Author Reply to Referee Comments

Referee #1:

This is an impressive model developed by the authors, and a comprehensive writeup that addresses a great deal of questions that many such papers generally do not answer. The CISM 2.1 model is an impressive piece of computational engineering. The experiments they present are similarly impressive given the time scales involved, the speeds presented, and the depth of detail able to be resolved. I enjoyed reading the discussion on the Greenland experiment, and the tradeoffs between the efficiency and accuracy of DIVA.

A: Thank you for your kind comments. We appreciate that you enjoyed reading the paper and found the model to be well engineered.

I note that I was unable to install CISM as the only linux stations to which I have access are those maintained by my academic department, and it is a departmental policy not to allow root (sudo) access on these computers, which prevents the use of package managers. (This type of situation may not be uncommon and the authors may want to think about providing alternative installation instructions.)

A: We are sorry that you were unable to install the code on a Linux station without root privileges. The CISM documentation gives instructions for an Ubuntu build using the apt-get package manager, but the experience of software engineers on the project is that building the code without a package manager is fraught with difficulties. We would suggest either asking your system administrator for help with the build, or requesting an account on the Cheyenne supercomputer at NCAR. The CISM/CESM developers at NCAR would be glad to sponsor a Cheyenne account for anyone who requests it for research.

I do not have any general criticisms for this manuscript and I found it to be well-written and comprehensive. I do however have a number of specific comments/questions that I feel the authors should consider before publication.

Thank you for the constructive comments, which helped us clarify a number of points for readers.

eqn 13: I think it would be better if the actual boundary conditions were stated, rather than/in addition to the form of pressure used in the BCs.

A: The lateral boundary condition at vertical cliff faces in given (in weak form) by the third term of eq. 9. The net pressure p = p_out - p_in (from eq. 11-12) is the same p that is inserted in the integral over gamma_L in eq. 9. We have reworded the text to make this more clear.

eqn 31: note that Arthern et al (2015) avoids the double integral by raising the (s-z') term to various powers.

A: We agree that the notation of Arthern et al. (2015) is more elegant. We rewrote these equations in terms of the integrals F1 and F2, with F2 replacing omega/H, and added the Arthern paper to the references.

page 12

line 20: and initial tau_b? and initial \eta?

A: The initial tau_b depends on beta and u_b, or equivalently on beta_eff and u_mean. The initial eta(z) depends on H, T(z), u(z), and v(z). We know H and T, and we start with the current guess for u and v. At code startup, this guess is u = v = 0. Otherwise, we initialize (u,v) from the previous time step. We added a sentence to clarify that u and v either are zero (at the start of the run) or are taken from the previous time step.

page 14

line 14: are you in fact using the L2 norm, which is independent of resolution/number of nodes? or is the norm simply the root mean square of a vector, which I believe is called the l-2 (script lowercase l) norm?

A: We changed the notation to l_2-norm. The answer actually depends on a solver option; it is possible to choose either an absolute error threshold, or a threshold that is normalized by the number of nodes. But that is probably more detail than is needed here, and we agree that the lower-case notation is more appropriate.

lines 16-18: how is the user notified upon nonconvergence?

A: The number of nonlinear iterations is written to standard output (or to a log file). Optionally, the number of linear iterations per nonlinear iteration can also be written to standard output. When the code fails to converge, a warning message is written to output, and it is up to the user to decide whether the non-converged solution is good enough. We added text as follows: "The number of nonlinear iterations per solve—and optionally, the number of linear iterations per

nonlinear iteration—are written to standard output. In the event of non-convergence, a warning message is written to output, but the run continues."

eqn (37): the text says this is solved in each column, but you show the full Laplacian term, which you explain later as approximated as vertical only – but upon reading eqn 37 this is not yet stated.

A: We agree that this is confusing. We revised eq. 37 to show only the vertical diffusion term, explaining in the text below that the horizontal diffusion term is assumed to be negligible in comparison.

eqn 44: I don't see how this follows from (42) and (43). I do see how such a relation would follow is effective strain rate is always proportional to effective stress, which it is, but this is not implied by (43).

