Direct numerical simulation of differential scalar diffusion in three-dimensional stratified turbulence

Ann E. Gargett\textsuperscript{1}, William J. Merryfield\textsuperscript{2} and Greg Holloway\textsuperscript{3}

\textsuperscript{1} Center for Coastal Physical Oceanography
Old Dominion University
Norfolk VA USA

\textsuperscript{2} Canadian Centre for Climate Modelling and Analysis
Victoria B.C. Canada

\textsuperscript{3} Institute of Ocean Sciences
Sidney B.C. Canada

accepted: February 3, 2003
Abstract:

The potential for differential turbulent transport of oceanic temperature (T) and salinity (S) is explored using three-dimensional direct numerical simulations of decaying stratified turbulence. The simulations employ a realistic molecular diffusion coefficient for T, and one for a “salt” scalar S that is 10 times smaller. Initially, a uniformly stratified medium is disturbed by a turbulent burst whose initial energy is assigned a range of values. In each instance, transports of T integrated over the subsequent decay of the burst exceed those of S. The more energetic cases occupy parameter ranges similar to, and exhibit spectral characteristics that are essentially indistinguishable from, those of direct observations of turbulence in the stratified ocean interior. In these cases, the turbulent diffusivity of T exceeds that of S by 6-22%. These simulations underestimate the degree of differential diffusion between T and salinity S (which has a molecular diffusivity 100 times less that T); thus at the Reynolds numbers attained by the simulations these results constitute lower bounds for differential diffusion associated with sporadic turbulence in the ocean. The simulation results are consistent with previous laboratory and two-dimensional numerical experiments, and suggest that the assumption of equal turbulent diffusivities for T and S, commonly used in circulation modelling and in interpreting oceanic mixing measurements, should be reconsidered.

1. Introduction

The assumption that turbulence mixes all scalar water properties similarly permeates many aspects of the measurement and interpretation of small-scale turbulence and its effects in the ocean. All methods presently used to estimate oceanic turbulent vertical (more accurately diapycnal) diffusivity from observations make this assumption. Measurements of temperature microscales are interpreted (using the framework of Osborn and Cox 1972) as a diffusivity appropriate not only for temperature but for all scalars, as are diffusivity
estimates made from microscale shear measurements via the method of Osborn (1980). Measurements of the vertical spread with time of purposeful releases of an inert tracer (most commonly sulphur hexafluoride: Ledwell and Watson 1991), are often interpreted as “the” turbulent vertical diffusivity characteristic of the release location (unless the background water column is unstable to double diffusive processes, St. Laurent and Schmitt 1998). However there is growing evidence that even in doubly-stable water columns, turbulence may preferentially mix a scalar with higher molecular diffusivity relative to one with lower molecular diffusivity. This process, termed differential diffusion, is a potential complication for these now-standard conversions from ocean measurements to the parameterizations incorporated in ocean models. As well, since chemical species are expected to behave like salinity (S) which has a molecular diffusivity 100 times smaller than that of temperature (T) in seawater, relatively slower diffusion of S would further decrease already low estimates (Lewis et al. 1984) of turbulent resupply of biologically essential “new” nutrients to the ocean euphotic zone. Accounting for differential diffusion may be particularly important in settings such as high latitude oceans and estuaries, where density structure is dominated by salinity.

Reviewing the process of differential diffusion in an oceanic context, Gargett (2003) discusses existing laboratory evidence, as well as ocean observations which exhibit possible effects of the differential diffusion of T and S, the basic stratifying scalar properties of seawater. However laboratory experiments are often considered “un-oceanic” because of the sometimes artificial nature of turbulent forcing, sensitivity of some results to seemingly minor details of the physical set-up, etc. Moreover, oceanic observations of effects that would be expected to result from differential diffusion may be suggestive, but can’t be viewed as proof. Numerical simulation offers a third method of investigating differential diffusion, a method with potential to provide the quantitative measures of differential fluxes that are necessary to test interpretation of oceanic observations.
To date, differential diffusion has been modelled by methods such as the linear eddy technique of Kerstein (1988), as well as by direct numerical simulation (DNS). The Kerstein technique simulates the stirring effects of turbulent eddies by a one-dimensional stochastic rearrangement process, while treating molecular diffusion by deterministic application of Fick’s Law. The fact that the velocity field is not calculated, coupled with the restriction in dimension, allows this method to resolve all relevant spatial scales for free shear “flows” with Reynolds number up to $O(10^4)$. Despite the sacrifice of three-dimensionality, the linear-eddy model has been successful in capturing qualitative features of mixing in a variety of flows. Kerstein (1990) applies this method to the case of H$_2$ and propane mixing in a round jet: while his results are in qualitative agreement with measurements by Bilger and Dibble (1982), there are significant quantitative discrepancies, possibly associated with missing effects of three-dimensionality.

The alternate “brute-strength” technique of DNS ideally resolves spatial scales from those at which energy is input to turbulence, through the dissipation scales of all scalars. The scale range involved depends upon the turbulence outer (energy-containing) scale, the turbulent kinetic energy dissipation rate $\varepsilon$, and the scalar Schmidt number $Sc_\theta \equiv \nu/D_\theta$, where $\nu$ is the fluid kinematic viscosity and $D_\theta$ the molecular diffusivity of scalar $\theta$. The turbulence outer scale is determined by large-scale properties of the flow, as imposed either by external forcing or natural instabilities, and by the density stratification $\partial \rho / \partial z$, normally quantified by the buoyancy (Brunt-Väisälä) frequency $N \equiv \left(-g \rho_0^{-1} \partial \rho / \partial z \right)^{1/2}$ or the associated buoyancy period $T_N = 2\pi/N$, where $g$ is acceleration due to gravity and $\rho_0$ is a reference density. In the stably-stratified ocean interior, it is generally accepted that the outer scale is roughly the length scale $L_\varepsilon \equiv (\varepsilon/N^3)^{1/2}$ that arises from an assumed balance between inertial and buoyancy forces (Dougherty 1961: Ozmidov 1965). Expressed in terms of wavenumber, DNS should thus resolve a lowest wavenumber $k_L \sim L_\varepsilon^{-1}$ which generally decreases with $\varepsilon$ (Re) and increases with stability.
The required high wavenumber resolution is determined by the scalar Schmidt number. Both of the oceanographically important scalars \( T \) and \( S \) are characterized by Schmidt numbers that are substantially larger than 1 (although \( \text{Sc}_T \) is more commonly called the Prandtl number (Pr), we adopt a generalized notation). Taking \( v \sim 10^{-6} \text{ m}^2\text{s}^{-1} \), \( D_T \approx 1.4 \times 10^{-7} \text{ m}^2\text{s}^{-1} \) and \( D_S \approx D_T/100 \) as characteristic of seawater, \( \text{Sc}_T \equiv v/D_T \approx 7 \) and \( \text{Sc}_S \equiv v/D_S \approx 700 \). For scalars with \( \text{Sc}_\theta >> 1 \), Batchelor (1959) argued that scalar gradient variance is in a steady state where dissipation by molecular forces is balanced by generation through deformation by a locally plane rate-of-strain \( \gamma \sim (\varepsilon/v)^{1/2} \) characteristic of the velocity shear field at wavenumbers above the Kolmogoroff wavenumber \( k_v = (\varepsilon/v^3)^{1/4} \) (Kolmogoroff 1941). This hypothesized scalar gradient balance predicts that a wavenumber

\[
\kappa_{\text{Br}} = \left( \frac{\varepsilon}{v D_\theta^2} \right)^{1/4} = \text{Sc}_{\theta}^{1/2} k_v > k_v \text{ for } \text{Sc}_{\theta} > 1
\]

now known as the Batchelor wavenumber, is characteristic of the smallest scales of gradient variance of scalar \( \theta \). Scalars with \( \text{Sc}_\theta >> 1 \) thus increase the high-wavenumber resolution required of DNS, sometimes substantially, as in the \( \text{Sc}_S \approx 700 \) case of \( S \) in seawater.

Inevitably, the DNS approach to the study of scalar differential diffusion faces computational limitations on the ranges of \( \text{Re} \) and \( \text{Sc}_\theta \) that can be considered, limitations which have been addressed in a variety of ways. By restricting attention to scalars with \( \text{Sc}_\theta \leq 1 \), computational requirements are eased. For isotropic turbulence in an unstratified fluid, Yeung and Pope (1993) were able to reach a Taylor-scale Reynolds number of \( \text{Re}_\lambda = 38 \) with a 64\(^3\) simulation. Their results show that two passive scalars with identical initial conditions but different molecular diffusivities decorrelate rapidly over short time scales, say \( t < 5 T_\epsilon \) where \( T_\epsilon \) is an eddy turnover time. Decorrelation then slows, but (statistically) complete decorrelation is eventually achieved by \( \sim 50 T_\epsilon \).
Spectral decomposition reveals that incoherency develops first at the smallest scales (highest wavenumbers), presumably as a result of the direct dependence of such scales on the (different) molecular diffusion coefficients, then spreads to larger scales (smaller wavenumbers). Yeung (1996) extends such calculations to $\text{Re}_\lambda = 160$ using a $(256)^3$ grid.

In order to treat the high Schmidt number case of the oceanic scalars $T$ and $S$, Merryfield et al. (1998) were forced to reduce the dimension of calculation to two. While their results revealed the action of differential diffusion, they are vulnerable to the criticism that two-dimensional turbulence has energy cascade properties that differ fundamentally from those of three-dimensional turbulence. The present work addresses this potential weakness by carrying out fully three-dimensional simulations. To do so with available computer resources, we consider a scalar $S$ (which we call “salt”) that is characterized by a diffusivity ratio $D_s/D_T = 0.1$ rather than the value $D_s/D_T = 0.01$ characteristic of temperature and salinity in seawater.

Sec. 2 describes the formulation of the DNS and the numerical techniques used, while Sec. 3 presents an overview of model results, examining the time evolution of significant diagnostic quantities over the range of forcing that was achievable. Since the impetus for this work is the potential implication(s) for mixing in the ocean, Sec. 4 addresses the relevance of the simulations to the oceanic case, comparing model diagnostics and spectra with similar results obtained from ocean microscale measurements. We next quantify the amount of differential diffusion observed over the range of simulations (Sec.5(a)), and (Sec.5(b)) compare our three-dimensional results with extensions of previous two-dimensional results of Merryfield et al. (1998), in order to assess the effect of the larger oceanic diffusivity ratio. Finally, Sec.6 discusses the physical mechanism underlying differential diffusion of high Schmidt number scalars, as suggested by the observed time evolution of scalar flux spectra in our simulations.
2. Model formulation

We consider the Boussinesq equations:

\[
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho_o} \nabla p + \frac{\rho}{\rho_o} \mathbf{g} + \nu \nabla^2 \mathbf{u} \tag{2}
\]

\[
\nabla \cdot \mathbf{u} = 0 \tag{3}
\]

\[
\frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla) T = D_T \nabla^2 T \tag{4}
\]

\[
\frac{\partial S}{\partial t} + (\mathbf{u} \cdot \nabla) S = D_S \nabla^2 S \tag{5}
\]

where \( \mathbf{u} = (u, v, w) \) is velocity, \( T \) is temperature, \( S \) is salt, \( p \) is pressure, \( \rho \) is density, \( \mathbf{g} \) is gravitational acceleration, \( \nu \) is viscosity, and \( D_T \) and \( D_S \) are molecular diffusion coefficients for temperature and salt. \( T \) and \( S \) are related to density \( \rho \) through a linear equation of state:

\[
\rho = \rho_o [1 - \alpha(T - T_o) + \beta(S - S_o)] \tag{6}
\]

where \( \alpha \) and \( \beta \) are coefficients of thermal expansion and haline contraction, and \( \rho_o, T_o \) and \( S_o \) are constant reference values. \( T \) and \( S \) are “active” scalars: their variation produces variation in \( \rho \), hence affects the motion of the fluid through the term involving \( \rho \) in (2).

