Dear editor,

In this second round of revision, we have addressed the comments of the reviewers and incorporated the changes that were suggested to us to the best of our ability. This has helped us improve our manuscript, and we thank the reviewers and your editorial work.

However, while all three reviewers were pointing out similar issues in the first submission, reviewer 1 is currently the only reviewer that does not seem satisfied with some of the issues originally raised. These issues can be summarised in two points, following their report:

1. There is uncertainty on how we derive the transition temperature. This is a point we addressed multiple times during the first review, and following that, we added a new figure (Fig.1) as well as a new method section. In the new revision we plan to submit, we also added a column in Table 1 with the inception temperatures for all simulations. Despite this, Reviewer 1 is expressing a similar concern as during the first round of revision. We note that, on the contrary, Reviewer 3 is content with the addition of Fig.1 which "adds a lot of aid for understanding, as well as precision".
2. Reviewer 1 seems to perceive the drift in inception temperatures for different abruptly applied forcing levels, which we argue is due to time-dependent effects, as scientific uncertainty. This is a point we already addressed in response to the first review. We refer the inception as "close to 0°C", as this is a hypothesis which is easier to test for other modelling centres, and furthermore, as we show the scientific uncertainty has in fact a minimal effect on our results. This is a point we addressed in the previous revision and when answering to Reviewer 1, notably by adding Figure 7. Contrary to what Reviewer 1 is pointing out, we do not believe we "downplay the uncertainty around the limit on the upper bound of climate sensitivity" at any point in our paper.

We believe that Reviewer 2 and 3 are both satisfied with the current state of our paper, as they recommend the paper to be accepted after some minor revisions that we have implemented. Reviewer 1 is expressing concerns regarding certain points of our paper, which we hope we have further clarified and strengthened in this iteration.

We are looking forward to your feedback and we would like to thank you for your time and consideration.

Best regards,
Martin Renoult
Johannes Hörner, Navjit Sagoo, Thorsten Mauritsen

In the text below we reply point by point to each reviewer comment. The reviewers text is blue and our answer in black font.

Reviewer 1:

- Abstract l. 10: Please specify "... approximately zero degree Celsius in MPI-ESM1.2 and CESM,…".

We note that the single simulation we have done with CESM1.2 is not nearly as comprehensive as the set of simulations performed with MPI-ESM1.2, therefore we only mention MPI-ESM1.2 in the abstract (L.6).

- Abstract l. 13: The authors still do not account for the uncertainty in the inception point temperature in their main estimate of the upper bound of ECS. This has to enter somehow. Please, at least add one more sentence about the limitation and indicate that the range could be different when using another inception temperature.

While there is a scientific uncertainty on the inception temperature which we have not tested, we believe what the reviewer is discussing here is the displacement of the inception temperature towards colder temperatures as the initial forcing increases, which is due to time-dependent effects discussed in our study. Regarding how the upper bound on ECS varies if the inception temperature differs, we added a sentence in the abstract on that aspect.

- Introduction: Please add the main motivation behind constraining the upper bound of ECS here. This was written in the response to reviewers only, which I found quite useful (see below). Then, please add references to the statements made in the response.
"We emphasise that we are constraining the upper bound of ECS here, not the best estimate value, which is necessarily lower. Much larger values on the upper bound are often given, and with some lines of evidence the upper bound is basically infinite. Therefore, even providing a value of 10 K brings valuable information to the community wide effort of constraining ECS."

We have added this comment in the introduction of our manuscript.

- l.70: You can add you argument here (or already in the introduction), that this is necessary because not all modelling centers share their failed LGM experiments. This would give a nice justification for your claim at the end of the manuscript.

Done.

- Table 1 and text: The name "abrupt50ppm" stands in contrast to the other runs called e.g. "1/8xCO2". But actually all runs use an abrupt change in CO2 at the start of the simulation, if I understand correctly. I would suggest to remove the "abrupt" from the simulation name, as otherwise the reader is misled to believe that a different procedure was taken in this run.

The simulation name was changed to “50ppm”.

- Figure 1: I like this schematic, but please indicate that there is uncertainty in the "True point of snowball Earth inception", for example as a note in the figure caption. Currently, the schematic makes a very strong statement that 0°C is the one and only ("True") snowball Earth inception point, which is misleading.

