Reviewer1:

The authors well responded to my comments and the revised manuscript reads well. This paper provides important insights into the driving mechanisms behind back-arc extension. Before accepting the manuscript, I would only suggest the following minor comments:

1. the importance of the manuscript is about analyzing the stress field evolution. However, the simulated stress field is only in a too small figure. I zoomed to 400% to Fig. 2 and still struggled to see the values of the stress field as written in the text. Please either make a separate figure of this or really zoom into the lithospheric domain that is discussed. For instance at 12 Myr, does the horizontal stress field shows contrasting extensional vs compressional stress on the two sides of the thinned area and in the upper vs lower part of the lithosphere? Can you comment this? Authors have edited the stress field (zoomed in further and used a different colour bar to make the stress clearer). The region that undergoes thinning is no longer thinning but healing at 4 Myr, and the stresses are not driving any more deformation after that, so the horizontal stress then does not need to be the peak extensional horizontal stress. We didn't show the time point just before the thinning, because the main purpose of this figure is to show the overall evolution of the subduction zone and not the full history of the stress field (we note that the history of this stress is summarised in Figure 11).

2. Fig. 4: there are plots of the viscosities based on the figure caption, but those are too small and not clear. Thank you for the suggestion. The authors have reduced the numbers of the snapshots and split figure 4 into 2 figures to make everything easier to read.

3. Please add scale and coordinates in fig. 10c. Edited in what is now fig.11.

Conclusions: Please also state here that your conclusions are derived from 2D simulations with a mobile upper plate. Edited in the text (Line 396-397).

Attila Balazs

Reviewer2:

I previously reviewed the original submission (2nd reviewer) and am now reviewing the original submission. I'm sorry for the delay with both of my reviews! It's been a really busy period. My main concerns centered around the mechanism descriptions, upper plate rheology, and the very rapid subducting plate velocities. The authors have worked on addressing each of these concerns but, in my opinion, a bit more work is needed for this to be sufficient. I'd therefore recommend "moderate" revisions. Below are my line-by-line comments:

Before the line-by-line: In code availability, you state that "the results of our models are available from the corresponding author upon reasonable request". I recommend sharing the full input files in a permanent citable repository, as is now standard practice (and I think the Solid Earth data policy requires this). See recent geodynamics Fluidity modeling studies that do this (e.g., Chen et al., 2022, https://doi.org/10.1029/2022GC010757). OK, thank you, the authors will do this.

Abstract: Confusing to write "Mode EH" and "Mode EF" in the abstract - just describe. Authors have edited these references in the text (Line 3-5 in the revised version).

Line 48: How is it "based on the work of Garel et al." Do you mean that you started with their models and then added the hot regions? I recommend being more precise. Yes, that's what the authors mean. In section 2.1, the authors have described it by introducing Garel's model firstly and then the hot region. Thank you for your suggestion which gives us a chance to have a look at this text again. For a more precise description, the authors have edited Section 2.1 (Line 60-61 in the revised version) and hope it would be clearer to readers.

Eq. 4 and surrounding text: The pressure description is still confused. The "full" stress is made of the full pressure (lithostatic + dynamic) and the deviatoric stress. Thank you for pointing out the potential confusion. 'σ_{ij}' is the stress tensor, where the lithostatic pressure is implicitly removed since only the lateral density variations are applied on the right-hand side of equation 2. So Eq.4 should use dynamic pressure. We have edited lines 90-91 in the revised version to make the description of the pressure and stress clearer.

L100-105: Where are these parameters taken from? Dry or wet olivine? Needs a reference. These parameters are from experimental estimations of dry olivine (Hirth and Kohlstedt, 2003; Karato and Wu, 1993; Karato, 1997). Authors have added these references in the text (Line 105-106 in the revised version).

Eq. 9 and surrounding text: One of my previous comments related to justifying the yield stresses imposed in the models. I understood this is a common approach – and that it's intended to mimic brittle failure – but I was more interested in how you justify these parameter values (cohesion, friction coefficient). Can you provide further details / reference(s) here? Many papers don't provide such details but, in the case of your study, it's especially important as the focus is on breaking the plate. Otherwise, you could show tests showing that the first-order conclusions do not depend on mu = 0.2 vs., e.g., mu = 0.4 (as you state in your original rebuttal). The friction coefficient of 0.2 in our models is intermediate between lower values of previous subduction models (Di Giuseppe et al., 2008; Crameri et al., 2012). Authors have added this note and references into the text (Line 120).

