J. Fluid Mech. (2015), *vol.* 775, *pp.* 464–500.
doi:10.1017/jfm.2015.305

Diapycnal diffusivity, turbulent Prandtl number and mixing efficiency in Boussinesq stratified turbulence

Hesam Salehipour[1,†] **and W. R. Peltier**[1]

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

(Received 19 December 2014; revised 23 April 2015; accepted 24 May 2015; first published online 26 June 2015)

In order that it be correctly characterized, irreversible turbulent mixing in stratified fluids must distinguish between adiabatic 'stirring' and diabatic 'mixing'. Such a distinction has been formalized through the definition of a diapycnal diffusivity, K_ρ (Winters & D'Asaro, *J. Fluid Mech.*, vol. 317, 1996, pp. 179–193) and an appropriate mixing efficiency, $\mathscr{E}$ (Caulfield & Peltier, *J. Fluid Mech.*, vol. 413, 2000, pp. 1–47). Equivalent attention has not been paid to the definitions of a corresponding momentum diffusivity K_m and hence an appropriately defined turbulent Prandtl number $Pr_t = K_m/K_\rho$. In this paper, the diascalar framework of Winters & D'Asaro (1996) is first reformulated to obtain an 'Osborn-like' formula in which the correct definition of irreversible mixing efficiency $\mathscr{E}$ is shown to replace the flux Richardson number which Osborn (*J. Phys. Oceanogr.*, vol. 10, 1980, pp. 83–89) assumed to characterize this efficiency. We advocate the use of this revised representation for diapycnal diffusivity since the proposed reformulation effectively removes the simplifying assumptions on which the original Osborn formula was based. We similarly propose correspondingly reasonable definitions for K_m and Pr_t by eliminating the reversible component of the momentum production term. To explore implications of the reformulations for both diapycnal and momentum diffusivity we employ an extensive series of direct numerical simulations (DNS) to investigate the properties of the shear-induced density-stratified turbulence that is engendered through the breaking of a freely evolving Kelvin–Helmholtz wave. The DNS results based on the proposed reformulation of K_ρ are compared with available estimations due to the mixing length model, as well as both the Osborn–Cox and the Osborn models. Estimates based upon the Osborn–Cox formulation are shown to provide the closest approximation to the diapycnal diffusivity delivered by the exact representation. Through compilation of the complete set of DNS results we explore the characteristic dependence of K_ρ on the buoyancy Reynolds number Re_b as originally investigated by Shih *et al.* (*J. Fluid Mech.*, vol. 525, 2005, pp. 193–214) in their idealized study of homogeneous stratified and sheared turbulence, and show that the validity of their results is only further reinforced through analysis of the turbulence produced in the more geophysically relevant Kelvin–Helmholtz wave life-cycle ansatz. In contrast to the results described by Shih *et al.* (2005) however, we show that, besides Re_b, a vertically averaged measure of the gradient Richardson number Ri_b may equivalently characterize the turbulent mixing at high Re_b. Based on the dominant driving processes involved in irreversible mixing, we categorize the intermediate (i.e. $Re_b = O(10^1–10^2)$) and high

† Email address for correspondence: h.salehipour@utoronto.ca

(i.e. $Re_b > O(10^2)$) range of Re_b as 'buoyancy-dominated' and 'shear-dominated' mixing regimes, which together define a transition value of $Ri_b \sim 0.2$. Mixing efficiency varies non-monotonically with both Re_b and Ri_b, with its maximum (on the order of 0.2–0.3) occurring in the 'buoyancy-dominated' regime. Unlike K_ρ which is very sensitive to the correct choice of $\mathscr{E}$ (i.e. $K_\rho \propto \mathscr{E}/(1-\mathscr{E})$), we show that K_m is almost insensitive to the choice of $\mathscr{E}$ (i.e. $K_m \propto 1/(1-\mathscr{E})$) so long as $\mathscr{E}$ is not close to unity, which implies $K_m \approx Ri_b Re_b$ for the entire range of Re_b. The turbulent Prandtl number is consequently shown to decrease monotonically with Re_b and may be (to first order) simply approximated by Re_b itself. Assuming $Pr_t = 1$, or $Pr_t = 10$ (as is common in large-scale numerical models of the ocean general circulation), is also suggested to be a questionable assumption.

Key words: mixing and dispersion, ocean circulation, stratified turbulence

1. Introduction

The idea of a turbulent diapycnal diffusivity, K_ρ, is often introduced to represent the irreversible flux of mass across isopycnal surfaces in a density-stratified fluid which experiences irreversible turbulent mixing. It may be defined most generally as the ratio of a scalar flux to its vertical gradient as:

$$K_\rho = -\frac{\text{flux}}{\text{gradient}}. \tag{1.1}$$

As noted by Winters & D'Asaro (1996) and Barry *et al.* (2001), the fundamental difference between models of K_ρ arises due to the different definitions adopted for both the scalar flux and its associated gradient in (1.1). At present, the representations proposed by Osborn & Cox (1972) and Osborn (1980) are still widely employed in the parameterization of small-scale diapycnal mixing in large-scale numerical models of the ocean general circulation (e.g. Danabasoglu *et al.* 2012) and also to infer K_ρ from oceanographic experimental measurements (e.g. Waterhouse *et al.* 2014). Given the approximate nature of these models (Davis 1994; Mashayek & Peltier 2013) and the fact that they do not distinguish between reversible 'stirring' and irreversible 'mixing' (as discussed in Winters *et al.* 1995 and Caulfield & Peltier 2000), an accurate definition for K_ρ using an isoscalar coordinate system was proposed by Winters & D'Asaro (1996) (hereinafter referred to as WD96). One goal of the present work is to analytically recast their definition of K_ρ into a more familiar 'Osborn-like' expression and to clarify its dependence on an accurately defined irreversible mixing efficiency, $\mathscr{E}$, and buoyancy Reynolds number Re_b.

Mixing efficiency, $\mathscr{E}$, may be defined as the 'useful' expenditure of total mechanical energy that is invested in irreversible mixing (Caulfield & Peltier 2000), which for a Boussinesq flow leads to a monotonic increase of a background potential energy (BPE) of the system (Winters *et al.* 1995; Tailleux 2009) that is to be distinguished from the available potential energy that can be converted into kinetic energy (Lorenz 1955). It has become common practice to assume a canonical value of 0.2 for an assumed constant value of mixing efficiency, through which the diapycnal diffusivity is estimated based on the approximate formulation of Osborn (1980). However, there is continuing debate on the relevance of this value to the full range of accessible

turbulent regimes in geophysical flows (Ivey, Winters & Koseff 2008). This issue has therefore prompted the development of alternative expressions for scalar diffusivity based on a variety of length scale arguments (e.g. see Ivey & Imberger 1991; Barry *et al.* 2001; Shih *et al.* 2005) that may be employed to interpret the measurements of turbulence related processes in environmental flows (Dunckley *et al.* 2012; Bluteau, Jones & Ivey 2013; Bouffard & Boegman 2013). The modest reformulation of K_ρ, to be presented in what follows, is a direct mathematical consequence of the governing Boussinesq equations and will lead us to advocate the use of an 'Osborn-like' formula with no limiting approximations, provided that a correct definition of $\mathscr{E}$ is employed. Moreover, this reformulation will be shown to reinforce the importance of the previously presented length scale arguments suggesting that K_ρ should depend strongly upon Re_b.

The buoyancy Reynolds number, Re_b (or as it is also sometimes referred to, the turbulence intensity parameter), measures the extent of the inertial subrange that is characterized by the ratio of the Ozmidov to the Kolmogorov length scales (to the power 4/3) (e.g. see Smyth & Moum 2000*b*). This dimensionless parameter has been suggested to be one of the two dynamically relevant parameters required for the characterization of stratified turbulence (e.g. see Brethouwer *et al.* 2007 and Ivey *et al.* 2008). The second parameter required for such characterization may be thought to be the gradient Richardson number Ri_b (Linden 1979; Fernando 1991) which characterizes the relative strengths of density stratification and velocity shear. These two parameters are respectively defined as:

$$Re_b = \frac{\varepsilon_k}{\nu N^2}, \quad Ri_b = \frac{N^2}{S^2}, \tag{1.2a,b}$$

where ε_k is the viscous dissipation of kinetic energy (to be precisely defined in what follows), ν is kinematic viscosity, N is the Brunt–Väisälä frequency and S is the vertical shear. Here we employ a bulk measure of the gradient Richardson number in which N^2 and S^2 are individually vertically averaged over the same appropriate length scale.

In addition to K_ρ, the momentum diffusivity (or eddy viscosity), K_m, is also required to characterize the subgrid-scale diffusion of momentum and it too must therefore be parameterized in large-scale numerical models of geophysical flows (Klymak, Legg & Pinkel 2010; Danabasoglu *et al.* 2012). A key parameter which measures the relative importance of these two turbulent diffusivities is the turbulent Prandtl number $Pr_t = K_m/K_\rho$. Venayagamoorthy & Stretch (2010) have recently studied the variations of Pr_t in the special case of homogeneous sheared and stratified turbulence in which they distinguished reversible from irreversible contributions to K_m by drawing an analogy to a formulation in the stationary turbulence limit (based on simplified energy and scalar balance equations) and proposing a new formulation for K_m for non-stationary homogeneous stratified and sheared turbulent flows. A second primary goal in this paper is to propose an expression for Pr_t applicable to any stratified Boussinesq flow at sufficiently high Re.

Any useful parameterization of K_ρ and K_m (or Pr_t) should rely on a study which not only isolates the signature of irreversible mixing on turbulent diffusivity, but also properly identifies the aforementioned parametric dependences. In this regard, direct numerical simulation (DNS) provides the ideal tool to analyse stratified turbulent mixing events. Nonetheless, previous DNS efforts (e.g. Smyth, Moum & Caldwell 2001; Shih *et al.* 2005; Smyth, Nash & Moum 2005; Mashayek & Peltier 2013) have

not consistently employed adequately precise definitions of these important quantities, namely $\mathscr{E}$, K_ρ and K_m (or Pr_t). Here, by use of the phrase 'adequately precise' we mean definitions that fully distinguish between the reversible and irreversible contributions to the scalar and momentum flux related processes. Furthermore, aside from the study by Shih *et al.* (2005) (hereinafter referred to as SKIF), who focused on the special case of homogeneous sheared and stratified turbulence, to the best of our knowledge there has been no other DNS study aimed at characterizing the dependence of $\mathscr{E}$, K_ρ and K_m on the dynamically relevant parameters of stratified turbulence (e.g. Re_b and Ri_b). In this paper, our intention is to begin to address both these aspects of the stratified turbulence problem through the analysis of an extensive series of three-dimensional DNS dataset of the much more geophysically relevant case of inhomogeneously stratified and sheared turbulence that is generated after a laminar two-dimensional Kelvin–Helmholtz (KH) wave 'breaks' at high initial Reynolds numbers (to be denoted henceforth as the KH ansatz). Nonetheless, we foresee a significant need for further numerical and experimental effort focusing on stratified turbulence with a wider range of turbulent generation mechanisms in order to fully establish the parametric dependences of these critical characteristics of stratified turbulence.

The remainder of this paper consists of two distinct parts. In the first part in §§ 2 and 3, we first review the available models that continue to be employed in the estimation of K_ρ and K_m, following which we propose improved expressions for the diapycnal diffusivity K_ρ and eddy viscosity K_m (and hence Pr_t) in stratified Boussinesq turbulence. Arguments presented in § 2.3 essentially transform the problem of parameterizing diapycnal diffusivity into the problem of correctly determining irreversible mixing efficiency. The second part of this paper, in § 4, describes our numerical methodology and provides the results of an extensive series of DNS analyses of turbulence characteristics based on the KH ansatz. While the first part of the paper relies solely upon the structure of the underlying Boussinesq equations (and is therefore independent of flow configuration), the second part is specific to our choice of model problem. Concluding remarks that follow from the analyses we have performed are offered in § 5.

2. Scalar diffusivity and mixing efficiency

Our primary goal is to moderately build upon and to extend the previous work of WD96 in order to derive a revised expression for diapycnal diffusivity. This reformulation will lead to an expression of Osborn (1980) form, but one that eliminates all the assumptions on which it was originally based, which should therefore constitute a useful basis for improved parameterizations of diapycnal diffusivity. In all that follows we will be providing non-dimensional representations for various models of K_ρ by normalizing them by the molecular diffusivity, κ, i.e. $\widetilde{K}_\rho = K_\rho/\kappa$. This non-dimensionalization will prove useful for the purpose of comparing K_ρ for fluids characterized by different molecular Prandtl numbers. Because the turbulent diffusivity is inevitably several orders of magnitude greater than its molecular counterpart, this non-dimensional diffusivity would be expected to reach high values of $O(10^1$–$10^4)$, regardless of the calculation methodology.

2.1. *Existing models of scalar diffusivity*

We begin by reviewing four different methods of calculating K_ρ, some of which are most commonly employed in numerical and experimental studies. These methods

include: (i) the Prandtl mixing length model denoted by K_ρ^{ml}, (ii) the Osborn formula, K_ρ^{osb}, (iii) the Osborn–Cox formula K_ρ^{cox} and finally (iv) the Winters–D'Asaro formula, K_ρ^*.

The Prandtl mixing length model is sometimes referred to as the 'direct' method of calculating K_ρ (Ivey *et al.* 2008; Bouffard & Boegman 2013). In fact, the historical difficulties associated with measuring and interpreting the turbulent density flux $\overline{\rho' w'}$ has led to its 'indirect' inference from the simplified turbulent kinetic energy (TKE) evolution equation and the balance equation for temperature variance, respectively in the models proposed by Osborn (1980) and Osborn & Cox (1972). In DNS studies of stratified sheared turbulence (e.g. see SKIF and Mashayek & Peltier 2013), because this density flux can be directly computed, K_ρ^{ml} derived on the basis of the mixing length model is often considered to be the most accurate value for K_ρ. However, the fidelity of the mixing length model critically depends on whether the turbulent density flux can be considered the appropriate 'flux' in (1.1). In fact, it was demonstrated by WD96 that the 'flux' that should be employed in the calculation of K_ρ is the 'diapycnal flux' ϕ_d, which includes only the irreversible and diffusive destruction of small-scale density variance. Unlike the turbulent density flux, the diapycnal flux may be precisely obtained by evoking the concept of adiabatic sorting, which continuously rearranges the evolving three-dimensional instantaneous density field into a background stably stratified density profile. This enables a clear distinction to be made between the reversible and irreversible processes that are active in flow evolution. Our proposed reformulation for K_ρ to be presented in this paper also accounts for this important distinction as it is simply a variation upon the Winters–D'Asaro formulation.

2.1.1. *Prandtl mixing length model*

The concept of 'mixing length', as first introduced by Taylor (1915), is associated with an average distance in the cross-stream direction that a fluid particle travels before being mixed. The mixing length theory of Prandtl (1925) employs this concept to relate the mean flow properties to the turbulent deviations from it. Thus, by following this approach K_ρ may be estimated in the simplest manner by the ratio of the turbulent density flux to the vertical gradient of the mean density profile. Namely, in non-dimensional form:

$$\widetilde{K}_\rho^{ml} = -\frac{1}{\kappa}\frac{\langle \overline{\rho' w'} \rangle}{\langle \mathrm{d}\overline{\rho}/\mathrm{d}z \rangle}, \tag{2.1}$$

in which the deviations from the horizontal mean are denoted by a prime (ρ' or w') and the overbar represents horizontal averaging, while $\langle . \rangle$ represents vertical averaging. The choices for the 'flux' and 'gradient' terms in (1.1) should be clear for this model from (2.1). As will be reviewed in what follows, the models proposed by Osborn (1980) and Osborn & Cox (1972) are both based on the expression in (2.1), but in these models the density flux is replaced by other flow characteristics.

2.1.2. *Osborn model*

Osborn (1980) simplified the TKE equation by assuming that the turbulence is stationary in time and the flow is homogeneous in space, and on this basis derived a three-way balance between the shear production of the background flow $\mathbb{P}$, the turbulent buoyancy flux $\mathbb{B}$, and viscous dissipation ε_k as $\mathbb{P} = \mathbb{B} + \varepsilon_k$. These terms are defined respectively as:

$$\mathbb{P} = -\left\langle \frac{\mathrm{d}\overline{u}}{\mathrm{d}z}\,\overline{u'w'} \right\rangle, \tag{2.2}$$

$$\mathbb{B} = \frac{g}{\rho_0}\langle \overline{\rho' w'} \rangle, \tag{2.3}$$

$$\varepsilon_k = 2\nu \langle \overline{s_{ij} s_{ij}} \rangle, \tag{2.4}$$

where $s_{ij} = (\partial u_i/\partial x_j + \partial u_j/\partial x_i)/2$ is the strain rate tensor. Note that ε_k is a positive-definite quantity which represents the dissipation of kinetic energy at smallest scales due to the finite molecular viscosity of the fluid.

