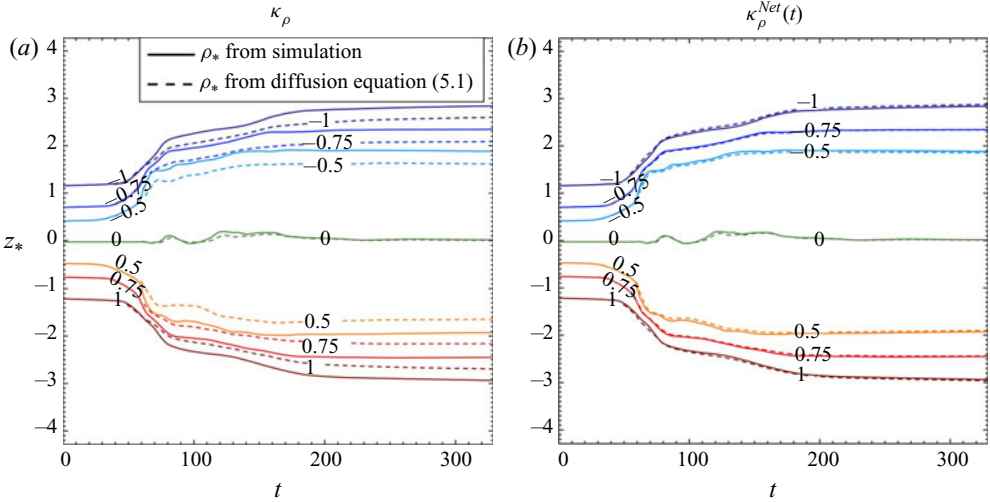
FIGURE 23. Evolution the density profiles calculated from (5.1) using the constant prescribed value of diffusivity (*a*) and a time-dependent value of the diffusivity as calculated from (B 3).
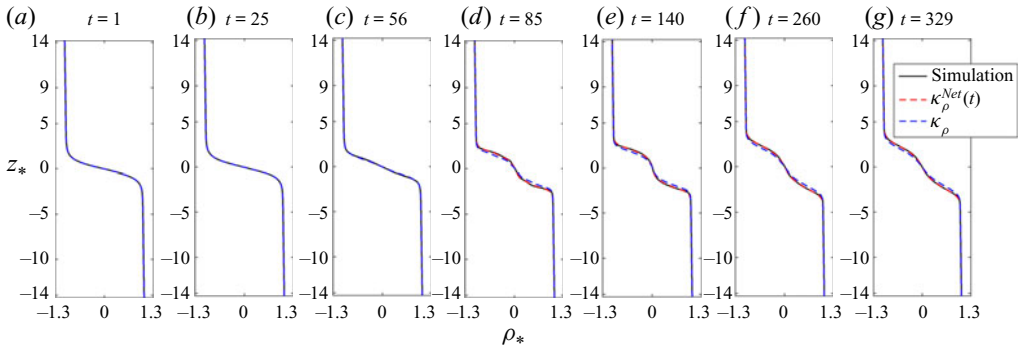

FIGURE 24. Background density profiles as calculated from simulation outputs (solid black line), calculated from (5.1) using the time-dependent, net diffusivity $\kappa_\rho^{Net}(t)$ (dashed red lines), and calculated from (5.1) using the prescribed, constant diffusivity κ_ρ (dashed blue lines).

Note that z_* is a monotonically decreasing function of ρ_*, so the inverse function $\rho_*(z_*, t)$ is well-defined. Now for any field $\phi(\boldsymbol{x}, t)$ define

$$\Phi(\rho_*, t) = \int_{\rho(\boldsymbol{x}, t) \geq \rho_*} \phi(\boldsymbol{x}, t) \, dV. \qquad (C\,3)$$

