Response to comments from Reviewer #2,

##General comments:

I am not an expert in the regional geology of the Zagros fold-and-thrust belt, which is why I can only comment on the modeling aspects of the study.

> *We would like to thank Reviewer #2 (Lorenzo Giuseppe Candioti) for careful reading and constructive review of the manuscript.*

The authors present 2D thermo-mechanical (one-way coupled) numerical models of tectonic extension and subsequent compression applied to the Zagros fold-and-thrust belt. The model features pre-existing basement faults and a salt layer that acts as décollement horizon. Varied model parameters include (1) the thickness and rheology of the salt layer, (2) the geometry of pre-existing basement faults, and (3) the horizontal velocity of the basement during convergence. Generally, the study is well written and logically organized. However, I think the current version of the manuscript needs further editing. In particular, (1) some aspects could be discussed in more detail and (2) the research question and main contribution of the study could be presented in a clearer way.

Introduction: The main insights from previous studies on thin- and thick-skinned tectonics are summarized and introduced well. In addition, the work of Kiss et al. 2020, Spitz et al. 2020, and Humair et al. 2020 could be introduced here as well. They also presented 2D and 3D geodynamic models that highlight the importance of tectonic inheritance for the evolution of fold-and-thrust belts. I would like to read about the observations that all these previous models did not capture. This would help putting this study into perspective.

> *Thanks for comment. We added a new paragraph to the introduction:*
>
> *"Various 2D numerical modeling studies have investigated the evolution of fold-and-thrust belts and salt-bearing basins. Nilforoushan et al. (2013) demonstrated the influence of geothermal gradients and basement mineralogy on fault geometry and basement reactivation in the Fars arc, emphasizing the role of weak salt horizons in mechanical decoupling. Heydarzadeh et al. (2020) analyzed factors such as sedimentation rates, erosion, and salt layer properties in Dehdasht basin, highlighting the importance of balanced surface processes and deformation rates. Humair et al. (2020) conducted simulations to study the interaction of folding and thrusting during Swiss Jura and Canadian Foothills fold-and-thrust belt evolution, focusing on the effects of layer-parallel shortening and initial geometrical perturbations. Their work showed that the magnitude of these perturbations influences whether folding or thrusting predominates, affecting the structural evolution and asymmetry of anticlines. Spitz et al. (2020) conducted 3D thermo-mechanical numerical simulations to investigate the influence of laterally variable inherited structures on fold-and-thrust belt evolution and nappe formation on Helvetic nappe system. The study demonstrated the fundamental importance of tectonic inheritance on fold-and-thrust belt evolution, with strain localization, folding, and nappe transport controlled by initial geometrical and mechanical heterogeneities. Almost all studies have focused on examining the collisional phase and deformation resulting from compression in the fold belts and the Fars Arc, while the earlier extensional history and its effect on later deformation have received less attention (e.g., Granado and Ruh, 2019)."*

Geological setting: Not being an expert in the regional geology of the Zagros, it seems to me that a lot of research has been conducted on the study area. What could be more focussed on is what exactly remains poorly understood about the geological evolution of the Fars Arc. Also, the particular geologic problem this study is addressing should be clearly outlined in an additional paragraph, and how the presented models will help gaining new insights.

> *Thanks for the comment. In this part of the paper, we introduce the geological setting of the Fars arc and mention the tectonic history of the Zagros fold-thrust belt and the stratigraphic column of the study area based on previous studies. As you suggested, we added an additional paragraph at the end of the geological setting section concerning your comment:*
>
> *"Many studies have been conducted in the Zagros fold-thrust belt and the Fars arc. However, several critical aspects of the geological evolution of the Fars Arc remain poorly understood. Specifically, the precise mechanisms and timing of basement involvement, the interaction between basement faults and salt décollements during tectonic inversion, and the relative contributions of thin-skinned versus thick-skinned tectonics to the overall structural evolution are not fully resolved (see Mouthereau et al., 2006; 2012). We employ a numerical model that simulates the complete tectonic history of the Fars Arc, including an initial extensional phase followed by a compressional phase."*

Results: I think it would help the reader to interpret the results, if the figures were larger. In the current state, I find it challenging to identify all the important details the models seem to predict. It might also be interesting to see stress, temperature, or viscosity fields for at least the reference model. This would make it a lot easier to identify rheological boundaries and structures.

