Response to comments from editor,

Dear authors,

Before i can accept the manuscript for publication, i need to ask you for another careful check regarding the description of the numerical model and the modelling approach. Please see my comments below, but i may well still have missed some points and would thus ask you to read especially section 3 critically.

We would like to thank the editor (Susanne Buiter) for careful reading and a constructive review of the manuscript.

The question to keep in mind is if enough information is provided for someone to rerun the models.

To the best of our knowledge, we provide all necessary information on geometry, boundary conditions, and mechanical and rheological constraints to reproduce the numerical experiments presented here. In either case, we carefully went through the manuscript again, with focus on section 3, and answered the points raised below to the best of our knowledge.

I also support the point made by reviewer #2 as this is an opportunity to clarify an essential characteristic of your models. I therefore would like you to explicitly address the question raised by reviewer #2 in your manuscript:

"The difference between models that do not include a rifting phase and those presented here could still be discussed in more detail. If modeling the rifting phase is important for the convergence model, what do the presented models predict that pure convergence models with prescribed weak zones fail to predict?"

Thanks for mentioning the lack of emphasis on this key feature of the presented models. Summarized in a simplified manner, including a previous extension phase allows to investigate the effects of structural and geometric features related to this extensional phase on the structural and mechanical evolution during compression. Obviously, any feature may also be implemented in an initial model before shortening without extension, but the strain-related and structural constraints would be less dynamic. The implementation of predefined faults helps reconstructing the specific location of inherited faults in the Zagros. We added a sentence to the introduction and added a paragraph in the discussion, describing in more detail the importance of a rifting phase for the obtained results:

"Incorporating a rifting phase into the model setup result in more realistic initial conditions for the convergence stage, featuring variations in crustal thickness, rift-related sedimentary basins and the presence of weak zones."

"An essential characteristic of the numerical experiments presented here is the implementation of a rifting phase prevailing crustal shortening. Earlier numerical studies on tectonic inversion demonstrated that strain-related weakening of the crustal basement during an extensional tectonic phase has an important effect on the localization of deformation during consequent convergence (Ruh and Vergés, 2018). Furthermore,

the mechanical strength of sedimentary deposits filling rift-related basins has been shown to influence the structural evolution during tectonic inversion, where weak syn-rift deposits favour the development of hanging wall by-pass structures (Granado and Ruh, 2019). In contrast, various studies mimic the extensional phase by prescribing inherited structures and geometric configurations comparable to rifted margins instead of conducting an extensional phase when investigating the tectonic inversion of fold-and-thrust belts (e.g. Buiter and Pfiffner, 2003; Nilforoushan et al., 2013; Bauville and Schmalholz, 2015; Kiss et al. 2020). However, the implementation of an extension phase allows for a dynamic formation of basement steps and graben geometry based on the thermo-mechanical model characteristics such as the visco-plastic/brittle transition and the evolution of the salt decoupling along the basement top (Fig. 4a-c and 5a,b). Model 7 demonstrates that a certain way of strain localization is needed to form narrow shear bands similar to basement faults during extension (Fig. 9c). While other studies impose specific boundary conditions (Ruh and Vergés, 2018) or thermal variations (Ruh and Vergés, 2018), we introduced predefined geometries to localize extensional normal faults, controlling the position of basement deformation (Fig. 3b)."

In addition, could you contrast the following observations (here in the summary from the abstract) for the models with prescribed weak zones with the model in which inherited faults were not present (Model 7)?

"The inverted basement faults form large foreland-verging fault-propagation anticlines in the sedimentary cover, while the thick salt layer promotes the growth of second-order detachment anticlines accompanied by both fore- and back-limb thrust faults."

Thank you. According to your suggestions, the text has been revised:

"Experiments without prescribed basement faults result in dispersed brittle/plastic deformation during rifting and convergence and an effective mechanical decoupling along the salt horizon."

Model 7 without prescribed weak zones shows a substantial difference to models that have prescribed weak zones. Could you add an explanation for the structural style exhibited in Model 7? My guess is that the salt layer is so weak that it decouples basement shortening from shortening above the salt layer, thus effectively leading to a thin-skinned style in the topmost layer. Fault-propagation folds can however still form and the role of salt is still pronounced. The low rheological strength of the salt layer will also explain the tendency to form conjugate thrusts (pop-ups) rather than vergence in one direction. The topography that is obtained at the base of the salt layer is interesting. Does it form because the lower crust behaves viscous, thus leading to a form of isostatic compensation?

Thanks for your comment, we added a paragraph describing the specific points raised in the discussion, section 5.1:

"All experiments with prescribed weak zones within the basement show intense inversion along these inherited structures (e.g., Fig. 4). A particular experiment is Model 7, where no basement faults were prescribed (Fig. 9c,d). The structural style observed in Model 7 is driven by the low rheological strength of the salt layer, which decouples basement

shortening from the overlying layers, resulting in thin-skinned deformation of the cover sequence (Fig. 9d). This weak salt layer promotes the formation of fault-propagation folds and conjugate thrusts (pop-ups) rather than a clear structural vergence, as observed by a large pop-up structure at x ≈ 230 km. The topography of the top of basement is likely caused by a critical wedge geometry defined by the viscous strength of the lowermost modelled crust."