Simple vector calculus implies that

$$\frac{\partial}{\partial t} \Phi(\rho_*, t) = \int_{\rho(\boldsymbol{x}, t) \geq \rho_*} \phi_t \, dV + \int_{S(\rho_*)} \phi \rho_t \frac{dA}{|\boldsymbol{\nabla}\rho|}, \qquad (C\,4)$$

and

$$\frac{\partial}{\partial \rho_*} \Phi(\rho_*, t) = -\int_{S(\rho_*)} \phi \frac{dA}{|\boldsymbol{\nabla}\rho|}, \qquad (C\,5)$$

where the surface $S(\rho_*)$ is the boundary of the volume $\{x|\rho(x,t) \geq \rho_*\}$ and $\mathrm{d}A$ is the area element on that surface. Note that it follows from (C 4) and (C 5) with $\phi = 1/A_d$ that

$$\frac{\partial}{\partial t} z_*(\rho_*, t) = \frac{1}{A_d} \int_{S(\rho_*)} \rho_t \frac{\mathrm{d}A}{|\nabla \rho|}, \tag{C 6}$$

and

$$\frac{\partial}{\partial \rho_*} z_*(\rho_*, t) = -\frac{1}{A_d} \int_{S(\rho_*)} \frac{\mathrm{d}A}{|\nabla \rho|}. \tag{C 7}$$

The next step is to consider $z_*(\rho_*, t)$, rather than ρ_*, as an independent variable, i.e. to regard ρ_* as a function of z_*, and t. Again, using (C 4) and (C 5),

$$\frac{\partial}{\partial t} \Phi(\rho(z_*, t), t) = -\left(\int_{S(\rho_*)} \phi \frac{\mathrm{d}A}{|\nabla \rho|} \right) \frac{\partial \rho_*}{\partial t} + \int_{\rho(x,t) \geq \rho_*} \phi_t \, \mathrm{d}V + \int_{S(\rho_*)} \phi \rho_t \frac{\mathrm{d}A}{|\nabla \rho|}. \tag{C 8}$$

Now note that the condition on $\partial \rho_*/\partial t$ for $z_*(\rho_*, t)$ to remain constant in time is that

$$-\int_{S(\rho_*)} \rho_t \frac{\mathrm{d}A}{|\nabla \rho|} + \frac{\partial \rho_*}{\partial t} \left(\int_{S(\rho_*)} \frac{\mathrm{d}A}{|\nabla \rho|} \right) = 0. \tag{C 9}$$

Hence using (C 9) to substitute for $\partial \rho_*/\partial t$ in (C 8) it follows that, regarding Φ as function of the two variables z_* and t,

$$\frac{\partial}{\partial t} \Phi(z_*, t) = -\left(\int_{S(\rho_*)} \phi \frac{\mathrm{d}A}{|\nabla \rho|} \right) \frac{\int_{S(\rho_*)} \rho_t \, \mathrm{d}A/|\nabla \rho|}{\int_{S(\rho_*)} \mathrm{d}A/|\nabla \rho|}$$
$$+ \int_{\rho(x,t) \geq \rho_*} \phi_t \, \mathrm{d}V + \int_{S(\rho_*)} \phi \rho_t \frac{\mathrm{d}A}{|\nabla \rho|}. \tag{C 10}$$

A natural definition of the mean value of ϕ over a z_*-coordinate surface is given by

$$\langle \phi(z_*, t) \rangle_{z_*} = \frac{1}{A_d} \frac{\partial}{\partial z_*} \Phi(z_*, t) = \frac{\int_{S(z_*)} \phi \, \mathrm{d}A/|\nabla \rho|}{\int_{S(z_*)} \mathrm{d}A/|\nabla \rho|} = -\frac{1}{A_d} \frac{\partial \rho_*}{\partial z_*} \int_{S(z_*)} \phi \, \mathrm{d}A/|\nabla \rho|, \tag{C 11}$$

where the final equality follows from (C 7), noting that $\partial z_*/\partial \rho_* = \{\partial \rho_*/\partial z_*\}^{-1}$, where both partial derivatives are evaluated for constant t. (Defining a mean of one quantity over isosurfaces of a different quantity, as suggested here, is a technique often employed in atmospheric sciences (e.g. Butchart & Remsberg 1986; Nakamura 1995).) To reflect the fact that z_* is from now on going to be an independent variable, rather than ρ_*, the notation $S(\rho_*)$ is replaced by $S(z_*)$ (but recall that $\rho = \rho_*$ on the surface $S(z_*)$).