As written above and in the text (e.g. L.100, from L187), there is no “uncertainty” in the temperature, apart from variability. What is perceived as uncertainty and Fig.1 is due to time-dependent effects, which should not exist because fast transitions cause by strong abruptly applied forcing are not relevant for a snowball Earth.

- l. 98-99: The formulation of this sentence is a bit unclear. As I do not see the added value of this sentence, I would suggest to remove it.

This sentence is important for the methods, so we rephrased it.

- l. 100 - 105: The fact that you get a different inception point in different runs, tells us that there is quite some uncertainty. You then go on to say that, therefore, you need additional experiments. But this still doesn't help the reader to understand HOW you used those additional runs. Right now, to me it seems that you have just taken the inception point of the run with the slowest transition to a snowball Earth. Which is okay, it just has to be clarified, and the remaining uncertainty has to be discussed.
You could simply continue this paragraph with a few more sentences here. Which requires further clarification, is that the new schematic in Fig. 1 hints to an approach, where the authors fitted a line through the points of lowest feedback and then calculate where that line crosses 0. I don't think this is what is done. Again, this is just an issue of missing detail on how the overall inception temperature of the model is derived from all of the runs.

L.90 we wrote “In the real climate system, the climate would slowly cool down as the forcing from the Sun or greenhouse gases change only slowly.”, which is further developed in the following paragraph and which we believe makes it evident that slow runs need to be used to correctly identify the inception temperature, in conjunction with Fig.1. Indeed, the inception temperature is therefore based on the slowest transition, as described in our Methods. There is no uncertainty due to the faster transitions happening at lower temperatures, as this is due to time-dependent effects (which are not uncertainties).
The approach to estimate the inception temperature is explained between L.91 and L.99, with further details added in the revised version of the manuscript. We note that the dashed line in Figure 1 is added for illustration purposes only, we do not fit a line to the transition points of the different runs. This is now explained in the caption of Figure 1.

- l. 142-143: This statement is only true for a slight range between -5 and -10 K temperature anomaly, so this condition must also be written in the text. I assume you want to highlight that the maximum (positive) values for the cloud feedback are achieved for the strongest forcings. So please be precise.

This is specified earlier at L.140 “during the initial 10°C of cooling. This strengthening arises from…”

- l. 148: The surface albedo feedback "starts to exceed" the combined other feedbacks at -15 to -20 K. If you only write "exceeds", it sounds like this only true during that temperature range. But it actually starts at these temperatures and becomes even stronger at colder temperatures.

Done.

- Fig. 4: The table here shows the years of the first instability. I don't know why, as this year is never again referred to in the text. It would be much more interesting to show the inception temperature you calculate from each of the runs in this table.

The year of instability is useful for other modelling centres aiming to perform the experiments. However, we also added the inception temperature for each run.

- l. 172: The authors write that the "temperature at which the climate transits towards a snowball Earth state is broadly similar across the different CO2 forcing", but they never actually show these temperatures anywhere, and from Fig. 4 it seems like there is a spread of at least 10 K among the runs. Again, simply showing the numbers for each of the run would be helpful for the argument.

Added from comment above. The temperatures are in Fig.4 and in Table 1.

- l. 180: I agree with one of the other reviewers that the geometric argument is very hand-wavy. Furthermore, this line of argument presumes that the snowball Earth instability would be reached when the sea-ice edge enters the sub-tropics in every model (and the real world). As this is not proven here (in other models the instability could be reached earlier or later), the line of argument is flawed and I would suggest to remove this sentence.

We believe it is important to keep a simple and illustrative argument in order to obtain a ball park estimate of the transition temperature. Central to this argument is that transition happened around 30°, which appears to be the case in the climate models we have access to. However, for instance in the simple Budyko model, the transition happens at 45°, but in this model there were no clouds, ocean heat transport, etc. Thus, in the simply Budyko model the transition temperature could be warmer than 0°C. If true, that would imply a far more constrained ECS. Hence it is interesting to retain this simple argument here, such that it can be tested with other climate models and any possible difference in temperature they may occur.

- l. 183ff: I would still suggest that the final global mean temperature of the 50 ppm run is very close to the inception temperature that the authors are looking for. Indeed, it is stated that the run shows signs of instability around 0°C of global mean temperature. But this is not shown anywhere, and just stated in the text so the reader simply has to believe what the authors say. This has to be shown! I also do not understand, why the 50 ppm run is not included in Fig. 4, as it would be very helpful. The 50 ppm run together with the 1/8xCO2 run are probably the most insightful runs, when it comes to defining the inception temperature.