Osborn (1980) further employed the definition of the flux Richardson number as $R_f = \mathbb{B}/\mathbb{P}$, which in conjunction with the simplified balance equation resulted in $\mathbb{B} = R_f/(1 - R_f)\varepsilon_k$. In order to obtain an estimate for diapycnal diffusivity that could be inferred on the basis of oceanographic measurements of turbulent dissipation, Osborn also employed the Prandtl mixing length theory (2.1) and replaced the buoyancy flux with $R_f/(1 - R_f)\varepsilon_k$. The resulting expression for K_ρ in our non-dimensional form can be written as:

$$\widetilde{K}_\rho^{osb} = Pr \left(\frac{R_f}{1 - R_f} \right) \frac{\varepsilon_k}{\nu N^2} = \gamma Pr\, Re_b, \tag{2.5}$$

in which $Re_b = \varepsilon_k/(\nu N^2)$ is the buoyancy Reynolds number, where $N = \sqrt{g/\rho_0 \langle \mathrm{d}\overline{\rho}/\mathrm{d}z \rangle}$ denotes the vertically averaged Brunt–Väisälä frequency and $Pr = \nu/\kappa$ is the molecular Prandtl number.

For practical purposes and in order to close the above parameterization for K_ρ, Osborn (1980) further proposed $R_f \leqslant 0.17$ (or $\gamma = R_f/(1 - R_f) \leqslant 0.2$) based on the theoretical work of Ellison (1957). Conventionally thereafter, a constant value of $\gamma = 0.2$ in (2.5) has been used as the 'Osborn formula' for both parameterizing diapycnal mixing in large-scale general circulation models of the oceans (St. Laurent, Simmons & Jayne 2002; Danabasoglu *et al.* 2012) and also for the purpose of inferring K_ρ from microstructure measurements (St. Laurent *et al.* 2012; Waterhouse *et al.* 2014).

2.1.3. *Osborn–Cox model*

The other widely used expression in oceanography for inferring scalar diffusivity of temperature K_θ was proposed by Osborn & Cox (1972). Note that for temperature stratified flows, as in the ocean thermocline for example, it is usually assumed that $K_\theta = K_\rho$. This approximate model is based on a simplified balance equation for temperature variance θ'^2 which assumes stationary homogeneous turbulence with a temperature gradient only in the vertical direction (Gregg 1987). The key result from the simplified balance equation is an approximate expression for $\overline{\theta' w'}$ which is employed in Osborn & Cox (1972) as the required 'flux' term in (1.1). The final result is reproduced below in the non-dimensional form (Gregg 1987):

$$\widetilde{K}_\theta^{cox} = \frac{\langle \overline{|\nabla \theta'|^2} \rangle}{\langle [\mathrm{d}\overline{\theta}/\mathrm{d}z]^2 \rangle}. \tag{2.6}$$

It is worthwhile noting that the right-hand side of (2.6) is known as the 'Cox number'. Thus in this paper, due to our non-dimensionalization of scalar diffusivity, $\widetilde{K}_\theta^{cox}$ may also be interpreted as the Cox number.

2.1.4. *Winters–D'Asaro model*

The precise formulation for scalar diffusivity associated with the diabatic transfer of fluid parcels across any isoscalar surface (termed 'diascalar' diffusivity) was proposed by WD96, whose expression follows directly from the conservation equations under the Boussinesq approximation. Their expression for diascalar diffusivity is reproduced here in our non-dimensional form by taking the scalar field to represent the density field (i.e. diapycnal diffusivity):

$$\widetilde{K}^*_\rho = \frac{\langle |\nabla \rho|^2 \rangle_{S_*}}{[\mathrm{d}\rho/\mathrm{d}z_*]^2}, \tag{2.7}$$

in which the squared magnitude of the density gradient, $|\nabla \rho|^2$, is averaged over an adiabatically re-stratified surface S_*. Furthermore, in the reference state of minimum potential energy, that is attainable through adiabatic rearrangement of fluid parcels, the reference vertical position of an infinitesimal element at $(\boldsymbol{x}, t)$ with density $\rho(\boldsymbol{x}, t)$ is defined as $z_*(\boldsymbol{x}, t)$ (Winters *et al.* 1995). One can also define a 'diapycnal flux' ϕ_d as:

$$\phi_d(z, t) = -\kappa \frac{\langle |\nabla \rho|^2 \rangle_{S_*}}{\mathrm{d}\rho/\mathrm{d}z_*}, \tag{2.8}$$

and infer the associated diapycnal diffusivity as $\widetilde{K}^*_\rho = -\langle \phi_d \rangle / (\kappa \langle \mathrm{d}\rho/\mathrm{d}z_* \rangle)$. Using ϕ_d as the required 'flux' in (1.1), enables one to correctly distinguish between the reversible and irreversible contributions to diapycnal diffusivity, in clear contrast to all of the previous models. However, the specific form of diapycnal flux as calculated by (2.8) has not been directly employed (to the best of our knowledge) in the literature, most probably due to the complications involved in evaluating the surface integral over an isoscalar surface of the adiabatically sorted and temporally evolving density field. We will return to this formulation in the next section.

2.2. *Theoretical preliminaries*

The governing conservation equations subjected to the Boussinesq approximation can be written in dimensional form as:

$$\frac{\mathrm{D}u_i}{\mathrm{D}t} = -\frac{1}{\rho_0}\frac{\partial p}{\partial x_i} - \frac{g}{\rho_0}\rho \delta_{i3} + \nu \frac{\partial^2 u_i}{\partial x_j^2}, \tag{2.9}$$

$$\frac{\partial u_i}{\partial x_i} = 0, \tag{2.10}$$

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = \kappa \frac{\partial^2 \rho}{\partial x_j^2}, \tag{2.11}$$

in which repeated indices imply summation ($i, j = 1, 2, 3$), g is gravitational acceleration (directed downward in the vertical direction $x_3 \equiv z$) and ρ_0 is a constant reference density.

The state of minimum potential energy is obtained after adiabatically 'sorting' the three-dimensional density field into a vertical profile with a decreasing upward density (i.e. the background density profile $\rho_*(z, t)$). The BPE, $\mathscr{P}_B$, corresponding to this reference state may be defined by employing either of the following forms (Winters *et al.* 1995; Caulfield & Peltier 2000):

$$\mathscr{P}_B = \frac{g}{\rho_0}\overline{\langle \rho(\boldsymbol{x}, t) z_*(\boldsymbol{x}, t) \rangle} = \frac{g}{\rho_0}\langle \rho_*(z, t) z \rangle, \tag{2.12}$$

in which again $\langle f \rangle$ and $\bar{f}$ respectively denote vertical and horizontal averaging of a general field f. Also note that $\mathscr{P}_B$ denotes the averaged BPE per unit mass and is therefore normalized by ρ_0.

Winters *et al.* (1995) derived $\mathrm{d}\mathscr{P}_B/\mathrm{d}t$ by taking the time derivative of (2.12) and further by replacing $\partial\rho/\partial t$ from (2.11) (note that the volume average of the other resulting term including $\partial z_*/\partial t$ is identically zero) to write:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{P}_B = \frac{g}{\rho_0 V}\int_V z_*(-\boldsymbol{u}\cdot\boldsymbol{\nabla}\rho + \kappa\nabla^2\rho)\,\mathrm{d}V. \tag{2.13}$$

Furthermore, by applying integration by parts to each of the terms inside the volume integral, the resulting evolution equation for BPE was shown to be reduced to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{P}_B = \mathscr{F}_{adv} + \mathscr{F}_{diff} + \mathscr{M} + D_p, \tag{2.14}$$

$$\mathscr{F}_{adv} = -\frac{g}{\rho_0 V}\oint_S \psi\boldsymbol{u}\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S, \tag{2.15}$$

$$\mathscr{F}_{diff} = \frac{\kappa g}{\rho_0 V}\oint_S z_*\boldsymbol{\nabla}\rho\cdot\hat{\boldsymbol{n}}\,\mathrm{d}S, \tag{2.16}$$

$$\mathscr{M} + D_p = \frac{\kappa g}{\rho_0 V}\int_V -\frac{\mathrm{d}z_*}{\mathrm{d}\rho}|\boldsymbol{\nabla}\rho|^2\,\mathrm{d}V, \tag{2.17}$$

in which a supplementary field variable ψ has been introduced as $\psi = \int^{\rho} z_*(\hat{\rho})\,\mathrm{d}\hat{\rho}$ (which yields $\boldsymbol{\nabla}\psi = z_*\boldsymbol{\nabla}\rho$) in order to incorporate the implicit dependence of z_* on ρ. Also, the integration is carried out over the control volume V which encloses the entire turbulent flow, bounded by its control surface S.

Notice that (2.14) only employs the Boussinesq approximation and includes no further assumptions. Based on this expression for $\mathrm{d}\mathscr{P}_B/\mathrm{d}t$, the BPE may change due either to surface fluxes across S or material changes within volume V. In addition, if the system is not closed, the surface fluxes are non-zero and might include both advective and diffusive processes, respectively denoted by $\mathscr{F}_{adv}$ and $\mathscr{F}_{diff}$. These surface fluxes represent reversible processes that are sign-indefinite and could support larger-scale stirring of the flow. On the other hand, the volume integral, denoted by $\mathscr{M} + D_p$, represents a diffusive destruction of small-scale density variance which is positive definite by construction because the adiabatic sorting of the density field ensures $\mathrm{d}z_*/\mathrm{d}\rho < 0$, implying that the small-scale processes involved in this destruction lead to an increase in BPE.

For an incompressible Boussinesq fluid, the turbulent diapycnal mixing, $\mathscr{M}$ in (2.14), represents the rate of irreversible energy conversion from internal energy into BPE (Tailleux 2009), which turns out to be identical to the irreversible conversion of available potential energy (or APE, defined as the difference between total potential energy $\mathscr{P}$ and the BPE, i.e. $\mathscr{P}_A = \mathscr{P} - \mathscr{P}_B$) into fluid internal energy, denoted by ε_p (i.e. $\mathscr{M} = \varepsilon_p$ as discussed by Tailleux 2009). The laminar analogue of $\mathscr{M}$ is denoted by D_p which characterizes the diffusive transfer of motionless fluid internal energy into BPE. This distinction between the laminar and turbulent conversions of internal energy into BPE is only important for low Reynolds number flows in which D_p provides comparable contributions to $\mathscr{M}$, while for high Reynolds number flows D_p is truly negligible. Also note that for a laminar flow in its reference state of potential energy where $\boldsymbol{\nabla}\rho = \mathrm{d}\rho/\mathrm{d}z$ and $z_* \equiv z$, the turbulent diapycnal mixing would be zero (i.e. $\mathscr{M} = 0$) and therefore D_p follows from (2.17) as $D_p = \kappa g\Delta\rho A/(\rho_0 V)$ where $\Delta\rho = \rho_{max} - \rho_{min}$ and A denotes a horizontal surface area of the laminar isopycnal surface within the control volume V.

2.3. *A revised formulation of scalar diffusivity K_ρ*

Our goal in the remainder of this section is to relate the exact formulation of diascalar diffusivity, proposed in WD96, to the definition of mixing efficiency proposed by Caulfield & Peltier (2000), thereby obtaining a more useful expression for K_ρ in terms of irreversible mixing efficiency (to be precisely defined in what follows). For this purpose, we should first note that for any high Reynolds number flow, D_p is several orders of magnitude smaller than irreversible mixing $\mathscr{M}$ and thus (2.17) can be written as:

$$\mathscr{M} = \frac{\kappa g}{\rho_0 V} \int_V - \left(\frac{\mathrm{d}\rho_*}{\mathrm{d}z}\right)^{-1} |\nabla \rho|^2 \, \mathrm{d}V, \tag{2.18}$$

in which we have replaced $\mathrm{d}z_*/\mathrm{d}\rho$ by $(\mathrm{d}\rho_*/\mathrm{d}z)^{-1}$ which essentially follows from the definitions of z_* and ρ_* as stated in the previous sections. In this regard it is important to note that on a given isopycnal surface, the scalar field $z_*(\boldsymbol{x}, t)$ is invariant by definition and therefore z_* is a unique function of ρ such that $z = z_*(\rho)$. Alternatively, at a given reference height, the density field is invariant by definition and therefore ρ is a unique function of z_* such that $\rho_* = \rho(z_*)$. These expressions imply that $\mathrm{d}z_*/\mathrm{d}\rho = (\mathrm{d}\rho_*/\mathrm{d}z)^{-1}$.

Because the volume integral in (2.18) is coordinate-independent, we choose $\mathrm{d}V$ in (2.18) to represent the infinitesimal volume of fluid enclosed between two adiabatically re-stratified isopycnal surfaces such that $\mathrm{d}V_* = \mathrm{d}S_* \mathrm{d}n$. Here $\mathrm{d}S_*$ is the infinitesimal surface area of an isopycnal surface and $\mathrm{d}n$ denotes its infinitesimal perpendicular (i.e. diapycnal) distance from the adjacent isopycnal (see figure 1 of WD96 for a depiction of $\mathrm{d}V_*$, $\mathrm{d}S_*$ and $\mathrm{d}n$). Note that in the reference state of minimum BPE, $\mathrm{d}n \equiv \mathrm{d}z$ where $\mathrm{d}z$ is the infinitesimal vertical distance. As a result, the volume integral in (2.18) can be split into a surface integral and a line integral which may be rewritten as:

$$\frac{\kappa g}{\rho_0 V} \int_V - \left(\frac{\mathrm{d}\rho_*}{\mathrm{d}z}\right)^{-1} |\nabla \rho|^2 \, \mathrm{d}V = \frac{\kappa g}{\rho_0 V} \int_z \int_{S_*} - \left(\frac{\mathrm{d}\rho_*}{\mathrm{d}z}\right)^{-1} |\nabla \rho|^2 \, \mathrm{d}S_* \, \mathrm{d}z. \tag{2.19}$$

Moreover, inasmuch as ρ_* changes equally across an adiabatically re-stratified isopycnal by construction, $\mathrm{d}\rho_*/\mathrm{d}z$ becomes constant, by definition, on $\mathrm{d}S_*$ and thus it may be taken outside of the above surface integral, which yields:

$$\frac{\kappa g}{\rho_0 V} \int_z - \left(\frac{\mathrm{d}\rho_*}{\mathrm{d}z}\right)^{-1} \left[\int_{S_*} |\nabla \rho|^2 \, \mathrm{d}S_*\right] \mathrm{d}z = \frac{g}{\rho_0 L_z} \int_z \underbrace{-\kappa \left(\frac{\mathrm{d}\rho_*}{\mathrm{d}z}\right)^{-1} \langle |\nabla \rho|^2 \rangle_{S_*}}_{\phi_d(z,t)} \mathrm{d}z, \tag{2.20}$$

where $\langle . \rangle_{S_*}$ denotes averaging over an adiabatically re-stratified surface, identical to the notation employed in (2.7). Furthermore, the resulting integrand in the above expression is recognized as the diapycnal flux, ϕ_d, defined in (2.8).

The above equalities lead to a useful representation of irreversible mixing $\mathscr{M}$ as being proportional to the vertically averaged rate of diapycnal flux $\langle \phi_d \rangle$, or:

$$\mathscr{M} = \frac{g}{\rho_0} \langle \phi_d(z, t) \rangle = \left(\frac{\mathscr{E}}{1 - \mathscr{E}}\right) \varepsilon_k. \tag{2.21}$$

Note that the second equality in this equation follows from the definition of instantaneous mixing efficiency as (Caulfield & Peltier 2000):

$$\mathscr{E} = \frac{\mathscr{M}}{\mathscr{M} + \varepsilon_k}. \tag{2.22}$$

It is also clear from (2.21) that, by definition, the irreversible mixing $\mathscr{M}$ represents the flux of a force, while the averaged diapycnal flux $\langle\phi_d\rangle_z$ may be considered a mass flux, analogous to the connection between the buoyancy flux $(g/\rho_0)\langle\overline{\rho' w'}\rangle$ and the density flux $\langle\overline{\rho' w'}\rangle$.

Based on (2.21), the irreversible diapycnal flux $\phi_d(z, t)$ can be represented entirely in terms of $\mathscr{E}$ and ε_k. Following WD96, for calculating the exact diapycnal diffusivity as discussed in § 2.1.4, we may also employ $\phi_d(z, t)$ as the required 'flux' in (1.1) and divide it by the vertical gradient of the sorted background density profile ρ_* (similar to the discussion in connection with (2.7)). The resulting diffusivity is denoted by K_ρ^* and can be written in dimensionless form as:

$$\widetilde{K}_\rho^* = -\frac{1}{\kappa}\frac{\langle\phi_d(z,t)\rangle}{\langle \mathrm{d}\rho_*/\mathrm{d}z\rangle} = Pr\left(\frac{\mathscr{E}}{1-\mathscr{E}}\right)\frac{\varepsilon_k}{\nu N_*^2} = \Gamma\, Pr\, Re_{b*}, \tag{2.23}$$

in which N_*^2 is the buoyancy frequency calculated using the sorted density profile ρ_* and thus the buoyancy Reynolds number Re_{b*} is accordingly defined. Moreover, $\Gamma = \mathscr{E}/(1-\mathscr{E})$ should be considered as a flux coefficient, which may be compared to its analogue $\gamma = R_f/(1-R_f)$ in the Osborn formula (2.5) and is defined based on the true mixing efficiency $\mathscr{E}$ not the flux Richardson number R_f.

