J. Fluid Mech. (2009), *vol.* 621, *pp.* 1–21.
doi:10.1017/S0022112008004308 Printed in the United Kingdom

Theory for differential transport of scalars in sheared stratified turbulence

P. RYAN JACKSON[1]† AND CHRIS R. REHMANN[2]

[1]Applied Ocean Physics and Engineering Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA
pjackson@whoi.edu

[2]374 Town Engineering Building, Department of Civil, Construction, and Environmental Engineering, Iowa State University, Ames, IA 50014, USA
rehmann@iastate.edu

(Received 3 January 2008 and in revised form 4 September 2008)

Scalars with different molecular diffusivities can be transported at different rates in a strongly stratified, weakly turbulent flow. Rapid distortion theory (RDT) is used to examine the mechanisms responsible for differential diffusion of scalars in a sheared stratified flow. The theory, which applies when the flow is strongly stratified, predicts upgradient flux and its wavenumber dependence, which previous direct numerical simulations have shown to be important in differential diffusion. The net effect of shear on differential diffusion depends on the Grashof number, or the relative importance of buoyancy and viscous effects. RDT also allows the effects of the density ratio, Schmidt number, Lewis number, scalar activity and mean shear to be examined without the high computational cost of direct numerical simulation. RDT predicts that differential diffusion will increase with increasing density ratio, but only at low Grashof number. When the Lewis number is fixed, the Grashof number below which differential diffusion occurs decreases with increasing Schmidt number, and when one of the Schmidt numbers is fixed, differential diffusion decreases with increasing Lewis number. Also, differential transport of passive scalars increases when the Schmidt number of the scalar stratifying the flow increases.

1. Introduction

Ocean circulation models and oceanographic studies of diapycnal mixing often assume that heat and salt have equal eddy diffusivities in a diffusively stable turbulent flow. However, because the molecular diffusivity of heat is much larger than the molecular diffusivity of salt, the two scalars can be transported at different rates in strongly stratified weakly turbulent flows (e.g. Turner 1968; Altman & Gargett 1990; Jackson & Rehmann 2003*b*). In the vast regions of the ocean interior where weak intermittent turbulence prevails (Merryfield 2005), this phenomenon of differential diffusion can be important for the interpretation of vertical mixing in the ocean (Gargett 2003), the generation of finestructure (Jackson & Rehmann 2003*a*; Martin & Rehmann 2006), ocean modelling (Merryfield, Holloway & Gargett 1998), and the propagation of intrusions (Hebert 1999). We extend an analytical model developed by

† Present address and email address for correspondence: US Geological Survey, Illinois Water Science Center, 1201 W. University Ave, Urbana, IL 61801, USA. pjackson@usgs.gov

Jackson *et al.* (2005) to examine the effect of shear and the properties of the scalars on differential diffusion.

Previous work has shown upgradient (or countergradient) fluxes to be important in differential diffusion. Rehmann (1995) and Nash & Moum (2002) compared the mixing time to the time for decay of the turbulence; when the mixing time is large, the turbulence will decay, and restratification will occur before the scalar can be irreversibly mixed. Direct numerical simulations (DNS) have provided information on the mechanisms leading to differential diffusion. The two-dimensional simulations of Merryfield *et al.* (1998) exhibited a maximum in the flux difference after the temperature flux and salt flux switched from downgradient to upgradient. By examining cospectra (or flux spectra), Gargett, Merryfield & Holloway (2003) showed that preferential transport of the scalar with the larger diffusivity requires larger upgradient fluxes of the more slowly diffusing scalar at high wavenumbers. Rapid distortion theory (RDT) for unsheared turbulence captures many of the phenomena, including the temporal evolution of the fluxes and the cospectra, observed in DNS (Jackson *et al.* 2005). Merryfield (2005) summarized the mechanisms in a schematic (his figure 4). When the flow is unstratified, a parcel of fluid containing heat and salt cannot restratify, and salt is transported more efficiently because it remains more concentrated than temperature. However, for stratified flow, parcels can restratify, and the net transport of temperature can be larger if the more efficient restratification of salt counteracts the transport during the downgradient phase.

Upgradient fluxes and differential diffusion depend on the properties of the flow and the properties of the fluid. Properties of the flow include the shear rate $\mathrm{d}\overline{U}/\mathrm{d}x_3$, the buoyancy frequency N, the background temperature gradient $\mathrm{d}\overline{T}/\mathrm{d}x_3$, the background salinity gradient $\mathrm{d}\overline{S}/\mathrm{d}x_3$, and a length scale L of the initial energy spectrum. Properties of the fluid include the kinematic viscosity ν, the molecular diffusivity of temperature D_T, the molecular diffusivity of salt D_S, and α and β, the coefficients of thermal expansion and haline contraction. In terms of dimensionless parameters, quantities such as the ratio d of the eddy diffusivities of salinity and temperature can be expressed as

$$d = \frac{K_S}{K_T} = f(Ri_g, Gr, Sc_T, Sc_S, Le, R_\rho), \tag{1}$$

where $Ri_g = N^2/(\mathrm{d}\overline{U}/\mathrm{d}x_3)^2$ is the gradient Richardson number, $Gr = NL^2/\nu$ is the Grashof number, $Sc_T = \nu/D_T$ is the Schmidt number for temperature, $Sc_S = \nu/D_S$ is the Schmidt number for salinity, $Le = D_S/D_T$ is the Lewis number, and $R_\rho = (-\alpha \mathrm{d}\overline{T}/\mathrm{d}x_3)/(\beta \mathrm{d}\overline{S}/\mathrm{d}x_3)$ is the density ratio. The activity of scalars – that is, whether the scalars change the fluid's density – will also be important. Jackson *et al.* (2005) related the Grashof number to $\varepsilon/\nu N^2$ (where ε is the rate of dissipation of turbulent kinetic energy), a parameter often called the buoyancy Reynolds number and used to describe the intensity of turbulence in a stratified fluid.

Because mean shear affects upgradient fluxes, it will also affect differential diffusion. Direct numerical simulations (e.g. Holt, Koseff & Ferziger 1992), experiments (e.g. Komori & Nagata 1996), and theory (e.g. Hanazaki & Hunt 2004) have all shown that shear reduces the magnitude of upgradient fluxes. According to DNS and RDT, shear dampens scalar flux oscillations (Hwang, Yamazaki & Rehmann 2006), and it can also lead to persistent upgradient fluxes (Gerz & Schumann 1991; Holt *et al.* 1992; Hanazaki & Hunt 2004). Laboratory experiments have lacked the measurements at large times needed to confirm these two results from DNS and RDT.

Although the most important case for oceanography has Schmidt number $Sc_T = \nu/D_T = 7$ and $Sc_S = \nu/D_S = 700$ (with Lewis number $Le = 0.01$), other values are often used because of practical constraints. For example, the DNS of Gargett *et al.* (2003), Merryfield (2005) and Smyth, Nash & Moum (2005) used $Sc_T = 7$ and $Sc_S = 50$–70 so that the small scales can be resolved, whereas the laboratory experiments of Hebert & Ruddick (2003) used passive dyes with Schmidt numbers of 2300 and 21000 in salt stratification to avoid complications with heat losses. Studies of the effect of the Schmidt number, Lewis number, and activity of the scalars are important for evaluating findings from DNS and experiments properly. For example, RDT shows that Schmidt number constraints lead DNS to underestimate oceanic differential diffusion at low Grashof numbers; RDT also suggests that when the Schmidt numbers increase (or the molecular diffusivities decrease), differential diffusion occurs at lower Grashof number (Jackson *et al.* 2005).

The effect of the density ratio, which measures the contributions of each scalar to the density gradient, on differential diffusion is unclear. Effects of density ratio on the ratio of eddy diffusivities were small in laboratory experiments on nearly homogeneous turbulence with $R_\rho \approx 0.25$ and $R_\rho \approx 5$ (Jackson & Rehmann 2003*b*; Martin & Rehmann 2006), and RDT calculations support this finding (Jackson *et al.* 2005). However, in the DNS of Merryfield (2005) greater differential diffusion occurred when the density ratio was small. Density ratio dependence in the DNS of Smyth *et al.* (2005) and the closure theory of Canuto *et al.* (2002) was consistent with that in Merryfield (2005). Theory for differential diffusion in a large-scale internal wave environment predicts that the diffusivity ratio will increase with increasing density ratio (Holloway 2006).

We extend the RDT of Jackson *et al.* (2005) to include mean shear and to study the effects of the dimensionless parameters on differential diffusion. RDT involves approximations, but it complements other approaches well. It applies for the weakly turbulent strongly stratified flows in which differential diffusion occurs, and it can produce spectra and upgradient fluxes. Jackson *et al.* (2005) showed that differential diffusion can be predicted using RDT. In § 2, RDT is developed for homogeneous turbulence with constant shear and buoyancy frequency, and in § 3, the effects of shear, density ratio, Schmidt numbers, Lewis number and activity of the scalars are discussed. The main findings are summarized in § 4.

2. Rapid distortion theory

2.1. *Assumptions*

In RDT, nonlinear terms in the equations for fluctuating momentum and scalars are neglected. The RDT equations are derived by linearizing the equations for the fluctuating quantities, choosing a coordinate system that deforms with the mean shear, and introducing a Fourier representation for the quantities (Rehmann & Hwang 2005). The equations for the Fourier amplitudes are used to obtain equations for the evolution of spectra. Spectra are obtained from a numerical solution and integrated to compute fluxes.

