Overall comment:

Note that Fig. 10 and some values in sections 5 and 6 have slightly changed due to using an older version of a data file. Now all time scales, diffusivities and R(0) are consistent throughout the paper. Note also that for better readability we have divided Section 3 (Theory and Methods) into subsections. For further clarity we have also reorganized Section 5 (The effects of the mean flow on eddy diffusivities) with respect to subsections.

Reviewer 2

We thank reviewer 2 for their very helpful comments, suggestions and raised questions. Please find our point-by-point responses to your comments below. Reviewer comments are shown in black and our responses are shown in blue. Edited text in the manuscript is shown in purple. Unchanged text that has been copied in for completeness is shown in gray.

Analyzing the drifter and numerical Lagrangian data, this manuscript studies the horizontal 2D diffusivity tensor with a particular focus on the scale effect.

Also, it is interesting to know the departure of the single-and pair-diffusivity relation from the theory.

However, I am confused about issues in the manuscript, such as the definition of average.
We agree with the reviewer that the definition of average should be clarified in this study. To define the background mean flow we use a time mean, where the time interval is one year, which is larger than the typical eddy timescales in the study region. A certain amount of spatial coarse-graining is also implied since in the observations the time mean is based on 1/3° resolution (OSCAR product) and in the simulation the time mean is based on 1/10° resolution (POP simulation).

We have added a more specific description to the definition of the background mean flow in L52-53 in the introduction and also added further clarifications on the background mean flow in the methods sections (L207-211). A table (Table 1) summarizing the description of the different flow components used in this study was added.

L50-51: ... where the overline denotes an average over intervals in time and space that are larger than typical eddy scales, and ' the deviation from that average, while Q describes the sources and sinks of the tracer T.

L210-214: In the observational drifter data set $\overline{U}$ is the time mean velocity of the 1/3° OSCAR surface currents interpolated to the drifter locations, while in the POP simulation $\overline{U}$ is the time mean of the 1/10° Eulerian currents interpolated to the numerical drifter locations (Table 1). Note the current altimeter-derived products, such as the OSCAR product, could have effective resolutions in space and time more close to about 100 km and 30 days (Ballarotta et al. 2019) and thereby also imply a spatial coarse-graining.

I will recommend this manuscript's publication after addressing the following questions and comments.

1) Definition of mean.
Please see response above to the definition of averages used in this study.

Line 113: " ... since it should be independent from the mean flow"
This statement depends on what "mean" means. If a time average to decompose roughly the internal waves and mesoscale eddies then the mesoscale mean flow also introduces pair diffusion.
We agree that clarification for the averaging and mean definition is needed. Please see responses

above to the definition of the background mean flow. We have edited the research question (L116-118) to clarify that the background mean definition does not resolve the motions induced by the eddies.

L116-118: How do the diffusivities depend on eddy-mean flow decompositions considering both single and pair particle statistics, and can the pair particle diffusivity emerge as an alternative to the single particle diffusivity since it should be independent from the mean flow that does not resolve the motions induced by the eddies?

Equation (12): I am confused about the definition of U. A five-day average is used?
For the OSCAR product yes, for the POP simulation the resolution is daily. We have clarified this information by adding a table (Table 1) that states the spatial and temporal resolution of products used.

In parameterizations, Eulerian spatial average should be considered, this manuscript uses time mean of Lagrangian data. They may not have a direct link.
We use the time mean of Eulerian data together with the Lagrangian trajectories that provide space-time information. The theoretical underpinning to estimate Lagrangian diffusivities in a spatially inhomogeneous mean flow and the connection to the Reynolds-averaged tracer transport equation (1) was developed by Davis (1987, 1990). We have already pointed this out in the introduction (L95-99).

L96-100: Davis (1987) and Davis (1991) devised the underlying theory how to compute the diffusivity tensor in principle in the presence of an inhomogeneous background mean flow where, instead of diagnosing the statistics from the absolute Lagrangian velocities and displacements, diffusivities are computed from residual velocities and displacements after the Eulerian mean has been subtracted.

For further clarification we have added the following lines in the methods section:
L200-201: In that sense, the diffusivity introduced by Davis (1987) is a mixed Eulerian-Lagrangian quantity that can directly be related to the eddy tracer fluxes in a diffusive parameterization, as derived in Davis (1987).

With respect to the Eulerian spatial averaging, please see our response on the definition of the background mean flow above.

Line 506: only time average is used, right? Then how can we know the spatial information of the deformation radius? This looks like a pure conjecture.
We use a time average, but spatial information is also included (please see our response on the definition of the background mean flow above). To avoid confusion we have changed the sentence which now reads as:
L591-594: This relates to the research question, which addresses the role of smaller scale motions, which are detected by drifters but not by current altimeter products in influencing diffusivities and anisotropy, the study highlights the significant contribution of these unresolved motions.

2) Inhomogeneity influences the diffusivity a lot, especially for the current Lagrangian data, which captures information of different locations at different times. Is there any way to check the homogeneity of the residual velocity field?
We agree that inhomogeneity of the residual velocity field could influence our diffusivity calculations. To address this, we applied a bootstrapping method to ensure the robustness of our results. This approach allows us to capture any variations of the residual velocities by including them in the error calculations. Therefore, even if there are inhomogeneities in the residual velocities, they are reflected in the resulting error estimates, providing a reliable measure of robustness despite any underlying

variability. In principle, one could test the homogeneity assumption by dividing the data into spatial bins and recalculating the diffusivity for these bins to obtain the spatial distribution in diffusivities as in Griesel et al. (2010, 2014). However, there is not enough observational data for binning.

3) Figure 3: It is strange to observe that the total-mean does not equal eddy?
It is not strange because in the observations the difference total-mean are the small scale processes not resolved by the OSCAR product. To improve clarity we have now changed the figure labels. The label 'eddy only' for the observations was changed to u_{OSCAR}. In addition, the different velocity components are explained in Table 1 for consistency and clarity.

4) Line 417: What about the parameterization effect, such as the unavoidable numerical dissipation, that enters the difference between measurement and numerics?
We believe that numerical dissipation and diffusion do not impact our results. As we show in Fig. 2 the velocity fields from the POP simulation compare well with the velocity fields from the OSCAR product (L289). In addition we added a new Fig. 3 that compares the kinetic energy levels, which also match(L290-293). The numerical drifters in the POP simulation are advected and not diffused, so the overall impact of numerical diffusion on eddy diffusivity is negligible.

L288: Regardless, as Fig. 2 illustrates, the horizontal distributions of kinetic energies are similar.

L288-293: There is visible interannual variability for the region (Fig. 3a) that is largely driven by the variability of the Agulhas rings. While there are discrepancies between OSCAR and POP particularly around 2004, the area-averaged energy levels in 1996 in POP (red solid and dashed lines in Fig. 3a) and in 2016 for the OSCAR product (black solid and dashed lines in Fig. 3a) are similar, and so are the total kinetic energies of POP and OSCAR averaged along the drifter locations (red and cyan diamonds in Fig. 3a respectively).