# Shear coincidence: implications of the statistics of ocean turbulence microphysics for global diapycnal mixing

Ali Mashayek[1], Brendan Barry[2], Matthew Alford[3], Laura Cimoli[4], and Colm-cille Caulfield[5]

[1]Imperial College London
[2]National Oceanography Centre, Southampton
[3]University of California San Diego, Scripps Institution of OCeanography
[4]University of California San Diego, Scripps Institution of Oceanography
[5]University of Cambridge

November 30, 2022

## Abstract

It is well established that small scale turbulent mixing induced by breaking of waves in the interior of the ocean plays a significant role in sustaining the deep ocean circulation and in regulation of tracer budgets such as those of heat, carbon and nutrients. There has been significant progress in fluid mechanical understanding of the physics of breaking internal waves. Connection of the microphysics of such turbulence to the global ocean, however, is significantly underdeveloped. We offer a theoretical-statistical approach, heavily informed by observations, to make such a link and then by employing climatological information show that in the global ocean, regions of optimal turbulent mixing coincide with regions that have a desirable balance of stratification and velocity shear. This optimality depends critically on the statistics of turbulent patches. Energetic mixing zones exhibit efficient bulk mixing that induces significant vertical density fluxes, while quiet zones (with small background turbulence levels), while efficient in mixing, exhibit minimal vertical fluxes. The transition between the less energetic to more energetic zones, quantifications of which we argue depends critically on turbulence statistics, implies upwelling and downwelling of deep waters may be stronger than previously estimated, which in turn has direct implications for the ocean overturning circulation as well as for the global budgets of heat, carbon, nutrients, and other tracers. Impact Statement Waves similar to those observed at the beach exist throughout the ocean interior and are induced by tides, winds, currents, eddies, and other processes. Similar to beach waves, internal waves can also roll up and break. Widespread internal-wave breaking helps drive the ocean circulation by upwelling the densest waters that form in polar regions and sink to the ocean abyss. They also play an important role in transport and storage of heat, carbon, and nutrients. In this work we show how well-understood concepts in wave physics can be used in conjunction with statistics of observed ocean turbulence to improve significantly our understanding of the impact of small-scale mixing on the global ocean, and thereby on the climate system.

*Flow* (2021), **1** E1
doi:10.1017/flo.2021.1

RESEARCH ARTICLE

# Shear coincidence: implications of the statistics of ocean turbulence microphysics for global diapycnal mixing

A. Mashayek[1*], B.B. Cael[2], M.H. Alford[3], L. Cimoli[3], and C.P. Caulfield[4]

[1]Imperial College London, London, UK
[2]National Oceanography Centre, Southampton, UK
[3]Scripps Institution of Oceanography, La Jolla, USA
[4]University of Cambridge, Cambridge, UK
*Corresponding author E-mail: mashayek@ic.ac.uk

**Received:** XX XX 2021; **Revised:** XX XX 202X; **Accepted:** XX XX 202X

**Keywords:** Ocean mixing, stratified turbulence, turbulence statistics

## Abstract

It is well established that small scale turbulent mixing induced by breaking of waves in the interior of the ocean plays a significant role in sustaining the deep ocean circulation and in regulation of tracer budgets such as those of heat, carbon and nutrients. There has been significant progress in fluid mechanical understanding of the physics of breaking internal waves. Connection of the microphysics of such turbulence to the global ocean, however, is significantly underdeveloped. We offer a theoretical-statistical approach, heavily informed by observations, to make such a link and then by employing climatological information show that in the global ocean, regions of optimal turbulent mixing coincide with regions that have a desirable balance of stratification and velocity shear. This optimality depends critically on the statistics of turbulent patches. Energetic mixing zones exhibit efficient bulk mixing that induces significant vertical density fluxes, while quiet zones (with small background turbulence levels), while efficient in mixing, exhibit minimal vertical fluxes. The transition between the less energetic to more energetic zones, quantifications of which we argue depends critically on turbulence statistics, implies upwelling and downwelling of deep waters may be stronger than previously estimated, which in turn has direct implications for the ocean overturning circulation as well as for the global budgets of heat, carbon, nutrients, and other tracers.

**Impact Statement** Waves similar to those observed at the beach exist throughout the ocean interior and are induced by tides, winds, currents, eddies, and other processes. Similar to beach waves, internal waves can also roll up and break. Widespread internal-wave breaking helps drive the ocean circulation by upwelling the densest waters that form in polar regions and sink to the ocean abyss. They also play an important role in transport and storage of heat, carbon, and nutrients. In this work we show how well-understood concepts in wave physics can be used in conjunction with statistics of observed ocean turbulence to improve significantly our understanding of the impact of small-scale mixing on the global ocean, and thereby on the climate system.

## 1. Introduction

Turbulence induced by breaking waves in the ocean interior plays a key role in regulating the global ocean circulation and budgets of climatically important tracers such as heat, carbon and nutrients [Talley et al., 2016]. Internal waves are primarily excited at the bottom of the ocean through interaction of tides, currents and eddies with bottom topography and at the surface through the wind stress acting on the sea-surface [Garrett and Kunze, 2007, Nikurashin and Ferrari, 2013, Legg, 2021, Alford, 2020]. Figure 1a,b shows sample regions in the Southern Ocean where all such sources are in play, exciting an energetic internal wave field. When the vertical shear induced by the waves is sufficiently strong (panel c), it leads to wave breaking and turbulence (panels d and e). Ocean turbulence is highly intermittent in time and space and thus difficult to sample sufficiently to allow for accurate quantification of properties of interest such as turbulent mixing. Thus, sampling poses a monumental challenge to connecting our understanding of physics of turbulence, which occur on scales $O(10^{-3} - 10^{-1})$m, to the larger scale regional and global implications, which are relevant on scales $O(10^6 - 10^7)$m. In this sense, this problem is analogous to cloud physics: both involve micro-physics with leading order global impacts, both occur on scales much smaller than typical climate model resolutions and thus need to be parameterized, and both have been notoriously difficult to parameterize and contribute significantly to inaccuracies in model solutions. In this paper we provide a novel methodology, based on recent progress in our understanding of the microphysics of mixing [Mashayek et al., 2021b, hereafter MCA21] and statistics of wave-induced turbulence [Cael and Mashayek, 2021, hereafter CM21], to connect the physics of small scale turbulent mixing to large scale deep ocean circulation and upper ocean fluxes of tracers.

## 2. Physics of wave-induced turbulence

Our understanding of the physics of wave-induced density stratified turbulent mixing has progressed significantly over the past few decades [Peltier and Caulfield, 2003, Ivey et al., 2008, Caulfield, 2021]. A key question is how the total power available to turbulence, $\mathcal{P}$ is partitioned into mixing, $\mathcal{M}$, defined as a net vertical irreversible density flux, and dissipation into heat (due to the seawater viscosity), the rate of which is referred to as $\varepsilon$.

Figure 1d shows the life cycle of a canonical wave's breakdown into turbulence. The top row shows the mixing of dense water with the lighter overlying water. The bottom row shows the density flux which is positive once averaged over the turbulence life cycle, implying an upward flux of dense waters raising the center of mass. The turbulent mixing may be approximated by

$$\mathcal{M} \approx \kappa N^2 \approx \Gamma \varepsilon, \tag{1}$$

where $N^2 = -\frac{g}{\rho_0}\partial_z \rho$ represents the density stratification, $g$ is the gravitational constant and $\rho_0$ is a reference density [Osborn, 1980][1]. The turbulent diffusivity, $\kappa$, is an input parameter in climate models to account for the subgrid-scale unresolved turbulent mixing. The turbulent flux coefficient, $\Gamma = \mathcal{M}/\varepsilon$, is often taken to be a constant value of 0.2 for historical reasons and arguably for the lack of a universally accepted parameterization for it. However, it is well known to be highly variable [Mashayek and Peltier, 2013b, Gregg et al., 2018], with appreciable consequences for ocean circulation [de Lavergne et al., 2015, Mashayek et al., 2017c, Cimoli et al., 2019]. While $\mathcal{M}$ is often the quantity of interest from a physical perspective, it is difficult to observe directly and is often inferred from $\varepsilon$ (via Eq. 1) which itself is inferred from micro-scale shear (i.e. spatial gradients of velocity) measured by microstructure probes on profiling instruments, gliders, or moorings. Thus, accurate quantification of $\Gamma$ is key to inferring ocean mixing from direct observations.

The density flux in the bottom row of Figure 1d is composed of large scale features associated with a primary overturning eddy which is the source of energy to $\mathcal{M}$ and $\varepsilon$, and smaller features comprise the turbulence cascade of eddies. Three basic turbulence scales have proven useful for parameterization of $\Gamma$ based on observable quantities. The Kolmogorov scale, $L_K = (\nu^3/\varepsilon)^{1/4}$, represents the scale below which viscous dissipation takes kinetic energy out of the system, the Ozmidov scale, $L_O = (\varepsilon/N^3)^{1/2}$, is the maximum (vertical) scale that is not strongly affected by stratification, and the Thorpe scale, $L_T$, is a geometrical scale characteristic of vertical displacement of notional fluid parcels within an overturning turbulent patch [Dillon, 1982, Smyth and Moum, 2000, Mashayek et al., 2017a] – $\nu$ is the kinematic

[1] Note that while the caveats associated with this approximation are important[Mashayek et al., 2013, Mashayek and Peltier, 2013a], they don't change the primary conclusions of this work.

viscosity of seawater. Panel (e) in Figure 1 shows time evolution of these scales for the turbulence event shown in panel (d).

Together, panels d and e show the initial accumulation of available potential energy (APE) from background shear into the primary billow (large $L_T$) followed by the growth of smaller eddies within the main billow upon feeding on its APE source (increase in $L_O$). As turbulence grows, the scale at which energy is taken out of the system, $L_K$, decreases. The phase where $L_O \sim L_T$ is the most efficient transfer of energy from the overturn to the smaller scales, is the richest, highest dynamic range turbulence (marked by the largest gap between $L_O$ and $L_K$), and is thus the most efficient phase of the flow where (instantaneous) mixing efficiency is defined as $\mathcal{M}/(\mathcal{M}+\varepsilon) = \Gamma/(1+\Gamma)$ [Peltier and Caulfield, 2003].

The ratio of $L_O$ and $L_K$, often expressed as the buoyancy Reynolds number $Re_b = (L_O/L_K)^{4/3}$, has been widely used to quantify $\Gamma$ [Bouffard and Boegman, 2013, Mashayek et al., 2017c, Gregg et al., 2018, Monismith et al., 2018] and also to establish the global scale impacts of variations in $\Gamma$ [de Lavergne et al., 2015, Mashayek et al., 2017c, Cimoli et al., 2019]. However, some of these efforts as well as others have also highlighted the inherent deficiency of $Re_b$ since it only includes instantaneous information on the turbulence scales $L_O$ and $L_K$ while not 'knowing' anything about either the energy containing scale $L_T$ [Gargett and Moum, 1995, Mashayek and Peltier, 2011c, Mashayekhi, 2013, Mater and Venayagamoorthy, 2014, Mashayek et al., 2021b] or any time dependence of the flow, although evidence is accumulating that 'history matters' in stratified mixing [Caulfield, 2021]. An alternate method for inferring mixing, employed when direct inference of $\varepsilon$ is not available, is to assume that $R_{OT} = L_O/L_T$ is a constant (taken to be between 0.6 and 1) and $\Gamma = 0.2$ [Dillon, 1982, Thorpe, 2005]. This assumption allows for an indirect inference of $\varepsilon$, and thereby of $\Gamma$. This method is inherently deficient (although in practice might be the only available option) since it only 'knows' about the energy containing scale and not the the dissipation scale, and so knows nothing about the dynamic range of the turbulence.