It is very interesting and useful to compare the Osborn formula in (2.5) and our newly derived expression for K_ρ in (2.23). In fact, they resemble each other as they both linearly depend on a buoyancy Reynolds number and a flux coefficient. Although in (2.23) Re_{b*} is defined using the sorted density profile, Re_b and Re_{b*} would be almost identical (i.e. $Re_{b*} \approx Re_b$) if the vertically averaged values of $\mathrm{d}\rho/\mathrm{d}z$ and $\mathrm{d}\rho_*/\mathrm{d}z$ are used in calculating N^2 and N_*^2 respectively (e.g. see Smyth *et al.* 2001 for similar definitions of averaged profiles in which they average numerator and denominator individually). However, the major difference between these formulations concerns the definitions of flux coefficients, namely $\gamma = R_f/(1-R_f)$ in the Osborn formula and $\Gamma = \mathscr{E}/(1-\mathscr{E})$ in ours. In other words, it is only by using the correct definition of mixing efficiency $\mathscr{E}$ instead of the flux Richardson number R_f that one is able to obtain a mathematically accurate calculation of K_ρ. While the Osborn formula has been derived using several important assumptions that are often criticized (e.g. see Ivey *et al.* 2008; Mashayek & Peltier 2013), this new formula relies upon the validity of the Boussinesq approximation to the governing equations and assumes only that Re is high enough to yield $\mathscr{M} \gg D_p$. This assumption is expected to be valid for all geophysical flows. For these reasons, it is perhaps most appropriate to refer to (2.23) as the 'generalized Osborn' formula.

It is also worthwhile mentioning that, as was noted in WD96, the Osborn–Cox model in (2.6) is also very similar in form to the formulation proposed by Winters–D'Asaro in (2.7). However, (2.7) is based on the gradient of the total scalar field averaged over an isoscalar surface, while (2.6) is formulated based on the gradient of scalar fluctuations averaged over the entire control volume (again the differences between density gradients in their denominator become negligible if vertical averaging is employed).

Effectively, the new formulation for K_ρ in (2.23) might be considered as an essential but otherwise slight modification to the Osborn formula itself. This modification includes a correct definition of mixing efficiency as introduced in (2.22) based on instantaneous irreversible mixing $\mathscr{M}$ (or diapycnal flux ϕ_d) which isolates the irreversible small-scale diffusive processes that lead to the destruction of density variance and hence result in diapycnal mixing. The assumption that the

mixing efficiency may be represented in terms of R_f (i.e. employing the Osborn formula), whether R_f is defined as $R_f = \mathbb{B}/\mathbb{P}$ (Osborn 1980) or more generally as $R_f = \mathbb{B}/(\mathbb{B} + \varepsilon_k)$ (Ivey & Imberger 1991), would inevitably involve the fundamental problem of improperly including the reversible processes in the calculation of diapycnal diffusivity.

In addition, as (2.23) makes explicit, it is physically expected that the true diapycnal diffusivity has to depend on the buoyancy Reynolds number, for the entire range of accessible Re_b. As explained earlier, (2.23) follows directly from the governing conservation equations and is therefore independent of the physical process that has generated the turbulence so long as that process conforms to the Boussinesq approximations and occurs at sufficiently high Re. In fact, based on the length scale argument of Ivey & Imberger (1991), it has been previously suggested that K_ρ might depend on $Re_b = \varepsilon_k/(\nu N^2)$ (e.g. see Ivey *et al.* 2008). The evidence for this dependence has also been documented for a variety of different physical processes that lead to stratified turbulent mixing. For example in their laboratory experiments, Ivey, Winters & De Silva (2000) studied benthic boundary layer turbulence energized by internal waves impinging on a uniformly sloping bottom, while Barry *et al.* (2001) and Rehmann & Koseff (2004) respectively focused on the turbulent mixing generated by a horizontally oscillating vertical grid and a towed grid. Furthermore, studies on the differential diffusion of salinity and temperature, including the laboratory experiments of Jackson & Rehmann (2003) and Martin & Rehmann (2006) and the numerical investigation of Smyth *et al.* (2005), have all suggested empirically that the measured scalar diffusivities are connected to Re_b. Furthermore, SKIF employed DNS to study the special case of homogeneous stratified and sheared turbulence and also found empirically that K_ρ may be represented in terms of Re_b^n. Based on the analysis presented herein, we are now in a position to formally generalize the length scale arguments of Ivey & Imberger (1991) and to assert that the dependence of K_ρ on Re_b is a direct consequence of the Boussinesq equations.

Besides the explicit linear dependence of K_ρ on Re_b in (2.23), K_ρ might also depend implicitly on Re_b due to a possible relation between Γ (or $\mathscr{E}$) and Re_b. For example, SKIF applied a least-squares power-law fit to their dataset with $Re_b > 100$ and showed that $K_\rho \propto Re_b^{1/2}$, which according to (2.23) evidently implies $\Gamma \propto Re_b^{-1/2}$. This latter implication has also been empirically noted by SKIF using R_f as a measure of mixing efficiency. Furthermore, Lozovatsky & Fernando (2012), based on their observational data with $Re_b > 10^4$ for stratified flow of the atmospheric surface boundary layer, also suggested a $-1/2$ power-law relation between R_f and Re_b. We will return to the discussion of such power-law relations in § 4.3.

3. Momentum diffusivity and turbulent Prandtl number

The momentum diffusivity K_m (also commonly referred to as the eddy viscosity) is conventionally defined as the ratio of the Reynolds stress (turbulent momentum flux) to the vertical shear following the mixing length ansatz as (cf. (2.1)):

$$\widetilde{K}_m^{ml} = -\frac{1}{\nu}\frac{\langle \overline{u'w'} \rangle}{\langle \mathrm{d}\overline{u}/\mathrm{d}z \rangle}, \tag{3.1}$$

in which $\widetilde{K}_m = K_m/\nu$ is non-dimensionalized by the kinematic viscosity ν.

Similar to Osborn (1980) who derived an expression for scalar diffusivity (2.5) (as reviewed in § 2.1.2), Crawford (1982) also employed the simplified TKE equation

and derived a formulation for momentum diffusivity (denoted by K_m^{Cr}) based on its definition in (3.1) by employing the flux Richardson number, denoted R_f, to replace the momentum flux term. In dimensionless form, this formulation can be represented as:

$$\widetilde{K}_m^{Cr} = \left(\frac{1}{1-R_f}\right)\frac{\varepsilon_k}{\nu S^2} = \left(\frac{1}{1-R_f}\right) Ri_b Re_b, \tag{3.2}$$

where $Ri_b = N^2/S^2$ denotes the vertically averaged gradient Richardson number.

As pointed out recently by Venayagamoorthy & Stretch (2010), these formulations of K_m include reversible contributions to the momentum fluxes in the production terms of the TKE equation and therefore do not represent K_m with adequate precision. In order to address this issue in the context of statistically homogeneous sheared and stratified turbulence, Venayagamoorthy & Stretch (2010) also employed the simplified balance equation of TKE by assuming stationary turbulence as in Crawford (1982) (i.e. $\mathbb{P} = \mathbb{B} + \varepsilon_k$ as discussed in § 2.1.2), but they replaced the buoyancy flux term in terms of K_ρ^{ml} as defined in (2.1). In order to extend their expression for K_m to non-stationary homogeneous flows, they simply replaced K_ρ^{ml} with K_ρ^{cox} (the latter is defined in (2.6), which becomes identical to K_ρ^* in (2.7) for the special case of homogeneous turbulence as discussed in Venayagamoorthy & Stretch 2006). In what follows, we will nevertheless address this issue by following a different and more generic approach.

3.1. *A revised formulation of momentum diffusivity K_m*

Our aim in this subsection is to propose a new expression for K_m (denoted by K_m^*) which distinguishes between reversible and irreversible contributions to the momentum diffusivity of a Boussinesq turbulent flow; a distinction which is essential to properly defining this quantity. For this purpose we note that the total energy that is extracted from the mean flow due to a turbulent mixing event, with a rate $\mathbb{P}$, is either retained in the system even after the turbulence decays (we define this component as the 'reversible' component, $\mathbb{P}_r$, because it is capable in principle of initiating further mixing events) or is eventually dissipated irreversibly into internal energy (denoted by $\mathbb{P}_i$). This decomposition can be written as:

$$\mathbb{P} = \mathbb{P}_r + \mathbb{P}_i, \tag{3.3}$$

in which $\mathbb{P}$ is associated with the shear production term in the TKE equation due to the interaction between the Reynolds stresses and the mean and fluctuating components of the flow as defined in (2.2).

To further establish the definition of $\mathbb{P}_i$ and $\mathbb{P}_r$ in (3.3), we invoke a Reynolds decomposition of the total kinetic energy reservoir $\mathscr{K} = \langle \overline{\boldsymbol{u}\cdot\boldsymbol{u}} \rangle/2$, into its background mean $\overline{\mathscr{K}} = \langle \overline{u}^2 \rangle/2$ and turbulent $\mathscr{K}' = \langle \overline{\boldsymbol{u}'\cdot\boldsymbol{u}'} \rangle/2$ components, as well as the decomposition of total potential energy into APE and BPE as discussed in § 2.2, i.e.

$$\mathscr{K} = \overline{\mathscr{K}} + \mathscr{K}', \tag{3.4}$$

$$\mathscr{P} = \mathscr{P}_A + \mathscr{P}_B. \tag{3.5}$$

Furthermore, the evolution equations for the energy budgets associated with $\mathscr{K}$, $\mathscr{K}'$ and $\mathscr{P}_A$ may be written as (Caulfield & Peltier 2000):

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{K} = -\mathbb{B} - \varepsilon_k, \tag{3.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{K}' = -\mathbb{B} + \mathbb{P} - \varepsilon_k', \tag{3.7}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{P}_A = \mathbb{B} - \mathscr{M}, \tag{3.8}$$

in which $\varepsilon_k' = 2\nu\overline{\langle \mathsf{s}_{ij}'\mathsf{s}_{ij}'\rangle}$ (cf. (2.4)) and $\mathsf{s}_{ij}' = (\partial u_i'/\partial x_j + \partial u_j'/\partial x_i)/2$ is the disturbance strain rate tensor.

The 'irreversible' component of the shear production, $\mathbb{P}_i$, is identified by considering the evolution equation for the total available energy, $E_A = \mathscr{K} + \mathscr{P}_A$, namely:

$$\frac{\mathrm{d}}{\mathrm{d}t}E_A = \frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{K} + \mathscr{P}_A) = -(\mathscr{M} + \varepsilon_k). \tag{3.9}$$

The above equation implies that the total energy that is available for irreversible mixing at larger scales of motion, in the form of both total kinetic energy and APE, must dissipate at the smallest scales due to either (i) viscous dissipation of fluid filaments (ε_k) or (ii) a diffusive destruction of small-scale density variance and therefore of potential energy (ε_p). Here we have exploited the fact that $\varepsilon_p = \mathscr{M}$ for incompressible Boussinesq flows as explained in § 2.2 (see Tailleux 2009 for further discussion).

We therefore attribute the irreversible monotonic decrease of E_A in (3.9) to the irreversible component of shear production, $\mathbb{P}_i$, that is manifested at dissipation scales. Thus $\mathbb{P}_i$ is defined as:

$$\mathbb{P}_i = \mathscr{M} + \varepsilon_k. \tag{3.10}$$

Furthermore, the 'reversible' component of the shear production, $\mathbb{P}_r$, is identified by considering the energy budget of the stratified turbulence (defined as the sum of TKE and APE, see Smyth, Carpenter & Lawrence 2007), which evolves based on the following equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{K}' + \mathscr{P}_A) = -\mathscr{M} - \varepsilon_k' + \mathbb{P}. \tag{3.11}$$

By employing (3.3) and (3.10) and after adding and subtracting mean flow viscous dissipation $\overline{\varepsilon} = \nu\langle(\mathrm{d}\overline{u}/\mathrm{d}z)^2\rangle = \varepsilon_k - \varepsilon_k'$, the reversible shear production, $\mathbb{P}_r$, is defined as:

$$\mathbb{P}_r = \frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{P}_A + \mathscr{K}') - \overline{\varepsilon}. \tag{3.12}$$

In other words, the reversible component of $\mathbb{P}$ is the rate of change in the sum of APE and TKE reservoirs due to the reversible energy transfer to the mean flow kinetic energy after subtracting the irreversible dissipation of mean flow kinetic energy ($\overline{\varepsilon}$), because the latter is supplied by the irreversible component of $\mathbb{P}$. Therefore the energy that was noted earlier as being retained after a mixing event is completed, is only associated with the mean flow kinetic energy that is capable of initiating further mixing events. In addition, notice that $\mathbb{P}_r$ does not contribute to the irreversible loss of E_A in (3.9). It, rather, contributes to the reversible energy exchanges between the APE–TKE energy reservoir and the mean flow kinetic energy ($\overline{\mathscr{K}}$).

For characterizing K_m^*, only the irreversible component of such shear production (i.e. $\mathbb{P}_i$), must be employed, whereby K_m^* is defined as:

$$\widetilde{K}_m^* = \frac{\mathbb{P}_i}{\nu S^2}. \tag{3.13}$$

Similar to the general definition of K_ρ in (1.1), K_m may also be defined as the ratio of a 'flux' to a 'gradient' (with a negative sign). Therefore (3.13) is obtained by considering $-\mathbb{P}_i/S$ as the appropriate irreversible momentum 'flux' and the vertical shear S, as the required 'gradient'.

Given the definition of $\mathbb{P}_i$ in (3.10), there are two alternative ways of incorporating previous concepts regarding irreversible mixing of the scalar field into the definition of $\widetilde{K}_m^*$. First, by employing the definition of irreversible mixing efficiency in (2.22), $\mathbb{P}_i$ may be re-written as:

$$\mathbb{P}_i = \frac{\varepsilon_k}{1-\mathscr{E}}. \tag{3.14}$$

Second, by evoking the relationship between $\mathscr{M}$ and ϕ_d in (2.21) and recalling the definition of $\widetilde{K}_\rho^*$ in (2.23), which effectively implies $\widetilde{K}_\rho^* = \mathscr{M}/(\kappa N_*^2)$, $\mathbb{P}_i$ may be alternatively expressed as:

$$\mathbb{P}_i = (\kappa N_*^2)\widetilde{K}_\rho^* + \varepsilon_k, \tag{3.15}$$

in which N_*^2 is again defined based on the adiabatically sorted density profile.

The above two expressions for $\mathbb{P}_i$ in (3.14) and (3.15) are identical. Nonetheless we have provided both forms because each manipulation has been exploited previously in Crawford (1982) and Venayagamoorthy & Stretch (2010) with certain limitations in each case as reviewed earlier. In what follows we will use each of these alternative forms of $\mathbb{P}_i$ to derive corresponding alternative expressions for $\widetilde{K}_m^*$ based on (3.13) which facilitates comparison with the previous formulae of Crawford (1982) and Venayagamoorthy & Stretch (2010).

By employing $\mathbb{P}_i$ based on either (3.14) or (3.15), $\widetilde{K}_m^*$ may be written in dimensionless form as either of the following expressions:

$$\widetilde{K}_m^* = \left(\frac{1}{1-\mathscr{E}}\right)\frac{\varepsilon_k}{\nu S^2}, \tag{3.16}$$

$$\widetilde{K}_m^* = \left(\frac{Ri_b^*}{Pr}\right)\widetilde{K}_\rho^* + \frac{\varepsilon_k}{\nu S^2}, \tag{3.17}$$

in which $Ri_b^* = N_*^2/S^2$ is defined based on the sorted background density profile. Obviously, by equating $\widetilde{K}_m^*$ based on (3.16) and (3.17), the revised expression for $\widetilde{K}_\rho^*$ in (2.23) is obtained.

It is important to note that (3.16) closely resembles the formulation of Crawford (1982) in (3.2) except for the following critical differences. (i) The new formula for K_m employs the correct definition of mixing efficiency $\mathscr{E}$ rather than the flux Richardson number R_f (refer to Peltier & Caulfield 2003 and Tailleux 2009 for a discussion of the intrinsic differences between $\mathscr{E}$ and R_f). Moreover, (ii) the expression for K_m in (3.16) does not involve any assumptions concerning the flow dynamics, as opposed to (3.2) which is derived based on the same limitations as those of the Osborn formula for K_ρ.

In addition, (3.17) also closely resembles the homogeneous formulation of Venayagamoorthy & Stretch (2010) (see their equation (2.8)) although it significantly extends their expression to any inhomogeneously stratified Boussinesq flow at sufficiently high Re.