L120: "like viscous dashpots in series" -> "as is the case for viscous dashpots in series". Authors have edited it in the text (Line 125).

L132-136: As also identified by both reviewers, the subduction velocities are extremely large: this is likely to promote excessive extension (as stresses scale approx. with velocities). Given this, I think more details are needed: You say that there is a local peak of "> 10 cm/yr". What is the local peak, actually? Thank you for this question. The actual peak velocity varies between models, here it is about 38 cm/yr in the reference model. And does extension always coincide with this peak? These details should be added to the main text. Yes, we note that in cases with extension this coincides closely with this peak (this is shown in fig.9 in the revised version), they both happen just before the slab tip reaches the more viscous lower mantle. We didn't emphasise this point because the extension doesn't always appear at the same time as the velocity peaks, but around it (a little before or after), so we think they cannot be strictly related, but we have qualified the text on line 141 to point to this broad correlation in time.

Figure 2: Stress should be in MPa (i.e., +/- 350 MPa). There is no velocity vector scale on the bottom panels. And what is "velocity component" (velocity magnitude?) We have changed the stress scale to MPa. Sorry, this figure does not show a velocity component, authors have edited 'velocity component' to 'velocity magnitude' in the figure. And why is it limited to 3.5 cm/yr (when you've stated that velocities are much higher)? In the reference model, the peak velocity is not that high. The high velocities always appear in models with weaker overriding plate and older (more negativey buoyant) subducting slabs. Also we should note, since the very high peak velocities are only very brief features, that this relatively higher velocit is between the outputs we have shown and so we didn't catch that the peak velocity.

Figure 4: The new panels (viscosity profiles) are way too small to read. The viscosity panels are probably also too small. Break the figure up into 2? Thank you for the suggestion. Author have reduced the numbers of the snapshots and split figure 4 into 2 figures to make everything easier to read.

L198: Can you quote some trench retreat velocity magnitudes? That's so readers can compare with other models/observations. Thank you for this suggestion. Authors have quoted the trench retreat rate in bracket. (Line 203)

L241: Saying that, in the first mechanism, the extensional stress is generated from the speed difference between the trench and the OP is not strictly true. That would thicken/thin the plate interface, which is not what you are talking about. We prefer to keep the current wording, noting that the 'trench' could be considered to be on the OP side of the plate interface. We are looking at the largest scale processes acting on the OP, and believe in this context this description is correct.

243: "a little more" – too informal. Authors have edited the whole sentence to 'We further investigate these aspects to find which are most important in our models.' (Line 248)

Figure 10: Panel B looks bad – can you plot in the same format as panel A. Also, did you integrate over the same depth range? Or the depth of the plate (which varies with age)? If the latter, I recommend dividing by this plate thickness to get the average plate stress. We have altered panel B of old figure 10 / new figure 11 to be in the same format as panel A. We integrated the stress over the depth of the plate which varies with age. Thank you for your suggestion and we have edited the figure to display the average plate stress.

253-255: I don't think this is correct (similar point to my original review). Yes, this shows that an in-plate stress increase is not the main extension initiation; rather it's the weakening in the upper plate. Agreed. But it does not discriminate between extension driven by force transmitted through the interface and that driven by underlying tractions. Both induce horizontal extension in the OP. Thanks for pointing it out. Authors have deleted 'applied at the OP/SP interface' in Line 264 and hoped that would reduce confusion.

354: Clarke, Stegman, Muller (2008, PEPI) is a nice paper about this. Their work should be referenced. We added this reference into the text. (Line 359)

Sect. 4.4 (Limitations): Bullet points 1 ("2D modeling") and 2 ("Simplications of the models") overlap (as 2-D is also a modeling simplification). More importantly, and following previous comments, I think one of the biggest simplifications is the combination of a very high subduction velocity and a relatively low yield stress (which both promote extension). I think this should be acknowledged. Thank you for these suggestions. To solve the overlapped classification, we have added 'other' before 'simplification of the models' (Line 380). We also have added the high subduction velocity into the limitation (Line 392-393), but we do not consider the maximum yield stress of 10 GPa as being particularly low (10 GPa) nor the friction coefficient.