Tracer-based estimates of eddy-induced diffusivities

Michael Haigh *, Luolin Sun, Igor Shevchenko, Pavel Berloff

Department of Mathematics, Imperial College London, London, United Kingdom

ARTICLE INFO

Keywords:
Tracer diffusivities
Double-gyre basin
Quasigeostrophic flows

ABSTRACT

This study provides estimates of the mean eddy-induced diffusivities of passive tracers in a three-layer, double-gyre quasigeostrophic (QG) simulation. A key aspect of this study is the use of a spatial filter to separate the flow and tracer fields into small-scale and large-scale components, and we compare results with those obtained using Reynolds temporal averaging. The eddy tracer flux is related to a rank-2 diffusivity tensor via the flux-gradient relation, which is solved for a pair of tracers with misaligned large-scale gradients. We concentrate on the symmetric part of the resulting diffusivity tensor which represents irreversible mixing processes. The eigenvalues of the symmetric tensor exhibit complicated behaviour, but a particularly dominant and robust feature is the positive/negative eigenvalue pairs, which physically represent filamentation of the tracer concentration. The large off-diagonal diffusivity tensor component is the primary contributor to the eigenvalue polarity, and since this is such a prevalent feature we argue that the (horizontal) eddy-induced diffusivity should always be treated as a full $2 \times 2$ tensor. Diffusivity magnitudes are largest in the upper layer and in the eastward jet region, where the eddying flow is strongest. After removing the rotational part of the eddy tracer flux, typical mean diffusivities (eigenvalues) in the upper-layer are on the order of $10^3$ m$^2$s$^{-1}$ in the jet region and $10^2$ m$^2$s$^{-1}$ elsewhere. We also confirm that the time-mean of the diffusivity calculated from instantaneous fluxes is not the same as the diffusivity associated with the time-mean fluxes.

1. Introduction

Mesoscale oceanic eddies play an important role in distributing tracers about the ocean, but resolving such transport in ocean circulation models is often unfeasible, especially in simulations that require long time integrations. This issue is likely to persist for the foreseeable future, and therefore alternative methods for representing eddy-induced tracer transport in ocean models are required. Due to its simplicity, a parameterisation of eddy tracer fluxes is often based upon the flux-gradient relation,

$$\mathbf{f} = -K \nabla C$$

In this case the eddy tracer flux $\mathbf{f}$ is related to the gradients of the large-scale tracer field $C$ via a diffusivity $K$. This study focuses on the quantification of bulk diffusivities diagnosed using (1), in a three-layer double-gyre quasigeostrophic simulation.

In ocean circulation models, the diffusivity is often assumed to be constant in both space and time, representing isotropic and homogeneous diffusion. However, there is strong evidence to suggest that neither assumption is valid [Berloff et al., 2002; Kamenkovich et al., 2015; Rypina et al., 2012]. In particular, there is growing awareness that mixing in the presence of zonal mean flows is highly anisotropic with zonal diffusivities potentially an order of magnitude larger than meridional ones [Kamenkovich et al., 2009]. In the case of eastward jets, theories and experiments predict suppression of across-stream mixing on the jet core and enhanced mixing on the jet flanks [Abernathey et al., 2010; Ferrari and Nikurashin, 2010; Klocker and Abernathey, 2014]. The former is due to wave propagation against the mean flow, whereas the latter is associated with the presence of critical layers. Overall, it is clear that diffusion of tracers is both anisotropic and inhomogeneous, and by seeking a diffusivity tensor $K$ the present study is able to provide further evidence and quantifications of this.

There is much debate surrounding the treatment of the eddy tracer flux $\mathbf{f}$ in the context of the flux-gradient relation (1). Since the eddy tracer flux sits inside the divergence operator in the tracer evolution equation, some studies [Eden et al., 2007; Eden and Greatbatch, 2009; Eden, 2010] argue that the rotational part should be removed before relating it to the large-scale tracer gradient. Moreover, the rotational part of the flux often dominates the divergent part [Marshall and Shutts, 1981] and therefore leads to diffusivities whose dynamically active part...
is obscured by a dominant inactive part. Other studies have noted that removal of the rotational flux may limit negative diffusivities [Bachman et al., 2015], but there are issues with regards to the non-uniqueness of the Helmholtz decomposition [Fox-Kemper et al., 2003]. We will present bulk diffusivities calculated using both divergent and full eddy tracer fluxes.

A key aspect of this study is the use of a spatial filter to separate the small- and large-scale flow and tracer fields [Lu et al., 2016]. The main benefit of this approach is that it allows us to relate the local eddy tracer flux to the local large-scale tracer gradient. Other studies [Bachman and Fox-Kemper, 2013; Bachman et al., 2015, 2017; Eden et al., 2007; Eden and Greatbatch, 2009; Eden, 2010; Medvedev and Greatbatch, 2004] use a Reynolds time-mean or zonal-mean to separate the scales, thus losing this locality benefit in either a spatial or a temporal coordinate. Moreover, Reynolds averaging leads to a reduced-dimension diffusivity, but we argue that a realistic diffusivity for an isopycnal model should have full two-dimensional spatial dependence and temporal dependence.

