

On eddy transport in the ocean. Part I: The diffusion tensor

Michael Haigh ^{*1}, Luolin Sun¹, James C. McWilliams², and Pavel Berloff^{1,3}

¹*Department of Mathematics, Imperial College London, London, United Kingdom*

²*Department of Atmospheric and Oceanic Sciences, University of California Los Angeles, CA, 90095*

³*Institute of Numerical Mathematics of the Russian Academy of Sciences, 119333 Moscow, Russia*

Keywords: Eddy transport, Eddy diffusion, Tracer parameterisations.

Abstract

This study provides an interpretation of isopycnal eddy transport for mass and passive tracers in double-gyre eddy-resolving oceanic circulation. This paper focuses on a transport/diffusion tensor representation of the eddy tracer flux, and a companion paper will focus on advective eddy-induced tracer and mass transports. We use a spatial filter to separate the large and small scales, which leads to results distinct from those obtained via a temporal Reynolds eddy decomposition. To work towards a parameterisation, we relate the eddy tracer flux to the large-scale tracer gradient via the transport tensor $\mathbf{K}$. The symmetric part of $\mathbf{K}$ is the diffusion tensor, $\mathbf{S}$, which parameterises diffusive fluxes and whose mixing properties are determined by the signs of its eigenvalues. The eigenvalues of $\mathbf{S}$ are robustly of opposite sign (polar) and thus quantify filamentation of the tracer via both up- and down-gradient fluxes. Given the prevalence of polar eigenvalues – which are also obtained for Reynolds eddy fluxes – representing their associated effects should be a target of future eddy tracer transport closures. Given the inherent inhomogeneity and anisotropy of the eddy-induced transport, we argue that a full transport tensor is better suited to this task than scalar coefficients or diagonal tensors. The diffusion axis, which represents the direction of preferential mixing, tends to align with the large-scale velocity vector and contours of large-scale relative vorticity and layer thickness. Strong shears can inhibit this alignment. We show that the large-scale velocity gradient matrix may be suitable for parameterising the transport tensor, in particular at depth. Furthermore, since entries of $\mathbf{K}$ and $\mathbf{S}$ exhibit probabilistic distributions when conditioned on certain large-scale flow features, we suggest that a stochastic closure for the eddy transport would be most suitable.

*Corresponding author. Email address: m.haigh15@imperial.ac.uk. (M. Haigh)

1 Introduction

Mesoscale eddies play a leading role in the transport of tracers in the ocean, but computational limitations mean that such small-scale flows are often not dynamically resolved in numerical models. As a result the associated eddy tracer transport is also not resolved, leading to inaccurate estimates of tracer distributions in simulations, especially those with integrations over long time scales. The prevailing solution to this issue is to parameterise eddy tracer fluxes via a transport tensor which acts upon the large-scale tracer field. When diagnosing the transport tensor in eddy-resolving simulations a common method is to separate it into its symmetric and antisymmetric components – these encompass diffusive and advective fluxes, respectively. In this study we present a physical interpretation of the transport tensor with a focus on its symmetric diffusion component. The interpretation of the antisymmetric advection tensor will follow in a companion study.

An important challenge faced by oceanographers is the reconciliation of transport tensors obtained via different methods. Lagrangian methods (e.g., Berloff et al. [2002], Kamenkovich et al. [2015], Rypina et al. [2012], Ying et al. [2019], Zhurbas and Oh [2004], Zhurbas et al. [2014]) consider the dispersion of synthetic particles or real-world drifters which, by construction, leads to a symmetric tensor, whereas Eulerian methods can additionally diagnose the advection tensor. Lagrangian and Eulerian methods generally obtain similar estimates for the amplitude of transport coefficients [Abernathy et al., 2013], and they both find that the associated transport is highly anisotropic and inhomogeneous [Bachman et al., 2020, Berloff et al., 2002, Eden et al., 2007, Kamenkovich et al., 2015, Rypina et al., 2012]. A range of studies have sought to understand the transport anisotropy in zonal jets [Abernathy et al., 2010, Ferrari and Nikurashin, 2010, Klocker and Abernathy, 2014] and have found that the cross-jet eddy transport is potentially an order of magnitude smaller than the along-jet eddy transport. This suppression in cross-jet transport is due to fast flows on the jet core which cause particles/tracers to sample many eddies, leading to dispersion effects that cancel. Maxima in transport are found on the jet flanks and beneath the jet, and are associated with critical layers where the eddy phase speed equals the mean flow speed. Since in this study we consider a high-resolution double-gyre system, our results help to build upon the understanding of eddy transport in jets.

With Eulerian methods the transport tensor is defined by assuming a relationship between eddy tracer fluxes and the large-scale tracer gradient. A key step in this approach is the separation of the small-scale (eddy) and large-scale fields. Motivated by its mathematical simplicity, most studies use a Reynolds decomposition with which the large-scale fields are usually a time mean or a zonal mean (e.g., [Bachman and Fox-Kemper [2013], Bachman et al. [2015, 2017b], Eden [2010], Eden and Greatbatch [2009], Eden et al. [2007], Medvedev and Greatbatch [2004]). In this study we use a spatial filter [Bachman et al., 2017a, Fox-Kemper and Menemenlis, 2008, Lu et al., 2016, Nadiga, 2008, Stanley et al., 2020] to separate the scales which allows us to relate the local large-scale tracer gradient to the local eddy tracer fluxes, in contrast to Reynolds averaging which leads to a loss in locality in either space or time.

There is an ongoing debate regarding the treatment of the eddy tracer flux. Some authors recommend removing the dynamically inert rotational part of the eddy tracer flux [Eden, 2010, Eden and Greatbatch,

[2009], [Eden et al., 2007] whereas others argue that the full flux should be considered [Bachman and Fox-Kemper, 2013], [Bachman et al., 2020]. One argument for using the full flux is to maintain the theoretical link to the transport tensor in parcel excursion theory [Taylor, 1921]. However, since the dynamically inert rotational eddy tracer flux dominates the divergent part [Marshall and Shutts, 1981], we argue that it should be removed to eliminate the obscuring effects it has upon the dynamically active part of the transport tensor. Since the focus of the present study is on the physical interpretation of this tensor we will use the divergent flux, and our motivations will be expanded upon in the main text.

This study builds upon the quantitative analysis of [Haigh et al., 2020], hereafter HSSB20, who used the same quasigeostrophic model and spatial filter as are used in the present study. HSSB20 presented essential statistics pertaining to the eigenvalues of the symmetric diffusion tensor and found that the amplitudes were in broad agreement with other studies. More significantly, it was found that the eigenvalues are robustly arranged in opposite-signed (polar) pairs, which physically represent diffusive eddy fluxes which are down-gradient in one direction, and up-gradient in the perpendicular direction. Polar eigenvalues, also found recently by [Stanley et al., 2020], are interpreted as representing filamentation of the tracer concentration. The present study extends HSSB20 by providing a deeper physical interpretation of the eddy transport tensor with a particular focus on its symmetric component.

This study is organised as follows. In section 2 we define the quasigeostrophic ocean model, the tracer model, and outline our method for obtaining the eddy-induced tracer transport tensor. Then in section 3 we present the transport tensor and discuss its essential qualities. The focus of this study, the symmetric diffusion tensor, will be analysed in section 4. In section 5 we consider results for Reynolds eddies. Lastly, in section 6 we summarise our results and discuss some important implications.

2 The model

2.1 Quasigeostrophic ocean model

Three-layer double-gyre mid-latitude dynamics are simulated using a quasigeostrophic (QG) ocean model. The QG potential vorticity (PV) equation for each layer is

$$\frac{\partial q_k}{\partial t} + J(\psi_k, q_k) + \beta \frac{\partial \psi_k}{\partial x} = \nu \nabla^4 \psi_k - \delta_{3k} \gamma \nabla^2 \psi_k + \frac{\delta_{1k}}{\rho_1 H_1} W. \quad (1)$$

Here $k = 1, 2, 3$ is the layer index and δ_{ij} denotes the Kronecker delta, so that the wind forcing $W(x, y)$ is active only in the upper layer and bottom friction – which is governed by $\gamma = 4 \times 10^{-8} \text{ s}^{-1}$ – is active only in the bottom layer. The beta-plane planetary vorticity gradient is $\beta = 2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$, the eddy viscosity is $\nu = 20 \text{ m}^2 \text{ s}^{-1}$ and the upper-layer density is $\rho_1 = 10^3 \text{ kg m}^{-3}$. The mean thicknesses in the upper, middle and lower layers are $H_1 = 250 \text{ m}$, $H_2 = 750 \text{ m}$ and $H_3 = 3 \text{ km}$, respectively. The Jacobian operator $J(\psi_k, q_k)$ represents nonlinear advection of the QG PV anomalies, q_k , by a non-divergent velocity

with streamfunction, ψ_k . The streamfunctions and the PV are related via the elliptic equations

$$q_1 = \nabla^2 \psi_1 + s_1(\psi_2 - \psi_1), \quad (2)$$

$$q_2 = \nabla^2 \psi_2 + s_{21}(\psi_1 - \psi_2) + s_{22}(\psi_3 - \psi_2), \quad (3)$$

$$q_3 = \nabla^2 \psi_3 + s_3(\psi_2 - \psi_3). \quad (4)$$

The stratification parameters, s_1 , s_{21} , s_{22} and s_3 , are selected such that the first and second Rossby deformation radii are 40 km and 20.6 km, respectively. The model is set up in a square domain with side length $L = 3840$ km and the zonal and meridional coordinates are $x, y \in [0, L]$.

Other model details include the asymmetric, tilted wind stress forcing which is defined in HSSB20 and is used in other QG-based studies (e.g., [Berloff 2015](#)). Also defined in HSSB20 are the partial-slip boundary conditions for which we use a boundary sub-layer lengthscale of 120 km [Berloff and McWilliams, 1999](#). To simulate the QG PV equations we use the CABARET scheme [Karabasov et al., 2009](#) with the domain uniformly discretised on a 1025^2 grid, which corresponds to a grid resolution of 3.75 km.

2.2 Eddy tracer transport

Tracer dynamics are governed by the advection-diffusion equation,

$$\frac{\partial C}{\partial t} + \nabla \cdot (\mathbf{u}C) = \nu \nabla^2 C + F. \quad (5)$$

We have omitted the layer subscript k since the same equation and following methods will be applied in each layer. In equation [\(5\)](#), C is the tracer field and $\mathbf{u} = (u, v) = (-\psi_y, \psi_x)$ is the horizontal velocity vector, where subscripts x and y represent zonal and meridional derivatives, respectively. External forcing is represented by F , which we take to be a relaxation of the large-scale tracer field back to its initial profile. The motivations for and details of this large-scale restoring force will be given later in this section.

The aim of this study is to examine the role of eddies in tracer transport, in which case it is necessary to separate the flow and tracer fields into large-scale and small-scale (eddy) components. In previous studies of eddy transport, scales are usually separated using Reynolds averaging and then the effects of the mean eddy tracer fluxes are examined. In contrast, we will use a spatial filter [Bachman et al., 2017a](#), [Fox-Kemper and Menemenlis, 2008](#), [Lu et al., 2016](#), [Nadiga, 2008](#), [Stanley et al., 2020](#) to separate the scales, leading to small-scale and large-scale components with full spatio-temporal dependence. At each grid point the large-scale component of a given field is obtained by averaging the field over the area covered by a moving square filter. That is, given a discrete field snapshot ϕ_{ij} , where i, j denote the grid point, its large-scale component is defined as

$$\bar{\phi}_{ij} = \frac{1}{w^2} \sum_{m=i-l}^{i+l} \sum_{n=i-l}^{i+l} \phi_{mn}, \quad \text{where } l = \frac{w-1}{2}. \quad (6)$$

Here w is the odd filter width, which is reduced in size near the boundaries so that the filter remains square. The small-scale field is $\phi'_{ij} = \phi_{ij} - \bar{\phi}_{ij}$. Herein we drop the subscript i, j notation. Note because we do not use a Reynolds decomposition, we in general have that $\bar{\bar{\phi}} \neq \bar{\phi}$ and $\overline{\phi'} \neq 0$.

With this scale separation we are able to relate the local (in time and space) large-scale tracer field to the local eddy tracer flux, a property which is lost with Reynolds averaging. Spatial filtering also lends itself well to the development of parameterisations that depend on the grid spacing of coarse-resolution models. In addition, as ocean models become eddy-permitting the temporal scale separation between resolved and unresolved flows – which is required for the temporal Reynolds decomposition to be valid – diminishes [Nadiga, 2008](#), in which case spatial filtering becomes the more appropriate method. The spatial filtering method can be extended by including a weighting function, by using a circular filter, or by additionally filtering in time. These upgrades can be considered in the future, but we use the filter in [\(6\)](#) since we first want to understand the model for the most primitive filters before introducing more parameters. We use a filter width of 112.5 km which is roughly three times the first deformation radius, and corresponds to $w = 31$. As shown in HSSB20, for larger/smaller filter sizes the amplitude of the transport tensor is greater/smaller, but its qualitative behaviour is unchanged. [Lu et al. 2016](#) additionally diagnose the time-mean and deviation parts of the large-scale and small-scale fields in order to separate the effects of transient and stationary eddies.

Applying the spatial filter to the tracer equation [\(5\)](#) yields the evolution equation for the large-scale tracer field:

$$\frac{\partial \bar{C}}{\partial t} + \nabla \cdot (\bar{\mathbf{u}} \bar{C}) + \nabla \cdot \mathbf{f} = \nu \nabla^2 \bar{C} + \bar{F}, \quad (7)$$

where

$$\nabla \cdot \mathbf{f} = \overline{\nabla \cdot (\mathbf{u} C)} - \nabla \cdot (\bar{\mathbf{u}} \bar{C}). \quad (8)$$

This is the divergence of the *eddy tracer flux* $\mathbf{f}$, and is the eddy forcing of the large-scale tracer field that would hypothetically be unresolved in a non-eddy resolving ocean model. Such a form for the eddy tracer flux and its divergence is more common in the context of large-eddy simulation [Lilly, 1967](#). It is possible to separately consider the Leonard flux, cross fluxes and eddy-eddy flux that contribute to $\mathbf{f}$ [Fox-Kemper and Menemenlis, 2008](#), [Leonard, 1975](#), but in this study we consider $\mathbf{f}$ as a whole.

The filter commutes with the divergence operator everywhere in the domain apart from within half a filter width's distance of the boundaries. Because of this non-commutativity near the boundaries, we do not have an explicit expression for $\mathbf{f}$ at every grid point. This is not an issue because we are only interested in the divergent part of $\mathbf{f}$, which we denote $\mathbf{f}_{\text{div}}$, and may calculate using the known field $\nabla \cdot \mathbf{f}$. Using $\mathbf{f}_{\text{div}}$ permits an analysis of the dynamically active part of the eddy tracer flux by omitting obscuring effects of the dynamically inert rotational part. In the next subsection we discuss the method and other motivations for our decision to focus on $\mathbf{f}_{\text{div}}$.