We adopt Cartesian coordinates \((x, y, z)\), with vertical unit vector \( \mathbf{\hat{z}} \parallel \mathbf{g} \), and express \( \mathbf{u} \) in terms of toroidal \((Z)\), poloidal \((W)\) and horizontal mean \((\bar{u}(z))\) components according to

\[
\mathbf{u} = \text{curl}(Z \mathbf{\hat{z}}) + \text{curl}^2(W \mathbf{\hat{z}}) + \bar{u}(z)
\]

\[
= (Z_y + W_{xz} + \bar{u}, -Z_x + W_{yz} + \bar{v}, -(W_{xx} + W_{yy})) \tag{7}
\]

(e.g. Moffatt 1978) which satisfies (3) exactly for arbitrary \( Z, W \) and \( \bar{u} \). This decomposition provides a convenient means for representing non-divergent
three-dimensional vector fields, and eliminates \( p \) from the governing equations. It has previously been applied in simulating thermal convective motions in spherical geometry (e.g. Glatzmaier 1984), and is adapted here for use with three-dimensional Cartesian coordinates. Mathematical properties of toroidal and poloidal functions are summarized in Chandrasekhar (1961).

Equations for \( Z \) and \( W \) are obtained from the z-components of curl and \( \text{curl}^2 \) of (2), using

\[
(curlu)_z = \nabla_h^2 Z, \quad \text{and} \quad \left(curl^2u\right)_z = \nabla_h^2 \psi
\]

where \( \nabla_h^2 = \partial_x^2 + \partial_y^2 \) and

\[ \psi \equiv \nabla^2 W \] (9)

An equation for \( \bar{u} \) is obtained by horizontally averaging (2).

We write

\[
T(x,y,z) = N_T z + T'(x,y,z)
\] (10)

\[
S(x,y,z) = -N_S z + S'(x,y,z)
\] (11)

where \( N_T > 0 \) and \( -N_S < 0 \) are uniform (stable) background vertical gradients of \( T \) and \( S \), and \( T' \) and \( S' \) represent departures from that state. Scaling length by \( d \equiv (\nu D_T/g \Gamma)^{1/4} \), time by \( d^2/D_T \), temperature by \( N_T d \) and salt by \( N_S d \), we obtain

\[
\frac{\partial}{\partial t} \nabla_h^2 Z = \hat{z} \cdot \text{curl}(u \times \text{curl} u) + \sigma \nabla^2 (\nabla_h^2 Z)
\] (12)

\[
\frac{\partial}{\partial t} \nabla_h^2 \psi = \hat{z} \cdot \text{curl}^2(u \times \text{curl} u) + \sigma \nabla^2 (\nabla_h^2 \psi) + \sigma \nabla^2 (-T' + S'R_p^{-1})
\] (13)

\[
\frac{\partial}{\partial t} \bar{u} = -u \times (\nabla \times u) + \nu \nabla^2 \bar{u}
\] (14)

\[
\frac{\partial}{\partial t} T' = -w - u \cdot \nabla T' + \nabla^2 T'
\] (15)

\[
\frac{\partial}{\partial t} S' = w - u \cdot \nabla S' + \tau \nabla^2 S'
\] (16)
where $\sigma = \nu / D_T \equiv Sc_T$, $\tau = D_S / D_T = Le^{-1}$ is the inverse Lewis number, $R_p = \alpha N_T / \beta N_S$ is a gradient form of the buoyancy ratio (Turner 1973), defined positive for doubly stable stratification, as here. In (15) and (16), the terms proportional to $w$ represent the production of $T'$ and $S'$ due to advection across the background gradients.

We consider a cubic computational domain with dependent variables, periodic in $x$, $y$ and $z$, describing departures from a quiescent state. Equations (9) and (12)-(16) are solved pseudospectrally, with de-aliasing according to the 2/3 rule (Canuto et al. 1988). Dissipation terms are represented in Fourier space by exponential integration factors. Temporal integration is via the leapfrog method, stabilized by a Robert filter (Robert 1966) with filter parameter 0.002. Because $S$ is a factor of 10 less diffusive than $T$, the need to resolve the smallest spatial scales of $S'$ dictates the overall computational cost. This cost is mitigated in most runs by allocating to $S'$ twice the spatial resolution used for $u$ and $T'$ (Table 1). In solving (16) spectrally, velocity variables are interpolated to the finer grid by padding the low-resolution three-dimensional $u$ spectrum with zeros, whereas $S'$ in (13) is obtained by truncating the high-resolution spectrum at the lower resolution.

Although the equations solved numerically are non-dimensional, we will occasionally present results in dimensional form to enable comparison with ocean measurements. In such instances we adopt typical values of $\nu = 10^{-2}$ cm$^2$ s$^{-1}$, $D_T = 1.4 \times 10^{-3}$ cm$^2$ s$^{-1}$, $g = 980$ cm s$^{-2}$, $\alpha = 2 \times 10^{-4}$ °C$^{-1}$, and $N_T = 10^{-4}$ °C cm$^{-1}$ (1°C/100m). These choices yield dimensional length scale $d = 0.92$ cm, domain size $6\pi d = 17.3$ cm, and buoyancy frequency $N \approx 6 \times 10^{-3}$ s$^{-1}$.

Initially, a specified velocity field acts upon previously undisturbed stratification ($T' = S' = 0$). Two types of initial $u$ are considered. In the first instance, Fourier coefficients of $W$, $Z$ and $\vec{u} = (\vec{u}, \vec{v}, 0)$ are assigned initial values drawn randomly from normal distributions, and scaled according to $W_k = C_E k^{-4}$, and $Z_k = (u_k, v_k) = 0$, where $k$ is wave vector, $k = |k|$, and the factor $C_E$ is Table 1. Properties of model runs. For comparison with two-dimensional results of Merryfield et al. (1998), initial Reynolds number $Re_0 \equiv U_0 L_0 / \sigma$ and Froude
number \( Fr_0 \equiv U_0/NL_0 \) are computed using their definitions of eddy velocity scale \( U_0 \equiv (2E_0)^{1/2} \) and eddy length scale \( L_0 \equiv \pi(E_0/Z_0)^{1/2} \), with non-dimensional energy \( E_0 = \frac{1}{2} \int |\psi_0|^2 dx \) and \( Z_0 = \frac{1}{2} \int |
abla \psi_0|^2 dx \), evaluated at \( t=0 \).

<table>
<thead>
<tr>
<th>Run</th>
<th>T resolution(^a)</th>
<th>S resolution(^a)</th>
<th>( E_0 )</th>
<th>Type</th>
<th>Re(_0)</th>
<th>Fr(_0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3.0</td>
<td>((64)^3)</td>
<td>((64)^3)</td>
<td>(10^3)</td>
<td>anisotropic</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>A4.0</td>
<td>((64)^3)</td>
<td>((64)^3)</td>
<td>(10^4)</td>
<td>anisotropic</td>
<td>31</td>
<td>12</td>
</tr>
<tr>
<td>A4.5</td>
<td>((128)^3)</td>
<td>((256)^3)</td>
<td>(10^4.5)</td>
<td>anisotropic</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>A5.0</td>
<td>((128)^3)</td>
<td>((256)^3)</td>
<td>(10^5)</td>
<td>anisotropic</td>
<td>99</td>
<td>39</td>
</tr>
<tr>
<td>I5.0H</td>
<td>((160)^3)</td>
<td>((320)^3)</td>
<td>(10^5)</td>
<td>isotropic</td>
<td>79</td>
<td>48</td>
</tr>
<tr>
<td>I5.0M</td>
<td>((128)^3)</td>
<td>((256)^3)</td>
<td>(10^5)</td>
<td>isotropic</td>
<td>79</td>
<td>48</td>
</tr>
<tr>
<td>I5.0L</td>
<td>((96)^3)</td>
<td>((192)^3)</td>
<td>(10^5)</td>
<td>isotropic</td>
<td>79</td>
<td>48</td>
</tr>
</tbody>
</table>

\(^a\) collocation points with respect to \( x, y \) and \( z \)

adjusted to yield the desired initial energy. This formulation yields an initial velocity field in which the root-mean-square (rms) vertical velocity is approximately twice the rms horizontal velocity components. In the second instance, \( W_k = C_E k^{-4} \) as before, but \( Z_k = C_E k^{-3} \), and \( (\bar{u}_k, \bar{v}_k) = C_E k^{-2} \), chosen to yield an initial velocity field that is statistically isotropic. Most of the runs reported here use the anisotropic initial velocity field, which apportions a larger fraction of kinetic energy to vertical motions than does an isotropic field. However, as did Siegal and Domaradzki (1994), we observe that in all except the weakest initial energy run, the velocity field quickly achieves dissipation-scale isotropy (see Sec.3(c)). As well, runs with initially isotropic conditions but the same initial energy level produce measures of differential diffusion which differ only slightly from those of the initially anisotropic cases. In both cases \( E(k) \propto k^{-3} \), so initial kinetic energy is strongly concentrated at the lowest wavenumbers. This initial energy spectral slope was chosen to be much steeper than the \(-5/3\) slope characteristic of an inertial subrange because (i) the largest spatial scale of our simulation corresponds to wavenumbers only slightly smaller than the low-
wavenumber end of the velocity dissipation range for achieved values of $\varepsilon$, and
(ii) one expects steeper slopes at the beginning of an instability, when there has been as yet no cascade to smaller scales.

We focus on results from the first four model runs listed in Table 1: these are referenced as, for example, A3.0 where “A” denotes anisotropic initial conditions and 3.0 is the exponent $p$ in non-dimensional initial energy $E_0 = 10^p$. With the computational resources available, the maximum feasible value of $p$ was 5; higher initial energies would lead to S microstructure on spatial scales too small for us to resolve. The second tier of runs listed in Table 1 are used in Appendix 1 to document the degree to which model results are sensitive to spatial resolution, using a case of isotropic (I) forcing with $E_0 = 10^5$: H is the run with the highest attained resolution, M and L are runs with lower resolutions.

All computations use $\alpha N_T = \beta N_S$ ($R_\rho = 1$), so that T and S contribute equally to the stable stratification. With $\sigma = 7$, our fast-diffusing scalar is representative of temperature in seawater. Finite computing resources limited the value of $\tau$ which could be achieved to $\tau = 0.1$. Since salinity S is ten times less diffusive than our slow-diffusing salt scalar, our results may be expected to provide a lower bound for the potential effects of differential diffusion in the oceanic context.

3. Evolution of simulated three-dimensional turbulence

We first describe general characteristics of the turbulent fields generated in the simulations, using flow visualizations and time series plots of various volume-averaged quantities. To enable the direct comparisons between present DNS results and oceanic observations that are the focus of the following section, dissipation-scale parameters presented here are calculated as they are routinely calculated from ocean microscale measurements, rather than as they could be calculated from the more complete information available from the simulations.
Figure 1: Left: Model fields of non-dimensional S, -T and density $\rho$ ($\rho = SR_p^{-1} - T$, with $R_p = 1$), sampled at the vertical mid-plane of the calculation domain during Run A4.5. Right: Vertical profiles of instantaneous (red) and stable-sorted (blue) density at the mid-point of the box, and resulting estimates of Thorpe displacement scale $L_T$ (both vertical coordinate and Thorpe displacement are in units of number of grid points: one grid point = $6\pi d/n \approx 0.13$ cm). Panel numbers at far left are used to reference data to stages of flow development. Time $t_N$ shown between the two sets of panels is time non-dimensionalized by the buoyancy period $T_N$ characteristic of the background (linear) stratification.
(a) Visualizations of the evolution of $T$ and $S$ fields

We describe qualitative characteristics of the turbulent fields generated in the simulations using two forms of visualization to illustrate time evolution of the two scalar fields during Run A4.5, a run in which $E_0$ is large enough that the initial flows are “turbulent” by several standards (see Sec.4), yet small enough that both scalar fields are well-resolved throughout the evolution (see Appendix 1).

In Figure 1, colour-coded panels show the fields of non-dimensional $S$, $-T$ and density $\rho = S - T$ in the vertical plane at the middle of the computational box (note that background gradients have been added for this presentation, although the equations solved are those for the departure fields $T'$ and $S'$). To the right of each set of panels in Fig.1 are two vertical profile plots, the first showing vertical profiles at the box mid-point of (red) instantaneous density and (blue) these density values sorted (with minimum displacement) to stability. The minimum displacements necessary to produce the stable density profile from the instantaneous profile are the Thorpe displacement scales $L_T$ (Thorpe 1977), shown (far left) to illustrate the scales of density inversions driven by the kinetic energy originally supplied to the box. Time evolution of the two tracers would be identical if they had identical molecular diffusivities. Instead, more rapid smoothing of $T$ fluctuations is evident at all but the very earliest stages of scalar evolution.