We added the 50ppm run onto Fig. 4. Contrarily to what the reviewer is suggesting, the 50ppm is not the most insightful run, because it does not enter an instability. The 50ppm run is computationally expensive to run, so it had to be stopped manually.

- l. 183: the authors responded to my and one of the other authors previous comment that 50 ppm is rather 1/6 of pre-industrial CO2 and not 1/4, by saying that they choose to stick with 1/4, because it fits better with their other runs. This is not a "choice" the authors can make! 50/284 = 0.176, which is roughly 1/6 and only 70% of what is actually 1/4 of pre-industrial CO2. Given the logarithmic nature of CO2, this is not an insignificant difference. The authors should not defy the actual numbers, just because it fits better into the list of other runs, especially since there seems to be no other real reason for doing that.

We changed “1/4” to “1/6” at L.183.

-l. 199-200: Does this mean that you take the value for the slowest run that transitions to a snowball Earth? And what exactly is this value? Again, "close to 0°" is not precise. Please provide the actual number that was used as the transition temperature. And if the authors simply used zero instead of deriving the number from their runs, then this should also be clarified.

Yes, this is the approach we took. This is specified in the Methods Section 2.1, in Fig.1, and between L.199-201. The value is indeed close to 0°C, as seen in Fig.4. In the case of 1/8xCO2, the value is -0.3°C. Throughout the manuscript, we use “close to 0°C” as we believe this is a much simpler message, relates to the physically relevant process, and also more inviting for other modelling centres to replicate the experiments. As shown in Fig.7, the exactitude of the transition temperature is not needed for our question, as it has limited impact on predicted upper bound on ECS.

- Fig. 6: The authors responded to my previous comment on Fig. 5 ("How does Fig. 5 show that the instability is around 0°C?) that they added an explanation on how to read a Gregory plot in the Method section. This was not the point of the comment. The authors should ideally add a graphical derivation of the inception point for these simulations. In their response they write that it is approximately -5°C in Fig. 6, but in line 209, they write that it is at -24 K, which translates to approximately -10°C. So what is true? From looking at Fig. 6, it seems the latter is the correct number. Again, my main concern is the missing scientific clarity and detail in explaining and showing how these numbers are derived.

Fig.1 is a graphical derivation and an explanation of the inception temperature. It is further explained in L.98-99 "By definition it is also the year where the tangential, which is the total climate feedback, would be strictly zero, as it becomes positive after, at lower temperatures.", a sentence that we rephrased based on the reviewer's suggestion.
The inception temperature is expected to vary between either 0°C or -5°C between models (Fig.7), as pointed out by the reviewer. At L.209, we wrote "at 1/128xCO_2 is around a temperature of -24 K, but considering the time-dependency effects discussed in this study". It is because of time-dependent effects that the temperature translates to approximately -10°C and not 0°C, as one can see from Fig.4 (and now in Table 1) that under strong forcing, the temperature can be displaced by around 10°C.

- l.209: Please specify in the text how exactly you go from -24 K to -14 K (or from -10°C to 0°C) for deriving the actual inception temperature in CESM1.2. This will let the reader get a better chance in building their own opinion on how reliable this estimate is. Please also discuss the uncertainty in this approach in more detail! CESM is not MPIESM, and it is not at all given that the instability temperature would progress to higher values in exactly the same way as it does for MPIESM.

See the previous comment above. While we indeed do not know if time-dependent effects would behave similarly in all models, they are thought to be connected to the large heat capacity of the ocean and slow mixing processes, which are two components that should be at least present in other coupled climate models.

- l. 221: "close to" or actually 0°C? Why not be precise? You must have used a specific value.

Given the methodology and the overall goals of the study, we believe this is the best way to describe the estimate. We note that we also test the impact of a -5°C error on the transition temperature, and find a relatively small effect (Fig.7). Given this there is not much point in using a higher precision.

- l. 228: The formulation is not accurate. You do not "apply an uncertainty", but rather test the outcome of using another value for the inception temperature.

We changed the sentence.