A: We simplified the derivation of equation 47 as follows (from p. 15, line 22 to p. 16, line 6 of the original manuscript):
The effective stress (cf. (3)) is defined by
$$\tau_e^2 = \tfrac{1}{2}\tau_{ij}\tau_{ij} \quad (43)$$
The stress tensor is related to the strain rate and the effective viscosity by
$$\tau_{ij} = 2\eta\dot{\varepsilon}_{ij} \quad (44)$$
Dividing each side in (44) by 2 \eta, substituting in (42), and using (43) gives
$$\Phi = \frac{\tau_e^2}{2\eta} \quad (45)$$

page 18

line 12: how do you ensure T does not go above the melting point?

A: This is a general property of IR, which is monotonic for tracers. To clarify that this is a property of IR, we replaced the statement on line 12 with the following text in the third bullet (on tracer monotonicity) in Section 3.3.1: "Thus, T never rises above the melting point under advection." The paper by Dukowicz and Baumgardner gives a good explanation for readers interested in the details.

line 17: there is something confusing about this, why is a vertical remapping needed? is this an alternative to casting the equations in (x,y,\sigma) coordinates and simply solving the equations, which will already contain a vertical advection term relating to the coordinate transform?

A: Basically yes. IR is implemented as an arbitrary Lagrangian-Eulerian (ALE) method. Mass transport is carried out in each layer, resulting in layers that no longer follow sigma coordinates. The vertical remapping transfers material between layers to restore the desired sigma coordinates. We revised the text as follows: "... top and bottom ice surfaces. The new vertical layers generally do not have the desired spacing in σ coordinates. For this reason, a vertical remapping scheme is applied to transfer ice volume, internal energy, and other conserved quantities between adjacent layers, thus restoring each column to the desired σ coordinates while conserving mass and energy. (This is a common feature of arbitrary Lagrangian-Eulerian (ALE) methods.)"

eqn (62). I have wondered about this – is it proven this is actually a stability limit with BP or SSA or DIVA? and I'm not referring to a von neumann analysis as this only takes a linearisation of leading order terms into account. The instability occurs because flux of mass is proportional to the gradient of thickness; but when membrane stresses are present, no matter how strong the bed, if the surface is steep enough they will begin to matter, and I do not believe flux would grow unbounded. This is not something proven; but neither, I believe, is the stability criterion stated for the equations considered.

A: We are unaware that this diffusive CFL condition is an actual stability limit for BP, DIVA, or SSA. Our experience suggests that the relevant stability limit in the presence of membrane stresses is usually advective CFL. To sharpen this point, we reworded the text as follows: "... is subject to this diffusive CFL. The stability of Glissade's BP, DIVA, and SSA solvers, however, is limited by (65) in practice; (66) is too restrictive." We are unable, however, to provide a detailed theoretical explanation.

eqn 72. This is unclear, as it is stated as a thinning rate but is always positive.

A: Yes, this is confusing. We changed the wording to describe dH/dt as a rate of thickness change, rather than a thinning rate. Since dH/dt < 0 when defined in this way, we added a minus sign to eq. 72.

eqn 76. Similarly the RHS is always positive, yet I think it is meant to refer to thinning.

A: We revised the description of dH/dt and added a minus sign to eq. 76.

eqn 79. What is the accumulation rate?

A: There is no accumulation rate for this experiment; this is now stated in the problem description.

section 4.1 it would actually be interested to see how the results compare when the DIVA balance is used.

A: It could be informative to do this comparison, but we do not know the analytical solution when using DIVA. We plan to do a more thorough comparison of DIVA to the SIA and BP solvers in a future paper.

section 4.2. I have always thought that ISMIP-A and -C have always been a bit binary – flat and sliding, or bumpy and frozen. I understand if you do not have the time, but if you do I feel it would be very interesting to examine BP/DIVA comparison in a situation with *some* topography and *some* (if slow) sliding, as this is perhaps closer to conditions tested in realistic models.

