Response to Reviewers

We thank both reviewers again for the time they dedicated to reviewing our manuscript and the helpful comments they both provided. We have changed the manuscript following the suggestions of reviewer one. Please find our point-by-point response below.

Response to Reviewer 1

M1 - Timescales: The authors state that the (effective) timescales are set to zero (L 99), and this shortcoming is discussed in several parts of the manuscript. While I definitely appreciate that, I am wondering why the authors still state tipping probabilities in the year 2500, and I don't think their current rationale holds (ice sheets have longer timescales, also considering their potential to trigger further subsequent tipping elements). Also, the authors cannot (and also do not claim) to make a difference between triggering and realization of the tipping element.

Given these shortcomings and in line with reviewer 2, I recommend to get rid of the timescale notion (in particular the 2500) where timescales are actually not taken into account. Rather, the authors could speak of crossing tipping points for equilibrium temperatures (that are different for each of the SSPs) because that is what the authors essentially do.

As such, I also think that the difference between tipping probabilities between (i) Delayed, (ii) Instantaneous tipping and (iii) Instantaneous tipping + carbon feedbacks is a very neat and helpful differentiation and analysis. Still, since the authors cannot take into account the timescales of tipping, (i) could be seen as a lower bound and (ii) as an upper bound of tipping probabilities – this could even be framed like this in the paper (if the authors agree). The current framing as instantaneous and delayed tipping, however, again doesn't really make sense because it gives the impression that timescales are investigated but they are not. Therefore, a different wording is needed.

Summarized, I think the many places in the manuscript that still indicate these timescales should be rewritten (in particular abstract, conclusion/discussion, and figure captions).

Thank you for reconsidering this point. We agree that concentrating on the 2500 time horizon might be confusing and doesn't add much additional information. Therefore, we do not use this notion any more, but speak of the probabilities in a more general sense as probabilities of triggering given a certain SSP.

We also renamed the "probability of delayed triggering" to "probability of equilibrium triggering", recognising the point the reviewer raised about this being the response to equilibrium temperatures. However, we still think that it makes sense to define instantaneous and equilibrium triggering by the different assumption we make about the effective timescale

of the tipping elements. This allows for a more meaningful discussion of the source of the uncertainty expressed by the difference of the two estimates, namely the unknown effective timescale. We mention that the two estimates can be interpreted as a lower/upper bound of triggering probabilities in L 109, if TE interactions are not considered. Adding the carbon TEs shows that the upper bound of triggering probabilities (instantaneous triggering) is increased by TE interactions.

The different probability estimates are now introduced as follows (L97-L111):
Instead, we adopt the most simple case for which the effective timescales are zero, i.e., a TE is triggered instantaneously once its threshold temperature is crossed. To show that this approach would lead to an overestimation of probabilities of triggering for emission scenarios producing a temperature overshoot, we also discuss the case of equilibrium triggering. Here, we assume that the effective timescales of the TEs are long compared to the overshoot time, i.e., a TE is only triggered if the stabilized temperature at the end of the model period exceeds the threshold temperature. The real probability of triggering will be somewhere between the probability of instantaneous triggering and the probability of equilibrium triggering, but remains unknown.

To analyse how carbon TEs and our assumption about the effective timescale of TEs affect the probabilities of triggering, we derive three estimates of the probability of triggering with different degrees of conservatism: equilibrium triggering, instantaneous triggering, and instantaneous triggering including the effect of carbon TEs. Distinguishing between the probabilities of instantaneous and equilibrium triggering allows us to estimate the magnitude of the uncertainty in the triggering probability resulting from not knowing the effective timescale of the TEs. The probability of instantaneous triggering can be interpreted as an upper bound and the probability of equilibrium triggering as a lower bound on the probability of triggering if interactions between TEs are ignored. The third probability estimate allows us to investigate how much the upper bound of the triggering probabilities could be increased by carbon TEs.

We explain in more depth how those probability estimates are derived in the methods section (L264 - L270) and discuss the implications of the different estimates in L422-L429.

M2 – calibration: L 112-121: I understand that the authors chose the Leach et al. (2021) calibration of FaIR, which is unfortunately not calibrated to the IPCC AR6 runs. While this is a shortcoming, I understand that this cannot be changed anymore at this point in time. That's fine with me, but I think a clear statement (and maybe a brief limitation discussion, how large are differences to be expected) of this in the methods is required.

