Response to Reviewer 1: Our comments are provided in blue. Text modifications are provided in green.

This paper introduces a numerical method to jointly simulate long-term and short-term evolution of faults, including dynamic rupture and fault localization and growth, in elasto-plastic media with viscous regularization. Such models have emerged in recent years to tackle important questions at the interface between earthquake research and long-term crustal deformation research. This work is an interesting contribution to those efforts. The results illustrate how models with ideal plasticity (constant friction) can generate earthquakes, despite the absence of explicit weakening of fault friction.

Thank you for your thorough and constructive feedback on our manuscript. We sincerely appreciate your recognition of the significance of our work. In our revisions, we will refine the descriptions and improve clarity to ensure that key concepts, methodologies, and results are effectively communicated. We are confident that these improvements will strengthen the manuscript and align it with the high standards of SE.

Thank you again for your valuable insights.

Sincerely,
Yury Alkhimenkov, Lyudmila Khakimova and Yury Podladchikov

My main suggestions are

1. Parts of the text and results (e.g. line 5, “Finer temporal discretization leads to sharper stress drops …”) give the impression that the simulations have not reached numerical convergence yet. If that is the case, I think you should keep refining the space and time discretization until the results converge (i.e. until there is negligible changes upon further refinement) and discuss only converged results. A focus on converged results can have a substantial impact on the statistics of stress drops and other physical quantities. If this requires new and more expensive simulations, it qualifies as major revision.

We acknowledge the importance of discussing converged results in the traditional sense. However, elasto-plasticity is a highly nonlinear problem—analogous to turbulence in the Navier-Stokes equations—where numerical convergence, as typically defined, is only achievable under specific conditions.

In our study, we show **trends**. For example, at low spatial and temporal resolution, the simulations do not show any stress drops. However, simulations with high spatial and temporal resolutions exhibit similar stress drops (both in number and amplitude). We agree that in the spatial convergence test, full convergence may not be reached; therefore, we have revised the wording and removed "convergence" from that section. Nevertheless, we clearly demonstrate the trend, which is the primary goal of this first study in this direction.

We deliberately present results that are not fully converged to illustrate trends, as no prior work, to our knowledge, has performed elasto-plastic simulations with sufficient temporal and spatial resolution to resolve stress drops. Without such a comparison, the existence of stress drops under a static friction coefficient might be questioned. As far as we know, this study is the first to resolve numerical stress drops under these conditions.

A study achieving spatial convergence will require significantly more computational power and presents a challenge that warrants a separate investigation (separate article).

Regarding the statistics of stress drops, our main analysis is based on high-quality simulations with sufficiently fine temporal and spatial resolution. We also include a lower-resolution simulation (which is of poor quality and clearly did not converge) to demonstrate the impact of temporal resolution on the results.

Modifications in the text: We removed the word “convergence” and replaced it with “trends,” along with some minor rephrasing for better flow.

2. But I wonder if this lack of convergence is only apparent. With each refinement, are you also changing the value of the artificial viscosity (regularization parameter)? If that is the case, maybe you should instead keep the viscosity fixed in convergence studies. Unless there is a good reason to scale the viscosity to the mesh size, but that should be explained in the paper and it should be done in a way that guarantees convergence.

This is an excellent point. Indeed, with each refinement, we adjust the artificial viscosity (regularization parameter) proportionally to maintain a constant shear band thickness across different spatial and temporal resolutions. This rescaling is crucial for preserving the physical consistency of the localization process. The rationale for this approach is explained in detail in our previous study:

Y. Alkhimenkov, L. Khakimova, I. Utkin, Y. Podladchikov (202X), Resolving strain localization in frictional and time-dependent plasticity: Two- and three-dimensional numerical modeling study using graphical processing units (GPUs), Journal of Geophysical Research: Solid Earth.

We will clarify this point in the revised manuscript to ensure transparency in our methodology. We will remove the word "convergence" and replace it with "trend" in the manuscript.

Note that we re-scale the viscosity damper proportionally to the resolution in each simulation to maintain the physical thickness of the shear bands \cite{alkhimenkov2024resolving}.

3. I found it very interesting that a bulk plasticity model with constant friction can generate earthquakes, because this is in contrast to fault friction models that are common in the computational earthquake dynamics community (earthquakes on pre-existing faults cannot be simulated without frictional weakening). This is not new, though, and it would be great to make more connections to existing related theoretical results. In particular, I find

the work by Le Pourhiet (2013 https://doi.org/10.2113/gssgfbull.184.4-5.357) contains very insightful explanations of “structural weakening” in plastic models and plenty of useful references.