Hanazaki & Hunt (2004) discussed the conditions for validity of RDT for sheared stratified flow in detail. Estimating the terms in the governing equations for eddies of size λ and velocity u_λ, they showed that if the eddy Froude number $Fr = u_\lambda/N\lambda$ is small, then RDT is valid for a stratified flow (Hanazaki & Hunt 1996) and that the additional condition $Ri_g < O(1)$ ensures that RDT holds for sheared stratified flow. For low and moderate Reynolds number (i.e. values typical of laboratory experiments and

numerical simulations), a small Froude number $Fr = u/N\ell$, based on the velocity scale u and length scale ℓ of the large eddies, guarantees a small eddy Froude number, but for high Reynolds number, the condition $Fr_\lambda \ll 1$ can be violated for the small scales even when $Fr \ll 1$ (Hanazaki & Hunt 1996). These considerations are important for differential diffusion because fluxes are computed by integrating the spectra over all wavenumbers. At small times, differences in flux spectra appear at large wavenumbers, but later the differences spread to small wavenumbers as well (e.g. figure 12 of Gargett *et al.* 2003). In any case, because differential diffusion occurs in turbulence with low Reynolds number (e.g. Jackson & Rehmann 2003*b*), RDT should be well suited to study the key mechanisms and the effect of various dimensionless parameters.

Hanazaki & Hunt (2004) also discussed the applicability of RDT at large times. The Froude number decreases with time in grid-turbulence experiments (Liu 1995). At large times, a stratified flow can consist of a vortex mode and a wave mode – that is, a combination of nonlinear horizontal motions with large time scales and linear waves with small time scales (Riley & Lelong 2000). RDT captures the wave mode, and although RDT does not reproduce the vortex mode well, the conditions required for the vortex mode do not occur in many laboratory experiments and DNS (Hanazaki & Hunt 2004). In fact, the change in dynamics of wakes that signals the appearance of the vortex mode occurs at non-dimensional times $t' = Nt$ ranging from 20 for spheres (Spedding, Browand & Fincham 1996) to 30 for grids (Liu, Maxworthy & Spedding 1987). We use these findings to set the maximum integration time in the next section.

2.2. *Application to differential diffusion*

We consider homogeneous turbulence with constant mean shear, stable linear profiles of two active scalars, and a third, passive scalar with a linear profile. This case extends previous work: Hunt, Stretch, & Britter (1988) solved the inviscid one-scalar case, whereas Hanazaki & Hunt (1996) examined the effect of viscosity and diffusion for a single scalar in unsheared flow. Hanazaki & Hunt (2004) added mean shear to the one-scalar system, and Jackson *et al.* (2005) studied an unsheared flow with three scalars of different diffusivities. For a three-scalar system with mean shear, the equations for the Fourier amplitudes of the three velocities and scalars are

$$\frac{\mathrm{d}\hat{u}_1}{\mathrm{d}t'} = Ri_g^{-1/2}\left[\frac{2\kappa_1^2}{k^2} - 1\right]\hat{u}_3 + \left[\frac{\kappa_1 k_3}{k^2}\right](\hat{u}_S - \hat{u}_T) - Gr^{-1}k^2\hat{u}_1, \tag{2a}$$

$$\frac{\mathrm{d}\hat{u}_2}{\mathrm{d}t'} = Ri_g^{-1/2}\left[\frac{2\kappa_1\kappa_2}{k^2}\right]\hat{u}_3 + \left[\frac{\kappa_2 k_3}{k^2}\right](\hat{u}_S - \hat{u}_T) - Gr^{-1}k^2\hat{u}_2, \tag{2b}$$

$$\frac{\mathrm{d}\hat{u}_3}{\mathrm{d}t'} = Ri_g^{-1/2}\left[\frac{2\kappa_1 k_3}{k^2}\right]\hat{u}_3 + \left[\frac{k_3^2}{k^2} - 1\right](\hat{u}_S - \hat{u}_T) - Gr^{-1}k^2\hat{u}_3, \tag{2c}$$

$$\frac{\mathrm{d}\hat{u}_T}{\mathrm{d}t'} = -\frac{R_\rho}{R_\rho + 1}\hat{u}_3 - (GrSc_T)^{-1}k^2\hat{u}_T, \tag{2d}$$

$$\frac{\mathrm{d}\hat{u}_S}{\mathrm{d}t'} = \frac{1}{R_\rho + 1}\hat{u}_3 - (GrSc_S)^{-1}k^2\hat{u}_S, \tag{2e}$$

$$\frac{\mathrm{d}\hat{u}_C}{\mathrm{d}t'} = -\frac{R_\rho}{R_\rho + 1}\hat{u}_3 - (GrSc_C)^{-1}k^2\hat{u}_C, \tag{2f}$$

where κ_j are wavenumbers made dimensionless by a length scale L, $k^2 = \kappa_1^2 + \kappa_2^2 + k_3^2$, and $k_3 = \kappa_3 - (\mathrm{d}\overline{U}_1/\mathrm{d}x_3)\kappa_1 t$. The temperature T, salinity S, and passive scalar C are normalized such that $\hat{u}_T = g\alpha\hat{T}/N$, $\hat{u}_S = g\beta\hat{S}/N$ and $\hat{u}_C = g\alpha_C\hat{C}/N$, where α_C is

a constant with dimensions of C^{-1}. The passive scalar is taken to mirror the active temperature scalar in terms of initial distributions but without affecting the buoyancy; that is, the passive scalar is distributed so that $\alpha_C \mathrm{d}\bar{C}/\mathrm{d}x_3 = \alpha \mathrm{d}\bar{T}/\mathrm{d}x_3$.

As in Rehmann & Hwang (2005) and Jackson *et al.* (2005), the RDT equations (2*a*)–(2*f*) are used to write a set of equations for the evolution of the spectra by employing a discrete Fourier transform of the two-point correlations. The initial conditions imposed on this system include energy spectra for isotropic turbulence:

$$E_{ij}(\underline{\boldsymbol{\kappa}}, 0) = \frac{E(k)}{4\pi k^2}\left(\delta_{ij} - \frac{\kappa_i \kappa_j}{k^2}\right). \tag{3}$$

Following Townsend (1980), the energy spectrum function is taken to be

$$\frac{E(k)}{q_0^2 L} = \frac{1}{3\sqrt{2\pi}} k^4 \exp\left(-\tfrac{1}{2}k^2\right), \tag{4}$$

where $q_0^2/2$ is the initial turbulent kinetic energy. Assuming no initial scalar fluctuations requires that all the remaining spectra are set to zero. The system of ordinary differential equations for the evolution of the spectra was solved numerically using the DLSODA solver, a variant of the Livermore solver for ordinary differential equations, which uses a dense or banded Jacobian and the backward differentiation formula methods (Gear methods) when the system is stiff and Adams methods (predictor–corrector) when the system is not stiff (Hindmarsh 1983). The numerical solution was tested against analytical and numerical results from previous work to ensure adequate spatial and temporal resolution. Once the vertical scalar flux spectra E_{S3}, E_{T3} and E_{C3} were computed, they were integrated in space to yield the vertical fluxes for salt $F_S = \overline{u_S' u_3'}$, temperature $F_T = -\overline{u_T' u_3'}$, and the passive scalar $F_C = -\overline{u_C' u_3'}$, where primes denote fluctuating quantities and the overbars denote ensemble averages. Except for runs examining the evolution of scalar spectra in § 3.1 (simulated until $t' = 3$), all runs were simulated until $t'_{max} = 30$, and the diffusivity ratio was calculated as

$$d = \frac{K_S}{K_T} = R_\rho \frac{\int_0^{t'_{max}} F_S \mathrm{d}t'}{\int_0^{t'_{max}} F_T \mathrm{d}t'}. \tag{5}$$

For two passive scalars, the density ratio was taken to be unity because the initial distributions of the two passive scalars were identical.

The diffusivity ratio computed with (5) depends on t'_{max}, the time over which the fluxes are integrated. Most of the contribution to the overall differential diffusion occurs in the first few buoyancy periods; for example, in the DNS of Merryfield (2005), the diffusivity ratio converged within the first four buoyancy periods ($t' \approx 25$). Nevertheless, in RDT, small oscillations in the fluxes continue to contribute to the transport after this time. Because RDT underestimates the rate of decay of vertical fluxes compared to DNS (Hanazaki & Hunt 1996), we ended the calculations at a specified value of t' to avoid a low bias in the diffusivity ratio. When comparing the RDT to the experiments of Jackson & Rehmann (2003*b*), Jackson *et al.* (2005) integrated the fluxes over a period equal to the waiting time between stir sets in the experiments. Here we integrated until $t' = 30$, the non-dimensional time by which grid turbulence collapses completely (Liu *et al.* 1987). This integration period, which is similar to that in Merryfield (2005), captures the most important buoyancy periods with respect to differential diffusion, while not allowing a significant influence of the long-time behaviour of RDT.

2.3. *Relevance for natural flows*

The RDT model described in the previous subsection allows the effect of dimensionless parameters on differential diffusion to be examined. However, three issues – the initial conditions, the range of scales represented, and the assumption of homogeneous turbulence – could be discussed further to understand how the model might apply to natural flows. The initial conditions specified above include a spectrum for the velocity field, but zero density fluctuations. In fact, Hunt *et al.* (1988) and Hanazaki & Hunt (1996) found that RDT predicts a strong dependence of fluxes on the initial conditions – in particular, PE_0/KE_0, the ratio of the initial fluctuating potential energy and the initial fluctuating kinetic energy (i.e. $q_0^2/2$). This ratio is difficult to measure directly at the beginning of a turbulent event in the ocean. Comparing RDT to laboratory experiments suggests the ratio of initial energies is small. After accounting for decay, Hunt *et al.* (1988) obtained qualitative agreement with laboratory measurements by choosing $PE_0/KE_0 = 0$, whereas Hanazaki & Hunt (1996) used $PE_0/KE_0 \approx 0.15$ in comparing to results on grid turbulence. In contrast, Gerz & Yamazaki (1993) and Hwang *et al.* (2006) have studied the extreme case of $PE_0/KE_0 \to \infty$ as a model of turbulence generated from density fluctuations remaining from a turbulent patch whose velocity fluctuations have decayed. Although setting $PE_0/KE_0 \neq 0$ would yield quantitatively different results and examining differential diffusion in buoyancy-generated turbulence would be an interesting extension, the choice $PE_0/KE_0 = 0$ corresponds well enough to previous measurements and represents a logical starting point for studying the effect of other parameters on differential diffusion.

