

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} \epsilon_{,k} - \dot{\epsilon}_{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:

$$\sigma_{xx}^{\text{INT}} = \frac{1}{L_y} \int_0^{L_y} (-p + \tau_{xx}) \, dy.$$

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 (Alkhimenkov 2024 resolving).

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 al (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

Response to Reviewer 2: Our comments are provided in blue. Text modifications are provided in green.

The study provides a physics-based explanation for stress drops and earthquake nucleation using a simple elasto-plastic model, avoiding the need for complex frictional laws. The findings emphasize the importance of plastic deformation in fault slip mechanics, an aspect often overlooked in traditional models. The paper employs high-resolution 2D numerical simulations, carefully analyzing temporal and spatial resolution effects on stress evolution and earthquake nucleation. This methodology using GPU-based parallelization has great potential in achieving high-resolution earthquake modeling. I believe both the conclusions and the novel methodology used in this study should be promptly communicated to specialists and general audience.

We sincerely appreciate your thoughtful and constructive feedback on our manuscript. Your recognition of the significance of our findings and the potential impact of our GPU-based methodology is greatly valued.

In our revisions, we will focus on enhancing clarity and precision in our descriptions to ensure that the key concepts, methodologies, and results are effectively conveyed to both specialists and a broader audience. We are confident that these refinements will further strengthen the manuscript and improve its readability.

Thank you again for your valuable insights and for recognizing the relevance of our work.

Sincerely,
Yury Alkhimenkov, Lyudmila Khakimova and Yury Podladchikov

However, I found the writing (both texts and figures) quality can still be improved.

First, the text is in many locations verbose and repetitive, often explaining the same concept multiple times in slightly different ways. For example, the sentence line 275-279 is directly repeated in the next paragraph line 278-280. The phrase "finer temporal/spatial resolution leads to sharper stress drops" appears in sections 4.1, 4.2, 4.4, 4.5, and 4.9.1. The discussion (section 5) reiterates results (e.g., resolution impact, regularization, plasticity) rather than synthesizing new insights. The same applies to figures. For example, Fig. 16a is essentially the same as Fig. 14a, and is actually not described or mentioned in the text. Fig. 16b-d are largely overlapping. With only three sentences (line 343-348) describing this figure, I suggest the panels to be merged.

Second, the writing lacks conciseness. Many sections could be rewritten in a more direct and streamlined manner. Streamline results by grouping related findings (e.g., combine resolution tests 4.4-4.6 into one subsection with subheadings for temporal/spatial/regularization effects). Multiple figures show stress drop evolution

with slightly different grid resolutions, but the key insights do not change significantly. Additionally, the captions are overly descriptive, without highlighting the key takeaways (Figs. 2-17). Combine redundant figures to highlight the key messages. Use subplots to highlight contrasts (e.g., low vs. high resolution in Figs. 2-3) rather than separate figures.

I will add more suggestions on conciseness in line-by-line comments.

We agree with the reviewer that the quality of the text can be improved and made more concise. We have addressed the specific suggestions provided by the reviewer and revised the manuscript accordingly to reduce redundancy, streamline descriptions, and improve clarity. Additionally, we have reorganized sections where appropriate and adjusted figures to better highlight key insights.

Major comments:

1. Through section 4.4-4.6 and Figs. 5-8 the authors claimed that detailed convergence tests have been conducted. However, the convergence is not visually observable from the figures. I wonder if the authors have metrics to quantify the convergence and if the converging rate matches the analytical expectation. In addition, it is suggested in section 4.6 that the results are sensitive to the regularization parameter η_{vp} . If η_{vp} needs to be adjusted per model resolution, I wonder if this then still indicates the existence of model convergence. And if so, whether the authors could quantify the result sensitivity on η_{vp} .

We thank the reviewer for this insightful comment. Plasticity is a highly nonlinear problem— analogous to turbulence in fluid dynamics—where traditional numerical convergence may not always be observed. Instead, what we assess is the statistical similarity of shear band patterns across different resolutions.

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 sections. Nevertheless, we clearly demonstrate the trend, which is the primary goal of this first study in this direction.

Regarding the regularization parameter η_{vp} , it must be adjusted to maintain a consistent shear band thickness across different model resolutions. We have now clarified this in the revised manuscript. Additionally, we note that a more detailed discussion on this topic is available in our previous publication in JGR: Solid Earth (Alkhimenkov et al, 2014).