We work towards parameterising the divergent eddy tracer flux by invoking the flux-gradient relation,

$$\mathbf{f}_{\text{div}} = -\mathbf{K} \nabla \bar{C}, \quad (9)$$

which relates the eddy tracer flux to the large-scale tracer gradient via a 2×2 tensor $\mathbf{K}$. The tensor $\mathbf{K}$ is commonly referred to as the diffusivity, but we shall refer to it as the *transport tensor* in order to illuminate

the idea that it encompasses both diffusive and advective transport. For a single tracer equation (9) is underdetermined, so we introduce a second tracer. Denoting the two tracers as C_1 and C_2 , and denoting their associated divergent eddy fluxes as $\mathbf{f}_1 = (f_1^{(u)}, f_1^{(v)})$ and $\mathbf{f}_2 = (f_2^{(u)}, f_2^{(v)})$, inverting (9) gives

$$\mathbf{K} \equiv \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \frac{1}{d} \begin{pmatrix} f_1^{(u)} & f_2^{(u)} \\ f_1^{(v)} & f_2^{(v)} \end{pmatrix} \begin{pmatrix} -\overline{C}_{2,y} & \overline{C}_{2,x} \\ \overline{C}_{1,y} & -\overline{C}_{1,x} \end{pmatrix}, \quad (10)$$

where

$$d = \overline{C}_{1,x}\overline{C}_{2,y} - \overline{C}_{1,y}\overline{C}_{2,x} \quad (11)$$

is the determinant of the matrix of large-scale tracer gradients.

In the results presented in this study, a single pair of tracers is used to determine $\mathbf{K}$ as a function of space and time. The local, instantaneous values of $\mathbf{K}$ are non-unique with respect to the choice of the tracer pair (Kamenkovich et al., 2020, Sun et al., 2021) and the implications of this are discussed in section 6. We solve the flux-gradient relation for a pair of tracers with initially linear profiles of the form

$$C_0 = \frac{ax + by + c}{\sqrt{a^2 + b^2}}, \quad (12)$$

where a , b and c are constants. As will be discussed in more detail in the next subsection, our choice for a , b and c is arbitrary, since the transport tensor is the same for all tracers with linear initial conditions. We additionally tested tracers with wave-like initial profiles and found that although the resulting transport tensor is distinct, the same qualitative patterns persist.

To avoid the singularity in equation (10) when inverting the flux-gradient relation, i.e., to avoid $d = 0$, the tracers must have misaligned large-scale gradients. To maintain misalignment throughout the simulation, the term F in the tracer evolution equation (5) represents relaxation of the large-scale tracer field back to its initial profile. That is,

$$F = r(C_0 - \overline{C}), \quad (13)$$

where r is the relaxation rate. We use a relaxation timescale of 5 days which is large enough to avoid the singularity at $d = 0$ while not notably affecting the results. Other time scales were tested and our conclusions remain the same. (Bachman et al., 2015) derived a method for removing the effect of relaxation on the tracer inversion method, but this does not apply to our model for two reasons. First, their method applies to Reynolds-averaged eddy-eddy fluxes, whereas our un-averaged eddy tracer fluxes are more complicated. Second, our relaxation targets only the large-scale tracer field, whereas the method of (Bachman et al., 2015) is designed for the case of full tracer field relaxation.

2.3 Divergent fluxes and non-uniqueness

We calculate the transport tensor $\mathbf{K}$ for the divergent eddy tracer flux $\mathbf{f}_{\text{div}}$. We obtain $\mathbf{f}_{\text{div}}$ by using the Helmholtz decomposition (Lau and Wallace, 1979) which separates a vector into its divergent, rotational and

harmonic parts. Specifically, we use the Fourier analysis and cyclic reduction method (Swarztrauber, 1974) to invert the Poisson equation

$$\nabla \cdot \mathbf{f}_{\text{div}} = \overline{\nabla \cdot (\mathbf{u}C)} - \nabla \cdot (\overline{\mathbf{u}}\overline{C}) = \nabla^2 \phi, \quad (14)$$

yielding the divergent potential ϕ . The divergent eddy tracer flux is then

$$\mathbf{f}_{\text{div}} = \nabla \phi. \quad (15)$$

The Helmholtz decomposition is not unique in a bounded domain (Fox-Kemper et al., 2003) due to a dependence on the boundary conditions. Although this may be an argument against use of the divergent flux, a unique decomposition is reached by imposing physically motivated zero normal flux boundary conditions (Maddison et al., 2015). Another potential argument against the use of the divergent flux is the loss of the theoretical link to the transport tensor obtained via parcel excursion theory (Taylor, 1921). We also note that it may be unfeasible to remove rotational fluxes in real observations because they are mostly local. Despite these issues, we argue it is best to remove the rotational component of the eddy tracer flux because it is dynamically inert, i.e., it does not influence the dynamics, and because it can be orders of magnitude larger than the divergent part (Marshall and Shutts, 1981). Indeed, HSSB20 found that the rotational part of the flux $\mathbf{u}C' + \mathbf{u}'\overline{C} + \mathbf{u}'C'$ is typically two orders of magnitude larger than its divergent part. Thus, inclusion of the rotational flux leads to a transport tensor whose dynamically active part is obscured by a dominant inactive part which inhibits proper physical interpretation.

In general, the transport tensor $\mathbf{K}$ for a pair of independent tracers is non-unique, meaning that it depends on the tracers used in the inversion. This issue exists for all studies of transport tensors for passive tracers, but has rarely been analysed quantitatively (Sun et al., 2021). In this study, because we remove the rotational eddy fluxes, $\mathbf{K}$ does not depend on the initial linear gradients (a and b) of the tracer fields. To expand on this, we first note that the constant c in the tracer initial conditions (12) only contributes to the rotational eddy tracer flux because the flow is non-divergent. Second, the linearity of the flux-gradient relation implies that the same $\mathbf{K}$ would be obtained for any tracer pairs that are linear combinations of one another (i.e., of the form $AC_1 + BC_2$, for $A, B \in \mathbb{R}$). These two properties mean that $\mathbf{K}$ is the same for all tracers with linear initial conditions, which has been confirmed numerically. We also simulated tracers with spatially non-linear, large-scale profiles and find that the resulting transport tensor is different, but retains the same qualitative behaviour. See the further discussion of non-uniqueness in section 6.

2.4 The filtered QG solution

The QG solution is spun up from rest until statistical equilibrium. After this the tracers and flow evolutions are simulated for one more year with data on which our analysis is based saved daily. QG double-gyre flow is a well-studied problem (e.g., Shevchenko and Berloff, 2016, 2015) so flow features will only be discussed briefly. The upper-layer contains a fast, meandering eastward jet characterised by a sharp meridional PV

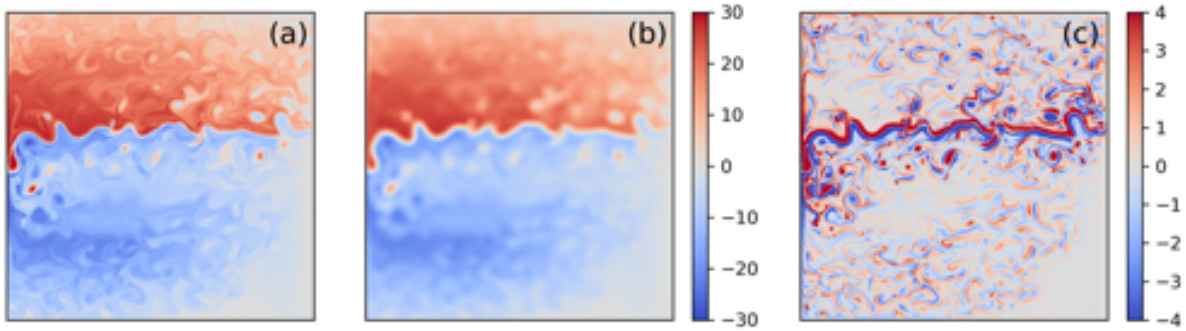


Figure 1: (a) The upper-layer PV anomaly q , (b) its large-scale part $\bar{q}$ and (c) its small-scale part q' . The PV here is dimensionless. Panels (a) and (b) share a colorbar. In this figure, and all following figures showing fields in a given layer, the x -axis represents longitude and the y -axis represents latitude.

gradient. The jet separates broad recirculation gyres that feed western boundary currents that in turn feed the jet. In the middle layer a weak eastward jet is present, but the PV field is more homogenized, with its meridional gradient dominated by the planetary vorticity. In the lower layer, there is little notable instantaneous eastward flow, but we do observe alternating latent jets.

To illustrate the effects of the spatial filter, in figure 1 we show snapshots of (a) the upper-layer PV anomaly, (b) its large-scale part and (c) its small-scale part. As in Berloff (2018), the small-scale PV field q' exhibits a ribbon of opposite-signed PV anomaly that straddles the jet and augments the relatively weak cross-jet PV gradient in the large-scale field $\bar{q}$. Swirls and filaments are ubiquitous in q' and have little to no signal in $\bar{q}$. Vortices, in particular the large ones, can have a signal in $\bar{q}$, but their amplitude and associated gradients are augmented by q' .

3 The transport tensor

In figure 2 we show snapshots of each element of the transport tensor $\mathbf{K}$ in the upper layer. Figure 3 presents the time means, denoted by $\langle \cdot \rangle$, of the same elements. The amplitude of $\mathbf{K}$ is typically on the order of $10^3 \text{ m}^2 \text{ s}^{-1}$, and is similar to estimates from a range of studies (Abernathey et al., 2010, Eden and Greatbatch, 2009, Klocker et al., 2012b, Zhurbas and Oh, 2004, Zhurbas et al., 2014). The instantaneous and time-mean transport tensors are vastly different, but there are certain spatial features that are common between the two. Each element of $\mathbf{K}$ is largest in amplitude in the eastward jet region where the flow is most energetic, a feature which persists in the other layers but is less prevalent as the jet representation weakens with depth. In and near the eastward jet the diagonal entry K_{11} is mostly positive, whereas large positive K_{22} is most common on the jet flanks. In particular K_{22} tends to be large and positive on the equatorward flank. A similar observation is made by other studies (Eden, 2006, Lu et al., 2016, Marshall et al., 2006). Positive

transport coefficients are usually interpreted as representing down-gradient fluxes, but whether the whole tensor truly quantifies locally down-gradient fluxes is unclear since the off-diagonal elements are significant.

A particularly notable feature is the dipole K_{21} pattern that straddles the downstream section of the jet. Dependent on the sign of the zonal tracer gradient, K_{21} exhibits a tendency for tracer flux divergence/convergence near the jet core and convergence/divergence further away from the jet. This divergence/convergence pattern is achieved by zonal advection by the eddy flow, which augments advection by the large-scale jet. Note that K_{21} achieves this via a divergent meridional flux, rather than a zonal one, which is a consequence of the Helmholtz decomposition combined with the differential tracer flux convergence, the latter caused by the meridionally sheared zonal velocity. Meanders in the jet mean the flow may be meridionally oriented, in which case the along-jet flux previously parameterised by K_{21} is instead parameterised by K_{12} . This gives rise to the alternating sign pattern in K_{12} and $\langle K_{12} \rangle$ on the jet.

In the middle layer each element of $\mathbf{K}$ (not shown) has many qualities in common (e.g., cross-jet sign change in K_{21} and alternating sign pattern in K_{12}) with the upper-layer $\mathbf{K}$, but these only become clear after time averaging. Similarly, instantaneous $\mathbf{K}$ in the lower layer (not shown) has few notable spatial features, other than large amplitudes in the jet region and western boundary current. We show the lower-layer time-mean $\mathbf{K}$ in figure 4. Somewhat similar to in the layers above, $\langle K_{21} \rangle$ changes sign near the jet (in the lower layer the time-mean signature of the jet is a broadly meandering flow) with negative/positive values to the north/south. We have that $\langle K_{22} \rangle$ in the lower layer (and in the middle layer) is large and positive at the latitude of the jet core, whereas in the upper layer large $\langle K_{22} \rangle$ is concentrated in the equatorward flank of the jet. This suggests that meridional cross-jet transport is stronger at depth and weaker near the surface. This will be expanded upon in section 4.2

Lastly, it is significant that K_{12} and K_{21} are entirely distinct from each other, since Lagrangian single-particle dispersion-based methods diagnose equal K_{12} and K_{21} by construction (Rypina et al., 2012). Furthermore, K_{12} and K_{21} are at times neglected in oceanographic studies (e.g., Eden et al., 2007), but our results show that they are similar in amplitude to the diagonal components, suggesting that they ought to not be neglected.

4 The diffusion tensor

We separate the transport tensor $\mathbf{K}$ into its symmetric and antisymmetric parts since these are associated with distinct physical processes. These tensors are

$$\mathbf{S} = \frac{1}{2} (\mathbf{K} + \mathbf{K}^T) \quad \text{and} \quad \mathbf{A} = \frac{1}{2} (\mathbf{K} - \mathbf{K}^T), \quad (16)$$

where superscript T denotes the transpose. We refer to the symmetric tensor $\mathbf{S}$ as the *diffusion tensor* since its associated fluxes drive irreversible mixing processes and transfer variance between the large and small scales. We refer to the antisymmetric tensor $\mathbf{A}$ as the *advection tensor* since it encompasses effects equivalent to advection by a 2D non-divergent velocity field (Griffies, 1998). The decomposition of $\mathbf{K}$ into

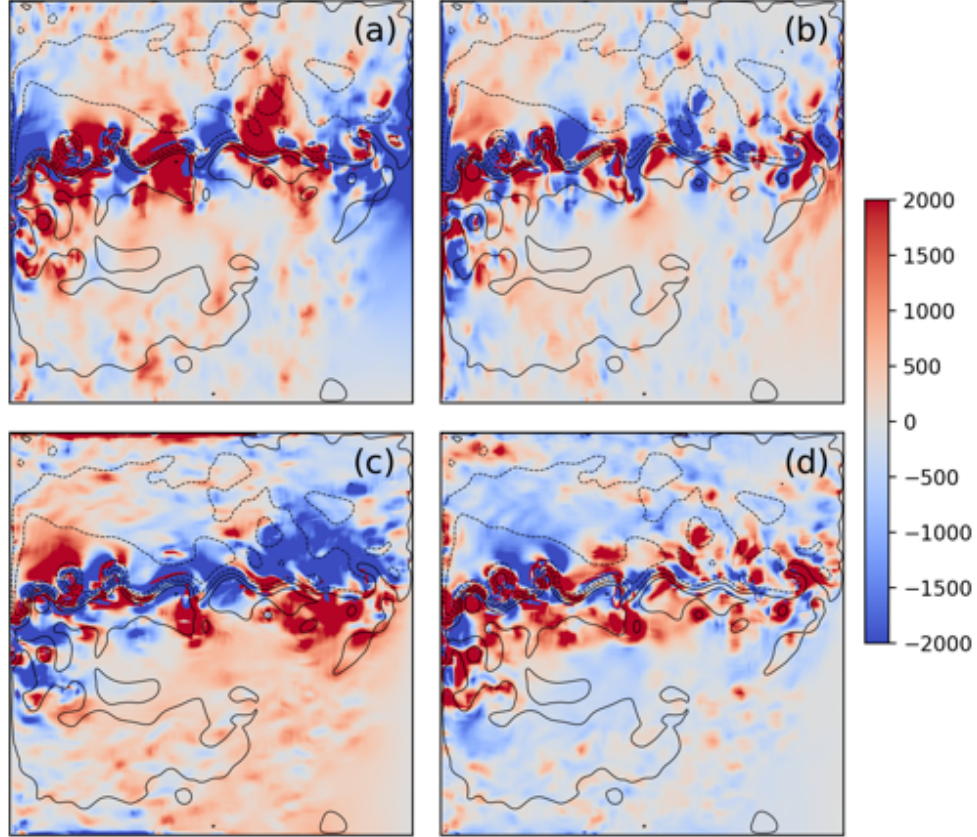


Figure 2: Snapshots of the elements of $\mathbf{K}$ for the upper layer. Plotted are (a) K_{11} , (b) K_{12} , (c) K_{21} and (d) K_{22} . Units are $\text{m}^2 \text{s}^{-1}$. The black solid/dashed lines represent positive/negative contours of the instantaneous large-scale streamfunction.