> *Thanks for pointing this out. We request the journal to input the figures in landscape, to make them better visible. Furthermore, we added a new figure 5, where viscosity and stress of the reference model are discussed individually:*
>
> *"Patterns of the second invariant of the stress tensor after the extensional phase and after full convergence indicate an increase in stress with depth down to ~30 km (y = 50 km), where the brittle-to-ductile transition begins (Fig. 5a,d). The lower part of the basement displays low stresses given its lower viscosities and ductile nature (Fig. 5b,e). Viscosity plots furthermore indicate the position of the low-viscous décollement and basement thrusts."*

Discussion: Agreements between the presented results and previous studies are discussed well. This suggests that the models are capable of making some realistic predictions for the evolution of the Fars Arc. I think the discussion would benefit from highlighting the advantages of the presented models compared to previous models and how they help gaining new insights into the evolution of fold-and-thrust belts in general. For example, one novel aspect of the models presented here seems to be that rifting is modeled prior to collision. I would like to read some *discussion on that particular model feature. Why is it important and what advantages does it bring compared to models that only focus on collision?*

> *Thanks for the comment, we have added a paragraph for clarification of the impact of rifting phase in fold thrust belt numerical modeling:*
>
> *"The consideration of rifting prior to collision represents a new aspect that is commonly ignored in previous studies related to the Zagros fold-and-thrust belt. This approach provides a more comprehensive understanding of tectonic evolution by allowing the investigation of pre-collisional structural configurations and their influence on subsequent deformation patterns. Including rifting before convergence is crucial as it sets the initial conditions that significantly impact structural evolution during collision (Buiter and Pfiffner 2003; Ruh et al., 2018; Granado and Ruh, 2019). This approach offers several advantages: It provides insight into structural inheritance, as rifting creates pre-existing weaknesses and fault systems that play a crucial role during subsequent compressional phases. Our model demonstrates how these inherited structures influence strain localization and deformation styles, providing insights into the evolution of complex geological features such as the Fars Arc."*

##Specific comments:

Lines 71-76: Seems like more recent studies favor the latter hypothesis. Are both hypotheses still equally supported by all the data collected so far?

> *We modified this sentence with respect to a comment by reviewer #1:*
>
> *"In the Fars Arc, the activity of inherited faults has affected the progression of deformation towards the foreland, with the Mountain Front Fault being related to basement thrusting (Bahroudi and Koyi, 2003; Mouthereau et al., 2006; Mouthereau et al., 2007a; Yamato et al., 2011; Ruh et al., 2014; Najafi et al., 2021)."*

Lines 80-82: It would help the reader to better understand the general relevance and importance of this study, if it was clearly stated what is poorly understood and why it is crucial to close this knowledge gap, here. Especially, since the previous paragraph outlined a certain degree of agreement in the community on the role of the décollement layer and the basement faults in that region.

> *We have added a paragraph for clarification of the impact of rifting phase in fold thrust belt numerical modeling:*
>
> *"Understanding these processes is crucial for improving our geological models of the region, which has significant implications for hydrocarbon exploration. Although there is a general consensus on the importance of the décollement layer and basement faults, the detailed dynamics and their broader impact on regional tectonics require further investigation in during inversion tectonics."*

Line 181: Is the material density assumed to be constant?

Yes, we ignored density changes due to thermal expansion or compression because their impact in the crust is minimal in the absence of phase transformations. We mentioned it at the corresponding place in the text.

Line 185: Why are contributions from viscous dissipation and radiogenic heat excluded? Are material conductivity and heat capacity constants? Their values do not seem to be provided.

These additional heat production mechanisms in the crust are indeed not included in our current model. The relatively short numerical runtime of our modeling means the cumulative effect of these heat sources would be minimal. These parameters are set as constants for each rock type in our model. Our primary interest is in the mechanical deformation, where small variations in thermal properties would have minimal impact on the overall results. For most rock types, we use a thermal conductivity of 2.5 W/(m·K) and a heat capacity of 1000 J/(kg·K), which are typical values for crustal rocks. We added the values in Table 1 for clarity.

Line 194: typo -> visco-elasto-plastic/brittle

Thanks, has been corrected.