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

2. The definition of “nucleation” is not clear. As a major topic as appeared on the tile, the term should be strictly defined. The interseismic loading and stress drop have been described, but it is not clear how nucleation makes the transition in between. Given that the time marching scheme in this study is implemented via strain increment, I wonder if more insights on the temporal behaviors of these processes (interseismic, nucleation, stress drop) can be added.

This is an excellent question, and we fully agree that the term *nucleation* should be explicitly defined in the manuscript. We think that in this early study we do not study nucleation process but only earthquake triggering. Therefore, we removed the word nucleation from the manuscript.

Triggering

Stress drop manifest the jump between the two quasi-static solutions and perhaps indicate the absence of a static transition between the stress state at the end interseismic period and onset of a new interseismic period.

This transition can be modeled with simplified linear elasto-dynamics by setting the difference in strain just before and after a stress drop as an initial condition for wave propagation.

3. Throughout the paper the simulated stress drop has been linked to “earthquake magnitude”. However, it is known from both numerical models and seismic observations that stress drop and magnitude are not (strongly) related. I found this extrapolation from varied stress drop to varied magnitude, and hence the stated link to the G-R law, too speculative. The authors also need to be careful when linking stress drop to seismic events.

We understand the reviewer’s concern and have revised the manuscript to adjust the wording accordingly. We acknowledge that stress drop and earthquake magnitude are not necessarily strongly correlated, as supported by both numerical models and seismic observations. In our revisions, we have taken care to avoid overgeneralizing this relationship and have refined our discussion regarding links to the Gutenberg-Richter law. We do not link stress drops to earthquake magnitudes in the revised manuscript. Regarding the distribution of stress drops, we now only describe it as G-R-like behavior.

Additionally, we emphasize that our study presents one of the first detailed numerical models demonstrating earthquake generation with a constant friction coefficient. This fundamentally differentiates our approach from previous models and highlights the novel aspects of our findings.

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.

The periodic nature of stress drops, interspersed with slower periods of strain accumulation, mirrors the earthquake-like cycle seen in nature.

The paper needs a consistent coordinate system. Although the authors prefer a dimensionless computational system, it is not properly introduced in section 3.4. With $x, y \in [0, L_x] \times [0, L_y]$ stated at the beginning of section 3.5, the readers might be confused if the followed equations 22-27 were expressed in dimensionless coordinates or not. To add more confusion, the authors used no axis labels (Figs. 1), “Grid Cells (-)” (Figs. 2-3, 9-11), “x(-), y(-)” (Figs. 5-8, 12-13, 17), “x, y” (Figs. 15) in different figures, and in many cases without ticks. I suggest the authors to unify the expression and add ticks to the axis for better reference and comparison.

We understand the reviewer’s concern and have improved the description of the initial model setup to ensure clarity in the coordinate system and its dimensional consistency. We have also revised the notation in the manuscript to make it more consistent.

Regarding the figures, we carefully evaluated different formatting options and found that our current choices balance readability and necessary detail without overloading the visuals. While full unification of all figures would make some of them too dense and difficult to interpret, we have made adjustments where possible to improve clarity and consistency, including adding axis labels and ticks where appropriate.

subsection{Model configuration, boundary conditions, and non-dimensionalization}

The computational domain is a square with dimensions $(x, y \in [0, L_x] \times [0, L_y])$. All simulations in this study are performed using a simple initial model configuration and non-dimensional equations. To ensure a consistent dimensionless framework, we define the following characteristic scales: length $(l^* = L_x)$, time $(t^* = 1/a)$, and pressure $(p^* = G_0)$. Here, (L_x) represents the domain size in the (x) -direction, and (a) denotes the background strain rate. Deformation evolves over timescales inversely proportional to the initial background strain rate (a_0) at $(t = 0)$. The ratio of cohesion (c) to the pressure scale (p^*) is defined as $r = \frac{c_0}{G_0}$.

For all computations, we set the coefficient of internal friction to $(\mu = 0.6)$. Pure shear boundary conditions are applied by prescribing velocities at all boundaries in the dimensionless framework:

Minor comments:

1. Line 11: the usage of “decay” not accurate. Decay usually refers to a temporal process. It is not clear the decay is with what.

We agree with the reviewer and have revised the text to use a more precise term that accurately describes the intended meaning.

The histogram of stress drop amplitudes shows a non-Gaussian distribution, characterized by a broad peak with long tails on both sides,

2. Line 13: across which “scale” (temporal or spatial)?

We have clarified in the text whether the scale refers to temporal or spatial.

This "solid turbulence" behavior suggests that stress is redistributed across spatial and temporal scales, with implications for understanding the variability ...