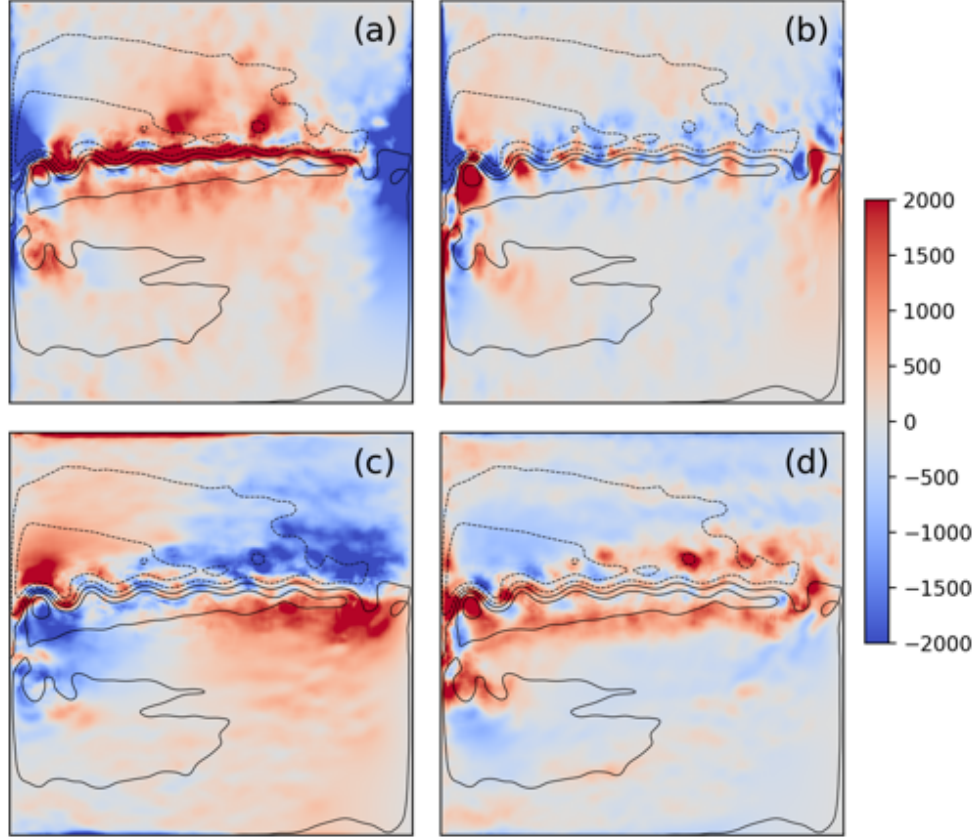


Figure 3: Annual-mean elements of $\mathbf{K}$ for the upper layer. Plotted are (a) $\langle K_{11} \rangle$, (b) $\langle K_{12} \rangle$, (c) $\langle K_{21} \rangle$ and (d) $\langle K_{22} \rangle$. Units are $\text{m}^2 \text{s}^{-1}$. The black solid/dashed lines represent positive/negative contours of the time-mean large-scale streamfunction.

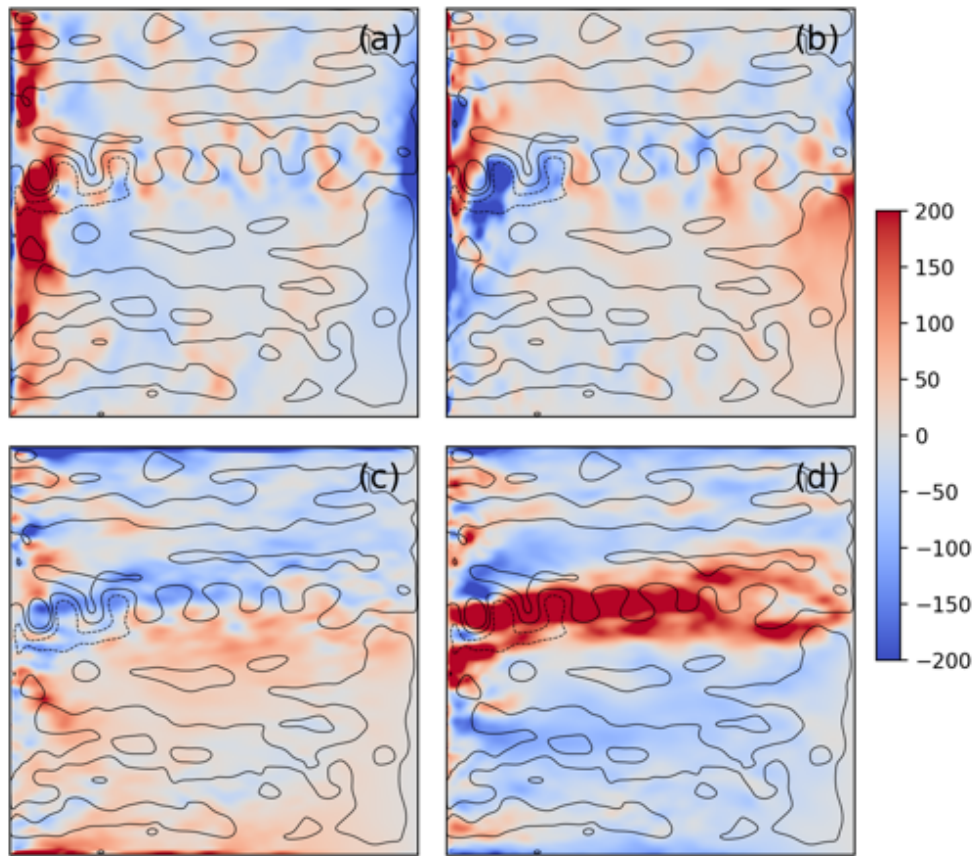


Figure 4: The same as figure 3 but for the lower layer.

$\mathbf{S}$ and $\mathbf{A}$ separates the eddy tracer flux $\mathbf{f}_{\text{div}} = -\mathbf{K}\nabla\overline{C}$ into the diffusive flux $\mathbf{f}_{\text{diff}} = -\mathbf{S}\nabla\overline{C}$ and the skew flux $\mathbf{f}_{\text{skew}} = -\mathbf{A}\nabla\overline{C}$. By design we have that $\mathbf{f}_{\text{div}} = -\mathbf{K}\nabla\overline{C}$ is purely divergent, but the diffusive and skew fluxes will in general have non-zero rotational components that cancel with one another. Unlike with the full eddy tracer flux $\mathbf{f}$ these rotational components do not dominate the divergent components. For the remainder of the study we focus on the diffusion tensor. An interpretation of the advection tensor will follow in a companion paper.

The entries $S_{11} = K_{11}$ and $S_{22} = K_{22}$ are shown in previous figures 2, 3 and 4. We show the time-mean S_{12} , that is the off-diagonal component of $\mathbf{S}$, in figure 5 for all layers. Also shown is the point-wise standard deviation of S_{12} , namely S_{12}^{sd} , in each layer. The dominant spatial patterns of instantaneous and time-mean $S_{12} = S_{21}$ in the upper two layers are similar to, but weaker than, the cross-jet sign dipole pattern of K_{21} . Thus the diffusive flux partly contributes to the tracer flux convergence/divergence near the jet, and the skew flux augments the convergence/divergence. In the lower layer, the instantaneous S_{12} (not shown) is large near the jet and western boundary current, but otherwise has no particularly clear spatial patterns. The time-mean S_{12} is similar to $\langle K_{21} \rangle$ with its noticeable cross-jet sign change pattern, but the contribution from $\langle K_{12} \rangle$ destroys this near the western boundary. In each layer the standard deviation S_{12}^{sd} is similar to the standard deviations of elements $S_{11} = K_{11}$ and $S_{22} = K_{22}$. In the upper two layers, S_{12}^{sd} is largest in the jet region and the subtropical western boundary current, whereas in the lower layer S_{12}^{sd} is largest in the western boundary currents. Thus regions of high standard deviation coincide with regions of strong eddy activity.

4.1 The diffusion eigenvalues

Since the diffusion tensor $\mathbf{S}$ is real and symmetric, it has real eigenvalues which represent diffusivities in the direction of their respective eigenvectors. We obtain the eigenvalues of $\mathbf{S}$ by rotating the coordinate system through an angle α where

$$\alpha = \frac{1}{2} \tan^{-1} \left(\frac{2S_{12}}{S_{11} - S_{22}} \right). \quad (17)$$

This rotation diagonalises $\mathbf{S}$:

$$\mathbf{S}' = \mathbf{R}^T \mathbf{S} \mathbf{R} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad (18)$$

where $\mathbf{R}$ is the rotation matrix. The eigenvalues of $\mathbf{S}$ are the diagonal entries of $\mathbf{S}'$ and are given by

$$\lambda_1 = S_{11} \cos^2 \alpha + S_{22} \sin^2 \alpha + 2S_{12} \cos \alpha \sin \alpha, \quad (19)$$

$$\lambda_2 = S_{11} \sin^2 \alpha + S_{22} \cos^2 \alpha - 2S_{12} \cos \alpha \sin \alpha. \quad (20)$$

If we ensure that the correct quadrant is selected when calculating α , then we have that $\lambda_1 \geq \lambda_2$. The angle α is referred to as the *diffusion angle*. The vector $\mathbf{v}_1 = (\cos \alpha, \sin \alpha)$, referred to as the *diffusion axis*, is the eigenvector corresponding to the most positive eigenvalue, λ_1 . The second eigenvector $\mathbf{v}_2 = (\sin \alpha, -\cos \alpha)$ is perpendicular to the first.

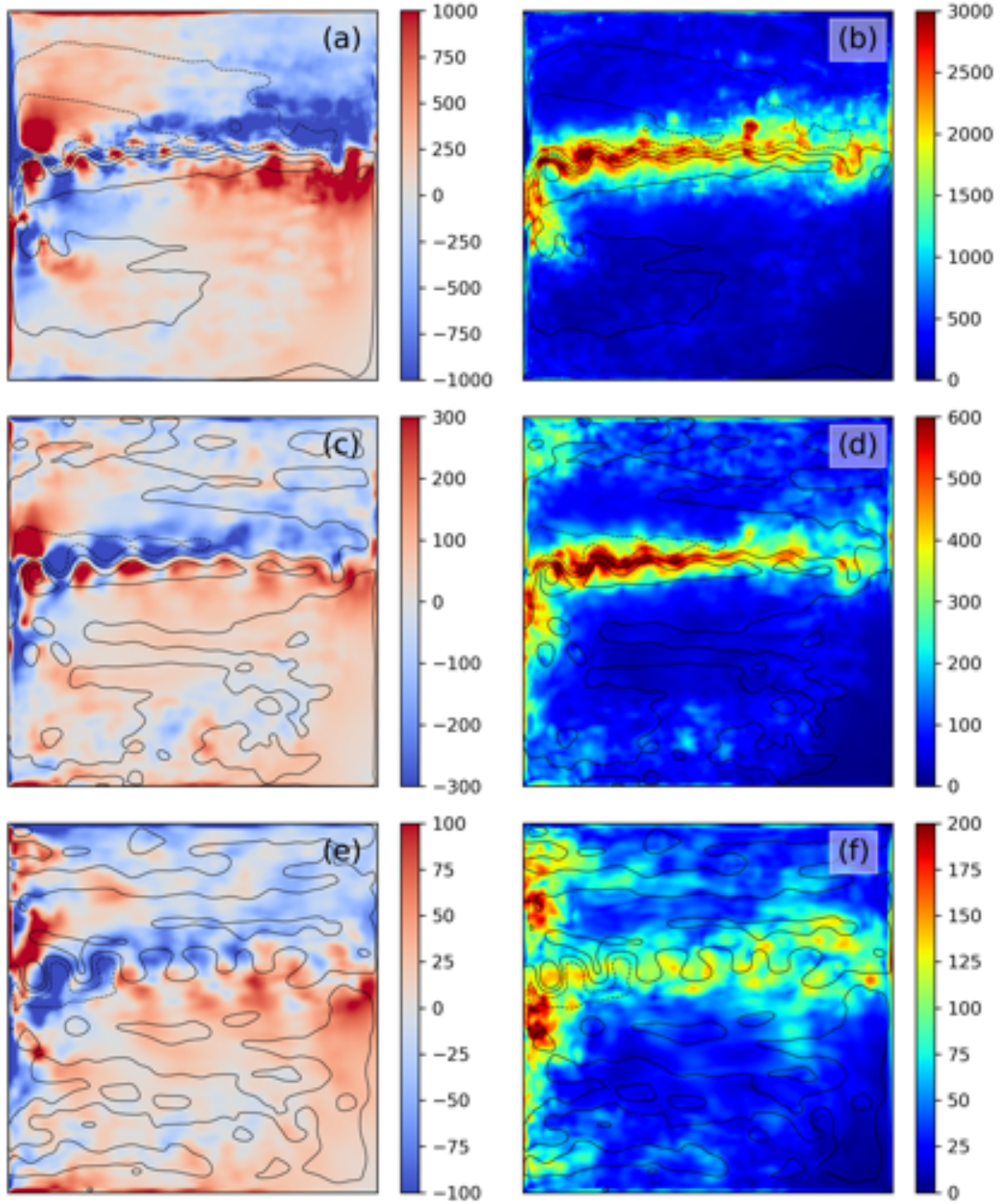


Figure 5: The time-mean and standard deviation of the off-diagonal diffusion tensor component S_{12} in each layer. Plotted are (a, c, e) the time mean, $\langle S_{12} \rangle$; (b, d, f) the standard deviation, S_{12}^{sd} . These are for the upper (a, b), middle (c, d) and lower (e, f) layers. Units are $\text{m}^2 \text{s}^{-1}$ in all panels. Contours represent the time-mean large-scale streamfunction.