A related issue is the range of wavenumbers that can develop. As Jackson *et al.* (2005) discussed, the wavenumber range is set by the initial spectrum because a linearized model such as RDT cannot produce an energy cascade. In DNS, however, non-zero spectral values appear at higher wavenumbers for scalars with low diffusivity than for scalars with high diffusivity because of differences in Batchelor wavenumbers. While RDT cannot generate new scales, a closer match between RDT predictions and DNS results and field measurements could be achieved by specifying initial scalar spectra with different wavenumber ranges.

The model described in § 2.2 applies to homogeneous turbulence, but turbulence and mixing in the interior of oceans and lakes usually occurs in patches (e.g. Baker & Gibson 1987; Gregg 1987). The degree of inhomogeneity can be measured by the ratio of the integral scale ℓ of the turbulence and the scale L_p of the patch. When the patch is much larger than the energy-containing eddies (i.e. $\ell/L_p \to 0$), then homogeneous turbulence is a good approximation. In laboratory experiments, $\ell/L_p \approx 0.05$ at $t' \approx 4$, when stratification limits the growth of a patch (De Silva & Fernando 1992). To account for the inhomogeneity, we could extend the WKB theory of Nazarenko, Kevlahan & Dubrulle (1999), who considered localized turbulence subjected to weakly non-uniform distortion from irrotational strain; in this case, the theory successfully explained the experimental observations that the small-scale turbulence generates a large-scale flow that limits the distortion. Applied to stratified flow, the WKB theory would allow the feedback between the turbulence and mean velocity and density profiles to be studied.

3. Results and discussion

In this section, results from many runs are considered in order to evaluate the effect of the Richardson number, Grashof number, density ratio, Schmidt numbers, Lewis

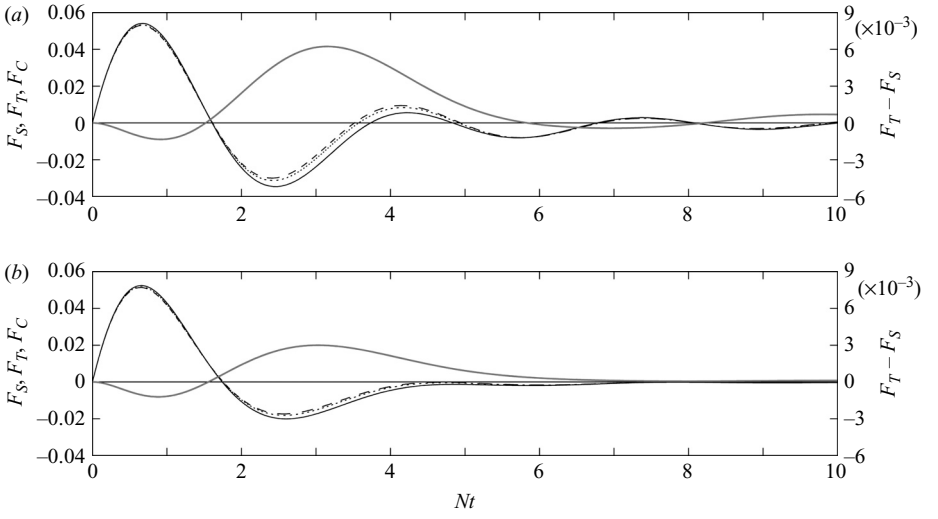

FIGURE 1. Temporal evolution of the scalar fluxes for $Gr = 10$. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$. The solid line is the salt flux F_S ($Sc = 700$), the dashed line is temperature flux F_T ($Sc = 7$), and the dotted line is passive scalar flux F_C ($Sc = 10$). Flux differences $F_T - F_S$ (grey line) are shown on the right axes.

number and scalar activity. Predictions from RDT are also compared with results from previous numerical and experimental studies.

3.1. *Scalar fluxes and spectra*

The excess upgradient flux of the lower molecular diffusivity scalar (salt) is the main cause of differential diffusion. The scalar fluxes oscillate between downgradient and upgradient (figure 1). At small times, the initial turbulent kinetic energy sustains downgradient flux, and in this phase, temperature exhibits smaller downgradient flux because its larger diffusivity leads to greater damping – especially at higher wavenumbers (figure 2*a*). After a quarter buoyancy period ($Nt \approx \pi/2$), the scalars restratify. Upgradient fluxes appear at small scales first (figure 2*b*) and eventually spread to the entire spectrum (figure 2*c*); again, the magnitude of the salt flux is larger than the magnitude of the temperature flux. When the total fluxes become downgradient at about $Nt = 3.5$ in the unsheared case (figure 1*a*), the temperature flux is larger because of the excess upgradient flux of salt. The fluxes continue to oscillate in time until they are damped by viscosity and diffusion. As Jackson *et al.* (2005) noted, the evolution of vertical scalar fluxes and spectra predicted by RDT resembles that seen in DNS (Merryfield *et al.* 1998; Gargett *et al.* 2003).

Consistent with previous work, RDT predicts that mean shear reduces the magnitude of upgradient fluxes of all scalars, increases damping of the oscillating fluxes, and produces persistent upgradient fluxes. Early in the evolution, shear affects the scalar fluxes very little in a flow with $Ri_g = 1$ (figures 1, 2*a*, 3*a*). Later, although the fluxes are similar, upgradient fluxes in the spectra are smaller in the sheared case, and they appear at higher wavenumbers (figures 2*b*, 3*b*). A similar reduction in fluxes and suppression of upgradient fluxes at large scales occurred in the experiments of Komori & Nagata (1996). Still later, upgradient fluxes appear at all wavenumbers, but the spectral values are much smaller in the sheared flow (figures 2*c*, 3*c*). For $Nt > 2$, the fluxes are smaller in the sheared case (figure 1*b*), and they mostly remain upgradient. Persistent upgradient fluxes at small scales are common in sheared and unsheared

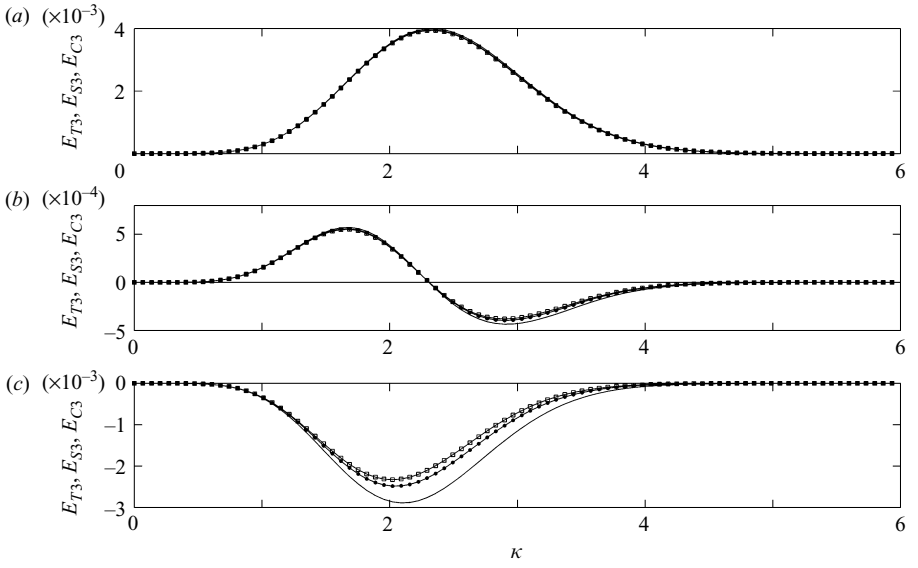

FIGURE 2. Temporal evolution of the nondimensional vertical scalar flux spectra for $Gr = 10$ and $Ri_g \to \infty$. (*a*) $Nt = 0.34$; (*b*) $Nt = 1.49$; (*c*) $Nt = 2.65$. The solid lines are salt flux spectra E_{S3} ($Sc = 700$), the squares are temperature flux spectra E_{T3} ($Sc = 7$), and the filled circles are passive scalar flux spectra E_{C3} ($Sc = 10$).

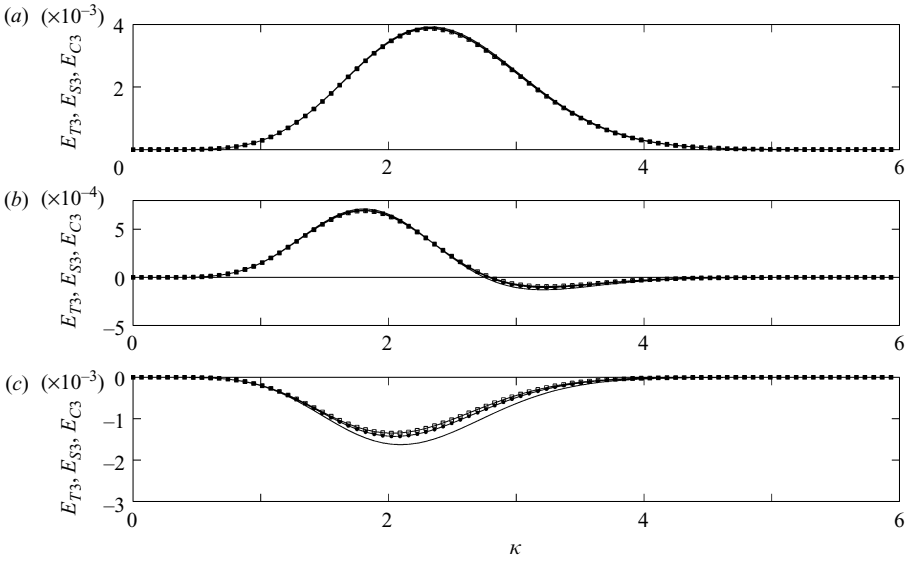

FIGURE 3. Temporal evolution of the nondimensional vertical scalar flux spectra for $Gr = 10$ and $Ri_g = 1$. (*a*) $Nt = 0.34$; (*b*) $Nt = 1.49$; (*c*) $Nt = 2.65$. The line designations are the same as in figure 2.

flows (Gerz & Schumann 1991; Komori & Nagata 1996). Gerz & Schumann (1991) argued that the total flux can be persistently upgradient when large-scale oscillating fluxes cancel and leave only the small-scale upgradient fluxes. Persistent upgradient fluxes in sheared flow may be attributed to reduced upgradient fluxes at large scales.

