We thank both reviewers for their time and the helpful comments which we address in turn, below. In general, we are confident that we can provide a revised manuscript that will satisfy their concerns and be appropriate for final publication in Earth System Dynamics.

Reply to Reviewer 2:

Smith and co-authors present the first description of new idealized Earth system model simulations conducted with the UKESM, which is used here in its fully coupled configuration, including dynamic Greenland and Antarctic ice sheets. The simulations encompass the Tier 1 experiments of the Tipping Point Model Intercomparison Project (TIPMIP) but extend substantially beyond them by incorporating additional global warming levels and branched simulation pathways. Running a model of this complexity, particularly with fully dynamic ice sheets, is a significant achievement requiring expertise across multiple disciplines. Nevertheless, the model setup remains somewhat unsatisfactory, as the ice sheets are initialized directly from present-day conditions, resulting in noticeable model drift even in unforced scenarios. The authors acknowledge these limitations and point to ongoing research aimed at addressing them.

The manuscript is well written and provides a concise overview of the simulations, with a particular emphasis on high-latitude processes. However, it remains largely descriptive and does not dive deeply into underlying mechanisms or dynamics. This is also reflected in the figures, which primarily present simple time series and omit visualizations of the more complex dynamics governing these overshoot scenarios. Given this, I wonder whether the manuscript may be more appropriately suited to *Earth System Science Data* rather than *Earth System Dynamics*.

Thanks for reviewing and the positive comments. As we state, our group has a suite of papers under review and in progress with more detailed analysis of individual aspects of these simulations but we think an overview paper such as this is an important introduction and overview in its own right. Although we acknowledge that it does not go into great physical detail of the results, our paper does include analysis and interpretation of the simulations and we feel that it is better suited to ESD than ESSD.

Beyond these general comments, I have two major points that I would like the authors to address:

1. Definition of reversibility.

The manuscript refers repeatedly to reversibility and irreversibility of sea ice, ice sheets, and other components, based on the path-(in)dependence of variables during ramp-up versus ramp-down phases. However, these responses are highly transient and differ fundamentally from the traditional concept of reversibility associated with hysteresis, which concerns (quasi) steady-state behavior. I encourage the authors to provide a more thorough definition of reversibility, one that explicitly accounts for the transient nature of the experiments and the characteristic internal timescales of different components (e.g., sea-ice versus ice-sheets).

This is a very good point. An earlier version of our paper did in fact include a discussion of exactly this issue which was cut for the sake of brevity. We will gladly take the opportunity to reinstate some of the material that addresses this topic.

2. Focus on high latitudes despite ad-hoc ice-sheet initialization.

The strong emphasis on high-latitude and ice-sheet responses is difficult to reconcile with the highly idealized ice-sheet initialization, which the authors themselves acknowledge as a major limitation. This is not a criticism of their general approach; coupling dynamic ice-sheets within fully complex ESMs remains extremely challenging due to their large inertia, long equilibration timescales, and the high computational cost of such simulations. However, given the substantial biases introduced, particularly for the Antarctic ice sheet, it is unclear whether the resulting ice-sheet behavior can be meaningfully interpreted. As there is no straightforward solution to the initialization problem, I recommend reducing the emphasis on ice-sheet responses and instead providing a more detailed analysis of other Earth system components, such as the AMOC, ocean circulation in general, or terrestrial changes.

We're glad that the Reviewer agrees with us that there are currently no easy answers to this conundrum. However, we believe that there are nevertheless scientifically robust conclusions that can be drawn from the behaviour of the ice sheets in these simulations with careful analysis, and do not believe we have gone beyond those limits here.

As part of addressing a related point by the other Reviewer we will include both PI and ZE-0 lines on all figures to more explicitly demonstrate that the background trend in ice sheet evolution is not significantly biasing the climate simulation. These figures will also demonstrate in particular that the trend does not affect ice sheet surface mass balance (SMB) which is an extremely important aspect of the ice sheet simulation and comprises half of the paper's material focusing on ice sheets.