Figure 6 shows snapshots of the two diffusion eigenvalues in all three layers. The eigenvalues are largest in the upper layer, smallest in the lower layer, and are generally stronger in the jet region or near the western boundary of each layer. For either eigenvalue distinctive spatial patterns in and near the jet (other than the increased amplitude) are not observable. The diffusion eigenvalue amplitudes, e.g., upper-layer values on the order of $10^3 \text{ m}^2 \text{ s}^{-1}$, are broadly in agreement with diffusivities obtained in other studies (Abernathy et al., 2010, Eden and Greatbatch, 2009, Klocker et al., 2012b, Zhurbas and Oh, 2004, Zhurbas et al., 2014). A deeper set of statistics covering the quantitative behaviour of diffusion eigenvalues was given in HSSB20. In each layer and in most areas of the ocean the eigenvalues are of opposite sign – we refer to these as *polar eigenvalues* – and these exist when $S_{12}^2 > S_{11}S_{22}$. Polar eigenvalues persist throughout time so these snapshots illustrate robust behaviour. In addition, we find that the eigenvalues of $\langle \mathbf{S} \rangle$ are most often pairs of opposite sign.

The diffusive eddy tracer flux can be expressed as

$$\mathbf{f}_{\text{diff}} = -\mathbf{S}\nabla\overline{C} = -|\nabla\overline{C}|(\lambda_1 \cos(\alpha - \omega)\mathbf{v}_1 + \lambda_2 \sin(\alpha - \omega)\mathbf{v}_2), \quad (21)$$

where $\omega = \tan^{-1}(\overline{C}_y/\overline{C}_x)$ is the orientation of the large-scale tracer gradient. The diffusive flux can be interpreted as the sum of two fluxes, one in the direction of the diffusion axis $\mathbf{v}_1$ scaled by $\lambda_1 \cos(\alpha - \omega)|\nabla\overline{C}|$ and another in the direction of $\mathbf{v}_2$ scaled by $\lambda_2 \sin(\alpha - \omega)|\nabla\overline{C}|$. If an eigenvalue is positive/negative then its flux contribution is oriented down/up the large-scale tracer gradient. For polar eigenvalues $\mathbf{f}_{\text{diff}}$ is the sum of an up-gradient and a perpendicular down-gradient flux, such that the prevailing effect is filamentation in the tracer concentration (Ledwell et al., 1998). This is the general case, but if the diffusion axis $\mathbf{v}_1$ aligns with the large-scale tracer gradient $\nabla\overline{C}$, then $\mathbf{f}_{\text{diff}}$ is a directly down-gradient flux and induces mixing. If $\mathbf{v}_1$ is perpendicular to $\nabla\overline{C}$ then $\mathbf{f}_{\text{diff}}$ is a directly up-gradient flux, driving frontogenesis.

Polar eigenvalues have rarely been examined in previous studies (Bachman et al., 2020, Eden and Greatbatch, 2009, Stanley et al., 2020), since negative eigenvalues are often considered undesirable for modelling purposes (Fox-Kemper et al., 2013) due to potential model instabilities. Given that $\mathbf{f}_{\text{div}} = -\mathbf{K}\nabla\overline{C}$ reproduces (ignoring numerical errors) the eddy tracer flux diagnosed from a stable QG solution, negative eigenvalues in $\mathbf{S}$ need not be problematic. Mathematically, the eigenvalues are of opposite sign if $S_{12}^2 > S_{11}S_{22}$. Thus, we have polar eigenvalues if the diagonal entries of $\mathbf{S}$ are of opposite sign, or if the off-diagonal component S_{12} is large relative to the diagonal ones – these conditions are satisfied in approximately equal measure. Similarly, it can be shown that an increase in $|S_{12}|$ always positively/negatively contributes to the first/second eigenvalue, and thus always contributes to the eigenvalue polarity. This notion has important consequences because many previous studies have often considered either a diagonal tensor or a scalar diffusivity, which can significantly affect the interpretation of the mixing. We therefore argue that studies of eddy diffusion (or, more generally, transport) should consider the most general tensor, and not be biased against diagnosing negative eigenvalues.

We interpret polar eigenvalues as being due to approximate area conservation of a Lagrangian volume element, a consequence of the advecting flows ($\overline{\mathbf{u}}$ and $\mathbf{u}'$) being non-divergent. That is, if there is a positive

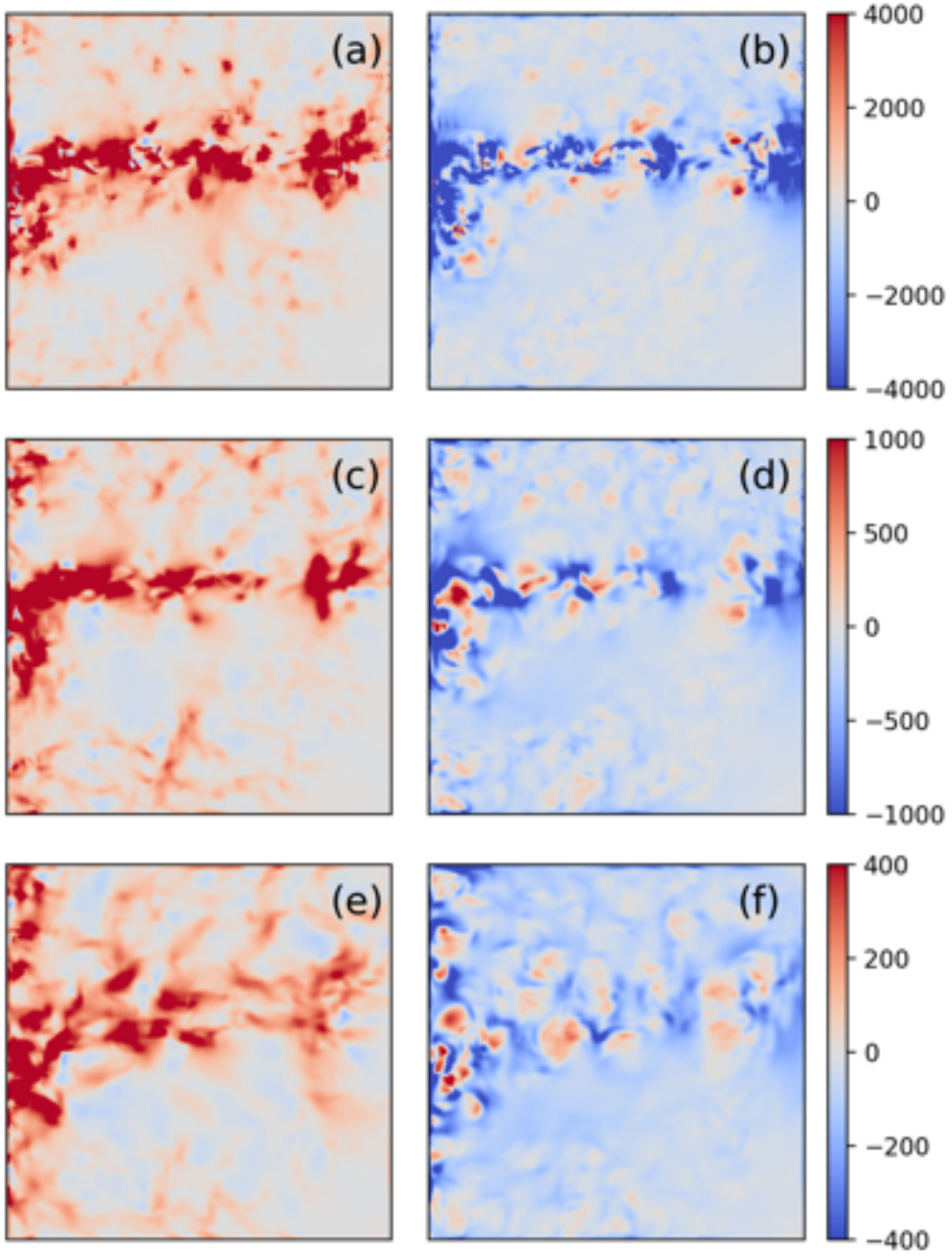


Figure 6: Snapshots of the first diffusion eigenvalue λ_1 (a, c, e) and the second diffusion eigenvalue λ_2 (b, d, f) in the upper layer (a, b), the middle layer (c, d) and the lower layer (e, f). Units are $\text{m}^2 \text{s}^{-1}$ in each panel.

spreading in one direction, then there must be a compensating negative spreading (contraction) perpendicular to this in order to prevent a large change in the volume of a Lagrangian volume element from one time step to the next. This same notion can be interpreted in the context of spreading ellipses (Kamenkovich et al., 2015, Rypina et al., 2012). For a diagonal diffusion tensor, the ellipse’s major and minor axes align with the coordinate axes. Inclusion of S_{12} rotates the ellipse, lengthens the major axis and shortens the minor axis (makes it more negative). Examination of the eddy tracer flux expression (21) can also provide insight. Recalling that $\mathbf{K}$ and $\mathbf{S}$ are obtained for a tracer pair, if the eddy flux of one tracer is oriented up-gradient and for the other tracer it is down-gradient, then polar eigenvalues are guaranteed. This, though, is not a necessary condition, and we find that polar eigenvalues exist when the both tracer fluxes are up/down-gradient.

4.1.1 Anisotropy

Isopycnal diffusion is most commonly implemented as being isotropic, i.e., with no preferred direction. However, there is growing evidence (Kamenkovich et al., 2009, 2015, Rypina et al., 2012) that tracer transport/diffusion is highly anisotropic, in particular in eastward jets (Ferrari and Nikurashin, 2010, Klocker and Abernathey, 2014, Klocker et al., 2012a, Smith, 2005). Since we observe eigenvalues which are robustly of opposite sign, our results also suggest that mixing is not isotropic. To examine this more closely we consider the mixing anisotropy, defined as the ratio of the eigenvalues, $r = \lambda_1/\lambda_2$. Similar to the findings of (Bachman et al., 2020), the anisotropy has no significant spatial pattern. In figure 7 we present a histogram of the anisotropy r evaluated over one year in the upper layer. The frequency distribution in the lower layers is very similar, and the anisotropy distribution for the eigenvalues of the time-mean $\mathbf{S}$ is also the same. Due to the eigenvalue polarity the anisotropy is most often negative and most commonly lies in the integer interval $[-1, 0]$ for which we must have $\lambda_1 > 0$, $\lambda_2 < 0$ and $|\lambda_1| < |\lambda_2|$. Such values are most common in areas where the eigenvalues are small, and we note that the domain-mean sum of both eigenvalues is positive. The frequency for $r \approx 1$ is very small, indicating that mixing due to $\mathbf{S}$ is explicitly not isotropic.

4.1.2 Inhomogeneity

It is clear the $\mathbf{S}$ and its eigenvalues are spatially dependent and therefore the diffusive eddy transport is inhomogeneous. We can quantify the inhomogeneity by calculating lengthscales associated with the diffusion eigenvalues. For eigenvalue $i = 1, 2$, measures of the local zonal and meridional lengthscales are

$$l_{ix} = \lambda_i \left(\frac{\partial \lambda_i}{\partial x} \right)^{-1} \quad \text{and} \quad l_{iy} = \lambda_i \left(\frac{\partial \lambda_i}{\partial y} \right)^{-1}, \quad (22)$$

respectively. This metric produces lengthscales near infinity when the derivative is near zero, but we exclude such points that from the histogram we show in figure 8 by limiting the range to $[0, 300]$ km. The results in figure 8 are for the upper layer, evaluated over one year. The lengthscales l_{1x} , l_{1y} , l_{2x} and l_{2y} , which span multiple orders of magnitude, have distributions similar to one another. The fact that the eigenvalues vary

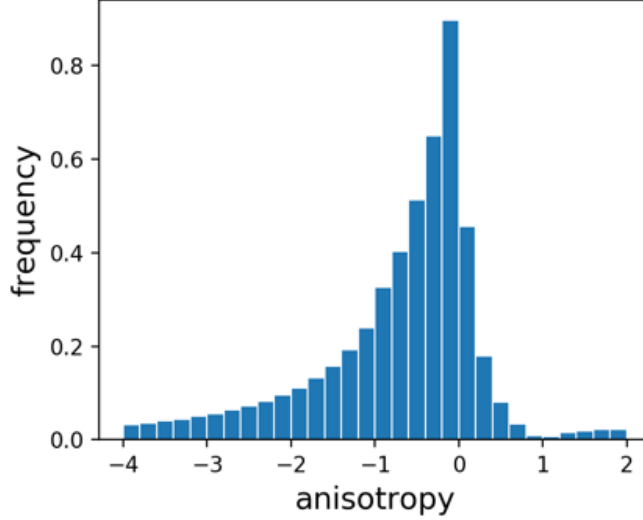


Figure 7: Histogram of the diffusion anisotropy, r , equal to the ratio of the eigenvalues. The y -axis represents the normalised frequency. The data are evaluated over one year. This is for the upper layer, but results in the layers below are similar.

on such small scales indicates that the mixing due to $\mathbf{S}$ is highly inhomogeneous. For λ_1 lengthscales in $[20, 40]$ km are most common whereas for λ_2 lengthscales in $[0, 20]$ km are most common. In the two layers below, such short lengthscales are still most common, but the overall frequency distribution (not shown) is closer to being uniform. For the time-mean $\mathbf{S}$ and its associated eigenvalues, short lengthscales (< 40 km) are again most common for both eigenvalues.

4.2 The diffusion angle

Recall the definition of the diffusion angle α in equation (17), such that $\alpha \in [-\pi/2, \pi/2]$ and is π -periodic. The diffusion axis is the first eigenvector of $\mathbf{S}$, that is $\mathbf{v}_1 = (\cos \alpha, \sin \alpha)$, and corresponds to the most positive of the two eigenvalues. In figure 9 we summarise the basic dependence of the diffusion axis on the elements of $\mathbf{S}$. In most studies which consider the diffusion angle the corresponding diffusion tensor is calculated from Reynolds-averaged eddy-eddy fluxes [Abernathy et al., 2013, Bachman et al., 2020] or ensemble-averaged particle trajectories [Kamenkovich et al., 2015, Rypina et al., 2012], and it is generally found that the diffusion axis is aligned with the mean velocity vector and is perpendicular to the mean PV gradient. This alignment can be attributed to sharp PV gradients such as those in eastward jets that suppress mixing across the gradient, [Ferrari and Nikurashin, 2010], or flow shears that enhance mixing in the direction of the flow [Jones and Young, 1994, Young et al., 1982].

In our study α is derived from highly variable instantaneous eddy fluxes and is therefore highly variable itself. As a result it is simpler to observe its overarching behaviour in the time-mean framework. Since

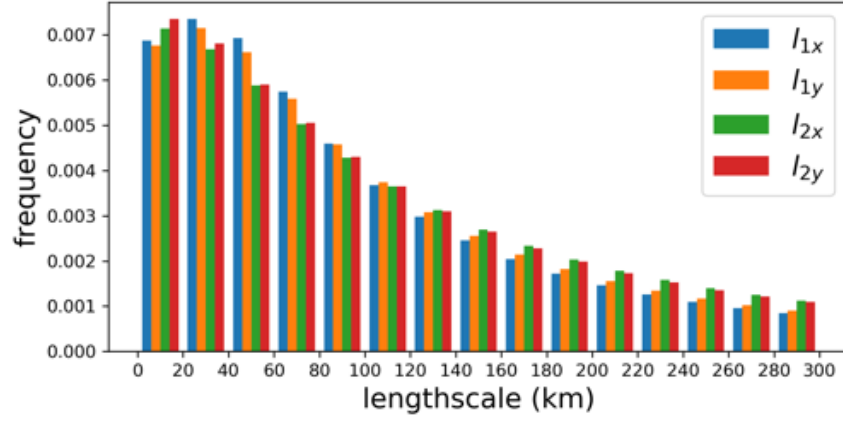


Figure 8: Histogram of the lengthscales associated with the diffusion eigenvalues, λ_1 and λ_2 , a measure of their inhomogeneity. See the legend for details. The x -axis represents the lengthscale in units km and the y -axis represents the normalised frequency. The data are collected for a year in the upper layer.