3.2. *Turbulent Prandtl number* Pr_t

A non-dimensional parameter of fundamental concern is the ratio of turbulent momentum diffusivity to the diapycnal diffusivity of density as characterized by the turbulent Prandtl number $Pr_t = K_m/K_\rho$. Using the newly proposed expressions for K_ρ in (2.23) and for K_m in (3.16), Pr_t becomes:

$$Pr_t = \frac{Ri_b}{\mathscr{E}}. \tag{3.18}$$

In the absence of careful definitions for K_ρ and K_m, Pr_t is commonly defined as the ratio of K_m^{ml} in (3.1) to K_ρ^{ml} in (2.1) based on the mixing length argument, which results in $Pr_t^{ml} = Ri_b/R_f$. The new expression for Pr_t in (3.18) also resembles this basic definition of turbulent Prandtl number except that $\mathscr{E}$ appears in place of R_f. As a result of this slight but important modification, Pr_t represents only the ratio of irreversible components of scalar and momentum diffusivities, which again critically depends on the accurate definition of irreversible mixing efficiency.

4. Direct numerical simulations of inhomogeneously stratified, shear-generated turbulence

In this second part of the paper, we will illustrate the results obtained using the newly proposed formulations for K_ρ, K_m and Pr_t, and in the case of K_ρ will compare this with the complete set of models we have reviewed. A large dataset of original DNS results associated with the KH ansatz will be employed to achieve this goal. First, in § 4.1 we describe our choice of problem configuration as well as the numerical methodology to be employed. The results obtained on the basis of these analyses are then consolidated in §§ 4.2–4.5. In appendix A, we will comment on the other model problem that has been employed for the study of shear-induced turbulence in stratified flows (e.g. in SKIF), which is referred to as homogeneous stratified and sheared turbulence.

4.1. *Design of the suite of DNS analyses*

Our intention is to provide a detailed investigation of the stratified turbulence that is generated through the transition of a primary shear instability in the form of a Kelvin–Helmholtz wave which thereafter 'breaks' to generate the turbulence of interest. As initial conditions, the velocity and density fields are set to the following one-dimensional initial profiles:

$$\overline{U}_0(z) = U_0 \tanh\left(\frac{z}{d}\right), \quad \overline{\rho}_0(z) = \rho_0 - \Delta\rho \tanh\left(\frac{z}{d}\right), \tag{4.1a,b}$$

which are both assumed to have equal thickness $2d$. Also U_0 and ρ_0 denote a reference velocity and density while $\Delta\rho$ describes a fixed density difference across the shear layer. The coordinate system is chosen such that x, y and z are aligned with the streamwise, spanwise and vertical directions respectively.

We non-dimensionalize the governing Boussinesq equations in (2.9)–(2.11) by choosing ρ_0, U_0 and d (half of the shear layer thickness) as the characteristic density, velocity and length scales which become:

$$\frac{\mathrm{D}u_i}{\mathrm{D}t} = -\frac{\partial p}{\partial x_i} - Ri_0\rho\delta_{i3} + \frac{1}{Re}\frac{\partial^2 u_i}{\partial x_j^2}, \tag{4.2}$$

$$\frac{\partial u_i}{\partial x_i} = 0, \tag{4.3}$$

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = \frac{1}{Re\,Pr}\frac{\partial^2 \rho}{\partial x_j^2}. \tag{4.4}$$

The three non-dimensional control parameters which appear in the above non-dimensional form of the governing equations, namely the Reynolds, Richardson and Prandtl numbers are respectively defined as:

$$Re = \frac{U_0 d}{\nu}, \quad Ri = -\frac{g}{\rho_0}\frac{\partial \rho}{\partial z}\left(\frac{\partial u}{\partial z}\right)^{-2}, \quad Pr = \frac{\nu}{\kappa}. \tag{4.5a–c}$$

The minimum value in the initial profile of Ri (not to be confused with its vertically averaged bulk measure defined in (1.2) as Ri_b), obtained from the initial density and velocity profiles in (4.1), occurs at the inflection point of the velocity profile (located at $z = 0$). We will denote this specific choice of gradient Richardson number by Ri_0, which appears as the coefficient of the non-dimensional buoyancy term in (2.9). We also retain our non-dimensional representations of diapycnal diffusivity and momentum diffusivity as introduced in the previous section, namely $\widetilde{K}_\rho = K_\rho/\kappa$ and $\widetilde{K}_m = K_m/\nu$ in which $\kappa = \nu/Pr$ and $\nu = U_0 d/Re$ based on our choice of Pr and Re in the non-dimensionalized Boussinesq equations. We also note in passing that it is common practice in the oceanographic literature (e.g. Smyth *et al.* 2001) to choose $2\rho_0$, $2U_0$ and $2d$ as the characteristic density, velocity and length scales for the non-dimensionalization. Therefore our Re should be multiplied by four in order to make direct comparisons to these earlier studies.

The DNS-based analysis of stratified turbulent flows to be discussed in the main body of this paper include those reported in Salehipour, Peltier & Mashayek (2015) (with $Re = 6000$, $Ri_0 = 0.12$ and $Pr = 1, 2, 4, 8, 16$) as well as an additional series of new simulations which are all listed in table 1. In the subset of these simulations discussed in Salehipour *et al.* (2015), the focus was upon variations in the molecular Prandtl number at a fixed Reynolds number of $Re = 6000$ and a fixed minimum initial gradient Richardson number (at the inflection point) of $Ri_0 = 0.12$. The new set of simulations includes various Ri_0 at a fixed Re and Pr ($Re = 6000$, $Pr = 1, 8$) and various Re at a fixed Ri_0 and Pr ($Ri_0 = 0.16$, $Pr = 1, 8$). All these simulations have been performed using the modern spectral element code, Nek5000, developed at the Argonne National Laboratory (Fischer, Kruse & Loth 2002; Fischer, Lottes & Kerkemeier 2008). The governing Navier–Stokes equations are discretized using the weak Galerkin formulation by employing Nth-order Lagrange polynomial interpolants following the $P_N - P_{N-2}$ formulation of Maday, Patera & Rønquist (1990) for the velocity (N) and pressure ($N - 2$) spaces respectively. In terms of the temporal discretization, a third-order semi-implicit operator-splitting method is employed (Maday *et al.* 1990; Fischer 1997). Besides higher-order accuracy, this code is ideally suited for our high-Re, high-Pr numerically demanding simulations due to its perfect scaling up to thousands of processors (we have successfully employed up to 65 536 cores on a BlueGene/Q machine). Table 1 also includes the streamwise (L_x), spanwise (L_y) and vertical (L_z) extents of the computational domain as well as the number of elements in each direction (i.e. K_x, K_y, K_z). The domain is periodic in the x and y directions, while the top and bottom boundaries are assumed to be free-slip and impermeable for velocity and insulated (zero density flux) for the density field. For

Re	Ri_0	Pr	L_x	L_y	L_z	L_z^c	N	K_x	K_y	K_z	K_z^c
6 000	0.12	1	14.27	3	30	10	10	64	13	66	44
6 000	0.12	2	14.27	3	30	10	10	64	13	66	44
6 000	0.12	4	14.27	3	30	10	10	90	18	88	62
6 000	0.12	8	14.27	3	30	10	10	127	25	116	88
6 000	0.12	16	14.27	3	30	10	10	180	36	156	124
6 000	0.005	1	14.27	3	50	30	10	64	13	154	132
6 000	0.01	1	14.27	3	30	20	10	64	13	104	88
6 000	0.02	1	14.27	3	30	20	10	64	13	104	88
6 000	0.04	1	14.27	3	30	20	10	64	13	104	88
6 000	0.08	1	14.27	3	30	20	10	64	13	104	88
6 000	0.16	1	14.27	3	30	10	10	64	13	66	44
6 000	0.20	1	14.27	3	30	10	10	64	13	66	44
6 000	0.02	8	14.27	3	30	10	10	127	25	116	88
6 000	0.04	8	14.27	3	30	10	10	127	25	116	88
6 000	0.10	8	14.27	3	30	10	10	127	25	116	88
6 000	0.14	8	14.27	3	30	10	10	127	25	116	88
6 000	0.16	8	14.27	3	30	10	10	127	25	116	88
6 000	0.18	8	14.27	3	30	10	10	127	25	116	88
6 000	0.20	8	14.27	3	30	10	10	127	25	116	88
4 000	0.16	1	14.27	3	30	10	10	47	10	53	33
4 000	0.16	8	14.27	3	30	10	10	94	19	91	65
8 000	0.16	1	14.27	3	30	10	10	80	16	69	55
8 000	0.16	8	14.27	3	30	10	10	158	32	139	109
12 000	0.16	1	14.27	3	30	10	10	108	22	100	74

TABLE 1. Details of the three-dimensional numerical experiments in which the total grid points is about $N^3 K_x K_y K_z$. Also, L_z^c represents the height of a central region of the domain which is further refined with K_z^c elements. Outside L_z^c, the adjacent elements of the grid are gradually stretched by a factor of 1.25 %.

further details concerning the numerical set-up, grid design, domain size, resolution requirements and initial perturbations the interested reader is referred to Salehipour *et al.* (2015).

For all the figures to be presented in this paper, we have adopted a special annotation to differentiate between the simulations that have different Pr and Ri_b. All the runs at $Pr = 1$ are marked by a circle (○) while those at $Pr = 8$ are shown by a triangle (△). Single runs at $Pr = 2, 4, 16$ are also shown respectively by $\times$, $+$ and $*$. Furthermore, the marker size was chosen to approximately represent the vertically averaged gradient Richardson number, Ri_b, such that lower Ri_b corresponds to smaller symbols. These distinctions are not intended to delineate these runs, but rather to facilitate our discussions to follow.

Unless otherwise stated explicitly, in all the results to be discussed in what follows the bulk measure of the gradient Richardson number, Ri_b, will be calculated as $Ri_b = N^2/\langle(\mathrm{d}\overline{u}/\mathrm{d}z)^2\rangle$ (i.e. vertical shear is first squared and then vertically averaged) in which $N^2 = -Ri_0\langle \mathrm{d}\overline{\rho}/\mathrm{d}z\rangle = Ri_0\Delta\rho/L_z$ due to our non-dimensionalization. Appendix B compares this preferred form of Ri_b with an alternative definition for Ri_b in which shear is first vertically averaged over the shear layer and then squared. Based on all of the DNS cases, an empirical relationship between the two definitions is also suggested in appendix B, which may serve as a mapping function between these two forms of Ri_b.

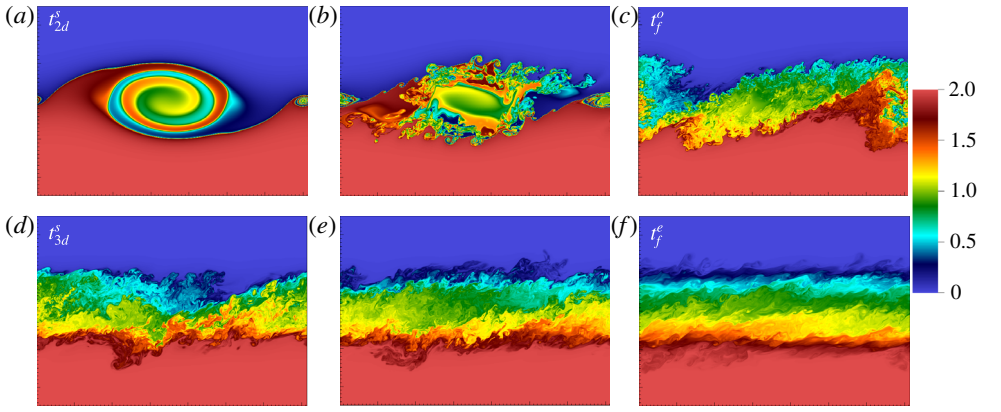

FIGURE 1. Illustrations of density field evolution. Streamwise-vertical cross-section at the mid-point in the cross-stream y direction for different time snapshots t^s_{2d}, $t \in [t^s_{2d} \quad t^o_f]$, t^s_{3d}, $t \in [t^s_{3d} \quad t^e_f]$ and t^e_f. For the definition of these times refer to the text.

Figure 1 illustrates an example DNS analysis in which the density field for a case with $Re = 6000$, $Pr = 8$ and $Ri_0 = 0.16$ is shown in different panels corresponding to different snapshots of time during flow evolution. Note that the time-dependent flow field undergoes various stages of evolution, beginning with the growth of the two-dimensional Kelvin–Helmholtz instability whose kinetic energy saturates at a time denoted by t^s_{2d} in figure 1(*a*). The onset of the fully turbulent regime does not begin until $t^o_f > t^s_{2d}$, at which time the kinetic energy of inherently three-dimensional flow components finally exceeds that of the two-dimensional flow (figure 1*c*). For the time period of $t^s_{2d} < t < t^o_f$ (figure 1*b*), three-dimensional secondary instabilities grow and facilitate the transition to turbulence. At a time $t^s_{3d} > t^o_f$, three-dimensional kinetic energy has saturated relative to the two-dimensional flow component and thus there is hardly any visible signature of the original two-dimensional KH wave in the flow field (figure 1*d*). We have realized that it is critical not to define t^s_{3d} as the time when the kinetic energy of the three-dimensional flow is highest because this point may occur when the kinetic energy of the two-dimensional KH wave continues to overwhelm that of the induced three-dimensional turbulence and is therefore not associated with the fully turbulent phase. The period of sustained or decaying turbulence ends at a time denoted by t^e_f, which we have chosen to define based on a threshold minimum Re_b. The flow field begins to re-laminarize (figure 1*f*), usually around $Re_b = 20$ in agreement with previous suggestions in the literature (e.g. see Stillinger, Helland & Atta 1983; Smyth & Moum 2000*b*). For the purpose of all of the analyses to be discussed in this paper, t^e_f is, rather, defined based on a threshold minimum value of $Re_b = 1$ in order to also encompass the mixing properties of weakly turbulent flows with extremely low Re_b. In addition, we will entirely focus upon the period of fully turbulent flow (i.e. $t^s_{3d} \leqslant t \leqslant t^e_f$, depicted in figure 1*d–f*) in order to calculate mixing efficiency as well as the scalar and momentum diffusivities.

In terms of the non-dimensional control parameters, D_p (as introduced in § 2.2) may be rewritten as $D_p = Ri_0 \Delta\rho/(RePrL_z)$ where L_z is the vertical extent of the computational domain. In all the DNS cases tabulated in table 1, $\Delta\rho = 2$, $L_z = 30$, $Ri_0 \sim O(10^{-1}\text{–}10^{-2})$, $Pr \sim O(10^0\text{–}10^1)$ and $Re \sim O(10^3\text{–}10^4)$. Thus $D_p \sim O(10^{-9}\text{–}10^{-6})$, which is two to five orders of magnitude smaller than the irreversible mixing, $\mathscr{M}$,

that is usually of order $\mathscr{M} \sim O(10^{-4})$ (also non-dimensional). During the final stages of the KH life cycle however, the flow begins to re-laminarize and turbulent mixing decays to zero (i.e. $\mathscr{M} \sim D_p \sim 0$ when $Re_b \sim O(10^0$–$10^1)$). This may seem to suggest that the revised formulation for $\widetilde{K}_\rho$ in (2.23) might not be applicable because (2.18) is not strictly satisfied when $\mathscr{M} \sim D_p$. Nevertheless, since turbulent mixing is negligible during this stage in any event (i.e. $\mathscr{M} \sim 0$), our proposed framework still holds and predicts that both mixing efficiency and diapycnal diffusivity would decay to zero.

4.2. *Scalar diffusivity estimation: a comparative study*

Different estimations of diapycnal diffusivity K_ρ that are commonly employed in the literature will be compared here with the 'correct' value of scalar diffusivity K_ρ^*. As discussed in § 2.3, the 'correct' formulation of K_ρ^* was proposed in the context of the diascalar framework of WD96 which has been mathematically reformulated in this paper into the form of a more familiar 'Osborn-like' expression. Therefore K_ρ^* may be considered to have been calculated based on either method.

Figure 2 contains the individual comparisons made between each non-dimensional estimation of scalar diffusivity and $\widetilde{K}_\rho^* = K_\rho/\kappa$. Figures 2(*d*) and 2(*c*) investigate the accuracy of the Osborn (1980) (with $R_f = 0.17$ or $\gamma = 0.2$, see (2.5)) and Osborn & Cox (1972) (see (2.6)) models for estimating diapycnal diffusivity. As far as the mixing length model is concerned, $\widetilde{K}_\rho^{ml}$ in (2.1) has been calculated using two alternative assumptions for the perturbation density flux $\overline{\rho' w'}$. Namely (i) $\widetilde{K}_{\rho'}^{ml}$ in figure 2(*a*) involves the conventional Reynolds decomposition ($f = \bar{f} + f'$) whereas (ii) $\widetilde{K}_{\rho_{3d}}^{ml}$ in figure 2(*b*) employs a density flux that is only based on the inherently three-dimensional component of a triple Reynolds decomposition ($f = \bar{f} + f_{2d} + f_{3d}$). Here f denotes either density or vertical velocity fields and $\bar{f}$ isolates the horizontal mean from the $2d$ and $3d$ components which respectively represent the spanwise averaged and fully three-dimensional components of a flow field (refer to Caulfield & Peltier 2000 for precise definitions).