This paper is organised as follows. In section 2 we describe the three-layer quasigeostrophic model, the tracer model, and the method for obtaining eddy-induced tracer diffusivities. In section 3 we discuss typical spatial patterns of the diffusivities and we present bulk diffusivity values for a range of experiments. Lastly, we discuss and compare our results with the findings of other studies in section 4.

2. Methodology

2.1. Quasigeostrophic ocean model

We use a three-layer quasigeostrophic (QG) model to simulate mid-latitude, double-gyre dynamics. In each layer the quasigeostrophic potential vorticity (QGPV) equation is

\[
\frac{\partial \psi_k}{\partial t} + J(\psi_k, q_k) + \beta \frac{\partial \psi_k}{\partial x} = \nu \nabla^2 \psi_k + \frac{\delta_{i\beta}}{\rho_i H_i} W_k,
\]

where \( k = 1, 2, 3 \) denotes the layer index. Here \( \delta_i \) denotes the Kronecker delta, such that the wind forcing \( W \) is only active in the top layer and bottom friction, governed by \( r = 4 \times 10^{-8} \text{ s}^{-1} \), is only active in the bottom layer. Also, \( J(\psi, \cdot) \) is the Jacobian operator; \( \beta = 2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1} \) is the planetary vorticity gradient; \( \nu = 20 \text{ m}^2 \text{ s}^{-1} \) is the eddy viscosity; \( \rho_1 = 10^3 \text{ kg m}^{-3} \) is the upper layer density. The potential vorticity anomalies \( \psi_k \) are related to the streamfunctions \( \psi_k \) via

\[
\begin{align*}
q_1 &= \nabla^2 \psi_1 + s_1 (\psi_2 - \psi_1), \\
q_2 &= \nabla^2 \psi_2 + s_2 (\psi_3 - \psi_2) + s_{21} (\psi_1 - \psi_2), \\
q_3 &= \nabla^2 \psi_3 + s_3 (\psi_2 - \psi_3),
\end{align*}
\]

where \( s_1, s_{21}, s_{22} \) and \( s_3 \) are the stratification parameters. These are selected such that the first and second deformation radii are 40 km and 20.6 km, respectively. The square basin has side length \( L = 3840 \text{ km} \) such that \( 0 < x, y < L \), and the layer depths are \( H_1 = 250 \text{ m}, H_2 = 750 \text{ m} \) and \( H_3 = 3 \text{ km} \). The asymmetric tilted wind forcing is defined by

\[
W(x, y) = A \begin{cases} 
\frac{\pi A}{2L} \sin \left( \frac{\pi y}{y_0} \right) & \text{for } 0 \leq y < y_0, \\
\frac{\pi A}{2L} \sin \left( \frac{\pi (y - y_0)}{1 - y_0/L} \right) & \text{for } y_0 \leq y < L,
\end{cases}
\]

where

\[ y_0 = \frac{L}{2} + B \left( x - \frac{L}{2} \right). \]

The wind stress amplitude is \( \tau = 0.8 \text{ N m}^{-1} \), the asymmetry parameter is \( A = 0.9 \) and the wind tilt parameter is \( B = 0.2 \).

The QGPV equations are simulated using the CABARET scheme [Karabasov et al., 2009] on a uniform 1025 grid, corresponding to a grid resolution of 3.75 km. On the lateral boundaries we use partial-slip conditions given by

\[
d \frac{\partial \psi_k}{\partial t} + \frac{\partial \psi_k}{\partial n} = 0,
\]

where \( n = 120 \text{ km} \) is a boundary sub-layer lengthscale and \( n \) represents the coordinate normal to the boundary. The model is span up from rest for 20 years before tracers are initialised.

2.2. Method for estimating diffusivities

After spin up of the QG model, tracers are initialised whose dynamics are governed by the advection-diffusion equation,

\[
\frac{\partial C}{\partial t} + \nabla \cdot (uC) = \nu \nabla^2 C + F.
\]

Here \( C \) represents the tracer concentration and \( u = (u, v) = (-\partial \psi / \partial y, \partial \psi / \partial x) \) is the horizontal flow vector. The forcing \( F \) represents relaxation of the large-scale tracer field back to its initial profile, with a relaxation timescale of 5 days, the motivation for which will become clear shortly. In this case the large-scale field is defined using a square spatial filter, as will be outlined in section 2.4.

