Response to reviewer remarks on “Coupling framework (1.0) for the Úa (2023b) ice sheet model and the FESOM-1.4 z-coordinate ocean model in an Antarctic domain” by Ole Richter, Ralph Timmermann, G. Hilmar Gudmundsson, and Jan De Rydt.

We thank the editor and reviewers for their remarks. Our response is in blue text.

Reviewer #1:

This paper describes the coupling framework for the Úa ice sheet and FESOM ice-ocean models. This coupling is particularly interesting because both models use unstructured meshes and mesh refinement to resolve better local processes in otherwise global or regional setups. The paper is well-structured and written, and I expect it to be of substantial interest to GMD readers. I find the paper quite mature, and I recommend publishing it once the authors have addressed my suggestions and minor concerns.

Sequential coupling: You use a sequential coupling, which essentially doubles the runtime of the whole system, at least at first guess. Yet, there is no discussion of why you don't use a parallel approach. You need to give at least the reason or justification for why you went with the sequential approach. An estimate of the performance benefits of using a parallel coupling would also benefit the paper.

The choice of a sequential approach goes back to us expecting that the wallclock runtime of one component (here: ocean) would be substantially larger than that of the other (here: ice). With the ice and ocean configurations employed here, this turned out not to be true, but we decided to continue with the working system. We acknowledge that parallel coupling is an important development for Ua-FESOM in future studies. We will add a comment to the manuscript that includes an estimation of the expected speedup.

I don't like your names for sections 2.4 and 2.5, as they could be more descriptive; please consider giving these more descriptive names.

We agree and adapt the suggestion from Reviewer 2 and will rename the sections to “FESOM-to-Ua step” and “Ua-o-FESOM step”.

L116: The description of the background mesh left me hanging. How do you decide what regions "could possibly unground during the simulation period"? You describe this later, but more details are in order here. Even saying that it's simulation-specific and will be described in more detail later would suffice (if this is the case).

We acknowledge that more detail and a reference is needed at this point in the text. We adapted the text following the reviewer's suggestion.

In section names for sections 3 and 3.3, you use "verification" when "evaluation" would be more fitting. You can only verify the model results if you have an analytical solution.

We agree and will change the wording to evaluation. We have also changed all occurrences of verification to evaluation in the text (only in line 338).

In section 3.1, I would appreciate more justification for choosing the IceOcean1ra experiment. I expect that there are several idealised experiments in the MISOMIP framework, so you should say why you chose this particular one.

There are two idealised ice-ocean experiments in the MISOMIP protocol (Asay-Davis et al. 2016): “IceOcean1: retreat and re-advance without dynamic calving” and “ IceOcean2: retreat and re-advance with dynamic calving”. As we do not include variable calving front positions in this first version of the model, only one experiment remains. We will include this information under section 3.1 in the revised manuscript.

L160: "... with the later described pan-Arctic setup" should be "... with the pan-Arctic setup described later".

We thank the reviewer for the correction.

L161: "summaries" should be "summarised".

We thank the reviewer for the correction.

In section 3.3, I miss a reference to what results other modelling groups get for the IceOcean1ra setup. Are your results similar to those of others? Is there a large spread in the results in general?

We agree that it would be very informative to compare our results with the statistics of the MISIMIP intercomparison. It is one of the main motivations of the MIPs to provide means for evaluation of new models and drive model development. However, the results of the MISOMIP1 intercomparison have not yet been published. Ua-FESOM results have been provided to the working group of MISOMIP1 and will be included in the comparison.

In section 4.4, you say that you use three different machines to run the model. Such a setup is very unusual, and it would be nice to have more details. Is everything automated, or is there some manual work involved? Do all the machines have access to a shared storage area, or do you need to copy data between machines? Why did you choose this setup?

We acknowledge that running different components of the coupling framework on different infrastructures is somewhat unusual, but not unprecedented (e.g. done for the FESOM-RIMBAY setup presented by Timmermann and Goeller, 2017) The benefit is that each component can be run on its most suitable infrastructure. Specifically, FESOM requires a massively parallel machine to complete model runs in a reasonable amount of time, while Úa requires a MATLAB environment that is typically not available on hpc systems. The whole machinery is fully automated using shell scripts. Relevant data is copied across the machines, which uses only a very small fraction of the time required for the coupling steps. More details and the motivation are given earlier in section 2.3 Coupling approach:

"*At each coupling step the models are restarted from their final state at the end of the previous timestep, through the use of restart files. This procedure allows users to run each component of the model system on its most suitable infrastructure. For example, in its configuration presented here, FESOM runs in massively parallel mode on an NHR (Nationales Hochleistungsrechnen) computer, Úa on the AWI supercomputer Albedo and the coupler on an AWI desktop machine.*"

In the revised manuscript we will add information about automation and data storage to section 2.3 and refer to it from 4.4.