Line 196-199: The rotational and advection terms are missing in the objective derivative. How is rigid body rotation included in the model? This is quite crucial for folding simulations (see Schmalholz et al. 2001).

The numerical model in this study employs a simplified form of the objective co-rotational time derivative for visco-elastic stresses, as shown in equation (5):

$$\frac{D\tau_{ij}}{Dt} = \frac{\tau_{ij} - \tau_{ij}^{old}}{\Delta t}$$

This simplification is a common practice in geodynamic modeling, as seen in works by Gerya and Yuen (2007) and Moresi et al. (2007). While this form does not explicitly show the rotational and advection terms, it is still capable of capturing the essential physics of visco-elastic deformation in many geological settings.

We added the references in the text.

Line 200: Where does the prefactor 0.5 come from? Is this an additional weakening factor? I suggest calling this viscosity eta_dis or eta_visc, for clarity. Also, shouldn't there be a factor that accounts for the conversion of the experimental 1D flow law to a flow law for stress tensor components?

The 0.5 factor is applied as a simplification, independent of the specific deformation type (e.g., shear, plane strain). The equation for effective viscosity is derived from a commonly used form in laboratory experiments: $\dot{\varepsilon} = A.\sigma^n.exp\,(\frac{-Q}{RT})$*. With the assumption of* $\sigma =$

$2\eta\dot{\varepsilon}$ *(as we use the second invariant,i.e. deviatoric components), this equation converts to*

$$\eta = 0.5 \cdot \frac{1}{A_D} \cdot \sigma_{II}^{(1-n)} \cdot exp\left(\frac{Q}{RT}\right)$$

We use η_{disl} *in the equations.*

Line 206: Is delta_t_e the Maxwell relaxation time? Why is it set to 1000 yrs? Why isn't the physical time step used here?

The selection of Δt_e *at 1000 years is typically a compromise. It is selected based on the need for numerical stability, accuracy, and the ability to capture the stress history effectively over geological timescales.*

Using the physical time step directly for updating visco-elastic stresses would not be practical. The physical time step in numerical models is often chosen based on the smallest relevant timescale for the problem at hand, which can be much shorter than the timescales over which elastic effects are significant. The use of a separate elastic time step, rather than the physical time step of the simulation, allows for a controlled implementation of elasticity in the model.

We added the information "Maxwell timestep" in the text.

Lines 207-209: Z should be dimensionless, but it is not. Where does this formulation come from and why is it used?

Thanks, It was a typo in the numerator, we mistakenly used a plus sign instead of multiplication. Has been corrected.

$$Z = \frac{\Delta t_e \cdot G}{\eta + \Delta t_e \cdot G}.$$

Z is derived from rheological studies, particularly those involving Maxwell-type visco-elastic models (e.g., Gerya and Yuen, 2007, Moresi et al., 2007, Moresi et al.,2003). This parameter is crucial for interpolating between the elastic and viscous responses of the material within a given time step, thus ensuring numerical stability and consistency in stress updates. The references have been added.

Line 212: There are brackets missing in the equation. Eta_num does not seem to be used in any other equation noted here. Where is it used in the algorithm?

Thanks, It was a typo in the numerator, we mistakenly used a plus sign instead of multiplication. Has been corrected.

$$\eta_{num} = \eta_{disl} \cdot Z = \frac{\eta_{disl} \cdot \Delta t_e \cdot G}{\eta_{disl} + \Delta t_e \cdot G}.$$

The term η_{num} *is introduced as a numerical (visco-elastic) viscosity, derived from the viscous viscosity* η_{disl}*, which is important for solving the set of equations governing these processes and is used to stabilize the numerical solution by adjusting the effective viscosity based on the time step and the plastic potential.* η_{num} *would be integrated into this framework to ensure that the visco-elasto-plastic behaviour is accurately represented (Gerya and Yuen, 2007).*

Lines 218-219: Sigma_xx and Sigma_xy have not been introduced before, are those components of the total stress tensor? In this formulation it seems that stress tensor components are increased if stresses are below the yield criterion. If this formulation is only valid in case of yielding, maybe using the mathematical notation of cases (curly brace) in the equation would make this formulation clearer.