Since we are interested in eddy-induced tracer transport, we split the flow and tracer fields into large-scale (denoted by \( \ast \)) and small-scale (denoted by \( \cdot \)) components, such that, for example, \( C(x, y, t) = \overline{C}(x, y, t) + \sigma(x, y, t) \). As discussed above, this scale separation will be defined using a spatial filter. Use of a spatial filter represents a significant novelty of this study - results using this filter will be compared with results obtained using standard Reynolds temporal averaging. After separating the scales, we can rewrite the tracer equation as

\[
\frac{\partial C}{\partial t} + \nabla \cdot ((\overline{C} \ast u) + \nabla \cdot f) = \nu \nabla^2 C + F + \overline{C} \ast f.
\]

where \( f = \nabla \cdot \overline{u} \ast C \) is the eddy tracer flux. This represents fluxes that would not be resolved in a coarse-resolution or reduced-dimension model, and is often parameterised by invoking the flux-gradient relation which relates the eddy tracer flux to the large-scale tracer gradient:

\[
f = -K \nabla \overline{C}.
\]

For a \( 2 \times 2 \) tensor diffusivity tensor \( K \), the flux-gradient relation (11) is under-determined. We solve this issue by using a pair of tracers \( C_1 \) and \( C_2 \) with corresponding eddy fluxes \( f_1 = (f_1^x, f_1^y) \) and \( f_2 = (f_2^x, f_2^y) \). It can then be shown that the diffusivity is given by

\[
K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} f_1^x f_2^y - f_1^y f_2^x \\ \tau_{1,2} & \tau_{1,3} & \tau_{1,4} \end{pmatrix} \begin{pmatrix} -\tau_{2,1} & \tau_{2,3} \end{pmatrix},
\]

where \( d = \tau_{1,2} \tau_{2,4} - \tau_{1,4} \tau_{2,2} \) is the determinant of the matrix of large-scale tracer gradients, and where subscripts \( x, y \) denote zonal and meridional derivatives, respectively. Our method for computing the diffusivity motivates the relaxation of the large-scale tracer field back to its initial profile, thus avoiding the singularity at \( d = 0 \) associated with alignment of the large-scale tracer gradients. We opted for a relaxation timescale of 5 days as it is sufficient to avoid the matrix singularity and simultaneously have only a weak effect on the tracer dynamics. In section 3 we discuss the dependence that the results have on the relaxation timescale. We stress that the eddying tracer field is not relaxed. As an alternative to the solution (12), Bachman et al. [2015] suggests over-determining (11) by including more tracers and solving using the least-squares method, but we argue that consideration of tracer pairs is sufficient since our results are not notably sensitive to the specific pair


2.3. Diffusivity tensor properties

We split the diffusivity into its symmetric and antisymmetric parts:

\[ S_{ij} = \frac{1}{2} \left( \nu \left( K_{ij} + K_{ji} \right) \right) \quad \text{and} \quad A_{ij} = \frac{1}{2} \left( K_{ij} - K_{ji} \right), \]

where \( i, j = 1 \) or 2. The antisymmetric matrix \( A \) represents advection of the large-scale tracer field by a non-divergent velocity \( \mathbf{u}' = \left( -\partial A_{12} / \partial y, \partial A_{21} / \partial x \right) \) [Griffies, 1998; Plumb and Mahlman, 1987]. Our focus is the symmetric matrix \( S \), which represents irreversible diffusion. This can be diagonalised through rotation by the angle

\[ \alpha = \frac{1}{2} \tan^{-1} \left( \frac{2S_{12}}{S_{11} - S_{22}} \right). \]

The diagonalised matrix is then

\[ S' = R^T S R = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \]

\[ R = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \]

is the rotation matrix. The components of \( S' \) are the eigenvalues of \( S \) and are defined as

\[ \lambda_1 = S_{11} \cos^2 \theta + S_{22} \sin^2 \theta + 2S_{12} \cos \theta \sin \theta, \]

\[ \lambda_2 = S_{11} \sin^2 \theta + S_{22} \cos^2 \theta - 2S_{12} \cos \theta \sin \theta. \]

We can impose that \( \lambda_1 > \lambda_2 \) by selecting the appropriate quadrant when calculating \( \alpha \). It can be shown that this works by considering the standard eigenvalue definition for the matrix \( S \).

The eigenvalues \( \lambda_1 \) and \( \lambda_2 \) represent diffusivities in the direction of their respective eigenvectors, which are orthogonal. If we define the major axis of diffusion to be the first eigenvector, \( \mathbf{e}_1 = \pm \left( \cos \theta, \sin \theta \right) \), then \( \lambda_1 \) quantifies diffusion in the direction of \( \mathbf{e}_1 \) and \( \lambda_2 \) quantifies diffusion normal to \( \mathbf{e}_1 \). It is the aim of this study to provide bulk estimates of \( \lambda_1 \) and \( \lambda_2 \), as well as highlight their broad spatial patterns. Deep physical analysis of the diffusivities and their orientation is left for a later study.

2.4. Description of experiments

A key aspect of this study is the use of a square spatial filter used for decomposing the flow and tracer fields. With this method we are able to relate the local eddy tracer flux to the local large-scale tracer gradient. Standard Reynolds averaging is unable to maintain this locality. Using the filtering method for a discrete field \( \phi_n \) and odd filter width \( w \), the large-scale field is defined as

\[ \varphi_x = \frac{1}{w} \sum_{n=-l}^{l} \sum_{m=-l}^{l} \phi_{nm}, \]

where \( l = \frac{w - 1}{2} \).

