Dear Ludovic Räss,

Thanks for the fast second review round.

Below are the two review responses to this review round. We hope these responses, the revised manuscript, and the tracked changes answer all remaining concerns.

Review 1

Dear authors,

I greatly appreciate the new experiments that you have performed, which I think raises the relevance of this paper. My main concerns are the following:

Thanks for the positive feedback.

- About the summary of how common ice sheet models solves non-linearity: Contrary to what is stated in the new version of the manuscript, I am almost certain that Elmer does have an Armijo step-size control. To give a fair summary I recommend that you contact the developers of these codes.

  Thanks for stating your concerns. We already contacted Fabien Gillet-Chaulet. Fabien confirmed that no Armijo step size control has been implemented. He added that he implemented the strategy to use first a few Picard iterations and then Newton's method for a nonlinear sliding law. We added this information (line 25).

- So you need the artificial viscosity to get a well-posed directional derivative but perform the experiments without an artificial viscosity. Could you expand on how this is possible?

  We think that is a weakness of our mathematical approach: We use the infinite-dimensional function spaces instead of the finite-dimensional finite element spaces. The infinite-dimensional approach is more general as constants do not depend on mesh resolution or used finite elements, but this approach has higher requirements on regularity. In the finite-dimensional analysis, e.g. [Hirn2013], the artificial diffusion term is not necessary. We added this information (lines 165 and 166). Our initial thought was that the artificial diffusion term could improve convergence properties, e.g., convergence speed or stability, for Newton's method. However, we did not observe this.

- Table 4: Are these the right headings?

  We hope that we made the headings for Tables 2 and 4 more clear by writing 'Number of iterations for solving the full-Stokes equations 'in' each time step.' instead of 'before'. We also clarified the headings of Tables 3 and 5.

- Clarify the summary - many readers only read the summary an the abstract

  We clarified the aspects mentioned below. Additionally, we added more information about the method of minimizing a function and determining approximately exact step sizes. We also stated all the experiments that we did shortly. We shifted the last parts, which were an outlook to the outlook section and details about the time measurement of the three-dimensional experiment, to that subsection.

- The first sentence in the summary: We conclude that our simulations are similar to the results of Pattyn et al. Similar in which sense? Make it clear that you mean that the solution is similar.

  We changed the sentence: 'We conclude that our simulations are similar to the results in [Pattyn2008].' to 'We observe that our calculated solutions are similar to those in [Pattyn2008].' (line 345).

- In the summary: "the exact step sizes verify convergence of the Newton method" - i don't understand this sentence.

  We replaced 'verify' with 'ensure' (line 346).

- The discussion on the convergence on Newtons method an its relation to δ is interesting. Can you expand this? Please summarize the findings of Hirn in more detail.

  We added a paragraph to discuss the experiments from [Hirn2013] in more detail (lines 214-219). However, these experiments are less relevant for glacier simulations, as they have 0 on the right-hand side of the full-Stokes equations instead of $\rho \boldsymbol{g}$.

- A good sanity check could be to try to run another ice sheet model to see how the Newton method converge for different δ.

  We agree that this step is necessary for improving Newton's method in ice sheet models. However, [Hirn2013] stated that a good choice of δ depends on the nonlinear exponent N, the domain Ω, and the mesh size. Moreover, [Hirn2013] did consider Dirichlet zero boundary conditions in the theoretical analysis. The sliding boundary condition could also need another choice of δ. Furthermore, this study considered the residual error compared to the solution. This approach could consider slow-flowing parts of the glacier not enough. Finally, we tried to change δ in *ISSM*. However, to our understanding, we would have to change every use of δ in the source code as this variable seems not changeable in the *MATLAB* interface. We also can not just state a value for δ. We would have to make a lot of comparisons between solutions with different domains, boundary conditions, and mesh resolutions. In conclusion, we think that this could be quite interesting, but it is out of the scope of this manuscript. We plan to do this in future work in Angelika's group together with a scientific software engineer, as this requires substantial changes in the foundation of the code structure.

Review 2

Dear authors,

Thank you for providing an updated version of your manuscript and for adding the new experiment. The manuscript draft looks in much better shape now, and we could further proceed with minor revisions.

Thanks for the positive feedback.

- Please implement all suggestions and modifications requested or suggested by the reviewer at this stage.

  We added all suggestions as explained in the review response except implementing different values of δ in a larger ice sheet model. Please see our argumentation for why we did not implement different choices of δ into an ice sheet model.

- Ideally, it would be very interesting and enlightening to also try to run another ice sheet model to see how the Newton method converge for different δ.

  We agree that this step is necessary for improving Newton's method in ice sheet models. However, [Hirn2013] stated that a good choice of δ depends on the nonlinear exponent N, the domain Ω, and the mesh size. Moreover, [Hirn2013] did consider Dirichlet zero boundary conditions in the theoretical analysis. The sliding boundary condition could also need another choice of δ. Furthermore, this study considered the residual error compared to the solution. This approach could consider slow-flowing parts of the glacier not enough. Finally, we tried to change δ in *ISSM*. However, to our understanding, we would have to change every use of δ in the source code as this variable seems not changeable in the *MATLAB* interface. We also can not just state a value for δ. We would have to make a lot of comparisons between solutions with different domains, boundary conditions, and mesh resolutions. In conclusion, we think that this could be quite interesting, but it is out of the scope of this manuscript. We plan to do this in future work in Angelika's group together with a scientific software engineer, as this requires substantial changes in the foundation of the code structure.

- Please also reach out to the Elmer devs to check what they have implemented and refer to anything relevant in the introduction.

  Thanks for stating your concerns. We already contacted Fabien Gillet-Chaulet. Fabien confirmed that no Armijo step size control has been implemented. He added that he implemented the strategy to use first a few Picard iterations and then Newton's method for a nonlinear sliding law.

- Moreover, I would suggest an alternative title such as "Assessing the benefits of approximately exact step sizes for Picard and Newton solver in simulating ice flow (FEniCS-full-Stokes v.1.3.1)" which better reflects the investigative work you are performing given there are quite some remaining open questions.

  We changed the title. To reflect that there are still open questions, we wrote ' 'We think the' reason for this behavior is that the discretized minimum of the convex functional and the root of the full-Stokes equations are slightly different.' instead of ' 'The' reason for ...' (line 207).

Thank you and best regards, Ludovic Räss