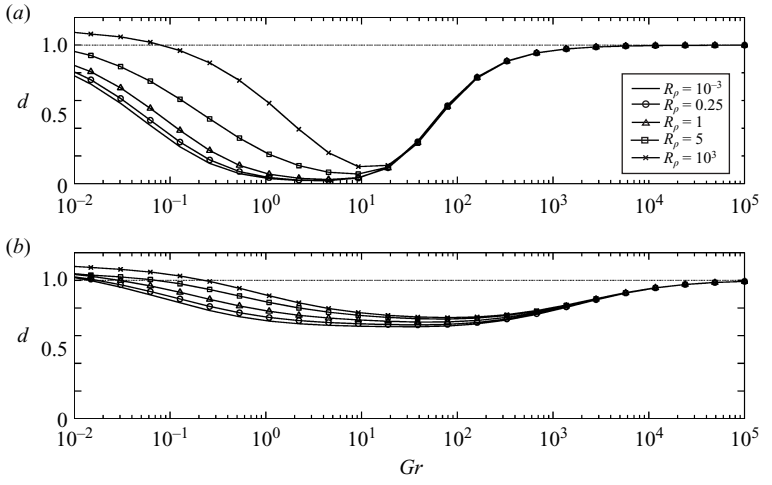

FIGURE 4. Diffusivity ratio as a function of Gr for various density ratios. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$.

Although the passive scalar is not influenced directly by gravity, it experiences both downgradient and upgradient fluxes (figures 1–3). Upgradient fluxes of the passive scalar arise because of the upgradient fluxes of the active scalar. In the cases shown in figure 1, flux of the passive scalar, which has a Schmidt number close to that of temperature, follows the temperature flux closely. Increasing the Schmidt number of the passive scalar above that of the active scalar with the smaller diffusivity does not increase the upgradient fluxes of the passive scalar because the upgradient fluxes are controlled by the active scalars. For unsheared, stratified flow without viscosity or diffusion, Hanazaki (2003) used RDT to show that when the initial spectrum of density fluctuations is zero (as it is in our case), the ratio of the fluxes of buoyancy and the passive scalar depends only on the ratio of N^2 and the mean gradient of the passive scalar; differential transport of the active and passive scalars occurred only under certain initial conditions. In our case, viscosity and diffusion alter the time development and render the total fluxes of temperature and passive scalar different even in conditions that would suggest equal eddy diffusivities in unsheared inviscid non-diffusive flow.

3.2. *Effects of density ratio*

The effects of the various parameters, such as density ratio, on differential diffusion can be examined by integrating the fluxes and computing the diffusivity ratio d using (5). In the base case of unsheared flow with $R_\rho = 1$, the diffusivity ratio is near unity at high Grashof number, for which effects of viscosity and diffusivity are less important (figure 4*a*). As the Grashof number decreases to moderate values, the total flux of temperature exceeds that of salt (i.e. $d < 1$). At low values of Gr, the flow is so heavily damped that the flux is non-zero for a short time only, and the effects of restratification are smaller. Therefore, the differential transport is smaller, and d increases again; the increase at low Gr is discussed further in § 3.3.

The density ratio affects the diffusivity ratio most at low Grashof numbers (figure 4). When the density ratio increases from 10^{-3} to 10^{3}, the diffusivity ratio can increase by as much as a factor of 16 for the unsheared case and 25 % for the flow with $Ri_g = 1$.

The largest effects of the density ratio on differential diffusion occur at Grashof numbers below the Grashof number associated with the minima in the diffusivity ratio curves – that is, as molecular diffusion becomes more important. Both Merryfield (2005) and Holloway (2006) argued that as R_ρ increases, the more diffusive scalar increasingly controls the restratification, and differential transport is smaller (i.e. d increases), as in figure 4.

Shear decreases the effect of the density ratio on the magnitude of the differential diffusion and increases the range of Gr over which density ratio influences differential diffusion (figure 4*b*). Increasing shear reduces the effect of R_ρ because shear reduces the magnitude of upgradient fluxes and dampens the fluxes, as discussed in § 3.1. This result is consistent with the closure theory of Canuto *et al.* (2002), which predicts the greatest effect of the density ratio at large gradient Richardson numbers (i.e. weakly sheared flows). In sheared flows, effects of the density ratio appear over two more decades in Gr; as discussed in § 3.5, shear makes molecular diffusion effects important at higher Gr relative to the unsheared case.

The present results agree with the results in Jackson & Rehmann (2003*b*), Martin & Rehmann (2006) and Jackson *et al.* (2005), who showed that the density ratio affected differential diffusion little for parameters typical for geophysical flows. The present work is also consistent with recent DNS (Merryfield 2005; Smyth *et al.* 2005) and theory (Holloway 2006) in some ways. All show that d increases with increasing R_ρ, and all predict the greatest effect of the density ratio on differential diffusion at low values of $\varepsilon/\nu N^2$, which Jackson *et al.* (2005) related to the Grashof number as $Gr = C_1\varepsilon/\nu N^2$, with $C_1 \approx 7$. However, among the studies that address high Gr or high $\varepsilon/\nu N^2$, the simulations of Merryfield (2005) show considerable effect of density ratio even at high $\varepsilon/\nu N^2$, whereas the present work and the work of Holloway (2006) do not. Increasing the Lewis number to 0.1, as in the simulations of Merryfield (2005), does not lead to noticeable effects of the density ratio at higher Gr. The differences between the results from RDT and DNS could be due to nonlinear processes not modelled in RDT.

3.3. *Effects of Schmidt number*

For unsheared flow, varying the Schmidt numbers while holding the Lewis number constant affects the magnitude of the diffusivity ratio little but changes the range of Grashof number over which differential diffusion occurs (figure 5*a*). Over a wide range of Schmidt number, the minima in diffusivity ratio vary little, but as the Schmidt numbers decrease, differential diffusion occurs for larger Gr. If Gr_c is the Grashof number at which the diffusivity ratio first deviates from unity, then a power-law fit to the results gives $Gr_C \propto {Sc_T}^{-1}$, except at high Schmidt number (Jackson 2006, figure 3.8). This result is consistent with results from Shih *et al.* (2005) and Ivey, Winters & De Silva (2000), who predicted that the value of $\varepsilon/\nu N^2$ below which molecular diffusion dominates scalar mixing is proportional to Sc^{-1}. This Schmidt-number dependence suggests that even though turbulence in the atmosphere is more intense than in the ocean (Joseph *et al.* 2004), differential diffusion may still be important because the Schmidt numbers of the scalars are less than unity.

Shear modifies these results (figure 5*b*). As observed in § 3.2, shear increases the range of Grashof number over which differential diffusion occurs. Also, shear causes the magnitude of differential diffusion to depend on the Schmidt numbers, especially for $Sc_T > 1$. We might expect differential diffusion in flows with lower Schmidt numbers to be more sensitive to shear because the molecular diffusivity would affect larger scales, and past work has shown that shear reduces upgradient fluxes at large

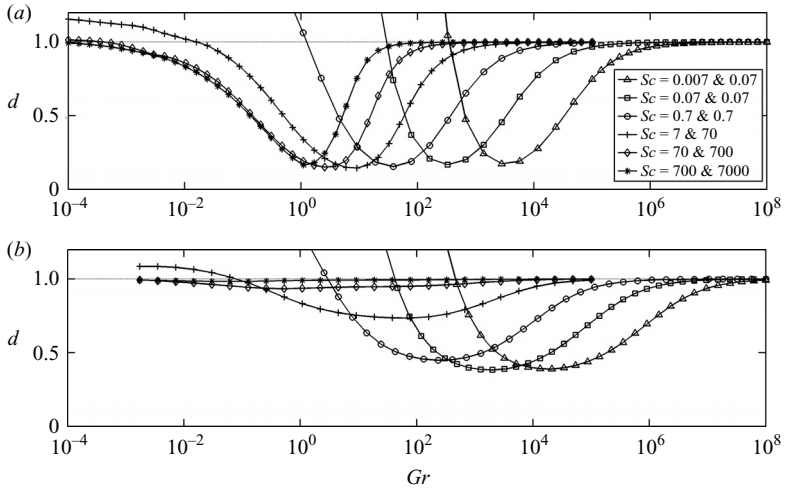

FIGURE 5. Diffusivity ratio as a function of Gr for various Schmidt number pairs. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$. All runs had $Le = 0.1$ and $R_\rho = 1$.

scales while leaving upgradient fluxes at small scales mostly unchanged (e.g. Komori & Nagata 1996). However, Hanazaki & Hunt (2004) observed that shear suppressed small-scale vertical motions and enhanced large-scale motions; this result is consistent with the smaller differential diffusion (i.e. larger diffusivity ratios) in sheared flows with high Schmidt numbers.