Near the boundaries, the width \( w \) is reduced to the appropriate size, and the average is evaluated over a smaller square range. The small-scale field is defined as \( \psi_y = \varphi_y - \varphi_y \). This method could be built upon by additionally filtering the data in time, by using alternative filter shapes (e.g., a circle) or by including a weighting function in equation (18). However, we opted to use the square spatial filter due to its simplicity and because only one parameter is required. Results using this filter will be compared with results obtained using the standard Reynolds time-mean.

We will present diffusivities from four experiments. Three experiments, namely F15, F31 and F61, use the spatial filter defined in (18) to separate the scales; here 15, 31 and 61 refer to the filter widths in grid point dimensions and correspond to physical filter widths of 52.5 km, 112.5 km and 225 km, respectively. In each of these three experiments we treat the eddy tracer flux in four distinct ways: (1) the flux \( f \) is untreated; (2) only the divergent part of \( f \) is retained; (3) only the eddy- eddy flux, \( \mathbf{u}' \cdot \mathbf{c}' \), is retained; (4) only the divergent part of the eddy eddy flux is retained. For short, we refer to these sub-experiments as \( f, f_{\text{div}}, \mathbf{u}' \cdot \mathbf{c}' \) and \( (\mathbf{u}' \cdot \mathbf{c}')_{\text{div}} \), respectively. In the fourth experiment, referred to as REYNOLDS, large-scale fields are defined using Reynolds temporal averaging, and small-scale fields defined as the deviation. In REYNOLDS, the time-mean flux-gradient relationship is solved rather than the instantaneous one, in which case we need only consider fluxes \( f \) and \( f_{\text{div}} \) since these are equivalent to \( \mathbf{u}' \cdot \mathbf{c}' \) and \( (\mathbf{u}' \cdot \mathbf{c}')_{\text{div}} \), respectively. We consider the divergent part of the eddy tracer flux since the rotational part does not influence the tracer dynamics. Furthermore, the rotational part of the eddy flux typically dominates the divergent part [Marshall and Shutts, 1981], and can consequently dominate diagnosed diffusivities [Eden et al., 2007]. We use the Helmholtz decomposition, with zero normal flow and zero tangential flow boundary conditions, to calculate divergent fluxes.

Each tracer is initialised with a linear profile,

\[ C_0 = \frac{ax + by + cz}{\sqrt{x^2 + y^2 + z^2}} \]

and its large-scale component is relaxed back towards its initial profile with a relaxation timescale of 5 days. Note that in each experiment, the same smaller filter width (52.5 km) is used when relaxing the large-scale tracer field. We do this for two reasons. First, the filtering process is time-consuming and the demands grow with the filter size. Second, this allows us to run all experiments on data attained from a single simulation, rather than from four different ones which may diverge from one another. All results presented in this study use the same pair of tracers; for the first tracer field we use \( (a, b, c) = (1, -4, 4) \) and for the second we use \( (a, b, c) = (-2, 3, 2) \). To test the robustness of our results, we additionally simulated two more tracers (leading to 6 pairs of tracers in total), and find that diffusivities are not notably sensitive to the specific pair used. Furthermore, for divergent fluxes, the diffusivity tensors \( \mathbf{K} \) calculated using different tracer pairs are indistinguishable from one another. This tracer-independence is due to the linearity of the large-scale tracer profiles. A thorough analysis of the uniqueness of \( \mathbf{K} \) will be presented in a later study.

3. Results

After initialisation of the tracer field, a further \( \sim 25 \) days are required before the eddy tracer field becomes statistically steady. The QG dynamics and tracer equations are then simulated for a further year. In this section we start by discussing the essential qualitative behaviour of the diffusivity eigenvalues obtained from these simulations. We then move onto presenting time-mean and spatial-mean diffusivities.

3.1. Qualitative behaviour

In Fig. 1 we present instantaneous \( \lambda_1 \) and \( \lambda_2 \) from experiment F31 for the flux \( f_{\text{div}} \). Note the dashed rectangle which outlines the jet region, inside which we calculate separate statistics which will be presented shortly. Across almost the entire domain the first/second eigenvalue is positive/negative such that they represent filamentation of the tracer concentration [Ledwell et al., 1998]. Such polarity of the diffusivity eigenvalues would have an important and yet unknown effect on tracer clustering at the surface [Klyatskin and Koshe, 2017]. In each layer the diffusivities are largest in the jet region where the eddy flow is strongest. Diffusivities calculated using the full eddy tracer flux \( f \) are typically two orders of magnitude larger than those for the divergent

\[ ^1 \text{We use } S \text{ to denote the symmetric part, and } A \text{ for the antisymmetric part. Their components are denoted } S_{ij} \text{ and } A_{ij}. \]
flux, but the broad qualitative behaviour of $\lambda_1$ and $\lambda_2$ (but not $\alpha$) is unchanged. In experiments F15 and F61 we observe qualitative behaviour similar to Fig. 1, and we conclude that the positive/negative diffusivity pair is a robust feature of our results. This highlights possible limitations of scalar, homogeneous or isotropic diffusivity closures, which are unable to encapsulate the nontrivial behaviour that we observe.