Figure 2 illustrates the fact that the models investigated herein are accurate only within a factor of approximately four. In general, there is a fundamental problem concerning the use of 'turbulent' density flux, $\overline{\rho' w'}$, in the mixing length model in order to estimate the diapycnal diffusivity. Namely, the density flux in the usual vertical coordinate system would inevitably and improperly include reversible processes in the estimation of K_ρ. For example, the direct field measurements of the instantaneous density flux $\overline{\rho' w'}$ (e.g. Moum 1990; Fleury & Lueck 1994) attest to the fact that the interpretation of such measurements is non-trivial since this flux becomes as frequently positive (down-gradient) as negative (counter-gradient). Furthermore, figure 16 in Mashayek & Peltier (2013) shows negative values for K_ρ derived on the basis of the use of (2.1) using $\overline{\rho' w'}$, which is clearly unphysical. Due to its logarithmic scale, figure 2(*a*) does not show the negative values of $\widetilde{K}_{\rho'}^{ml}$, but the large scatter away from the correct estimation of diapycnal diffusivity should be seen as indicative of the same fundamental problem. Although use of a triple Reynolds decomposition to obtain $\overline{\rho_{3d} w_{3d}}$ is empirically seen to produce values for this flux which are strictly positive, $\widetilde{K}_{\rho_{3d}}^{ml}$ is still unable to capture the complete irreversible diapycnal flux ϕ_d as defined in (2.8) or (2.21) because there exists a clear offset between $\widetilde{K}_{\rho_{3d}}^{ml}$ and $\widetilde{K}_\rho$ which points to the fact that $\overline{\rho_{3d} w_{3d}}$ systematically underestimates ϕ_d.

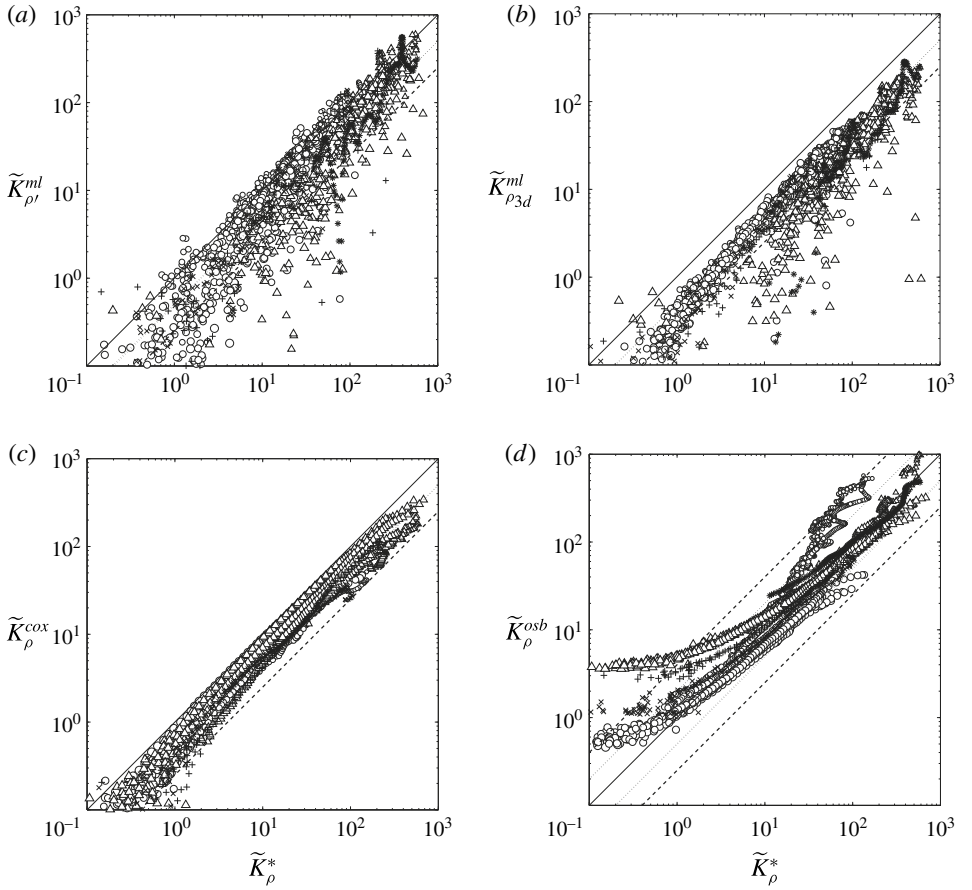

FIGURE 2. Comparing the accurate representation of scalar diffusivity $\widetilde{K}^{*}_{\rho}$ (2.7) or (2.23) with: (*a*) mixing length model (2.1) based on a full perturbation buoyancy flux $\widetilde{K}_{\rho'}$ or (*b*) based on inherently three-dimensional buoyancy flux $\widetilde{K}_{\rho_{3d}}$; (*c*) the Osborn–Cox model $\widetilde{K}^{cox}_{\rho}$ (2.6); and (*d*) the Osborn formula $\widetilde{K}^{osb}_{\rho}$ (2.5) (with $R_f = 0.17$). The dashed (dotted) lines show a factor of four (two) under/over-estimation. See figure 3 caption for description of symbols.

Although the Osborn–Cox model also underestimates the actual scalar diffusivity in figure 2(*c*), it seems to provide the best estimates among the complete set of approximate models. Not only is $\widetilde{K}^{cox}_{\rho}$ (or the Cox number) close to the actual values of $\widetilde{K}^{*}_{\rho}$ (i.e. $0.25\widetilde{K}^{*}_{\rho} \leqslant \widetilde{K}^{cox}_{\rho} \leqslant \widetilde{K}^{*}_{\rho}$), but it also correctly follows the trend in $\widetilde{K}^{*}_{\rho}$ (i.e. the independent series of data points are all individually aligned with the solid line of $\widetilde{K}^{cox}_{\rho} = \widetilde{K}^{*}_{\rho}$). In fact, the Osborn–Cox estimate would be accurate (within an approximately $\pm 50\,\%$ offset), if it were multiplied by a factor of 2. This is based on figure 2(*c*) in which the DNS data encompass the line $\widetilde{K}^{cox}_{\rho} = 0.5\widetilde{K}^{*}_{\rho}$ (dotted line in figure 2*c*).

The Osborn model in figure 2(*d*), on the other hand, does not follow the trend in $\widetilde{K}^{*}_{\rho}$. In fact, $\widetilde{K}^{osb}_{\rho}$ deviates noticeably from the line of $\widetilde{K}^{osb}_{\rho} = \widetilde{K}^{*}_{\rho}$ at both extreme limits of diapycnal diffusivity. This inaccuracy has been commonly remarked upon as being associated with the assumptions employed in the derivation of the Osborn formula

(Mashayek, Caulfield & Peltier 2013; Mashayek & Peltier 2013), in addition to the accepted deviations of mixing efficiency from any canonical constant value (e.g. Ivey *et al.* 2008). Although geophysical flows do not obey the simplifying assumptions of Osborn (1980), we showed in § 2.3 that for any Boussinesq stratified fluid with high Re, an expression for K_ρ similar to the Osborn formula is indeed accurate, provided that $\mathscr{E}$ replaces the flux Richardson number R_f. It is worthwhile noting that, even if a 'variable' flux Richardson number were to be used in the Osborn model (2.5), there would still exist an inconsistency between $\widetilde{K}_\rho^{osb}$ and $\widetilde{K}_\rho^*$ due to an inherent inability of R_f to isolate the diabatic process of irreversible mixing from the adiabatic process of reversible stirring (Peltier & Caulfield 2003; Tailleux 2009). Nonetheless, the major discrepancy between $\widetilde{K}_\rho^{osb}$ and $\widetilde{K}_\rho^*$ in figure 2(*d*) should be ascribed to the assumption of *constant* mixing efficiency. The issue of constant versus variable mixing efficiency is a pressing one because, based on (2.23), $\widetilde{K}_\rho^* \propto \mathscr{E}/(1-\mathscr{E})$ which indicates that extreme variations of $\widetilde{K}_\rho^*$ will occur in the near vicinity of $\mathscr{E}=0$. This implies that using $\mathscr{E}=0.17$ (or $\Gamma=0.2$) would overestimate $\widetilde{K}_\rho^*$ by several orders of magnitude if the stratified shear flow is being mixed quite inefficiently (e.g. when the flow is either very weakly or highly stratified).

For the remainder of this paper, we will primarily employ the accurate formulation of diapycnal diffusivity $\widetilde{K}_\rho^*$ obtained from either (2.7) or (2.23). To simplify notation, we will drop the asterisk and simply refer to $\widetilde{K}_\rho$ as the diapycnal scalar diffusivity. The dimensionless notation (denoted by $\widetilde{K}$) will be retained, which implies division by molecular diffusivity, κ, and kinematic viscosity, ν, for $\widetilde{K}_\rho$ and $\widetilde{K}_m$ respectively.

4.3. *Turbulent scalar diffusivity and mixing efficiency*

Figure 3 compiles the data derived from all of the simulations whose characteristics have been tabulated in table 1 and illustrates the variations of $\widetilde{K}_\rho$ with respect to the buoyancy Reynolds number $Re_b=\varepsilon_k/(\nu N^2)$. The data points for each DNS experiment are associated with the fully turbulent phase of the flow evolution which belong to $t\in[t_{3d}^s \quad t_f^e]$. For the purpose of generating this figure (and the similar figures to be discussed in what follows), as noted in § 4.1 we have computed t_f^e based on $Re_b=1$ in order to also include the re-laminarized flow whose turbulent diapycnal diffusivity approaches zero. It is important to mention that not all of the DNS runs have achieved $Re_b=O(10^0$–$10^1)$. Especially for the weakly stratified cases, in which the turbulence suppressing influence of the buoyancy force is considerably weaker than other more highly stratified situations, the turbulence persists for much longer periods (i.e. $t_f^e \gg t_{3d}^s$). Thus in these cases, as will be illustrated later, the simulations require a much greater expenditure of computational resources in order to allow the full influence of the buoyancy effects to activate so as to significantly diminish the turbulence intensity.

The magnitude of diapycnal diffusivity, K_ρ, in the interior of lakes and oceans is typically of $O(\kappa)$ and $O(10^2\kappa)$ respectively (Ivey *et al.* 2008). These values however increase to $O(10^2\kappa)$ (for lakes) and $O(10^4\kappa)$ (for oceans) as we approach abyssal regions with rough topography (Ledwell *et al.* 2000). In addition, there are other 'hot spots', such as coastal zones and equatorial near-surface currents, where K_ρ may be as large. The range of our non-dimensionalized diapycnal diffusivity $\widetilde{K}_\rho=K_\rho/\kappa$ in figure 3 reaches $O(10^3)$ due to the higher molecular Prandtl numbers of $O(10)$ investigated in this work. For instance, if we take the molecular diffusivity of temperature-stratified seawater as $\kappa_T=10^{-7}$ m^2 s^{-1}, and its molecular Prandtl

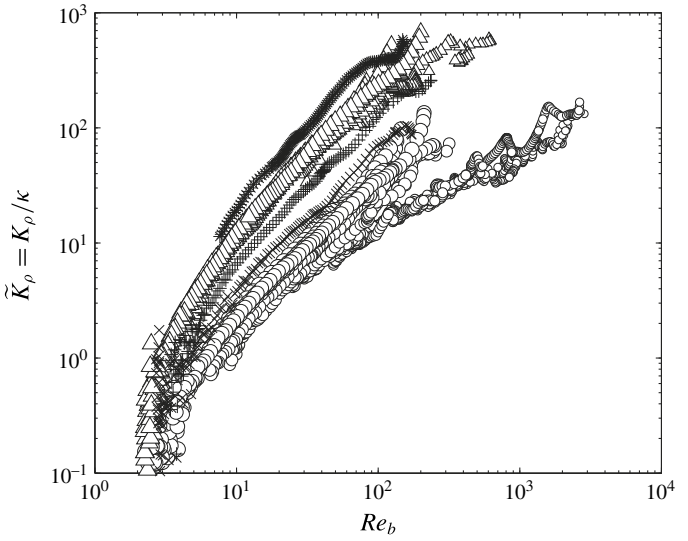

FIGURE 3. Variation of non-dimensional diapycnal scalar diffusivity $\widetilde{K}_\rho = K_\rho/\kappa$ with the buoyancy Reynolds number $Re_b = \varepsilon_k/\nu N^2$. $\widetilde{K}_\rho$ has been calculated based on (2.7) due to WD96 or our proposed reformulation in (2.23). This figure illustrates a compilation of 24 DNS runs with various combinations of Re, Ri_0 and Pr as tabulated in table 1 and are all associated with the fully turbulent phase of the flow evolution as defined in the text for the KH ansatz; ○, $Pr = 1$; ×, $Pr = 2$; +, $Pr = 4$; △, $Pr = 8$; ∗, $Pr = 16$. Bigger (smaller) markers are associated with higher (lower) Ri_b. Time increases roughly as Re_b decreases.

number to be $Pr = 7$, then the DNS results at $Pr = 8$ in figure 3 suggest that the range of dimensional turbulent diapycnal diffusivity K_ρ reaches approximately $K_\rho \approx 10^{-4}$ m^2 s^{-1}, which lies within the range of estimated turbulent diffusivities in the ocean (Waterhouse *et al.* 2014).

It is useful to compare our DNS results for K_ρ with those reported in figure 1 of SKIF. Based on their numerical results, SKIF categorized the turbulent mixing into three regimes in terms of Re_b: (i) a 'molecular' regime for $Re_b < 7$, (ii) an 'intermediate' regime corresponding to stationary turbulence for $7 < Re_b < 100$ and (iii) an 'energetic' regime for $Re_b > 100$. Note that the DNS results of SKIF are based on a constant value of $Pr = 0.72$ whereas figure 3 includes DNS runs with various Pr. Thus we have normalized $\widetilde{K}_\rho$ by Pr in figure 4, thereby effectively plotting K_ρ/ν (where K_ρ is dimensional) versus Re_b. Despite the intrinsic differences between the geophysically relevant flows we are analysing (e.g. see Smyth *et al.* 2001 for a discussion concerning such relevance) and the more highly constrained homogeneous problem of SKIF and their calculation methodology for K_ρ (which is based on the mixing length model), there exist striking similarities between our figure 4 and their figure 1. In order to better highlight the similarities, figure 4 also includes the approximation of diapycnal diffusivity due to Osborn (1980), $\widetilde{K}_\rho^{osb} = 0.2\,Pr\,Re_b$, as discussed in § 2.1.2 (shown by a dashed line), as well as the line of best fit taken from SKIF for their 'energetic' regime, namely $\widetilde{K}_\rho = 2\,Pr\,Re_b^{1/2}$ (shown by a solid line).

Similar to SKIF, a transition to a noticeably different regime of turbulent mixing is also evident in our simulations at approximately $Re_b = O(10^2)$. However, unlike SKIF,

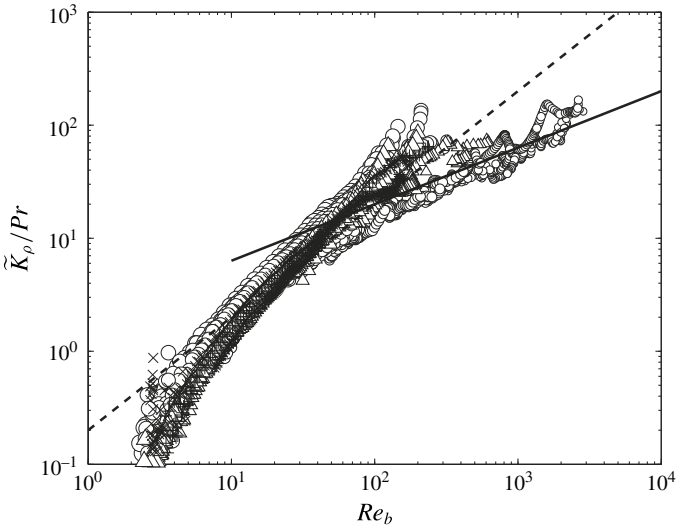

FIGURE 4. Variation of non-dimensional diapycnal scalar diffusivity $\widetilde{K}_\rho$ (after being normalized by Pr) with $Re_b = \varepsilon_k/\nu N^2$ for all of the DNS cases investigated. – – –, Osborn formula of $\widetilde{K}_\rho = 0.2\,Pr\,Re_b$ (2.5); ——, $\widetilde{K}_\rho = 2\,Pr\,Re_b^{1/2}$ best fit in SKIF. For symbol conventions, refer to figure 3.

our data that reside in the range of $7 < Re_b < 100$ are associated with dynamically evolving flow fields that bear minimal resemblance to 'stationary' turbulence (e.g. see our figure 1). In fact, our results show that the major difference between DNS cases whose Re_b exceeds 100 and those below $Re_b \approx 100$ is associated with their stratification level. The onset of turbulence for the 'weakly stratified' simulations with the minimum initial Richardson numbers of $Ri_0 = 5 \times 10^{-3}$, 1×10^{-2} and 2×10^{-2} (at $Pr = 1$ and $Re = 6000$), occurs well above $Re_b = 100$, while most of the other runs that are relatively more 'strongly stratified' belong to the regime with $Re_b < 100$. For these reasons, it is more instructive to denote the 'intermediate' and 'energetic' regimes of SKIF respectively as corresponding to '*buoyancy-dominated*' and '*shear-dominated*' mixing regimes based on the competing dominance of density stratification and velocity shear as characterized by Ri_b. Furthermore, the 'molecular' regime in which molecular diffusion dominates the turbulent diffusion (i.e. $\widetilde{K}_\rho < 1$) seems to occur at $Re_b = O(10^0)$, with $Re_b = 7$ as a reasonable characteristic value (in agreement with SKIF) for the transition from the 'molecular' regime to the 'buoyancy-dominated' regime.