A second view of the differences in scalar field evolution is provided by Figure 2, which documents the time behaviour of $S$ and $T$ iso-surfaces that are initially located at mid-depth in the computational box: numbers on the figure correspond to the panel numbers in Fig.1. Both iso-surfaces are flat and featureless at $t = 0$ (not shown). Development of spatial structure in the iso-surfaces is initially similar for both tracers: however these structures diverge with time. As expected, the $S$ iso-surface develops much finer spatial structure than is ever observed on the $T$ iso-surface. The $T$ iso-surface also loses any evidence of
vertical inversions (revealed by detached “tendrils” of the surface) and becomes smooth more rapidly than does the S iso-surface.

Figure 2: Time evolution of S and T iso-surfaces that were initially horizontal and located at the vertical mid-point of the computational box. Numbers noted between pairs of surfaces correspond to the panel numbers of the vertical-plane realizations shown in Fig.1.
Figure 3: Time histories of volume-averaged scalar variances $\sigma_T^2$ (solid) and $\sigma_S^2$ (dashed) for $T$ and $S$ respectively, from simulations (a) A3.0 (b) A4.0 (c) A4.5 (d) A5.0, shown in order of increasing initial energy $E_0$, as listed in Table 1. Evolution is shown as a function of $t_N$, time non-dimensionalized by the buoyancy period $T_N$ characteristic of the background (linear) stratification.

(b) Scalar variances

Figure 3(a-d) shows the time evolution of scalar variances $\sigma_T^2 = \langle(T')^2\rangle$ and $\sigma_S^2 = \langle(S')^2\rangle$ (where angle brackets denote averaging over the computational volume) as a function of a buoyancy-scaled run time $t_N \equiv t/T_N$ : while all simulations ran beyond $t_N \sim 1.2$, further change was extremely slow, hence is not shown. Here, as in subsequent sets of plots, the simulations are arranged in order of increasing initial energy $E_0$, from (a) to (d). Both scalar variances start at zero, rise to a peak as the imposed velocity field generates variance through overturning motions, then decline monotonically to zero as overturning motions collapse and both turbulent and molecular diffusion smooth
out the small-scale structures. This behaviour can be related to the previous qualitative descriptions of Run A4.5 through the labels along the upper horizontal axis of Fig. 3(c), which correspond to the panel numbers in Figs.1 and 2. As expected, scalar variance magnitudes increase as initial energy increases. Note however that in all cases, even the weakest (Fig.3(a)), T and S variances do not show identical time evolution, straightforward evidence of differential diffusion of the two scalars.

(c) Turbulent kinetic energy dissipation rates

Most of the observational information about ocean turbulence has come from instruments carried on vertical profilers. In this configuration, orthogonal airfoil probes (Osborn 1974) provide measurements of \( (\partial u/\partial z)^2 \) and \( (\partial v/\partial z)^2 \), where the averaging denoted by double overbars is over vertical interval \( \Delta z \). These measured variances are transformed to two independent estimates of turbulent kinetic energy dissipation rate \( \varepsilon \) using the isotropic formula

\[
\varepsilon_{ij} = 7.5 \nu (\partial u_i / \partial x_j)^2, \quad i \neq j,
\]

assuming that the vertical averaging produces an appropriate ensemble (single overbar) average. We calculate similar estimates, respectively termed \( e_u \) and \( e_v \), from volume \( (\langle \rangle \) averages of the simulation shear variances \( \langle (\partial u/\partial z)^2 \rangle \) and \( \langle (\partial v/\partial z)^2 \rangle \), using averaging over the two
Figure 4: Top curves in each panel: time histories of estimates of the (logarithm of) turbulent kinetic energy dissipation rate, made from isotropic formulae using volume-averaged vertical gradients of vertical velocity, \( \log e_v (+) \), and horizontal velocity, \( \log e_h (x) \). The horizontal line corresponds to \( \log(\varepsilon_c = 25\nu N^2) \): the criterion \( \varepsilon > \varepsilon_c \) is frequently used to designate “active” turbulence. Lower curve in each panel: time history of \( A_e \equiv e_h/e_v \), a measure of dissipation scale isotropy. Simulations (a) A3.0 (b) A4.0 (c) A4.5 (d) A5.0 are in order of increasing initial energy \( E_0 \), see Table 1. Evolution is shown as a function of \( t_N \), time non-dimensionalized by the buoyancy period \( T_N \) characteristic of the background (linear) stratification.

horizontal dimensions as well as the vertical dimension to partially compensate for the restricted vertical extent of the computational domain relative to the value of \( \Delta z \sim 1-2 \) m usual for the observational estimates. In subsequent calculations of
derived quantities involving \( \varepsilon \), we use \( \varepsilon = e_h = 0.5(e_u + e_v) \), as is customary with field observations. For comparison with a smaller number of oceanic measurements of \( (\partial w/\partial z)^2 \) using a pitot tube on a vertical profiler (Moum 1990), we also use the alternate isotropic formula \( \varepsilon_i = 15\nu(\partial u_i/\partial x_i)^2 \) (no summation over \( i \)) to compute an estimate \( e_w \) from \( <(\partial w/\partial z)^2> \). Shown as the upper set of curves in each panel of Figure 4 are the logarithms of \( e_h \) (solid) and \( e_w \) (dashed). The heavy horizontal line in each upper panel corresponds to a value of \( \log \varepsilon_c \), where \( \varepsilon_c = 25\nu N^2 \). The condition \( \varepsilon > \varepsilon_c \) is frequently used to designate “active” turbulence, i.e., motions still able to overturn density surfaces and cause downgradient buoyancy flux (note however that literature values of the coefficient in the expression for \( \varepsilon_c \) range from 8 (Ivey and Imberger 1991) to 30 (Gibson 1980)). As \( E_0 \) increases, the time extent during which \( \varepsilon > 25\nu N^2 \) also increases. The ratio \( A_e \equiv e_h/e_w \), shown as the lower curve of each set, provides a rough estimate of the degree to which dissipation scales achieve isotropy, which would be characterized by \( A_e = 1 \). While \( e_w \) exceeds \( e_h \) at \( t = 0 \) due to the anisotropic initial conditions, the difference is quickly erased. In all but the most weakly forced run, \( e_h \approx e_w \) by \( t_N \sim 0.1 \).

(d) Scalar dissipation rates (Cox numbers)

Another traditional field measure of ocean turbulence is the Cox number (Osborn and Cox 1972), defined as the ratio of the turbulent vertical diffusivity \( K_\theta \) of a scalar to its molecular diffusivity \( D_\theta \). Using vertically-averaged vertical \( T' \) gradient variance \( (\partial T'/\partial z)^2 \) measured by fast thermistors on vertical profilers, an isotropic estimate of the Cox number is determined from

\[
C_{x_T} = \frac{\chi_T}{2D_T N_T^2} = \frac{K_T}{D_T}
\]

where \( \chi_T = 6D_T (\partial T'/\partial z)^2 \) is the dissipation rate of \( T \) variance, assuming dissipation-scale isotropy (Gregg 1987). We compute \( C_{x_T} \) (again averaging over
Figure 5: Top curves in each panel: time histories of estimates of Cox numbers, $C_{xT}$ (solid) and $C_{xS}$ (dashed), derived from volume-averaged scalar vertical gradients. Lower curve in each panel: time histories of scalar isotropy ratios $A_T$ (solid) and $A_S$ (dashed) defined in the text, Sec.3(d). Simulations (a) A3.0 (b) A4.0 (c) A4.5 (d) A5.0 are in order of increasing initial energy $E_0$, see Table 1. Evolution is shown as a function of $t_N$, time non-dimensionalized by the buoyancy period $T_N$ characteristic of the background (linear) stratification.

the computational volume of the simulation), as well as an isotropy metric

$$A_T = 2 < (\partial T'/\partial z)^2 > / ( < (\partial T'/\partial x)^2 > + < (\partial T'/\partial y)^2 > ),$$

defined such that $A_T = 1$ if gradient variance is indeed isotropic, as assumed in deriving turbulent diffusivity $K_T$ from $C_{xT}$. Although a similarly defined Cox number for salinity has been out of
reach of ocean microscale measurements (until perhaps recently: Nash and Moum 2002), we also routinely calculate \( Cx_S \) and the associated salinity isotropy metric \( A_S \). The upper pair of curves in Figure 5 (a-d) are \( Cx_T \) (solid) and \( Cx_S \) (dashed); the lower pair are the associated isotropy metrics. Simulations with initial energy \( E_0 \geq 10^{4.5} \) required computation of the S field at double the resolution of the T field (for discussion of the effects of numerical resolution on the various derived parameters, see Appendix 1). However because of computer storage limitations, these double-resolution (2xS) fields were saved less frequently than the single-resolution (1x) fields of both T and S. In Fig.5(c & d), values of \( Cx_S \) calculated with the (2xS) field are shown as circles superimposed on curves drawn using the more plentiful values computed from the (1xS) field.

As \( E_0 \) increases, there is an increase in the maximum values attained by \( Cx_T \) and \( Cx_S \) (note scale changes), accompanied by a decrease in the time required to achieve the maxima, ie to accomplish the spectral transfers that “fill out” the high wavenumbers of the scalars. \( A_T \) and \( A_S \) are generally near 1 through early stages of the simulations when scalar gradient variances and scalar dissipation rates are increasing. As the relevant variances begin to decrease, first \( A_T \), then later \( A_S \) rise above 1, indicating a decay stage in which vertical gradients exceed horizontal, even at dissipation scales.

4. Oceanic relevance of simulation results

The limited size of our computational domain raises questions about the relevance of the present three-dimensional DNS results to the ocean environment. Thus before proceeding to examine the extent of differential diffusion that is present in the simulations, we first compare descriptions of the velocity and scalar fields “observed” in the simulations with those available from ocean microprofiler observations.
(a) Microscale metrics

Some typical microscale metrics are listed in Table 2. Values of the maximum value of $C_{xT}$ achieved in runs A4.0 through A5.0 span a range of Cox numbers typically reported for the stratified interior of the ocean: see for example Gregg (1987). In all but the most weakly forced run (A3.0), at the times when maximum values of $C_{xT}$ were achieved, the ratio $\varepsilon_m / \nu N^2$ equalled or exceeded the value of 25 often quoted as critical (Stillinger et al. 1983) for the ability of turbulence to produce a downgradient buoyancy flux. The largest values of $\varepsilon / \nu N^2$ in Gregg’s (1987) compilation from the stratified ocean interior are comparable to the range (79-120) estimated for Run A4.5. Values of $\varepsilon / \nu N^2$ estimated for Run 5.0 (400-1800) are comparable to typical values (though smaller than the largest values) in more extensive data sets examined by Moum (1996a) and Smyth et al. (2001). Thus the present simulations achieve values of $C_{xT}$ and relative magnitudes of nonlinear and buoyancy terms (as quantified by values of $\varepsilon / \nu N^2$) that appear comparable to those of a large fraction of observations from the stratified interior of the ocean.

**Table 2.** Values of maximum achieved temperature Cox number ($\max C_{xT}$), and the ratio ($\varepsilon_m / \nu N^2$) at the time of $\max(C_{xT})$ for the primary model runs.