Note that this implies that $A_d^{-1} \int_{S(\rho_*)} \phi \, \mathrm{d}A/|\nabla \rho| = \langle \phi \rangle_{z_*} (\partial \langle \rho \rangle_{z_*} /\partial z_*)^{-1}$, so (C 10) is equivalent to

$$\frac{1}{A_d} \frac{\partial}{\partial t} \Phi(z_*, t) = \left(\langle \rho_t \rangle_{z_*} \langle \phi \rangle_{z_*} - \langle \phi \rho_t \rangle_{z_*} \right) (\partial \langle \rho \rangle_{z_*} /\partial z_*)^{-1} + \frac{1}{A_d} \int_{\rho(x,t) \geq \rho_*} \phi_t \, \mathrm{d}V. \tag{C 12}$$

Now differentiating with respect to z_* gives, using (C 11) as an identity that holds for any field,

$$(\langle \phi \rangle_{z_*})_t = \langle \phi_t \rangle_{z_*} + \frac{\partial}{\partial z_*} \left(\left(\langle \rho_t \rangle_{z_*} \langle \phi \rangle_{z_*} - \langle \phi \rho_t \rangle_{z_*} \right) \left(\frac{\partial \langle \rho \rangle_{z_*}}{\partial z_*} \right)^{-1} \right). \tag{C 13}$$

C.2. *The standard case: $\phi = \rho$*

This is the case considered by Nakamura (1996) and Winters & D'Asaro (1996). Note that (C 11) implies the identity

$$\langle f(\rho)\psi \rangle_{z_*} = f(\langle \rho \rangle_{z_*}) \langle \psi \rangle_{z_*} = f(\rho) \langle \psi \rangle_{z_*} \qquad (C\,14)$$

for any function $f(\cdot)$ and for any quantity ψ. In particular $\langle \rho_t \rangle_{z_*} \langle \rho \rangle_{z_*} = \langle \rho\rho_t \rangle_{z_*}$ and hence from (C 13)

$$(\langle \rho \rangle_{z_*})_t = \langle \rho_t \rangle_{z_*} . \qquad (C\,15)$$

Because $\langle \rho \rangle_{z_*}$ and ρ are the same, and ρ_* is the notation used for the value of ρ on the ρ-surface which is labelled by z_*, ρ_* will be used instead of $\langle \rho \rangle_{z_*}$.

The standard result on effective diffusivity follows if ρ is the solution of an advection–diffusion equation, with incompressible flow field $\boldsymbol{u}(\boldsymbol{x}, t)$ and diffusivity κ_ρ. Then from the above results,

$$\partial\rho_*/\partial t = \frac{1}{A_d} \frac{\partial}{\partial z_*} \left(\int_{\rho(\boldsymbol{x},t) \geq \rho_*} (\kappa_\rho \nabla^2 \rho - \boldsymbol{u} \cdot \nabla \rho) \, \mathrm{d}V \right) = \frac{1}{A_d} \frac{\partial}{\partial z_*} \left(\int_{S(z_*)} \kappa_\rho |\nabla \rho| \, \mathrm{d}A \right),$$
$$(C\,16)$$

where the last equality follows by first using the divergence theorem to transform the volume integral to a surface integral. Then, using the fact that for the term including $\boldsymbol{u}$ the factor ρ is constant on the surface $S(z_*)$ and can be taken outside the integral, the incompressibility condition can be applied so that the integral vanishes. The final step is to use (C 7) to introduce a factor $\partial\rho_*/\partial z_*$, giving

$$(\rho_*)_t = \frac{\partial}{\partial z_*} \left(\frac{1}{A_d^2} \left[\int_{S(z_*)} \kappa_\rho |\nabla \rho| \, \mathrm{d}A \right] \left[\int_{S(z_*)} \frac{\mathrm{d}A}{|\nabla \rho|} \right] \frac{\partial\rho_*}{\partial z_*} \right) = \frac{\partial}{\partial z_*} \left(K_\rho \frac{\partial\rho_*}{\partial z_*} \right), \qquad (C\,17)$$

where the effective diffusivity K_ρ is defined by the second equality and hence, via (C 7) and (C 11), is given by

$$K_\rho = \kappa_\rho \left\langle |\nabla \rho|^2 \right\rangle_{z_*} (\partial z_*/\partial\rho_*)^2. \qquad (C\,18)$$