We agree this needs to be made clearer. We now discuss this point more thoroughly, starting from L 430:
The probabilities of triggering derived by us might be slightly overestimated, since by the climate sensitivity of FaIRv2.0.0 is not well constrained towards its upper limit (Leach et al., 2021). This is a result of this version of FaIR not being calibrated to match the IPCC range of climate sensitivity. This has been fixed in later versions of the model we were not aware of when conducting this study. Nevertheless, we regard this possible overestimation to be

small, since the median climate sensitivity of Fairv2.0.0 agrees well with the latest IPCC estimate (Forster et al., 2021).

Minor comments:

1) The authors find that the tipping effects from the feedbacks of the carbon tipping elements are relatively minor (around 3% tipping probability difference). That is a nice finding and in itself conservative because the authors only take the Amazon and the permafrost into account but not the ice sheet feedback (on long timescales of course) nor the feedbacks from sea-ice, etc.. While this is of course not necessary for this paper, the authors could briefly state that their estimate is conservative and may be higher if additional feedbacks would be taken into account.

This is also our understanding, we discuss this point in L 460:
Even though we find an increase in the probabilities of triggering caused by the additional carbon emissions from carbon TEs, this impact is not strong enough to trigger any tipping cascades and remains small compared to the scenario–dependence of tipping probabilities. Even under SSP2-4.5, which features the highest long-term increase in temperature of up to 0.91°C due to carbon emissions from carbon TEs, the additional probability of instantaneous triggering caused by this temperature increase is only 3 pp on average. Despite our finding that the impact from carbon TEs alone is too small to trigger tipping cascades, tipping cascades might still emerge as major physical interactions between TEs aside from carbon emissions are not accounted for in this study (Wunderling et al., 2021).

2) L 64: “Since the northern expansion and southern dieback … balance out”. I am good with leaving this out, but it is uncertain whether it balances out in the end and probably depends on the speed of anthropogenic climate change (I assume we are faster than northern growth).

We agree that there is a lot of uncertainty whether the two will balance out or not, we rely on the estimate of Armstrong et al. (2023) for making this statement. Nevertheless, even if the two won’t cancel out, their net emissions will likely be small compared to what could be expected from permafrost degradation.

Some additional and very new literature that may be worthwhile considering in the current version of the manuscript:
3) Around L 70: Here, a new review on (no) permafrost tipping points could be cited:
Nitzbon, J., Schneider von Deimling, T., Aliyeva, M., Chadburn, S.E., Grosse, G., Laboor, S., Lee, H., Lohmann, G., Steinert, N.J., Stuenzi, S.M. and Werner, M., 2024. No respite from permafrost-thaw impacts in the absence of a global tipping point. Nature Climate Change, pp.1-13.
4) L87: The paper backing up the science panel for the amazon reference (2021) is here:
Flores, B.M., Montoya, E., Sakschewski, B., Nascimento, N., Staal, A., Betts, R.A., Levis, C., Lapola, D.M., Esquível-Muelbert, A., Jakovac, C. and Nobre, C.A., 2024. Critical transitions in the Amazon forest system. Nature, 626(7999), pp.555-564.
5) L 97-99: I agree with the authors that timescales of tipping elements are still very difficult to assess; somewhat similarly to predicting the tipping times in the first place: Ben-Yami, M.,

Morr, A., Bathiany, S. and Boers, N., 2024. Uncertainties too large to predict tipping times of major Earth system components from historical data. Science Advances, 10(31), p.eadl4841.
6) L 455: Add the new GRIS tipping point paper by Bochow et al, 2023, Nature (basically a follow up of the earlier Robinson et al., 2012 paper): Bochow, N., Poltronieri, A., Robinson, A., Montoya, M., Rypdal, M. and Boers, N., 2023. Overshooting the critical threshold for the Greenland ice sheet. Nature, 622(7983), pp.528-536.
7) A study that couples FaIR with tipping probabilities under overhoots: Möller, T., Högner, A.E., Schleussner, C.F., Bien, S., Kitzmann, N.H., Lamboll, R.D., Rogelj, J., Donges, J.F., Rockström, J. and Wunderling, N., 2024. Achieving net zero greenhouse gas emissions critical to limit climate tipping risks. Nature Communications, 15(1), p.6192.

Thanks a lot for bringing up this new literature, it's highly appreciated! We cited the papers where we thought it was fitting, mostly following your suggestions:
Nitzbon et al., 2024 in L73
Flores et al., 2024 in L87
Bochow et al., 2024 in L452

Response to Reviewer 2

Thank you for taking the time to review our paper three times now. Your comments were very helpful for making this a better paper.