<table>
<thead>
<tr>
<th>Run #</th>
<th>$\max C_{xT}$</th>
<th>$\varepsilon_m / \nu N^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3.0</td>
<td>6</td>
<td>0.4</td>
</tr>
<tr>
<td>A4.0</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>A4.5</td>
<td>210</td>
<td>79-120*</td>
</tr>
<tr>
<td>A5.0</td>
<td>366</td>
<td>400-1800*</td>
</tr>
</tbody>
</table>

* In both these cases $C_{xT}(t_n)$ is a broad function, making the actual peak hard to define within a finite set of saved values. The range of $\varepsilon_m / \nu N^2$ values reported here come from points with $C_{xT}$ values within 10% of the absolute maximum value found in the saved data set.
(b) Velocity spectra

As mentioned above, initial anisotropy of dissipation rates appears to be rapidly erased, with $A_e$ approaching 1 by $t_N \approx 0.1$ in all but the most weakly forced run (A3.0). At $t_N = 0.070$ (panel 3 in Figs. 1 and 2), when we have the first full suite of numerical fields from run A4.5, we compare velocity dissipation spectra with appropriate “universal” one-dimensional Kolmogorov forms, as determined by Gargentt et al. (1984, henceforth referred to as GON) from airfoil probe (transverse velocity gradient) and heated-film probe (longitudinal velocity gradient) measurements taken in high Reynolds number geophysical turbulence. Since GON provided only graphical presentations, Appendix 2 provides a table of universal values for both transverse ($\phi_{ij}$) and longitudinal ($\phi_{ii}$) velocity spectral densities non-dimensionalized by Kolmogoroff scaling $\left(\epsilon_k V^5\right)^{1/4}$, as functions of one-dimensional (radian) wavenumber non-dimensionalized by Kolmogoroff wavenumber $k_v$. Solid curves in Figure 6 are various representations of both universal one-dimensional velocity gradient spectra, the transverse spectrum $k^2\phi_{ij}(k)$, such that

$$\left(\frac{\partial u}{\partial x_i}\right)^2 = \int_0^\infty k^2 \phi_{ij}(k) dk \quad i \neq j,$$

(upper curves) and the longitudinal spectrum $k^2\phi_{ii}(k)$ such that

$$\left(\frac{\partial u}{\partial x_i}\right)^2 = \int_0^\infty k^2 \phi_{ii}(k) dk \quad \text{(no summation over i)}$$

(lower curves).

Transverse spectra computed from the numerical fields $\partial u/\partial z$ and $\partial v/\partial z$ are shown as superimposed triangular symbols in Figure 6. At $t_N = 0.07$ in Run A4.5 (Fig.6(a, b)) these values lie slightly above the universal transverse spectral curve, a characteristic also observed in the early stages of Run A5.0 (not shown). At the same non-dimensional time, overall agreement with the universal transverse spectral curve is even closer for an isotropically forced run (Run I5.0H, shown in Fig.6(c,d)), suggesting that the transverse velocity spectral excess in the anisotropic runs is at least partly a remnant effect of anisotropic
initial conditions. It should be emphasized, however, that had any of the “observational estimates” of the transverse spectra in Fig.6 resulted from microprofiler airfoil probe shear measurements, the degree of agreement with the universal form would have been considered quite satisfactory (cf. Oakey 1982, his Fig. 4; Moum 1990, his Fig. 7).

Figure 6: One-dimensional longitudinal (lower curves, circles) and transverse (upper curves, triangles) velocity gradient spectral densities $k^2 \phi(\hat{k})$, normalized by Kolmogoroff scaling for velocity and wavenumber and displayed as functions of $\hat{k} = k/k_\nu$, one-dimensional (radian) wavenumber normalized by Kolmogoroff wavenumber. The curves are universal forms derived from the data of Gargett et al. (1982: see Appendix 2): symbols are spectra computed from numerical gradients "observed" at $t_N = 0.07$ in model runs A4.5 (upper panels) and I5.0H (lower panels). The lefthand panels of each set present a form frequently used in the literature, while the righthand panels show variance-preserving forms.

The “observed” longitudinal spectrum of $\partial w/\partial z$ in Run A4.5 (Fig.6(a,b)), lies more distinctly above its universal curve than does the transverse spectrum (while the longitudinal spectrum of Run I5.0H (Fig.6(c,d)) again agrees more
closely with the isotropic form). In assessing the degree of (dis)agreement between the universal longitudinal spectra and the numerical “data”, it should be recognized that reliable measurements of the longitudinal spectrum are considerably more rare than those of the transverse spectrum. Electronic noise contaminated high wavenumbers in the original oceanic measurements of the longitudinal spectrum (Grant et al. 1962, and to a lesser degree Nasmyth 1970). Other than the observations that GON used to define a universal form of the longitudinal spectrum at the high wavenumbers characteristic of gradient spectra, the only oceanic measurements are those of Moum (1990) who used a pitot tube mounted on a vertical profiler to measure a longitudinal spectrum of $\partial w/\partial z$, as well as orthogonal airfoils to provide two independent estimates of the transverse spectrum. Moum’s values of the ratio $R_b = \phi_{ij}/\phi_{ii}$ of transverse to longitudinal spectra (his Fig.8) have been redrawn here, with permission, as Fig.7(a). These measurements, originally presented in the frequency space shown, have been roughly aligned in the wavenumber space of the simulations (Fig.7(b,c)) using a value of $f \sim 30\text{Hz}$ given by Moum (1990) as the high wavenumber end (approximated by $\hat{k} \sim 0.1$) of the inertial subrange. At frequencies above $\sim 30\text{ Hz}$, the ocean ratios fall steeply as a result of noise in the pitot tube measurement. However this noise does not affect the measurements at low frequencies (wavenumbers), where observed values of $R_b$ lie significantly below the value ($4/3$) that is predicted for an inertial subrange, a value that was actually observed over nearly a decade in wavenumber in the strongest turbulence reported by GON (their Fig. 16). Since the airfoil spectra ($\phi_{ij}$) approximate the universal transverse spectral curve (as shown in Moum (1990), his Fig.7), the low ratio value observed at low wavenumbers must arise because the measured longitudinal spectrum $\phi_{ii}$ substantially exceeds its universal curve\(^1\), exactly the behaviour observed in the numerical $\phi_{ii}$ spectra of Fig. 6(a,b). The spectral ratios of Run A4.5 (Fig.7(b)) generally approach the

\(^1\) especially since the micro-profiler measurement of $\partial w/\partial z$ is, if anything, underestimated at low vertical wavenumbers by a high-pass filter effect of vehicle response to vertical velocity variations on scales greater than its length
Figure 7: The ratio $R_\phi = \phi_i / \phi_l$ of transverse to longitudinal velocity spectral densities. (a) Oceanic estimates of $R_\phi$ from Moum (1990: his Fig. 8, modified with permission), presented originally as a function of frequency. Numerical estimates of $R_\phi$ from (b) Run A4.5 and (c) Run I5.0H, as functions of normalized vertical wavenumber $\hat{k}$. Solid curves in (b) and (c) represent ratios calculated using universal forms as described in Sec. 4(a). Determination of the horizontal position of the measurements in (a) relative to that in (b) and (c) is described in the text (Sec. 4(b)).

“universal” ratio curve determined by GON (solid curve) only at the very highest wavenumbers, falling progressively below it at lower wavenumbers because of anisotropy in the initial velocity field. When the initial energy is both larger and isotropic, agreement between simulated ratios and the universal ratio curve improves at high wavenumbers and extends towards lower wavenumbers, as seen in Fig. 7(c) (Run I5.0H), where $R_\phi$ brackets the inertial subrange value of 4/3 at the lowest resolved wavenumber. Comparison between the observed and
computed ratios is somewhat tenuous, since the resolved wavenumber ranges barely overlap. However the numerical “observation” of spectral ratios from Run A4.5 falling below the universal ratio curve (because of anisotropy in the initial velocity field) is consistent with the observed oceanic ratios reported by Moum (1990).

Overall, in the early stages of development of turbulence in the present simulations, we find that both transverse and longitudinal velocity gradient spectral forms agree with universal curves to a degree similar to that reported for typical open ocean observations. Similar comparisons are not possible at later stages of the decay, because \( \varepsilon \) values fall below the noise level (\( \sim 10^{-6} \text{ cm}^2\text{s}^{-3} \)) typical of present ocean measurements.

(c) Scalar spectra

We now compare the spectra of scalars T and S (actually calculated from the fluctuation fields \( T' \) and \( S' \)), as measured during evolution of the numerical flow, with “universal” forms proposed by Batchelor (1959) and Kraichnan (1968) for scalars with generalized \( Sc > 1 \). A useful review of these forms and the degree to which laboratory and oceanic measured spectra agree (or more frequently do not agree) with these forms is provided by Bogucki et al. (1997). Here we use their expressions for the one-dimensional scalar spectrum \( E_{10}^{\vartheta} \), as derived from the Batchelor and Kraichnan three-dimensional forms and normalized by Batchelor scaling \( E_{\text{Bo}}^{\vartheta} = \chi_{\vartheta}(\nu/\varepsilon)^{1/2}k_{\text{Bo}}^{-4} \) (their eqns. (29) and (30)).

**Batchelor:**

\[
\frac{E_{10}(\hat{k}_{\vartheta})}{E_{\text{Bo}}^{\vartheta}} = q\hat{k}_{\vartheta}^{-1} \exp(-q\hat{k}_{\vartheta}^2) + \pi^{1/2}q^{3/2}(\text{erf}(q^{1/2}\hat{k}_{\vartheta}) - 1)
\]  

**Kraichnan:**

\[
\frac{E_{10}(\hat{k}_{\vartheta})}{E_{\text{Bo}}^{\vartheta}} = q\hat{k}_{\vartheta}^{-1} \exp(-(6q)^{1/2}\hat{k}_{\vartheta})
\]
In both forms $\hat{k}_\theta = k / k_{B\theta}$, where $k$ is one-dimensional radian wavenumber, $k_{B\theta}$ is the Batchelor wavenumber (1) for scalar $\theta$, $\chi_\theta$ is the dissipation rate of scalar variance and $q$ is a constant, originally defined as relating $\gamma$, the least principal rate of strain of the velocity field, to $(\varepsilon / \nu)^{1/2}$, the rate of strain of Kolmogoroff-scale eddies, via

$$\gamma = - q^{-1} \left( \frac{\varepsilon}{\nu} \right)^{1/2}$$  \hspace{1cm} (19)

(Batchelor 1959). At low wavenumbers, both forms (17) and (18) exhibit a $\hat{k}_\theta^{-1}$ viscous-convective subrange: major differences lie in the position and form of the high-wavenumber roll-off.

To allow unbiased assessment of the goodness-of-fit of scalar spectra to either Batchelor or Kraichnan forms, both $\varepsilon$ and $\chi_\theta$ must be measured concurrently and independently. This tends to be a rather rare occurrence in observational (microprofiler) data sets, due to conflicting fall-speed requirements for resolution of small-scale temperature and velocity gradients. Oceanic data sets that do provide both $\varepsilon$ and $\chi_T$ require $q$ values significantly larger than the value $q_B \approx 2$ originally estimated by Batchelor (1959) in order to fit observed $T$ spectra to the Batchelor shape. For example, using concurrent measurements of $\varepsilon$ and $\chi_T$, Oakey (1982) found a best fit to the $T$ gradient spectrum in offshore waters using $q_B = 3.7$. With similar measurements in higher Reynolds number coastal flows, Gargett (1985) found values in the range $4 < q_B < 12$, and further concluded that $q_B$ could not be considered constant, but was dependent upon the stage of flow evolution as described by a turbulent Reynolds number $Re_t \equiv \varepsilon / \nu N^2$. Recent measurements by Nash and Moum (2002) find scalar spectra best fit by the Kraichnan form, with values of $q_K$ that varied within a factor of 2 of a mean value of 7.5.

In recent years, information about scalar spectral shapes has also begun to come from direct numerical simulations, at Reynolds numbers that continue to increase with computing capabilities. Bogucki et al. (1997) found that DNS spectra of passive scalars with $3 < Sc < 7$ in statistically steady homogeneous
low Re turbulence were best fit by the Kraichnan form with $q_K = 5.26 \pm 0.25$.

Smyth (1999) reports DNS of turbulence resulting from time-dependent Kelvin-Helmholtz instability, thought to be the most common generation mechanism for turbulence in the stratified interior of the ocean. His results also suggest that the Kraichnan form best describes the shape of the scalar spectrum in all but the final stages of decay. However the effective value $q_e$ of the "constant" $q_K$ that is required to best fit the Kraichnan spectral form is found to vary dramatically over the evolution of the flow, as a result of non-constant strain. Both early and late stages of evolution were characterized by large values of $q_e$, intermediate stages by values between 6 and 8.