Recently, MCA21 [Mashayek et al., 2021b] proposed a simple parameterization for $\Gamma$ based on basic physical grounds that agreed well with observational data:

$$\Gamma = A\frac{R_{OT}^{-1}}{1+R_{OT}^{\frac{1}{3}}}, \tag{2}$$

where $A$ is a constant ranging between $1/2$ and $2/3$. Crucially, this parameterization captures the fundamental time-dependent nature of shear-driven mixing events, allowing for variation in $R_{OT}$ and hence $\Gamma$ at different stages in the life cycle of a mixing event. Figure 2a, reproduced from MCA21, shows the agreement with a combination of data from ~50,000 turbulent patches gathered from six different field campaigns that sampled turbulence in different geographical locations around the global, at different depths and turbulence regimes, and from turbulence induced by different processes (see *Materials* for a brief description of data). Eq. 2 reduces to $A/R_{OT}$ in the limit of young turbulence ($R_{OT} \ll 1$)) and to $A/R_{OT}^{-4/3}$ in the limit of (older) decaying turbulence ($R_{OT} \gg 1$)). While the decaying turbulence limit of Eq. 2 was suggested by others in the past, as discussed in MCA21, the young turbulence limit and the transition between the two limits were formulated by MCA21.

A nice feature of Eq. 2 is that it allows for the two ranges to merge smoothly at $R_{OT} \sim 1$. This limit corresponds to efficient mixing when there exists an optimal balance between the stratification and energy available to turbulence. In the young turbulence range, stratification is relatively high and sufficient energy has not yet transferred to turbulent eddies to work against the stratification and mix. So, while $\Gamma$ can be very large, it does not imply much mixing: it is large since $\varepsilon$ is very small, not because $\mathcal{M}$ is large. In fact, in the limit of laminar flow, $\Gamma \to \infty$. On the other hand, in the decaying phase of turbulence, stratification is somewhat eroded and so $\mathcal{M}$ is weaker as there is less to mix, while $\varepsilon$ is still finite as even an unstratified flow can have significant $\varepsilon$. Thus, in this limit $\Gamma \to 0$. It is the $R_{OT} \sim 1$ limit in which the right balance of stratification and power exists and optimal mixing occurs.

Since this intermediate phase appears to lead to 'optimal' mixing, neither too hot nor too cold but 'just right', MCA21 referred to this as 'Goldilocks mixing'.

In this paper we will show that all three (time-dependent) phases of turbulence life cycle, importantly including the young turbulence limit, are key to connecting the small-scale physics of mixing to the large-scale ocean dynamics. While the Goldilocks mixing phase is when most of the effective turbulent flux occurs, the young and weakly turbulence patches are important since most of the ocean interior is relatively 'quiet' with intermittent bursts of turbulence. We will argue that it is essential to account for the less/non turbulent regions whereas historically parameterizations have primarily focused on energetic turbulence. This will necessitate careful analysis of the statistics of turbulent patches.
On dimensional grounds, recently Mashayek et al. [2021a, hereafter MCBC21] argued that Eq. (2) may also be expressed in terms of the buoyancy Reynolds number $Re_b = (L_O/L_K)^{4/3} = \varepsilon/\nu N^2$, and the gradient Richardson number $Ri = N^2/S^2$, where $S$ is the vertical shear of horizontal velocity:

$$\Gamma = A\frac{R_{OT}^{-1}}{1+R_{OT}^{\frac{1}{3}}} = A\frac{{Re_b^*}^{1/2}Ri^*}{1+Re_b^*}, \tag{3}$$

where

$$Re_b^* = \frac{Re_b}{Re_{bm}}, \qquad Ri^* = \frac{Ri}{Ri_m}, \tag{4}$$

where $Ri_m$ and $Re_{bm}$ are the values of $Ri$ and $Re_b$ where the Goldilocks mixing phase occurs, i.e. where $R_{OT} \sim 1$. MCBC21 argued that by normalizing $Re_b$ and $Ri$ with their values evaluated at $R_{OT} \sim 1$, similar to Eq. (2), Eq. (3) represents two limits of the young/growing turbulence ($Re_b^* \ll 1$) and fossilizing/decaying turbulence ($Re_b^* \gg 1$). The two limits are smoothly connected at $Re_b^* \approx Ri_* \approx R_{OT} \sim 1$, i.e. at some critical/marginal values of $Re_b$ and $Ri$ at which $R_{OT} \sim 1$, within the Goldilocks mixing phase.

Using oceanic datasets and building on earlier works, MCBC21 argued that both $Re_{bm}$ and $Ri_m$ can vary (potentially co-dependently) from experiment to experiment: this is not surprising as intensity of stratified turbulence is known to be dependent on both Reynolds and Richardson numbers. Of relevance to this work, however, is the mere notion that there exists a critical Richardson number that corresponds to $R_{OT} \sim 1$. So, the fact that $R_{OT} \sim 1$ corresponds to optimal mixing for individual patches (as discussed) and also to regions of optimal mixing on global scale (to be discussed) implies that in such regions there exists an optimal balance between stratification and shear to keep $Ri$ at, or at least near, a marginal value or critical value $Ri_m$.

## 3. Statistics of turbulence

### *3.1. Statistics of patches*

Figure 2b shows the histogram of the data used in panel a separated into four quartiles in terms of $\varepsilon$. The distribution shows that most turbulent patches lie within a factor of 3 of $R_{OT} = 1$ and the larger the $\varepsilon$, the closer to 1 the peak of $R_{OT}$ lies. This suggests the majority of turbulent patches are in this phase of optimal or Goldilocks mixing. Such a clustering of data has been widely reported in the past [Dillon, 1982, Thorpe, 2005, Mater et al., 2015, Mashayek et al., 2017a] and suggests that out of the three phases of energetic ocean turbulence events, namely 'young' growth, Goldilocks mixing, and 'old' decay, the intermediate phase spans the larger fraction of the turbulence life cycle. This is shown quantitatively in panels c,d of Figure 2 where life-cycle-averaged properties are plotted from turbulence life cycles simulated by direct numerical simulations such as that in Figure 1. Panels c,d, together, show that for sufficiently energetic turbulence (i.e. sufficiently high Reynolds number[2] and sufficiently low

[2]Reynolds number in the simulations is defined as $Re = Ud/\nu$, where U is a characteristic flow velocity scale and $d$ is the characteristic vertical scale of shear.

$Ri$), a larger proportion of the turbulence life cycle corresponds to the Goldilocks phase rather than to the growth and decay phases.

The parameterization (2) describes individual turbulent patches. A turbulent region within the ocean, such as that shown in Figure 1b-c based on an observationally-tuned numerical simulation, hosts many patches in different stages of their evolution. MCA21 argued that an appropriate value for a bulk $\Gamma$ for a region, such as those in Figures 1b-c, may be constructed for such a region through

$$\Gamma_B = \frac{\mathcal{M}_{tot}}{\varepsilon_{tot}} \approx \frac{\sum_{i=1}^{n} \Gamma_i \times \varepsilon_i}{\sum_{i=1}^{n} \varepsilon_i} \tag{5}$$

where $n$ represents the number of patches in a region of interest. For example, typical resolution of climate models is $O(100\text{ km})$ in the horizontal and $O(100\text{ m})$ in the vertical direction. Models that employ a time variable mixing parameterization[3], employ $\Gamma = 0.2$ to construct a diffusivity for each grid cell. The more appropriate value, however, would be the $\varepsilon$-weighted $\Gamma$ average, i.e. $\Gamma_B$, which would rely on the statistical distribution of patches within the grid cell and their individual $\Gamma$. By applying Eq. 5 to four datasets, MCA21 showed that $\Gamma_B$ can be close to the Goldilocks mixing (i.e. $A/2 \approx 1/3$ according to Eq. 2) when the region of study has the right balance of power and stratification (where $R_{OT}$, $Re_b^*$, $Ri^* \sim 1$), thereby comprising mostly Goldilocks mixing patches. However, regions that host a higher percentage of young or weak turbulent patches can have $\Gamma_B \sim O(1)$ since young patches have large $\Gamma$ as shown in Figure 3 in the limit of small $R_{OT}$.[4]

### 3.2. Statistics of continuous profiles

Bulk mixing therefore depends on statistics of turbulent patches. However, weakly/non-turbulent waters reside in between intermittent turbulent patches. Recently Cael and Mashayek [2021, hereafter CM21], showed that $\varepsilon$ data from over 750 full depth microstructure profiles from 14 field experiments (covering a wide range of depths, geographical locations, and turbulence-inducing processes; see CM21 for details), are well-described by a log-skew-normal distribution (Figure 4a), which has the form

$$f(\varepsilon; \xi, \omega, \alpha) = \frac{2}{\omega\varepsilon}\phi\Big(\frac{\log\varepsilon - \xi}{\omega}\Big)\varphi\Big(\alpha\frac{\log\varepsilon - \xi}{\omega}\Big), \tag{6}$$

where $f$ is a probability density function and $\phi$ and $\varphi$ respectively are the probability and cumulative density functions (PDF and CDF) of a standard Gaussian random variable. As discussed in CM21, the relationship between the log-skewness ($\theta$) of the distribution and the additional shape parameter $\alpha$ is one-to-one. The parameters $(\xi, \omega, \alpha)$ are related to the log-mean, log-standard deviation, and log-skewness $(\mu, \sigma, \theta)$ of a log-skew-normal random variable according to the equations:

$$\mu = \xi + \sqrt{\frac{2}{\pi}}\omega\delta, \tag{7}$$

where

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}, \qquad \sigma = \omega\sqrt{1 - \frac{2}{\pi}\delta^2}, \qquad \theta = (4-\pi)\Big(\sqrt{\frac{2}{\pi}}\delta\Big)^3 / (2(1 - \frac{2}{\pi}\delta^2)^{3/2}). \tag{8}$$

[3]Many models don't employ such a time-varying parameterization and even those that do take the tidal power as a given and merely adjust the diffusivity based on the evolving stratification.

[4]Young patches have large $\Gamma$ because of small $\varepsilon$ NOT due to large $\mathcal{M}$. Thus, while they don't contribute much to the net turbulent flux $\sum \Gamma_i \varepsilon_i$, they bias $\Gamma_B$ high.

The log-skew-normal distribution arises from the analogue of the Central Limit Theorem for log-normal variables: that is, the sum of log-normal variables converges to a log-skew-normal distribution [Wu et al., 2009]. For turbulence, the log-skew-normal distribution is thought to arise because the total, measured $\varepsilon$ results from a combination of multiple and/or not statistically steady turbulence-generating processes [Caldwell and Mourn, 1995], which individually and/or instantaneously have log-normally distributed dissipation rates.

### *3.3. Chronic undersampling: inevitability of a statistical approach to $\Gamma_B$*

The difficulty of sampling intermittent turbulence has long been recognized [Gregg, 1992], but the problem is worse than thought owing to the even more heavy-tailed nature of of a log-skew-normal $\varepsilon$ distribution than a lognormal one. This presents a severe challenge for constraining the average $\varepsilon$ value for an ocean volume from a limited sample set since many samples are required to resolve the disproportionately consequential tail. To illustrate, simulations like those in Figure 1d,e demonstrate that turbulent quantities such as $\varepsilon$ and $\Gamma$ vary substantially over the lifetime of an overturning event. This implies that a single measurement, or even a handful of measurements, of such an event can be a poor estimate of its time-integrated properties since they most likely miss the most energetic phase of turbulence for each event. This underscores the need for large measurement sets and a statistical approach to fill in the gaps and accurately estimate bulk turbulent properties. To highlight the sampling issue, we consider two examples: one based on sampling of an individual turbulence event; and one based on sampling of the collective global dataset discussed in Figure 4.