"The components of stress (σ_{xx}*is the normal stress component, and* σ_{xy} *is the shear stress component) and the viscosity are then updated as:"*

Also, we presented the equation using the mathematical notation of cases (curly braces):

$$\sigma_{xx}^{new} = \begin{cases} \sigma_{xx}\dfrac{\sigma_y}{\sigma_{II}},\ if\ \sigma_{II} > \sigma_{yield} \\ \sigma_{xx},\ if\ \sigma_{II} \le \sigma_{yield} \end{cases},$$

$$\sigma_{xy}^{new} = \begin{cases} \sigma_{xy}\dfrac{\sigma_y}{\sigma_{II}},\ if\ \sigma_{II} > \sigma_{yield} \\ \sigma_{xy}, if\ \sigma_{II} \le \sigma_{yield} \end{cases},$$

Line 223: This seems non-standard, especially for power-law rheology. Why are calculations not performed on the Eulerian grid?

This approach is advantageous for complex rheologies and geometries, such as power-law behavior, because it allows for an increased resolution tracking of material properties and interfaces. The flexibility of the Lagrangian markers ensures that the non-linear and history-dependent nature of visco-elasto-plastic materials is accurately captured. To us this has been standard so far.

Line 225: Formatting. 10^25 appears as 1025

Thanks, has been corrected.

Lines 234 f.: How are basement faults parameterized in the models?

Thanks for pointing this out. We added another line in Table 1 listing the parameters of the basement faults.

Table 1: What is the underlying assumption for choosing a fluid pressure ratio of 0.4? Values for Cohesion seem to be very low even before softening. What was the motivation to choose

such low values and to reduce them further as a function of strain? Why are the sediments parameterized by the same flow law but different densities and friction angles?

The choice of a fluid pressure ratio (λ) of 0.4 in the model is based on typical values for hydrostatic fluid pressure as a first-order approximation. The rheological properties of the sediment are based on quartzite. According to the stratigraphic column of the Fars Arc, the nature of the sediments layers varies depending on the tectonic history. These sediments include evaporites and shales, which are weaker than clastics. Either way, these rocks mainly deform in a brittle manner.

Lines 243-245: Between the extension and the convergence period in the Zagros there seems to be a 180 Myr period of inactivity which can have an impact, especially on the thermal field and the dynamics in the asthenosphere below, which may in turn have control on the lithospheric deformation. We have shown this in a series of publications (starting with Candioti et al. 2020). As far as I understand, the model switches from extension to convergence instantly. Why is the rifting period included but the passive margin period excluded in the models presented here?

We ignored the phase of tectonic quiescence to simplify the modelling approach. We added a sentence in the corresponding section:

"The phase of tectonic quiescence between rifting and convergence is ignored here for simplification of the model setup."

Lines 371-375: A figure showing the stress and temperature field of the described models would be helpful to support the line of argumentation here.

The corresponding line were removed during the revision process.

Line 461-462: This is likely a result of the one-way thermomechanical coupling and the strain weakening. I suspect that this frictional weakening of already weak lithologies promotes immediate material failure under compression. In that case, visco-elastic stresses cannot be build up to significantly high values and then be released when shear zones form. I would also not expect to see this effect in Fig 10e. Instead, this may explain the signal pattern (if vx_b < vx) in Fig. 11: In absences of stress built-up and release, the only signal recorded is crustal thickening. As the belt grows, more and more force is necessary to drive the convergence at constant speed. I suspect that if the lithologies were stronger and shear heating would be considered, stresses would build up to higher values and then drop once a shear zone forms. This might then also be visible in at least Fig. 11 (compare to Fig. 11b in Candioti et al. 2021). How do the values for forces compare to estimates for collision zones in general?

As stated by the reviewer, it is a competition between vertical growth increasing strength and weakening of faults decreasing it, as outlined in the text. As for a comparison, there is not much information on fold-and-thrust belts, and in mantle-scale experiments, the mantle lithosphere adds to boundary force by a large part. However, we added a reference from analogue modelling comparing the geometry of the temporal evolution of total energy consumed:

> *"A similar pattern of boundary force evolution was also observed from analogue models, however, with more acute variations for specific internal deformation of the compressed sand pile (McBeck et al., 2018)."*

Line 491: Missing word „of"

> *Thanks, has been corrected.*

Line 497: Missing word „to"

> *Thanks, has been corrected.*

Lines 595-596: The diapir in the rifting model is hardly visible. An enlargement or generally larger figures would help identifying the diapirs in the models. A brief discussion about earlier work (e.g., Fernandez & Kaus 2014) would be suitable here.