The observation that the critical Grashof number is inversely proportional to Sc_T suggests plotting the diffusivity ratio as a function of $GrSc_T = NL^2/D_T$. In both the sheared and unsheared cases, the data collapse well for $Sc_T < 1$, but not for $Sc_T > 1$ (figure 6). The parameter NL^2/D_T can be interpreted as a ratio of a diffusion time L^2/D_T and a restratification time N^{-1}. When NL^2/D_T decreases, more diffusion occurs during a restratification period, and differences between the fluxes become apparent. Similarly, the Schmidt number can be interpreted as the ratio of the time λ^2/D_T for diffusion of the scalar and the time λ^2/ν for diffusion of momentum across a length λ. When $Sc_T < 1$, molecular diffusion of the scalar affects a larger range of scales than molecular diffusion of momentum; therefore, scalar diffusion controls the vertical flux, and the curves collapse when plotted against NL^2/D_T. When $Sc_T > 1$, viscosity affects a larger range of scales and limits the vertical velocity fluctuations contributing to the flux, and the times for diffusion of momentum and diffusion of the scalar (i.e. Gr and Sc_T) both become important.

As the Grashof number decreases, d increases to values above unity. The diffusivity ratio is computed from turbulent fluxes; if fluxes from pure molecular diffusion were included, the total diffusivity ratio would approach the Lewis number, as in the experiments of Jackson (2006). The increase in d can be explained qualitatively by reconsidering the fluxes in figure 1. As explained in § 3.1, the salt flux exceeds the temperature flux at small times. If the flow is heavily damped (small Gr), restratification will be unimportant, and d can exceed unity. In fact, in the limit $Gr \to 0$, the diffusivity ratio can be computed from (2) and (5) as

$$d = \frac{Sc_S(1 + Sc_T)}{Sc_T(1 + Sc_S)} = Le^{-1}\frac{(1 + Sc_T)}{(1 + Sc_S)}, \tag{6}$$

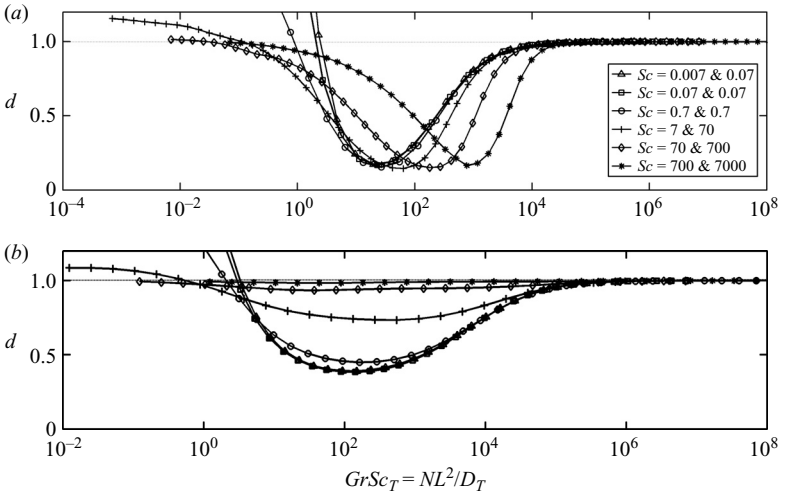

FIGURE 6. Diffusivity ratio as a function of NL^2/D_T for various Schmidt number pairs. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$. All runs had $Le = 0.1$ and $R_\rho = 1$.

which shows that $d > 1$ for low Gr when $Sc_S > Sc_T$. This result holds for active and passive scalars in sheared and unsheared flows because the viscous and diffusive terms in equation (2) dominate.

3.4. *Effects of Lewis number*

As the difference between molecular diffusivities increases (i.e. the Lewis number decreases), differential diffusion increases (figure 7). The Lewis number was varied by holding Sc_T constant and changing Sc_S. The curves for different Le differ most at intermediate Grashof number. For the unsheared cases at $Gr = O(10)$, the diffusivity ratio is 1 for identical scalars ($Le = 1$), and it approaches zero for scalars with widely different molecular diffusivities. The critical Grashof number varies little because Sc_T is fixed. The case with $Ri_g = 1$ mimics the unsheared case, but with less overall differential diffusion and a shift to higher Grashof number. RDT probably underestimates differential diffusion because in real turbulence, the Batchelor wavenumbers separate as the Lewis number decreases and greater restratification of the scalar with low molecular diffusivity will increase differential diffusion. In contrast, RDT neglects the cascade, and in our application, the range of scalars is set by the initial velocity spectrum, as discussed in § 2.3.

Resolution requirements have so far constrained DNS to use Lewis numbers larger than the heat–salt value of $O(10^{-2})$, and RDT can help to estimate the implications. The direct simulations of differential diffusion of Merryfield (2005) and Smyth *et al.* (2005) used Schmidt numbers of 7 and 70 ($Le = 0.1$) and 7 and 50 ($Le = 0.14$), respectively. The theory of Holloway (2006) examined the effect of the Lewis number on differential diffusion for $Le = 0.1$ and $Le = 0.01$; the observation of less differential diffusion for $Le = 0.1$ than for $Le = 0.01$ is consistent with intuition and RDT. Gargett *et al.* (2003) estimated that differential diffusion was about 20–40 % smaller for their three-dimensional DNS with $Le = 0.1$ than the two-dimensional DNS of Merryfield *et al.* (1998) with $Le = 0.01$. However, RDT for an unsheared flow simulated until $t' = 240$ shows that the two cases have diffusivity ratios within 10 % for $Gr > 600$ and 40 % for $Gr > 300$, but for low Grashof numbers, the diffusivity ratio for heat and salt is greatly underpredicted if theory or simulations use $Le = 0.1$ (Jackson

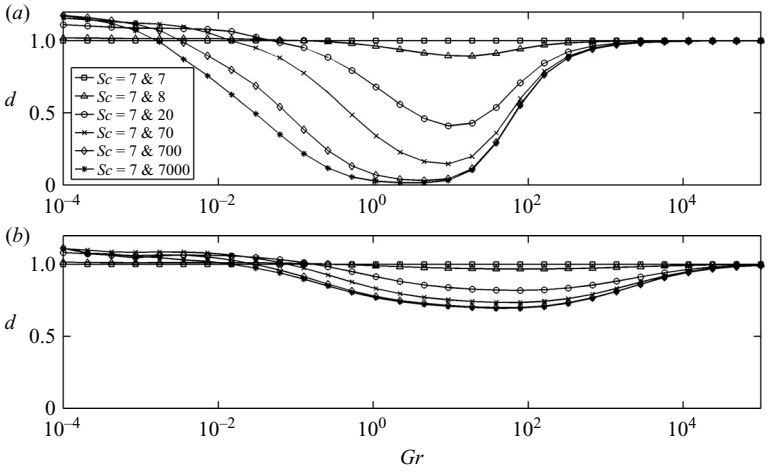

FIGURE 7. Diffusivity ratio as a function of Grashof number for various Lewis numbers. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$. All runs had $R_\rho = 1$.

et al. 2005). For $Le = 0.1$ and $Le = 0.01$, the present work, in which the fluxes were computed until $t' = 30$, predicts a factor of 5 difference in the diffusivity ratio minima for an unsheared flow and only a 4 % difference in the case of $Ri_g = 1$.

RDT shows that differential diffusion can also be significant even when the Schmidt numbers differ by less than one order of magnitude. Oceanographers, limnologists and laboratory experimentalists regularly assume equal mixing between a tracer and the scalar of interest (e.g. Ledwell, Watson & Law 1993; Komori & Nagata 1996) even though the molecular diffusivities of the tracer and the scalar of interest are rarely equal. For example, sulphur hexafluoride, which is used in ocean tracer release experiments, has a Schmidt number between 611 and 2263 (King & Saltzman 1995) over the range 5–30 °C, and it is used to measure vertical mixing of salt ($Sc = 700$, $0.9 > Le > 0.3$) and heat ($Sc = 7$, $0.01 > Le > 0.003$). Dyes are also used to estimate mixing in natural waters. Disodium fluorescein ($Sc = 1200$) and Rhodamine WT dyes ($Sc = 2000$) are frequently used to estimate vertical mixing of heat ($0.0035 > Le > 0.006$) and biological and chemical constituents ($0.06 > Le > 2$; see Socolofsky & Jirka 2002). Disodium fluorescein dye is also used in the laboratory to track salt for optical mixing measurements ($Le = 0.6$; e.g. Komori & Nagata 1996). In these cases with $Le < 1$, the flux of the scalar of interest can be underestimated.

3.5. *Effects of scalar activity*

The discussion in the previous sections focused on active scalars, such as heat and salt, which affect the density and the dynamics of the flow. To examine the effects of scalar activity, a case with two active scalars with Schmidt numbers of 7 and 700 (heat and salt) was compared to two cases with two passive scalars with Schmidt numbers of 7 and 700 and a third, active scalar – either heat or salt. Some features of the diffusivity ratio can be expected from the discussion in previous sections (figure 8): At high Gr, properties of the scalars become unimportant, while at low Gr, the diffusivity ratio approaches the limit in (6). When $Ri_g = 1$, effects of the scalar properties are seen at higher Gr, but differential diffusion is reduced for most Grashof numbers.