DNS simulations, such as that shown in Figure 1 or those shown in Figure 2, clearly demonstrate that turbulent quantities such as $\mathcal{M}$, $\varepsilon$, and $\Gamma$ vary substantially over the lifetime of an overturning event. This implies that a single measurement, or even a handful of measurements, of such an event are highly likely to be a poor estimate of its time-integrated properties. Figure 3a shows the uncertainty and bias associated with random samples of one[5] simulation's history (Re = 6000, Ri = 0.16) for different sample sizes. Uncertainty is quantified by the normalized standard deviation of the sample mean for collections of samples taken at random times along the temporal history of the overturning event. Bias is quantified by the ratio of the median sample mean from these collections of randomly timed samples to the true time-averaged mean of each property. Because all the properties shown are positively skewed quantities, the sample means tend to miss the peak values and thus underestimate the time-averaged value; because these skews are so large, means of different collections of random samples vary by e.g. more than a factor of two for fewer than ~16 measurements. Turbulent measurements are snapshots of events; it is not possible to make tens of measurements of the same turbulent event. These appreciable biases and uncertainties for characterizing individual events thus underscore the importance of the statistics of turbulent properties.

The heavy-tailed nature of the $\varepsilon$ distribution in Figure 4a presents a similar challenge for constraining the average $\varepsilon$ value for an ocean volume from a limited sample set. Figure 3b shows the same as Figure 3a but for random draws from a log-skew-normal distribution whose parameters are fit to match the $\varepsilon$ data from the experiments described above ($\xi = -24.8$, $\omega = 3.91$, $\alpha = 5.89$) using the same procedure as in CM21. We discard values above a chosen $\varepsilon_{max}$ threshold of $10^{-5}$ m$^2$/s$^3$; as with any parameterization for the PDF of $\varepsilon$ supported on $(0, \infty)$, the log-skew-normal allows for a non-zero probability density for unmeasurably and/or unphysically large $\varepsilon$ values, which are either too large to be measured by sampling probes and/or yield impossibly large $L_O$ values. $O(1000)$ samples are needed to make the underestimation bias less than ~10%, and $O(100)$ samples are needed to make the standard deviation of sample means less than the true mean (i.e. $< 100\%$ relative uncertainty). (Larger/smaller values of $\varepsilon_{max}$ increase/reduce these sample size numbers.) This further underscores the need for large measurement sets and a statistical approach to estimate bulk turbulent properties accurately.

[5]Other simulations yielded similar results.

## 4. A recipe for $\Gamma_B$

### *Ingredients: $\Gamma(R_{OT})$ for individual patches + statistics of $L_O$, $L_T$, and $R_{OT}$*

Here we aim to exploit the log-skew-normality of $\varepsilon$ and the statistics of $R_{OT}$ to construct a parameterization for $\Gamma_B$ based on the total power and the mean stratification for the region (or the grid cell) for which $\Gamma_B$ is sought. To this end, first we note that $L_O \propto \varepsilon^{1/2}$ by definition; the former lies within the $R_{OT}$-based $\Gamma$ parameterization in Eq. 2 and the latter is captured by Eq. 6.[6] Because $L_O$ divides $\varepsilon$ by another random variable and then takes a square root of their quotient, this distribution is also preserved for $L_O$ (Figure S1; Table S1; also see below). Here we pair the parameterizations for the statistical distribution of $\varepsilon$ in Eq. 6 and $\Gamma = f(R_{OT})$ in Eq. 2 to obtain a parameterization for $\Gamma_B$.

If, as described above, $\varepsilon$ is log-skew-normally distributed because it is the sum of log-normal random variables (call these $\epsilon_i$), then $L_O$ must also be so distributed. If we pass the $N^3$ to the summed-over $\epsilon_i$, such that $\varepsilon/N^3 = \sum_i (\epsilon_i/N^3)$, this will introduce a correlation into the $\epsilon_i/N^3$ being summed over, which does not affect the log-skew-normality of their sum [Hcine and Bouallegue, 2015]. For the $\epsilon_i$ this introduces another variable multiplicative factor and these each should still then be log-normal, so altogether $\sum_i (\epsilon_i/N^3)$ is still the sum of log-normal random variables. As the skew-normal distribution is insensitive to multiplicative transformations – we may write any skew-normal random variable $s$ as $s = \mu + \sigma(\delta|n_1| + n_2\sqrt{1-\delta^2})$ [Pourahmadi, 2007] so multiplying by some factor $m$ just changes $\mu \longrightarrow m\mu$ and $\sigma \longrightarrow m\sigma$ – the log-skew-normal must be equivalently insensitive to exponentiation. Thus, if $\varepsilon$ is log-skew-normally distributed, then so is $L_O = (\varepsilon/N^3)^{\frac{1}{2}}$. We indeed find that all the eight datasets used here, as well as the aggregated dataset, are well-described by a log-skew-normal distribution (Kuiper's statistic $V = 0.021$ for the combined dataset and as well as the median across the individual datasets; Figure S1, Table S1).

In contrast, there is not a clear description of the probability distribution of $L_T$. While calculating $L_T$ is straightforward conceptually, it involves subjective choices that bias its distribution [Mater et al., 2015]. Thus, it is less practical to identify the 'true' distribution for $L_T$ from collections of $L_T$ measurements as is the equivalent identification of an underlying distribution for $\varepsilon$. It has recently been argued that the size of turbulent overturns, which are thought to have a correspondence with $L_T$, should be power-law distributed [Smyth et al., 2019], but we find little to no evidence of this in the datasets we use here (Table S2). Either $L_T$ is not proportional to patch size in the data sets examined here, or the patch sizes in these datasets are not power-law distributed [Clauset et al., 2009]. The $L_T$ probability distribution for all the datasets used here is unimodal in log-space, such that at least the bulk of the distributions are qualitatively much closer to log-normally distributed than power-law distributed (Figure S1).

Regardless, it is the probability distribution of $R_{OT} = L_O/L_T$ that is of interest here, for which a parameterization of $L_T$'s probability is not necessary. Instead, one needs a suitable description of the *conditional* distribution of $L_T$ for a given $L_O$ value, i.e. $P(L_T|L_O)$. Somewhat surprisingly, we find that the log-skew-normal is again an excellent description of $R_{OT}$'s probability distribution ($V = 0.010$ for the combined dataset and the median $V = 0.027$ across the eight individual datasets; Table S1), but that in this case the log-skewness is *negative*. $R_{OT}$'s log-skewness is in fact only positive for the IH18 dataset (skewness $\tilde{\mu}_3 = 0.48$). Thus, the log-skew-normal distribution is, similar to $\varepsilon$ and $L_O$, a satisfactory parameterization of $R_{OT}$'s probability distribution, but by and large with a different sign in log-skewness. The question then becomes how $L_O$ and $L_T$ are related such that $R_{OT}$'s distribution is: i) log-skew-normal; ii) with a negative log-skewness; and iii) peaked at $O(1)$. In the absence of a theory for $L_T$'s probability distribution, we derive an empirical construction of $P(L_T|L_O)$ that recapitulates the probability distribution of $R_{OT}$ with simulated $L_T$ values. As $L_T$ appears to be approximately log-normally distributed (Figure S1), it is justifiable to base such a construction on a statistical scaling relationship, with multiplicative fluctuations.

[6] Historically, $\varepsilon$ has often been described as log-normally distributed based on a simple argument from multi-stage subdivisions of an initial flux for 3D homogeneous isotropic turbulence [Gurvich and Yaglom, 1967], but the log-normal distribution has also been long recognized as a quantitatively inaccurate description of measured distributions [Yamazaki and Lueck, 1990].

The conditions (i-iii) above can occur for $R_{OT}$ if $L_T$ scales *heteroscedastically*, i.e. the fluctuations around how $L_T$ scales with $L_O$ are themselves also a function of $L_O$, and the coefficients of the scaling relationship are close to unity. If $L_T \sim \zeta L_O^{\beta}$ with $\zeta \approx 1$ and $\beta \approx 1$, then most values of $L_O/L_T$ will be $O(1)$. If this scaling relationship is tight when $L_O$ is large (i.e. small fluctuations and low probability of small $L_T$), this will make cases where $L_O/L_T$ is very large unlikely. Conversely, if this scaling relationship is weaker when $L_O$ is small (i.e. larger fluctuations and comparatively higher probability of larger $L_T$), this will make cases where $L_O/L_T$ is very small less unlikely. Together these heteroscedastic effects can change the sign of the log-skewness.

As both $L_O$ and $L_T$ are random variables, their scaling relationship can be captured by model II regression; the heteroscedasticity of this relationship can then be captured by quantile regression, which estimates the conditional quantiles of the response variable [Koenker and Hallock, 2001], applied to the residuals. First applying model II regression to the combined and log-transformed $(L_O, L_T)$ dataset (Figure 5b), we find the best-fit scaling for these data to be $L_T = 1.24 L_O^{1.01}$. (We use least-squares cubic regression [York, 1966] but other standard methods such as bisector or major axis yield similar results.) Then applying quantile regression to the residuals, we indeed find that the fluctuations in $L_T$ around this scaling relationship increase as $L_O$ decreases. This decrease is characterized by the relationship $r = r_o + r_1 \log_{10}(L_O)$, where $r$ is the amplitude of the residuals around the best-fit scaling, $r_o$ is the residual amplitude when $L_0 = 1$ m, and $r_1$ captures how the residuals decrease as $L_O$ increases. We calculate separate relationships for positive and negative residuals, via quantile regression. (We use the 10th and 90th percentiles to compute $r_o$ and $r_1$ but our results are not sensitive to this choice.) We can then recover the joint $(L_O, L_T)$ distribution and hence the $R_{OT}$ distribution by simulating the $L_T$ accordingly for a given $L_O$. The simulated $L_T$ values yield good agreement with the empirical $R_{OT}$ distribution ($V = 0.028$). We define $L_T = \eta \times 1.24 L_O^{1.01}$, where $\eta$ is a log-normal random variable with $\mu = 0$ and $\sigma \propto L_O$. This results in a conditional distribution for $L_T$ that matches the estimated scaling relationships, and is thus an approximation of the joint distribution of the data shown in Figure 5b.

Physically, this scaling behavior has an intuitive interpretation. $L_T$ scaling with $L_O$ with an exponent close to unity, such that the peak in the probability density function for $R_{OT}$ is $O(1)$, occurs because for most turbulent overturns the displacement of fluid parcels in that patch ($L_T$) will be of the same order of magnitude as the vertical scale that they can be displaced given the background stratification ($L_O$), as discussed above. If an overturn has a small $L_O$, this could either be because it is a small overturn, meaning it has a similarly small $L_T$, or a young overturn with a larger $L_T$ but still low turbulent dissipation. If $L_O$ is large, however, $L_T$ would be reasonably expected to be similarly large. This asymmetry in the life cycle of turbulent overturns produces this skewed, heteroscedastic scaling in Figure 5b.

### 4.1. *The recipe*

Altogether, these pieces can be assembled into a 'recipe' to compute a bulk flux coefficient as follows. To calculate $\Gamma_B$ for a coarse resolution grid cell of a general circulation model, which is large enough to enclose sufficient statistics, we imagine a total power $P$ is available for the grid cell (from one or more sources, such as tides, winds, etc.). Then

1. We assume an initial value of 0.2 for $\Gamma_B$ and using $P = \mathcal{M}_B + \mathcal{E}$ infer an initial value for $\mathcal{E}$; $\mathcal{M}_B$ and $\mathcal{E}$ are the total mixing and dissipation in the cell.
2. Distribute $\mathcal{E}$ over a log-skew-normal distribution of $\varepsilon$ for individual patches so then $\Sigma\varepsilon = \mathcal{E}$ and individual $\varepsilon$ values are random log-skew-normal sample draws.
3. Calculate $L_O$ for these patches by transforming their $\varepsilon$ values and a specified background stratification into a distribution for $L_O$.
4. Simulate $L_T$ values corresponding to these $L_O$ values according to the heteroscedastic scaling relationship between $L_O$ and $L_T$.
5. Calculate $\Gamma$ for these patches based on the $R_{OT}$ parameterization and their simulated $L_O$ and $L_T$ values (with a small background diffusivity added; see below).