C.3. *The non-standard case: $\phi \neq \rho$*

The question is how to organise (C 13) in this case. The key difference is that ϕ varies on the coordinate surfaces defined by z_*. Therefore define $\phi' = \phi - \langle \phi \rangle_{z_*}$ to capture this variation. Assume that ϕ satisfies an advection–diffusion equation, i.e. $\phi_t + \boldsymbol{u} \cdot \nabla \phi = \kappa_\phi \nabla^2 \phi$, allowing the diffusivity of ϕ to be different from κ_ρ, and also either a constant or a function of time. Consider

$$\langle \phi_t \rangle_{z_*} = \frac{1}{A_d} \frac{\partial}{\partial z_*} \left(\int_{S(z_*)} [\kappa_\phi \nabla \phi - \phi\boldsymbol{u}] \cdot \mathrm{d}\boldsymbol{A} \right)$$
$$= \frac{1}{A_d} \frac{\partial}{\partial z_*} \left(\int_{S(z_*)} [\kappa_\phi \nabla \phi' - \phi'\boldsymbol{u}] \cdot \mathrm{d}\boldsymbol{A} + \int_{S(z_*)} \kappa_\phi \nabla \langle \phi \rangle_{z_*} \cdot \mathrm{d}\boldsymbol{A} \right). \qquad (C\,19)$$

The term inside the brackets on the right-hand side of the first equality comes from writing the volume integral of $\kappa_\phi \nabla^2 \phi - \boldsymbol{u} \cdot \nabla \phi$ as a surface integral using the divergence theorem. In the second equality, the term containing $\boldsymbol{u} \langle \phi \rangle_{z_*}$ is zero by the same reasoning as applied in the previous section. In the second term in brackets on the right-hand side the quantity $\nabla \langle \phi \rangle_{z_*}$ can be rewritten as $(\partial \langle \phi \rangle_{z_*} /\partial\rho_*)\nabla \rho$ and then, again, using the same reasoning as

in the previous section, the entire term can be written in terms of K_ρ. Inserting in (C 13) implies that

$$(\langle\phi\rangle_{z_*})_t = \frac{\partial}{\partial z_*}\left(\frac{\kappa_\phi}{\kappa_\rho}K_\rho\frac{\partial\langle\phi\rangle_{z_*}}{\partial z_*}\right) + \frac{\partial}{\partial z_*}\left(-\langle\phi'\rho_t\rangle_{z_*}\frac{\partial z_*}{\partial\rho_*} + \frac{1}{A_d}\int_{S(z_*)}[\kappa_\phi\nabla\phi' - \phi'\boldsymbol{u}]\cdot\mathrm{d}\boldsymbol{A}\right).$$

(C 20)

Now consider the two quantities in the bracket in the second term on the right-hand side, which can be combined as a surface integral, using (C 7) and (C 11), plus the fact that $\mathrm{d}\boldsymbol{A} = \nabla\rho\,\mathrm{d}A/|\nabla\rho|$, to give

$$-\langle\phi'\rho_t\rangle_{z_*}\frac{\partial z_*}{\partial\rho_*} + \frac{1}{A_d}\int_{S(z_*)}[\kappa_\phi\nabla\phi' - \phi'\boldsymbol{u}]\cdot\mathrm{d}\boldsymbol{A}$$

$$= \frac{1}{A_d}\int_{S(z_*)}\frac{-\kappa_\rho\phi'\nabla^2\rho + \phi'\boldsymbol{u}\cdot\nabla\rho + \kappa_\phi\nabla\phi'\cdot\nabla\rho - \phi'\boldsymbol{u}\cdot\nabla\rho}{|\nabla\rho|}\mathrm{d}A$$

$$= \frac{\partial z_*}{\partial\rho_*}\langle(\kappa_\phi\nabla\phi'\cdot\nabla\rho - \kappa_\rho\phi'\nabla^2\rho)\rangle_{z_*}.$$

(C 21)

Hence, substituting into (C 19),

$$(\langle\phi\rangle_{z_*})_t = \frac{\partial}{\partial z_*}\left(\frac{\kappa_\phi}{\kappa_\rho}K_\rho\frac{\partial\langle\phi\rangle_{z_*}}{\partial z_*}\right) + \frac{\partial}{\partial z_*}\left(\frac{\partial z_*}{\partial\rho_*}\langle(\kappa_\phi\nabla\phi'\cdot\nabla\rho - \kappa_\rho\phi'\nabla^2\rho)\rangle_{z_*}\right).$$

(C 22)

This is the governing equation for $\langle\phi\rangle_{z_*}$. Some checks on the form of the above equation are as follows.