A: We agree that a test with intermediate properties between ISMIP-A and -C would be useful and interesting for model evaluation. We would prefer not to develop such a test for this paper, because the paper is already long and because it could take some time to determine what are the most appropriate or realistic parameter values to represent "some" topography and "some" sliding. And the results would likely call for further analysis.
    As mentioned above, we have started working on a paper doing further comparisons of DIVA to BP and other solvers. That paper will include, for example, some problems with realistic temperature profiles. We will work on adding a hybrid ISMIP A/C test as well.

page 28

line 22. "outside the RACMO ice sheet footprint". unclear what you mean.

A: RACMO uses a mask to delimit the ice sheet extent, and provides an SMB only within the masked region. Outside this region we prescribe a negative SMB to inhibit ice advance, since the ice would advance unrealistically if we set SMB = 0. We added the following sentence to the text: "RACMO2 provides a SMB only for the region included within its ice sheet mask; outside this region, we prescribe a negative SMB." We removed the word "footprint."

page 29

line 27-29. This isn't really a parameter change so much as an adjustment of model physics.

A: We replaced the first sentence in line 27 by: "Some of the model physics is constrained to make the simulation more robust."

p30, line 32-34. there is a subtlety here and it is something not made clear from the description of DIVA. are you saying that the discrepancy is due to vertical variation of temperature? if so, then I would ask whether DIVA accounts for depth-dependent temperature. It would be easy to implement as viscosity is vertically resolved, and if not i would recommend implementing this for a better DIVA-BP comparison (though, of course, re-running your Greenland tests for this paper would likely be too difficult).

A: Our DIVA solver does, in fact, account for depth-dependent temperature, in the sense that $T(z)$ appears (via the flow factor A) in the expression for depth-dependent viscosity $eta(z)$, which is integrated vertically to obtain the depth-average viscosity in the stress balance. What we mean to say is that in the presence of a vertical temperature profile, there can be large vertical gradients in viscosity (relative to a case with uniform temperature), which could affect the accuracy of the approximation in a way that is not considered in test problems such as ISMIP-HOM A and C. This is another factor we would like to study in the future paper mentioned above. We modified the text as follows: "... in slow-sliding regions. However, their analysis did not consider the effects of vertically varying temperature. In CISM simulations of real ice sheets, the depth-integrated viscosity depends on the vertical temperature profile (cf. (2) and (24)), with dynamic effects that are not included in test problems with a uniform temperature and flow factor."

Referee #2:

The paper by William H. Lipscomb et al. in my opinion is a well written and extensive description of the features of the new version 2.1 of the Community Ice Sheet Model (CISMv2.1) and fits very well in the scope of GMD. The reader is provided with a detailed layout of the approximations (and their numerical implementation) to the 3D full stokes flow equation used in the model as well as the user choices regarding model parameterizations of features such as basal sliding or calving. The fundamental equations and their implementation in the model are well documented allowing the reader to appreciate the physical aspects of ice flow incorporated in CISMv2.1.

A: Thank you for your positive comments on the paper and for your many constructive suggestions below.

Before providing a more specific review of the paper I want to mention 2 general minor points which might improve the manuscript.

1. The authors mention that CISMv2.1 so far is limited to Greenland applications and that support for Antarctic model settings is deferred to future model releases. Reading the manuscript, I got the feeling that all major features required for modelling the Antarctic Ice Sheet (AIS) are included in CISMv2.1, e.g. calving and a simplified scheme prescribing ice shelf melt rates. Maybe one or two sentences as to which features are missing to make AIS simulations feasible (e.g. forcing interfaces, melt rate parameterizations, grounding line migration schemes) and whether AIS simulations with CISM are a near term option or require more extensive work would be helpful.

A: We agree that most features needed for AIS simulations are already present, and that this should be pointed out. Perhaps the most important feature not included in CISM2.1 is a grounding-line parameterization (GLP), which is described by Leguy et al. in a paper to be submitted in the near future. Once that paper has been submitted, a GLP will be included in the next CISM release. The reviewer might be aware that two of us (Lipscomb and Leguy) submitted the results of CISM Antarctic simulations for the initMIP-Antarctica project (Seroussi et al., 2018, in prep), using a GLP along with new inversion methods. Thus a developmental version of CISM is Antarctic-capable, but requires more testing and documentation before being included in a supported code release.