We thank the reviewer for suggesting this valuable reference, which we have now cited, along with additional relevant studies. We acknowledge that, from a theoretical perspective, structural softening in plasticity models with a static friction coefficient has been analyzed in previous works. However, to our knowledge, no computational studies have systematically examined stress drops in detail—particularly the occurrence of multiple stress drops, their statistical properties, and their dependence on numerical parameters. Our study represents one of the first efforts in this direction.

From a theoretical perspective, such stress drops were predicted and analyzed by, e.g., \cite{vermeer1990orientation} and \cite{le2013strain}.

Minor comments:

Line 34, “Recent studies ….”: You can also cite old seminal studies by Joe Andrews. In the 1976 paper (https://doi.org/10.1029/JB081i020p03575) where he basically opened the era of computational earthquake dynamics by introducing slip-weakening rupture simulations, he also realized that friction models were insufficient and introduced simulations with plasticity in the bulk. He was clearly very far ahead of his time. Renewal of this topic had to wait his 2005 paper (https://doi.org/10.1029/2004JB003191), which motivated the papers in computational earthquake dynamics that you cite. There is also important literature on plastic fracture dynamics in the fracture mechanics community; you can find many cited in Ben Freund’s book and in Gabriel et al (2013).

We thank the reviewer for highlighting these important references, which we have now incorporated into the manuscript. While many studies employ non-constant friction laws, often supplemented with bulk plasticity, our approach is fundamentally different. Our model relies solely on ideal plasticity with a static friction coefficient—without any additional weakening mechanisms. We demonstrate that this simple mechanical framework is sufficient to generate earthquakes, but only if the simulations are conducted at sufficiently high temporal and spatial resolution. Simply adding plasticity is not enough; resolving plastic deformation and capturing stress drops with adequate numerical precision is crucial.

One of the first computational earthquake dynamics models with slip-weakening rupture simulations was introduced by \cite{andrews1976rupture}. Recent studies have suggested that plasticity plays a crucial role in the nucleation of earthquakes, particularly through off-fault plasticity mechanisms (e.g., \cite{andrews2005rupture}).

Line 42: You can cite the earliest 3D studies of dynamic rupture with plasticity, e.g. Ma (2008 https://doi.org/10.1029/2008GC002231), Ma and Andrews (2010, https://doi.org/10.1029/2009JB006382)

We thank the reviewer for suggesting these valuable references, which we have now incorporated into the manuscript.

\cite{ma2008physical, ma2010inelastic} conducted some of the earliest studies on dynamic rupture with plasticity.

Line 59: It would be useful to emphasize in this sentence that the friction coefficient is assumed constant (no softening/hardening, “ideal plasticity”).

Yes, we have now explicitly emphasized in the manuscript that the friction coefficient is constant, with no softening or hardening, corresponding to an ideal plasticity framework.

The friction coefficient is assumed to be constant in all simulations, with no hardening or softening, which corresponds to an ideal plasticity model.

Line 69: the word “static” can be removed (one could misinterpret the sentence as implying that there is a dynamic coefficient and it’s not constant).

As per the reviewer's suggestion, we have removed the word "static" to avoid potential misinterpretation.

We utilize the simplest pressure-sensitive ideal plasticity model with constant in time and space friction coefficient.

Line 114: should equation 13 involve the elastic strain instead of the total strain? Are you assuming plasticity also during dynamic stages of the simulation? If not, this assumption needs to be justified.

We fully agree with the reviewer and have revised this equation. Yes, there should be only elastic strain.

\begin{equation}\label{eq1}
\frac{\partial \sigma_{ij}}{\partial t} = C_{ijkl}^e \, ({\dot{\varepsilon}}_{kl} - {\dot{\varepsilon}}^{pl}_{kl}),
\end{equation}

Line 162, “This re-scaling process is iterated over "pseudo-time" …”: explain this in more detail. Make sure the description of the methods is complete enough to guarantee reproducibility.

We have improved the explanation in the manuscript to provide more detail and ensure clarity. Additionally, we have included relevant references where this methodology is further explained.

To achieve this, the equations are written in their residual form and iterated over "pseudo-time" until convergence is reached.

Line 165: show also the continuum equations describing the modified rheology assumed, so that readers don't have to go look for it in previous papers.