> *Thanks for spotting this. The figures will be in landscape orientation and therefore better visible. Thanks for suggestion, the studies are discussed:*
>
> *"Fernandez and Kaus (2014) showed with numerical modeling that pre-existing salt diapirs can significantly influence the pattern and growth of three-dimensional folds and fold patterns, accelerating fold formation and localizing deformation, highlighting the important role of diapirism in structural evolution during tectonic processes."*

Figure 12: I have the impression that the models are generally dominated by faulting whereas the reconstruction seems to show more folding dominated deformation. It would be interesting to see a movie that shows the folding and thrusting in one of these models.

> *That is a good point. We added now GIF movies of all experiments in the Supplementary Material which can be accessed from "Data availability" section of the Manuscript.*
>
> *The GIF movies of all experiments that support the findings of this study are available in Figshare with the identifier: https://figshare.com/s/38141397f97519f7dc31.*

Line 643-644: The depth of the brittle-ductile transition does not only depend on material parameters and the temperature, but also on strain rate (stress) among other variables. Depending on local conditions, this depth can vary for the same material parameters. It should therefore be generally possible to get similar depths of the brittle-ductile transition for different material parameters at different conditions. Hence, it might not be the best justification for the choice of flow law parameters here.

> *Thank you for the comment. We have revised this part of the discussion accordingly:*
>
> *"By examining the deformation within the basement and the earthquake depths, we estimate the transition zone from brittle to ductile behaviour at around 30 km. While this*

depth is influenced by multiple factors, including material parameters, temperature, and strain rate (stress), our approximation is based on typical conditions relevant to the region. This suggests that diabase rheology, combined with the applied geotherm, is appropriate for modeling the basement under these specific conditions. Mouthereau et al. (2006) on the other hand concluded that diabase might be too weak to reproduce the observed topography in the Fars arc. However, they apply a crustal thickness of 45 km, which implies increased temperature conditions and thus a weaker diabase detachment."

Figure 13: A description of panel d is missing.

Thanks, the figure caption has been corrected.

"Figure 13: (a) Seismicity of Zagros fold-and-thrust belt reported by ISC during the years 2000 to 2023 with magnitudes ML> 4 and local temporary network data by Tatar et al (2004) with magnitudes ML<4 , superimposed on a shaded relief map derived from Global Topography. (b, c, d) Projection of earthquakes along profile A–B after removing fixed depths on geological cross section and the reference model (Model 1) after 15 Myr of convergence."

References:

Candioti, L. G., Schmalholz, S. M., & Duretz, T. (2020). Impact of upper mantle convection on lithosphere hyperextension and subsequent horizontally forced subduction initiation. Solid Earth, 11(6), 2327-2357.

Candioti, L. G., Duretz, T., Moulas, E., & Schmalholz, S. M. (2021). Buoyancy versus shear forces in building orogenic wedges. Solid Earth, 12(8), 1749-1775.

Fernandez, N., & Kaus, B. J. (2014). Influence of pre-existing salt diapirs on 3D folding patterns. Tectonophysics, 637, 354-369.

Kiss, D., Duretz, T., & Schmalholz, S. M. (2020). Tectonic inheritance controls nappe detachment, transport and stacking in the Helvetic nappe system, Switzerland: insights from thermomechanical simulations. Solid Earth, 11(2), 287-305.

Humair, F., Bauville, A., Epard, J. L., & Schmalholz, S. M. (2020). Interaction of folding and thrusting during fold-and-thrust-belt evolution: Insights from numerical simulations and application to the Swiss Jura and the Canadian Foothills. Tectonophysics, 789, 228474.

Schmalholz, S. M., Podladchikov, Y. Y., & Schmid, D. W. (2001). A spectral/finite difference method for simulating large deformations of heterogeneous, viscoelastic materials. Geophysical Journal International, 145(1), 199-208.

Spitz, R., Bauville, A., Epard, J. L., Kaus, B. J., Popov, A. A., & Schmalholz, S. M. (2020). Control of 3-D tectonic inheritance on fold-and-thrust belts: insights from 3-D numerical models and application to the Helvetic nappe system. Solid Earth, 11(3), 999-1026

We integrated the listed references into the manuscript.