The polarity of diffusivity eigenvalues is rarely observed in previous studies [Bachman et al., 2020; Eden and Greatbatch, 2009], but there are simple reasons that can explain why this is the case. First, Lagrangian methods (e.g., Rypina et al., [2012]; Ying et al., [2019]; Zhurbas and Oh [2004]; Zhurbas et al., [2014]; Rypina et al., [2012]; Ying et al., [2019]; Zhurbas and Oh [2004]; Zhurbas et al., [2014]) lead to a symmetric diffusivity tensor which, by construction, has non-negative eigenvalues. Second, in studies which use Eulerian methods (e.g., Abernathey and Marshall [2013]; Eden et al., [2007]; Marshall et al., [2006]; Abernathey and Marshall [2013]; Eden et al., [2007]; Marshall et al., [2006]), the diffusivity is rarely treated as a full $2 \times 2$ (in the case of 2 dimensions) tensor, but rather as a diagonal tensor or a scalar. In these cases, the off-diagonal diffusivity component $S_{12}$ is neglected, which consequently reduces the likelihood of observing polar eigenvalues. This is because the term $2S_{12}\cos\alpha \sin\alpha$ is always positive, and therefore $S_{12}$ always makes a positive/negative contribution to the first/second eigenvalue (see equations 16 and 17). More precisely, the diffusivity eigenvalues are of opposite sign if $S_{12} > S_{11}S_{22}$. Thus, polar eigenvalues are obtained if $S_{12}$ is sufficiently large, or if $S_{11}$ and $S_{22}$ are of opposite sign. As an example, for the eigenvalue snapshots shown in Fig. 1, we calculated the frequency with which these conditions are satisfied (in the upper layer). At (approximately) 60% of grid points $S_{12} > |S_{11}S_{22}|$, at 20% of grid points $S_{11}S_{22} < 0$, and at 72% of grid points $S_{12} > S_{11}S_{22}$. This suggests that the eigenvalue polarity is predominantly due to the large off-diagonal tensor component $S_{12}$.

Fig. 1 also depicts the local major axis of diffusion, i.e., the unit eigenvector of $\lambda_1$, in order to motivate the idea that the behaviour of the diffusivity orientation is nontrivial. The major axis is plotted every 30 grid points. The dashed rectangle outlines the jet region.
diffusion orientation, but observations indicate that in the jet region the major axis of diffusion is not necessarily aligned with the mean flow. This perhaps contradicts results from previous studies (e.g., Abernathey et al., [2013]; Kamenkovich et al., [2015]; Rypina et al., [2012]; Abernathey et al., [2013]; Kamenkovich et al., [2015]; Rypina et al., [2012]), so it is therefore unclear to what extent our results agree with other studies of cross-jet mixing [Abernathey et al., 2010; Klocker and Abernathey, 2014; Ferrari and Nikurashin, 2010]. However, previous studies of the diffusion orientation make use of averaged tracer fluxes or ensemble-averaged particle trajectories, and these approaches can eliminate variability in the major axis of diffusivity. Thus, our results may not disagree with those of previous studies, but further analysis is required and this will be presented in a later study.

Before considering mean diffusivities, we briefly relate our results to classic eddy diffusivity theory [Prandtl, 1925; Taylor, 1921; Vallis, 2017], in which we have $K \sim \sigma^2 \tau$ where $\sigma^2$ is the eddy velocity variance (the time-mean eddy kinetic energy) and $\tau$ is the Eulerian decorrelation timescale. We may use this to estimate the extent to which spatial variability in the eigenvalues $\lambda_1$ and $\lambda_2$ is due to variations in the eddy intensity. In Fig. 2 we plot the normalised eigenvalues, i.e.,

$$\tilde{\lambda}_1 = \frac{\lambda_1}{\langle \sigma \rangle^2} \quad \text{and} \quad \tilde{\lambda}_2 = \frac{\lambda_2}{\langle \sigma \rangle^2}$$

(20)

in the upper layer. Firstly, the normalised diffusivities are largest away from the jet, suggesting that the decorrelation timescale $\tau$ is short inside the jet region relative to other regions in the domain.\(^2\) The effect of normalising the diffusivities in the lower layers is much the same as in the upper layer. Also, in the lower layers $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are of the same order as in the upper layer, so although normalisation by $\sigma^2$ eliminates the variability between layers, some layer-wise variability persists.

3.2. Quantitative behaviour

The remainder of this study focusses on the bulk diffusivities. To quantify the bulk diffusivity we use two measures: the mean diffusivity,

$$\langle \lambda \rangle = \frac{\langle \lambda_1 + \lambda_2 \rangle}{2}$$

(21)

and the mean absolute diffusivity,

$$\langle |\lambda| \rangle = \frac{\langle |\lambda_1| + |\lambda_2| \rangle}{2}$$

(22)

where the angular brackets denote a year-long time-mean and a spatial mean, which is either a domain mean or a jet-region mean (see Fig. 1). The polarity of the eigenvalues means that we expect distinct estimates for $\langle \lambda \rangle$ and $\langle |\lambda| \rangle$, and their disparity could be a measure of such polarity. Before calculating the bulk diffusivities, any values that lie more than three standard deviations away from the mean are capped. This prevents spurious values - which are due to alignment of large-scale tracer gradients - from dominating the mean.