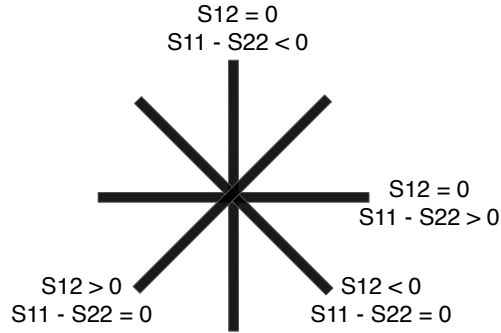


Figure 9: Diagram of the diffusion axis v_1 orientation and its dependence on the signs of S_{12} and $S_{11} - S_{22}$. As an example, if $S_{12} > 0$ and $S_{11} = S_{12}$, then the diffusion axis is oriented south-west to north-east, for which $\alpha = \pi/4$.

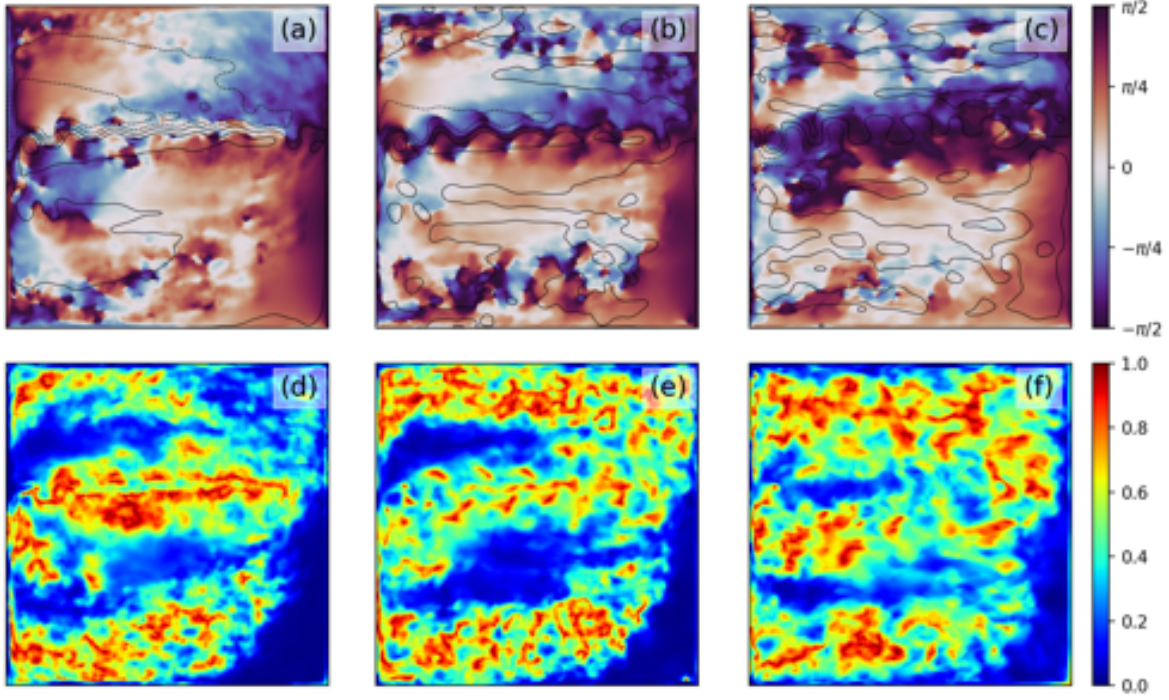


Figure 10: (a, b, c) Half the time-mean of the angle 2α and (d, e, f) the circular variance of 2α in the upper (a, d), middle (b, e) and lower (c, f) layers. The statistics are evaluated over a year. In the panels (a, b, c) contours represent the time-mean large-scale streamfunction.

the diffusion angle α is angular data standard averaging methods are not suitable, so instead we appeal to so-called circular statistical methods. Given a time series of an angle $\varphi \in [-\pi, \pi]$, its mean is calculated by finding the angle of the mean of the unit vectors associated with each φ instance. Thus, denoting each discrete instance of φ as φ_n , the mean is

$$\langle \varphi \rangle = \text{Arg} \left(\sum_{n=1}^N \exp(i\varphi_n) \right), \quad (23)$$

where Arg denotes the complex argument. We can also calculate the circular variance of φ . This is

$$\varphi_{\text{var}} = 1 - R_\varphi \quad \text{where} \quad R_\varphi = \frac{1}{N} \sqrt{\left(\sum_{n=1}^N \cos \varphi_n \right)^2 + \left(\sum_{n=1}^N \sin \varphi_n \right)^2} \quad (24)$$

is the mean resultant length of the vectors associated with each φ_n . We also have to account for the fact that α is π -periodic rather than 2π -periodic, which we do by calculating the mean and variance of 2α . In figure 10 we present the half the time-mean of 2α and the circular variance of 2α , which for short we refer to as just the mean ($\langle \alpha \rangle$) and variance (α_{var}) of α .

In the upper layer the mean of α has some clear behaviour, in particular the cross-jet pattern which indicates that the $\mathbf{v}_1$ tends to be leant into the shear of the jet. This average behaviour is only observed

sufficiently downstream in the jet, away from the western boundary. The diffusion axis leaning into the shear of the jet is expected given the strong dipole K_{21} profile that straddles the jet, which leads to large S_{12} and consequently α near to $\pm\pi/4$. There is a transition on the jet core as $\langle\alpha\rangle \approx 0$ switches sign, at which point the instantaneous $\mathbf{v}_1$ has a tendency to be aligned with the jet core. Similar to the upper layer, $\mathbf{v}_1$ in the lower layers is on average leant into the shear of the jet. Instead of becoming more aligned with the upper-layer jet core when near it, however, $\mathbf{v}_1$ is instead more likely to be perpendicular to the jet, such that $\langle\alpha\rangle$ is near to $\pm\pi/2$ under the jet core. This is due to $S_{12} \approx 0$ and $S_{22} > S_{11}$, i.e., meridional mixing dominating zonal mixing, which is not the case in the upper layer. Our results therefore suggest that there is reduced mixing across the jet core in the upper layer, and enhanced mixing beneath the jet. This agrees with findings of other studies (Ferrari and Nikurashin, 2010; Klocker and Abernathey, 2014; Klocker et al., 2012a) where the mixing maxima beneath the jet is attributed to the presence of a critical layer. With vertical structure in our model represented by only three layers, we cannot confidently attribute our results to a critical layer.

In the upper layer the variance of α is large in three distinct regions: the jet, the northern section of the subpolar gyre and the southern section of the subtropic gyre. In between these regions the variance is low. In the middle layer α_{var} is large on the jet core and weak towards the centres of the gyres. Just as in the upper layer, the high variance on the jet core is likely due to shifting and meandering of the jet, which still has a notable signal in the middle layer. In the lower layer the behaviour of α_{var} is again different, with a broad low-variance jet region, indicating that roughly meridional orientation of $\mathbf{v}_1$ persists throughout time.

4.3 Relation with the large-scale flow

Previous studies of the diffusion axis predicted that it is aligned with the mean flow and is perpendicular to the mean PV gradient (Bachman et al., 2020; Rypina et al., 2012). These relationships exist because PV gradients can suppress mixing across the gradient (Ferrari and Nikurashin, 2010) and flow shears can enhance mixing in the direction of the flow (Jones and Young, 1994; Young et al., 1982). Both effects increase along-flow mixing relative to cross-flow mixing. We quantify the relation between the diffusion axis orientation α and the large-scale flow by considering the angle between the diffusion axis $\mathbf{v}_1$ and (i) the large-scale velocity vector $\bar{\mathbf{u}}$, (ii) the large-scale relative vorticity (RV) gradient $\nabla\bar{\zeta}$ and (iii) the large-scale thickness gradient $\nabla\bar{h}$. In figure 11 we present histograms for these three angles in all three layers, evaluated over one year of data. In all layers $\mathbf{v}_1$ is commonly aligned with $\bar{\mathbf{u}}$, with a monotonic decrease in the frequency as the angle increases through $[0, \pi/2]$ (for our choice of bins). Similarly, in all layers $\mathbf{v}_1$ is commonly perpendicular to $\nabla\bar{\zeta}$ and is least likely to be oriented down/up the gradient. We opt to show alignments with the RV gradient rather than the PV gradient since the latter has a slightly weaker, albeit similar, relationship with the diffusion axis. The diffusion axis $\mathbf{v}_1$ aligns with the thickness gradient almost as strongly as it does with the RV gradient, except in the lower layer where baroclinicity is weakest. Overall these results agree with other studies, but the alignment tendencies that we observe are relatively weak, in particular in comparison

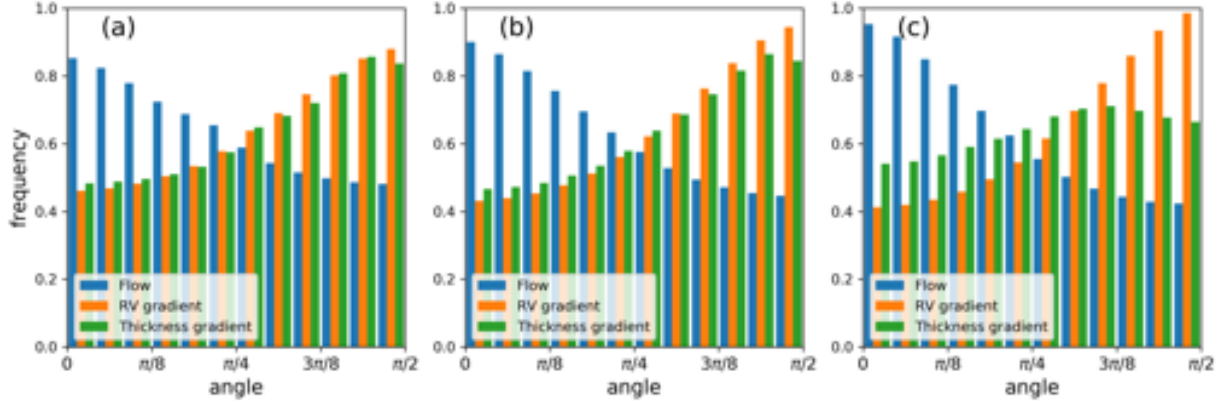


Figure 11: Histogram for the angle between the diffusion axis and (i) the large-scale velocity vector (blue), (ii) the large-scale RV gradient (orange) and (iii) the large-scale thickness gradient (green). Data is collected for a year in (a) the upper layer, (b) the middle layer and (c) the lower layer.

to Bachman et al. [2020] who studied Reynolds eddies. We give two reasons for this difference. First, α depends on instantaneous eddy fluxes which contribute to cross-flow and along-PV gradient mixing more so than their time-mean or zonal-mean values do. Second, we consider the divergent eddy tracer flux, whereas Bachman et al. [2020] retain its dominant rotational component.

Above we relate the orientation of preferential mixing to the large-scale flow, but parameterisation of eddy diffusion also requires knowledge of how the amplitude of the diffusion/transport relates to the large scales. For an initial test we link to the studies Eyink [2001], Meneveau and Katz [2000], Nadiga [2008] which proposed closures of the form $(\nabla \bar{\mathbf{u}})^T \nabla \bar{\xi}$ for the eddy fluxes of various active tracers ξ , each derived by Taylor expansion of the eddy flux. This implies using minus transpose of the large-scale velocity gradient matrix for the corresponding transport tensors. Because the derivations of Eyink [2001], Meneveau and Katz [2000], Nadiga [2008] can be applied to passive tracers, we first test the relation between $\mathbf{K}$ and $(\nabla \bar{\mathbf{u}})^T$. For this we use correlations evaluated over all grid points and all time instances. We find that $(\nabla \bar{\mathbf{u}})^T$ overall has a weak negative correlation with the transport tensor. In the lower layer the entry-wise correlations are: -0.15 for $(K_{11}, \bar{u}_x)$; -0.41 for $(K_{12}, \bar{v}_x)$; -0.27 for $(K_{21}, \bar{u}_y)$; and -0.10 for $(K_{22}, \bar{v}_y)$. Moving upwards the correlations becomes weaker, with largest values ~ 0.2 in the middle layer and ~ 0.1 in the upper layer. Thus, the approximation $\mathbf{K} \propto -(\nabla \bar{\mathbf{u}})^T$ has reasonable accuracy in the lower layer only. This implies that minus the strain rate tensor $\mathbf{E} = (\nabla \bar{\mathbf{u}} + (\nabla \bar{\mathbf{u}})^T)/2$, which has equal and opposite eigenvalues, could be used for $\mathbf{S}$ in the lower layer. The correlation between S_{12} and $\bar{u}_y$ and $\bar{v}_x$ illuminates how shear influences the diffusion axis $\mathbf{v}_1$. That is, the correlation implies that as shears grow stronger, as does S_{12} which drives the diffusion angle towards $\pm\pi/4$, which will in general inhibit alignment between $\bar{\mathbf{u}}$ and $\mathbf{v}_1$, having an effect opposite to shear dispersion Jones and Young [1994], Smith, [2005], Young et al. [1982]. We also considered point-wise correlations between $\mathbf{K}$ and $(\nabla \bar{\mathbf{u}})^T$ and found the resulting fields to be an ocean of negative

correlations with localised regions of positive correlation. We observe little consistency in these correlation patterns for the different tensor entries and the different layers.

For a broader search of potential closures for the amplitude of the eddy transport we considered numerous linear and nonlinear combinations [Anstey and Zanna, 2017] of the large-scale shear deformation $\bar{\theta} = \bar{u}_y + \bar{v}_x$, stretching deformation $\bar{\sigma} = \bar{u}_x - \bar{v}_y$ and relative vorticity $\bar{\zeta} = \bar{v}_x - \bar{u}_y$. We additionally considered the large-scale PV, thickness and their gradients, but these did not perform better. Motivated by their rotational invariance we seek to relate the diffusion eigenvalue amplitudes $|\lambda_1|$ and $|\lambda_2|$ and the Frobenius norms $\|\mathbf{S}\|$ and $\|\mathbf{K}\|$ to the large-scale flow. To be explicit, $\|\cdot\|$ denotes the square root of the sum of the squared elements and $\|\mathbf{S}\| = (\lambda_1^2 + \lambda_2^2)^{1/2}$. Of the numerous rotationally invariant combinations of $\bar{\theta}$, $\bar{\sigma}$ and $\bar{\zeta}$ considered, we find that the combination $\Omega = (\bar{\theta}^2 + \bar{\sigma}^2 + \bar{\zeta}^2)^{1/2} = \sqrt{2}\|(\nabla\bar{\mathbf{u}})^T\|$ performs best.