6. Calculate $\mathcal{M}$ for individual patches based on their $\Gamma$
7. Sum over $\mathcal{M}$ of all patches to compute $\mathcal{M}_B$.
8. Calculate $\Gamma_B = \mathcal{M}_B/\mathcal{E}$
9. Back to step one and iterate until $\Gamma_B$ converges (usually within a few iterations).

We repeat this procedure many times and take the average of the repetitions to account for the stochasticity in the process; in practice for oceanographically relevant values the variability between repetitions is negligible. In step 4 we also add a small (regularizing) background diffusivity of $\kappa_{background} = 10^{-6.5}$ m$^2$/s to account for the background wave field processes not encapsulated in the $R_{OT}$ parameterization. This value, which is close to the molecular diffusion coefficient of heat in seawater, was chosen according to the inflection point on the probability density function of $\kappa$ from the combined dataset described above, which separates the low tail of the distribution from the bulk of the data (Figure S2). Including this $\kappa_{background}$, mixing for each patch becomes $M_{background}+M_{patch} = \Gamma_{total}\epsilon = \kappa_{background}N^2+\Gamma_{param}\epsilon$ and so $\Gamma_{total} = \kappa_{background}N^2+\Gamma_{param}$, where $\Gamma_{param}$ is the $R_{OT}$-parameterized $\Gamma$ value. The advantage of adding the background mixing is that it allows for $\Gamma$ to tend to large values in the limit of weak turbulence, i.e. when there is not much turbulence due to lack of power and/or overly strong stratification. From a physical perspective, in the limit of laminar flow, $\Gamma \to \infty$ since $\varepsilon$ vanishes but $\mathcal{M}$ remains finite. Note that Figure 2a shows high values of $\Gamma$ for observed weakly turbulent young patches.

Steps 1-5 involve a degree of computational expense that might prove impractical if applied to every grid cell of an Earth System Model at every time step. It is, however, straightforward and computationally cheap to do the calculations a priori for ranges of power and stratification relevant to the oceans, and then employ a simple lookup table in the models. We employ this approach for the purpose of analysis to be discussed in the next section.

## 5. Regional/global implications

To highlight the importance of the appropriately estimated bulk $\Gamma_B$ for the larger scale circulation, we next apply our recipe to the the South Atlantic Ocean. This choice is made for four reasons: (I) it is a region of tidal-shear dominance; (II) it is home to the iconic Brazil Basin Tracer Release Experiment (BBTRE), the first deep observation of bottom generated enhanced turbulent mixing [Polzin et al., 1997, Ledwell et al., 2000, St. Laurent et al., 2001]; (III) there exist three-dimensional tidal power estimates which are constructed on theoretical grounds but have been verified to agree excellently with observations including BBTRE [de Lavergne et al., 2020]; and (IV) the BBTRE data were shown to correspond to small-scale marginally-unstable shear-induced turbulence, and hence consistent with the above-mentioned physical and statistical paradigms laid out in CM21, MCA21 and MCBC21.

We construct a regional map from the sum of the tidally-generated power and the contribution of the wind-induced near-inertial shear. The former, from de Lavergne et al. [2019, 2020], estimates the power in the internal tide field generated through the interaction of ocean tides with ocean bathymetry. The latter, constructed based on Alford [2020], estimates the power in the downward-propagating wind-induced internal wave field. Figure 5a-c show the power plotted on three density layers (shallow, intermediate depth, and deep).[7] The deep layer, as will be shown, corresponds to the peak cross-density (diapycnal) turbulent mixing that helps sustain the deep branch of the ocean circulation [de Lavergne et al., 2017]. The shallow layer approximately corresponds to the base of the subtropical wind-driven gyres across which turbulence mixes heat, carbon, and nutrients between the upper ocean and the deep ocean [Talley, 2011]. The patterns on the shallow layer reflect the seasonal atmospheric storm patterns at the surface and the rough topography at the bottom (from which internal tides radiate upwards). The patterns on the deep layer primarily represent the tidally-generated turbulence with little contribution

[7]Neutral density is often used instead of in-situ density as it removes the dynamically inconsequential compression of seawater due to increase in pressure with depth. It is also common to deduct 1000 when reporting seawater density as in Figure 5.

from the surface generated downward propagating waves. The intermediate layer 'feels' both top and bottom generated wave fields.

The power maps, together with climatological hydrographic information (i.e. $N^2$) from the World Ocean Circulation Experiment [WOCE; see Gouretski and Koltermann, 2004], enable us to calculate $\Gamma_B$ on the WOCE grid on which the power maps in panels a-b are constructed. The resolution of the grid is half a degree in the horizontal (~50km) and ~110m in the vertical. Applying our recipe to each grid cell with its local power and density stratification, we obtain the maps of $\Gamma_B$ shown in panels d-f of the figure. We normalized the maps by the value of 0.2, the value generically applied in climate models to split the power into $\mathcal{M}$ and $\varepsilon$[8]. Since the bulk turbulent diffusivity for each grid cell may be defined as $K \approx \Gamma_B N^2/\epsilon$, and the associated buoyancy flux is defined as $\mathcal{F} \approx KN^2$, the ratio $\Gamma_B/0.2$ is exactly equal to the ratios of $K$ and $\mathcal{F}$ calculated based on our recipe to their values when calculated using $\Gamma_B = 0.2$. Thus, panels d-f represent all three ratios and collectively show not only that mixing can vary by up to a factor of 4 regionally, but also that there are significant spatial patterns to the variations. For example, if one is interested in the flux of quantity $C$ between the surface ocean and the deep ocean (where $C$ can be heat, carbon or nutrients for example) panel d shows that the flux $KC$ can be overestimated or underestimated by more than 200% if physics and statistics of breaking waves are ignored.

In the deep ocean, internal wave-driven turbulence leads to the irreversible transformation of water masses which facilitates upwelling of dense waters and closure of the deep branch of ocean meridional circulation [Nikurashin and Vallis, 2011, Mashayek et al., 2015, de Lavergne et al., 2016b, Ferrari et al., 2016]. The net transformation along a density layer consists of a complex pattern of upwelling and downwelling defined based on topographic features, large scale stratification, and turbulence [de Lavergne et al., 2016a, Mashayek et al., 2017c, Cimoli et al., 2019, de Lavergne et al., 2017]. To show this, we note that water moves across density surfaces at the diapycnal velocity:

$$\tilde{\mathbf{e}} = \frac{\partial_z \mathcal{M}}{N^2}. \tag{9}$$

This quantity, plotted in Figure 5g on the deep density layer, is positive/negative where waters become lighter/denser–note that a variable $\Gamma_B$ based on our recipe is used for panel $g$. Specifically, $\tilde{\mathbf{e}}$ is positive when $\mathcal{M}$ decreases with depth, for example when there is surface-intensified mixing, or in the bottom boundary layer where a one-dimensional model would predict $\mathcal{M} \to 0$ towards the ocean floor [Ferrari et al., 2016] (what actually happens in real ocean boundary layers is not yet known). Conversely, $\tilde{\mathbf{e}}$ is negative and waters become denser when mixing increases with depth, for example in the ocean interior near rough topography, where mixing is enhanced toward the bottom. The net transformation over $\gamma = 28.1$ in panel $g$ is therefore the sum of sharp near boundary upwelling squeezed between the bottom topography and regions of intense downwelling. Panel $h$ shows the same map as panel $g$ but calculated with $\Gamma_B = 0.2$, while panel $i$ shows the difference between panels $g$ and $h$. Red regions in panel $i$ imply either an increase in local upwelling or a decrease in local downwelling whereas blue implies enhanced downwelling or diminished upwelling [see Cimoli et al., 2019, for a more detailed discussion of these patterns]. The change to $\tilde{\mathbf{e}}$ is of the same order of magnitude as $\tilde{\mathbf{e}}$ itself. Thus, the variations in $\Gamma_B$ exert a leading order control over the rate and patterns of deep upwelling and downwelling in the ocean. Panel $i$ shows an interesting asymmetric variations across mid-ocean ridges which has important implications for mixing across the ridge and the broader inter-hemispheric Atlantic Meridional Circulation as well as inter-based exchange of water masses.

The net water mass transformation rate across a density surface may be defined as the area integral of the local transformation $\tilde{\mathbf{e}}$ through

$$\mathcal{D}(\gamma_*^n) = -\iint_{A(\gamma_*^n)} \tilde{\mathbf{e}} \cdot \widehat{\mathbf{n}}\, dA, \tag{10}$$

[8]This is the approach used in models that employ an evolving tidal mixing parameterization rather than a fixed diffusivity.

where $\widehat{\mathbf{n}}$ is the unit vector normal to the neutral density surface $\gamma_*^n$, and $A$ is the surface area of the density surface. $\mathcal{D}$ is measured in Sverdrups where 1Sv=$10^6$ m$^3$/s. Figure 5j shows the net rate as a function of neutral density with the mean depth of the density layers shown in panel $k$. In the deep ocean $\Gamma_B$ enhances the net transformation rate by over 50% which is significant (although regionally the changes can be much higher as shown in panel $i$). In the upper ocean, however, the change to the net is also equally important.

Figure 5l shows the normalized histograms of $\Gamma_B/(A/2)$, where $A$ was the coefficient in Eq. 2, for various density levels. The density levels range from a shallow layer with mean depth of 490m and a total area of 0.98 of the sea surface area to a deep layer with mean depth of 3170m and an area of 0.63 of the sea surface area. Recalling that $R_{OT} \sim 1$ corresponds to $\Gamma_{goldilocks} = A/2$, the fact that the histograms center around $A/2$ implies that for most of the ocean, the statistical distribution of patches (within each grid cell of the climatological data) mostly comprises patches that are optimally mixing. This implies an optimal balance between stratification and shear (remembering $R_{OT} \sim Re_b^* \sim Ri^* \sim 1$ in the Goldilocks regions). Or in other words, regions of strong shear **coincide** with the most 'desirable' stratification to yield optimal mixing.[9] For completeness, panel $m$ shows the same distributions as in panel $l$ but normalized by 0.2.

## 6. Discussion

Most turbulent patches are characterized by an optimal balance between density stratification ($N^2$) and velocity shear ($S$). In this 'Goldilocks' state, where neither $N^2$ or $S$ are too weak nor too strong, turbulent mixing is most efficient. This was shown in the data presented in Figure 2a,b, in agreement with data published in other aforementioned works (see MCA21 for a review). Figure 5 showed that this result extends to the global scale as well: regions of high turbulence generation (hence enhanced shear) coincide with desirable stratification for optimal mixing. The coincidence of regions of Goldilocks mixing with energetically turbulent regions points to a deeper underlying physical implication: for such a balance to exist, the ocean circulation should restratify the turbulent zones at rates that optimize the balance of stratification and shear and keeps them at a desirable marginally-unstable state (where $R_{OT} \sim Re_b^* \sim Ri^* \sim 1$). If stratification is too strong for a given power injection, it will suppress turbulence and conversely, if the stratification is too weak in a high power injection region, it will get eroded rapidly. In the former scenario, energy builds up until a burst of a cascade of hydrodynamic instabilities [Mashayek and Peltier, 2012a,b] releases the stored available potential energy and moves the system to a near critical state. In the latter state, the larger scale circulation replenishes the stratification until the flow becomes marginally unstable. Therefore, the coincidence of stratification and shear (manifested as $Ri^* \sim 1$) is not accidental, but instead due to the fine balance between the processes that build stratification with those that erode it.

Given that ocean mixing is a key driver of ocean circulation itself, the variability of $\Gamma_B$ may provide a self-optimizing mechanism to the diffusively-driven branch of ocean circulation. It also provides the system with the ability to sustain internal variability driven by mixing. Such a feedback mechanism is entirely absent from all Earth System Models. Not only is its inclusion important for understanding the present day climate, but it could also be key to allowing the ocean and the climate system to adapt properly to changes in ocean stratification and power that are induced by changes to surface winds, air-sea buoyancy exchanges, and changes to tidal power (due for example to changes in sea levels and ice shelves in the polar regions). We conclude by highlighting the need for: (I) better understanding of how the Goldilocks ocean stratified mixing is linked with the larger scale horizontal and meridional ocean circulation patterns; and (II) what the implications are once this feedback is included in climate models.