Since the most common oceanic measurements are of small-scale scalar vertical gradients measured from vertical profilers, we focus here on spectra of vertical scalar gradients as functions of vertical wavenumber. Logarithmic plots of $\hat{k}_B^2 \hat{E}_\theta = k_z^2 E_{\theta}(k_z)/k_{B0}^2 E_{\theta_0}$, the Batchelor-scaled one-dimensional spectrum of $\partial \theta/\partial z$ as a function of non-dimensional vertical wavenumber $\hat{k}_B = k_z/k_{B0}$, are commonly used in both the observational and computational literature on scalar spectra: this presentation, while not variance-preserving, will be used here to facilitate comparisons.

We have found that the evolution of the scalar spectra generated in our simulations forms a coherent picture when organized by time relative to the time $t_\theta$ at which the scalar's dissipation rate peaks. Note that since $t_S > t_T$, ie the time required to achieve maximum dissipation rate is longer for $\theta = S$ than for $\theta = T$ (see Fig.5), a gradient spectrum of $\theta = S$ that is obtained at the same run time ($t_N$) as a $\theta = T$ spectrum occurs relatively earlier in scalar gradient evolution as a function of $t_N/t_\theta$.

Arranged in this scaled time $t_N/t_\theta$, scalar gradient spectra from all our turbulent simulations form a distinctive, consistent continuum of shapes, which is illustrated in Fig.8 using results from run A4.5. The spectral plots shown in Fig.8
Figure 8: Scalar gradient spectra calculated as functions of one-dimensional vertical wavenumber using $\theta = T'$ (left panels) and $\theta = S'$ (right panels) from Run A4.5 (see Fig.5c) at run times $t_N = 0.07$, (a) and (d); $t_N = 0.15$, (b) and (e); $t_N = 0.31$, (c) and (f). Spectra are normalized using Batchelor scaling (see text) and presented as functions of $\hat{k}_\theta = \left( k_z / k_{B\theta} \right)$, where $k_{B\theta}$ is the Batchelor wavenumber appropriate to the scalar in question. Relative to "universal" shapes suggested by Kraichnan (dashed curves) and Batchelor (solid curves), the time evolution of the "observed" scalar spectral shapes shows a consistent pattern when viewed in a framework (central axis) of time relative to $t_\theta$, the time of maximum scalar dissipation rate: for discussion, see Sec.4(c).
are arranged vertically in order of increasing (top to bottom) values of \( t_N/t_0 \), as annotated on the central bar: panels to the left of this bar are scalar spectra derived from \( T \), those to the right are derived from \( S \). In addition to the spectrum of the computational data (points), each panel contains “universal” Batchelor (solid) and Kraichnan (dashed) curves calculated respectively with \( q_B = 3.9 \) and \( q_K = 5.26 \), the best-fit values determined by Bogucki et al. (1997). The "earliest" scalar spectrum is the \((S)\)spectrum at \( t_N/t_0 = 0.4 \) (Fig. 8(d)). At this time \( t_N = 0.07 \) in Figs. 3, 5), when both scalar variance and dissipation rate are still increasing, the scalar spectrum approximates the Kraichnan form at intermediate wavenumbers. The significance of a slight high-wavenumber excess observed relative to both universal form is not clear (oceanic measurements offer no guidance in this range, because they are usually obscured by instrumental noise before such high wavenumbers are reached). However over a substantial range of the lowest resolved wavenumbers, there is clearly significant excess variance relative to both universal shapes. The excess of low wavenumber variance decreases with time, but is still evident in the \((T)\)spectrum at \( t_N/t_0 = 0.5 \) (Fig. 8(a)). We suggest that the presence of excess low-wavenumber scalar variance is evidence of a non-equilibrium stage in which there has been insufficient time for scalar variance formed at the largest overturning scale to cascade out of these scales to fill in the full range of available spatial scales. The spatial (and spectral) “distance” to the smallest available (dissipation) scale is larger for \( S \) than for \( T \); thus at the same point in run time, the \((S)\)spectrum at \( t_N/t_0 = 0.4 \) is further from equilibrium, hence exhibits a larger amount of excess low wavenumber gradient variance, than the \((T)\)spectrum at \( t_N/t_0 = 0.5 \). Slight excess low-wavenumber variance is still observed at \( t_N/t_0 = 0.8 \) (Fig. 8(e)), but has disappeared by \( t_N/t_0 = 1.2 \) (Fig. 8(b)), when the scalar \((T)\)spectrum is rather well described by a Batchelor spectrum, exhibiting a peak at slightly higher wavenumber and a somewhat steeper high wavenumber roll-off than the Kraichnan spectrum. For \( t_N/t_0 > 1 \), the period of decay that follows the peaking of dissipation rate, scalar spectra begin to diverge from the Batchelor form, first (Fig. 8(f)) by developing a progressive deficit of low wavenumber gradient
variance while maintaining a high wavenumber roll-off approximating that of Batchelor, finally (Fig.8(c)) by export of remaining variance to wavenumbers significantly exceeding the wavenumbers at which both universal spectral shapes roll off. These same general features of spectral evolution are also found in both runs A5.0 and I5.0H, although these results are less accurate than those for run A4.5 (for run A5.0 because spatial resolution of the S field is marginal, see Appendix 1; for run I5.0H because a paucity of saved sets makes determination of the times of maximum scalar dissipation rates inaccurate). Thus no truly universal form of the scalar variance spectrum is observed in these time-dependent simulations. Depending upon the stage of flow evolution, either Kraichnan and Batchelor forms can be adequate descriptions of the measured spectra. Far into the decay, neither shape is appropriate.

Lack of universality of scalar spectral shape during the process of growth and decay of a turbulent instability in a stratified fluid implies that the oceanographic practice of deriving a value for turbulent kinetic energy dissipation rate $\varepsilon$ by fitting a constant-q Batchelor spectral shape will be inaccurate both early in flow evolution, when the observed gradient spectrum exhibits excess low wavenumber energy relative to the Batchelor form, and late in flow evolution when it exhibits a deficit of such energy (and when moreover, the velocity shear spectrum is no longer isotropic). Errors at these later times are relatively unimportant because $\varepsilon$ is small (indeed such situations may be invisible to instruments with finite noise levels). However in the initial period when $\varepsilon$ is large but scalar gradient variance may still be small, hence strongly affected by noise at high wavenumbers, adjustment of the Batchelor curve towards the "enhanced" lower wavenumbers would result in overestimating $\varepsilon$.

To emphasize the degree of relevance of these results to oceanic measurements, Fig.9(a) reproduces (with permission) the collection of oceanic temperature gradient spectra used by Oakey (1982) to determine a value of $q_B = 3.7$ for the best-fit Batchelor curve seen in Fig.9(b). The panels below show similar plots of scalar gradient spectra calculated from the numerical "data" (using saved sets of both $T'$ and $S'$ from runs A4.0, A4.5, and A5.0), non-
Figure 9: Comparison between (upper panels) an ensemble of scalar (T) gradient spectra measured in the ocean (Oakey 1982, his Fig.9: reproduced with permission) and (lower panels) an ensemble of scalar gradient spectra determined from our three-dimensional numerical simulations (both T and S spectra from Runs A4.5 (t_n = 0.07, 0.15, 0.31); Run A5.0 (t_n = 0.04, 0.09, 0.14); Run I5.0H (t_n = 0.04, 0.14). The right-hand panel of each set displays an averaged version of the data on the left: the curves are Batchelor spectra with q_B = 3.7. For details, see Sec.4(c).

dimensionalized as described above. [There are offsets in both axes relative to the plots in Fig.9(a,b), as a result of Oakey's use of cyclic wavenumber (his \( \hat{k} \)) in place of radian wavenumber \( k_\theta \) in calculating the Batchelor-scaled wavenumber: the resulting x-offset has been removed by the relative horizontal alignment of the two sets of plots. A Batchelor curve with q_B = 3.7 is plotted in both right-hand panels, with scaling appropriate to the definition of non-dimensional wavenumber. Comparison of the two “data” sets shows that the fit of the ocean
observations to the Batchelor curve is closely matched by the simulations. This comparison suggests that the scatter in oceanic measurements of T gradient spectra results from sampling time-dependent turbulent flows, in parameter ranges comparable to those of our simulations, at random times during their growth and decay.

5. Quantifying differential diffusion

(a) Three-dimensional simulations

In order to examine differential vertical scalar fluxes, we first define vertical advective (correlation) fluxes $F_T \equiv \langle -w'T' \rangle$ of $(-T')$ and $F_S \equiv \langle w'S' \rangle$ of $S'$: in all presentations, fluxes are normalized by the background thermal conductive flux. Recalling that (non-dimensional) density $\rho = S - T$, positive values of both $F_T$ and $F_S$ are associated with downgradient density flux, normally associated with turbulent stirring, while negative values indicate countergradient density flux, leading to restratification.

Figure 10 shows the time evolution of volume-averaged $F_T$ for the primary runs (solid lines). $F_T$ starts at zero, rises to a positive (downgradient) maximum, then crosses zero to become weakly negative (countergradient). For the most weakly forced run (A3.0), the zero-crossing occurs at a non-dimensional time which is roughly one-quarter of a buoyancy period, moving to slightly longer times as initial energy increases. $F_S$ closely follows $F_T$: indeed, plotted on the same scale, $F_S$ is indistinguishable from $F_T$. However the two fluxes are not identical, as shown by the time series of flux difference $F_T - F_S$ (dashed lines). Initially the flux difference is negative, ie $F_T < F_S$. However the difference eventually reverses sign, at a time which is nearly equal to the time of sign reversal of the flux itself for the weakest initial forcing, but becomes somewhat shorter than the flux reversal time for all the more energetic initial conditions.

During the early “turbulent” phase when $F_T > 0$ and $F_S > 0$, the flux difference $(F_T - F_S) < 0$, hence S is being fluxed downgradient more efficiently.
Figure 10: Correlation flux $F_T \equiv \langle -wT' \rangle$ (solid) and correlation flux difference $(F_T - F_S)$ (dashed) between $F_T$ and $F_S \equiv \langle wS' \rangle$: all fluxes are non-dimensionalized by the background thermal conductive flux. Simulations (a) A3.0 (b) A4.0 (c) A4.5 (d) A5.0 are in order of increasing initial energy $E_0$, see Table 1. Evolution is shown as a function of $t_N$, time non-dimensionalized by the buoyancy period $T_N$ characteristic of the background (linear) stratification.

than $T$. Later, during a re-stratification phase when $F_T < 0$, $F_S < 0$, the flux difference $(F_T - F_S) > 0$, hence $S$ is being fluxed countergradient more efficiently than $T$. Note that these two stages cover most of the flow evolution, with the exception of a short period of time when $(F_T-F_S)$ has reversed sign while both fluxes are still weakly downgradient. The observed behaviours mean that during both dominant stages, an amount of $T$ variance is being “left behind” in the vertical relative to $S$ variance. Just such behaviour would result if the transfer of $T$
variance from the large scales to dissipation scales is more efficient than that of 
S, because T dissipation scales are “closer” in a spatial sense than the physically 
smaller S dissipation scales. Averaged over the flow evolution, less effective 
vertical transfer of S than T will occur if the excess of countergradient transport of 
S during the restratification phase exceeds its excess downgradient transport 
during the initial mixing phase.