None of the above transition points should be regarded as precise because a transition from one regime of mixing to another would be expected to occur smoothly. For example, Bouffard & Boegman (2013) suggested a new Pr-dependent regime between the 'molecular' and the 'intermediate' regimes of SKIF in which the diapycnal diffusivity varies with a different slope of $Re_b^{3/2}$. Our DNS results also support this slight change in the slope for this limited range of Re_b (especially for the higher-Pr simulations). However, we believe that any careful attempt to estimate the variations of $\widetilde{K}_\rho/Pr$ in figure 4 in terms of a power-law relation (i.e. $\widetilde{K}_\rho/Pr \propto Re_b^\alpha$) would essentially lead to a piecewise fragmentation of $\widetilde{K}_\rho$ as a function of Re_b which would be of limited utility.

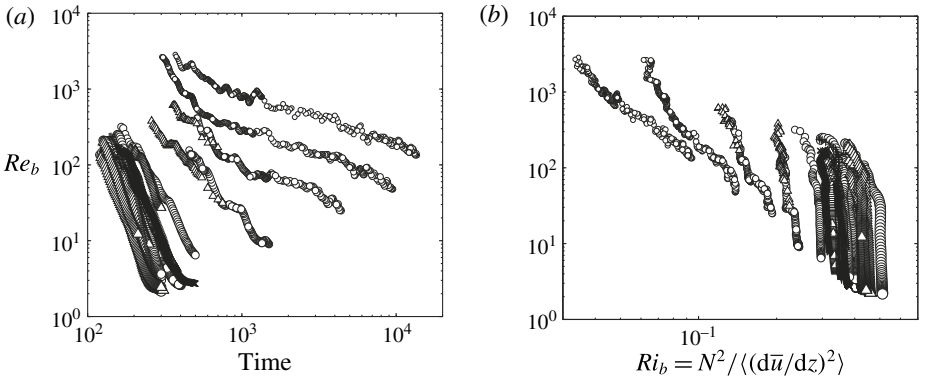

FIGURE 5. (*a*) Time variations of the buoyancy Reynolds number $Re_b = \varepsilon_k/(\nu N^2)$ for all of the DNS cases investigated. (*b*) The interdependence of Re_b and vertically averaged Richardson number $Ri_b = N^2/S^2$ for all of the DNS cases investigated. For symbol conventions, refer to figure 3.

Furthermore, a careful inspection of figure 4 indicates that at a fixed Re_b, the gradient Richardson number of the flow (in a vertically averaged sense as defined in (1.2)) as well as the molecular Prandtl number Pr also affect $\widetilde{K}_\rho/Pr$. However, to a first order, the buoyancy Reynolds number is clearly the single most relevant dimensionless parameter that may be employed to characterize the turbulent mixing and its associated diffusivity. This conclusion extends the suggestions of SKIF regarding the importance of Re_b in the categorization of mixing regimes, to a more geophysically relevant example of turbulent mixing. An important caveat to this, however, is that as a consequence of their focus upon the simplified model of homogeneous stratified sheared turbulence, SKIF did not capture the connection between Re_b and Ri_b in the transition from the 'buoyancy-dominated' to the 'shear-dominated' mixing regimes, to which we will return for further discussion in what follows.

The reason that we are not accounting for the effect of the molecular Prandtl number on K_ρ in figure 4 is that our available DNS data are insufficient to allow the full characterization of its effect in both regimes. It was shown in Salehipour *et al.* (2015) however that for the 'buoyancy-dominated' mixing regime, the irreversible mixing efficiency decreases as Pr increases, due to the emergence of smaller-scale secondary instabilities through which the three-dimensional turbulence is injected directly into smaller scales. As a result, the downscale cascade of energy is accelerated and furthermore the entrainment of the higher-speed flow, characteristic of the flanks of the shear layer, into the mixing region is weakened and consequently the irreversible mixing becomes less efficient. Based on our discussion in § 2.3, because mixing efficiency decreases for higher values of Pr, the diapycnal diffusivity is also expected to decrease, which is also manifested in figure 4. Nevertheless, characterizing the effect of Pr on mixing efficiency in the 'shear-dominated' regime, where the suppressing effects of buoyancy are relatively weak, requires long time integrations of an already high-resolution DNS, which is excessively expensive. The long time integrations required for the DNS cases in the 'shear-dominated' regime may be inferred from figure 5(*a*) which illustrates the time evolution of Re_b for all the simulations.

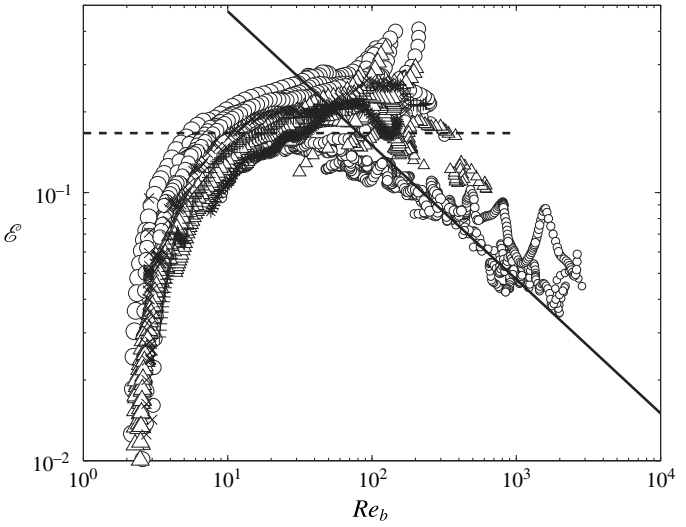

FIGURE 6. Variation of irreversible mixing efficiency $\mathscr{E}$ (2.22) with $Re_b = \varepsilon_k/\nu N^2$ for all of the DNS cases investigated. – – –, The constant value of $R_f = 1/6$ associated with $\gamma = 0.2$ in the Osborn formula (2.5); ——, $R_f = 1.5Re_b^{-1/2}$ best fit in SKIF. For symbol conventions, refer to figure 3.

Based on the reformulation of K_ρ proposed in (2.23) (following WD96), the variation in scalar diffusivity as a function of Re_b has to be associated with the inherent and intricate dependences of the mixing efficiency $\mathscr{E}$ (or flux coefficient Γ) on the governing control parameters. It is therefore expected, based on mathematical grounds (and not length scale arguments), that for any stratified turbulent flow subject to the Boussinesq approximations and at sufficiently high Re, $\widetilde{K}_\rho$ may vary linearly with Re_b if and only if the mixing efficiency is independent of Re_b. Likewise $\widetilde{K}_\rho$ would conceivably vary like $Re_b^{1/2}$ (or $Re_b^{3/2}$) if and only if $\mathscr{E}$ or Γ were to vary like $Re_b^{-1/2}$ (or $Re_b^{1/2}$). In other words, a simple picture for $\mathscr{E}$ which emerges after considering these piecewise power-law expressions and the formulation for K_ρ in (2.23), suggests that the mixing efficiency would be expected to increase within the lower range of Re_b like $\mathscr{E} \propto Re_b^{1/2}$, stay constant for some intermediate range of Re_b and then decrease for higher values of Re_b like $\mathscr{E} \propto Re_b^{-1/2}$. Of course as mentioned previously for K_ρ, $\mathscr{E}$ is expected to deviate from the above approximate view such that it would vary smoothly from one regime to the other.

Figures 6 and 7 illustrate the dependence of the irreversible mixing efficiency, $\mathscr{E}$ defined in (2.22), on respectively the buoyancy Reynolds number Re_b, and the vertically averaged gradient Richardson number Ri_b, again for the entire database of DNS experiments. Based on these figures, $\mathscr{E}$ varies (in an ensemble-averaged sense) non-monotonically with both Re_b and Ri_b, which agrees with previous DNS, experimental and observational investigations (Caulfield & Peltier 2000; Ivey *et al.* 2008; Bouffard & Boegman 2013; Mashayek *et al.* 2013). The bulk Richardson number associated with the transition between 'shear-dominated' and 'buoyancy-dominated' mixing regimes in figure 7 occurs at approximately $Ri_b \sim 0.2$ (also see figure 5*b* which illustrates the variations of Re_b with Ri_b). Note that this transition value appears to be independent of the Ri_b definition as discussed in

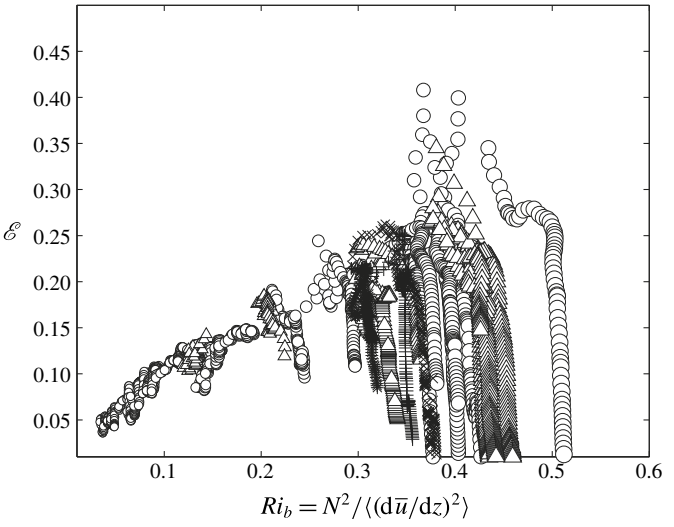

FIGURE 7. Variation of irreversible mixing efficiency $\mathscr{E}$ (2.22) with $Ri_b = N^2/S^2$ for all of the DNS cases investigated. For symbol conventions, refer to figure 3.

appendix B. Based on figure 7 (or figure 5*b*), in the 'buoyancy-dominated' and 'molecular' regimes, Ri_b remains almost constant as the turbulence ages (i.e. as the turbulence intensity measured by Re_b decays). This is in clear contrast to the 'shear-dominated' regime in which Ri_b increases as turbulence ages up to the transition value of about 0.2 after which it tends to stay constant. Note that the above observation concerning the characteristics of the 'shear-dominated' regime contradicts SKIF (e.g. see their figure 14*a*) who argued that Ri_b cannot exclusively characterize diapycnal diffusivity in this regime. It should be emphasized that the DNS study of SKIF is designed such that the background shear and stratification are forced to remain constant throughout the flow evolution, whereas in our DNS study Ri_b evolves freely (see e.g. figure 5*b*).

In the 'buoyancy-dominated' regime, the mixing efficiency may not be characterized solely by Ri_b because it becomes approximately constant while $\mathscr{E}$ decreases to zero. Nonetheless, it is particularly interesting to notice that the dependence of the mixing efficiency on Ri_b is evident in the 'shear-dominated' regime (see figure 7). This implies that in the 'shear-dominated' regime, Re_b and Ri_b are no longer independent parameters and there could be a unique mapping between them. Such a mapping would be important because it could suggest that for the 'shear-dominated' mixing regime, simple parameterizations for the viscous subrange energy dissipation ε_k may be obtained in terms of the Richardson number of the mean flow; a task which has long been attempted (e.g. see Large, McWilliams & Doney 1994). However, finding the unique relations between Re_b and Ri_b would require a much expanded set of DNS analyses of 'shear-dominated' mixing, which is beyond the scope of the current paper.

4.4. *The eddy diffusivity for momentum*

Focusing next on the eddy diffusivity for momentum K_m, figure 8 shows the variations of non-dimensional momentum diffusivity $\widetilde{K}_m = K_m/\nu$ with Re_b for the complete set

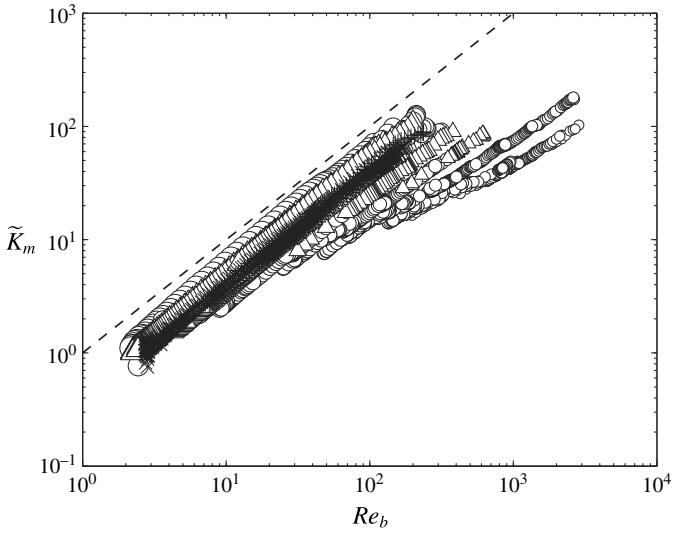

FIGURE 8. Variation of non-dimensional momentum diffusivity $\widetilde{K}_m = K_m/\nu$ with $Re_b = \varepsilon_k/\nu N^2$. $\widetilde{K}_m$ has been calculated based on our proposed formulation in (3.16) for all of the DNS cases investigated. For symbol conventions, refer to figure 3.

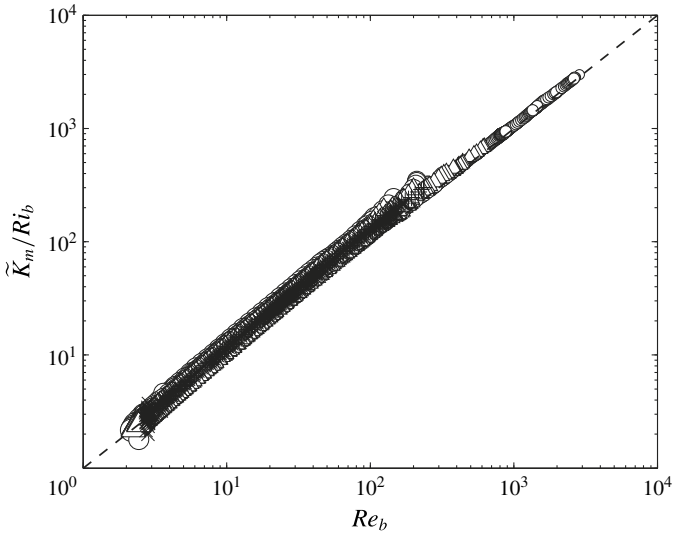

FIGURE 9. Similar to figure 8 but $\widetilde{K}_m$ is now normalized by $Ri_b = N^2/S^2$.

of DNS runs. $\widetilde{K}_m$ is calculated based on the representation proposed in (3.16) which distinguishes between the reversible and irreversible shear production fluxes. Note that $\widetilde{K}_m$ in (3.16) may be re-written as $\widetilde{K}_m = (1 - \mathscr{E})^{-1} Ri_b\, Re_b$ to explicitly include the buoyancy Reynolds number, Re_b. Normalizing $\widetilde{K}_m$ by Ri_b would collapse the data onto an almost straight line as illustrated in figure 9 which points to the fact that $\widetilde{K}_m/Ri_b \approx Re_b$ for the entire range of Re_b. This is not surprising considering the role of mixing efficiency in the formulation for $\widetilde{K}_m$.

In contrast to the scalar diffusivity $\widetilde{K}_\rho$ in (2.23), which is proportional to $\mathscr{E}/(1-\mathscr{E})$, the momentum diffusivity $\widetilde{K}_m$ in (3.16) is proportional to $1/(1-\mathscr{E})$. This important difference implies that K_m varies monotonically for the range of accepted mixing efficiencies associated with shear-induced stratified turbulence (i.e. $0 \leqslant \mathscr{E} \leqslant 0.3$). This should be contrasted with the dependence of K_ρ on mixing efficiency, which undergoes abrupt variations in the vicinity of $\mathscr{E} = 0$. Therefore the accurate representation of mixing efficiency, whether it is defined as either the flux Richardson number R_f or the mixing efficiency $\mathscr{E}$, or even assumed to be constant at 0.17 or 0.2, would affect K_m minimally while it might cause orders of magnitude difference in K_ρ. As a result, for turbulent flows with relatively small values of $\mathscr{E}$, the estimation of Crawford (1982) for K_m (reproduced in (3.2)) should be recognized to be as accurate as our proposed expression (3.16).