Table 1 presents bulk diffusivities from experiment F31. Bulk diffusivities are strongest in the upper layer, weakest in the lower layer, and span many orders of magnitude (1 - $10^8$ m$^2$ s$^{-1}$). The mean absolute diffusivity $\langle |\lambda| \rangle$ is consistently larger - typically by one or two orders of magnitude - than the mean diffusivity $\langle \lambda \rangle$. This disparity is due to the polarity of the eigenvalues, and therefore extreme $\langle |\lambda| \rangle$ values ought to be interpreted with this in mind. Diffusivities are consistently largest in the jet region, especially in the upper layer where the flow is fastest. Negative domain-mean diffusivities are obtained for the full flux $f$ in the lower layers, which is due to large negative trace values $\langle \lambda_1 + \lambda_2 \rangle$ outside of the jet region. Although such locally negative trace values exist in upper layer - and in the lower layers for other fluxes - these are less negative, and so we obtain positive domain means. We remind the reader that the bulk values, although not in general tracer-independent, vary very little when using alternative tracer pairs (we tested 6 tracer pairs in total). Furthermore, for divergent eddy fluxes the bulk diffusivities are the same for different tracer pairs, which is due to the linearity of the large-scale tracer gradients.

The qualitative behaviour of each diffusivity as shown in Fig. 1 for experiment F31 is preserved in experiments F15, F61. In Tables 2 and 3 we present mean diffusivities for experiments F15 and F61, respectively. For experiment F15, mean diffusivities are typically half those in experiment F31. This reduction may be expected since for the smaller filter less of the eddying tracer and flow fields are captured by $c$ and $u$, and instead are captured by $\mathcal{C}$ and $\mathbf{u}$. However, the larger filter width in experiment F61 does not lead to diffusivities that are consistently larger than those in experiment F31. In experiments F15 and F61, the negative $\langle |\lambda| \rangle$ values persist for the full flux $f$, and again we note that the equivalent bulk diffusivities are positive for the divergent flux $f_{\text{div}}$. Thus, our results agree with the suggestion by Bachman and Fox-Kemper [2013] that negative diffusivities could be due to contamination by the dominant rotational component of the tracer flux (provided that we use the full divergent eddy tracer flux rather than just the eddy-eddy component).

For experiment REYNOLDS the flux-gradient relation (11) is solved only once per sub-experiment, for the time-mean eddy tracer flux. As shown in Table 4, mean diffusivities are most commonly between $10^2$ and $10^3$ m$^2$ s$^{-1}$ with weakest values at depth. In layers 2 and 3 for the full flux $f$, both $\lambda_1$ and $\lambda_2$ are almost globally positive, leading to near-identical $\langle \lambda \rangle$ and $\langle |\lambda| \rangle$ estimates. For the divergent flux $f_{\text{div}}$, mean diffusivities are mostly drastically reduced outside of the jet region, as negative $\lambda_2$ values become more abundant, leading to small domain-mean diffusivities which are negative in the lower layers. Therefore, the effect of removing the rotational part from the eddy tracer flux is not the same as in experiments F15, F31 and F61. We also note that in experiment REYNOLDS, the weakening of the diffusivities with depth is not as pronounced as in the other experiments. Overall, an important - and perhaps expected - conclusion to be made from experiment REYNOLDS is that the time-mean diffusivity (e.g., from experiments F15, F31, F61) is not the same as the diffusivity associated with the time-mean fluxes.

Recall that in order to avoid the singularity in solving for the diffusivity tensor $K$, the large-scale tracer concentrations are relaxed back towards their initial profiles with a relaxation timescale of 5 days. The results we present are dependent on the relaxation rate with stronger/weaker relaxations leading to smaller/larger diffusivity estimates, but we stress that this sensitivity is not spurious. For vanishingly weak relaxations, we hypothesise that the diffusivity amplitudes would asymptote, provided the simulation time is short enough such that the tracer gradients remain misaligned.

4. Conclusions and discussion

In this study we have provided estimates of diffusivities of passive tracers in a mid-latitude, double-gyre ocean model. Tracers were advected in a high-resolution, three-layer quasigeostrophic simulation, after which the eddy tracer flux was related to its large-scale gradient via a diffusivity tensor. To separate the flow and tracer field scales, a spatial filter was used and results were compared to the output obtained using temporal Reynolds averaging. Importantly, via this comparison we showed that the time-mean diffusivity, which is calculated from instantaneous fluxes, is not the same as the diffusivity associated with the time-mean fluxes.