We show 2D histograms of Ω versus $|\lambda_1|$, $|\lambda_2|$, $\|\mathbf{S}\|$ and $\|\mathbf{K}\|$ for the lower layer in figure 12. The correlations for these pairs (evaluated over one year all grid points in the lower layer) are: 0.56 for $(|\lambda_1|, \Omega)$; 0.34 for $(|\lambda_2|, \Omega)$; 0.57 for $(\|\mathbf{S}\|, \Omega)$; and 0.61 for $(\|\mathbf{K}\|, \Omega)$. The same comparison in the upper layer yields the correlations: 0.37 for $(|\lambda_1|, \Omega)$; 0.34 for $(|\lambda_2|, \Omega)$; 0.41 for $(\|\mathbf{S}\|, \Omega)$; and 0.44 for $(\|\mathbf{K}\|, \Omega)$. These results indicate a weak positive correlation between Ω and the diffusion and transport magnitudes. Across all large-scale features considered, the first eigenvalue consistently correlates more strongly than the second. Previously we considered the relation $\mathbf{K} \propto -(\nabla\bar{\mathbf{u}})^T$ which suggests that the diffusion tensor magnitude $\|\mathbf{S}\|$ may be proportional to the strain magnitude $(\bar{\theta}^2 + \bar{\sigma}^2)^{1/2} = \sqrt{2}\|\mathbf{E}\|$. The strain magnitude was one of the large-scale features considered, but in comparison to Ω it does not correlate more strongly with $|\lambda_1|$, $\|\mathbf{S}\|$ or $\|\mathbf{K}\|$ and only has marginally stronger correlation with $|\lambda_2|$.

We conclude that the transport and diffusion tensors are related with the large-scale flow, but only weakly so. Given the weak relationship, our results suggest that a stochastic closure for $\mathbf{S}$ or $\mathbf{K}$ could be suitable. In particular the entries of $\mathbf{K}$ exhibit broad and approximate normal distributions when conditioned on the corresponding entries of $(\nabla\bar{\mathbf{u}})^T$, suggesting that normal distributions conditioned on the local $(\nabla\bar{\mathbf{u}})^T$ in coarse-grid model is a potential approach for parameterising $\mathbf{K}$. We have mostly focused on the lower layer, where the relations are strongest. The weaker relationship nearer the surface is due two factors. First, the correlations between $\mathbf{K}$ and $(\nabla\bar{\mathbf{u}})^T$ are due to the Leonard flux $\overline{\bar{\mathbf{u}}\bar{C}} - \bar{\mathbf{u}}\bar{C}$ [Leonard, 1975] and the cross flux $\overline{\mathbf{u}'\bar{C}}$. Second, at depth these two contributors to the eddy tracer flux are strongest, while nearer the surface the other two, $\overline{\bar{\mathbf{u}}\bar{C}'}$ and $\overline{\mathbf{u}'\bar{C}'}$, are just as important. Separate closures for the contributions made by these fluxes may be required for an accurate parameterisation of $\mathbf{K}$ near the surface. Finding a more accurate closure for $\mathbf{K}$ may require more sophisticated methods, such as a convolutional neural network [Bolton and Zanna, 2019] or a relevance vector machine [Zanna and Bolton, 2020].

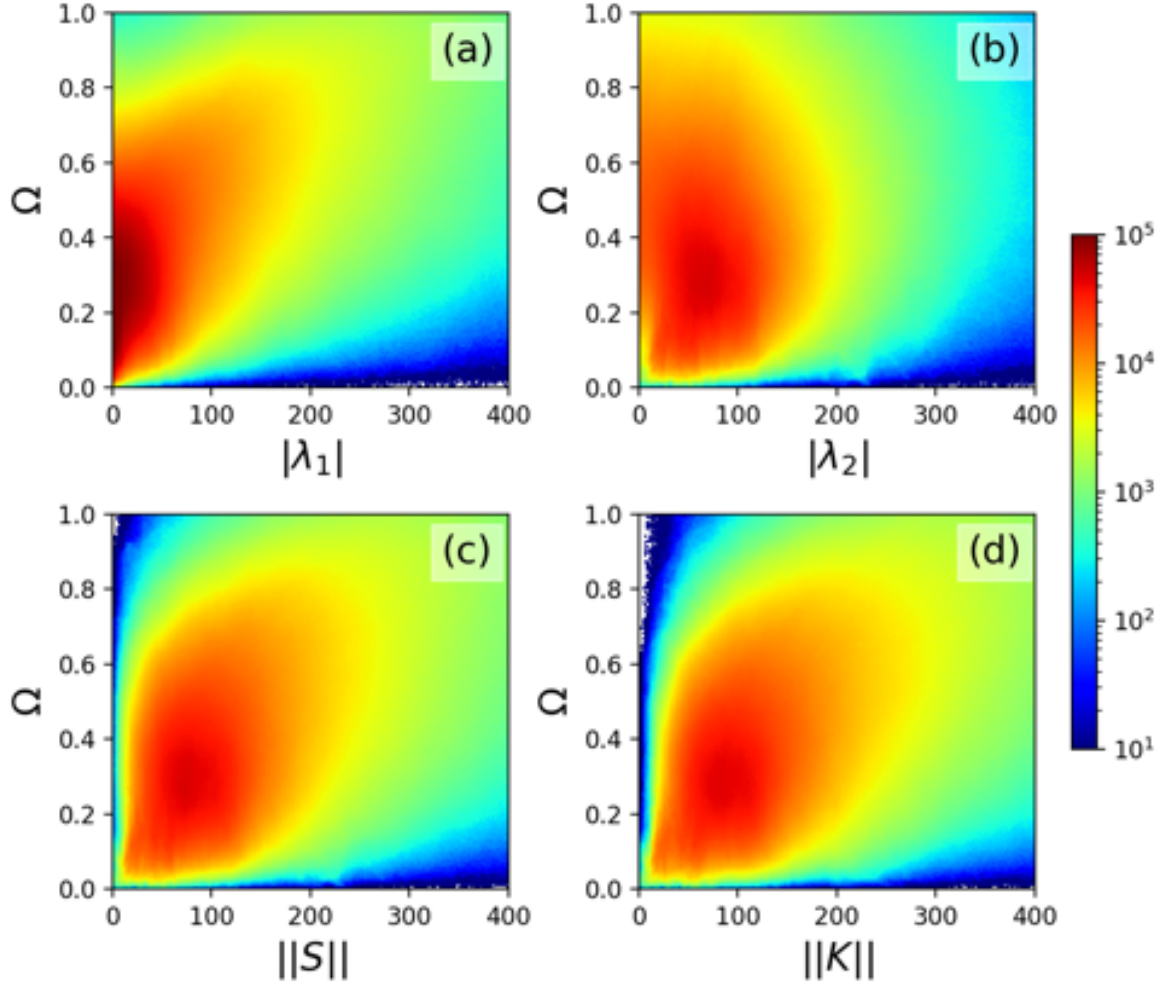


Figure 12: 2D histograms of (a) $|\lambda_1|$ versus $\Omega = (\bar{\theta}^2 + \bar{\sigma}^2 + \bar{\zeta}^2)^{1/2}$, (b) $|\lambda_2|$ versus Ω , (c) $\|\mathbf{S}\|$ versus Ω , and (d) $\|\mathbf{K}\|$ versus Ω . In (a) the correlation is 0.56; in (b) the correlation is 0.34; in (c) the correlation is 0.57; in (d) the correlation is 0.61. Units for $|\lambda_1|$, $|\lambda_2|$, $\|\mathbf{S}\|$ and $\|\mathbf{K}\|$ on the x -axes are $\text{m}^2 \text{s}^{-1}$. The large-scale flow feature Ω on the y -axes has been rescaled and has arbitrary dimensionless units. The axes are discretised using 200 uniformly spaced bins in each direction. The colorbar represents the number of grid points in each bin.

5 A comparison with Reynolds eddies

Here we consider the transport tensor for eddies defined as the deviation from the time-mean state, which we refer to as *Reynolds eddies*. The temporal Reynolds decomposition for a variable ϕ is

$$\phi(x, y, t) = \langle \phi \rangle(x, y) + \phi^\#(x, y, t), \quad (25)$$

where $\langle \cdot \rangle$ denotes the time mean and $\cdot^\#$ denotes the deviation from the time mean. The corresponding flux-gradient relation for the time-mean eddy tracer flux, is then

$$\langle \mathbf{u}^\# C^\# \rangle_{\text{div}} = -\mathbf{K}^{\text{Reyn}} \nabla \langle C \rangle, \quad (26)$$

where only the divergent part of the flux is retained. The resulting transport tensor $\mathbf{K}^{\text{Reyn}}$ is time-independent, which we deem unsuitable since any coarse-resolution model in which we want to include a parameterisation will have an evolving large-scale circulation.

We show the upper-layer $\mathbf{K}^{\text{Reyn}}$ in figure 13, which appears similar to $\langle \mathbf{K} \rangle$. There are, however, important differences, the first of which being that $\mathbf{K}^{\text{Reyn}}$ is roughly twice the amplitude of $\langle \mathbf{K} \rangle$. Second, K_{22}^{Reyn} is large and positive on the jet core, whereas $\langle K_{22} \rangle$ is large and positive on the jet flanks. Third, although K_{21}^{Reyn} exhibits a cross-jet dipole pattern similar to $\langle K_{21} \rangle$, its behaviour on and near the jet core is much more convoluted. As with the other entries, K_{11}^{Reyn} and $\langle K_{11} \rangle$ have notable differences, again most clear near the jet. For example, while $\langle K_{11} \rangle$ is consistently positive on the jet core, K_{11}^{Reyn} exhibits both positive and negative regions. We show the lower-layer $\mathbf{K}^{\text{Reyn}}$ in figure 14. In the lower layers $\mathbf{K}^{\text{Reyn}}$ and $\langle \mathbf{K} \rangle$ are more qualitatively similar than they are in the upper layer. Qualitative differences are most noticeable in the jet region and near the western boundary. Most notably, however, the amplitude disparity is significant with $\mathbf{K}^{\text{Reyn}}$ at least twice as large as $\langle \mathbf{K} \rangle$ in most of the layer.

As standard, the diffusion ($\mathbf{S}^{\text{Reyn}}$) and advection ($\mathbf{A}^{\text{Reyn}}$) tensors for Reynolds eddies are the symmetric and antisymmetric parts of $\mathbf{K}^{\text{Reyn}}$. In figure 15 we show the eigenvalues of $\mathbf{S}^{\text{Reyn}}$ in the upper layer, which we denote λ_1^{Reyn} and λ_2^{Reyn} . These can be compared with the upper-layer eigenvalues λ_1 and λ_2 of $\mathbf{S}$, as shown in figure 6 (a, b). It is clear that eigenvalue polarity persists for Reynolds eddies so this is a robust property not exclusive to space-filtered eddies. Given the amplitude disparity between $\mathbf{K}^{\text{Reyn}}$ and $\langle \mathbf{K} \rangle$ (and $\mathbf{K}$), it is not surprising that the amplitudes of λ_1^{Reyn} and λ_2^{Reyn} in all layers are larger than the amplitudes of λ_1 and λ_2 .

For Reynolds eddies the diffusion axis (not shown) exhibits behaviour vastly different to the average behaviour shown in figure 10 for space-filtered eddies. For example, in the upper layer where K_{22}^{Reyn} is large and positive across the entire jet region, the diffusion axis is typically between $-\pi/2$ and $-\pi/4$ ($\pi/4$ and $\pi/2$) to the north (south) of the jet core. Because of the convoluted behaviour with $K_{21}^\#$ on the jet core, we do not observe alignment between the diffusion axis and the flow on the jet core. This represents a subtle but important difference in comparison to results for space-filtered eddies.

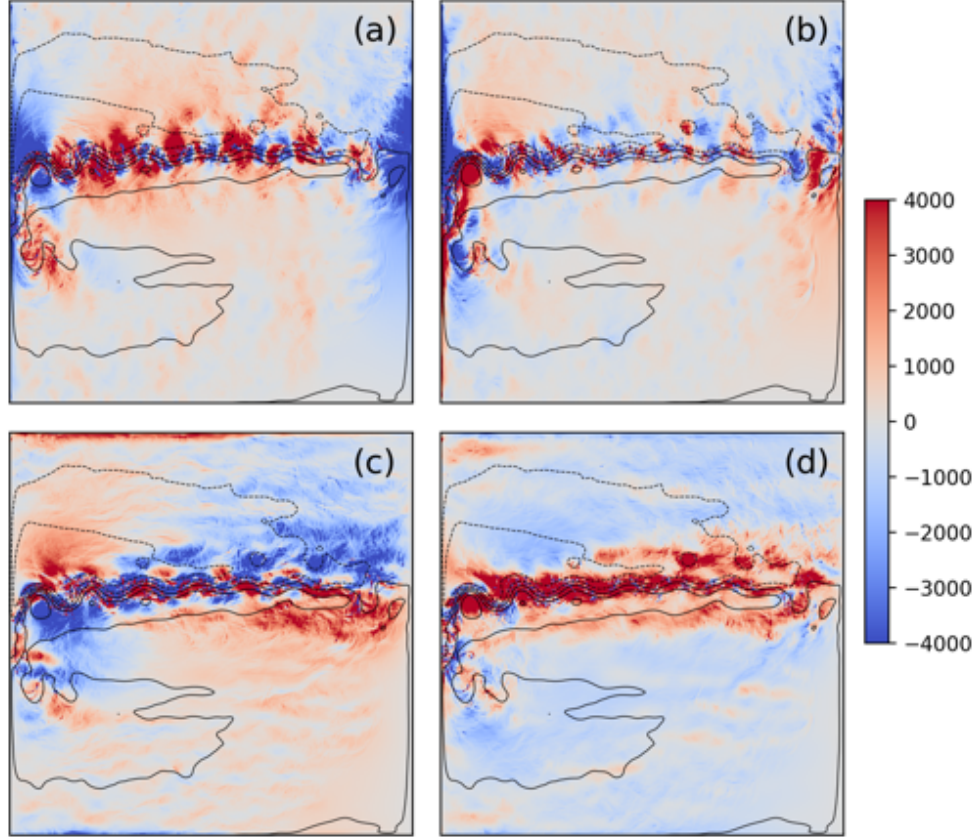


Figure 13: Elements of the transport tensor $\mathbf{K}^{\text{Reyn}}$ in the upper layer for Reynolds eddies. Plotted are (a) K_{11}^{Reyn} , (b) K_{12}^{Reyn} , (c) K_{21}^{Reyn} and (d) K_{22}^{Reyn} . Units are $\text{m}^2 \text{s}^{-1}$. The black solid/dashed lines represent positive/negative contours of the time-mean streamfunction.