You only mention Greenland in the "summary and conclusions" sections. You should either remove this or also mention it in the main text. As it is, it comes completely out of nowhere.

We agree and remove it from the text.

Reviewer #2:

Review of "Coupling framework (1.0) for the Úa (2023b) ice sheet model and the FESOM-1.4 z-coordinate ocean model in an Antarctic domain" by Ole Richter, Ralph Timmermann, G. Hilmar Gudmundsson, and Jan De Rydt.

High-resolution is crucial to simulate both the ocean and the ice dynamics at the Antarctic ice sheet margins. While a few ice sheet ocean models have emerged in the last few years, no coupled model had analysed the benefits of unstructured meshes in both components at the same time. By exploring these aspects, this study introduces a novel modelling tool with great potential for projections of future Antarctic mass loss. I therefore recommend this paper for publication, but I have three main moderate suggestions (and several minor ones) that will hopefully improve this article.

<u>Main comments:</u>

Model Description: I find it difficult to find the information on the model parameterisations because the models are first presented in section 2, then more information on the parameters are given in the description of the MISOMIP configuration (section 3), and there is another part in section 4 to state that only one aspect differs in the Antarctic configuration. I would find it much easier to read if all the model parameterisations were described in section 2, and if only the specificities of either the MISOMIP or the Antarctic configurations were described in sections 3 and 4.

The motivation for the original structure was to separate hard-coded model design from case-specific parameter choices. However, we acknowledge that this structure impairs readability and we have brought the descriptions together, closely following the reviewers suggestion. In the new manuscript we have moved the model parameter choices used for MISOMIP to sections 2.1 (model description Ua) and 2.2 (Model description FESOM). The distinction between model design and parameter choices is highlighted in the text. Deviations from the default parameterizations in the realistic case remain described under 4.2. (Model configuration, pan-Antarctic case).

Coupling method: Other Z-coordinate models previously corrected velocities to cope with abrupt changes in the ice shelf geometry during coupling steps. Favier et al. (2019) claimed that they avoided the generation of spurious barotropic waves by imposing a conservation of barotropic velocities across the step change in the ice-shelf geometry, which was likely first implemented by Asay-Davis in POPSICLES. Smith et al. (2019, their Appendix A2) noted that this method could not be applied when an entire water column was grounded and that it often led to unstable numerical artifacts when used with realistic ice shelf geometries in UKESM1.0-ice. Therefore, instead of artificially constraining barotropic velocities, they artificially forced the three-dimensional divergence field to be unchanged across the change in discretization, for just the first time step after coupling. This was done by adding artificial volume fluxes where necessary, which was claimed to prevent the formation of instabilities. Has anything similar been applied in the Úa-FESOM coupling? If not, can the authors show whether or not spurious barotropic waves develop at the coupling time step? My point is not to ask for a change in the coupling method, but to document it and discuss whether this is satisfactory.

During the design of the model and the experiments, we have accounted for the issue that large and abrupt changes in ice shelf geometry can cause spurious barotropic waves leading to model instability. A high vertical resolution and a large initial coupling step (20 years) result in only small changes in ice shelf geometry during the simulation period. Ice retreat or readvance rarely exceeds more than one layer, which equates to 10-30 m. We never experienced problems related to spurious barotropic waves. We acknowledge that this has been an important challenge for previous studies and will include a statement about how this influenced our model design at an early stage.

Demonstration: In section 4, the authors choose to focus on the Amundsen Sea and Pine Island glacier, which is clearly a region of interest and a region difficult to represent at the resolution of usual climate models. However, given the Antarctic configuration of Úa and the global configuration of FESOM, it is surprising not to mention the coupled model behaviour elsewhere. Is the model only good in the Amundsen Sea region? Even if this is the case, it is worth describing the biases and remaining challenges, at least briefly.

Yes, Pine Island has been chosen as it is arguably the most critical challenge for large scale coupled models. We agree that discussing model performance and biases in other regions would make the paper stronger. We've chosen to now also present FRIS as a prominent cold water ice shelf example, where the model performs well, and Totten Glacier as a warm water example, where the model has biases. In addition, we now also discuss spurious oscillations in ice thickness in most regions. The biases can be attributed to either the ocean or the ice model and recommendations to reduce them in future studies are given.

<u>Minor comments and edits:</u>

The title of subsections 2.4 and 2.5 should probably be "FESOM to Úa" and "Úa to FESOM".