We concentrated on the two eigenvalues of the symmetric part of the diffusivity tensor, which represent irreversible diffusion along their orthogonal eigenvector directions. First of all, the diffusivity eigenvalues exhibit complicated spatial patterns and can span many orders of

\(^2\) Note in particular the very large negative $\lambda_2$ in the south-eastern corner of the domain, where eddy activity is consistently weak.
magnitude. A robust feature, common to every layer, is the positive/negative eigenvalue pairs, which quantify filamentation of the tracer concentration [Ledwell et al., 1998]. There is strong evidence to suggest that this eigenvalue polarity is predominantly due to the large off-diagonal tensor component $S_{12}$. In other studies which use Eulerian methods (e.g., Abernathey and Marshall [2013]; Eden et al., [2007]; Marshall et al., [2006]), the diffusivity is rarely treated as a $2 \times 2$ tensor, but rather as a diagonal tensor or a scalar. As a result, such studies neglect the off-diagonal mixing terms and inevitably suppress eigenvalue polarity in their results. Moreover, although Lagrangian estimates (e.g., Kamenkovich et al., [2015], Klocker et al., [2012a, b], LaCasce et al., [2014], Rypina et al., [2012], Ying et al., [2019], Zhurbas and Oh [2004], Zhurbas et al., [2014]) may reliably predict diffusivity amplitudes, they do not capture the spatial complexity and eigenvalue polarity that we observe. Overall, we argue that studies of eddy-induced transport ought to consider the full $2 \times 2$ diffusivity tensor and consequently not restrict the possibility of positive/negative eigenvalue pairs.

Bulk diffusivities were defined to be the mean of the eigenvalues (or the mean of their absolute values), additionally averaged in space and time:

$$\lambda_1 \quad \lambda_2$$

Table 1
Experiment F31. Bulk diffusivities (units m$^2$ s$^{-1}$) in each of the three layers. We give the mean eigenvalue, $\langle \lambda \rangle$, and the mean absolute eigenvalue, $\langle |\lambda| \rangle$, for domain averages and jet/region averages, and for four tracer fluxes.

<table>
<thead>
<tr>
<th>Layer 1</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>6528</td>
<td>202</td>
<td>791</td>
<td>−1.0</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>46550</td>
<td>1642</td>
<td>3678</td>
<td>−56</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>145929</td>
<td>3488</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>633719</td>
<td>17721</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 2</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−317</td>
<td>17</td>
<td>116</td>
<td>5.0</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>717</td>
<td>323</td>
<td>248</td>
<td>39</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>23968</td>
<td>431</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>32628</td>
<td>1066</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 3</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−394</td>
<td>2.0</td>
<td>34</td>
<td>0.7</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>135</td>
<td>87</td>
<td>57</td>
<td>11</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>12264</td>
<td>227</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>16095</td>
<td>356</td>
</tr>
</tbody>
</table>

Table 2
The same as Table 1 but for experiment F15.

<table>
<thead>
<tr>
<th>Layer 1</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>5657</td>
<td>132</td>
<td>408</td>
<td>21</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>31838</td>
<td>764</td>
<td>2077</td>
<td>111</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>101639</td>
<td>2814</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>429123</td>
<td>13522</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 2</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−215</td>
<td>6.0</td>
<td>22</td>
<td>1.1</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>180</td>
<td>79</td>
<td>29</td>
<td>3.1</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>9028</td>
<td>193</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>11322</td>
<td>543</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 3</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−114</td>
<td>0.9</td>
<td>5.3</td>
<td>0.2</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>−6.0</td>
<td>23</td>
<td>8.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>3715</td>
<td>70</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>4998</td>
<td>116</td>
</tr>
</tbody>
</table>

Table 3
The same as Table 1 but for experiment F61.

<table>
<thead>
<tr>
<th>Layer 1</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>662</td>
<td>156</td>
<td>2236</td>
<td>82</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>15709</td>
<td>2632</td>
<td>10843</td>
<td>1025</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>124565</td>
<td>2954</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>337018</td>
<td>9364</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 2</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−523</td>
<td>24</td>
<td>414</td>
<td>11</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>1922</td>
<td>797</td>
<td>1213</td>
<td>263</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>49513</td>
<td>987</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>79686</td>
<td>2339</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 3</th>
<th>f</th>
<th>$f_{\text{div}}$</th>
<th>$u'c'$</th>
<th>$(u'c')_{\text{div}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \lambda \rangle$, domain</td>
<td>−753</td>
<td>2.6</td>
<td>159</td>
<td>1.4</td>
</tr>
<tr>
<td>$\langle \lambda \rangle$, jet</td>
<td>986</td>
<td>235</td>
<td>307</td>
<td>76</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, domain</td>
<td>30935</td>
<td>645</td>
</tr>
<tr>
<td>$\langle</td>
<td>\lambda</td>
<td>\rangle$, jet</td>
<td>40567</td>
<td>1023</td>
</tr>
</tbody>
</table>

Fig. 2. Instantaneous normalised diffusivities $\hat{\lambda}_1$ (left) and $\hat{\lambda}_2$ (right) in the upper-layer from experiment F31. Units are $2 \times 10^3$ s. Colour is strongly saturated in the south-eastern corner of the right-hand panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
time. The spatial mean was either a layer-wise domain-mean or a jet-region mean. We have also considered the effects of removing the rotational part from the eddy tracer flux, and the effects of removing the eddy-mean interaction terms from the tracer flux. Bulk diffusivities are systematically largest in the upper layer and in the jet region of each layer. For example, for the full eddy tracer flux \( f \) and a filter width of 112.5 km, the mean diffusivity is \( \frac{6000 \text{ m}^2 \text{s}^{-1}}{\text{in the upper-layer domain mean}} \) and \( \frac{45000 \text{ m}^2 \text{s}^{-1}}{\text{in the jet-region mean}} \). Removal of the dominant rotational part of the eddy tracer flux leads to diffusivities that are one to two orders of magnitude smaller. Since the diffusivities are of opposite sign in most of the domain, mean absolute diffusivities are commonly two orders of magnitude larger than the raw values, leading to extreme \( (\times 10^6 \text{ m}^2 \text{s}^{-1}) \) mean absolute diffusivities for the full eddy tracer flux.

