Author's response to Anonymous Referee 2 Comment 1

April 20, 2023

[1. Currently, the organization of the manuscript feels like a laundry list of simulations completed and a description of the results, without prioritizing the most important and revealing simulations and little discussion for what the results mean and how they relate to other studies. The study also seems to currently have about 10-15 areas of focus, making it challenging to see which end up actually being important. My recommendation would be to take all the description of simulations where numerical choices don't appear to have much influence on the results and move them into a supplement. They can be summarized in a few sentences perhaps at the end of the results section, but right now they add too much extra to the manuscript making it longer and hard to get through. This includes the sediment-hard bed transition, basal hydrology, basal hydrology instead of temperature ramp and the max time step.] Discussing numerical choices that do not significantly affect the results is equally important. However, we agree that the in-depth discussion can be moved to the supplement. Following the idea of the research questions, the results section will be restructured into the following three main themes: 1) strictly numerical aspects (Q1 and Q11), 2) numerical/modeling choices with a significant effect on the results (Q3, Q4, Q5, Q8), 3) numerical/modeling choices without a significant effect (Q2, Q6, Q9, Q10). As suggested, the third theme will only be a short summary with the details in the supplement.

[2. The way the results are currently described and presented also contributes to the challenge of reading through this manuscript. There are 15 tables, and it is difficult to understand what all the numbers in the tables mean. Figures 4 and 6 seem like a more intuitive way to present this information (though all the markers and line and shading in Figure 6 need explicit descriptions in the caption and perhaps a legend to be interpretable). There really shouldn't be more than a few tables in the main text of this manuscript, the rest should be relegated to a supplement.] Figures in the form of Figure 4 are not suitable to replace the tables because they only show the results of one parameter vector. The restructuring of the results section described above will remove two tables from the main manuscript, and, where possible, we will replace the tables with plots similar to Figure 6. Most information necessary to interpret Figure 6 was already in the caption. However, to make it even clearer, we slightly updated the caption to: *Percentage differences in event characteristics compared to the GSM base setup ($T_{ramp} = 0.0625$, $T_{exp} = 28$) for different basal temperature ramps at* 3.125 km *horizontal grid resolution (average of the* 5 *parameter vectors). The ramps are sorted from widest (first row) to sharpest (last row, see Fig. S25 for a visualization of all ramps). The different colors were added for visual alignment of the individual basal temperature ramps, and the horizontal bars represent the standard deviations. The shaded pink regions mark the numerical noise estimates (Tab. 5) and the black numbers in the title of each subplot represent the mean values of the base setup. No runs crashed and all runs had more than* 1 *surge event. The first* 20 kyr *of each run are treated as a spin-up interval and are not considered in the above. The x-axis is logarithmic. The exact values are given in Tab. S5.*

[3. The "such as" in the title of the paper is misleading. The only type of ice sheet instability discussed and tested in this paper is a thermal oscillatory instability (or B-P). A more accurate title would be simpler: "Numerical issues in modeling binge-purge type cyclic ice stream surging"] Changed to 'Numerical issues in modeling ice stream surge cycling'.

[4. Related to #3 and throughout - the term "binge-purge" oscillations has largely fallen out of favor in the ice sheet modeling community. These are more commonly called thermal oscillations or ice stream activation-stagnation cycles. It is OK to mention in the introduction that these have historically been referred to as binge-purge, but it isn't in keeping with the field to continue to refer to them as such throughout.] We are not aware of such a development [e.g., Roberts et al., 2016, Feldmann and Levermann, 2017, Ziemen et al., 2019, Schannwell et al., 2023]. Furthermore, the term *'thermal oscillations'* disregards the effect of basal hydrology and *ice stream activation-stagnation cycles* seems to be just a lengthy expression for *binge-purge cycles*. However, we will use *ice stream surge cycling* instead of *binge-purge cycling*.

[5. I am quite confused over how numerical noise is defined in this manuscript. It seems in section 3.1.3 that to quantify numerical noise that different solver choices are used. However, depending on the number of iterations occuring between the different tolerances (and the details of how the solver works) it would seem that this method could yield strongly different estimates for the "noise". Additionally, it is unclear why this is the correct bar for determining whether a change is "important", instead of a more physically meaningful quantity. Additionally, the rationale behind the set of simulations detailed in Table 6 is confusing. In a bitwise reproducible code, I don't see why the number of cores used in a simulation should have any influence on the simulations. This makes me concerned about the robustness of other simulations if the number of cores has such an important influence on the results. How reproducible are these results by other researchers? If the same setup is run on a different cluster architecture would the results be different?] Note that *These noise estimates set a minimum threshold for discerning physical significance of changes in surge characteristics due to physical model components.* We are unsure how a *physically meaningful quantity* could determine whether the modeled result is due to numerical aspects or a true physical phenomenon. However, we have tested the effect of different tolerances, and a short statement will be added to the revised draft. Additionally, we have added the following sentence to the description of the numerical noise: *By numerical noise, we refer to any non-physical differences in the model solution induced by numerical aspects such as rounding errors and convergence criteria of the numerical solver.*

When dealing with non-linear transitions such as surge onset and shutdown, small differences can accumulate and lead to a somewhat different result at the end of the model run. These small differences can be caused by, e.g., a different number of cores (see **online PETSc-FAQ link**). PETSc is used by PISM. So the results might be somewhat different on a different cluster. However, the whole point of introducing the numerical noise threshold is to identify what characteristics and relations are likely to be robust.

[6. In section 4 I see the summary of results, but very little discussion of what they mean (in some cases but not others). This seems to me to be the main missing piece of the manuscript to be useful to other researches, a discussion of how these results relate to the theory of thermal oscillations (from Schiavi, Mantelli, Robel, MacAyeal, etc.) and how they might relate to other ice sheet models.] It is not straightforward to compare our results to the existing theory because the theory in this context is not fully developed. However, a more in-depth discussion will be added to the revised draft.

References

Johannes Feldmann and Anders Levermann. From cyclic ice streaming to Heinrich-like events : the grow-and-surge instability in the Parallel Ice Sheet Model. *The Cryosphere*, 11:1913–1932, 2017. doi: 10.5194/tc-11-1913-2017.

William H.G. Roberts, Antony J. Payne, and Paul J. Valdes. The role of basal hydrology in the surging of the Laurentide Ice Sheet. *Climate of the Past*, 12(8):1601–1617, 2016. ISSN 18149332. doi: 10.5194/cp-12-1601-2016.

C. Schannwell, U. Mikolajewicz, F. Ziemen, and M.-L. Kapsch. Sensitivity of heinrich-type ice-sheet surge characteristics to boundary forcing perturbations. *Climate of the Past*, 19(1):179–198, 2023. doi: 10.5194/cp-19-179-2023. URL `https://cp.copernicus.org/articles/19/179/2023/`.

Florian Andreas Ziemen, Marie Luise Kapsch, Marlene Klockmann, and Uwe Mikolajewicz. Heinrich events show two-stage climate response in transient glacial simulations. *Climate of the Past*, 15(1):153–168, 2019. ISSN 18149332. doi: 10.5194/cp-15-153-2019.