This is evidently the case. A measure of the net difference between the 
two fluxes, from initiation of mixing until kinetic energy and scalar variances have 
effectively decayed away, is the cumulative flux ratio \( f \), defined as \( f = \Phi_T/\Phi_S \), 
where the cumulative fluxes

\[
\Phi_T = \int_0^{t_{\text{max}}} F_T dt \quad \text{and} \quad \Phi_S = \int_0^{t_{\text{max}}} F_S dt
\]

are time integrals of the normalized correlation fluxes over the entire simulation. 
For ease of later comparison with extensions to previous two-dimensional 
turbulence simulations (see following section), the three-dimensional results will 
be presented in the plane of initial \((t=0)\) Reynolds number \( \text{Re}_0 = \frac{U_0 L_0}{\sigma} \) and 
Froude number \( \text{Fr}_0 = \frac{U_0}{N L_0} \), defined by Merryfield et al. (1998) using eddy 
velocity scale \( U_0 \equiv (2E_0)^{1/2} \) and eddy length scale \( L_0 \equiv \pi \left( \frac{E_0}{Z_0} \right)^{1/2} \), with non-
dimensional energy \( E_0 = \frac{1}{2} \int |\psi_0|^2 dx dz \) and \( Z_0 = \frac{1}{2} \int |\nabla \psi_0|^2 dx dz \). In addition to 
the values of \( f \) (Fig. 11(a)) calculated for the set of primary simulations, 
Fig.11(b) shows the largest value of normalized flux difference attained during 
the run, \( \max(F_T - F_S) \), while Fig.11(c) presents the overall mixing efficiency \( \Gamma \equiv \frac{\Delta PE}{\Delta KE} \), defined as the ratio of net increase in potential energy \( \Delta PE = \sigma (\Phi_T + \Phi_S R_p^{-1}) \) divided by net decrease in kinetic energy \( \Delta KE \) \( \approx E_0 \) since kinetic 
energy decays almost completely by the end of each run). The region of the 
\( \{\text{Re}_0, \text{Fr}_0\} \) plane where \( f > 1 \) while simultaneously \( \max(F_T - F_S) > 1 \), ie the 
maximum flux difference exceeds the background thermal conductive flux (hence 
also the background conductive flux of salt), may be considered both 
differentially diffusive and turbulent. All of the simulation results are found in such 
a region, with the largest \( f \) values associated with the weakest turbulence. The
cumulative flux ratio $f$ is observed to approach 1 as the initial energy (hence $Fr_0$, $Re_0$, and turbulence intensity) increases. Unlike $f$, which decreases monotonically with increase in initial energy, both $\max(F_T - F_S)$ (Fig.11(b)) and the mixing efficiency (Fig.11(c)) exhibit maximum values at mid-range, a point to which we will return in the following section.

While the maximum mixing efficiency achieved in the set of three-dimensional runs (0.089 in Run A4.0) is somewhat lower than the value $\Gamma \sim 0.2$ widely considered to be “normal” for ocean turbulence, it is not outside the bounds of reported observational estimates, which range from 0.05 (Fleury and Lueck 1994) to 0.7 (Gargett and Moum 1995). Moreover the DNS results of Smyth (1999) suggest that mixing efficiency of turbulence generated by Kelvin-Helmholtz (K-H) instability varies with turbulent “age”, being high in a pre-turbulent roll-up stage when the strain field remains aligned with large scalar gradients, and less efficient after transition when the relative orientation between strain and scalar gradient fields is far from optimal. Since our simulations lack the initial, highly organized stage of a Kelvin-Helmholtz instability, average mixing efficiency would be expected to be biased towards the (lower) values typical of
fully turbulent post-transition stages of Kelvin-Helmholtz instabilities. In addition, Smyth et al. (2001) show that the mixing efficiency generated by K-H instability also decreases with increase in Prandtl number (our Sc) over the range $1 \leq \text{Pr} \leq 7$. Our net buoyancy flux includes that due to a scalar of molecular diffusivity ten times smaller than the smallest used by Smyth et al. (2001), a second reason that the maximum value of $\Gamma$ observed in our simulations may be expected to be smaller than theirs. Finally, because mixing efficiency involves the net flux of density, in the oceanic case where density depends on both $T$ and $S$, $\Gamma$ must be expected to be a function of the density ratio $R_\rho$, a dependence not addressed by our $R_\rho = 1$ simulations.

Another view of differential diffusion is given in Figure 12(a), which shows co-spectra between vertical velocity and individual scalars (scalar flux spectra) at three stages in the evolution of Run A4.5. In Fig.12(a), wavenumber ranges with downgradient fluxes are indicated by solid lines, and those with countergradient fluxes by dashed lines. Near $t=0$ (cf $t_N = 0.014$ in Fig.12), fluxes of both scalars are downgradient at all wavenumbers and $\langle wS' \rangle$ slightly exceeds $\langle -wT' \rangle$. As the run continues, turbulent kinetic energy continues to support downgradient fluxes at the largest scales while countergradient fluxes are observed to develop, first at the smallest scales, then extending to progressively larger scales. At the time $t_N = 0.07$ that corresponds to the fields of panel 3 in Figs.1 & 2, flux reversal has already begun at high wavenumbers. At this stage, the flux spectra resemble the (ensemble-averaged) temperature flux spectrum presented by Moum (1996b, his Fig. 3), in which clearly resolved downgradient flux at low wavenumbers appears to reverse sign at a wavenumber lower than that of the peak of the temperature spectrum (although resolution of countergradient flux at higher wavenumbers may be affected by the noise level of the $w$ measurement). In the simulations, reversed (countergradient) fluxes proceed to envelop almost all of the resolved range by a time ($t_N = 0.66$, panel 6 of Figs.1 & 2) which is substantially less than one buoyancy period.
Figure 12: (a) Volume-averaged flux cross-spectra $\langle -wT' \rangle$ (thick lines) and $\langle wS' \rangle$ (thin lines) as functions of vertical wavenumber $k_z$ (in units of $d^{-1}$) at three times during Run A4.5: upper (early) to lower (late) sets of curves correspond to $t_N = 0.014, 0.07,$ and $0.66$. Solid lines indicate downgradient (mixing) flux, while dashed lines represent countergradient (restratifying) flux. Countergradient fluxes develop first at the highest wavenumbers (small scales) and progress to lower wavenumbers (larger scales) with increasing time. (b) Volume-averaged spectrum of scalar coherency $\langle -T'S' \rangle / \sigma_T \sigma_S$ at the same times.

Figure 12(b) shows (minus the) coherency spectrum of $T'$ and $S'$, essentially the correlation coefficient of the $T'$ and $S'$ spectra, at the same times used in Fig. 12(a). This provides an opportunity to compare our results with studies of differential diffusion in unstratified turbulence (e.g. Yeung and Pope 1993) that have been mainly concerned with evolving correlations between
passive scalars of differing Sc. At $t_N = 0.014$, the two scalars are highly correlated at all scales, whereas at $t_N = 0.07$, coherency remains > 0.8, indicating that diffusion still has done little to decorrelate $T'$ and $S'$. We note that both these times lie within the initial regime where $F_T < F_S$ (see Fig. 10(c)). By $t_N = 0.66$, however, coherency has declined substantially, especially at high wavenumber, indicating that $T'$ and $S'$ have become largely decorrelated due to the more rapid molecular diffusion of $T'$. This time is well within the $F_T > F_S$ regime, reflecting the preferential restratification of the relatively compact $S'$ concentrations versus the more diffuse $T'$ concentrations. This tendency for decreasing coherency qualitatively resembles that seen by Yeung and Pope (1993).

Figure 13: Time evolution of velocity gradient spectra, relative to "universal" shapes, during Run A4.5: (a) $t_N = 0.070$ and (b) $t_N = 0.31$. Curves and symbols are as defined in Fig.6.

The combination of (i) loss of low wavenumber energy, through both the turbulent cascade and work against buoyancy forces, and (ii) gain of high wavenumber energy associated with small-scale (negatively buoyant)
restratifying motions causes the velocity dissipation spectra to peak at progressively higher wavenumbers as \( t_n \) increases. This evolution, characteristic of all the "turbulent" simulations, is illustrated in Fig.13 for Run A4.5. At the earliest time for which a save set is available (\( t_n = 0.07 \), Fig.13(a)), when scalar variances and dissipation rates are still increasing, both transverse and longitudinal scalar gradient spectra peak near their respective universal spectral peaks. By a time (\( t_n = 0.31 \), Fig.13(b)) that is well into the period of decay of scalar variances, the peaks of both gradient spectra lie closer to the Kolmogoroff wavenumber itself (\( \log \hat{k} = 0 \)).

(b) Two-dimensional simulations: estimating the effect of different Schmidt number

Since the parameter ranges and spectral characteristics of our simulations are comparable to those of a large fraction of the observations of ocean turbulence, our simulations should provide relevant estimates of the magnitude of oceanic differential diffusion if they can be scaled from the value of the molecular diffusivity ratio (\( \tau = 0.1 \)) used in our three-dimensional simulations to that (\( \tau = 0.01 \)) characteristic of temperature \( T \) and salinity \( S \) in seawater. Due to computational limitations, we could not run three-dimensional simulations with \( \tau = 0.01 \). Instead, as one accessible means of exploring the magnitude of differences that might be expected as a result of differences in \( \tau \), we repeated the two-dimensional simulations of Merryfield et al. (1998), replacing their value of \( \tau = 0.01 \) with the larger value of \( \tau = 0.1 \). Fig.14 presents these two-dimensional results, using the same metrics used in Fig.11 for the three-dimensional results. The upper panels repeat the results of Merryfield et al. (1998) for comparison with the new \( \tau = 0.1 \) simulations (lower panels). Both sets of simulations show similar qualitative behaviour for all three diagnostics. The \( \tau = 0.1 \) simulations show decreases of order 20-40% in both \( f \) and \( \max(F_T - F_S) \), causing a slight shrinkage in the differentially-diffusive region of the \( \{Re_0, Fr_0\} \)
Figure 14: Comparison of differential diffusion diagnostics from two-dimensional simulations run with (a-c): the molecular diffusivity ratio $\tau = 0.01$ typical of temperature (T) and salinity (S) in seawater, (d-f): the value $\tau = 0.1$ accessible to the present three-dimensional calculations. For details and discussion, see Sec. 5(b).

plane relative to the $\tau = 0.01$ simulations. The observed increase in mixing efficiency $\Gamma$ with $\tau$ is expected: since T and S contribute equally to density stratification, the amount of irreversible density transfer should increase if the “slow” scalar diffusion becomes faster, i.e. $\tau$ increases.

Quantitatively, the three-dimensional results show slightly less differential diffusion than two-dimensional runs with the same $\tau$ and roughly similar values of $\{Fr_0, Re_0\}$ (compare for example a value of $f = 1.28$ at $\{50,35\}$ in Fig.14(d) with a value between 1.11 and 1.22 in Fig. 11(a)). We believe that these differences result from the different energy cascade properties of two and three
dimensional turbulence. In two-dimensional turbulence, “turbulent” kinetic energy cascades to low wavenumbers, attaining a state in which the simulated velocity field consists of a single overturning cell, roughly the size of the box, that decays away very slowly under (nearly) molecular viscosity. An increase in initial energy $E_0$ increases the energy in this cell and hence its decay time (as illustrated in Fig.15(a)), prolongs the period of time during which (both) fluxes are downgradient, and thus increases mixing efficiency. In three dimensions, turbulent kinetic energy is instead transferred to high wavenumbers (smaller scales) where it can be quickly removed by viscous forces. The transfer rate $\varepsilon$ scales as $u^3/\ell \sim E_0^{3/2}/\ell$, where $u \sim E_0^{1/2}$ and $\ell$ are typical large-eddy velocity and length scales (Tennekes and Lumley 1972). Because the computational box size is fixed, we increase $E_0$ without the concomittant increase in spatial scale $\ell$ needed to maintain a constant energy transfer (dissipation) rate. The resulting
increase in \( \varepsilon \) is such that the decay time scale \( \tau_d \sim E_0 / \varepsilon \sim E_0^{-1/2} \) for the three-dimensional large eddies decreases as \( E_0 \) increases (as seen in Fig.15(b)).

Taking the induced change in potential energy to scale as a typical vertical displacement \( u \tau_d \sim E_0^{1/2}E_0^{-1/2} \), the mixing efficiency scales as \( \Gamma \sim \Delta PE/E_0 \sim u \tau_d / E_0 \sim E_0^{-1} \), and hence decreases with increasing initial energy. In the most weakly forced three-dimensional run (A3.0: Fig.4(a)), \( \varepsilon \) exhibits non-monotonic time behaviour, suggestive of the presence of wave-like motions that may be expected to have low mixing efficiency. As \( E_0 \) is increased, \( \varepsilon \) becomes monotonic with time (cf Fig. 4(c)) and the small-scale fields increasingly exhibit various properties considered characteristic of three-dimensional turbulent flows, as described in Sec. 4(a)-(c). It seems reasonable that the moderately energetic turbulence of runs A4.0 and 4.5 first results in more efficient net downgradient flux of buoyancy relative to that associated with the less energetic, more wave-like flows of run A3.0. However as \( E_0 \) continues to increase, the associated decrease in the turbulent decay time \( \tau_d \) eventually leads to decrease in \( \Gamma \), as observed in the most vigorous run (A5.0, Fig.5(c)). A very similar dependence of mixing efficiency on \( \{Re, Fr\} \) in experimental data is shown in Fig. 2 of Ivey and Imberger (1991).