- l. 229-231: I do not agree that this is a modest change, especially since you only tested one additional value. Furthermore, even in this example the difference is quite significant, as the lower bound of the 90% confidence interval for -5°C as inception temperature (4.7 K) is much less critical for CMIP6 models than the default lower bound (3.9 K). As written above, this uncertainty has to also enter the abstract, as this is probably the main outcome that readers will take away from this work.

We find it hard to imagine that the true transition temperature is off by as much as 5°C, but since the relationship is linear the results of this test can be scaled to any error estimate. Based on the available evidence we are not able to estimate the transition temperature uncertainty.

Furthermore, the more relevant number to perhaps compare to models is the median since this is the most probable value. And although we place our results in relation to the IPCC (2021) scientific assessment of climate sensitivity, we note that most CMIP6 models are below these.

- I agree with one of the other reviewers that there should be a conclusion section. Especially since the last section also still includes some results. There is a lot of value in the standard structure of a scientific paper, and I think that also here it would be useful to summarize and conclude in a separate section.

We note that the other reviewers are content with our manuscript structure.

- General: Please consistently use either "inception temperature" or "transition temperature". Otherwise this will only confuse the reader. Ideally, also describe how this relates to the "temperature of instability”.

We modified “inception temperature” to “transition temperature” in the text.

l. 379-381: Should also include what greenhouse gas concentrations are used in Voigt et al. (2011), I think they used pre-industrial values.

We added the CO_2 concentration which is indeed near pre-industrial (278 ppm).

We carefully considered the list of technical comments and we have applied the majority of them. From the list of technical corrections, we applied the following:

- l. 9: comma after "slowly"; "transition appears”
- l. 100: remove "Unfortunately"
- l. 137: "break down"

- l. 138 and at other locations: "..anomalies to pre-industrial" is poor use of the English language. I know it requires more words each time, but correct would be e.g. "We report anomalies of the global mean temperature compared to pre-industrial values in units of K." Or "We report global mean temperature anomalies with respect to the pre-industrial value in units of K."
- l. 151: "diagnostic ... is" or "diagnostics ... are"
- l. 177-180: Sentence very long and hard to understand. Please rewrite.
- l. 189: time-dependent
- l. 190: " we believe to be due to the..." -> "we speculate/hypothesize that this is due to the".
- l. 210-211: Sentence poorly formulated
- l. 235: remove "to"
- Fig. A1.: The second sentence in the caption is poorly formulated.
- l. 250: Remove "model". This was raised already in my previous review.

We have decided to not apply the following comments, as we believe they are discretionary suggestions that do not impact the clarity of the manuscript and that the original words convey the intended meaning clearly.

- l. 9: remove “Regardless”
- l. 102: remove "kind of"
- l. 130 "considered to be the main driver"
- l. 168: "Whereas" -> "While"
- l. 195: "... shown in Appendix ..."
- l. 198: "hardly" -> "not"
- l. 233: "This" -> "The"

Reviewer 2:

- Fig. 7: As mentioned before, I find this much more convincing now. Still, I wonder how much the relationship would change if the upper left point (CESM2) were left out. Specifically, what are the central values and uncertainty intervals for the upper limit of ECS (assuming inception temperatures of 0 degC and 5 degC, respectively) if this model is discarded? If they change substantially, this would imply that the relationship relies a lot on one model, and then this should be reported in the text. If not, it's fine.

As a tangent to this: I interpret "whether CESM2.1 is included or not" from l. 143 such that you experimented with including/excluding the single model-family ensemble, but not the simulation that is part of PMIP4 (i.e., the upper left point in Fig. 7). Is that correct? If that is not correct, then ignore my previous remark.

The relationship is not sensitive to the inclusion of CESM2. Renoult et al. (2023) showed that the PMIP4 emergent constraint has the same robustness and slope whether the PMIP4 ensemble includes CESM2, an apparent outlier with a very high ECS, or models from the CESM family (CESM1.2, 1.3 namely, and not CESM2), with ECS much closer to the original PMIP4 ensemble as published in Kageyama et al. (2021). In our manuscript, we do not include the CESM family because we only have access to the SSTs and we believe this could create confusion for the readers, since we use surface temperature, including land and sea-ice, and not SST in the paper (see Section 2.3). We acknowledge that we should have specified that the inclusion (or non-inclusion) of CESM2 and/or the CESM model family leads to the same slope and intercept for the emergent constraint relationship (which is shown in Renoult et al., 2023), which we have now added to the manuscript.