      To make this more clear, we added the following text at the end of Section 2: "CESM2 does not have an interactive Antarctic ice sheet, in part because of the many scientific and technical issues associated with ice sheet–ocean coupling. However, CISM2.1 includes many of the features needed to simulate marine ice sheets, and a developmental model version (including a grounding-line parameterization to be included in a near-future CISM release) has been used for Antarctic simulations."

Furthermore, since the model is freely available via github maybe there is an interactive development platform in which new features can be committed for dedicated developers (similar to e.g. the PISM approach).

A: Yes, this is now possible. After the paper was submitted, new CISM development moved to https://github.com/escomp/cism on the Earth System Community Modeling Portal. The latest code is publicly available, and new developers are welcome to check out the code, create branches, and make pull requests. We added text in the Code Availability section to point readers and potential developers to the new repository.

2. It would be interesting to see which implementations and aspects in CISMv2.1 provide a different or similar approach as compared to other models on the market (such as Sicopolis, PISM, ISSM, BISICLES etc.) both regarding the physics and performance of the model. Differences and similarities could be explicitly pointed out at relevant sections of the paper

which would allow potential future users of the model to quickly grasp the strengths and specialties of CISMv2.1.

A: We hesitate to do this throughout the paper, since these models are always evolving, and we are not familiar with many details of other models beyond what we have read in publications (which might not reflect current capabilities). We have tried to summarize the overall code development philosophy in the Introduction, and we hope that this summary, together with the code description that follows, will allow potential model users to decide whether CISM (possibly in combination with CESM) would be useful for their research.

In the following I will address specific points of the manuscript:

Page 2:

Line 32: Here, a condensed list of the main changes (pointing to the sections in the manuscript) between v2.0 and 2.1 would be helpful to give the reader familiar with CISMv2.0 a quick overview.

A: Thank you for the suggestion. We added the following sentence near the end of the Introduction (after the sentence beginning, "Changes between versions…":
"These changes include a depth-integrated higher-order velocity solver (Section 3.1.4), new parameterizations of basal sliding (Section 3.4), iceberg calving (Section 3.5), and sub-ice-shelf melting (Section 3.6), a new build-and-test structure (Section 4.6), and many small improvements in model numerics." We also added a reference to Price et al. (2015), a tech report that describes CISM2.0 as it existed soon after its release.

Page 5:
Line 2: but will be in the near future?

A: At the end of this sentence we added "but is under development." This is a hard problem, so we cannot promise this capability in the near future. It is a goal for CESM3. We also added here some text on CISM Antarctic capabilities, as suggested by Referee #1.

Page 9:
Line 27: may be faster? Can this be quantified? What is the definition of large problems here? Here a back of the envelope estimate might be useful.

A: By "large" we mean "whole-ice-sheet", e.g., all of Greenland. Since Glide has an implicit thickness solver whereas Glissade is explicit, this is difficult to quantify; it depends on the

diffusive CFL condition (Eq. 62 of the original manuscript) that applies to the ice sheet geometry, and hence the ratio in stable time step between Glide and Glissade (which might be roughly one order of magnitude). In practice, we now use Glissade for nearly all Greenland simulations, because of its higher-order velocity solvers and more advanced physics options.

For improved clarity, we modified the text as follows: "For small problems that can run on one processor, there is no particular advantage to using Glissade's local SIA solver in place of Glide. Glide's implicit thickness solver permits a longer time step and thus is more efficient for problems run in serial. For whole-ice-sheet problems, however, Glissade's parallel solver can hold more data in memory and may have faster throughput, simply because it can run on tens to hundreds of processors."

Page 13
Line 3-5: put in a reference to Figures 4,5,13 in which differences of BP and DIVA solution are visualised?

A: We would prefer not to refer here to Fig. 13. While that figure illustrates differences between DIVA and BP in Greenland simulations, we have not done enough analysis to show that the differences are due primarily to vertical variations in gradients of membrane stresses. (For example, some differences may be transients, but to show this would require a longer and more expensive BP simulation.) We agree that Figures 4 and 5 are relevant, and we reference them implicitly by referring to Section 4.2.

Page 14
Line 16: is there a flag pointing out that the run didn't converge? I guess it would be problematic e.g. in ensemble simulations or intercomparisons to mix converged runs with non-converged runs.