4.5. *Turbulent Prandtl number*

Numerical modelling of larger-scale fluid phenomena rely on subgrid-scale parameterizations to prescribe both K_ρ and K_m. It is often assumed that $Pr_t = K_m/K_\rho = 1$ on the basis of the notion that, for a turbulent flow, the effective diffusivities of mass and momentum are identical (e.g. see the parameterization for tidally induced mixing proposed by Klymak *et al.* 2010). A further example involves ocean general circulation models such as the Parallel Ocean Program (POP) (which is currently being employed as the ocean component of the Community Earth System Model (CESM1) of the US National Center for Atmospheric Research). In this model, the turbulent mixing associated with tidally induced internal waves in the ocean interior is parameterized following St. Laurent *et al.* (2002) and Simmons *et al.* (2004), in which K_ρ is assumed to diminish exponentially as a function of increasing height above the sea floor, with horizontal variations assumed to be related to the expected changes in tidal dissipation with longitude and latitude associated with bottom topography, and which also includes an assumption that the value of the flux coefficient is $\gamma = 0.2$ as suggested in Osborn (1980). The eddy viscosity K_m is then determined on the basis of the assumption that the turbulent Prandtl number is equal to 10 (Danabasoglu *et al.* 2012). The latter assumption appears to be inspired solely on the basis of the limited early observations of Peters, Gregg & Toole (1988) as reviewed in Large *et al.* (1994).

Figures 10 and 11 show the variations of $Pr_t = K_m/K_\rho$ as defined in (3.18) as a function of Re_b and Ri_b respectively. Similar to our argument for K_ρ and thus $\mathscr{E}$, inspection of figures 10 and 11 suggests that Pr_t may be estimated, to first order, solely on the basis of Re_b for all types of mixing regimes. Pr_t decreases monotonically with Re_b as illustrated in figure 10. In the 'molecular' regime, since the mixing efficiency and the scalar diffusivity approach zero while the momentum diffusivity remains finite, Pr_t becomes extremely large of $O(10^2)$. However, in the 'buoyancy-dominated' regime, Pr_t decreases to $O(10^0)$ after which its decrease becomes modest such that in the 'shear-dominated' regime Pr_t may be approximated by a value close to unity ($Pr_t \approx 0.8$–1). While assuming $Pr_t = 1$ seems to be more appropriate for the 'shear-dominated' mixing regime, the choice of $Pr_t = 10$ is likely to be more relevant to the 'molecular' regime and values in between, $1 \lesssim Pr_t \lesssim 10$, are likely to belong to the 'buoyancy-dominated' regime. Unfortunately, the sensitivity of the larger-scale numerical models to the choice of Pr_t has not been subjected to detailed analysis. As such, it is unclear whether a more refined characterization of Pr_t (e.g. based on molecular Pr or Ri_b) would be warranted at the current stage

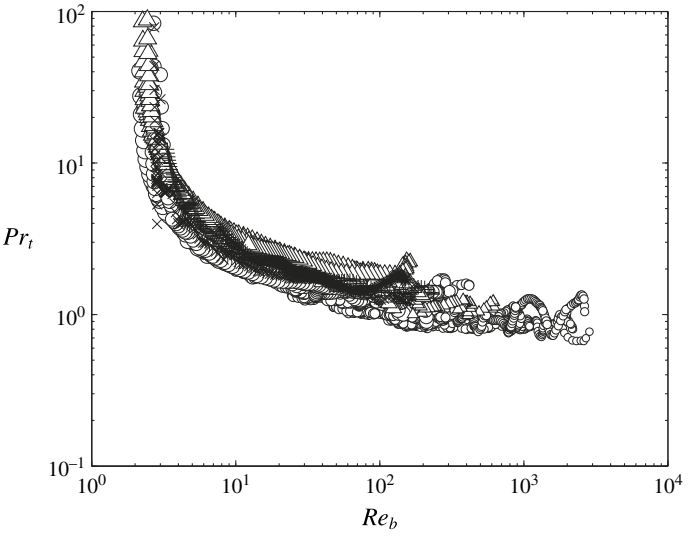

FIGURE 10. Variation of turbulent Prandtl number $Pr_t = K_m/K_\rho$ with $Re_b = \varepsilon_k/\nu N^2$. Pr_t has been calculated based on our proposed formulation in (3.18) for all of the DNS cases investigated. For symbol conventions, refer to figure 3.

of large-scale model development. What is clearly evident, however, is that the parameterizations embedded in larger scale numerical models should not simply rely on the assumption of a constant value of Pr_t. They should instead attempt to identify the relevant mixing regime and therefore employ a dynamic representation of Pr_t based on Re_b. Because the maximum Re_b in oceanic flows extends to $Re_b = O(10^5)$, further DNS analysis of weakly stratified turbulence will be necessary in order to develop an appropriate parameterization for Pr_t.

5. Summary and conclusion

Diapycnal diffusivity, K_ρ, is a means of characterizing the irreversible vertical mixing of density-stratified turbulence and plays a crucial role in estimating the amount of upwelling associated with the abyssal waters required for the closure of the meridional overturning circulation of the oceans (Wunsch & Ferrari 2004). Yet, it seems that unambiguous and accurate representations of this quantity have not been employed in the subgrid-scale parameterizations embedded in the large-scale circulation models of the oceans or for the purpose of interpreting oceanic microstructure measurements. Rather, canonical models due to Osborn & Cox (1972) or Osborn (1980) have been favoured for such purposes despite their well-established limitations, apparently due to their simplicity. In this paper, we have reformulated the diascalar representation of K_ρ due to WD96 into a more familiar 'Osborn-like' expression (see (2.23)) which essentially inherits the simplicity of the Osborn formula but avoids its limiting assumptions. The main difference however is that the correct representation of instantaneous mixing efficiency, $\mathscr{E}$, should replace the flux Richardson number R_f in the Osborn formula. Such a slight modification accurately distinguishes the irreversible process of mixing (at smallest scales) from the reversible process of stirring (at larger scales) as required for a correct representation of K_ρ. Furthermore, the new formulation formally generalizes the length scale arguments of

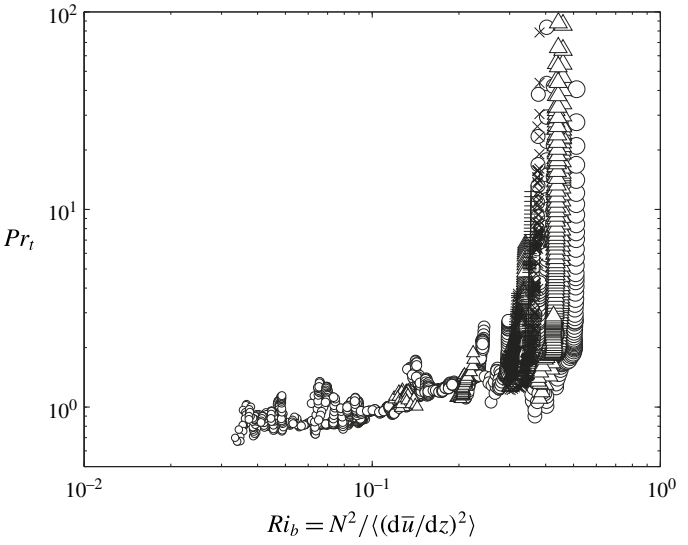

FIGURE 11. Similar to figure 10 but showing the variations of Pr_t with respect to $Ri_b = N^2/S^2$.

Ivey & Imberger (1991) and implies on mathematical grounds that K_ρ must depend on Re_b for a Boussinesq fluid at high Reynolds numbers.

Of equal importance for modelling purposes is an accurate representation of turbulent momentum diffusivity K_m (e.g. turbulence closure schemes in the Reynolds-averaged Navier–Stokes (RANS) equations rely on K_m as well as K_ρ). This quantity is commonly calculated based on estimations of K_ρ and the turbulent Prandtl number, $Pr_t = K_m/K_\rho$, which is occasionally assumed to be unity. Nonetheless, reasonable definitions for K_m and hence Pr_t which distinguish between reversible and irreversible processes for any Boussinesq stratified fluid have apparently been previously unavailable in the literature. Therefore we have notionally decomposed the total shear production rate of TKE into reversible and irreversible components, and have argued that only the 'irreversible' component should be employed in the definition of K_m (see (3.13)). We have defined this irreversible component to be manifested at smallest scales in terms of the rate of energy dissipation from both kinetic energy and available potentially energy reservoirs into the internal energy of the flow. For an incompressible Boussinesq fluid, this notional description of irreversible shear production can be interpreted as the sum of irreversible mixing and viscous dissipation (see (3.10)). We have derived alternative but otherwise identical formulae for K_m in (3.16) and (3.17) after replacing the irreversible mixing term ($\mathscr{M}$) by different expressions based on mixing efficiency or diapycnal diffusivity. This result extended the previous homogeneous formulation of Venayagamoorthy & Stretch (2010) to any general incompressible Boussinesq flow with high Reynolds numbers. Similarly for $Pr_t = K_m/K_\rho$, an adequately precise formulation has been proposed in (3.18).

In order to demonstrate the utility of the proposed formulations, we have employed a large set of DNS analyses to simulate the evolution of a Kelvin–Helmholtz instability and focused entirely on the three-dimensional turbulence that it engenders once the nonlinear wave 'breaks'. The DNS experiments include high initial Reynolds

numbers (mostly at $Re = 6000$) and a wide range of minimum initial Richardson numbers, Ri_0, for fluids characterized by several different values of the molecular Prandtl number Pr (mostly at $Pr = 1, 8$). The reported simulations represent an unprecedented series of DNS runs in the literature of stratified turbulence due to the computationally intensive requirements for such analyses. These requirements are either imposed by the high resolution demands of high-Re and high-Pr simulations or by the long-time integrations of low-Ri_0 simulations. We have also commented on the differences between the results we have obtained for turbulence generated by the breaking Kelvin–Helmholtz wave process and those for an alternative common configuration that has often been employed previously in investigations of stratified turbulence, which is referred to as homogeneous stratified and sheared turbulence. In those studies the background shear and stratification are forced to remain constant throughout flow evolution and thus the BPE does not rise monotonically as a result of irreversible mixing. Furthermore, by using the new formulation for K_ρ, we have proposed an expression in appendix A for calculating $\mathscr{E}$ in those studies instead of R_f.

Our accurate calculations of K_ρ have been compared with estimations based on the Prandtl mixing length theory, and both the Osborn–Cox and the Osborn formulae. Among these approximate models the Osborn–Cox formula was shown to provide the closest estimates, which could be further improved by introducing an *ad hoc* multiplicative factor of two. The major issue that arises by using a conventional Reynolds decomposition for representing the perturbation density flux in the mixing length model is that the irreversible representation of small-scale mixing becomes entangled with the reversible contributions, thus rendering the resulting estimates of diapycnal flux either negative on occasion or simply inaccurate. Even after employing a triple Reynolds decomposition, although the density flux seems to be positive-definite, the resulting diapycnal diffusivity systematically underestimates K_ρ which may suggest that the full diapycnal flux ϕ_d includes other irreversible processes that are unaccounted for by the three-dimensional density flux alone.

Given the close similarity between our proposed expression for K_ρ and that of the Osborn formula, the main problem with the way that this expression is currently being used relates to the assumption of a fixed mixing efficiency. This problem becomes particularly obvious in the limit of either low or high values of stratification (at a constant background shear) where the mixing efficiency is much lower than 0.2. Of course, the second issue, whose adverse impacts on estimating K_ρ are not as large, concerns the fact that R_f contains an inaccurate representation of irreversible mixing and thus leads to similar deficiencies, already mentioned for the mixing length model (in which the turbulent density flux was employed in lieu of the diapycnal flux). It is critical to emphasize that following the new formulation of K_ρ, the stationary assumption (among other assumptions) made by Osborn (1980) and also suggested by SKIF (in their intermediate range of Re_b to recover the Osborn predictions of K_ρ) is not required to arrive at a correct calculation of diapycnal diffusivity.

We have also investigated the dependence of K_ρ, K_m, $\mathscr{E}$ and Pr_t on the buoyancy Reynolds number (a measure of turbulent intensity) defined as $Re_b = \varepsilon_k/(\nu N^2)$ and raised the question as to whether this single parameter is able to uniquely characterize these important quantities. A compilation of all of our DNS data reinforces the previous categorizations of turbulent mixing based on Re_b as first suggested by SKIF. Unlike SKIF, who force the background Richardson number to remain equal to its initial value, in the KH ansatz this parameter is free to evolve as the shear and density layers expand. Therefore in this context we have found that turbulent mixing

may also be exclusively characterized by a measure of relative strength of the mean flow stratification and mean flow shear as, for example, captured by the gradient Richardson number, Ri_b (in a vertically averaged bulk sense), for high values of Re_b, implying a unique mapping between Ri_b and Re_b.

The period of flow evolution in the KH ansatz upon which we have focused begins with high values of Re_b and continues with a gradual decay in turbulence intensity (and thus Re_b) until the molecular diffusion dominates and turbulence has decayed. Within this period, Ri_b may either remain constant or increase depending on the turbulent regime of mixing. High values of $Re_b \gtrsim O(10^2)$ were only realized in weakly stratified simulations with $Ri_b \lesssim 0.2$. On the other hand, the intermediate range of $Re_b \approx O(10^1\text{–}10^2)$ occurred with $Ri_b \gtrsim 0.2$. As a result, we have re-labelled the 'intermediate' and 'energetic' regimes of mixing in SKIF as 'buoyancy-dominated' and 'shear-dominated' mixing regimes respectively, based on the dominance of density stratification relative to velocity shear rather than the range of a single parameter (i.e. Re_b). It is interesting to note that, most recently in the laboratory experiments on stratified turbulent shear flows down an inclined duct, Meyer & Linden (2014) have also suggested a transition to a different regime of mixing at approximately $Re_b = 100$ based on the length scales of the flow structures observed in their shadowgraph images. In addition, evidence for a different regime beginning at $Re_b \sim 100$ has also been found in isotropy statistics (Smyth & Moum 2000*a*) and in differential diffusion studies (Smyth *et al.* 2005).

Despite the inherent differences between our DNS study of inhomogeneously stratified shear-induced turbulence due to the breaking of a KH instability and the more idealized simulations of homogeneously stratified and sheared turbulence in SKIF (see appendix A regarding such differences), our figure 4 shares a striking similarity with figure 1 of SKIF. Based on this similarity, we may conjecture that at high Reynolds numbers, shear-driven stratified turbulence may be represented by a universal characteristic behaviour regardless of the details of the route to turbulence. This point must remain as a plausible speculation that deserves to be the subject of further investigation.

In the 'shear-dominated' regime as the turbulence ages and hence its intensity decays, mixing efficiency increases and concurrently Ri_b increases. These characteristics are in contrast with those of the 'buoyancy-dominated' mixing in which the mixing efficiency declines as turbulent intensity decays and moreover Ri_b stays relatively constant throughout the turbulent phase. In fact a DNS run with a relatively low Ri_0 begins its turbulent phase in the 'shear-dominated' regime, in which the mixing efficiency as well as Ri_b continue to rise secularly until the mixing regime undergoes a transition at approximately $Re_b \sim O(10^2)$ or $Ri_b \sim 0.2$, after which the irreversible mixing efficiency declines until the flow re-laminarizes while Ri_b stays constant. Thus mixing efficiency varies non-monotonically with both Re_b and Ri_b but in distinctly different ways. Furthermore, based on our DNS results, we have argued that to first order, both K_ρ and $\mathscr{E}$ may be characterized by only a single parameter (e.g. Re_b). Because in the 'shear-dominated' regime, Re_b and Ri_b seem to be closely related, Ri_b could be equivalently employed for such a purpose in this regime only.

Insofar as the momentum diffusivity is concerned, K_m varies monotonically with $\mathscr{E}$ because $K_m \propto 1/(1-\mathscr{E})$, where $0 \leqslant \mathscr{E} \leqslant 0.3$ in most shear-stratified flows. Therefore, so long as $\mathscr{E}$ is not close to unity, K_m should be almost independent of the choice of mixing efficiency (whether assumed constant or variable, or represented by $\mathscr{E}$ or R_f). In fact $K_m \approx Re_b Ri_b = \varepsilon_k/(\nu S^2)$ seems to be a satisfactory approximation for the entire range of Re_b and for all the DNS cases studied herein. This may be contrasted

to the delicate dependence of K_ρ on $\mathscr{E}$ (i.e. $K_\rho \propto \mathscr{E}/(1-\mathscr{E})$). The above conclusion implies that the momentum diffusivity may be considered to be equivalent to the shear Reynolds number, $K_m \approx \varepsilon_k/(\nu S^2)$, with the constant of proportionality being unity.

We have also shown that, to first order, Pr_t may be approximated solely by Re_b for all types of mixing regimes. Therefore, the range of Pr_t critically depends on the mixing regime at which the turbulent is active. In the 'molecular' regime, Pr_t becomes irrelevant since K_ρ approaches zero. In the 'buoyancy-dominated' regime, Pr_t is in the range $Pr_t \approx 1$–5, while it asymptotes to $Pr_t \approx 0.8$–1 for larger values of Re_b in the 'shear-dominated' regime.