We have added the relevant continuum equations describing the modified rheology in the revised manuscript to ensure clarity and self-containment.

Line 166: explain the rationale to set the viscosity value.

This is an important point. We have now included an explanation of the rationale behind selecting the viscosity value in the revised manuscript to clarify its role in the numerical framework.

The numerical viscosity is usually set to a small value. If this value is too high, the shear bands become very thick; conversely, if the value is too small, the thickness of the shear band is just one pixel. The correct value of the viscosity damper lies between these limits. In the following section, we examine how the choice of viscosity damper affects the solution.

Line 196, "integrated stress": define this quantity (integrated in space? in time? over what domain?)

We have now explicitly defined this quantity in the revised manuscript.

The integrated stress σ_{xx}^INT is computed over a vertical line segment using the following expression:
\begin{equation}\label{sigmaINT}
\sigma_{xx}^{\text{INT}} = \frac{1}{L_y} \int_{0}^{L_y} (-p + \tau_{xx}) \, dy.

\end{equation}

Line 237, "fault gouge ... fault plane": Which gouge? Which fault? These objects are not explicitly introduced in the model, I think you just mean "shear band" or "plastic zone" here.

We agree with the reviewer and have revised the text to use the appropriate terminology, replacing "fault gouge" and "fault plane" with "shear band" or "plastic zone" as appropriate.

The reduced apparent coefficient of friction is a consequence of plastic flow in the plastic zone, which allows slip to occur more easily along the shear band, despite the actual slip occurring along the Coulomb shear planes.

Section 4.5: Do you change the viscosity when you change N? Clarify.

Yes, the viscosity is adjusted when changing N, as explained in our previous manuscript (*Alkhimenkov et al 2014*, JGR: *Solid Earth*). However, we have now clarified this point in the revised manuscript to ensure transparency.

Note that we re-scale the viscosity damper proportionally to the resolution in each simulation to maintain the physical thickness of the shear bands \cite{alkhimenkov2024resolving}.

Line 254-255, “simulations with sufficient resolution produce stress drops and their amplitudes are similar”: The shapes of the curves are still different. Can you try even larger values of N to show convergence convincingly?

Yes, we agree that the shapes are still different. That’s why we no longer call it a convergence test but instead refer to it as "trends" upon mesh refinement. The word "convergence" has been replaced throughout the manuscript.

\subsection{Trends with increasing spatial resolution}

Line 297, “These results highlight the sensitivity of fault behavior to the dilatation angle”: relate to published results, or instanc Templeton and Rice (2008)

In the revised manuscript, we now relate this finding to published results, including the work of Templeton and Rice (2008).

These results highlight the sensitivity of fault behavior to the dilatation angle, emphasizing the importance of including dilatancy effects in models of fault mechanics and earthquake nucleation, as also suggested by \cite{templeton2008off}.

Line 308: Figure 14 seems to show lack of convergence. Clarify.

In Figure 14, we show two curves: low temporal resolution (red) and high temporal resolution (blue). The main idea is to illustrate the trend that with higher temporal resolution, we observe significantly more stress drops. We do not analyze convergence in this result.

Line 310, “simulation with fine temporal resolution and the lowest regularization”: This suggests that you are changing systematically the viscosity when you refine the simulations. Please clarify, explain that in detail.

This comment overlaps with our previous discussion on the rationale for selecting the viscosity value. In the revised manuscript, we have clarified that viscosity is systematically adjusted when refining simulations and provided a detailed explanation to ensure transparency.

Note that we re-scale the viscosity damper proportionally to the resolution in each simulation to maintain the physical thickness of the shear bands \cite{alkhimenkov2024resolving}.

Line 316, “dynamic rupture events, akin to the rapid stress release observed during seismic slip”: Are these events as fast as earthquakes (slip rate of m/s, rupture speeds of few km/s)? Is the inertial term important during these events?

This is an excellent comment, and we thank the reviewer for pointing it out. We can confirm that stress drop and stress release occur very rapidly, but we have not analyzed the slip rate and rupture speeds. We believe that such an analysis warrants a separate study and a dedicated publication.

We can definitely say that once a stress drop occurs, wave propagation begins, which is only possible due to inertia terms. However, we believe the reviewer raised another important point: “Do inertia terms play a role and affect stress drops?” In this simple model, inertia terms do not affect stress drops because stress drops correspond to a quasi-static solution.

However, we admit that in this first study, we only indicate this similarity and do not provide a detailed analysis of slip rate and rupture speeds in the present model.