While there are certainly significant differences between two and three dimensional turbulent flow fields, the associated cumulative flux ratios \( f \) exhibit comparable magnitudes and parameter sensitivity in both cases. The present comparison between two-dimensional simulations with different \( \tau \) values thus suggests that, in the three-dimensional case as well, estimates of flux ratios made via simulations with \( \tau = 0.1 \) may be within ~ 40\% of the effect that would arise with \( \tau = 0.01 \). Measures of differential diffusion from our three-dimensional runs with \( \tau = 0.1 \) should provide reliable lower bounds to the magnitude of such effects for salinity.
6. Discussion and conclusions

The three-dimensional direct numerical simulations reported here provide both qualitative and quantitative evidence that two active scalars with molecular diffusivities that differ by a factor of ten exhibit differential diffusion as we have defined it, i.e., larger turbulent diffusivity of the property having the larger molecular diffusivity.

Qualitative variation of the cumulative flux excess of the scalar (T) of higher molecular diffusivity is similar to that observed in the two-dimensional simulations of Merryfield et al. (1998): the flux ratio is largest at small values of $Re_0$ and $Fr_0$ and decreases monotonically towards 1 as $Re_0$ and $Fr_0$ increase. The magnitude of differential diffusion is found to be slightly smaller in a three-dimensional case compared to a two-dimensional case with the same initial $Re_0$ and $Fr_0$ and the same value of molecular diffusivity ratio. In those flows that exhibit properties characteristic of turbulence, the cumulative flux of T exceeds that of S by (6-22)%.

For similar values of $Re$ and $Fr$, a larger difference should be expected for the case of temperature and salinity in seawater, characterized by a molecular diffusivity ratio that is a factor of ten smaller than the value used in our simulations. Moreover, our present results indicate that overall mixing efficiency $\Gamma$ varies substantially over the parameter range of the simulations, being greatest where differential diffusion is largest. This suggests, as do Jackson and Rehmann (2002), that differential diffusion may have implications for another frequently used assumption in ocean microscale interpretation – that of constant mixing efficiency in stratified turbulence. Furthermore, in the presence of differential diffusion, mixing efficiency may also be expected to be a strong function of the stability ratio $R_p$, a dependence which requires further investigation.

Although our simulations are restricted in both spatial scale and forcing energy levels by available computer resources, we have demonstrated that both qualitative and quantitative features of our results closely resemble those
characteristic of all but the strongest turbulence observed in the stratified ocean interior. In particular:

(a) the range of numerical “turbulent” Reynolds number $Re_t \equiv \varepsilon / \nu N^2$ (hence $Fr_t = 1$, Gargett 1988) and temperature Cox numbers is typical of a large fraction of values determined from microprofiler measurements of velocity and temperature microscale gradients in the stratified interior;

(b) velocity dissipation spectra quickly achieve near-universal forms, to a degree consistent with observations in the ocean interior;

(c) while scalar spectra are actually “best” fit by either Batchelor or Kraichnan forms (or neither) depending upon evolution time, a composite of “Batchelor-scaled” scalar spectra formed from our results closely resembles a similar compilation of observed oceanic temperature spectra.

Having thus determined the magnitude of differential diffusion in our DNS results and demonstrated the likely applicability of our results to turbulence in the stratified ocean interior, we turn finally to the question of the fundamental roots of the differential diffusion phenomenon. We believe that both in the present simulations and elsewhere, the observed sense of differential diffusion, ie larger turbulent vertical diffusivity of the scalar with the larger molecular diffusivity, requires the net dominance of countergradient scalar fluxes at high wavenumbers (small scales). Our argument is illustrated schematically in Fig.16, which shows the significant stages in the time evolution of flux spectra $F_\theta(k)$ (similar to the simulation cross-spectra of Fig. 10) of two generalized scalars $T$ and $S$ that both contribute to stable stratification, but are characterized by $D_T >> D_S$, hence $k_{BT} << k_{BS}$. In the initial stages (Fig.16(a)) where both fluxes are downgradient (positive, for consistency with Fig.10), the two flux spectra differ only at high wavenumbers, where $F_S$ (dashed line) extends to higher wavenumbers than $F_T$ (solid line) because $k_{BT} << k_{BS}$. Although the magnitude of differences in the fluxes (given by the integral of the relevant flux spectrum over $0 < k < \infty$) are expected to be small, since the flux spectra fall off steeply above the Komogoroff wavenumber characteristic of velocity gradient smoothing, it is expected (and indeed we observe) that at this stage any difference will be in
Figure 16: Schematics of various phases in the evolution of flux spectra $F_T$ and $F_S$ associated with two generalized scalars $T$ and $S$ that both contribute to stable density stratification, but are characterized by $D_T >> D_S$, hence $k_{BT} << k_{BS}$. (a) Initial phase with downgradient fluxes of both scalars at all wavenumbers, $F_T > 0$, $F_S > 0$, and $F_T - F_S < 0$. (b) Intermediate phase in which $F_T > 0$, $F_S > 0$ still, but $F_T - F_S > 0$, as a result of countergradient fluxes that have started to develop at high wavenumbers. (c) Final phase in which countergradient fluxes spreading to progressively smaller wavenumbers have led to $F_T < 0$, $F_S < 0$, while $F_T - F_S > 0$. Countergradient fluxes at high wavenumbers, $F_T - F_S > 0$, must dominate if net differential diffusion is to have the sense observed in the present simulations and in various laboratory stratified flows, namely preferential turbulent vertical diffusion of the property with the larger molecular diffusion.

the sense of $F_S > F_T$ ($F_T - F_S < 0$), ie the downgradient turbulent flux is larger for the scalar of smaller molecular diffusivity. In a subsequent stage, illustrated in Fig.16(b), both flux spectra have become countergradient at high wavenumbers (cf Fig.12) due to the negative buoyancy of displaced fluid elements from which $T$ has preferentially diffused. While at this stage both fluxes remain positive, albeit smaller than in Fig.16(a), the flux difference reverses sign, ie $F_T - F_S > 0$, because the countergradient (-ve) portion of the spectrum is larger for $S$ than for $T$. Eventually, as shown in Fig.16(c), countergradient fluxes extend to all but (perhaps) the smallest wavenumbers; both fluxes have now become negative, with $|F_S| > |F_T|$. Thus during this stage, the flux difference remains positive, ie $F_T - F_S > 0$ as in the previous intermediate stage. Net differential diffusion can occur
only if the excess of upwards (downgradient) vertical transfer of $S$ during the original phase (Fig. 16(a)) is exceeded by its excess countergradient transfer during the re-stratification phases (Fig. 16(b) and (c)), during which the development and spread of countergradient fluxes from higher to lower wavenumbers are crucial features.

The above interpretation of high-wavenumber countergradient fluxes as the root of differential diffusion has a number of consequences. It suggests, first, that differential diffusion will occur in any situation where average countergradient fluxes are characteristic of the smallest scales, regardless of whether the flow is time-dependent (our simulations and the K-H simulations of Smyth et al. 2001) or statistically stationary (the laboratory measurements of Komori and Nagata 1996). Further, it suggests that there will be both qualitative and quantitative differences between an active scalar and a truly passive scalar, defining the latter as a scalar present at such low concentration that it does not produce significant density stratification in an otherwise unstratified fluid (a caveat necessary to ensure that the “passive” tracer is not highly correlated with an active tracer). In the case of such truly passive scalars, fluid particles tagged with a scalar anomaly have no dynamical “memory” of their initial positions, hence no tendency to return to it. With thus no physical process available for generation of countergradient flux at any scale, velocity variance at $k > k_v$ is only that left at the high wavenumber end of a turbulent cascade. Consequently, flux spectra of truly passive scalars $T_p$ and $S_p$, with $D_{T_p} > D_{S_p}$ may only resemble those of Fig. 16(a), hence $F_{T_p} < F_{S_p}$, and $F_{T_p}^p - F_{S_p}^p < 0$, ie the flux difference should have the sense opposite to that defined as differential diffusion. This prediction seems consistent with Kerstein’s (1990) calculations for truly passive scalars $H_2$ (= $T_p$, $Sc_{T_p} = 0.18$) and propane (= $S_p$, $Sc_{S_p} = 1.2$) mixing in a round jet. Defining a parameter $z \equiv \xi_{S_p} - \xi_{T_p}$, where $\xi$ is the ratio of species mole fraction to its original (nozzle) mole fraction, Kerstein finds $z > 0$, ie fluid “particles” lose more $T_p$ (higher molecular diffusivity, lower $Sc$ passive scalar) than $S_p$ (lower molecular diffusivity, higher $Sc$ passive scalar). Interpreted in the context of downgradient fluxes, the observation that $T_p$ is "left behind" implies that $S_p$ is
preferentially transferred in the vertical, ie that $F_{r_p} - F_{s_p} < 0$, when the two scalars are dynamically passive.

When instead scalars are dynamically active, buoyancy forces acting on as yet un-mixed fluid parcels provide a mechanism for the generation of countergradient fluxes, hence the potential for flux differences with the sense defined here as differential diffusion. Moreover, motions associated with the tendency (however small) of buoyant fluid particles to return towards their original location provide an additional source of velocity variance (as possibly evidenced by the appearance of shear variance in excess of universal spectral levels at wavenumbers near $k_v$ during the turbulent decay process, cf Fig.15(c)), prolonging the phase during which $F_T - F_S > 0$ (Fig.16(b,c)), and enhancing the amount of differential diffusion.

Studies of the physical basis of countergradient scalar fluxes are presently few. In two dimensions (vertical plane), Holloway (1988) predicts the phenomenon using closure theory constrained by entropy increase, while Gerz and Schumann (1996) present mechanistic models for both unstratified and stratified shear flows. Hanazaki and Hunt (1996) applied rapid distortion theory to the case of unsteady turbulence in uniformly stratified unsheared flow, and obtained analytical results that predict high-wavenumber countergradient buoyancy flux for scalars with $Pr > 1$. Their results suggest that the high-wavenumber countergradient phenomenon in stably-stratified systems can be explained largely in terms of linear processes (since rapid distortion is a linear theory). This conclusion contrasts with Yeung’s (1996) explanation of countergradient fluxes for passive scalars with $Sc < 1$, which was in terms of fundamentally nonlinear triadic interactions. Given our emphasis on the importance of high wavenumber restratifying fluxes to the sense as well as magnitude of differential vertical fluxes of $Sc > 1$ scalars, further work in this area is clearly essential.

Acknowledgments: Comments from two anonymous reviewers led to significant improvements in the manuscript. This research was supported by Fisheries and Oceans
Appendix 1: Effects of numerical resolution on derived parameters

A crucial question in DNS is whether all relevant flow structures have been adequately resolved. As a minimum requirement, the grid spacing should be roughly half the smallest wavelength expected in the fields to be computed. We estimate the smallest expected scales as the Batchelor scales associated with the scalars, and compare with grid resolutions of $\Delta x = 0.202$ cm ($17.3\text{cm}/128)x^{3/2}$ for $u$ and $T$ and $\Delta x_S = 0.101$ cm ($17.3\text{cm}/256)x^{3/2}$ for $S$, as used in the most strongly forced runs, A4.5 and A5.0 (the factors of $3/2$ arise from the $2/3$ spectral truncation used in dealiasing). Table A1.1 presents values of relevant parameters computed at two different times during these runs, first the time ($t = 0$) of maximum turbulent kinetic energy dissipation rate $\varepsilon$, and second the time at which the maximum value of the salt Cox number was achieved. Values of $\varepsilon$ are used to calculate the Kolmogoroff wavenumber $k_\nu$, hence the lengths scale $\ell_u = 2\pi/k_\nu$ that characterizes the small-scale limit of the velocity dissipation spectrum (Gargett et al. 1984). The smallest scales likely to be significant in the scalar dissipation spectra are then of order $\ell_T = \ell_u/\sqrt{Sc_T}$ and $\ell_S = \ell_u/\sqrt{Sc_S}$ for $T$ and $S$ respectively (cf eqn.(1)). At $t = 0$, $\ell_u > 2\Delta x$, and $\ell_T \sim 2\Delta x$, but $\ell_S < 2\Delta x_S$, suggesting that the salt field would not be well-resolved. However at $t = 0$ both $C_{xT} = 0$ and $C_{xS} = 0$, since there has been no time for scalar variance to cascade to dissipation scales; thus resolution of the scalar dissipation fields is not an issue at this time. The more appropriate calculation is that at time $t(\max C_{xS})$, when the Cox number of the least diffusive scalar achieves its maximum value. By this time in the simulations, decay of $\varepsilon$

| Table A1.1: Spatial scales required to resolve velocity ($u$), temperature ($T$) and salt ($S$) variables through their respective dissipation ranges at two times during |
Runs A4.5 and A5.0. The times chosen are \( t=0 \), at which the dissipation rate is maximum and \( t(\text{max } C_x S) \) at which maximum \( C_x S \) is attained.