The recipe proposed in this work, the code for which is available in the Supplementary Materials, may be used to tackle these problems in climate models as a first step. On regional scales, observations can be

[9] We use $A = 2/3$ in this study since it was justified on physical grounds by MCA21 and also was inferred through data regression as shown in Figure 3a and in MCA21. With this value, $\Gamma_{gold} = 1/3 \sim 1.7 \times 0.2$ which corresponds to green in panels $d - f$ of Figure 5. But note that other values of $A$ will not change the histogram in panel $l$ and the optimal mixing argument that we make.

parameterized in terms of the log-skew-normal distribution of CM21 and our recipe to improve mixing estimates, though uncertainties may be appreciable depending on sample size. Understanding how the $L_O - L_T$ scaling parameters change with space and time warrants further research as it is clearly critical to developing our recipe further, most importantly the scaling exponent relating $L_O$ and $L_T$ (Figure 4d).

## 7. Materials

Panels a-c in Figure 1 are produced from a higher-resolution of the observationally-tuned simulations of [Mashayek et al., 2017b] and panels d,e are based on simulations of [Mashayek et al., 2013]. Panels a,b of Figure 2 and panel b of Figure 4 are based on the data used in MCA21, a brief description of which is included in the Supplementary Materials. Panels c,d of Figure 2 are based on the simulations of [Mashayek and Peltier, 2011b,a, 2013a, Mashayek et al., 2013]. Figure 4a is reproduced from CM21, based on 14 microstructure field experiments discussed therein. In Figure 5, the tidal power is from [de Lavergne et al., 2020] and the wind power is based on [Alford, 2020] who provided estimates for the net wind work on the ocean, the portion of the power that is dissipated within the surface mixed layer, and the radiation efficiency of the downward propagating wavefield.

## References


M. H. Alford. Global calculations of local and remote near-inertial-wave dissipation. *Journal of Physical Oceanography*, 50(11):3157–3164, 2020.

D. Bouffard and L. Boegman. A diapycnal diffusivity model for stratified environmental flows. *Dynamics of Atmospheres and Oceans*, 2013.

B. Cael and A. Mashayek. Log-skew-normality of ocean turbulence. *Physical Review Letters*, 126(22): 224502, 2021.

D. R. Caldwell and J. N. Mourn. Turbulence and mixing in the ocean. *Reviews of Geophysics*, 33(2 S): 1385–1394, 1995. ISSN 19449208. http://dx.doi.org/10.1029/95RG00123.

C. P. Caulfield. Layering, Instabilities, and Mixing in Turbulent Stratified Flows. *Annual Review of Fluid Mechanics*, 53(1):113–145, 1 2021. ISSN 0066-4189. http://dx.doi.org/10.1146/annurev-fluid-042320-100458. URL https://doi.org/10.1146/annurev-fluid-042320-.

L. Cimoli, C. P. Caulfield, H. L. Johnson, D. P. Marshall, A. Mashayek, A. C. Naveira Garabato, and C. Vic. Sensitivity of Deep Ocean Mixing to Local Internal Tide Breaking and Mixing Efficiency. *Geophysical Research Letters*, 46(24):14622–14633, 12 2019. ISSN 0094-8276. http://dx.doi.org/10.1029/2019GL085056. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085056.

A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

C. de Lavergne, G. Madec, J. Le Sommer, A. J. G. Nurser, and A. C. Naveira Garabato. The impact of a variable mixing efficiency on the abyssal overturning. *Journal of Physical Oceanography*, (2015), 2015.

C. de Lavergne, G. Madec, X. Capet, G. Maze, and F. Roquet. Getting to the bottom of the ocean. *Nature Geoscience*, 9(12):857–858, 12 2016a. ISSN 1752-0894. http://dx.doi.org/10.1038/ngeo2850. URL http://www.nature.com/articles/ngeo2850.

C. de Lavergne, G. Madec, J. Le Sommer, A. J. G. Nurser, and A. C. Naveira Garabato. On the Consumption of Antarctic Bottom Water in the Abyssal Ocean. *Journal of Physical Oceanography*, 46(2):635–661, 2 2016b. ISSN 0022-3670. http://dx.doi.org/10.1175/JPO-D-14-0201.1. URL http://journals.ametsoc.org/doi/10.1175/JPO-D-14-0201.1.

C. de Lavergne, G. Madec, F. Roquet, R. M. Holmes, and T. J. McDougall. Abyssal ocean overturning shaped by seafloor distribution. *Nature*, 551(7679):181–186, 11 2017. ISSN 0028-0836. http://dx.doi.org/10.1038/nature24472. URL http://www.nature.com/articles/nature24472.

C. de Lavergne, S. Falahat, G. Madec, F. Roquet, J. Nycander, and C. Vic. Toward global maps of internal tide energy sinks. *Ocean Modelling*, 137:52–75, 5 2019. ISSN 14635003. http://dx.doi.org/10.1016/j.ocemod.2019.03.010. URL https://linkinghub.elsevier.com/retrieve/pii/S1463500318302890.

C. de Lavergne, C. Vic, G. Madec, F. Roquet, A. F. Waterhouse, C. Whalen, Y. Cuypers, P. Bouruet-Aubertot, B. Ferron, and T. Hibiya. A parameterization of local and remote tidal mixing. *Journal of Advances in Modeling Earth Systems*, 12(5):e2020MS002065, 2020.

T. M. Dillon. Vertical overturns: A comparison of Thorpe and Ozmidov length scales. *Journal of Geophysical Research*, 87(C12):9601, 1982. ISSN 0148-0227. http://dx.doi.org/10.1029/jc087ic12p09601.

R. Ferrari, A. Mashayek, T. J. McDougall, M. Nikurashin, and J.-M. Campin. Turning ocean mixing upside down. *Journal of Physical Oceanography*, (2016), 2016.

A. E. Gargett and J. N. Moum. Mixing efficiencies in turbulent tidal fronts: Results from direct and indirect measurements of density flux. *Journal of Physical Oceanography*, 25(11):2583–2608, 1995.

C. Garrett and E. Kunze. Internal tide generation in the deep ocean. *Annu. Rev. Fluid Mech.*, 39:57–87, 2007.

H. Gibson. Turbulence , Mixing , and Heat Flux in the Ocean Main Thermocline. 96(Cll):403–420, 1991.

V. Gouretski and K. P. Koltermann. {WOCE} global hydrographic climatology. *Berichte des BSH*, 35: 1–52, 2004.

Gregg. Statistics of Shear and Dissipation. *JPO*, 1992.

M. Gregg, E. D'Asaro, J. Riley, and E. Kunze. Mixing Efficiency in the Ocean. *Annual Review of Marine Science*, 10(1):443–473, 1 2018. ISSN 1941-1405. http://dx.doi.org/10.1146/annurev-marine-121916-063643. URL http://www.annualreviews.org/doi/10.1146/annurev-marine-121916-063643.

A. S. Gurvich and A. M. Yaglom. Breakdown of Eddies and probability distributions for small-scale turbulence. *Physics of Fluids*, 10(9), 1967. ISSN 10706631. http://dx.doi.org/10.1063/1.1762505.

M. B. Hcine and R. Bouallegue. Highly accurate log skew normal approximation to the sum of correlated lognormals. *arXiv preprint arXiv:1501.02347*, 2015.

G. N. Ivey, K. B. Winters, and J. R. Koseff. Density stratification, turbulence, but how much mixing? *Annu. Rev. Fluid Mech.*, 40:169–184, 2008.

R. Koenker and K. F. Hallock. Quantile regression. *Journal of economic perspectives*, 15(4):143–156, 2001.

J. R. Ledwell, E. T. Montgomery, K. L. Polzin, L. C. S. Laurent, R. W. Schmitt, and J. M. Toole. Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature*, 403:182–189, 2000.

S. Legg. Mixing by oceanic lee waves. *Annual Review of Fluid Mechanics*, 53, 2021.

A. Mashayek and W. R. Peltier. Turbulence transition in stratified atmospheric and oceanic shear flows: Reynolds and Prandtl number controls upon the mechanism. *Geophysical Research Letters*, 38: L16612, 5 PP, 2011a.

A. Mashayek and W. R. Peltier. Three-dimensionalization of the stratified mixing layer at high Reynolds number. *Physics of Fluids*, 23:111701, 2011b.

A. Mashayek and W. R. Peltier. Turbulence transition in stratified atmospheric and oceanic shear flows: Reynolds and Prandtl number controls upon the mechanism. *Geophysical Research Letters*, 38(16): L16612, 2011c.

A. Mashayek and W. R. Peltier. The 'zoo' of secondary instabilities precursory to stratified shear flow transition. Part 1 Shear aligned convection, pairing, and braid instabilities. *Journal of Fluid Mechanics*, 708:5, 2012a.

A. Mashayek and W. R. Peltier. The 'zoo' of secondary instabilities precursory to stratified shear flow transition. Part 2 The influence of stratification. *Journal of Fluid Mechanics*, 708:45, 2012b.

A. Mashayek and W. R. Peltier. Shear-induced mixing in geophysical flows: does the route to turbulence matter to its efficiency? *Journal of Fluid Mechanics*, 725:216–261, 2013a.

A. Mashayek and W. R. Peltier. Shear-induced mixing in geophysical flows: does the route to turbulence matter to its efficiency? *Journal of Fluid Mechanics*, 725:216–261, 2013b.

A. Mashayek, C. P. Caulfield, and W. R. Peltier. Time-dependent, non-monotonic mixing in stratified turbulent shear flows: implications for oceanographic estimates of buoyancy flux. *Journal of Fluid Mechanics*, 736:570–593, 2013.

A. Mashayek, R. Ferrari, M. Nikurashin, and W. R. Peltier. Influence of Enhanced Abyssal Diapycnal Mixing on Stratification and the Ocean Overturning Circulation. *J. Phys. Oceanogr.*, 45(10):2580–2597, 2015.

A. Mashayek, C. P. Caulfield, and W. R. Peltier. Role of overturns in optimal mixing in stratified mixing layers. *Journal of Fluid Mechanics*, 826:522–552, 2017a.

A. Mashayek, R. Ferrari, S. Merrifield, J. R. Ledwell, L. St Laurent, and A. N. Garabato. Topographic enhancement of vertical turbulent mixing in the Southern Ocean. *Nature communications*, 8:14197, 2017b.

A. Mashayek, H. Salehipour, D. Bouffard, C. P. Caulfield, R. Ferrari, M. Nikurashin, W. R. Peltier, and W. D. Smyth. Efficiency of turbulent mixing in the abyssal ocean circulation. *Geophysical Research Letters*, 44(12):6296–6306, 2017c.

A. Mashayek, C. Caulfield, L. Baker, and B. B. Cael. A generalized marginal stability criterion for shear-induced ocean interior diapycnal turbulent mixing. *Earth and Space Science Open Archive–Under review by GRL*, page 16, 2021a. http://dx.doi.org/10.1002/essoar.10507781.1. URL https://doi.org/10.1002/essoar.10507781.1.

A. Mashayek, C. P. Caulfield, and M. H. Alford. Goldilocks mixing in shear-induced ocean turbulence. *Journal of Fluid Mechanics*, (In Press; https://arxiv.org/abs/2108.05284):1—23, 2021b.

A. Mashayekhi. *Diapycnal Mixing in the Ocean: From Dissipation Scale to Large Scale Meridional Overturning Circulation*. PhD thesis, University of Toronto, 2013.

B. D. Mater and S. K. Venayagamoorthy. A unifying framework for parameterizing stably stratified shear-flow turbulence. *Physics of Fluids (1994-present)*, 26(3):36601, 2014.

B. D. Mater, S. K. Venayagamoorthy, L. S. Laurent, and J. N. Moum. Biases in thorpe-scale estimates of turbulence dissipation. Part I: Assessments from large-scale overturns in oceanographic data. *Journal of Physical Oceanography*, 45(10):2497–2521, 2015. ISSN 15200485. http://dx.doi.org/10.1175/JPO-D-14-0128.1.