- Methods section: I failed to make this point in the previous review: In the methods section it should be stated in isolation that the emergent constraint is derived from the PMIP4 multi-model ensemble, and how many models this includes. Right now, the fact that PMIP is used only clearly stated in the caption of Fig. 7, and PMIP is mentioned in a side note in the methods section.

We indeed failed to mention that we are using PMIP4 models in the methods section. It is now added to the manuscript.

- l. 143 – 146: I don't understand the statement because CESM2 is a *high*-ECS model (and CESM2.1 too, I assume?), so I don't understand how the fact that the relationship is robust against adding or leaving out CESM2.1 shows "that the *lower*-ECS CESM models are not necessary to preserve the quality of the relationship".

We acknowledge that we should have specified that the inclusion (or non-inclusion) of CESM2 and/or the CESM model family leads to the same slope and intercept for the emergent constraint relationship (which is shown in Renoult et al., 2023), which we have now added to the manuscript. Because the slope is the same in both cases, CESM2 and the other CESM models are exchangeable which is useful here because as described above we do not have access to the surface temperature of the lower-ECS CESM models.

Very minor

- l. 227: "a temperature of -24 K" is not correct, either temperature anomaly, or degC.

We modified and corrected this sentence based on this comment and a comment of the second reviewer.

- l. 253-255: grammar error in first half-sentence

This sentence was corrected.

- in the authors' response letter, an important point was raised:
"We emphasise that we are constraining the upper bound of ECS here. Much larger values on the upper bound are often given, and with some lines of evidence the upper bound is basically infinite. Therefore, even providing a value of 10 K with uncertainty brings valuable information to the community effort of constraining ECS."
I very much agree with this statement and believe that the paper would profit if this statement were included with the same clarity.

Although the reviewer specifies this does not require to be addressed, a similar point was raised by another reviewer, therefore we added this comment in the introduction of the manuscript, in order to emphasise the value of the method.

Reviewer 3:
I would omit the phrase "And as it happens" from the start of this sentence at L183. Similarly, L245 avoid starting a sentence with the word "Because". L251 rephrase the start of this sentence " This here proposed approach" possibly to "The described approach to provide an upper bound…"

The suggestions of the reviewer have been applied.

When referring to anomalies, where the authors use temperatures in K, I am uncomfortable with statements in which temperatures are reported with negative temperatures (in units of Kelvin) without reference to the temperature being an anomaly relative to pre-industrial, e.g., L227 ".. which at 1/128x CO2 is around a temperature of -24K.". Wherever the author presents a negative temperature & Kelvin unit please insert the word "anomaly" or words to that effect, e.g., L227 ".. which at 1/128x CO2 is around a temperature anomaly of -24K.". Address also L168 and L173. Personally, I would use the same temperature unit throughout and rely upon appropriate text to notify the reader that an anomaly is being presented.

A similar comment was raised by another reviewer so we changed the °C units in the manuscript to K and we now refer to "anomaly relative to pre-industrial" when appropriate.

Please also clarify if the land surface is held fixed in these experiments or if a dynamic vegetation model is used. I ask as the authors use an ESM (GCM with additional earth system components) and they acknowledge that ECS includes changes in atmosphere, ocean, and vegetation. Mauritsen et al., 2019, which they cite for model description, states that MPI-ESM1.2-LR (and I assume -CR) has a dynamically computed vegetation distribution. If vegetation is held fixed then this could be corrected at L71 , for example "Continental configuration, land surface, non-CO2 greenhouse gasses and orbital configurations are held fixed as in PI or LGM". If a dynamic vegetation model is implemented what extent is this coupled with an interactive carbon cycle (one-way coupling?) and what happens to the vegetated part of the land surface in these simulations? (interesting to think what would happen to plant photosynthesis and leaf water vapor conductance at low CO2 values) – could this cause the ESM to fail? If the land surface is held fixed then this needs to be specified as it would be useful for other modelling groups that would like to follow up on this work.

We thank the reviewer for this relevant information that we omitted. the vegetation model is dynamic and so not all land surfaces properties are held fixed, only ice sheets. The dynamic vegetation model reads greenhouse gas values but does not feedback on them. The vegetation dies off in our simulation, just as what could be expected in a true snowball Earth scenario considering the low CO_2 levels and temperatures dealt with even in the tropics.