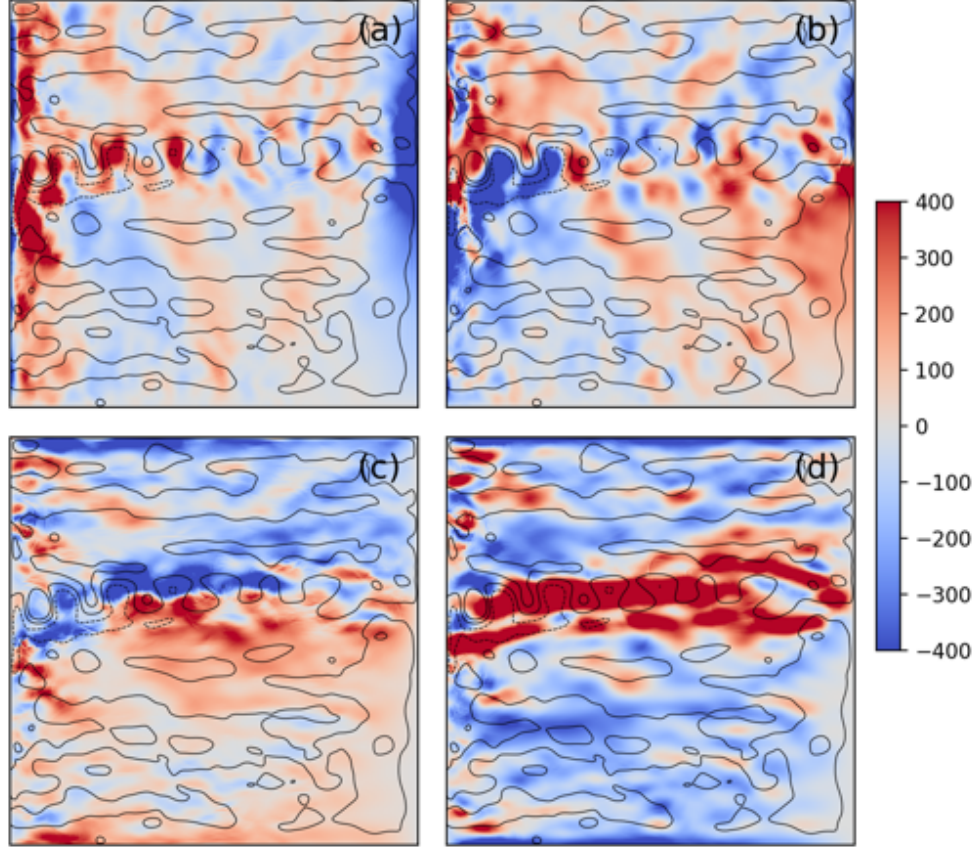


Figure 14: Elements of the transport tensor $\mathbf{K}^{\text{Reyn}}$ in the lower layer for Reynolds eddies. Plotted are (a) K_{11}^{Reyn} , (b) K_{12}^{Reyn} , (c) K_{21}^{Reyn} and (d) K_{22}^{Reyn} . Units are $\text{m}^2 \text{s}^{-1}$. The black solid/dashed lines represent positive/negative contours of the time-mean streamfunction.

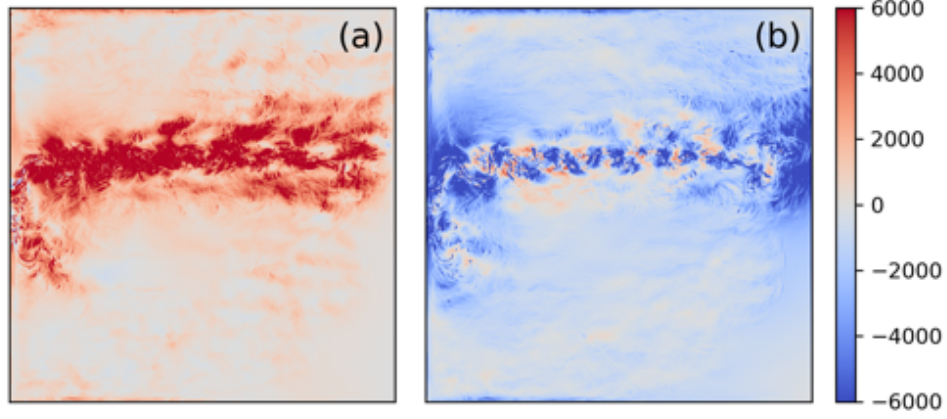


Figure 15: Diffusion eigenvalues (a) λ_1^{Reyn} and (b) λ_2^{Reyn} of the symmetric tensor $\mathcal{S}^{\text{Reyn}}$ in the upper layer. Units are $\text{m}^2 \text{s}^{-1}$.

Although $\langle \mathbf{K} \rangle$ and $\mathbf{K}^{\text{Reyn}}$ can have similar qualitative features, we conclude that they are fundamentally different. The most clear difference is in their amplitudes and in spatial features near the eastward jet and western boundary current. Similar differences in the diffusion eigenvalues were observed by HSSB20, and these simply imply that Reynolds eddies and spatial-filter eddies drive different transport effects. It is then natural to ask which scale separation leads to a transport tensor most suitable for parameterising missing eddy effects in a coarse-resolution ocean model. Given that modellers are computationally limited by spatial resolution, and not temporal resolution, we argue that using a spatial filter is more suitable, and leads to eddy fields more representative of those missing in a coarse-resolution simulation^{*}. In addition, as eddy-permitting ocean models become more common, temporal Reynolds averaging will become more unviable as the separation of timescales between resolved and unresolved flows diminishes [Nadiga, 2008]. Moreover, stationary eddy fields which are particularly strong in the jet region in our simulations will always be captured by the large-scales when Reynolds averaging is used, but, conversely, would not be simulated by coarse-resolution dynamics. Lastly, we reiterate that use of a spatial filter allows us to derive a transport tensor that evolves in time with the large-scale flow, thus retaining maximal information which is otherwise lost with Reynolds averaging.

6 Conclusion

In this study we considered the isopycnal eddy transport tensor $\mathbf{K}$ for passive tracers in an eddy-resolving ocean model. We use a three-layer, double gyre quasigeostrophic (QG) model such that our focus is on lateral eddy transport. Our main motivation is to better understand eddy effects on tracer transport so that

^{*}This is also an argument for not additionally filtering in time; temporal resolution isn't the pressing issue when it comes to simulating mesoscale flows.

we can improve parameterisations of such effects which are missing in coarse-resolution ocean models.

A key step in diagnosing $\mathbf{K}$ is the separation of scales in the tracer and flow fields. Previous studies commonly use a Reynolds average and consequently obtain a transport tensor that lacks full spatio-temporal dependence. Our novel approach involves using a spatial filter to separate the large and small (eddy) scales, which allows us to relate the local large-scale tracer field to the local eddy tracer fluxes via a transport tensor that has full spatio-temporal dependence. Another key aspect of our approach regards the treatment of the eddy tracer flux, $\mathbf{f}$. Some authors [Bachman et al., 2015, 2020] argue that the full eddy tracer flux should be used in order to maintain the theoretical link to parcel excursion theory [Taylor, 1921]. However, since the rotational component of $\mathbf{f}$ is dynamically inert, and since this component dominates the divergent component by two orders of magnitude [Marshall and Shutts, 1981], we opt to remove it. In this study we provided a physical interpretation of $\mathbf{K}$ with a focus on its symmetric part, i.e., the diffusion tensor $\mathbf{S}$. A follow-up paper will provide an analysis of its antisymmetric part, which we call the advection tensor.

The diffusion tensor $\mathbf{S}$ encompasses diffusive fluxes, such that if both its eigenvalues are positive, then the tracer tendency due to $\mathbf{S}$ is guaranteed to transfer variance to smaller scales. We found that the eigenvalues of $\mathbf{S}$ are robustly organised in opposite-signed (polar) pairs which physically represent filamentation in the tracer concentration. Negative eigenvalues in the diffusion tensor are often considered undesirable [Fox-Kemper et al., 2013] since the associated up-gradient fluxes can lead to singularities in the parameterised tracer evolution equation. However, we argue that polar eigenvalue pairs need not be a problem for multiple reasons. First, as in HSSB20 we find that on average the sum of the diffusion eigenvalues is positive, and although the amplitudes of the diffusive flux components depend on the tracer gradients, this indicates a tendency for the down-gradient fluxes to be larger than the up-gradient ones. Second, $\mathbf{S}$ evolves in time, and so negative diffusivities in a specific area may not persist indefinitely. Third, a coarse-resolution tracer transport model may still contain homogeneous diffusion that will prevent the development of singularities.

Our estimates of the diffusion tensor contribute to growing evidence that tracer mixing in the ocean is highly anisotropic and therefore must be parameterised by a tensor rather than a scalar. We have further shown that the components of this tensor have a non-trivial spatial structure, in particular in the jet region where the diagonal elements of $\mathbf{S}$ tend to be positive and where element S_{12} exhibits a distinct dipole pattern. The dipole pattern causes the diffusion axis, which quantifies the direction of preferential mixing, to be on average leant into the shear of the jet. This behaviour exists in all layers, but in the lower layers the orientation is closer to being meridional, in particular under the jet core. This means that at depth there is notable cross-jet mixing whereas in the upper layer the diffusion axis is more likely to be aligned with the jet when near the core. This represents an inhibition of cross-jet mixing [Ferrari and Nikurashin, 2010, Klocker and Abernathey, 2014]. In general areas of the ocean, it was shown that the diffusion axis has a tendency to be aligned with the large-scale velocity vector, or be perpendicular to the large-scale relative vorticity and thickness gradients, but these alignments are not extremely strong. Strong shears inhibit these alignments, but further exploration of this phenomenon is out of the scope of this study.

For the purposes of parameterisation it is necessary to relate the transport tensor to properties of the large-scale flow. We observed a negative correlation between $\mathbf{K}$ and $(\nabla \bar{\mathbf{u}})^T$, strongest in the lower layer and very weak in the upper layer. This implies the approximation $\mathbf{f}_{\text{div}} \propto (\nabla \bar{\mathbf{u}})^T \nabla \bar{C}$ for the eddy tracer flux, similar to the closures proposed for Reynolds stresses [Meneveau and Katz, 2000], eddy vorticity fluxes [Eyink, 2001] and eddy PV fluxes [Nadiga, 2008]. It was also shown that eddy diffusion and transport magnitudes correlate with the large-scale strain and relative vorticity magnitudes. Because these relations are quite weak, and because the entries of $\mathbf{K}$ conditioned on $(\nabla \bar{\mathbf{u}})^T$ are approximately normally distributed, we suggest that a stochastic closure for $\mathbf{K}$ would be most suitable.

This study is a report of the properties of $\mathbf{K}$ for a particular tracer pair. The tensor $\mathbf{K}$ is known to be non-unique with respect the pair choice [Kamenkovich et al., 2020, Sun et al., 2021]. We additionally diagnosed $\mathbf{K}$ for a pair of wave-like tracer fields and found that, although the new $\mathbf{K}$ is not exactly the same, none of our conclusions change. Thus, our main results such as polar eigenvalues, diffusion axis alignment with large-scale flow and correlations between $\mathbf{K}$ and large-scale flow all persist when $\mathbf{K}$ is computed for a different pair of tracers. Our results also show that $\mathbf{K}$ has substantial space-time variability. Our provisional interpretation is that the non-uniqueness and variability may be due to two factors that require further assessment:

1. eddies in the ocean are chaotic (i.e., exhibit sensitive dependence), and even deterministic, time-varying flows can induce *chaotic advection* of passive tracers [Pierrehumbert, 1991]; thus, $\mathbf{K}$ is a stochastic object subject to different realisations with different eddy and tracer fields.
2. the local flux-gradient relation (9) may not be a wholly general representation of tracer fluxes (e.g., there could be dependencies in the local tracer curvature field, $\nabla \nabla \bar{C}$, or the tracer history); thus, $\mathbf{K}$ is an *incomplete model* of eddy transport, in spite of its ability to locally fit any pair of independent tracers.

A tensor diffusivity representation of eddy fluxes should be viewed as sampled from a random distribution expressing the underlying chaos and/or incompleteness. This would not contradict the view that the eddy-dynamical process that generates the $\mathbf{K}$ distribution is itself independent of C . An interesting question is how similar the resulting distributions are for different C pairs.

Acknowledgements

MH and LS are supported by the NERC grant NE/T002220/1. PB gratefully acknowledges support by the NERC grants NE/R011567/1 and NE/T002220/1, the Leverhulme grant RPG-2019-024 and the Moscow Center for Fundamental and Applied Mathematics (supported by the Agreement of 075-15-2019-1624 with the Ministry of Education and Science of the Russian Federation). JM worked pro bono.

References

- R. Abernathey, J. Marshall, M. Mazloff, and E. Shuckburgh. Enhancement of mesoscale eddy stirring at steering levels in the southern ocean. *J. Phys. Oceanogr.*, 40(1):170–184, 2010. doi:[10.1175/2009JPO4201.1](https://doi.org/10.1175/2009JPO4201.1).
- R. Abernathey, D. Ferreira, and A. Klocker. Diagnostics of isopycnal mixing in a circumpolar channel. *Ocean Modell.*, 72:1 – 16, 2013. ISSN 1463-5003. doi:[10.1016/j.ocemod.2013.07.004](https://doi.org/10.1016/j.ocemod.2013.07.004).
- J.A. Anstey and L. Zanna. A deformation-based parametrization of ocean mesoscale eddy reynolds stresses. *Ocean Modelling*, 112:99 – 111, 2017. ISSN 1463-5003. doi:<https://doi.org/10.1016/j.ocemod.2017.02.004>.
- S.D Bachman and B. Fox-Kemper. Eddy parameterization challenge suite i: Eady spindown. *Ocean Modell.*, 64:12 – 28, 2013. ISSN 1463-5003. doi:[10.1016/j.ocemod.2012.12.003](https://doi.org/10.1016/j.ocemod.2012.12.003).
- S.D. Bachman, B. Fox-Kemper, and F.O. Bryan. A tracer-based inversion method for diagnosing eddy-induced diffusivity and advection. *Ocean Modell.*, 86:1 – 14, 2015. ISSN 1463-5003. doi:[10.1016/j.ocemod.2014.11.006](https://doi.org/10.1016/j.ocemod.2014.11.006).
- S.D. Bachman, B. Fox-Kemper, and B. Pearson. A scale-aware subgrid model for quasi-geostrophic turbulence. *Journal of Geophysical Research: Oceans*, 122(2):1529–1554, 2017a. doi:<https://doi.org/10.1002/2016JC012265>.
- S.D. Bachman, D.P. Marshall, J.R. Maddison, and J. Mak. Evaluation of a scalar eddy transport coefficient based on geometric constraints. *Ocean Modell.*, 109:44 – 54, 2017b. ISSN 1463-5003. doi:[10.1016/j.ocemod.2016.12.004](https://doi.org/10.1016/j.ocemod.2016.12.004).
- S.D. Bachman, B. Fox-Kemper, and F.O. Bryan. A diagnosis of anisotropic eddy diffusion from a high-resolution global ocean model. *Journal of Advances in Modeling Earth Systems*, 12(2):e2019MS001904, 2020. doi:[10.1029/2019MS001904](https://doi.org/10.1029/2019MS001904).
- P. Berloff. Dynamically consistent parameterization of mesoscale eddies. Part I: Simple model. *Ocean Modelling*, 87:1 – 19, 2015. ISSN 1463-5003. doi:<https://doi.org/10.1016/j.ocemod.2014.12.008>.
- P. Berloff. Dynamically consistent parameterization of mesoscale eddies. Part III: Deterministic approach. *Ocean Modelling*, 127:1 – 15, 2018. ISSN 1463-5003. doi:<https://doi.org/10.1016/j.ocemod.2018.04.009>.
- P.S. Berloff and J.C. McWilliams. Quasigeostrophic dynamics of the western boundary current. *Journal of Physical Oceanography*, 29(10):2607–2634, 1999. doi:[10.1175/1520-0485\(1999\)029<2607:QDOTWB>2.0.CO;2](https://doi.org/10.1175/1520-0485(1999)029<2607:QDOTWB>2.0.CO;2).
- P.S. Berloff, J.C. McWilliams, and A. Bracco. Material Transport in Oceanic Gyres. Part I: Phenomenology. *J. Phys. Oceanogr.*, 32(3):764–796, 2002. doi:[10.1175/1520-0485\(2002\)032<0764:MTIOGP>2.0.CO;2](https://doi.org/10.1175/1520-0485(2002)032<0764:MTIOGP>2.0.CO;2).

