Reviewer 1:

In this paper, the authors use a state-of-the-art ice sheet model to carry out (slightly compressed) glacial interglacial simulations on the Ekstrom ice catchment and ice shelf. They use a stokes ice-sheet model with a grounded to floating transition and step forward the ice-sheet model over time scales considerably longer than those generally considered in such simulations. By make use of two different linear system solvers, as well as testing the response to a highly uncertain physical parameter dealing with the nature of the sea bottom in the ice-shelf cavity (at least, this is my understanding, see below). The improved numerical solver shows great gains in terms of computational cost, and the bed geology of ice-shelf cavities is shown to play a large role in fluctuation of ice-sheet volume and potentially lead to hysteretic behaviour.

I believe this paper can potentially by published in this journal. The simulations they have done are impressive from a computational standpoint alone, but also raise interesting questions regarding our ability to model past behaviour of ice sheets when we know so little about the marine geology of ice-shelf cavities. I have a few general comments, and a number of detailed comments, however, that should ideally be addressed.

We thank the referee for the thoughtful and thorough review of our paper. We appreciate you taking the time to complete the reviews and welcome your helpful comments. We have revised the manuscript to address your review comments (see below). Throughout this response to review document your (referee review) comments are provided in regular, non-italic font text, our response comments are provided in red font (as here).

General Comments:

1. Though this is quite specific, it is quite important, and i would like to see it clarified, ideally in the response and in a revision. The central science result hinges around the effect of different bed strength. The methods seems to suggest that, between the "hard bed" and "soft bed" runs, the only difference between the bed frictional coefficient (C) is in areas where this CANNOT be inferred from an inversion of velocities as described in 3.4 ǎ˘AˇT in other words, bed within the current ice-shelf cavity â˘AˇT meaning in both experiments, C is identical in currently grounded areas. Is this correct???

Yes, this is correct. We are only investigating the effect of different geology underneath present-day ice shelves (floating parts), while the inferred distribution of the friction coefficient (C) for the present-day ice sheet (grounded parts) is the same in all simulations.

I ask because section 3.4 would imply this, though i could not find any other part of the paper that made this clear. If the only difference is indeed below currently floating ice, this has very strong implications; however, i fear that (a) i have misunderstood and (b) even if i have not, other readers might. This aspect of the methodology should be pointed out with crystalline clarity to the point that maybe even the experiment names should change to emphasise this. (And i should add if "hard bed" means 10-1 everywhere and soft bed means 10-5everywhere, then the results overall are not very surprising ⁀A̐T so this is why it is general point #1)

To make this more clear, we have added the following to the introduction: "Here, we present the first regional scale FS simulations investigating the effect of different ocean bed properties under contemporary ice shelves on ice sheet geometry over a glacial cycle."
We also added the following statement to the model initialisation: "The model is initialised to the present day geometry using the commonly applied snapshot initialisation in which the basal traction coefficient C is inferred under the grounded ice sheet by matching observed surface velocities with modelled surface velocities."

Please note that in the Experimental Design section, we also explicitly state that different values for basal traction are only applied underneath present day ice shelves: "We consider two end member basal property scenarios by prescribing either soft ocean bed conditions (mimicking sediment deposits) or hard ocean bed conditions (mimicking crystalline rock) under all present day ice shelves in the modelling domain."

2. A key scientific result put forth is that of hysteresis with a strong (ice shelf cavity?) bed, in that the grounding line (GL) does not return to its original position. This is used to argue that even without a retrograde slope (line 364) there can be hysteretic behaviour. I point out that there have been previous studies suggesting that a continuum of grounding lines were possible, but these (and other) authors later showed that correct treatment of grounding zone boundary layers removed this degeneracy, but this treatment involved resolving the grounding zone, the length of which scales inversely with bed strength. In the context of Stokes, Nowicki and Wingham (2008) found that with an effectively non-sliding bed, there was not a unique sliding solution in the presence of a frozen grounding zone. While the authors' results are interesting, they should allow for the possibility that (a) the grounding zone is not sufficiently resolved or (b) there is not a unique solution to the Stokes contact problem with an effectively non sliding bed (therefore raising the question of whether the model finds the physically correct one) rather than assuming that the model results are correct, and hysteresis of ice sheet is possible without retrograde beds.