Acknowledgements

We are grateful to B. Smyth and the anonymous reviewers whose constructive comments have further improved the paper. Computations were performed on the BG/Q supercomputer at the SciNet HPC Consortium of the University of Toronto. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund – Research Excellence; and the University of Toronto. The research of W.R.P. at the University of Toronto is sponsored by NSERC Discovery Grant A9627.

Appendix A. A note on studies of homogeneous stratified turbulence

Idealized studies of turbulence in which the turbulent flow is assumed to be homogeneous (Batchelor 1953), include sheared unstratified, stratified unsheared and stratified sheared turbulent flows (see Holt, Koseff & Ferziger 1992 for a review). In the case of homogeneous stratified and sheared turbulence, the background stratification and shear are both forced to remain time-invariant throughout the course of flow evolution. Here we intend to (i) clarify the relationship between the concept of irreversible mixing which is associated with a rise in the BPE in more realistic configurations and that in these idealized homogeneous studies, and furthermore (ii) propose a formulation for mixing efficiency $\mathscr{E}_h$ which could be employed instead of the flux Richardson number R_f, again in the context of homogeneous stratified turbulence analyses.

It was noted in § 2.2 that for a closed system in which the mass flux across the bounding surfaces is zero (for example by assuming a large enough control volume which isolates the physical process that leads to turbulent mixing away from the boundaries), the surface flux terms in (2.14) would vanish (i.e. $\mathscr{F}_{adv} = \mathscr{F}_{diff} = 0$) and therefore the BPE would increase monotonically in time as a consequence of irreversible mixing. Nonetheless, in the studies of homogeneous stratified and sheared turbulence, because the vertical boundary conditions are chosen to be shear-periodic (Gerz, Schumann & Elghobashi 1989), the system is not closed and hence $\mathscr{F}_{adv}$ and $\mathscr{F}_{diff}$ are non-zero. In these studies (e.g. see Gerz *et al.* 1989; Holt *et al.* 1992; SKIF), since the BPE is kept constant in time (i.e. $\mathrm{d}\mathscr{P}_B/\mathrm{d}t = 0$), (2.14) becomes,

$$\mathscr{M}_h = \mathscr{F}_{adv} + \mathscr{F}_{diff} - D_p. \tag{A 1}$$

Note that the amount of irreversible mixing in these studies (denoted by $\mathscr{M}_h$ above) cannot be obtained by computing $\mathrm{d}\mathscr{P}_B/\mathrm{d}t$, in contrast to the common practice for realistic flows in which a global adiabatic rearrangement of the density field (as described in Winters *et al.* 1995 and Caulfield & Peltier 2000) is required. Alternatively, $\mathscr{M}_h$ should be either directly calculated as per its definition in (2.17) or

inferred indirectly from (A 1). Either way, it remains unclear how irreversible mixing $\mathscr{M}_h$ can be assumed to be simply equal to the buoyancy flux for the purposes of calculating K_ρ as in SKIF. In their DNS study of homogeneous stratified and sheared turbulence, SKIF used the Prandtl mixing length model for calculating K_ρ, thereby implicitly assuming that the buoyancy flux is equivalent to irreversible mixing (or likewise implicitly assuming that R_f is equivalent to $\mathscr{E}$), which necessarily resulted in some negative values for R_f (see their figure 3).

Furthermore, in the same context of homogeneous stratified flows with a uniform background density gradient, it was shown by Venayagamoorthy & Stretch (2006) that the Osborn–Cox model (2.6) becomes equivalent to the exact formulation of Winters–D'Asaro (2.7). Therefore on the basis of the analysis provided in § 2.3, these models must also be equivalent to our newly proposed formulation in (2.23). Thus, it is possible to infer the exact value of mixing efficiency indirectly by setting $\widetilde{K}_\rho^{cox} = \widetilde{K}_\rho^{*}$ which results in:

$$\Gamma_h = \frac{1}{Re_b Pr}\left[\frac{\langle|\overline{\nabla\theta'|^2}\rangle}{\langle(\mathrm{d}\overline{\theta}/\mathrm{d}z)^2\rangle}\right]. \tag{A 2}$$

Using (A 2), the accurate representation of mixing efficiency in homogeneous stratified turbulence, $\mathscr{E}_h$, may be found as $\mathscr{E}_h = \Gamma_h/(1+\Gamma_h)$. In fact this definition of $\mathscr{E}_h$ was motivated in Venayagamoorthy & Stretch (2010) (their equation (2.10)) based on the original definition in Caulfield & Peltier (2000), although it is confusingly defined as the flux Richardson number! It would be interesting nevertheless to compare $\mathscr{E}_h$ with the common definition of R_f in the studies of homogeneous stratified turbulence (e.g. in SKIF) to investigate their inherent distinctions because we expect that R_f systematically underestimates $\mathscr{E}_h$.

Appendix B. Bulk measures of the gradient Richardson number

The bulk measure of the gradient Richardson number, Ri_b, was introduced in (1.2) and has been employed consistently throughout this paper. A more common definition of this parameter in physical oceanography is, however, based on the vertical shear being first averaged over the shear layer thickness before being squared and is denoted as $Ri_b^{\dagger}$ here. These two definitions are compared below:

$$Ri_b = \frac{g/\rho_0\langle\mathrm{d}\overline{\rho}/\mathrm{d}z\rangle}{\langle(\mathrm{d}\overline{u}/\mathrm{d}z)^2\rangle}, \tag{B 1}$$

$$Ri_b^{\dagger} = \frac{g/\rho_0\langle\mathrm{d}\overline{\rho}/\mathrm{d}z\rangle_\delta}{\langle\mathrm{d}\overline{u}/\mathrm{d}z\rangle_\delta^2} \tag{B 2}$$

in which $\langle f\rangle$ indicates vertical averaging over any arbitrary length scale while $\langle f\rangle_\delta$ specifies vertical averaging over δ, the evolving thickness of the stratified shear layer. This thickness is determined here as the difference between the highest and the lowest vertical positions at which $\mathrm{d}\overline{\rho}/\mathrm{d}z$ becomes greater than a threshold value of 10^{-3}. Furthermore, $\overline{f}$ denotes horizontal averaging. Note that (B 1) reproduces Ri_b in (1.2) with $N^2 = g/\rho_0\langle\mathrm{d}\overline{\rho}/\mathrm{d}z\rangle$ and $S^2 = \langle(\mathrm{d}\overline{u}/\mathrm{d}z)^2\rangle$.

The definition of Ri_b in (B 1) has been preferred in this manuscript over that in (B 2) simply because it is consistent with the method that the volume average of other quantities that involve a ratio, such as $\mathscr{E}$, K_ρ or K_m, is implemented, namely by averaging their numerator and denominator individually. However for completeness, figure 12 illustrates the interdependence of Ri_b defined based on

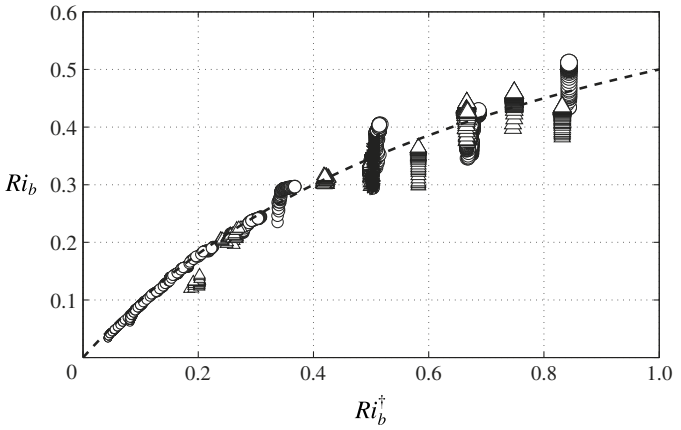

FIGURE 12. The interdependence of two vertically averaged bulk measures of the gradient Richardson number Ri_b as computed based on (B 1) and (B 2) for all of the DNS cases investigated. – –, Nonlinear least-squares fit.

either of the foregoing definitions in (B 1) and (B 2). The dashed line represents the nonlinear least-squares fit which takes the following form:

$$Ri_b = \frac{0.9Ri_b^\dagger}{0.8 + Ri_b^\dagger}. \tag{B 3}$$

The above relationship should serve as a transformation function to convert $Ri_b^\dagger$ to those that have been employed in generating figures 5(*b*), 7 and 11. It is worthwhile mentioning that, at the moment, it is very difficult to estimate either Ri_b or $Ri_b^\dagger$ in both observations and large-scale general circulation models. This difficulty is indeed due to their coarsely resolved velocity estimates as well as the inherent inability to accurately measure the shear layer thickness.

REFERENCES

BARRY, M. E., IVEY, G. N., WINTERS, K. B. & IMBERGER, J. 2001 Measurements of diapycnal diffusivities in stratified fluids. *J. Fluid Mech.* **442**, 267–291.

BATCHELOR, G. K. 1953 *The Theory of Homogeneous Turbulence*. Cambridge University Press.

BLUTEAU, C. E., JONES, N. L. & IVEY, G. N. 2013 Turbulent mixing efficiency at an energetic ocean site. *J. Geophys. Res.* **118** (9), 4662–4672.

BOUFFARD, D. & BOEGMAN, L. 2013 A diapycnal diffusivity model for stratified environmental flows. *Dyn. Atmos. Oceans* **61**, 14–34.

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

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

CRAWFORD, W. R. 1982 Pacific equatorial turbulence. *J. Phys. Oceanogr.* **12** (10), 1137–1149.

DANABASOGLU, G., BATES, S. C., BRIEGLEB, B. P., JAYNE, S. R., JOCHUM, M., LARGE, W. G., PEACOCK, S. & YEAGER, S. G. 2012 The CCSM4 ocean component. *J. Clim.* **25** (5), 1361–1389.

DAVIS, R. E. 1994 Diapycnal mixing in the ocean: the Osborn–Cox model. *J. Phys. Oceanogr.* **24** (12), 2560–2576.

DUNCKLEY, J. F., KOSEFF, J. R., STEINBUCK, J. V., MONISMITH, S. G. & GENIN, A. 2012 Comparison of mixing efficiency and vertical diffusivity models from temperature microstructure. *J. Geophys. Res.* **117**, C10008.

ELLISON, T. H. 1957 Turbulent transport of heat and momentum from an infinite rough plane. *J. Fluid Mech.* **2** (5), 456–466.

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

FISCHER, P. F. 1997 An overlapping schwarz method for spectral element solution of the incompressible Navier–Stokes equations. *J. Comput. Phys.* **133** (1), 84–101.

FISCHER, P. F., KRUSE, G. W. & LOTH, F. 2002 Spectral element methods for transitional flows in complex geometries. *J. Sci. Comput.* **17** (1–4), 81–98.

FISCHER, P. F., LOTTES, J. W. & KERKEMEIER, S. G. 2008 nek5000 Web page. http://nek5000.mcs.anl.gov.

FLEURY, M. & LUECK, R. G. 1994 Direct heat flux estimates using a towed vehicle. *J. Phys. Oceanogr.* **24** (4), 801–818.

GERZ, T., SCHUMANN, U. & ELGHOBASHI, S. E. 1989 Direct numerical simulation of stratified homogeneous turbulent shear flows. *J. Fluid Mech.* **200**, 563–594.

GREGG, M. C. 1987 Diapycnal mixing in the thermocline: a review. *J. Geophys. Res.* **92** (C5), 5249–5286.

HOLT, S. E., KOSEFF, J. R. & FERZIGER, J. H. 1992 A numerical study of the evolution and structure of homogeneous stably stratified sheared turbulence. *J. Fluid Mech.* **237**, 499–539.

IVEY, G. N. & IMBERGER, J. 1991 On the nature of turbulence in a stratified fluid. Part I: the energetics of mixing. *J. Phys. Oceanogr.* **21** (5), 650–658.

IVEY, G. N., WINTERS, K. B. & DE SILVA, I. P. D. 2000 Turbulent mixing in a sloping benthic boundary layer energized by internal waves. *J. Fluid Mech.* **418**, 59–76.

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

JACKSON, P. R. & REHMANN, C. R. 2003 Laboratory measurements of differential diffusion in a diffusively stable, turbulent flow. *J. Phys. Oceanogr.* **33** (8), 1592–1603.

KLYMAK, J. M., LEGG, S. & PINKEL, R. 2010 A simple parameterization of turbulent tidal mixing near supercritical topography. *J. Phys. Oceanogr.* **40** (9), 2059–2074.

LARGE, W. G., MCWILLIAMS, J. C. & DONEY, S. C. 1994 Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.* **32** (4), 363–403.

LEDWELL, J. R., MONTGOMERY, E. T., POLZIN, K. L., ST. LAURENT, L. C., SCHMITT, R. W. & TOOLE, J. M. 2000 Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature* **403**, 179–182.

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

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

LOZOVATSKY, I. D. & FERNANDO, H. J. S. 2013 Mixing efficiency in natural flows. *Phil. Trans. R. Soc. Lond.* A **371**, 20120213.

MADAY, Y., PATERA, A. T. & RØNQUIST, E. M. 1990 An operator-integration-factor splitting method for time-dependent problems: application to incompressible fluid flow. *J. Sci. Comput.* **5** (4), 263–292.

MARTIN, J. E. & REHMANN, C. R. 2006 Layering in a flow with diffusively stable temperature and salinity stratification. *J. Phys. Oceanogr.* **36** (7), 1457–1470.

MASHAYEK, A., CAULFIELD, C. P. & PELTIER, W. R. 2013 Time-dependent, non-monotonic mixing in stratified turbulent shear flows: implications for oceanographic estimates of buoyancy flux. *J. Fluid Mech.* **736**, 570–593.

MASHAYEK, A. & 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.

MEYER, C. R. & LINDEN, P. F. 2014 Stratified shear flow: experiments in an inclined duct. *J. Fluid Mech.* **753**, 242–253.

MOUM, J. N. 1990 The quest for K_ρ–preliminary results from direct measurements of turbulent fluxes in the ocean. *J. Phys. Oceanogr.* **20**, 1980–1984.

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

OSBORN, T. R. & COX, C. S. 1972 Oceanic fine structure. *J. Geophys. Astrophys. Fluid Dyn.* **3** (1), 321–345.

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

PETERS, H., GREGG, M. C. & TOOLE, J. M. 1988 On the parameterization of equatorial turbulence. *J. Geophys. Res.* **93** (C2), 1199–1218.

PRANDTL, L. 1925 Bericht über die entstehung der turbulenz. *Z. Angew. Math. Mech.* **5**, 136–139.

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

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.

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

SIMMONS, H. L., JAYNE, S. R., ST. LAURENT, L. C. & WEAVER, A. J. 2004 Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Model.* **6** (3), 245–263.

SMYTH, W. D., CARPENTER, J. R. & LAWRENCE, G. A. 2007 Mixing in symmetric Holmboe waves. *J. Phys. Oceanogr.* **37**, 1566–1583.

SMYTH, W. D. & MOUM, J. N. 2000*a* Anisotropy of turbulence in stably stratified mixing layers. *Phys. Fluids* **12** (6), 1343–1362.

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

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

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

STILLINGER, D. C., HELLAND, K. N. & VAN ATTA, C. W. 1983 Experiments on the transition of homogeneous turbulence to internal waves in a stratified fluid. *J. Fluid Mech.* **131**, 91–122.

ST. LAURENT, L., NAVEIRA GARABATO, A. C., LEDWELL, J. R., THURNHERR, A. M., TOOLE, J. M. & WATSON, A. J. 2012 Turbulence and diapycnal mixing in drake passage. *J. Phys. Oceanogr.* **42** (12), 2143–2152.

ST. LAURENT, L. C., SIMMONS, H. L. & JAYNE, S. R. 2002 Estimates of tidally driven enhanced mixing in the deep ocean. *Geophys. Res. Lett.* **29** (23), 2106.

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.

TAYLOR, G. I. 1915 Eddy motion in the atmosphere. *Phil. Trans. R. Soc. Lond.* A **215**, 1–26.

VENAYAGAMOORTHY, S. K. & STRETCH, D. D. 2006 Lagrangian mixing in decaying stably stratified turbulence. *J. Fluid Mech.* **564**, 197–226.

VENAYAGAMOORTHY, S. K. & STRETCH, D. D. 2010 On the turbulent Prandtl number in homogeneous stably stratified turbulence. *J. Fluid Mech.* **644**, 359–369.

WATERHOUSE, A. F., MACKINNON, J. A., NASH, J. D., ALFORD, M. H., KUNZE, E., SIMMONS, H. L., POLZIN, K. L., ST. LAURENT, L. C, SUN, O., PINKEL, R., TALLEY, L. D., WHALEN, C. B., HUUSSEN, T. N., CARTER, G. S., FER, I., WATERMAN, S., NAVEIRA GARABATO, A. C., SANFORD, T. B. & LEE, C. M. 2014 Global patterns of diapycnal mixing from measurements of the turbulent dissipation rate. *J. Phys. Oceanogr* **44**, 1854–1872.

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

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

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