Our bulk diffusivity estimates span many orders of magnitude, but they are in broad agreement with estimates from other studies. Rypina et al. (2012) used drifters and synthetic particles to estimate anisotropic diffusivities in the North Atlantic and found that the strongest mixing occurs in the Gulf Stream. The domain-mean diffusivity - averaging the contribution from both directions - was estimated to be approximately \(5000 \text{ m}^2 \text{s}^{-1} \), very close to our estimates from experiments F15 and F31 for the full eddy tracer flux. Such values are also similar to those attained by Zhurbas and Oh (2004), who considered both the Atlantic and Pacific Oceans, and similar to Zhurbas et al. (2014) who considered the entire global ocean. Eden and Greatbatch (2009) used the flux-gradient relation to diagnose diagonal diffusivity tensors of various tracers in the Atlantic Ocean. They found positive/negative diffusivity pairs in parts of the domain with diffusivity magnitudes generally between 0 and 5000 \( \text{m}^2 \text{s}^{-1} \). Our results suggest that if Eden and Greatbatch (2009) included the off-diagonal diffusivity tensor components in their method, then polar eigenvalue pairs would be a more common feature.

Many other studies have focussed on the Southern Ocean and the Antarctic Circumpolar Current (ACC), where the diffusivity tends to be weaker in comparison to a mid-latitude basin, due to the relatively small deformation radius. Similar to our approach, Lu et al., (2016) used a spatial filter to separate the flow scales in a Southern Ocean model. They found both positive and negative buoyancy diffusivities to be as large as 3000 \( \text{m}^2 \text{s}^{-1} \), but did not consider the most general case of a 2 \( \times \) 2 tensor diffusivity. Using the effective diffusivity of Nakamura (1996), Klocker et al., (2012b) found typical cross-ACC surface diffusivities between 500 and 1000 \( \text{m}^2 \text{s}^{-1} \) for a tracer advected by altimetry-derived geostrophic flow. Using the same method, Marshall et al., (2006) and Abernathey et al., (2010) calculated diffusivities of \( \approx 2000 \text{ m}^2 \text{s}^{-1} \) on the equatorward flank of the ACC, and reduced values of \( \approx 500 \text{ m}^2 \text{s}^{-1} \) in the ACC core. Such reduced diffusivities in the ACC core are predicted by the theory of Ferrari and Nikurashin (2010), by which eddies propagating against the mean-flow act to suppress mixing. This theory was corroborated by Klocker and Abernathey (2014) who explored the effects of systematically varying a zonal background flow. However, it is unclear to what extent such behaviour is exhibited in our experiments; in-depth analyses of the diffusivities and their orientation will follow in a later study.

The use of divergent over full eddy tracer fluxes is an actively discussed topic, but we can present arguments for removing the rotational part of the flux, in spite of non-uniqueness issues associated with the Helmholz decomposition (Fox-Kemper et al., 2003)\(^3\). First, use of the divergent component of the eddy tracer flux limits the influence of the dynamically inert rotational component, which dominates the full flux [Marshall and Shotts, 1981] and the resulting diffusivity. Second, our results imply that negative mean diffusivities are less likely to be obtained after removal of the rotational part of the eddy tracer flux, agreeing with the suggestion of Bachman and Fox-Kemper (2013). Third, for divergent fluxes and linear large-scale tracer gradients, the resulting diffusivity is tracer-independent. In general, however, our approach yields tracer-dependent \( K \) which Bachman et al., (2015) argued is best treated by overdetermining the flux-gradient relation (include more tracers) and solving using the least-squares approach. Overall, non-uniqueness of the diffusivity is a critical topic, and will therefore be comprehensively addressed in a later study. Furthermore, our work can be considered as foundational for systematic modelling of the eddy diffusivity tensor coefficients as random processes, following on from earlier studies [Berloff and McWilliams, 2003; Grooms, 2016].

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

Michael Haigh, Luolin Sun and Pavel Berloff are supported by the NERC grant NE/T002220/1. Igor Shevchenko and Pavel Berloff are supported by the Leverhulme grant RPG-2019-024. Pavel Berloff is additionally supported by the NERC grant NE/R011567/1 and the Royal Society grant IEC/R2/1811033.

### References


\(^3\) Maddison et al., (2015) derived a physically motivated decomposition of eddy PV uxes, but their logic may not be readily applied to passive tracers.