S. G. Monismith, J. R. Koseff, and B. L. White. Mixing Efficiency in the Presence of Stratification: When Is It Constant? *Geophysical Research Letters*, 45(11):5627–5634, 2018. ISSN 19448007. http://dx.doi.org/10.1029/2018GL077229.

M. Nikurashin and R. Ferrari. Overturning Circulation Driven by Breaking Internal Waves. *Journal of Physical Oceanography*, page DOI: 10.1002/grl.50542, 2013.

M. Nikurashin and G. Vallis. A theory of deep stratification and overturning circulation in the ocean. *Journal of Physical Oceanography*, 41(3):485–502, 2011.

T. R. Osborn. Estimates of the local rate of vertical diffusion from dissipation measurements. *Journal of Physical Oceanography*, 10:83–89, 1980.

W. R. Peltier and C. P. Caulfield. Mixing efficiency in stratified shear flows. *Annual Rev. Fluid Mech.*, 35:135–167, 2003.

K. L. Polzin, J. M. Toole, J. R. Ledwell, and R. W. Schmitt. Spatial variability of turbulent mixing in the abyssal ocean. *Science*, 276(5309):93–96, 1997.

M. Pourahmadi. Construction of skew-normal random variables: are they linear combinations of normal and half-normal? 2007.

W. Smyth, J. Nash, and J. Moum. Self-organized criticality in geophysical turbulence. *Scientific reports*, 9(1):1–8, 2019.

W. D. Smyth and J. N. Moum. Length scales of turbulence in stably stratified mixing layers. *Physics of Fluids*, 12:1327, 2000.

L. C. St. Laurent, J. M. Toole, and R. W. Schmitt. Buoyancy Forcing by Turbulence above Rough Topography in the Abyssal Brazil Basin*. *Journal of Physical Oceanography*, 31(12):3476–3495, 2001.

L. D. Talley. *Descriptive physical oceanography: an introduction*. Academic press, 2011.

L. D. Talley, R. A. Feely, B. M. Sloyan, R. Wanninkhof, M. O. Baringer, J. L. Bullister, C. A. Carlson, S. C. Doney, R. A. Fine, E. Firing, N. Gruber, D. A. Hansell, M. Ishii, G. C. Johnson, K. Katsumata, R. M. Key, M. Kramp, C. Langdon, A. M. MacDonald, J. T. Mathis, E. L. McDonagh, S. Mecking, F. J. Millero, C. W. Mordy, T. Nakano, C. L. Sabine, W. M. Smethie, J. H. Swift, T. Tanhua, A. M. Thurnherr, M. J. Warner, and J. Z. Zhang. Changes in Ocean Heat, Carbon Content, and Ventilation: A Review of the First Decade of GO-SHIP Global Repeat Hydrography. *Annual Review of Marine Science*, 8:185–215, 1 2016. ISSN 19410611. http://dx.doi.org/10.1146/annurev-marine-052915-100829. URL www.go-ship.org.

S. A. Thorpe. The turbulent ocean. *Cambridge University Press*, 2005.

Z. Wu, X. Li, R. Husnay, V. Chakravarthy, B. Wang, and Z. Wu. A novel highly accurate log skew normal approximation method to lognormal sum distributions. In *2009 IEEE Wireless Communications and Networking Conference*, pages 1–6. IEEE, 2009.

H. Yamazaki and R. Lueck. Why Oceanic Dissipation Rates Are Not Lognormal. *Journal of Physical Oceanography*, 20(12):1907–1918, 1990. ISSN 0022-3670. http://dx.doi.org/10.1175/1520-0485(1990)020<1907:wodran>2.0.co;2.

D. York. Least-squares fitting of a straight line. *Canadian Journal of Physics*, 44(5):1079–1086, 1966.

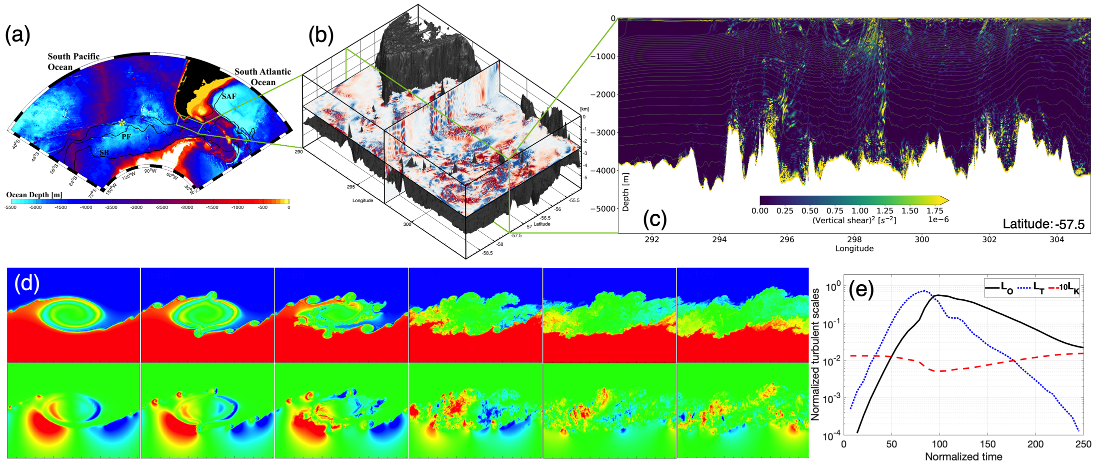


***Figure 1.*** *(a) A sector of the Southern Ocean in which strong westerly winds at the surface and interaction of Antarctic Circumpolar Current system (ACC; illustrated by the black lines) with rough topography create an energetic internal wave field. From an observationally forced and tuned high resolution numerical simulation. (b) A zoomed out 3D view of the wave field in the Drake Passage which is particularly energetic due to the geometric constriction on the ACC flow. Red/blue patterns represent positive/negative vertical motion. (c) Regions of high vertical shear (i.e. vertical gradient of horizontal velocity) which are susceptible to wave breaking and turbulence. (d) Turbulence life cycle of a shear-induced breaking internal wave; the top row shows the evolution of density (with heavier water in 'red' mixing with overlying lighter water in 'blue', resulting in a mixed density in 'green') while the lower row shows the evolution of the vertical turbulent flux of density. (e) Evolution of three key turbulent length scales: the overturn scale ($L_T$), the turbulent eddy scale ($L_O$), and the dissipation scale ($L_K$). See Materials for information on data sources.*

**Cite this article:** Mashayek *et al*. (2021). Shear coincidence: implications of the statistics of ocean turbulence microphysics for global diapycnal mixing. *Flow*, *1*, E1. doi:10.1017/flo.2021.1

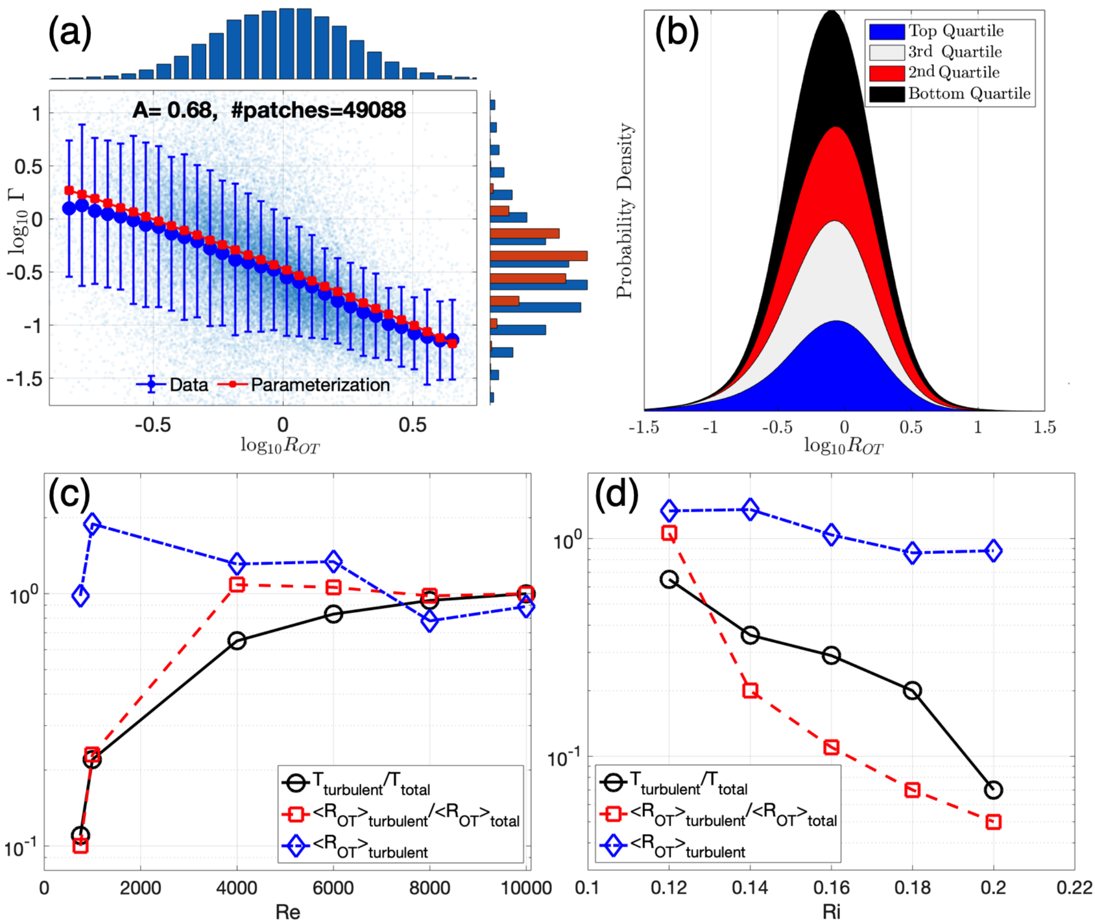


***Figure 2.*** *(a) Agreement of Eq. 2 with ∼49000 turbulent patches from six different datasets covering a variety of turbulent regimes and processes. The bar plot insets show the histogram of $R_{OT}$ (top axis) and $\Gamma$ for the parameterization (in red) and data (in blue) along the right vertical axis. The value of A is obtained through regression of data to Eq. 2. Reproduced from Mashayek et al. [2021b]–See Materials for a brief description of data. (b) probability density function (PDF) of $R_{OT} = L_O/L_T$ for the same data as in panel a, compartmentalised in terms of rate of dissipation of kinetic energy, $\varepsilon$. The mode of the PDF increases from 0.78 to 0.92 as additional $\varepsilon$-quartiles are included. Note the negative skewness of the log-transformed PDF. Panels (c) and (d) show the temporal fraction of turbulence life cycle as well as the ratio of $R_{OT}$ during turbulent phase of the flow to its mean value over the whole life cycle. Each symbol represents a life-cycle-averaged quantity from a direct numerical simulation (such as the one in Figure 1d) for the corresponding Reynolds and Richardson numbers. All cases in panel c are for $Ri = 0.12$ while all cases in panel d are for $Re = 6000$. Turbulent phase of the life cycles are defined as the times when $Re_b > 20$ [Gibson, 1991, Smyth and Moum, 2000]. See Materials for information on data sources.*