A: Please see the response to a similar question from Referee #1 above. We added a statement in the text that the number of nonlinear iterations per solve (and optionally, the number of linear iterations per nonlinear iteration) is written to standard output, and that a warning is written to output when convergence fails. It is left to the user to decide whether non-convergence is scientifically problematic.

Page 22
How is basal melting handled in partially filled ice shelf cells?

A: This is a good question, which reminded us of a related question: "How is sub-shelf melting handled in cells that contain the grounding line?" We added text at the end of Section 3.6 to answer both questions:

"Sub-shelf melting is applied only to cells that are floating based on the criterion $b < -\rho_i H / \rho_o$, where b is bed elevation (negative below sea level) and H is ice thickness. In partly filled cells at the calving front, basal melt is applied to the effective thickness H_{eff} rather than the grid cell mean thickness H (see Section 3.5). For example, a melt rate of 10 m/yr applied in a cell that is 10% full would reduce H by only 1 m/yr."

Page 23

Lines 25-30: How does this approach differ from (DeConto, R. M. Alley, R. B.Pollard & DeConto, 2015, Potential Antarctic Ice Sheet retreat driven by hydrofracturing and ice cliff failure)? Has it been tested against runaway marine ice sheet retreat, i.e. would it lead to realistic solutions in an Antarctic setting for present day climate conditions?

A: Both approaches are motivated by Bassis and Walker (2012), but our approach differs from Pollard et al. (2015) in the details. Notably, CISM does not combine cliff limiting with crevasse deepening and hydrofracture, as in Eq. A.3 of Pollard et al. Also, the details of horizontal wastage (Eq. A.4) are different. Pollard et al. compute a rapid horizontal retreat rate which is converted to a large thinning rate, whereas CISM simply relaxes the cliff surface elevation toward the upper limit of Bassis and Walker. In CISM Greenland and Antarctic experiments to date, cliff limiting has not led to runaway retreat; it simply thins vertical cliffs at marine margins.
      We added text as follows: "In experiments to date, this mechanism has not triggered the rapid ice sheet retreat seen by Pollard et al. (2015) in Antarctic simulations that combined cliff failure with hydrofracture. CISM2.1 does not simulate hydrofracture."

Page 24 (Standard Test Cases)

Is CISMv2.1 tested for reproducibility on long (e.g. tens of thousands of years) paleo simulations? Experience with other models has shown that e.g. dynamic choice of FFT's can lead to small changes in the results which build up on long time scales leading to a lack of reproducibility. A paleo-repro test would be very helpful to check whether the model choices ensure identical solutions for identical model settings. I'm aware that such a test is not standard in comparable models with paleo applications but wonder whether it has been looked at in CISMv2.1.

A: We have not looked specifically at long paleo simulations, but we have done many multi-millennial simulations of the kind described in Section 5.1. We routinely test for exact restart, and the model does yield identical solutions when run with identical settings, provided it is running on the same computer platform with the same processor count. We have made this more explicit with modified text in Section 2: "A special kind of I/O file is the restart file, which

includes all the fields needed to restart the model exactly. Whatever configuration options are chosen, model results are exactly reproducible (i.e., bit-for-bit) for a given computer platform and processor count, regardless of how many times a simulation is stopped and restarted."

Page 25

Line 17-21: Doesn't it make more sense to compare test cases applying the same time step? So either 1 yr or 5 yr in both applications.

A: Glide results are sufficiently converged with a 5-year time step, and a shorter time step changes the results very slightly. For this particular test, Glide is more accurate than Glissade for any time step that is sufficiently short to be stable in both models. To make this more clear, we moved the last sentence of the paragraph to immediately after "...with a time step of 1 yr", to make clear that the stability requirement is the reason for choosing a shorter time step.

Line 19: errors expressed in percent thickness change would be more instructive.

A: The errors vary with distance from the central dome, with the largest percentage errors near the edge, so it is not obvious how to translate the rms error to a percent error. But it is helpful to know H0 and R0 to have a sense of scale. We added text stating that for these tests, $H0 = 500\sqrt{2} \approx 700$ m and $R0 = 15\sqrt{2} \approx 21$ km.