Example C.1. When $\kappa_\rho = 0$ and $\kappa_\phi = 0$ both ρ and ϕ are materially conserved and therefore, since ρ-surfaces are material surfaces, there will be no change in the total amount of ϕ within each ρ-surface. Furthermore the volume within each ρ-surface remains constant. Therefore for each z_*, $\langle\phi\rangle_{z_*}$ is constant in time. Equation (C 22) is consistent with this.

Example C.2. When $\kappa_\rho \neq 0$ and $\kappa_\phi = 0$ then ϕ is materially conserved. The rate of change of the total amount of ϕ within the volume z_* is therefore given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{z_*}\phi\,\mathrm{d}V = \int_{S(z_*)}\phi(\boldsymbol{u}_S - \boldsymbol{u})\cdot\mathrm{d}\boldsymbol{A}$$

$$= \int_{S(z_*)}\phi(\boldsymbol{u}_S - \boldsymbol{u})\cdot\nabla\rho\frac{\mathrm{d}A}{|\nabla\rho|}$$

$$= \langle\phi(\boldsymbol{u}_S - \boldsymbol{u})\cdot\nabla\rho\rangle_{z_*}(\partial z_*/\partial\rho_*),$$

(C 23)

where $\boldsymbol{u}_S$ is the local normal velocity of the surface $S(z_*)$. Since $S(z_*)$ is a surface of constant ρ (but the value of ρ on the surface varies in time), on $S(z_*)$,

$$\rho_t + \boldsymbol{u}_S\cdot\nabla\rho = \kappa_\rho\nabla^2\rho + (\boldsymbol{u}_S - \boldsymbol{u})\cdot\nabla\rho = \alpha(t),$$

(C 24)

where $\alpha(t)$ is to be determined by the condition that the volume within the surface $S(z_*)$ stays constant, requiring that

$$\int_{S(z_*)}\boldsymbol{u}_S\cdot\mathrm{d}\boldsymbol{A} = \int_{S(z_*)}\boldsymbol{u}_S\cdot\nabla\rho\frac{\mathrm{d}A}{|\nabla\rho|} = \langle\boldsymbol{u}_S\cdot\nabla\rho\rangle_{z_*}(\partial z_*/\partial\rho_*) = 0.$$

(C 25)

Incompressibility implies (e.g. replace $\boldsymbol{u}_S$ by $\boldsymbol{u}$ in the preceding equation) that $\langle \boldsymbol{u} \cdot \nabla \rho \rangle_{z_*} = 0$. Therefore applying $\langle \cdot \rangle_{z_*}$ to (C 24) implies that

$$\alpha(t) = \langle \kappa_\rho \nabla^2 \rho \rangle_{z_*} \tag{C 26}$$

and it follows also from (C 24) that $(\boldsymbol{u}_S - \boldsymbol{u}) \cdot \nabla \rho = \langle \kappa \nabla^2 \rho \rangle_{z_*} - \kappa \nabla^2 \rho$. Additionally $\langle (\boldsymbol{u}_S - \boldsymbol{u}) \cdot \nabla \rho \rangle_{z_*} = 0$ implies that

$$\langle \phi (\boldsymbol{u}_S - \boldsymbol{u}) \cdot \nabla \rho \rangle_{z_*} = \langle \phi' (\boldsymbol{u}_S - \boldsymbol{u}) \cdot \nabla \rho \rangle_{z_*} = \langle \phi' (\langle \kappa_\rho \nabla^2 \rho \rangle_{z_*} - \kappa_\rho \nabla^2 \rho) \rangle_{z_*} = - \langle \phi' \kappa_\rho \nabla^2 \rho \rangle_{z_*}. \tag{C 27}$$

It follows from (C 23) that

$$((\langle \phi \rangle_{z_*})_t = \frac{\partial}{\partial z_*} \left(- \langle \phi' \kappa_\rho \nabla^2 \rho \rangle_{z_*} (\partial z_* / \partial \rho_*) \right), \tag{C 28}$$

consistent with (C 22) when $\kappa_\phi = 0$.

Example C.3. When (formally) $\kappa_\rho = 0$ but $\kappa_\phi \neq 0$, then as in Example 1 ρ-surfaces are material surfaces, and there is no advective flux of ϕ across ρ-surfaces. Therefore the rate of change of the total amount of ϕ within a ρ-surface is equal to the diffusive flux across the surface.