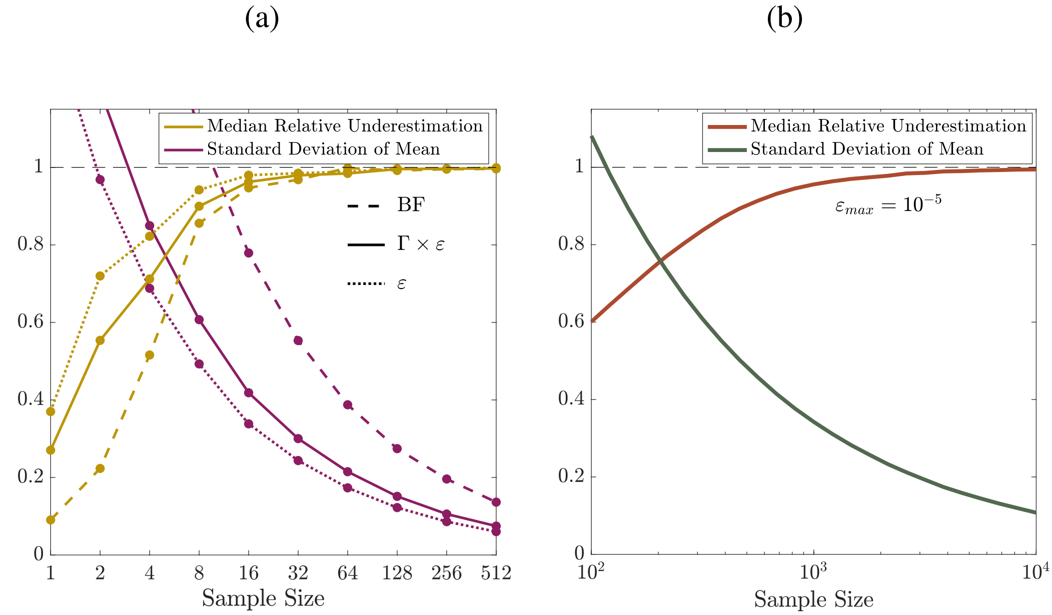


***Figure 3.*** *(a) Sampling bias and uncertainty for buoyancy flux (dashed lines), $\Gamma \times \varepsilon$ (solid lines), and $\varepsilon$ (dotted lines) for a single turbulent event. Purple lines indicate the normalized standard deviation, and yellow lines indicate the median relative underestimation, each as a function of sample size, estimated from bootstrap resampling a DNS simulation. For example, relative uncertainty is > 100% for the time-averaged buoyancy flux of this event with fewer than ~16 samples, and the median time-averaged buoyancy flux estimate with 4 or fewer samples underestimates the buoyancy flux by a factor of two or more. (b) Sampling bias and uncertainty for the mean $\varepsilon$ sampled from a log-skew-normal distribution with the parameters estimated from the dataset described in the text, discarding $\varepsilon$ values $>10^{-5}$ $m^2/s^3$. Green line indicates the normalized standard deviation, and orange line indicates the median relative underestimation, each as a function of sample size, estimated from bootstrap resampling.*

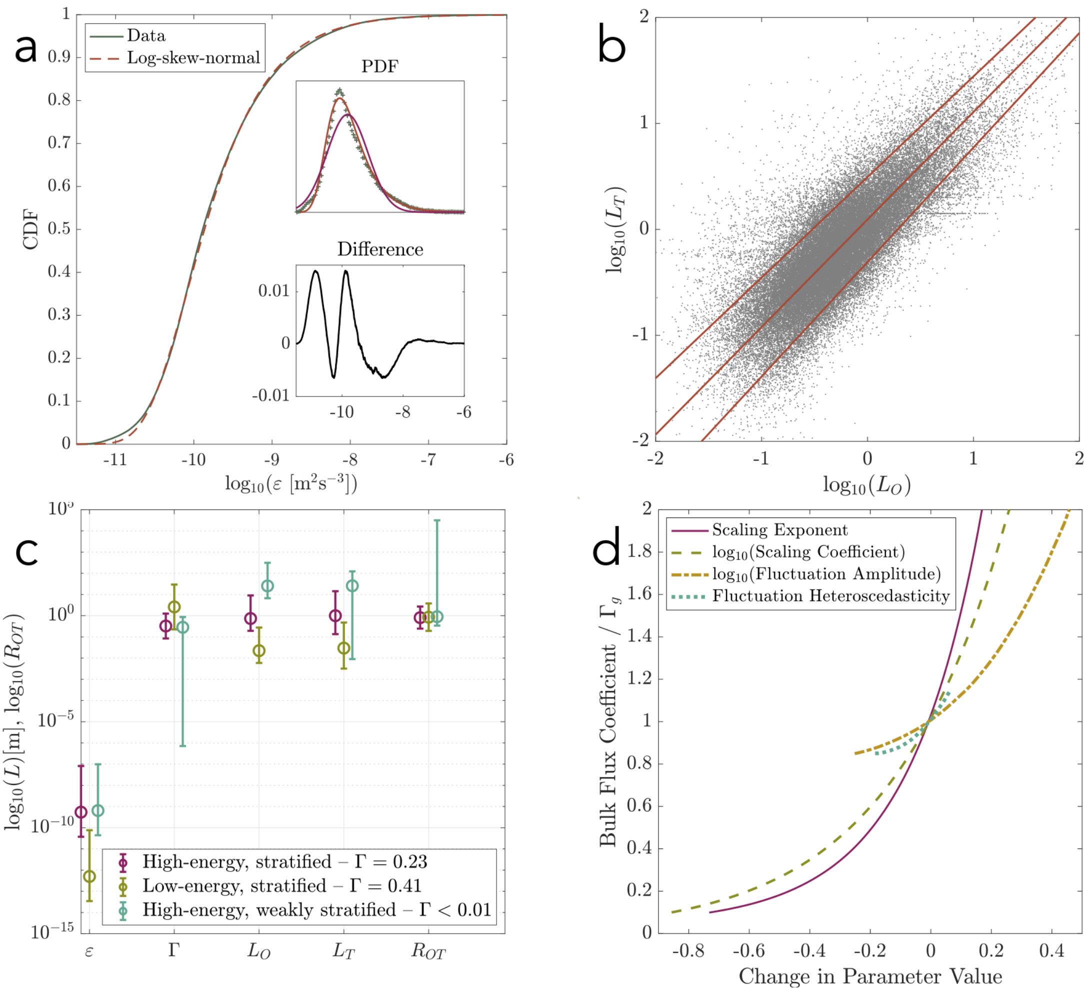


***Figure 4.*** *(a) Reproduced from CM21, $\varepsilon$ data from from over 750 full depth microstructure profiles from 14 field experiments (see Cael and Mashayek [2021] for more details) are excellently characterized by a log-skew-normal distribution. Cumulative distribution functions (CDFs) are shown in main plot; upper inset shows the corresponding probability density functions and that of a log-normal distribution (purple line) for comparison; lower inset shows the difference between the empirical and hypothesized log-skew-normal CDFs. (b) Regression of $\log_{10}(L_T)$ against $\log_{10}(L_O)$. Middle line captures the central scaling relationship and is estimated via model II regression as described in the text; outside lines capture the heteroscedasticity of the residuals and is estimated via quartile regression as described in the text. (c) (d) Sensitivity of the bulk flux coefficient to the different $L_T$-$L_O$ scaling parameters, perturbed from a baseline $\Gamma_g$ estimated from the combined global dataset described in the text. Larger fluctuations in $L_T$ when $L_O$ is large can increase the bulk flux coefficient, but it asymptotes to a constant value with decreasing fluctuations. Increasing either the scaling exponent or coefficient increases $L_T$ values for large $L_O$, thus increasing the bulk flux coefficient for large $\varepsilon$; decreasing either the scaling exponent or coefficient drives bulk mixing to zero as $R_{OT} \gg 1$ when $\varepsilon$ is large. See Materials for information on data sources.*

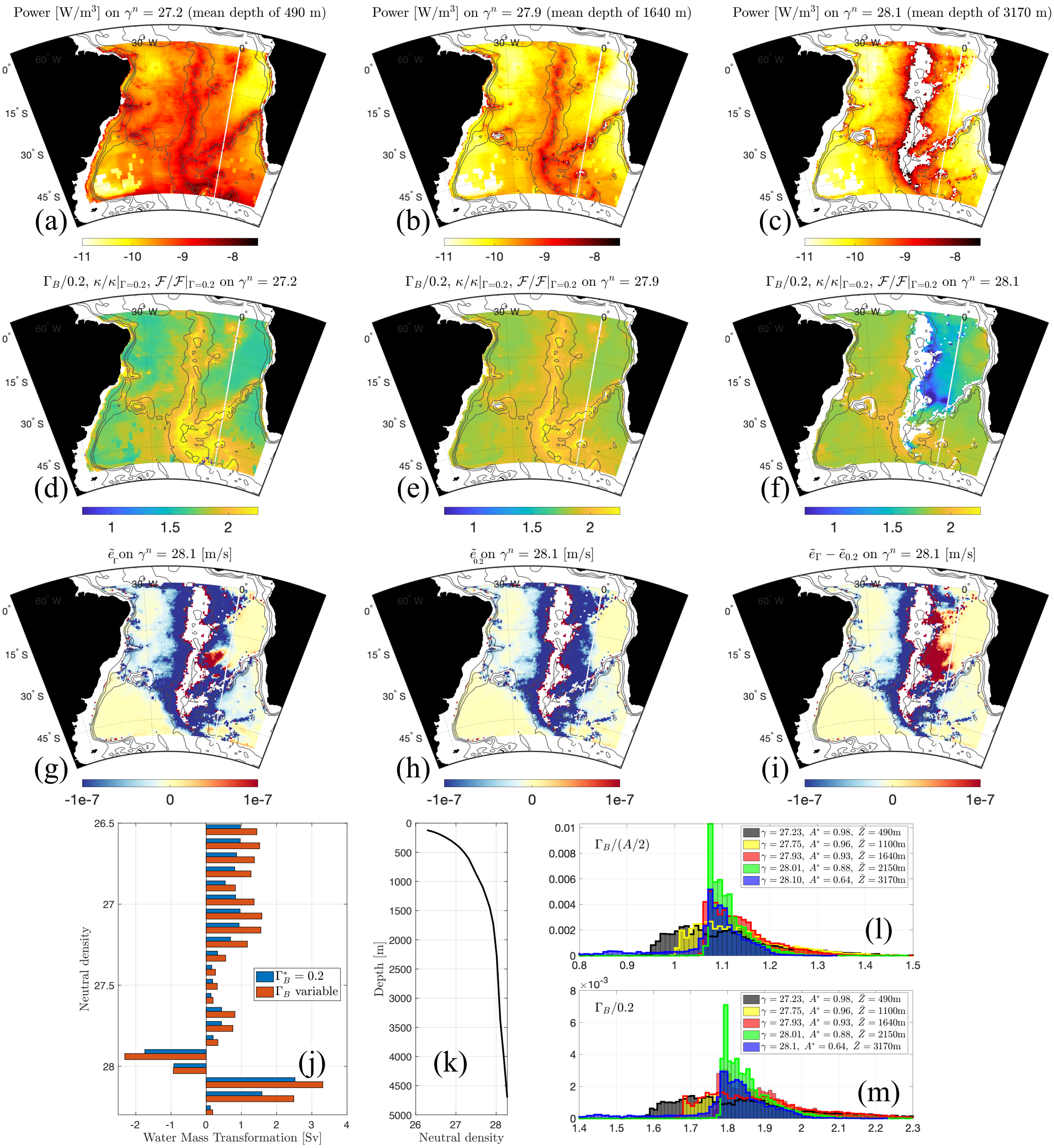


***Figure 5.*** *(a-c) The combined powers in the internal wave fields induced by tides and winds that is available to mixing and dissipation for the South Atlantic basin, plotted on various density levels. (d-f): $\Gamma$ normalized by $\Gamma^* = A/2$ based on Eq. 2, plotted on same density levels as in $a-c$. Since $K \approx \Gamma_B \epsilon / N^2$ and turbulent flux is $\mathcal{F} \approx KN^2$, $\Gamma_B/0.2$ is equal to ratios of $K$ and $\mathcal{F}$ based on $\Gamma_B$ calculated using our recipe to their values when $\Gamma_B = 0.2$ is used. (g): Local diapycnal velocity (see Eq. 9) showing upwelling (in red) and downwelling (in blue) on the deep 28.1 calculated using a variable $\Gamma_B$. (h) same as g but for the bulk $\Gamma_B = 0.2$. (i) difference between g and h. (j) Water mass transformation (i.e. integral of $\tilde{e}$, as per Eq. 10) calculated for $\Gamma_B = 0.2$ and variable $\Gamma_B$ calculated using our recipe. For reference, mean depth of isopycnals are also shown in panel $k$. (l,m): Histogram of the normalized $\Gamma_B$ on various density levels. The legend includes information on the mean depths of density levels as well as their total area normalized by the area at the sea surface. Histograms are normalized by $\Gamma_{goldilocks} = A/2$ (see Eq. 2) in panel $l$ and by 0.2 in panel $m$. See Materials for information on data sources.*