We thank the reviewer for pointing this out. We have added the following sentence to the discussion section of the modelled hysteresis to acknowledge this: "However, this result could also be caused by a combination of the non-uniqueness of the Stokes contact problem for non-sliding beds and an under resolving of the grounding line zone (e.g. Nowicki and Wingham, 2008)."

3. There are extensive mentions of ensemble modelling in the paper; while you do not say outright you are doing ensemble modelling, you don't say that you are not (aside from a mention that your approach is "complimentary" to it, line 310, which is confusing; it is not ensemble modelling because it does not vertically average?) I would argue you tested 2 end members of a (albeit important) physical parameter (the choice of solver is not a physical parameter), so perhaps you should be as clear as you can be that this is NOT an ensemble of experiments.

The reviewer is correct that we are not presenting ensemble simulations, but provide a pair of envelope simulations covering two extreme scenarios. We have added a sentence to the introduction to state this explicitly. It reads: "We do this by investigating end-member scenarios as opposed to ensemble modelling."

We also expanded on the statement in line 310 now explicitly stating that our approach of using a complex ice-physics model investigating end-member scenarios and ensemble modelling using simplified ice physics both have their advantages and disadvantages, but both are worth pursuing.

It now reads: "For example, radar isochrones for floating ice shelves could be incorporated more easily into the model tuning, because the FS approach does not apply a vertical average in these areas unlike ice models using a simplified force balance. We believe that ensemble modelling using simpler ice physics models and our approach of employing a complex ice-physics model and investigating end-member scenarios can both provide different new insights. Hence, both approaches should be pursued in future. This also holds for shallow ice approximation-FS hybrid approaches (Ahlkrona et al., 2016) which can build on the results shown here."

A series of 4 experiments are done, varying one of each: bed strength, and numerical solver ˘A˘T and the results of the paper are presented as dual: the effects of the hard bed, and the effectiveness of the solver. It is therefore confusing whether this paper is meant to be about numerics (in which case it might be better suited for a different journal) or about the scientific results. Both aspects are presented quite prominently making the message of the paper a bit unclear. If the paper is to be about science, then aspects dealing with numerical methods such as scaling should perhaps be in an appendix and not feature in the abstract (though i do have comments about these aspects as well).

The goal of the paper is indeed two-fold. One goal is to show that progress towards faster full-Stokes simulations is being made, so that in the near future simulations over longer time scales (>1,000 years) with this type of model should be possible for regional domains. As the scaling of the ParStokes solver is integral to the speed-up in computation time, we would like to keep this Figure in the main text.
The second goal of the paper is to present the effect of different geology under present-day ice shelves for ice-sheet evolution over a glacial cycle. The main scientific messages are however independent of the solver setup. We tried to make

this separation clear by putting technical and scientific results and their discussion in separate sections. However, to make our intentions clear to the reader from the beginning, we have rewritten the last paragraph of the introduction.

It now reads: "Here, we present the first regional scale FS simulations investigating the effect of different ocean bed properties under contemporary ice shelves on ice sheet geometry over a glacial cycle. We do this by investigating end-member scenarios as opposed to ensemble modelling. This means, we specify either very soft and slippery or very hard and sticky conditions under present-day ice shelves. The goal of the paper is hence two-fold. First, we present methodological advances by extending the feasibility of regional FS ice sheet simulations by an order of magnitude using the open source code Elmer/Ice (Gagliardini et al., 2013). We do this with a highly parallelised numerical scheme allowing to maintain a high mesh resolution (~1 km) and a freely evolving grounding line over glacial/interglacial timescales. Second, we present new scientific insights regarding the effect of different ocean bed properties seawards of today's grounding line and quantify its impact on the evolution of the entire catchment. This is done for the Ekström Ice Shelf catchment, Dronning Maud Land, East Antarctica (Fig. 1)."

Detailed comments:

line 73, data of shelf cavity ˘AˇT should point out this is only relevant to the present study *up to the farthest point of grounding line advance* in your experiments.