<table>
<thead>
<tr>
<th>Run</th>
<th>( \epsilon_w ) (cm(^2)s(^{-3}))</th>
<th>( k_u ) (rad/cm)</th>
<th>( \ell_u ) (cm)</th>
<th>( \ell_T ) (cm)</th>
<th>( \ell_S ) (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T=0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4.5</td>
<td>3.20 \times 10^{-3}</td>
<td>7.50</td>
<td>0.838</td>
<td>0.317</td>
<td>0.100</td>
</tr>
<tr>
<td>A5.0</td>
<td>5.60 \times 10^{-3}</td>
<td>8.66</td>
<td>0.726</td>
<td>0.274</td>
<td>0.087</td>
</tr>
<tr>
<td>( t(\text{max } C_x S) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4.5</td>
<td>2.04 \times 10^{-5}</td>
<td>2.13</td>
<td>2.94</td>
<td>1.11</td>
<td>0.353</td>
</tr>
<tr>
<td>A5.0</td>
<td>1.48 \times 10^{-4}</td>
<td>3.49</td>
<td>1.80</td>
<td>0.68</td>
<td>0.215</td>
</tr>
</tbody>
</table>

has led to larger values of \( \ell_u \), which eases the resolution requirements, and in both runs the scalars have Batchelor scales significantly larger than \( (T) \), or comparable to \( (S) \), twice their respective grid resolutions.

However the simplistic criterion used above may underestimate the effects of under-resolution, since the effective grid resolution of a numerical simulation depends upon the specific numerical methods used. Thus, we have addressed resolution effects in a second way, by examining the evolution of computed parameters as the base computational grid is refined from \( n^3 = (96)^3 \) through \( (128)^3 \) to \( (160)^3 \) (with \( S \) computed at twice the base resolution) in three simulations \( I5.0L – I5.0H \) run at the largest initial forcing energy. [The runs in Table A1.2 differ from Run A5.0 only in the isotropic initial conditions, resulting in lower initial Re. While for present purposes it would be desirable to make these comparisons for Run A5.0, we were unable to carry out a second \( (160)^3 \) simulation.] The first three parameters listed in Table A1.2, those used to

**Table A1.2:** Effects of numerical resolution on calculated variables. The first three parameters, measures of differential diffusion, are identical for all three
runs. The maximum CoxT is well-resolved by the lowest resolution (n=96) run; maximum CoxS is resolved with n=128.

<table>
<thead>
<tr>
<th>Run</th>
<th>I5.0L(n=96)</th>
<th>I5.0M(n=128)</th>
<th>I5.0H(n=160)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f$</td>
<td>1.08</td>
<td>1.08</td>
<td>1.08</td>
</tr>
<tr>
<td>$\text{max}(F_T - F_S)$</td>
<td>5.0</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>$\Gamma$</td>
<td>0.0137</td>
<td>0.0137</td>
<td>0.0137</td>
</tr>
<tr>
<td>$\text{max CoxT}$</td>
<td>210</td>
<td>207</td>
<td>210</td>
</tr>
<tr>
<td>$\text{max CoxS (2xS)}$</td>
<td>768</td>
<td>1464</td>
<td>1263*</td>
</tr>
</tbody>
</table>

* from comparison with values computed (frequently) from single-resolution S-fields, this value is likely an under-estimate by ~10-15% as a result of the infrequency with which double-resolution S-fields were saved, particularly in this run which was the largest undertaken.

quantify differential diffusion, are seen to be insensitive to resolution in the range examined. This lack of sensitivity arises because fluxes are carried predominantly by the larger scales of the flow which are well-resolved even with $n = 96$. The maximum value attained by CoxT is also seen to be insensitive to resolution in this range: this guarantees that $\varepsilon$ is also well-resolved, since velocity shear variance resides at scales significantly larger than those characterizing T gradient variance. The only parameter showing resolution sensitivity is the salt Cox number. The maximum value of CoxS doubles with an increase from $n = 96$ to $n = 128$, and would be expected to further increase, if anything, with subsequent increase to $n = 160$. Instead the max CoxS (2xS) value listed in Table A1.2 for $n = 160$ shows a slight decrease. This anomaly arises because the (2xS) save set occurred at a time that was slightly before the actual maximum, at a time coincident with the beginning of a plateau (not shown) between two subsequent CoxS values determined from the more frequent (1xS) values. The true maximum achieved during this run could only have occurred within this plateau and is unlikely to have exceeded the observed (2xS) level by more than 10-15%. We conclude that at the largest forcing amplitude used in any of our
simulations, a base resolution of \( n = 128 \) is adequate to resolve all of the scales relevant to scalar fluxes and both velocity and scalar variances for runs with isotropic initial forcing. However Run A5.0 has initial Reynolds number \( \text{Re}_0 \) that is 25% larger than Run I5.0 (Table 1), and the observed difference between near-peak \( \text{C}_{xS} \) as calculated from (1xS) and (2xS) fields (cf Fig. 5(d)) is nearly a factor of 2, suggesting that resolution of \((128)^3\) may be marginal for this run. For this reason we have concentrated analysis on Run A4.5, in which doubled resolution produces only minor change in near-peak \( \text{C}_{xS} \) (see Fig.5(d)).

Appendix 2: Tabular values of universal longitudinal and transverse velocity and shear spectra determined by Gargett et al. (1984)

Table A2 lists (logarithmic) universal values of \( \hat{\phi}_{ii} \) (longitudinal) and \( \hat{\phi}_{ij} \) (transverse) velocity spectral densities non-dimensionalized by the Kolmogoroff scaling \( (\nu^{5})^{1/4} \), as functions of \( \hat{k} \), the one-dimensional radian wavenumber non-dimensionalized by the Kolmogoroff wavenumber. For convenience, logarithmic values of the associated shear spectral densities are listed in the final two columns in the form \( \hat{k}^3 \hat{\phi}_{ii} \) and \( \hat{k}^3 \hat{\phi}_{ij} \) that preserves variance when plotted against \( \log \hat{k} \) (cf Fig.15). These forms were determined from observations in high \( \text{Re} \) turbulence generated by breaking internal waves in Knight Inlet, British Columbia: for details of the observational setting, instrumentation and processing, see Gargett et al. (1984). Universal forms were derived using only their Class 1 records, those taken at the highest \( \text{Re} \) and exhibiting a full decade of inertial subrange as revealed by the requirement that \( R_\phi = \phi_{ij} / \phi_{ii} = 4/3 \), a more stringent condition than the more usual observation of a \( k^{-5/3} \) subrange in an individual velocity component spectrum. Tabulated values are derived from a linear fit to spectral estimates within the inertial subrange \((-2.5 \leq \log \hat{k} \leq -1.2\)\), matched to a 6th order polynomial fit in the rolloff region.

**Table A2.** Non-dimensional (logarithmic) values of universal velocity spectral densities \( \hat{\phi}_{ii} \) and \( \hat{\phi}_{ij} \) as functions of one-dimensional (radian) wavenumber \( \hat{k} \),
determined from the observations of Gargett et al. (1984). Logarithmic values of
the associated shear spectral densities are listed in the final two columns in the
form $k^3 \hat{\phi}_{ii}$ and $k^3 \hat{\phi}_{ij}$ that preserves variance when plotted against $\log \hat{k}$.

```
\begin{tabular}{cccc}
  $\log \hat{k}$ & $\log \hat{\phi}_{ii}$ & $\log \hat{\phi}_{ij}$ & $\log k^3 \hat{\phi}_{ii}$ & $\log k^3 \hat{\phi}_{ij}$ \\
  \hline
  -2.5 & 3.88 & 4.01 & 0.0002 & 0.0003 \\
  -2.4 & 3.72 & 3.84 & 0.0003 & 0.0004 \\
  -2.3 & 3.55 & 3.67 & 0.0004 & 0.0006 \\
  -2.2 & 3.38 & 3.51 & 0.0006 & 0.0008 \\
  -2.1 & 3.22 & 3.34 & 0.0008 & 0.0011 \\
  -2.0 & 3.05 & 3.17 & 0.0011 & 0.0015 \\
  -1.9 & 2.88 & 3.01 & 0.0015 & 0.0020 \\
  -1.8 & 2.72 & 2.84 & 0.0021 & 0.0028 \\
  -1.7 & 2.55 & 2.67 & 0.0028 & 0.0038 \\
  -1.6 & 2.38 & 2.51 & 0.0038 & 0.0051 \\
  -1.5 & 2.22 & 2.34 & 0.0052 & 0.0069 \\
  -1.4 & 2.05 & 2.17 & 0.0071 & 0.0094 \\
  -1.3 & 1.88 & 2.01 & 0.0096 & 0.0128 \\
  -1.2 & 1.72 & 1.84 & 0.0131 & 0.0174 \\
  -1.1 & 1.55 & 1.69 & 0.0176 & 0.0247 \\
  -1.0 & 1.35 & 1.54 & 0.0225 & 0.0346 \\
  -0.9 & 1.13 & 1.36 & 0.0270 & 0.0452 \\
  -0.9 & 1.01 & 1.25 & 0.0289 & 0.0504 \\
  -0.8 & 0.88 & 1.14 & 0.0304 & 0.0552 \\
  -0.8 & 0.75 & 1.02 & 0.0314 & 0.0594 \\
  -0.7 & 0.60 & 0.90 & 0.0319 & 0.0628 \\
  -0.7 & 0.45 & 0.76 & 0.0318 & 0.0653 \\
  -0.6 & 0.29 & 0.62 & 0.0311 & 0.0666 \\
  -0.6 & 0.12 & 0.47 & 0.0298 & 0.0666 \\
\end{tabular}
```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.5</td>
<td>-0.05</td>
<td>0.31</td>
<td>0.0279</td>
<td>0.0652</td>
</tr>
<tr>
<td>-0.5</td>
<td>-0.24</td>
<td>0.15</td>
<td>0.0255</td>
<td>0.0625</td>
</tr>
<tr>
<td>-0.4</td>
<td>-0.44</td>
<td>-0.03</td>
<td>0.0227</td>
<td>0.0584</td>
</tr>
<tr>
<td>-0.3</td>
<td>-0.89</td>
<td>-0.43</td>
<td>0.0163</td>
<td>0.0468</td>
</tr>
<tr>
<td>-0.2</td>
<td>-1.40</td>
<td>-0.89</td>
<td>0.0101</td>
<td>0.0327</td>
</tr>
<tr>
<td>-0.1</td>
<td>-1.99</td>
<td>-1.42</td>
<td>0.0051</td>
<td>0.0190</td>
</tr>
<tr>
<td>0.0</td>
<td>-2.70</td>
<td>-2.06</td>
<td>0.0020</td>
<td>0.0087</td>
</tr>
<tr>
<td>0.1</td>
<td>-3.55</td>
<td>-2.84</td>
<td>0.0006</td>
<td>0.0029</td>
</tr>
<tr>
<td>0.2</td>
<td>-4.59</td>
<td>-3.80</td>
<td>0.0001</td>
<td>0.0006</td>
</tr>
<tr>
<td>0.3</td>
<td>-5.87</td>
<td>-4.99</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
References


Osborn, T. R., 1980: Estimates of the local rate of vertical diffusion from


