Reviewer 2, Second review

Dear Reviewer, thank you for your feedback. Please find below in blue, our replies to your comments.

The authors have provided some answers to the questions posed, and have made a few changes to the manuscript. I must say, however, that I am still unsatisfied with a few points that I noted in my first review and which, in my view, need to be clarified:

When I asked to clarify the choice made for the order of integration of the variables Ubar and eta with the modified FB (SE) scheme and the modified AB3-AM4 (SESM) scheme, I was not asking about the order of convergence, but about the order of integration of the two variables, i.e integrating Ubar before eta (which allows to use the newly available value for Ubar in eta integration) or conversely integrating eta before Ubar (which allows to use the newly available value for eta in Ubar integration). It has been noted that the choice of integrating Ubar before eta with the SE scheme is for historical reasons. However, it seems to me that the question of considering the inverse choice arises : this would remove the term involving $< Ubar^n + 1 >$ and $< Ubar^n >$ of the right-hand side of equation 12 : this is possibly a good thing, because as the authors note, the initialization for $< Ubar^n >$ is not well constrained (l124-126).

When we replied that the integration of $\langle\langle\overline{\mathbf{U}}\rangle\rangle$ in SE is done first for historical reason, we could have provided another answer that we just followed the original paper by Demange et al. (2019), as also mentioned in the footnote 2 of our manuscript. It is possible to do either way but that should not change the net outcome. The reviewer is right that the reverse way of integration is also possible, but it will lead to identical damping. The new equation equivalent to (12) will be simpler, but it will not lead to any practical advantages because $\langle\langle\overline{\mathbf{U}}\rangle\rangle$ is accumulated by summing the velocities appearing at the rhs of equation (8). Also the the last statement by the reviewer is a misunderstanding: we wrote that $\overline{\mathbf{U}}^n$ is not well constrained meaning that 3D transports are known at semi-integer time steps, so there is no uniquely defined way to reconcile 3d and barotropic transport. The solution we proposed is one way of reconciling and there is no issue with $\overline{\mathbf{U}}^n$. In addition, even though $\overline{\mathbf{U}}^n$ will be absent from an expression in place of (12), all velocities in that expression still depend on it. In summary, both ways of implementing SE (used by us and in Demange et al. (2019), as well as the one proposed by the reviewer) are possible and either can be followed as both will result in identical damping.

I asked to clarify the correction of the 3d variables after completion of the barotropic integration, because I had suspected that this implementation might be missing one of the two constraints which link the 2d and 3d dynamics, and which classically appear in split-explicit algorithms. As pointed out by the authors, it has been considered to correct the 3d fluxes (after having advected tracers) to be consistent with the instantaneous barotropic flux at n+1/2, but that this correction seems unnecessary (l151-155). Why not instead correct the 3d fluxes with the half sum of $< Ubar^n + 1 >$ and $< Ubar^n >$?

There are three possibilities of imposing the 'second correction' in our case. The first one is just to leave the 3d transports corrected with $\langle\langle \overline{\mathbf{U}} \rangle\rangle$ (no additional correction). The second one is to retrim 3d transports with the instantaneous barotropic flux at n+1/2 (proposed by us), and the third one to use the correction proposed by the reviewer. All these corrections are similar as they are based on the estimate of the barotropic transport at n+1/2, and they all provide coupling between 2d and 3d dynamics. These possibilities differ by the amount of temporal averaging, which is strongest in the first case, and absent in the second. Indeed, the barotropic transport $\langle\langle \overline{\mathbf{U}} \rangle\rangle$ is averaged over the entire interval from n to $n+1$, and the solution proposed by the reviewer is also an average over this interval. The first option and the option proposed by the reviewer are expected to provide a stronger feedback of forward-backward type, but as we already mentioned, we found that even the option proposed by us does not lead to noticeable changes compared to the first option. For this reason we did not consider the option mentioned by the reviewer. However, to demonstrate that it does not lead to noticeable changes, we added it and retested our simulations. As shown below, there is hardly any difference.