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{z_*} \phi \, \mathrm{d}V = \int_{S(z_*)} \kappa_\phi \nabla \phi \cdot \mathrm{d}\boldsymbol{A}. \tag{C 29}$$

Now write $\phi = \langle \phi(\rho) \rangle_{z_*} + \phi'$ and hence $\nabla \phi = (\partial \langle \phi \rangle_{z_*})/(\partial \rho) \nabla \rho + \nabla \phi'$. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{z_*} \phi \, \mathrm{d}V = \kappa_\phi \left(\frac{\partial \langle \phi \rangle_{z_*}}{\partial \rho} \int_{S(z_*)} \nabla \rho \cdot \mathrm{d}\boldsymbol{A} + \int_{S(z_*)} \nabla \phi' \cdot \mathrm{d}\boldsymbol{A} \right)$$

$$= A_d \kappa_\phi \left(\frac{\partial \langle \phi \rangle_{z_*}}{\partial \rho} \langle |\nabla \rho|^2 \rangle_{z_*} + \langle \nabla \phi' \cdot \nabla \rho \rangle_{z_*} \right) \frac{\partial z_*}{\partial \rho}. \tag{C 30}$$

Finally, differentiating with respect to z_*, dividing by A_d and rearranging the right-hand side,

$$(\langle \phi \rangle_{z_*})_t = \frac{\partial}{\partial z_*} \left(\kappa_\phi \langle |\nabla \rho|^2 \rangle_{z_*} \left(\frac{\partial z_*}{\partial \rho_*} \right)^2 \frac{\partial \langle \phi \rangle_{z_*}}{\partial z_*} + \langle \nabla \phi' \cdot \nabla \rho \rangle_{z_*} \frac{\partial z_*}{\partial \rho_*} \right)$$

$$= \frac{\partial}{\partial z_*} \left(\frac{\kappa_\phi}{\kappa_\rho} K_\rho \frac{\partial \langle \phi \rangle_{z_*}}{\partial z_*} \right) + \frac{\partial}{\partial z_*} \left(\kappa_\phi \langle \nabla \phi' \cdot \nabla \rho \rangle_{z_*} \frac{\partial z_*}{\partial \rho} \right), \tag{C 31}$$

again consistent with (C 22) in the relevant limit.

REFERENCES

ALEXAKIS, A. 2009 Stratified shear flow instabilities at large Richardson numbers. *Phys. Fluids* **21** (5), 054108.

AUCLAIR, F., BORDOIS, L., DOSSMANN, Y., DUHAUT, T., PACI, A., ULSES, C. & NGUYEN, C. 2018 A non-hydrostatic non-Boussinesq algorithm for free-surface ocean modelling. *Ocean Model.* **132**, 12–29.

BRIERLEY, A. S. & KINGSFORD, M. J. 2009 Impacts of climate change on marine organisms and ecosystems. *Curr. Biol.* **19** (14), R602–R614.

BROWNING, K. A. & WATKINS, C. D. 1970 Observations of clear air turbulence by high power radar. *Nature* **227** (5255), 260–263.

BUTCHART, N. & REMSBERG, E. E. 1986 The area of the stratospheric polar vortex as a diagnostic for tracer transport on an isentropic surface. *J. Atmos. Sci.* **43** (13), 1319–1339.

CANUTO, V. M., CHENG, Y. & HOWARD, A. M. 2011 Vertical diffusivities of active and passive tracers. *Ocean Model.* **36** (3), 198–207.

CARPENTER, J. R., BALMFORTH, N. J. & LAWRENCE, G. A. 2010 Identifying unstable modes in stratified shear layers. *Phys. Fluids* **22** (5), 054104.

CAULFIELD, C. P. & PELTIER, W. R. 1994 Three dimensionalization of the stratified mixing layer. *Phys. Fluids* **6** (12), 3803–3805.

CAULFIELD, C. P., YOSHIDA, S. & PELTIER, W. R. 1996 Secondary instability and three-dimensionalization in a laboratory accelerating shear layer with varying density differences. *Dyn. Atmos. Oceans* **23** (1), 125–138.

DEBREU, L., MARCHESIELLO, P., PENVEN, P. & CAMBON, G. 2012 Two-way nesting in split-explicit ocean models: algorithms, implementation and validation. *Ocean Model.* **49-50**, 1–21.

FRITTS, D. C. & RASTOGI, P. K. 1985 Convective and dynamical instabilities due to gravity wave motions in the lower and middle atmosphere: theory and observations. *Radio Sci.* **20** (6), 1247–1277.