At intermediate Grashof numbers, the effect of scalar activity can be significant. In particular, differential diffusion of two passive scalars depends on the active scalar. For the unsheared case (figure 8*a*), if the active scalar is salt, then differential diffusion

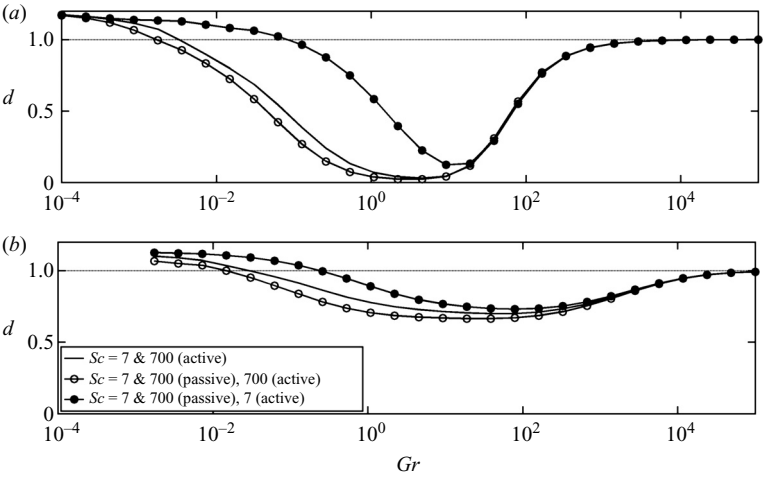

FIGURE 8. Diffusivity ratio as a function of Grashof number for active and passive scalar pairs. (*a*) $Ri_g \to \infty$ (unsheared); (*b*) $Ri_g = 1$. The runs with the passive scalar pairs have a third, active scalar that provides background stratification; for these runs, the diffusivity ratio is based on the fluxes of the two passive scalars.

increases (i.e. d decreases) by as much as 40 % compared to the case with active heat and salt, but if the active scalar is temperature, then differential diffusion decreases by as much as a factor of 8. The larger upgradient flux that occurs for salt, as discussed in § 3.1, leads to more differential diffusion of the passive scalars. The case with active heat and salt is bounded by the passive scalar cases because salt and heat contributed equally to the stratification ($R_\rho = 1$). These results are consistent with those in § 3.2, in which effects of large and small density ratio were considered (figure 4).

The dependence of differential transport of passive scalars on the active scalars suggests that the density ratio should affect transport of passive scalars (such as tracers, dyes, etc.) in flows with two active scalars, such as the ocean. In particular, these findings are important for studies of scalar transport in high-latitude oceans and estuaries, where salinity controls the stratification. Also, because Schmidt numbers typically range from 500 to 2500 for most scalars of interest (Socolofsky & Jirka 2002), biogeochemical scalar transport in lakes and other freshwater bodies is subject to differential diffusion. Errors can be introduced when estimating transport of such scalars using temperature microstructure methods; that is, if the eddy diffusivity of the scalar of interest is assumed to equal the eddy diffusivity of temperature, the error can exceed the typical uncertainty associated with microstructure methods.

3.6. *Effects of shear*

In the preceding sections, shear reduced the magnitude of upgradient fluxes, increased damping of the fluxes, and generated persistent upgradient fluxes. As a consequence, shear reduced the effect of the density ratio on differential diffusion, made the diffusivity ratio minima depend on the Schmidt number of the higher diffusivity scalar, increased the critical Grashof number and the range of Grashof number over which differential diffusion was important, and lessened the effect of scalar activity on the diffusivity ratio at intermediate Grashof numbers. However, these results are based on a comparison between the unsheared case and a case with $Ri_g = 1$. We now examine the effect of shear on differential diffusion by considering a range of Ri_g.

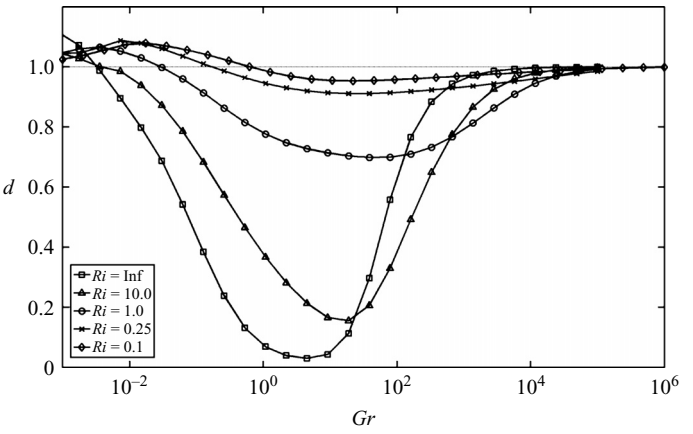

FIGURE 9. Diffusivity ratio as a function of Grashof number for flows with different Richardson numbers. All runs had $R_\rho = 1$.

The effect of shear depends on the Grashof number (figure 9). For $Gr < 30$, shear reduces differential diffusion. For $Gr > 30$, the effect of shear is more complex, but it generally increases differential diffusion compared to the unsheared case. In general, a larger Ri_g leads to a lower minimum diffusivity ratio. For example, decreasing Ri_g from 10 to 0.1 increases the minimum diffusivity ratio by a factor of nearly five. Therefore, a flow with significant differential diffusion may be altered to produce equal mixing by increasing the shear. This result could be important for laboratory experiments and industrial mixing processes (e.g. turbulent diffusion flames, see Bilger 1989) in which equal mixing rates of multiple scalars are desired and the flow is externally controlled. For the runs in figure 9, the critical Grashof number was largest for $Ri_g = 0.25$. Values of Ri_g between about 0.1 and 0.25 separate flows with growing and decaying turbulence in numerical simulations of the full governing equations (e.g. Shih *et al.* 2000).

To understand why shear affects differential diffusion differently at high and low Grashof numbers, the scalar fluxes were examined for $Gr = 1000$ and $Gr = 30$ for $Ri_g = 1.0$ and ∞. At both Grashof numbers, shear reduces the upgradient fluxes, produces nearly persistent upgradient fluxes, and dampens the flux oscillations (figure 10). However, for $Gr = 30$, shear increases the diffusivity ratio, whereas for $Gr = 1000$, shear decreases the diffusivity ratio. For low Grashof number, strong effects of viscosity lead to small scalar fluxes that dampen quickly even in unsheared flow (figure 10*a*). Shear further reduces oscillations in the flux (figure 10*b*); in particular, the smaller upgradient fluxes early in the evolution lead to a smaller difference between the cumulative scalar fluxes $\Phi_T = \int F_T \mathrm{d}t'$ and $\Phi_S = \int F_S \mathrm{d}t'$ (figure 11*a,b*) and a larger diffusivity ratio than in the unsheared case. Persistent upgradient fluxes later in the evolution contribute little to further decreasing the diffusivity ratio because the fluxes become small quickly. For high Grashof number, the fluxes oscillate for long times in the unsheared case (figure 10*c*). Shear mostly eliminates the downgradient fluxes for $Nt > \pi/2$ (figure 10*d*), and the persistent upgradient fluxes lead to a greater difference between the cumulative fluxes (figure 11*c*, *d*) and a smaller diffusivity ratio.

3.7. *Comparison with numerical and experimental results*

In this section, the present theory is compared with previous experimental and numerical studies of differential diffusion. Jackson *et al.* (2005) compared their

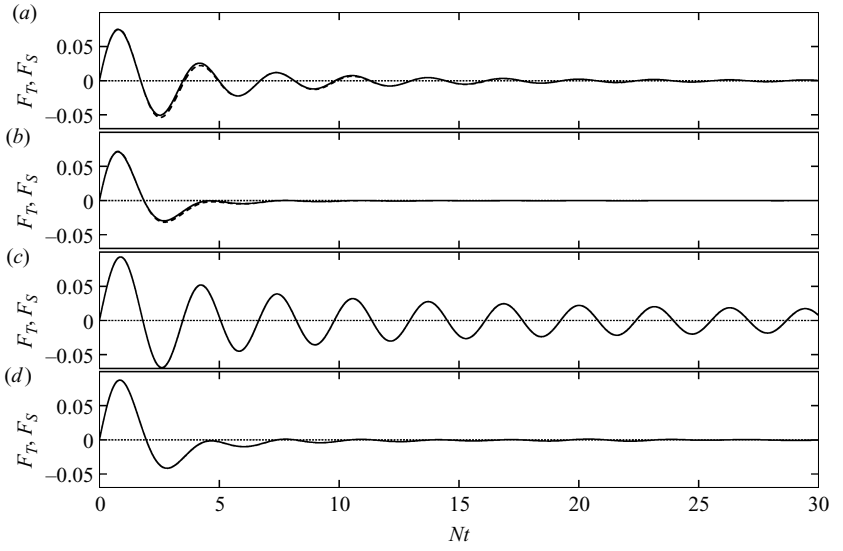

FIGURE 10. Temporal evolution of the normalized scalar fluxes for high and low Grashof number and sheared and unsheared flow: (*a*) $Gr = 30$, $Ri_g \to \infty$, (*b*) $Gr = 30$, $Ri_g = 1$, (*c*) $Gr = 1000$, $Ri_g \to \infty$, and (*d*) $Gr = 1000$, $Ri_g = 1$. The solid line is temperature and the dashed line is salt.

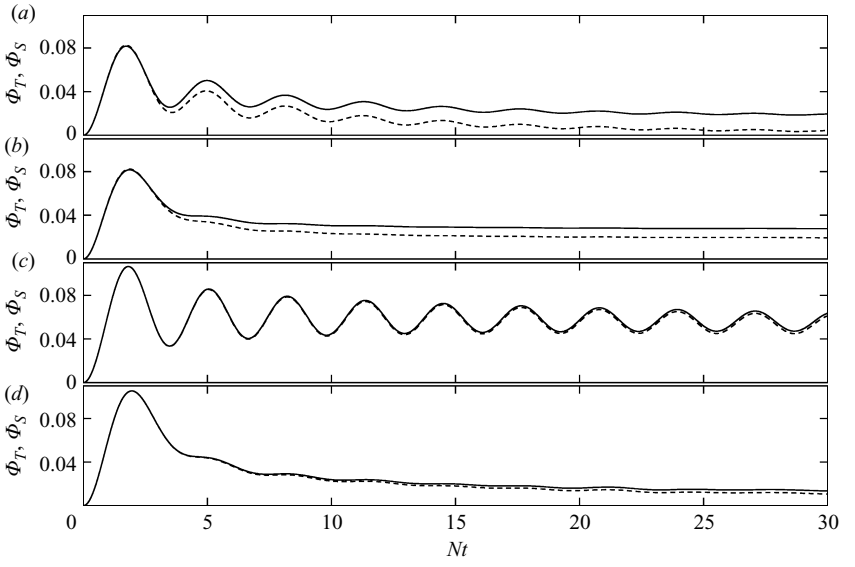

FIGURE 11. Temporal evolution of the normalized scalar flux integrated in time for high and low Grashof number and sheared and unsheared flow: (*a*) $Gr = 30$, $Ri_g \to \infty$, (*b*) $Gr = 30$, $Ri_g = 1$, (*c*) $Gr = 1000$, $Ri_g \to \infty$, and (*d*) $Gr = 1000$, $Ri_g = 1$. The solid line is temperature and the dashed line is salt.