- T. Bolton and L. Zanna. Applications of deep learning to ocean data inference and sub-grid parameterization. *Journal of Advances in Modeling Earth Systems*, 11(1):376–399, 2019. doi:<https://doi.org/10.1029/2018MS001472>.
- C. Eden. Thickness diffusivity in the southern ocean. *Geophysical Research Letters*, 33(11), 2006. doi:[10.1029/2006GL026157](https://doi.org/10.1029/2006GL026157).
- C. Eden. Anisotropic rotational and isotropic residual isopycnal mesoscale eddy fluxes. *J. Phys. Oceanogr.*, 40(11):2511–2524, 2010. doi:[10.1175/2010JPO4397.1](https://doi.org/10.1175/2010JPO4397.1).
- C. Eden and R.J. Greatbatch. A diagnosis of isopycnal mixing by mesoscale eddies. *Ocean Modell.*, 27(1):98 – 106, 2009. ISSN 1463-5003. doi:[10.1016/j.ocemod.2008.12.002](https://doi.org/10.1016/j.ocemod.2008.12.002).
- C. Eden, R.J. Greatbatch, and D. Olbers. Interpreting eddy fluxes. *J. Phys. Oceanogr.*, 37(5):1282–1296, 2007. doi:[10.1175/JPO3050.1](https://doi.org/10.1175/JPO3050.1).
- G.L. Eyink. Dissipation in turbulent solutions of 2d euler equations. *Nonlinearity*, 14(4):787–802, may 2001. doi:[10.1088/0951-7715/14/4/307](https://doi.org/10.1088/0951-7715/14/4/307).
- R. Ferrari and M. Nikurashin. Suppression of eddy diffusivity across jets in the southern ocean. *J. Phys. Oceanogr.*, 40(7):1501–1519, 2010. doi:[10.1175/2010JPO4278.1](https://doi.org/10.1175/2010JPO4278.1).
- B. Fox-Kemper and D. Menemenlis. *Can Large Eddy Simulation Techniques Improve Mesoscale Rich Ocean Models?*, pages 319–337. American Geophysical Union (AGU), 2008. ISBN 9781118666432. doi:<https://doi.org/10.1029/177GM19>.
- B. Fox-Kemper, R. Ferrari, and J. Pedlosky. On the indeterminacy of rotational and divergent eddy fluxes. *J. Phys. Oceanogr.*, 33(2):478–483, 2003. doi:[10.1175/1520-0485\(2003\)033<0478:OTIORA>2.0.CO;2](https://doi.org/10.1175/1520-0485(2003)033<0478:OTIORA>2.0.CO;2).
- B. Fox-Kemper, R. Lumpkin, and F.O. Bryan. Chapter 8 - lateral transport in the ocean interior. In Gerold Siedler, Stephen M. Griffies, John Gould, and John A. Church, editors, *Ocean Circulation and Climate*, volume 103 of *International Geophysics*, pages 185 – 209. Academic Press, 2013. doi:<https://doi.org/10.1016/B978-0-12-391851-2.00008-8>.
- S.M. Griffies. The gent–mcwilliams skew flux. *J. Phys. Oceanogr.*, 28(5):831–841, 1998. doi:[10.1175/1520-0485\(1998\)028<0831:TGMSF>2.0.CO;2](https://doi.org/10.1175/1520-0485(1998)028<0831:TGMSF>2.0.CO;2).
- M. Haigh, L. Sun, I. Shevchenko, and P. Berloff. Tracer-based estimates of eddy-induced diffusivities. *Deep Sea Research Part I: Oceanographic Research Papers*, page 103264, 2020. ISSN 0967-0637. doi:<https://doi.org/10.1016/j.dsr.2020.103264>.
- S. W. Jones and W. R. Young. Shear dispersion and anomalous diffusion by chaotic advection. *Journal of Fluid Mechanics*, 280:149–172, 1994. doi:[10.1017/S0022112094002880](https://doi.org/10.1017/S0022112094002880).

- I. Kamenkovich, P.S. Berloff, and J. Pedlosky. Anisotropic material transport by eddies and eddy-driven currents in a model of the north atlantic. *J. Phys. Oceanogr.*, 39(12):3162–3175, 12 2009. ISSN 0022-3670. doi:[10.1175/2009JPO4239.1](https://doi.org/10.1175/2009JPO4239.1).
- I. Kamenkovich, I.I. Rypina, and P.S. Berloff. Properties and origins of the anisotropic eddy-induced transport in the north atlantic. *J. Phys. Oceanogr.*, 45(3):778–791, 2015. doi:[10.1175/JPO-D-14-0164.1](https://doi.org/10.1175/JPO-D-14-0164.1).
- I. Kamenkovich, P. Berloff, M. Haigh, L. Sun, and Y. Lu. Complexity of mesoscale eddy diffusivity in the ocean. *Geophysical Research Letters*, n/a(n/a):e2020GL091719, 2020. doi:<https://doi.org/10.1029/2020GL091719>. e2020GL091719 2020GL091719.
- S.A. Karabasov, P.S. Berloff, and V.M. Goloviznin. Cabaret in the ocean gyres. *Ocean Modell.*, 30(2):155 – 168, 2009. ISSN 1463-5003. doi:[10.1016/j.ocemod.2009.06.009](https://doi.org/10.1016/j.ocemod.2009.06.009).
- A. Klocker and R. Abernathey. Global patterns of mesoscale eddy properties and diffusivities. *J. Phys. Oceanogr.*, 44(3):1030–1046, 2014. doi:[10.1175/JPO-D-13-0159.1](https://doi.org/10.1175/JPO-D-13-0159.1).
- A. Klocker, R. Ferrari, and J. LaCasce. Estimating suppression of eddy mixing by mean flows. *J. Phys. Oceanogr.*, 42(9):1566–1576, 2012a. doi:[10.1175/JPO-D-11-0205.1](https://doi.org/10.1175/JPO-D-11-0205.1).
- A. Klocker, R. Ferrari, J. LaCasce, and S. Merrifield. Reconciling float-based and tracer-based estimates of eddy diffusivities in the southern ocean. *J. Marine Res.*, 70:569–602, 07 2012b. doi:[10.1357/002224012805262743](https://doi.org/10.1357/002224012805262743).
- N.C. Lau and John M. Wallace. On the Distribution of Horizontal Transports by Transient Eddies in the Northern Hemisphere Wintertime Circulation’. *Journal of Atmospheric Sciences*, 36(10):1844 – 1861, 1979. doi:[10.1175/1520-0469\(1979\)036<1844:OTDOHT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<1844:OTDOHT>2.0.CO;2).
- J.R. Ledwell, A.J. Watson, and C.S. Law. Mixing of a tracer in the pycnocline. *Journal of Geophysical Research: Oceans*, 103(C10):21499–21529, 1998. ISSN 0148-0227. doi:[10.1029/98JC01738](https://doi.org/10.1029/98JC01738).
- A. Leonard. Energy cascade in large-eddy simulations of turbulent fluid flows. In F.N. Frenkiel and R.E. Munn, editors, *Turbulent Diffusion in Environmental Pollution*, volume 18 of *Advances in Geophysics*, pages 237 – 248. Elsevier, 1975. doi:[https://doi.org/10.1016/S0065-2687\(08\)60464-1](https://doi.org/10.1016/S0065-2687(08)60464-1).
- D.K. Lilly. The representation of small-scale turbulence in numerical simulation experiments. *Proceedings of the IBM Scientific Computing Symposium on Environmental Sciences*, pages 195–201, 1967.
- J. Lu, F. Wang, H. Liu, and P. Lin. Stationary mesoscale eddies, upgradient eddy fluxes, and the anisotropy of eddy diffusivity. *Geophys. Res. Letters*, 43(2):743–751, 2016. doi:[10.1002/2015GL067384](https://doi.org/10.1002/2015GL067384).
- J.R. Maddison, D.P. Marshall, and J. Shipton. On the dynamical influence of ocean eddy potential vorticity fluxes. *Ocean Modell.*, 92:169 – 182, 2015. ISSN 1463-5003. doi:[10.1016/j.ocemod.2015.06.003](https://doi.org/10.1016/j.ocemod.2015.06.003).

- 682 J. Marshall and G. Shutts. A note on rotational and divergent eddy fluxes. *J. Phys. Oceanogr.*, 11(12):
683 1677–1680, 1981. doi:[10.1175/1520-0485\(1981\)011<1677:ANORAD>2.0.CO;2](https://doi.org/10.1175/1520-0485(1981)011<1677:ANORAD>2.0.CO;2).
- 684 J. Marshall, E. Shuckburgh, H. Jones, and C. Hill. Estimates and implications of surface eddy diffusiv-
685 ity in the southern ocean derived from tracer transport. *J. Phys. Oceanogr.*, 36(9):1806–1821, 2006.
686 doi:[10.1175/JPO2949.1](https://doi.org/10.1175/JPO2949.1).
- 687 A.S. Medvedev and R.J. Greatbatch. On advection and diffusion in the mesosphere and lower thermosphere:
688 The role of rotational fluxes. *J. Geophys. Res.: Atmos.*, 109(D7), 2004. doi:[10.1029/2003JD003931](https://doi.org/10.1029/2003JD003931).
- 689 C. Meneveau and J. Katz. Scale-invariance and turbulence models for large-eddy simulation. *Annual Review*
690 *of Fluid Mechanics*, 32(1):1–32, 2000. doi:[10.1146/annurev.fluid.32.1.1](https://doi.org/10.1146/annurev.fluid.32.1.1).
- 691 B.T. Nadiga. Orientation of eddy fluxes in geostrophic turbulence. *Philosophical Transactions of*
692 *the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1875):2489–2508, 2008.
693 doi:[10.1098/rsta.2008.0058](https://doi.org/10.1098/rsta.2008.0058).
- 694 R. T. Pierrehumbert. Chaotic mixing of tracer and vorticity by modulated travelling rossby waves. *Geo-*
695 *physical & Astrophysical Fluid Dynamics*, 58(1-4):285–319, 1991. doi:[10.1080/03091929108227343](https://doi.org/10.1080/03091929108227343).
- 696 I.I. Rypina, I. Kamenkovich, P.S. Berloff, and L.J. Pratt. Eddy-induced particle dispersion in the near-surface
697 north atlantic. *J. Phys. Oceanogr.*, 42(12):2206–2228, 2012. doi:[10.1175/JPO-D-11-0191.1](https://doi.org/10.1175/JPO-D-11-0191.1).
- 698 I. Shevchenko and P. Berloff. Eddy backscatter and counter-rotating gyre anomalies of midlatitude ocean
699 dynamics. *Fluids*, 1(28), 2016. doi:[10.3390/fluids1030028](https://doi.org/10.3390/fluids1030028).
- 700 I.V. Shevchenko and P.S. Berloff. Multi-layer quasi-geostrophic ocean dynamics in eddy-resolving regimes.
701 *Ocean Modelling*, 94:1 – 14, 2015. ISSN 1463-5003. doi:<https://doi.org/10.1016/j.ocemod.2015.07.018>.
- 702 K.S. Smith. Tracer transport along and across coherent jets in two-dimensional turbulent flow. *Journal of*
703 *Fluid Mechanics*, 544:133–142, 2005. doi:[10.1017/S0022112005006750](https://doi.org/10.1017/S0022112005006750).
- 704 Z. Stanley, S. D. Bachman, and I. Grooms. Vertical structure of ocean mesoscale eddies with
705 implications for parameterizations of tracer transport. *Journal of Advances in Modeling Earth*
706 *Systems*, 12(10):e2020MS002151, 2020. doi:<https://doi.org/10.1029/2020MS002151>. e2020MS002151
707 10.1029/2020MS002151.
- 708 L. Sun, M. Haigh, I. Shevchenko, and P. Berloff. On non-uniqueness of the mesoscale eddy diffusivity.
709 *Submitted to J. Fluid Mech.*, 2021.
- 710 P.N. Swarztrauber. A direct method for the discrete solution of separable elliptic equations. *SIAM Journal*
711 *on Numerical Analysis*, 11(6):1136–1150, 1974. ISSN 00361429.

- 712 G. I. Taylor. Diffusion by continuous movements. *Proc. London Math. Soc.*, s2-20(1):196–212, 1921.
713 doi:[10.1112/plms/s2-20.1.196](https://doi.org/10.1112/plms/s2-20.1.196).
- 714 Y.K. Ying, J.R. Maddison, and J. Vanneste. Bayesian inference of ocean diffusivity
715 from lagrangian trajectory data. *Ocean Modelling*, 140:101401, 2019. ISSN 1463-5003.
716 doi:<https://doi.org/10.1016/j.ocemod.2019.101401>.
- 717 W. R. Young, P. B. Rhines, and C. J. R. Garrett. Shear-flow dispersion, internal waves and horizon-
718 tal mixing in the ocean. *Journal of Physical Oceanography*, 12(6):515–527, 1982. doi:[10.1175/1520-0485\(1982\)012<0515:SFDIWA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1982)012<0515:SFDIWA>2.0.CO;2).
719
- 720 L. Zanna and T. Bolton. Data-driven equation discovery of ocean mesoscale closures. *Geophysical Re-*
721 *search Letters*, 47(17):e2020GL088376, 2020. doi:<https://doi.org/10.1029/2020GL088376>, e2020GL088376
722 10.1029/2020GL088376.
- 723 V. Zhurbas and I.S. Oh. Drifter-derived maps of lateral diffusivity in the pacific and atlantic oceans in relation
724 to surface circulation patterns. *J. Geophys. Res.: Oceans*, 109(C5), 2004. doi:[10.1029/2003JC002241](https://doi.org/10.1029/2003JC002241).
- 725 V. Zhurbas, D. Lyzhkov, and N. Kuzmina. Drifter-derived estimates of lateral eddy diffusiv-
726 ity in the world ocean with emphasis on the indian ocean and problems of parameterisation.
727 *Deep Sea Research Part I: Oceanographic Research Papers*, 83:1 – 11, 2014. ISSN 0967-0637.
728 doi:<https://doi.org/10.1016/j.dsr.2013.09.001>.