Reviewer 1

The authors have carefully addressed most of the concerns and comments I proposed in the previous review report. Only a few issues remained for further revision and clarification.

1. Correct typos like Line 428 Fig. S4 (should be Fig. S6).

Thanks for spotting this, it has been corrected.

2. Add the percentage of the accumulated heat in each layer relative to the total amount in Table 4 like 0.7 (17%) and the author will note that despite the absolute value of the amount of deeper layer accumulated heat remaining nearly constant (ranging around 3*10^24J), the percentage of the heat accumulated in the deeper layer (below 2000m) generally decreases from b990 to b100. This is because in a scenario with a higher forcing (like b100) than b990, the fast upper ocean warming and ocean dynamic adjustment associated with AMOC weakening would primarily cause a larger warming in the upper 2000m but not below 2000m (diffusion is rather slow in transfer heat downward). Therefore, the percentage of the heat stored in the deeper layer decreases from b990 to b100. This is not related to the fact that the higher forcing cases had more time to accumulate heat in the deep ocean, which may be valid for the Indo-Pacific but not the Atlantic. The middle panel in Fig. 10 clearly illustrates this issue that between 2000-3000m, the b100 simulation displays a weaker warming than b990, this is linked to the AMOC-related ocean interior adjustment. The differences between the Indo-Pacific and Atlantic deeper layer response may explain this issue.

We tried to add the information on the relative amount of heat in Table 4, but that makes it difficult to read. We added Table S1 to the Supplementary, above Figure S6, reporting the relative values. A reference to the new table has been added to the text. We agree with the reviewer that the dynamical changes in the AMOC are a possible explanation; this was already stated at lines 489-493 and 541-546 (new version), which have now been rearranged to better make the point.

3. It may not be proper to name it "warming hole" in Line 500, which indeed indicates the role of opposite stratification change above and below 2000m that may also relate to AMOC change. This is an interesting result and may be worthy of more in-depth discussion.

We removed the name “warming hole” which indeed could have been misleading, and substituted that with “non-linear behaviour”. We added a sentence to stress the link of this phenomenon with the ocean dynamical adjustments at lines 491-493.

Reviewer 2

Review of "Multi-centennial evolution of the climate response and deep ocean heat uptake in a set of abrupt stabilization scenarios with EC-Earth3"

By Fabiano et al.,

Recommend: minor revision

General comment:
This study presents a thorough analysis of 1000-year abrupt stabilization simulations with EC-Earth3, offering some insights into long-term committed climate change. The simulations, spanning historical and SSP5-8.5 scenarios, reveal temperature increases exceeding Paris Agreement goals. Notably, only the 1990 simulation achieves stabilization below 1.5 degrees warming. The study emphasizes the importance of multi-centennial timescales, noting a decrease in climate feedback parameter magnitude and revealing variations in surface warming patterns. Precipitation changes, particularly in sub-tropical oceans and Mediterranean hotspots, highlight dynamic climate processes. The focus on deep ocean heat storage underscores its role in determining the final state of the climate system. Overall, the findings contribute to our understanding of climate dynamics and have implications for informing climate policies. Although long simulations with coupled climate models have been conducted and presented in many previous studies, the simulations from the present study have much higher resolutions and consider different warming scenarios. The paper exhibits a well-organized structure and skillful writing, with a coherent storyline that enhances the research's accessibility. Overall, I find the paper intriguing and recommend its publication after the necessary revisions.
I have also read previous review comments and the point-to-point reply to all review comments. In the previous review, the comprehensive comments provided by both reviewers have proven invaluable in refining and improving the manuscript during the revision process.

I have a few minor points that require further attention and revision:
1. One main concern is the forcing dependence of the climate feedback parameter: I still do not understand why the climate feedback parameter turns out to be more negative while increasing external forcing. This is different from the previous study of Jonah Bloch-Johnson et al. (2021) and Meraner et al. (2013). I am sure there are also other studies on state-dependent climate feedback and climate sensitivity. The authors have added one more section to discuss the forcing dependence of the climate feedback parameter. But I still cannot an answer in the manuscript or in the reply why the climate feedbacks behave like that. I am sure many things should be done and can be done in further studies on this issue, but it would be great if the authors would provide some explanations. A reasonable physical mechanism will be very important here since the authors argue this is one of the main novelty of the present research.