RDT results with the work of Jackson & Rehmann (2003*b*), Martin & Rehmann (2006), and Hebert & Ruddick (2003) and found good agreement between RDT and measurements of differential diffusion of heat and salt. RDT was also able to predict a large shift in the diffusivity ratio curve to low $\varepsilon/\nu N^2$ for passive dyes in a

salt stratification (Hebert & Ruddick 2003). The present work extends the work of Jackson *et al.* (2005) and focuses on comparing to the experimental work of Turner (1968), to the numerical studies of Merryfield (2005) and Smyth *et al.* (2005), and to the theory of Holloway (2006) and Canuto *et al.* (2002).

Directly comparing predictions of RDT to results from previous work is complicated by two issues. The first is the time t'_{max} over which the fluxes are integrated, which we take to be 30; this choice is explained in § 2.3. The second issue is the relationship between the Grashof number in RDT and a parameter of DNS and experiments (usually $\varepsilon/\nu N^2$). As stated in § 3.2, Jackson *et al.* (2005) related these parameters by $Gr = NL^2/\nu \cong C_1\varepsilon/\nu N^2$ and found that their RDT matched the data of Jackson & Rehmann (2003*b*) and Martin & Rehmann (2006) quite well if $C_1 = 7$. They argued that this result means that the length scale L is related to the Ozmidov scale, an important length scale for strongly stratified flow, by an $O(1)$ coefficient. In addition to the uncertainty introduced by the value of the coefficient of proportionality, further uncertainty comes from the values of $\varepsilon/\nu N^2$ computed for past work. The DNS of Gargett *et al.* (2003), Merryfield (2005), and Smyth *et al.* (2005) compute $\varepsilon/\nu N^2$ directly, but $\varepsilon/\nu N^2$ for the experiments of Turner (1968) was estimated (Nash & Moum 2002).

The magnitude of the diffusivity ratio and its dependence on $\varepsilon/\nu N^2$ predicted by RDT for unsheared flow is consistent with experiments, DNS, and theory (figure 12). All predict that the diffusivity ratio decreases with decreasing Grashof number (or $\varepsilon/\nu N^2$) from unity to quite low values. RDT slightly overpredicts the diffusivity ratio at high $\varepsilon/\nu N^2$ and underpredicts at very low $\varepsilon/\nu N^2$. Increasing the Lewis number to $O(10^{-1})$ reduced the difference between RDT and DNS at low $\varepsilon/\nu N^2$, though RDT still underpredicts d. These slight differences might be caused in part by taking a constant integration period which would bias the diffusivity ratio high at large $\varepsilon/\nu N^2$. Although RDT for $Le = 0.01$ differs at low $\varepsilon/\nu N^2$, the experiments of Turner (1968) and the theory of Holloway (2006) all apply to the heat–salt case; however, as Nash & Moum (2002) and Holloway (2006) discuss, several assumptions were required to obtain the quantitative curves.

For sheared flow, RDT predicts much less differential diffusion than observed in the flow studied by Smyth *et al.* (2005) (figure 13). In their simulations, Smyth *et al.* (2005) set the bulk Richardson number for the billows $Ri_0 = (\Delta b \delta_\rho)/(\Delta U)^2 = 0.1$ and 0.12 – where Δb is the buoyancy jump, ΔU is the velocity jump, and δ_ρ is the thickness of the density interface. These bulk Richardson numbers translate to average gradient Richardson numbers within the interface of about 0.05 to 1 based on the laboratory measurements of Strang (1997). Over the range $0.1 < Ri_g < 1$, RDT overpredicts the diffusivity ratio by 30–60 % with the best agreement at higher Ri_g (figure 13). Differences in the effect of shear might arise owing to differences in the flows; the present RDT applies to a linearly stratified flow with constant shear (or linear velocity profile), while Smyth *et al.* (2005) studied a two-layer flow with shear imposed across the interface.

As discussed in § 3.2, RDT and previous work predict that the diffusivity ratio increases with increasing density ratio. However, although the DNS of Merryfield (2005) and Smyth *et al.* (2005) and the theory of Holloway (2006) and Canuto *et al.* (2002) report effects of the density ratio on d, RDT predicts no significant dependence on the density ratio over the range of $\varepsilon/\nu N^2$ covered in previous studies. Only Merryfield (2005) found density-ratio dependence at large $\varepsilon/\nu N^2$. In the theory of Canuto *et al.* (2002), the effect of the density ratio on differential diffusion decreases with increasing shear, which is consistent with RDT.

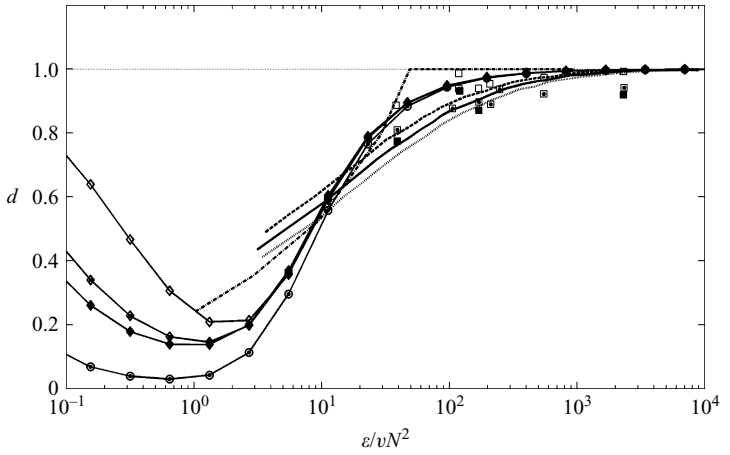

FIGURE 12. Diffusivity ratio as a function of $\varepsilon/\nu N^2$ for unsheared flows. Filled symbols have $R_\rho = 0$, open symbols have $R_\rho \to \infty$, and open symbols with dots have $R_\rho = 1$. The symbols and lines are as follows: circles, RDT ($Le = 0.01$); diamonds, RDT ($Le = 0.1$); squares, Merryfield (2005) ($Le = 0.1$); dash-dot line, Turner (1968) ($Le = 0.01$); solid line, Holloway (2006) ($Le = 0.1$, $R_\rho = 1$); dashed line, Holloway (2006) ($Le = 0.1$, $R_\rho = 2$); dotted line, Holloway (2006) ($Le = 0.1$, $R_\rho = 0.5$). The constant of proportionality of 7 between Gr and $\varepsilon/\nu N^2$ used by Jackson *et al.* (2005) has been applied to the present RDT results.

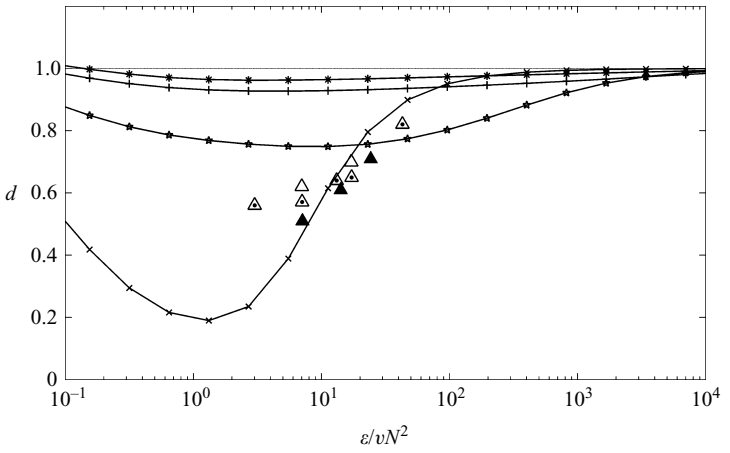

FIGURE 13. Diffusivity ratio as a function of $\varepsilon/\nu N^2$ and Richardson number for $Le = 0.14$: triangles, DNS of K-H billows (Smyth *et al.* 2005); lines, present RDT with $R_\rho = 1$: *, $Ri_g = 0.1$; +, $Ri_g = 0.25$; ☆, $Ri_g = 1$; ×, $Ri_g \to \infty$. DNS data symbols are designated by density ratio: filled symbols have $R_\rho = 0.2$, open symbols have $R_\rho = 5$, and open symbols with dots have $R_\rho = 1$.

The general agreement of the results from various studies suggests that differential diffusion may not depend strongly on the process generating the turbulence, the steadiness of the mixing, or shear in a two-layer flow. Merryfield (2005) drew a similar conclusion when comparing his work to the work of Jackson & Rehmann (2003*b*). However, the sensitivity of differential diffusion in a linearly stratified flow to uniform shear predicted by RDT and the lack of agreement with observations in a two-layer stratified shear layer (Smyth *et al.* 2005) suggests that this generalization may not hold for all flows.

4. Summary and conclusions

We used rapid distortion theory to study why differential diffusion occurs and how changes in properties of the scalars and the flow affect it. In particular, for homogeneous turbulence in a sheared, stratified flow with multiple scalars, we evaluated the effect of the density ratio, the Schmidt numbers of the scalars, the Lewis number, the scalar activity, and shear.

RDT captures the essential mechanism for differential diffusion – excess upgradient flux of the scalar with lower molecular diffusivity – identified by DNS (e.g. Gargett *et al.* 2003). Whereas differential transport increases during periods of upgradient flux, it decreases during periods of downgradient flux. Despite not being directly influenced by gravity, passive scalars can exhibit differential diffusion because of coupling with the active scalars in the flow; the active scalar with the larger Schmidt number sets the limit for the flux and ultimately the differential transport of the passive scalar. Shear reduces the magnitude of upgradient fluxes, increases damping of the flux oscillations, and generates persistent upgradient fluxes. The net effect on differential diffusion depends on the Grashof number, or the relative importance of buoyancy and viscous effects.

