Review of reviewer #3, Dr Deepak Cherian

This paper is a fairly thorough look at microstructure measurements from a small number of profiles (approx 10) spanning a transect across the Atlantic Ocean. The amount of analysis done is impressive, and the methodological summaries are a useful resource for anyone new to the field. However, having read the paper twice now, I have some major comments; and recommend major revisions.

Thanks for reading the paper thoroughly. Your feedback is much appreciated and we have used it to improve the paper as we explain below. Our comments are in blue.

1. A lot of effort is spent in explaining the methods, so much so that it is hard to draw strong conclusions. Some of the shear-strain ratio analysis, and the less compelling Thorpe scale based estimates might be better suited to supplementary material. The discussion section needs some context on how these results might be novel; or if not, how they support an existing body of work in to weak turbulence regimes. In general, drawing strong conclusions from a small number of turbulence profiles is hard; so it might be useful to bring in previous datasets from the NSF microstructure database to provide extra context.
   We have added some context in the introduction how our study differs from existing comparisons (lines 61-65). Additionally, we have added more discussion placing our results in context of existing work in the Section 8, lines 533-538, and 548-554.
   We do acknowledge that the amount of data is relatively limited (lines 555-559) and that it is a caveat in our study. However, we don't think bringing in more datasets will change the conclusions and focus of this paper. The focus of this paper is 1) to compare thermistor data and shear data in low turbulence regions, 2) compare strain rate and Thorpe estimates directly to observed profiles, and 3) understand if the triple decomposition can give insights for interpretation. We stressed these points more in the introduction.
   These main results do not necessarily require more data and compiling a large set of turbulence measurements could be a study by itself. In addition, those datasets will either be in different locations, or years or even decades apart, and thus not likely to improve the results from this dataset, by averaging between profiles. We therefore want to stay with a comprehensible number of profiles so we can study individual differences. We also added the following note to lines 559-561: *"Adding more profiles will undoubtedly lead to better comparisons between the direct microstructure observations and the indirect parameterizations, but it seems unlikely that it will change the key conclusions of this paper."*

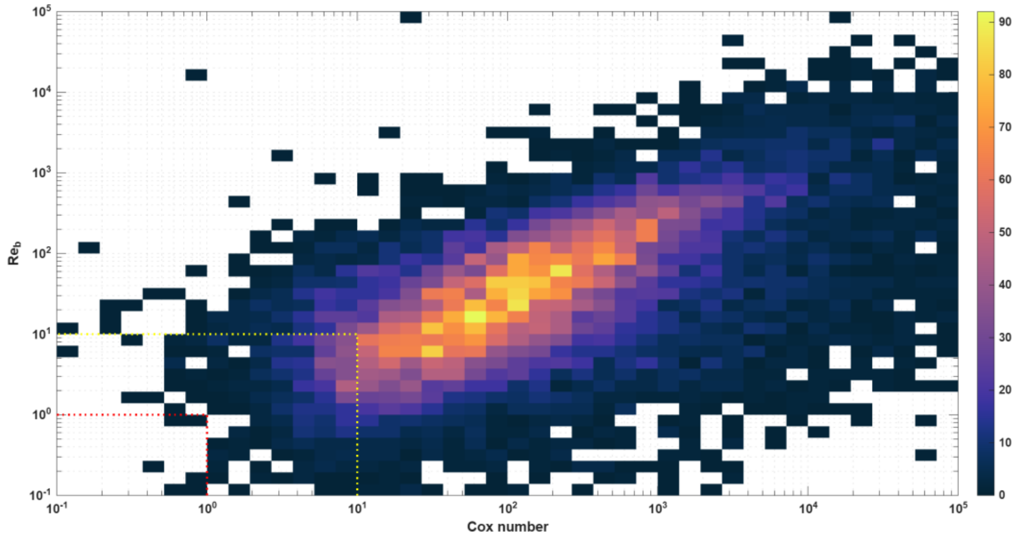
2. Much of the paper focuses on very low ε & χ values. It is not clear to me how the noise floor of the thermistor is being handled. At some point, there is no "overturning turbulence", and (to me) the only sensible thing to do is apply molecular diffusivities and set ε to 0. See Cherian et al (2020) for references (including Gregg et al 2012; Itsweire et al 1988). If you "discard" too many of these low turbulence segments, you will bias the mean. From the text it doesn't appear as if this is actually biasing the mean, but it would be good to be sure and to discuss the work of Itsweire and Gregg here.