For the GrIS, in section 3.3, the forced signal of mass change is clearly very much larger than any background trend, so we think that GrIS mass change can be considered as a robust response to the experimental protocol. Our simulated AIS mass change is clearly something that needs to be interpreted with more nuance. We see no significant trend in either SMB or basal mass balance (BMB) for AIS under the PI climate. The most apparent background trend attributable to our initialisation is that of the Mass Above Flotation (MAF) visible in Fig 7b. As noted in the text, this is attributable to the continuation of the currently-observed thinning of Pine Island and Thwaites Glaciers. Although significant in terms of the MAF of the ice sheet, this background trend is independent of GWL, relatively constant for all except the last few years of the simulations and only affects a geographically very-limited part of the ice sheet. This regional MAF trend does not significantly affect mass loss from the large floating shelves of the Ross and Filchner-Ronne which is the largest part of the forced mass loss of the AIS as a whole in this model, on these timescales (compare FIg 7a with 7b, especially the vertical scales), nor does is interact significantly with the AIS-integrated SMB which is the other key part of the whole-icesheet mass balance.

So, to the degree that we do interpret the forced evolution of the AIS, we believe what we have said and shown in this paper is physically meaningful. We certainly agree that there is a long way to go to develop more satisfactory ways of initialising ice sheet components for coupled ESM simulations, and that this significantly limits how we can relate these simulations to projections for the real world, particularly when considering their change in MAF and contribution to sea level rise. It is for this reason that in section 3.7 we go no further than concluding that

"the GMSL contribution from Antarctica should not be viewed as a simple function of global warming level. More so than Greenland on the timescales of our simulations, the AIS contribution arises from a complex interaction of the evolving boundary SMB and BMB with the internal dynamics of the ice sheet"

As for refocusing the paper to include more parts of the Earth System in the description, we must once again defer this to other parts of our wide group who have their own papers in progress analysing exactly these things, which we are not really at liberty to include here.

Minor comments:

L2: Please mention the range of GWLs here.

OK

L11: Mention after how many years of zero CO2 emissions, CO2 was removed again from the atmosphere.

OK

L67: branched off instead of spawned?

OK

L73: I would prefer to have Appendix A moved to the main text, to further improve readability and look up of simulation names.

OK

L96: Can you briefly explain why the negative emission rate was set to half the positive one and not the same?

Our wider experiment protocol (which Reviewer 1 would rather we say less about) includes testing the sensitivity of Earth System response to carbon sequestration at different rates. Using the same rate of negative emissions as positive ones led to some stability problems in UKESM, so in purely practical terms we got most simulation output for analysis, most quickly, from the half-rate experiments. For the level of detail and conclusions present in this overview paper, the rate of negative emission is not significant. Since Reviewer 1 would like to see less mention of the wider protocol for simulations not used in this paper, we will try to balance these two requests in the revised paper

L115: Is JULES also running on a 1° grid?

The JULES base longitude-latitude grid is identical to that of the atmosphere, (1.875◦ x 1.25◦). We will note this in the revised text

L120: Please briefly mention the solver for the ice-sheet dynamics (SIA, SSA, hybrid …)

We use the L1L2 solver (Cornford et al. 2013). We will note this in the revised text

L142: Are the ice-sheets in thermal equilibrium at time of branching?

We don't have the thermodynamics component activated in these simulations. The ice sheets maintain their initial internal temperature and effective viscosity fields throughout. We will note this in the revised text.

L295: Explicitly mention again that this refers to the first 550 yr of the simulation.

OK

L339: Typo: THe PI

We will fix this

Fig. 1: Can you add a panel (or a separate figure) depicting the AMOC evolution for these different model runs? In panel c, why is there an offset at the beginning of the simulations of PI and Up8?

We are reluctant to do this, as others in our group are preparing a paper solely focussing on the AMOC evolution in these simulations. It is a substantial topic in its own right, and we would rather not briefly allude to it here without doing it justice.

We think the offset between black and red lines in Fig 1c results from the 30 year running mean that has been applied to the timeseries, meaning that nothing can be plotted for the first 15 years of data for PI and Up8. This does highlight however that there appears to have been a problem applying the equivalent smoothing in panel b), since they do not have offsets. This will be checked and corrected in the revised manuscript.

Fig. 7: I find the term Mass Change slightly confusing here. Initially, I thought this refers to the net mass change, instead of the loss term (correct?).

Mass change refers to the cumulative, net mass change of the ice sheet. E.g. in 7a, negative numbers mean less ice than at the start of the simulation. 7a refers to the whole ice sheet, grounded ice and floating shelves combined, and 7b shows only the change in ice mass that would contribute to sea level rise (ie largely excluding the shelves) via the Mass Above Flotation definition. We will clarify this in the revised text.