RDT also allows a study of scalars with high Schmidt numbers without the high computational cost of DNS. Therefore, effects of parameters varied over wide ranges can be explored. Differential diffusion increases with increasing density ratio; in RDT the effects were confined to low Grashof number and large Richardson number, but DNS and other theory yield density ratio dependence at higher Gr. The Schmidt numbers control the Grashof number below which differential diffusion occurs; for fixed Lewis number, the critical value of Gr decreases with increasing Schmidt number. With shear, the magnitude of differential diffusion decreases sharply with increasing Schmidt number. When one of the Schmidt numbers is fixed, differential diffusion decreases with increasing Lewis number. Also, differential transport of passive scalars increases when the Schmidt number of the scalar stratifying the flow increases.

The present work and the work of Jackson *et al.* (2005) demonstrate the usefulness of RDT in studies of differential diffusion. Unlike simple eddy-diffusivity models and two-equation turbulence models, RDT can predict upgradient fluxes and their wavenumber dependence which are essential in differential diffusion. Therefore, as Jackson *et al.* (2005) recommended, RDT could be used to develop subgrid-scale models of weakly turbulent, strongly stratified flows in the interior of oceans and lakes. Because most biological and chemical scalars have high Schmidt numbers (Socolofsky & Jirka 2002; Gargett 2003), RDT would be a useful tool in modeling diapycnal biogeochemical transport in weakly turbulent stratified flows.

REFERENCES

Altman, D. B. & Gargett, A. E. 1990 Differential property transport due to incomplete mixing in a stratified fluid. In *Proc. 3rd Intl Symp. on Stratified Flows* (ed. E. J. List & G. H. Jirka), pp. 454–460. American Society of Civil Engineers, New York.

Baker, M. A. & Gibson, C. H. 1987 Sampling turbulence in the stratified ocean: statistical consequences of strong intermittency. *J. Phys. Oceanogr.* **17**, 1817–1836.

Bilger, R. W. 1989 Turbulent diffusion flames. *Annu. Rev. Fluid Mech.* **21**, 101–135.

Canuto, V. M., Howard, A., Cheng, Y. & Dubovikov, M. S. 2002 Ocean turbulence. Part II: Vertical diffusivities of momentum, heat, salt, mass, and passive scalars. *J. Phys. Oceanogr.* **32**, 240–264.

De Silva, I. P. D. & Fernando, H. J. S. 1992 Some aspects of mixing in a stratified turbulent patch. *J. Fluid Mech.* **240**, 601–625.

Gargett, A. E. 2003 Differential diffusion: an oceanographic primer. *Prog. Oceanogr.* **56**, 559–570.

Gargett, A. E., Merryfield, W. J. & Holloway, G. 2003 Direct numerical simulation of differential scalar diffusion in three-dimensional stratified turbulence. *J. Phys. Oceanogr.* **33**, 1758–1782.

Gerz, T. & Schumann, U. 1991 Direct simulation of homogeneous turbulence and gravity waves in sheared and unsheared stratified flows. *Turbulent Shear Flows 7*, pp. 27–45. Springer.

Gerz, T. & Yamazaki, H. 1993 Direct numerical simulation of buoyancy-driven turbulence in stably stratified fluid. *J. Fluid Mech.* **249**, 415–440.

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

Hanazaki, H. 2003 Effects of initial conditions on the passive and active scalar fluxes in unsteady stably stratified turbulence. *Phys. Fluids* **15**, 841–848.

Hanazaki, H. & Hunt, J. C. R. 1996 Linear processes in unsteady stably stratified turbulence. *J. Fluid Mech.* **318**, 303–337.

Hanazaki, H. & Hunt, J. C. R. 2004 Structure of unsteady stably stratified turbulence with mean shear. *J. Fluid Mech.* **507**, 1–42.

Hebert, D. 1999 Intrusions: What drives them? *J. Phys. Oceanogr.* **29**, 1382–1391.

Hebert, D. & Ruddick, B. R. 2003 Differential mixing by breaking internal waves. *Geophys. Res. Lett.* **30**, 1042–1045.

Hindmarsh, A. C. 1983 ODEPACK, A systematized collection of ODE solvers. In *Scientific Computing* (ed. R. S. Stepleman) pp. 55–64. North-Holland, Amsterdam.

Holloway, G. 2006 Statistically stationary differential diffusion in a large-scale internal waves-vortical modes environment. *Deep Sea Res. II* **53**, 116–127.

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.

Hunt, J. C. R., Stretch, D. D. & Britter, R. E. 1988 Length scales in stably stratified turbulent flows and their use in turbulence models. In *Stably Stratified Flow and Dense Gas Dispersion* (ed. J. S. Puttock), pp. 285–321. Clarendon.

Hwang, J. H., Yamazaki, H. & Rehmann, C. R. 2006 Buoyancy generated turbulence in stably stratified flow with shear. *Phys. Fluids* **18**, 045104.

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.

Jackson, P. R. 2006 Differential diffusion of scalars in sheared, stratified turbulence. PhD thesis, University of Illinois at Urbana-Champaign.

Jackson, P. R. & Rehmann, C. R. 2003*a* Kinematic effects of differential transport on mixing efficiency in a diffusively stable, turbulent flow. *J. Phys. Oceanogr.* **33**, 299–304.

Jackson, P. R. & Rehmann, C. R. 2003*b* Laboratory measurements of differential diffusion in a diffusively stable, turbulent flow. *J. Phys. Oceanogr.* **33**, 1592–1603.

Jackson, P. R., Rehmann, C. R., Saenz, J. A. & Hanazaki, H. 2005 Rapid distortion theory for differential diffusion. *Geophys. Res. Lett.* **32**, L10601, doi:10.1029/2005GL022443.

Joseph, B., Mahalov, A., Nicolaenko, B. & Leung Tse, K. 2004 Variability of turbulence and its outer scales in a model tropopause jet. *J. Atmos. Sci.* **61**, 621–643.

King, D. B. & Saltzman, E. S. 1995 Measurement of the diffusion coefficient of sulfur hexafluoride in water. *J. Geophys. Res.* **100**, 7083–7088.

Komori, S. & Nagata, K. 1996 Effects of molecular diffusivities on counter-gradient scalar and momentum transfer in strongly stable stratification. *J. Fluid Mech.* **326**, 205–237.

Ledwell, J. R., Watson, A. J. & Law, C. S. 1993 Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature* **364**, 701–703.

Liu, H.-T. 1995 Energetics of grid turbulence in a stably stratified fluid. *J. Fluid Mech.* **296**, 127–157.

Liu, Y. N., Maxworthy, T. & Spedding, G. R. 1987 Collapse of a turbulent front in a stratified fluid 1. Nominally two-dimensional evolution in a narrow tank. *J. Geophys. Res.* **92**(C5), 5247–5433.

Martin, J. E. & Rehmann, C. R. 2006 Layering in a flow with diffusively stable temperature and salinity stratification. *J. Phys. Oceanogr.* **36**, 1457–1470.

Merryfield, W. J. 2005 Dependence of differential mixing on N and R_ρ. *J. Phys. Oceanogr.* **35**, 991–1003.

Merryfield, W. J., Holloway, G. & Gargett, A. E. 1998 Differential vertical transport of heat and salt by weak stratified turbulence. *Geophys. Res. Lett.* **25**, 2773–2776.

Nash, J. D. & Moum, J. N. 2002 Microstructure estimates of turbulent salinity flux and the dissipation spectrum of salinity. *J. Phys. Oceanogr.* **32**, 2312–2333.

Nazarenko, S., Kevlahan, N. K.-R. & Dubrulle, B. 1999 WKB theory for rapid distortion of inhomogeneous turbulence. *J. Fluid. Mech.* **390**, 325–348.

Rehmann, C. R. 1995 Effects of stratification and molecular diffusivity on the mixing efficiency of decaying grid turbulence. PhD thesis, Stanford University.

Rehmann, C. R. & Hwang, J. H. 2005 Small-scale structure of strongly stratified turbulence. *J. Phys. Oceanogr.* **35**, 151–164.

Riley, J. J. & Lelong, M. P. 2000 Fluid motions in the presence of strong stable stratification. *Annu. Rev. Fluid Mech.* **32**, 613–657.

Shih, L. H., Koseff, J .R., Ferziger, J. H. & Rehmann, C. R. 2000 Scaling and parameterization of stratified homogeneous turbulent shear flow. *J. Fluid Mech.* **412**, 1–20.

Shih, L. H., Koseff, J. R., Ivey, G. N., & Ferziger, J. H. 2005 Parameterization of turbulent fluxes and scales using homogeneous sheared stably stratified turbulence simulations. *J. Fluid Mech.* **525**, 193–214.

Smyth, W. D., Nash, J. D. & Moum, J. N. 2005 Differential diffusion in breaking Kelvin–Helmholtz billows. *J. Phys. Oceanogr.* **35**, 1004–1022.

Socolofsky, S. A. & Jirka, G. H. 2002 *Environmental Fluid Mechanics Part I: Mass Transfer and Diffusion*, 2nd edn. Institut für Hydromechanik, Universität Karlsruhe, Germany.

Spedding, G. R., Browand, F. K. & Fincham, A. M. 1996 The long-time evolution of the initially-turbulent wake of a sphere in a stable stratification. *Dyn. Atmos. Oceans* **23**, 171–182.

Strang, E. J. 1997 Entrainment and mixing in stratified shear flows. PhD thesis, Arizona State University.

Townsend, A. A. 1980 *The Structure of Turbulent Shear Flow*. Cambridge University Press.

Turner, J. S. 1968 The influence of molecular diffusivity on turbulent entrainment across a density interface. *J. Fluid Mech.* **33**, 639–656.