Line 29: what makes them particularly useful?

A: We changed the text as follows: "CISM supports all six experiments, and here we show results for Experiments A and C.   These two tests are particularly useful for benchmarking higher-order models, since they gauge the accuracy of simulated 3D flow over a bed with large- and small-scale variations in basal topography and friction.

Page 26
Line 31/32: how does the model perform in the stream tests using a range of resolutions e.g. 4,8,16 km? This could be plotted in Figures 6-7.

A: While it would be possible to do this, we would prefer to keep the original figures for somewhat technical reasons. All the figures in Section 4 are generated with Python scripts included in the CISM2.1 release, so that any user can reproduce them out of the box. Since the Python script for the stream test is set up to plot one profile at a time, a plot with multiple profiles is not something a user could readily generate.

Alternatively, we could show multiple profiles on multiple panels, but we do not think the additional plots would contain enough information to justify the additional space. The existing figures serve the purpose of showing that the model can accurately simulate lateral shear margins, even when the margins are fairly sharp compared to the grid resolution as in the Raymond test.

Page 27
Line 18: is there a reason for the 6.8 km resolution?

A: This resolution was chosen by McAyeal et al. (1996), and the test is set up to have the same resolution they used in their paper. We added a few words to make this explicit.

Page 28
Line 22: maybe change to "[. . .], it is constrained by reanalysis at model boundaries and well validated against observations, therefore its SMB is more realistic compared to a global climate model."

A: We changed the wording as suggested.

Page 29
Line 3/4: Maybe change to: model parameters should always be thoroughly tested and reviewed depending on the intended application.

A: We replaced the last sentence with the following text: "Generally speaking, model parameters should always be tested and reviewed depending on the science application."

Line 11/12: why not use temperature and surface mass balance from the same climatology?

A: This was an oversight on our part. When we updated our SMB forcing with the latest data from Brice Noël, we neglected to update the surface temperature T_s. We note, however, that ice sheet evolution is much less sensitive to T_s than to SMB in these simulations, since T_s is used only to set the ice temperature at the upper boundary. We have found that a small (~1 degree) bias in T_s has a negligible effect on ice sheet dynamics, given that most vertical shear occurs in warmer layers near the bed. Thus, a modified T_s from a newer climatology would not significantly change the results shown. But we agree that it is better to be consistent, and we will use an updated T_s climatology for future simulations.

Page 30

Line 12: which is probably due to boundary conditions (e.g. geothermal heat flux) and the parameterization of subglacial hydrology?

A: Yes, those are the most likely explanations. We changed the text to "...representation of NEGIS, possibly due to a missing geothermal heat source or simplified subglacial hydrology."

Line 30: maybe "A more detailed investigation would be needed [. . .]"

A: We changed the wording as suggested.

Page 31

Line 31: "Simulated GL migration would differ [. . .] run at different resolution [...]"

A: We changed the text to "Simulated GL locations could differ, however, if the model were run at different resolution or with a grounding-line parameterization." We prefer "locations" to "migration" since it is not obvious that the grounding line would move from its initial location. In the hypothetical case that the model and the forcing were perfect, grounding lines might simply stay put at present-day locations.

Page 32

Line 10-20: I wonder whether this discussion and the associated Figure 17 can be omitted as the authors point out correctly, that the form and shape of the ice shelves would change given different parameter choices.

A: Although the results are parameter-dependent, we would prefer to leave this discussion in the paper because it illustrates some of the challenges in simulating Greenland's ice shelves. For example, the Kangerlussuaq figure shows that eigencalving methods, unaccompanied by strong melting, can overestimate shelf extent in enclosed bays where the flow cannot spread out. Figure 17 might be a useful reference for other modelers trying to simulate these shelves.

Line 22: "modelling challenge for Grenland" -> "[challenge for modelling ice shelves in general]"

A: We do think that some of the challenges discussed here are specific to Greenland's small ice shelves, coastal topography, and surrounding ocean temperatures. In our experience, it is also very challenging to simulate correct ice shelf extent for Antarctica, but this could be for a different set of reasons, so we hesitate to generalize.

Page 33

Line 3: "in both idealized and real world applications"

A: We changed the wording as suggested.