Correct. We added the following sentence: "For our simulations, this difference is only relevant up to the point of farthest grounding line advance."

line 174 and potentially elsewhere; please say something about the FEM basis functions in your scheme(s). It is important to establish that the basis functions are LBB conforming and that the solutions are exactly mass conserving (ie. not using penalty methods) ˘AˇT the latter perhaps not being as important for short term runs but very important for long term.
line 174 how many cells? how many DoFs?

We have added this information to the mesh generation and refinement section.

It now reads: "The 3D mesh consists of ~200,000 nodes and therefore ~800,000 degrees of freedom. We are using stabilised P1P1 elements and an algorithm that deduces a mass-conserving nodal surface to avoid artificial mass loss (Gagliardini et al., 2013)."

line 203: are you sure? all physical uncertainty? what about ice shelf crevassing weakening? Not to mention these physical parameters, if i understand correctly, are only varied in the ice-shelf cavity (see General Comment 1)

The reviewer is correct. We qualified this statement. It now reads: " Hence they also account for some uncertainties in model parameters, forcings, and physics of the applied ice sheet model."

line 223: since this is exceptional is it really of value for general knowledge?

We agree that there is not much value for the compute times using a now decommissioned system. We therefore removed the paragraph from the manuscript and also deleted the corresponding compute times in Figure 4.

Also:

a) it is odd to compare one solver on one system and another on another system. How about an additional test (only a few time steps) of both on the same system, with walltimes so a comparison can be made.

We agree that this would be the preferred option. However, larger jobs get priority on the SuperMUC-NG and hence queuing times would be much longer on this system for smaller jobs like the MUMPS jobs. We did however test whether absolute compute times for a few time steps between the systems are similar. As this was the case, we believe that the absolute numbers provided in Figure 4 are most informative.

b) ParStokes has great scaling but what about absolute time for a fixed core count on the same system compared against MUMPS?

We added the following statements to the solver setup comparison section (sec 4.1): "This speed-up is in part due to using more CPUs in the ParStokes simulations. When comparing absolute runtime of the scaling simulations, ParStokes provides faster computations for >168 CPUs. This means the minimum requirement for faster simulations with ParStokes is a supercomputer with more than 168 CPUs. The exact CPU number may however very well vary from system to system depending on the available hardware."

line 228: following on from comment on line 174, which of the two uses a stabilisation method? if not both, then what about the other one?

They both use stabilisation methods, but different ones (stabilized method for MUMPS and bubble stabilisation for ParStokes) for stability reasons. We added the following sentence: " When using the MUMPS solver, the stabilised method
is used, while for the ParStokes solver we use bubble stabilisation (Gagliardini et al., 2013). This results in slightly different systems that need to be solved."

Figure 8: Here or in an appendix you should show a similar plot comparing ParStokes and Mumps for soft and hard bed (whichever shows poorer agreement). It is important to establish that the effect of solver, while having a large difference on performance, has very small effects on Volume and GL position relative to the

effect of the physical parameter. If MUMPS and ParStokes differ, at most one is correct – if the difference is large, how are we to trust the physical results?

We do not think that this proposed Figure would add information that is not already there. We are showing grounded area differences in Figure 5 which is indirectly a measure for the differences in ice volume. In addition, Figures 6 and 7 show differences in grounding-line position and ice thickness between the solver setups for different time snapshots. We believe that these Figures provide a solid case that differences between the solver setups are small.

line 250: following what?

Changed to: "Following the period of volume gain, ..."

line 257: then there is a volume decrease – what causes it?

Apologies, but we could not find the statement of a volume decrease the reviewer is referring to.

line 289, some funny maths. How does increasing performance by a factor of 6 allow runs of 40,000yrs when using the MUMPS solver allowed less than 1000a? this is not a factor of 6. did you do something to increase the time step that was not mentioned?

I think we confused the reviewer here. These numbers have nothing to do with each other. The speed-up refers to the difference in computation time between MUMPS and ParStokes. MUMPS does allow computations of 40,000 years, they just take 3-6 times as long as with ParStokes. But given enough time, these type of runs are possible. To avoid this confusion we removed the 1,000 year maximum from the sentence.
It now reads: "The new setup allows 3D full-Stokes ice sheet simulations
on the regional scale over 40,000 years now in under a month's time."

line 349: explain what you mean by grounding line
We changes this to: "The earlier onset of grounding line motion in the retreat phase ..."