FUKAO, S., YAMANAKA, M. D., AO, N., HOCKING, W. K., SATO, T., YAMAMOTO, M., NAKAMURA, T., TSUDA, T. & KATO, S. 1994 Seasonal variability of vertical eddy diffusivity in the middle atmosphere: Part 1. Three-year observations by the middle and upper atmosphere radar. *J. Geophys. Res.* **99** (D9), 18973–18987.

GEYER, W. R., LAVERY, A. C., SCULLY, M. E. & TROWBRIDGE, J. H. 2010 Mixing by shear instability at high Reynolds number. *Geophys. Res. Lett.* **37** (22), L22607.

GEYER, W. R. & SMITH, J. D. 1987 Shear instability in a highly stratified estuary. *J. Phys. Oceanogr.* **17** (10), 1668–1679.

GRIFFIES, S. M. 2004 *Fundamentals of Ocean Climate Models*. Princeton University Press.

VAN HAREN, H., GOSTIAUX, L. 2010 A deep-ocean Kelvin–Helmholtz billow train. *Geophys. Res. Lett.* **37** (3), L03605.

JANJIC, Z. I., GERRITY, J. P. & NICKOVIC, S. 2001 An alternative approach to nonhydrostatic modeling. *Mon. Weath. Rev.* **129** (5), 1164–1178.

JIANG, N. 2019 A pressure-correction ensemble scheme for computing evolutionary Boussinesq equations. *J. Sci. Comput.* **80** (1), 315–350.

KLAASSEN, G. P. & PELTIER, W. R. 1985 The onset of turbulence in finite-amplitude Kelvin–Helmholtz billows. *J. Fluid Mech.* **155**, 1–35.

KLAASSEN, G. P. & PELTIER, W. R. 1989 The role of transverse secondary instabilities in the evolution of free shear layers. *J. Fluid Mech.* **202**, 367–402.

LAURITZEN, P. H. & THUBURN, J. 2012 Evaluating advection/transport schemes using interrelated tracers, scatter plots and numerical mixing diagnostics. *Q. J. R. Meteorol. Soc.* **138** (665), 906–918.

LEDWELL, J. R., WATSON, A. J. & LAW, C. S. 1993 Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature* **364** (6439), 701.

LI, D., BOU-ZEID, E. & DE BRUIN, H. A. R. 2012 Monin–Obukhov similarity functions for the structure parameters of temperature and humidity. *Boundary-Layer Meteorol.* **145** (1), 45–67.

LINCOLN, B. J., RIPPETH, T. P. & SIMPSON, J. H. 2016 Surface mixed layer deepening through wind shear alignment in a seasonally stratified shallow sea. *J. Geophys. Res.* **121** (8), 6021–6034.

LODER, T. C. & REICHARD, R. P. 1981 The dynamics of conservative mixing in estuaries. *Estuaries* **4** (1), 64–69.

MARMORINO, G. O. 1987 Observations of small-scale mixing processes in the seasonal thermocline, part II: Wave breaking. *J. Phys. Oceanogr.* **17**, 1348–1355.

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

NAGATA, K. & KOMORI, S. 2001 The difference in turbulent diffusion between active and passive scalars in stable thermal stratification. *J. Fluid Mech.* **430**, 361–380.

NAKAMURA, N. 1995 Modified Lagrangian-mean diagnostics of the stratospheric polar vortices. Part 1. Formulation and analysis of GFDL SKYHI GCM. *J. Atmos. Sci.* **52** (11), 2096–2108.

NAKAMURA, N. 1996 Two-dimensional mixing, edge formation, and permeability diagnosed in an area coordinate. *J. Atmos. Sci.* **53** (11), 1524–1537.

OFFICER, C. B. & LYNCH, D. R. 1981 Dynamics of mixing in estuaries. *Estuar. Coast. Shelf Sci.* **12** (5), 525–533.

PATTERSON, M. D., CAULFIELD, C. P., MCELWAINE, J. N. & DALZIEL, S. B. 2006 Time-dependent mixing in stratified Kelvin–Helmholtz billows: experimental observations. *Geophys. Res. Lett.* **33** (15), L15608.

PENNEY, J. & STASTNA, M. 2016 Direct numerical simulation of double-diffusive gravity currents. *Phys. Fluids* **28** (8), 086602.

PLUMB, R. A. 2007 Tracer interrelationships in the stratosphere. *Rev. Geophys.* **45** (4), RG4005.