3. Line 14: I doubt if the link to “magnitude” is proper here, see major comment 3.

We have revised the terminology and adjusted the wording to ensure accuracy and avoid misleading implications.

with implications for understanding the variability of stress drop magnitudes.

4. Line 49: I don't see the necessity of introducing heterogeneity here.
Heterogeneity was only mentioned in the discussion of model limitation once. I suggest the whole paragraph can be eliminated.

We agree with the reviewer and have removed the paragraph for consistency

5. Line 90: D/Dt should be defined here already.

We have now modified this in the text.

where the Jaumann rate of Cauchy stress, represented as $\frac{D}{Dt} \sigma_{ij}$, is provided in the following section and the deviatoric plastic strain rate is

6. Line 177: is the model a “square” ($L_x=L_y$)? I can find nowhere the values of L_x and L_y .

We have added the values in the text to clarify the model dimensions.

The computational domain is a **square** with dimensions $(x, y \in [0, L_x] \times [0, L_y])$.

Here, (L_x) represents the domain size in the (x) -direction, $(L_x=L_y=1)$ and (a) denotes the background strain rate.

7. Line 193: connecting to comment 6, what does 0.2 refers to? Is the equation dimensionless or not?

We have clarified in the text whether the equation is dimensionless and provided a proper explanation of the value 0.2.

The expression is the following in the dimensionless framework:

8. Fig 1: axes should be marked and labeled.

We have added axis labels and markings for clarity.

We introduce a circular inclusion in the non-dimensional pressure $\backslash(p \backslash)$, representing a localized increase with the highest value at the center of the model. The expression in the dimensionless framework is as follows:

9. Fig 2: it is not clear where the “three different stages” were visualized. Could you mark them in panel a?

We have now marked the three stages in panel (a) for better visualization.

10. Line 242: which red circle?

We have improved the description to clearly indicate that the first red circle is being referenced.

where ϵ_{t_1} corresponds to the total strain just before the stress drop (first red circle)

11. Section 4.2.1: Mohr’s circle analysis is not really where your novelty locates. The section can be condensed.

This is one of the first numerical studies where stress drops are explicitly analyzed, with a detailed explanation using Mohr’s circles. Reflecting the reviewer's comment, we have made the section more concise.

12. Sections 4.4-4.6: consider to merge

We have merged these sections to subsections for better organization and readability.

$\backslash$ subsection{Trends}

13. Line 261: I wonder if figure 9 shows a convergence. The results are not visually identical as claimed. Have you tried larger N? I also find the whole subsection a bit speculative. Terms like “too high” requires quantification

We have revised the text to be more concise and precise. Additionally, we acknowledge the reviewer’s concern and have clarified the interpretation of convergence while improving the wording of subjective terms. We removed word “convergence”.

Due to high regularization, the results are nearly identical and the thickness of the shear bands is the same in all panels. However, due to over-regularization, the stress drop is not visible.

14. Lines 278-280: repeating lines 275-277, should be eliminated.

We have removed these redundant lines to avoid repetition.

15. Fig 11: panel b is identical to Fig 10b. A different visualization is not needed.

We agree and have removed the redundant panel.

16. Section 4.8: this section is not pure results. Sentences like the first paragraph and lines 298-299 are either introduction or discussion materials and should have come earlier or later. I also wonder why the deformation mechanism in metallic material is mentioned in the end (line 304-306), which is largely off topic.

We removed lines 298–299. We believe that referencing deformation mechanisms in metallic materials is important to highlight the differences with rocks.

17. Fig 12: after “ $\psi = 5^\circ$ ”, the symbol of degree should be added, same below.

We agree and have added the degree symbol.

18. Line 323: is the set of notion (1-3) the same as (t1-t3) in section 4.3? Better keep consistent.

We agree with the reviewer and corrected the explanation.

Figure \ref{Seq_evolzoom_UxOK}a shows the displacement increment $\Delta u_x = u_x(t_3) - u_x(t_2)$, where $u_x(t_2)$ and $u_x(t_3)$ represent the displacement fields at the beginning and end of the interseismic period, respectively (the period between two high-amplitude stress drops. Similarly, the displacement increments $\Delta u_x = u_x(t_2) - u_x(t_1)$ during major stress drops are shown in Figures \ref{Seq_evolzoom_UxOK}c-d.

19. Line 330: “leading up to” implies there is a causal relationship between the “aseismic slip accumulation” in interseismic and the “stress drop event”. This is not proved by the authors. Moreover, the general recognition would link any aseismic slip deficit in the shear band to the following earthquake. More elaboration is needed here.

We agree and have revised the text to avoid implying causality without proper evidence. We have also elaborated on the relationship between aseismic slip and stress drops to ensure clarity.