   The reviewer makes a few points that we would like to address separately.

   First, the thermistor noise floor is discussed in lines 159-165 of the revised paper: *"In our processing routines we discard any data segment where the average spectral levels fall below a factor of 1.3 (Piccolroaz2021) of the well-characterised sensor noise curve (following RSI, see Technical Note 40 available at www.rocklandscientific.com). Only 0.49% of data-segments were discarded following that criterion. For the retained data, the distributions of ε_{\mu T} do not show a strong deviation from the expected log-normal distribution (Fig. 3a,c), as found for the shear sensor, suggesting that these estimates are still reliable and above the noise floor."*

   So based on the low amount of segments being discarded following this criterion, plus the distributions seen in Fig 3, we argue that the noise floor of the thermistors is not reached for the data that is used. So it is unlikely biasing the mean by discarding too many segments, which would be the case if only the shear that is used.

   Besides the handling of the data, we argue that the noise floor of the thermistors is not reached for the data that is used. At least, it doesn't show as a clear cutoff in the distribution of the data (Fig 3), while a cutoff is found for the shear data.



Second, we agree that at some point there is no overturning turbulence anymore. However, as can be seen in Fig 8, the diffusivities are still multiple orders above the molecular levels (smallest diffusivities are O(1e-6), but most of the time larger, whereas typical values of the molecular thermal diffusivity is O(1e-7)), and thus setting epsilon to zero does not make sense here.

A further analysis based on the buoyancy Reynolds number and the Cox number confirm this (see Figure). With only a minor part of the estimates having both the Reb and Cz below 10 (~6%). It is unlikely to significantly impact our results.

The dissipation rate and diffusivity would be biased high if we only used the shear probes (affected by noise floor). We believe that our approach of using a combined profile of shear and thermistor based dissipation estimates avoids the introduction of such bias, by using the most accurate and direct estimate available.

The "triple decomposition" analysis in Section 6 doesn't add much to the paper, and could be removed in my opinion. I completely agree with the conclusion (lines 510-520) that it is not useful over much of the data. But that is to be expected from an analysis method that relies on residuals. Combining that with uncertainties in estimating the very small isopyncal density gradient in the abyss, the resulting diffusivity values cannot be good. A formal uncertainty analysis should indicate that your residuals are indistinguishable from 0, but isn't worth the effort in my opinion. In my experience, it is quite hard to extract the signal even from very well sampled measurement campaigns like NATRE. Outside of spicy isopycnals, where there is very low isopyncal stirring, you will simply get junk if attempting to estimate a residual (this is easily visble in Ferrari and Polzin, 2005). If you'd really like to do the comparison, I'd recommend limiting yourself to regions where "spiciness" is large,

and avoiding diffusivity and comparing χ instead (if the χ residual proves to be significant).
We agree with the reviewer that this part of the manuscript can be improved. It is indeed indicated that the isopycnal diffusivity estimates don't have much value due to the highly variable results (l. 786-790). Therefore we moved this part about the isoneutral diffusivities to the appendices, so that an interested reader can still find it, but it is no longer part of the main text. We find it useful to keep these results to show that it is not useful in this situation and provide arguments why. However, we agree this does not fit the main results.
We have left Section 6 about the triple decomposition in place, as it supports the arguments made in other sections and the phenomena seen (l.469-486). These arguments pertain specifically to the areas where the triple decomposition does have a large signal, and thus the decomposition can be used.
Hopefully the reviewer can agree with this compromise.