Reviewer #2

I thank the author team for their interesting contribution. The study finds an interesting AMOC stability landscape in CLIMBER-X by systematically varying freshwater forcing and atmospheric CO2. I have not seen such a comprehensive study of AMOC stability across varying backgrounds before, and I think it provides an intriguing template to move beyond single-perturbation AMOC studies and broaden our understanding of the processes affecting AMOC stability on multi-millennial time scales. This merits publication in itself, but I think the authors could increase the relevance of their manuscript by expanding their discussion of the implications of their findings for our understanding of past or future AMOC states. I recommend publication with minor corrections.

We would like to thank the Reviewer for the positive appraisal of our work and the comments on our manuscript.
Following the Reviewers' suggestions below, we have now expanded the discussion on the implications of our results for the interpretation of AMOC states, both for the past (i.e. Pliocene) and for the future.

Main points

1. I did not understand how many and which simulations were run in total to explore the CO2-FWF space. Was each combination of CO2 and FWF run for initial 'on' and 'off' states? Were the 'pathway' simulations run transiently, or iteratively run into equilibrium?

To clarify which simulations were performed in the CO2-FWF space, we will provide more detailed information about the paths in Fig. A1 in an expanded caption:
Simulation pathways used to explore the stability of the four different AMOC states in the combined CO2 and freshwater space plotted on top of the AMOC stability landscape shown in Fig. 4. The stability of the Off AMOC state in (a) was explored with simulations starting from a large FWF of +0.5 Sv and then gradually decreasing the FWF until the AMOC recovers, for all levels of CO2. The stability of the Strong AMOC state in (d) was tracked in simulations starting from a large negative FWF of -0.5 Sv and then gradually increasing the FWF. For the investigation of the stability of the (b) Weak and (c) Modern AMOC states, the starting point were pre-industrial conditions, marked by the black dot. The black arrows indicate the primary path through the CO2 and FWF space, from which subsequent experiments with varying FWF in different directions are initialized (green arrows). Since the Strong AMOC state is not stable for CO2 lower than 280 ppm for the FWF range shown in the figure, the stability of the Modern AMOC state in (c) for CO2 lower than pre-industrial is diagnosed from simulations initialized with a large negative FWF of -0.5 Sv, similarly to what done in (d) for the Strong AMOC state. The rate of change of the forcing in all the experiments is 0.02 Sv/kyr for FWF and 2 %/kyr for CO2.

2. The existence of a 'strong' AMOC mode above 370 ppm is a key result of the study. The authors already compare their results to other model studies and mention that it shows the potential for stronger-than-present AMOC states in the past. Given that this CO2 threshold is close to both current and Pliocene CO2 concentrations, and that the Arctic might have been seasonally ice-free in the Eemian, it would be interesting to discuss this result also in the context of observations and proxy data. Is this solution

purely theoretical or is there evidence that such a circulation pattern has existed in the past?

There is indeed some proxy-based evidence from the Pliocene suggesting a stronger AMOC in the higher CO_2 world of the Pliocene. We will add the following to the revised manuscript:
Our results suggest that a different mode of AMOC (the Strong mode) was possible during past warm climate conditions. The Pliocene was the most recent period in Earth's history with elevated atmospheric CO_2 concentrations of ~400 ppm (Martínez-Botí et al., 2015; Seki et al., 2010), which, according to our results, would be high enough to push the AMOC to a Strong state. There is indeed proxy-based evidence of a stronger-than-present AMOC in the Pliocene (Raymo et al., 1996; Ravelo and Andreasen, 2000) with an increased northward ocean heat transport in the Atlantic (Dowsett et al., 1992), which is consistent with sea surface temperature reconstructions for this period showing warmer conditions in the North Atlantic (McClymont et al., 2020). Climate models also tend to produce a stronger AMOC under mid-Pliocene conditions, although with considerable spread (Zhang et al., 2021; Weiffenbach et al., 2023).
The Eemian is probably not very relevant in the context of our paper, as the climate change is induced mainly by a very different orbit at that time, and the effect of orbital configuration on AMOC is not considered in our paper.

3. Besides the 'strong' AMOC state, the study shows that the 'off' state is also stable for high CO_2 concentrations. If I understand correctly, this is the conclusion from simulations that were initialised with an 'off' state at high CO_2. What forcing is required to transition from a 'strong' AMOC state into an 'off' state under high CO_2?

Yes, the stability of the Off state is derived from simulations starting with an initial large FWF hosing of +0.5 Sv. This is now more explicitly stated in the caption of Fig A1, as explained in the response to comment 1 above. Actually, as shown in Figs. 4 and 5, already a freshwater flux of 0.2 Sv is sufficient to bring the AMOC to an Off state under any CO_2 concentration.

4. What forcing is required to tip from a 'modern' into a 'strong' AMOC? What are the climatic impacts of this shift? Do the authors think that tipping into an 'off' or 'strong' AMOC is possible under persistent anthropogenic climate change?

If we understand this correctly, here the Reviewer is concerned about transient future climate change. This can of course only partly be addressed with our equilibrium experiments presented in the paper. Nevertheless, it is clear from our results that both the Off and Strong AMOC states are possible outcomes of ongoing global warming, and whether one or the other state will eventually be reached depends probably largely on the rate and amplitude of future warming.
To at least partly address this question, we will add the following to the discussion:
It is hence in principle possible that slightly different future global warming trajectories could lead in one case to an irreversible (on multi-centennial time scales) AMOC shutdown and in another case to a transient AMOC weakening followed by a transition into a Strong AMOC state, eventually resulting in fundamentally different climate conditions in the North Atlantic. Transient model simulations under future emission scenarios will have to be performed to explore this possibility.

Minor points

Page 12: Please add an explanation of the grey and black arrows and the black dots to the figure caption or a legend.

We will explain this figure in much more detail, as explained in the response to the main comment 1 above.

Page 14: Please add a caption for Fig B2.

We apologize for having overlooked the missing caption and will add the following caption to Fig. B2:
Change in the net freshwater flux into the ocean as a function of global temperature change in transient historical and future simulations under the SSP2-4.5 scenario until the year 2300 CE for (a) the northern North Atlantic and Arctic (north of 50° N) and (b) the whole Atlantic ocean. The solid line is for CLIMBER-X results and the circles represent CMIP6 model results.