Meraner, K.⋆ , T. Mauritsen, and A. Voigt, 2013: Robust increase in equilibrium climate sensitivity under global warming, Geophysical Research Letters, 40, 5944–5948.

Thanks for the comment. We first point out that Figure 5 has been updated since the previous version did not include the correction of the TOA fluxes for the model imbalances. The changes are minor and do not modify the overall behaviour observed.

Bloch-Johnson et al. (2021) study a multi-model set of idealized abrupt NxCO2 simulations. We first notice that a few individual models in Bloch-Johnson et al. (2021) go in the opposite direction, although only one model shows a negative sensitivity to both forcing and warming, which is consistent with our result. Also, the forcing of b990 and b025 (1.25x, 1.5x), which show the strongest non-linearity, is in a range which is not covered in Bloch-Johnson et al. (2021) and Meraner et al. (2013).
Apart from this, we are cautious against a direct comparison due to differences in our experimental setup, regarding both the initialization and the external forcing. Regarding the first point, our simulations start from different climate states (from historical and SSP5 scenario), rather than the pre-industrial climate. Regarding the forcing, our simulations also include other GHGs and aerosols, following historical+SSP5 forcing. If we expect the GHGs contribution to be in line with the CO2 one, aerosols might have a different impact, especially on the lower forcing cases. However, this alone is unlikely to explain the observed difference in the feedback parameter, since it would require an extremely large sensitivity of the feedbacks to the forcing agent, while most studies indicate at most a moderate sensitivity (Salvi et al. 2022, Richardson et al. 2019). For b990, if we consider an aerosol forcing around -1.0 W/m2 (with respect to a total of 1.8 W/m2 of instantaneous GHG forcing), the observed difference in the feedback parameter with respect to b100 would require an aerosol feedback of about -3 W/m2/K, which is not supported by current evidence.

We extended the discussion on this at lines 304-317. Further analysis is required to assess possible physical mechanisms, including exploring individual feedback contributions and performing a more quantitative estimate of the impact of other forcing agents, which we will leave to future studies.

2. I must say that I enjoyed reading the review comments from reviewer #2. She has raised many interesting and important research questions. I am wondering whether the authors could add one more discussion section, where you could discuss some unsolved questions or raise new open questions for further studies.

Following this suggestion, we briefly extended the last paragraph to present some pathways for future research at lines 547-556.

3. Another issue is whether the simulations reach a final equilibrium of the whole climate system. The authors have mentioned the “quasi-equilibrium” as defined by Li et al. (2013). I think the simulations reach the “quasi-equilibrium” state, where surface temperature is almost stabilized while TOA equals the deep-ocean heat uptake. I would suggest the authors move the reply to reviewer #2 about the “quasi-equilibrium” into the manuscript.

Thanks for the suggestion, we added a comment on the quasi-equilibrium at lines 246-250 of the revised version.

4. I would suggest the authors use different experiment names. The names of ‘b990’, ‘b025‘,…,’b100’ reflect the starting time of the simulations, but not the forcing difference. I would suggest the authors use an experiment name that could directly reflect the key information of the experiment, such as ‘1.25xCO2’, ‘1.5xCO2’,’2xCO2’,…,’4xCO2’, or the direct CO2 concentrations.

The reviewer’s suggestion would certainly help show the actual CO2 concentration of the runs, but we prefer leaving it as it is, because the starting year is a key information in our setup. In fact, in our setup also other GHGs and aerosol concentration change following the CMIP6 historical (b990) and SSP5-8.5 scenario (all other runs). The complete forcing set is then only completely defined through the starting year, which also gives information about the branching time (the simulations start directly from the historical/ssp585 simulations).