#Detailed comments:

Please explain somewhere that the numerical model generates shear zones, that you refer to as faults (which many will view as a more discrete feature).

Thanks for the comment. We have mentioned to this point in the initial paragraph of the results section:

"While the numerical model is based on a viscous formulation and thus develops localized shear zones, we refer to them as faults if they form due to the implemented Drucker-Prager failure criterion (see equation 11)."

3.1 Governing equations
Please justify why it is appropriate to exclude heat production from the thermal equation for a model that includes the entire continental crust. How representative is a linear geotherm for continental crust?

Thanks for this comment. There are several reasons for this simplification. One is the geometry and the boundary conditions of the model domain, that impede thermal diffusion and advection. Hence, a simplified temperature description was preferred. We have added more information in the respective section:

"Additional heat production, such as radioactive heating and shear heating, is not activated in the presented experiments due to the geometrical constraints of the model setup and related boundary conditions that affect the diffusion of such secondary heat production."

The continental crust can exhibit variations due to heat sources like radioactive decay, mantle heat flow, or tectonic activity. Based on our simplification for modeling, a linear initial geotherm is assumed, which is a reasonable approximation for the continental crust:

"Applying a linear geotherm is a reasonable simplification for the uppermost ~50 km of the continental lithosphere (Hasterok and Chapman, 2011; Goes et al., 2020)."

3.2 Rheological model
Line 235, gravitational acceleration g has an index in equation 2, but misses the indes in the symbol explanation. I assume it is the vertical component of gravity. Please correct.

Thanks, we have changed to "g_i" ($g_1 = 0$; $g_2 = 9.81 \frac{m}{s^2}$), How important is elasticity for your models? Could you add an estimate of the Maxwell relaxation time and the Deborah number?

We have added more information about the important of elasticity:

"Elasticity plays a key role in capturing short-term stress accumulation and release, crucial for fault and fold behavior. The Maxwell model allows the simulation of both immediate elastic response and long-term viscous flow, ensuring that important transient phenomena such as fault reactivation and seismic activity are accurately represented during both extension and convergence."

Regarding the Maxwell time and the Deborah number, we added the following statements:

"Given the numerical time step of Δt_e = 1000 years, material undergoing deformation at viscosities below 3.16 × 10^{21} Pa·s can be considered predominantly viscous, while deformation at viscosities above 3.16 × 10^{21} Pa·s is to a significant part elastic and thus reversible. Elastic relaxation time varies between ~1 year and 1 million years, depending on the viscosity of the material, which results in Deborah numbers of 10^{-7} – 0.1 for a deformation period of 10 million years."

Table 1: please add the value for shear modulus G

Thanks, we have added the value of G in the text because it is constant for all markers type:

"G is the elastic shear (100 GPa for all materials here)".

Please describe how strain weakening of angle of internal friction phi and cohesion c is achieved.

We added a sentence in the corresponding section:

"Strain weakening is implemented by a linear decrease of the values for frictional angle (φ) and cohesion (c) between a lower and upper strain threshold, defined by $\varepsilon_w^0 = 0.1$ and $\varepsilon_w^1 = 1$."

3.3 initial geometrical setup
Please give the justification for using a 30 km thick crust.

The continental crust in the Zagros is 40 – 45 km thick (Jimenez-Munt et al., 2012; Taghizadeh-Farahmand et al., 2015). However, given the model setup with the induced shear zone (velocity boundary condition) at the base, we are interested in the thickness of the deformed basement. We added a sentence in the corresponding section:

"The initial marker distribution defines, from bottom up, 1) a 30-km-thick crustal basement layer, given the depth of basement crustal detachment (Vergés et al., 2011; Kendall et al., 2019)"

"Eulerian cell initially contains 16 randomly distributed Lagrangian markers carrying rock information and properties". The number of markers in a cell will change because of their movement through the grid. Please describe the policy for empty cells and whether a population reduction is applied to cells with many markers.

We apply a common way to deal with this problem. If no Lagrangian markers are available to be interpolated onto a node, the old values of these nodes are used, until a marker moves into their area of influence again. This concerns all parameters that are important to solve the system of equations such as density, viscosity, elasticity, thermal parameters, etc. However, given the large number of Lagrangian markers and the incompressible manner of the resulting velocity field (what goes out goes in), it is highly unlikely that a cell remains empty. We added a sentence in the Initial geometrical setup section:

"If the finite spatial domains of a specific node become empty of any Lagrangian marker, the previous interpolated parameters are applied for solving the system of equation of this particular node."

And in closing, if the reviews by the two reviewers were useful for your manuscript, and i hope they were, i would appreciate an acknowledgement of the reviewers.

"We would like to express our sincere gratitude to the reviewers, Dr. Frédéric Mouthereau and Dr. Lorenzo Giuseppe Candioti, for their time and effort in reviewing our paper. Their insightful comments and suggestions have significantly improved the quality of the manuscript. Additionally, we extend our appreciation to Dr. Susanne Buiter for her comments and editorial guidance throughout the review process."

With best wishes,
Susanne Buiter