Line 360, “characterized by a sharp peak followed by a gradual decay”: I see instead a broad peak and two long tails on both sides.

We fully agree with the reviewer that our initial description was inaccurate. We have revised the text to more accurately describe the observed behavior.

The distribution of stress drop amplitudes is notably non-Gaussian, characterized by a broad peak with long tails on both sides, indicating that while small stress drops are more common, larger stress drops still occur with significant probability.

Line 357, “This insight aligns with the Gutenberg-Richter law”: but here the distribution is truncated at low values too. Show a log-log plot to check if the upper tail is really a power law analogous to the G-R law.

We thank the reviewer for this comment. Reflecting on this (and the second reviewer's comment), we have significantly revised our interpretation because stress drop magnitude may not be strongly related to seismic event magnitude according to the G-R law. Instead, we now only indicate that our results resemble the G-R law but require further analysis. We believe a detailed analysis of a different histogram is needed—specifically for seismic event amplitudes (rather than stress drop amplitudes). In such an analysis, we will follow the reviewer's suggestion and include a log-log plot in a future study.

This insight suggests a resemblance to the Gutenberg-Richter-like law, which describes the frequency-magnitude distribution of earthquakes; however, a more detailed analysis is required to establish a direct connection, particularly from a plastic deformation perspective.

Section 5.1, “the nature of stress drops”: relate your results to insights from existing theory, e.g. Le Pourhiet (2013 https://doi.org/10.2113/gssgfbull.184.4-5.357)

We relate our results to one of the earliest studies, Vermeer (1990). In the revised manuscript, we have also incorporated additional relevant studies, including Le Pourhiet (2013), as suggested by the reviewer.

From a theoretical perspective, such stress drops were predicted and analyzed by, e.g., \cite{vermeer1990orientation} and \cite{le2013strain}.

Lines 395+, “3D simulations … with zero regularization …. convergence tests performed in 3D”: are you suggesting that simulations converge even without regularization? If so, why is regularization needed? Clarify.

Regularization is necessary to control the physical thickness of shear bands. In our 3D simulations, we observed a form of “trend” (we removed the word “convergence”) in which the general shear band patterns remained similar across different resolutions. However, we did not analyze numerical convergence in the traditional sense. We have clarified this point in the revised manuscript.

These simulations provide valuable insights into how strain localization and stress drops manifest in fully 3D domains. The tests performed in 3D, both in temporal and spatial resolutions, show similar trends with the results of the 2D simulations presented in this study.

Line 409, “closely mirrors the earthquake cycle seen in nature”: Do you find multiple stress drop happening on the same “fault” (shear band) or do they occur each time on a different segment of the fault?

This is an excellent point, and we thank the reviewer for highlighting it. Indeed, as the simulation progresses, multiple shear bands develop, and stress drops can occur repeatedly on the same shear band. We have now incorporated this important clarification into the revised manuscript.

The periodic nature of stress drops, interspersed with slower periods of strain accumulation, closely mirrors the earthquake-like cycle seen in nature. As the simulation progresses, multiple shear bands develop, and stress drops can occur repeatedly on the same shear band, rather than always initiating on new segments. This behavior closely resembles natural faulting processes, where strain localization leads to repeated cycles of stress accumulation and release along pre-existing fault structures.

Section 5.6: there is redundancy with previous sections, which could be avoided.

We removed this section.

Line 469, “that plasticity should be considered alongside traditional frictional models in future earthquake simulations”: This is already the case in published work, e.g. Erickson et al (2017 https://doi.org/10.1016/j.jmps.2017.08.002), Preuss et a (2020 https://se.copernicus.org/articles/11/1333/2020/), Simpson (2023 https://doi.org/10.1016/j.tecto.2023.230089). Rephrase and add references.

We have rephrased this statement in the manuscript and incorporated the suggested references to accurately reflect existing work in the field.

Second, the results confirm previous studies highlighting the important role of plastic deformation in fault weakening and rupture, suggesting that plasticity should be considered alongside traditional frictional models in future earthquake simulations.

Other studies highlighting the importance of plasticity in earthquake physics modeling include \cite{erickson2017finite}, \cite{preuss2020characteristics}, and \cite{simpson2023emergence}.

We would like to thank the reviewer again for valuable comments, which helped us improve the quality of the manuscript.

Sincerely,
Yury Alkhimenkov, Lyudmila Khakimova and Yury Podladchikov