Our simulation results also demonstrate the material’s behavior during interseismic periods, where displacement gradually accumulates without significant stress drops. As shown in Figure \ref{Seq_evolzoom_UxOK}, displacement increments Δu_x during interseismic periods increase progressively as loading continues. This behavior mirrors the slow, aseismic slip observed between seismic events in fault zones. The gradual buildup of displacement during interseismic periods reflects the loading of the fault system, while the rapid displacement during stress drops corresponds to seismic slip. However, our model does not explicitly establish a causal link between aseismic slip accumulation and subsequent stress drops, which requires further investigation.

20. Line 343: add reference to “solid turbulence”.

We have modified the text.

The histogram of stress drop amplitudes shown in Figure~\ref{Hist} provides a

quantitative representation of the frequency and magnitude of stress drops occurring during the simulations. 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. This distribution resembles turbulence-like spectrum, where a few large events (bursts) coexist with numerous smaller fluctuations. For solid systems, this phenomenon was first analyzed by \cite{poliakov1994fractal}, who explored the multi-fractal characteristics of shear bands in elasto-plastic media, that makes it similar to the fluid turbulence.

21. Fig 16: panel a is the same as Fig 14a, consider to remove. Panels b-d should be merged.

Figure 16, panel (d), contains four curves, all of which are different, whereas Figure 14a has only two curves. We believe that the merged panel (panel d) is more intuitive while still preserving the end-member cases that are well-separated—panels (b) and (c).

22. Line 356-357: the discussion on “magnitude” and the link to G-R law is not necessarily true. See major comment 3. Have you tried to calculate the magnitude of the events and check if it fits G-R law? It might be difficult because the simulations were 2D.

We agree with the reviewer’s concern. We have revised the discussion to avoid overgeneralizing the relationship between stress drops and earthquake magnitude, acknowledging the limitations of 2D simulations in directly comparing to the G-R law. The calculation of seismic events magnitudes is a topic for 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.

23. Line 364: how is your “nucleation” defined? Does it align with others’ such as Rubin & Ampuero 2005? I also find it unfair that this term comes too late and is not extensively described in the results. It is one of the key features in your title and should be well addressed.

We have removed “nucleation” from the article.

24. Line 376: how is your “seismic event” defined? Do you see fast (m/s) slips in your model? Does your definition align with others’?

Seismic event corresponds to the wave propagation which is explained in the text section 4.3. In our model the transition from one quasi-static solution to another occurs as a “jump” (very fast, one or several loading increments) but we do not analyzed how fast they are in dimensional quantities in this early study.

Stress drop manifest the jump between the two quasi-static solutions and perhaps

indicate the absence of a static transition between the stress state at the end of an interseismic period and onset of a new interseismic period.

This transition can be modeled with simplified linear elasto-dynamics by setting the difference in strain just before and after a stress drop as an initial condition for wave propagation.

25. Section 5: reiteration of results should be removed so that the whole section can be streamlined.

We have condensed Section 5 to reflect the reviewer's comment.

26. Section 5.2: can you comment more on how the regularized parameter η_{vp} influence the results? Do you have a quantification for this? Also see major comment 1.

We have expanded the discussion of how η_{vp} affects the results and, where possible, provided a more quantitative assessment of its impact.

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.

Note that we re-scale the viscosity damper proportionally to the resolution in each simulation to maintain the physical thickness of the shear bands (Alkhimenkov et al., 2024).

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.

27. Line 396: you claimed that the 3D results you published earlier showed good agreement with the results in this paper. Could you elaborate more? I would expect clear differences between 2D and 3D simulations. Many numerical studies show that the third dimension has impacts on nucleation and rupture that are not negligible.

We did not analyze nucleation in 3D but rather observed that the general pattern of shear band patterns and the presence of stress drops occur in both 2D and 3D. However, we acknowledge that the magnitude of stress drops and rupture dynamics may differ in 3D. A proper analysis of these differences is an important topic for future research. 3D simulations are very computationally expensive and along deserve a separate study.

28. Sections 5.4-5.5: these comparisons to nature and previous models are useful. I wonder if you can further comment on how the weakening process occurred in your model differentiate itself from that in rate-and-state friction. Do they predict similar features such as some slip-weakening distance? Such insights would be inspiring.

We agree that relating our results to rate-and-state friction models would be valuable. However, to make a precise comparison, a more detailed analysis is required, which may be best addressed in a separate study. This comparison deserves a separate manuscript.

29. Section 5.6: largely repeating section 4.2.1, can be removed.

We removed this section.

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