RHINES, P. B. & YOUNG, W. R. 1983 How rapidly is a passive scalar mixed within closed streamlines? *J. Fluid Mech.* **133**, 133–145.

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

SALEHIPOUR, H., PELTIER, W. R. & MASHAYEK, A. 2015 Turbulent diapycnal mixing in stratified shear flows: the influence of Prandtl number on mixing efficiency and transition at high Reynolds number. *J. Fluid Mech.* **773**, 178–223.

SCINOCCA, J. F. 1995 The mixing of mass and momentum by Kelvin–Helmboltz billows. *J. Atmos. Sci.* **52** (14), 2509–2530.

SEINFELD, J. H. & PANDIS, S. N. 1998 *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. John Wiley & Sons.

SHCHEPETKIN, A. F. & MCWILLIAMS, J. C. 2005 The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* **9** (4), 347–404.

SHETE, M. H. & GUHA, A. 2018 Effect of free surface on submerged stratified shear instabilities. *J. Fluid Mech.* **843**, 98–125.

SHU, C.-W. 1999 High order ENO and WENO schemes for computational fluid dynamics. In *High-Order Methods for Computational Physics* (ed. T. J. Barth & H. Deconinck), pp. 439–582. Springer.

SHUCKBURGH, E. & HAYNES, P. 2003 Diagnosing transport and mixing using a tracer-based coordinate system. *Phys. Fluids* **15** (11), 3342–3357.

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

SMYTH, W. D., NASH, J. D. & MOUM, J. N. 2005 Differential diffusion in breaking Kelvin–Helmholtz billows. *J. Phys. Oceanogr.* **35** (6), 1004–1022.

SPARLING, L. C., KETTLEBOROUGH, J. A., HAYNES, P. H., MCINTYRE, M. E., ROSENFIELD, J. E., SCHOEBERL, M. R. & NEWMAN, P. A. 1997 Diabatic cross-isentropic dispersion in the lower stratosphere. *J. Geophys. Res.* **102** (D22), 25817–25829.

SUBICH, C. J., LAMB, K. G. & STASTNA, M. 2013 Simulation of the Navier–Stokes equations in three dimensions with a spectral collocation method. *Intl J. Numer. Meth. Fluids* **73** (2), 103–129.

TAILLEUX, R. 2009 On the energetics of stratified turbulent mixing, irreversible thermodynamics, Boussinesq models and the ocean heat engine controversy. *J. Fluid Mech.* **638**, 339–382.

TERAMOTO, T. 1993 *Deep Ocean Circulation: Physical and Chemical Aspects*. Elsevier.

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

TILMES, S., MÜLLER, R., GROOSS, J.-U., NAKAJIMA, H. & SASANO, Y. 2006 Development of tracer relations and chemical ozone loss during the setup phase of the polar vortex. *J. Geophys. Res.* **111** (D24), D24S90.

TOMCZAK, M. 1981 A multi-parameter extension of temperature/salinity diagram techniques for the analysis of non-isopycnal mixing. *Prog. Oceanogr.* **10** (3), 147–171.

VAQUER-SUNYER, R. & DUARTE, C. M. 2008 Thresholds of hypoxia for marine biodiversity. *Proc. Natl Acad. Sci.* **105** (40), 15452–15457.

WARHAFT, Z. 2000 Passive scalars in turbulent flows. *Annu. Rev. Fluid Mech.* **32**, 203–240.

WINTERS, K. B. & D'ASARO, E. A. 1996 Diascalar flux and the rate of fluid mixing. *J. Fluid Mech.* **317**, 179–193.

WOODS, J. D. 1968 Wave-induced shear instability in the summer thermocline. *J. Fluid Mech.* **32** (4), 791–800.

WUNSCH, C. & FERRARI, R. 2004 Vertical mixing, energy, and the general circulation of the oceans. *Annu. Rev. Fluid Mech.* **36** (1), 281–314.

YANG, Q., EASTER, R. C., CAMPUZANO-JOST, P., JIMENEZ, J. L., FAST, J. D., GHAN, S. J., WANG, H., BERG, L. K., BARTH, M. C., LIU, Y., et al. 2015 Aerosol transport and wet scavenging in deep convective clouds: a case study and model evaluation using a multiple passive tracer analysis approach. *J. Geophys. Res.* **120** (16), 8448–8468.

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