We agree. The reviewer #1 also commented on this and we have changed the titles to the suggestion of the reviewer.

Section 2.4: So nothing is done to conserve mass? I mean that the mass of meltwater injected into FESOM is not the same as the mass of ice lost by Úa. How strong is the imbalance at the scale of Antarctica?

Ua-FESOM does not strictly conserve mass. The paper discusses this for the idealised experiment (L. 215-221), including a quantification and recommendations for future development. We have given a thorough evaluation for an established test case that represents one of the most rapidly changing systems today. We expect our estimate to be an upper limit for the inaccuracies for present day ice-ocean systems. In general, evaluation of realistic applications will depend on the specific research questions of future studies that use Ua-FESOM. However, we agree with the reviewer, that pan-Antarctic inaccuracies in mass loss could be of interest to many and we will include and discuss this metric for the realistic case in the revised manuscript.

L. 123-125: this should probably be moved to the ocean model description.

We agree. This detail is related to only the ocean model and we have moved it to the ocean model description. We now also highlight that the free-slip condition along the grounding line deviates from the default FESOM-1.4 configuration.

L. 130-131: If I understand correctly, the ice shelf front interpolated to FESOM isn't vertical in case of a vertical front in Úa, right? Doesn't this create spurious melting and currents at the Front?

In sigma coordinate ocean models sloping ice fronts are an issue including the artefacts mentioned by the reviewer. Here, however, we use the z-coordinate flavour of FESOM, that is able to represent vertical cliff faces. The manipulations described in the text at L. 130-131 do not smooth the slope of the ice front. Thin ice regions (less than 10 m) are seen as open ocean regions by FESOM. This would only act to make the front steeper. How well different coordinate systems represent melting near the ice front is not clear (Malyarenko et al. 2019). We will clarify this point in the text.

Table 1: Is the 10-30 m of vertical resolution in sub-ice shelf cavity due to the use of partial steps (Adcro` et al., 1997) or to the coarser vertical resolution at depth?

Due to a coarser vertical resolution at depth. We will clarify this in the text.

L. 215-221 & Figure 4: it may be worth explaining that the inaccuracy is estimated from the difference between the fluxes in FESOM and the fluxes in Úa.

We agree and will add this information to the text.

Section 3.2 and Figure 4: In their description of the MISOMIP protocol, Asay-Davis et al. (2016) write "Models using volume or mass fluxes will need a strategy for removing mass in the open ocean to compensate for the volume of meltwater that enters the domain". Do I understand correctly that no such correction is applied? Furthermore, I think that another inaccuracy is the one due to the absence of volume conservation. For example, in the absence of melting beneath the ice shelf, if the grounding line retreats due to a reduction of the ice flow at the grounding line, sea level should drop and the ocean volume should remain constant, but I don't believe that this is the case with the proposed coupling method. Is there a way to estimate this and plot the inaccuracy in Figure 4b?

FESOM only incorporates the meltwater as a virtual salinity flux, while the ocean volume changes only according to the change in cavity geometry. Therefore, volume conservation is not an issue in our setup. We do acknowledge the fact though that on long time scales and for big excursions of the grounding line position an accurate computation of sea-level evolution would require an assessment of possible inaccuracies caused by this approach.

L. 246: "Wessem et al." should be "Van Wessem et al.".

We corrected this mistake.

Figure 7, about "Blue box visualises the resolution of a quarter-degree ocean model": most so-called quarter-degree global ocean models have a Mercator grid to ensure a nearly isotropic resolution by having Δx = RE cos(lat) Δlon with Δlon = 0.25°, and Δlat varying with latitude so that and $\Delta y = \Delta x$ everywhere (e.g., Spence et al., 2014; Storkey et al., 2018). Such quarter-degree ocean models therefore have a resolution of 7.2km at 75°S. The blue squares in Figure 7d seem larger than that, and in any case, this should be clarified in the figure Caption.

We have added information about how we calculated the quarter-degree resolution to the caption.

L. 279: not clear whether Cd is the drag coefficient/turbulent momentum exchange coefficient (also involved in the drag seen by the ocean dynamics) or a heat/salt turbulent exchange coefficient only used in the three equations at the ice shelf base (sometimes referred to as Γ or St).

Our discussion refers to the turbulent momentum exchange coefficient used in the computation of friction velocity for heat/salt turbulent exchange at the ice shelf base. We will clarify this in the manuscript.

Acknowledgements: “Funding by the EU Horizon 2020 project PROTECT (grant no. 869304) has been indispensable for this study and is gratefully acknowledged” does not seem to follow the standards expected by EU.

We have changed the format closely following the EU standards. The new acknowledgement reads: “*This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 869304 (PROTECT).*”