*Flow* (2021), **1** E1
doi:10.1017/flo.2021.1

FLOW CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

# Supplementary Materials
# Shear coincidence: implications of the statistics of ocean turbulence microphysics for global diapycnal mixing

A. Mashayek[1]*, B.B. Cael[2], M.H. Alford[3], L. Cimoli[3], and C.P. Caulfield[4]

[1]Imperial College London, London, UK
[2]National Oceanography Centre, Southampton, UK
[3]Scripps Institution of Oceanography, La Jolla, USA
[4]University of Cambridge, Cambridge, UK
*Corresponding author. E-mail: mashayek@ic.ac.uk

Received xx xxx xxxx

### *Turbulence patch data*

The six oceanic datasets employed for the purposes of analyses in Fig 2b and Fig 4b of the main text are the same as those employed by MCA21. Here we provide a brief description and refer to MCA21 for a more comprehensive discussion. The Tropical Instability Wave Experiment (TIWE) dataset includes turbulent patches sampled at the equator at $140^o$W in the shear-dominated upper-equatorial thermocline, between 60m and 200m depths, spanning both the upper and lower flanks of the Pacific Equatorial Undercurrent [Lien et al., 1995, Smyth et al., 2001]. The FLUX STAT (FLX91) experiment sampled turbulence at the thermocline (~350-500m depth), in part generated through shear arising from downward-propagating near-inertial waves, about 1000 km off the coast of northern California [Moum, 1996, Smyth et al., 2001]. The IH18 experiment measured full-depth turbulence (up to ~5300m deep) primarily generated by tidal flow over the Izu-Ogasawara Ridge (western Pacific, south of Japan), a prominent generation site of the semidiurnal internal tide that spans the critical latitude of 28.88N for parametric subharmonic instability [Ijichi and Hibiya, 2018]. The Samoan Passage data are measurements of abyssal turbulence generated by hydraulically-controlled flow over sills in the depth range 4500-5500m in the Samoan Passage, an important topographic constriction in the deep limb of the Pacific Meridional Overturning Circulation [see Alford et al., 2013, Carter et al., 2019, we use data from the latter]. The BBTRE data are from turbulence induced by internal tide shear in the deep Brazil Basin (~2500-5000m depth) and were acquired as a part of the original Brazil Basin Tracer Release Expermient (BBTRE; Polzin et al. [1997]), recently re-analyzed by Ijichi et al. [2020]. Also re-analyzed by Ijichi et al. [2020], we use the data from DoMORE which focused on flow over a sill on a canyon floor in the Brazil Basin [Clément et al., 2017, Ijichi et al., 2020].

| Dataset | $V(L_O)$ | $V(R_{OT})$ |
|---|---|---|
| Combined | 0.021 | 0.010 |
| BBTRE | 0.022 | 0.028 |
| DoMORE | 0.021 | 0.039 |
| FLX91 | 0.023 | 0.015 |
| Hawaii | 0.011 | 0.021 |
| IH18 | 0.008 | 0.027 |
| IWISE | 0.037 | 0.033 |
| SP | 0.015 | 0.017 |
| TIWE | 0.021 | 0.031 |

***Table 1.*** *Kuiper's statistic V for the log-skew-normal approximation of the $L_O$ and $R_{OT}$ distributions of each of the datasets and the combined dataset described in the text.*

| Dataset | Exponent | $n_{tail}$ (percentile) | $p$-value |
|---|---|---|---|
| Combined | 2.20 ± 0.01 | 12519 (80th) | <0.01 |
| BBTRE | 2.04 ± 0.03 | 5754 (73rd) | <0.01 |
| DoMORE | 3.27 ± 0.57 | 178 (98th) | 0.08 |
| FLX91 | 3.85 ± 0.24 | 598 (82nd) | 0.03 |
| Hawaii | 2.81 ± 0.12 | 1010 (90th) | 0.09 |
| IH18 | 2.95 ± 0.05 | 605 (94th) | **0.13** |
| IWISE | 2.64 ± 0.21 | 225 (82nd) | 0.07 |
| SP | 2.57 ± 0.05 | 3133 (76th) | <0.01 |
| TIWE | 3.65 ± 0.53 | 322 (72nd) | <0.01 |

***Table 2.*** *Scant evidence for power-law distributions in the Thorpe scale $L_T$ in the datasets used in this study. Left column gives the datasets described in the text and the combined dataset that aggregates all of these. Second column gives the power-law exponent estimate and its uncertainty, as calculated by the method described in Clauset et al. [2009]. The same method selects a minimum $L_T$ value above which a power-law tail is fit; the third column gives the number of samples above, and the percentile above, this minimum $L_T$ value in each case. The fourth column gives the probability that these tail data are power-law distributed, with the p-values above the rule-of-thumb significance criteria of 0.1 suggested by Clauset et al. [2009] in bold. The power-law hypothesis is rejected in all cases except IH18; in this case, the exponent is much steeper (~3) than those described by Smyth et al. [2019] and only ~6% of the data fall into this power-law tail. The cases with borderline p-values ($0.1 > p > 0.01$) also have steep exponents and/or small tail sample sizes. Altogether these results show that the power-law parameterization is a poor description of the total $L_T$ distribution considered here.*

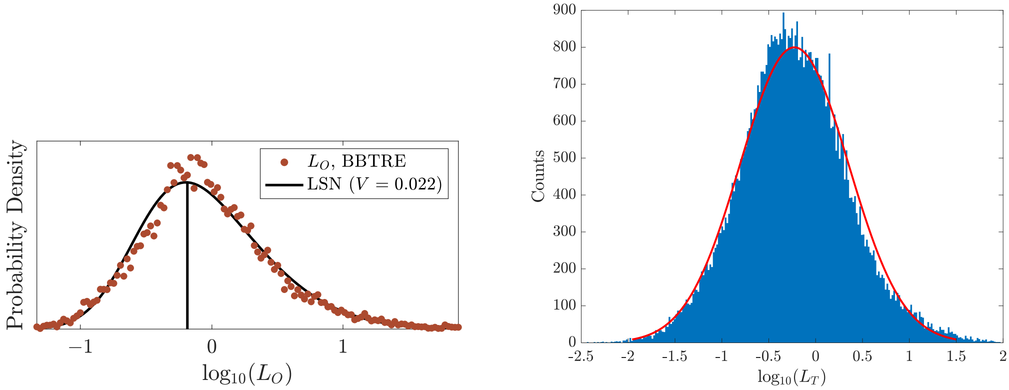


***Figure 1.*** *Left: Log-skew-normal (LSN) fit to $L_O$ data from the Brazil Basin Tracer Release Experiment (BBTRE), shown as probability density. Orange points are the data histogram, with number of bins equal to the (integer-rounded) square root of the sample size. Vertical line added to visualize skewness. V is Kuiper's statistic (see text), evaluated on cumulative probability density. For this log-skew-normal distribution, the log-skewness $\tilde{\mu}_3 = 0.48$. Right: Probability distribution of combined $L_T$ dataset, and the log-normal with the same log-mean and log-standard deviation overlaid.*

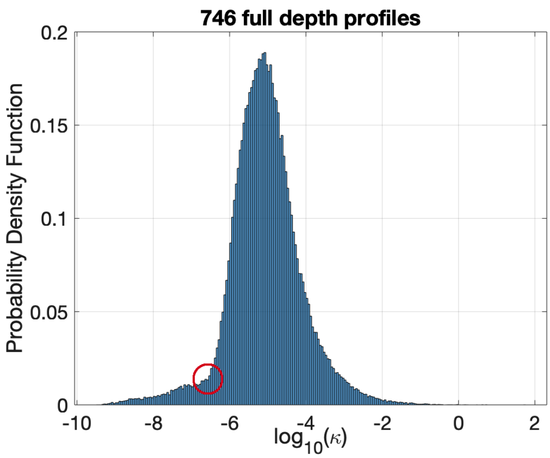


***Figure 2.*** *Probability density function of diffusivity $\kappa$ for the combined dataset described in the text based on the work of Cael & Mashayek 2021 Cael and Mashayek [2021]. Circle highlights inflection point in cumulative distribution at $10^{-6.5}$ from which we take a background diffusivity $\kappa_{background} \sim 10^{-6.5}$, such that mixing for each patch becomes $M_{background} + M_{patch} = \Gamma_{total}\epsilon = \kappa_{background}N^2 + \Gamma_{param}\epsilon$ and so $\Gamma_{total} = \kappa_{background}N^2 + \Gamma_{param}$, where $\Gamma_{param}$ is the $R_{OT}$-parameterized $\Gamma$ value.*

## References


M. H. Alford, J. B. Girton, G. Voet, G. S. Carter, J. B. Mickett, and J. M. Klymak. Turbulent mixing and hydraulic control of abyssal water in the Samoan Passage. *Geophysical Research Letters*, 40(17): 4668–4674, 2013.

B. Cael and A. Mashayek. Log-skew-normality of ocean turbulence. *Physical Review Letters*, 126(22): 224502, 2021.

G. S. Carter, G. Voet, M. H. Alford, J. B. Girton, J. B. Mickett, J. M. Klymak, L. J. Pratt, K. A. Pearson-Potts, J. M. Cusack, and S. Tan. A spatial geography of abyssal turbulent mixing in the samoan passage. Technical Report 4, 2019.

A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

L. Clément, A. M. Thurnherr, and L. C. St. Laurent. Turbulent mixing in a deep fracture zone on the Mid-Atlantic Ridge. *Journal of Physical Oceanography*, 47(8):1873–1896, 2017. ISSN 15200485. http://dx.doi.org/10.1175/JPO-D-16-0264.1.

T. Ijichi and T. Hibiya. Observed variations in turbulent mixing efficiency in the deep ocean. *Journal of Physical Oceanography*, 48(8):1815–1830, 2018. ISSN 15200485. http://dx.doi.org/10.1175/JPO-D-17-0275.1. URL www.ametsoc.org/PUBSReuseLicenses.

T. Ijichi, L. St. Laurent, K. L. Polzin, and J. M. Toole. How Variable Is Mixing Efficiency in the Abyss? *Geophysical Research Letters*, 47(7):1–9, 2020. ISSN 19448007. http://dx.doi.org/10.1029/2019GL086813.

R.-C. Lien, D. R. Caldwell, M. C. Gregg, and J. N. Moum. Turbulence variability at the equator in the central {P}acific at the beginning of the 1991–1993 {E}l {N}ino. *J. Geophys. Res.: Oceans*, 100 (C4):6881–6898, 1995.

J. N. Moum. Efficiency of mixing in the main thermocline. *J. Geophys. Res.*, 101(C5):12–57, 1996.

K. L. Polzin, J. M. Toole, J. R. Ledwell, and R. W. Schmitt. Spatial variability of turbulent mixing in the abyssal ocean. *Science*, 276(5309):93–96, 1997.

W. Smyth, J. Nash, and J. Moum. Self-organized criticality in geophysical turbulence. *Scientific reports*, 9(1):1–8, 2019.

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


**Cite this article:** Mashayek *et al*. (2021). Supplementary Materials
Shear coincidence: implications of the statistics of ocean turbulence microphysics for global diapycnal mixing. *Flow*, *1*, E1. doi:10.1017/flo.2021.1