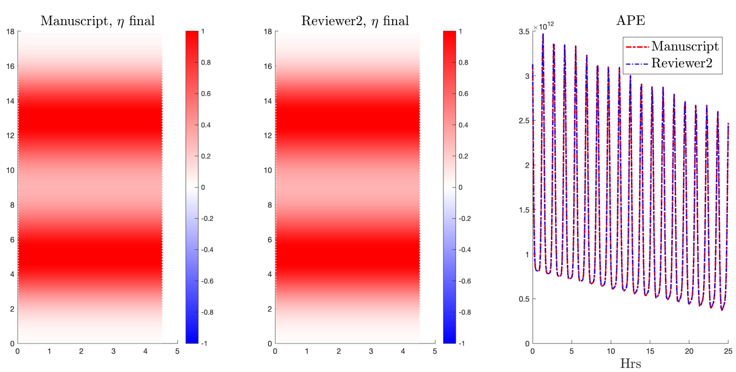

Figure 1: Comparison of elevation (m) distribution snapshots for the SE solver under the two different re-trimming (after 3 days) and their total available potential energy density (m^5/s^2) time series. The number of barotropic subcyles $M = 30$ and $\theta = 0.14$. Here triangular mesh with sides of 10km and baroclinic time-step of $\tau = 5$min is used.

For a reminder, this test case (Section 4.2 of the manuscript) uses a simple surface gravity wave (SGW) setup where we simulate a channel (of same geometry (as Section 4.1 of the manuscript) with an initial elevation distribution which is meridionally gaussian, i.e., $\ln(\eta/A) = -(y - y_{mid})^2/\sigma^2$ where $A = 3$ meters is the amplitude and $\sigma = 200$ km is the half-bell width. The temperature is set at $T = 20^o$c, the velocities are initialized to 0, and the simulation is run for 3 days with baroclinic time-step $\tau = 5$ mins. As seen in Figure 1 there's hardly any difference in the dissipation levels between the two re-trimming methods.

This could strengthen the coherence between the 2d and 3d dynamics, making the 3d fluxes $U_k^{n+1/2}$ 'feel' the value $< Ubar^n + 1 >$ and $< Ubar^n >$.

The 3d fluxes still 'feel' the 2d fluxes as the values of every sub-step still goes into its correction through $<< Ubar >>$. See the discussion above.

Such retrimming would parallel the treatment of layer thicknesses stated at line 135. It would also give some confidence in initializing the 2d with the final state of the previous time step, as this 2d state will have been used to correct the 3d variables. Such retrimming could be seen as the simplest analogue in this time-staggered context of the second constraint which classically appear in split-explicit algorithms.

While it is true that the reviewer's proposal would match the thickness treatment as per l135, we also state in l135 itself that this is not generally the case. While Zstar is independent of 3d flux divergences, layer algorithms like ztilde will not be. Furthermore, some layer algorithms by design would be operating with only part of the flux divergences depending on the desired objective. Therefore, in general, one will not be able to parallel the flux treatment with the thickness treatment, and neither would it be desirable.

These points are at the basis of the mechanics of the mode-splitting algorithm. I would be surprised if they did not affect the stability properties of the algorithm. It is quite possible, however, that the dissipation values used for the numerical experiments reported in the paper are so large that to counter the consequences of any inconsistencies in the algorithm. In my opinion, a more detailed study of the algorithm's stability would be desirable. However, I can imagine that this is not the purpose of the paper, which is essentially to show that 'it works'.

While exciting, exploring all types and possibilities of mode-splitting is not the objective of this paper. Indeed, our objective has always been to have an improved solver for FESOM2 that 'works'. It has been made clear from the title and throughout the manuscript that this work explores a split-explicit approach for ocean model FESOM2. The decisions made were centered around and motivated by the performance of FESOM2. And in regards to FESOM2, the new solver shows sufficient stability and abundant improvement over the previous approach across all targeted categories - dissipation, speed, and scalability.