

Response to Reviewer 1

We thank Reviewer 1 for the detailed technical review. We accept the recommendation for major revision and have addressed each point through additional numerical, analytical, and sensitivity analyses. The additional analyses make explicit the physical quantities used throughout the study: amplification is evaluated as an increase in near-critical modal and nonlinear growth tendency, localization is assessed relative to the intrinsic nucleation stability length and by equal-stage spatial profiles, and scaling is expressed through the input-only ratio $\chi = t_c V_s / D_c$. We performed independent time-step and grid convergence tests, linear-stability analysis, reduced-layer-thickness sensitivity tests, perturbation-size tests, equal-stage comparisons, and alternative temporal-response tests. The corresponding changes in text are highlighted in yellow.

Reviewer comment: This manuscript examines whether a finite delay in state evolution can change the way earthquake nucleation develops in space and time. Moving from the authors previous block-based models to a spatially distributed fault model is potentially useful. But, I do not think the current results are sufficient to support the main claims about a new constitutive mechanism, a distinct high-velocity nucleation regime, or a general scaling relationship. At present, the study mainly shows that an imposed delay amplifies a strongly forced response in one particular numerical configuration. According to me, several important physical and mathematical issues therefore need to be addressed before the novelty can be properly evaluated. I recommend major revisions.

Response: We accept this suggestion. To make the physical basis of the study explicit, we performed the requested convergence, stability, perturbation-scale, and equal-stage analyses. These analyses distinguish three complementary quantities. Amplification is quantified as the delay-induced increase in near-critical modal and nonlinear growth tendency, rather than by the crossing of a single velocity threshold. Localization is evaluated relative to the intrinsic critical wavelength and by profiles compared at matched center velocity and matched accumulated slip. Scaling is expressed through the prescribed-state ratio $\chi = t_c V_s / D_c$. Temporal refinement also shows that finite-window peak velocity is not the sensitive diagnostic in the subcritical 2 km Master case: converged $V_{\max}$ values for $t_c = 0$ and 0.12 s differ by only about -0.01 to -0.02% . The main evidence is therefore based on characteristic roots, the critical wavelength L_c , post-transient logarithmic growth rates, perturbation-size sensitivity, and equal-stage spatial structure.

Reviewer comment: The Meaning of the proposed delay needs to be explained much more clearly. In Equation 4, both velocity and state are taken from an earlier time, while healing remains instantaneous and the current state is still used to calculate friction. This is one possible mathematical choice, but it is not clear why it is the appropriate description of contact renewal or gouge reorganization.... Classical rate-and-state friction already includes gradual state adjustment through the characteristic slip distance. I think the authors therefore need to explain what additional physical process the new response time represents, why it should be a fixed time rather than a slip-dependent quantity, and whether other reasonable descriptions of delayed adjustment would produce the same behaviour. Without this foundation, the delay currently appears more phenomenological than constitutive.

Response: We accept this suggestion. We have expanded the physical interpretation of t_c and tested several alternative response architectures. The parameter t_c is treated as a phenomenological coarse-grained response time for an unresolved internal reorganization process whose adjustment is controlled primarily by elapsed time; it is distinct from the slip-controlled renewal scale D_c and is not presented as a uniquely calibrated contact-scale constant. Equation (4) is written as retarded ageing-law feedback $g[V(t-t_c), \theta(t-t_c)]$; because the healing contribution is the constant 1, delaying the arguments of g is mathematically equivalent to delaying the complete right-hand side. A causal exponential time relaxation, $t_c \, dm/dt = V\theta - m$, produces essentially the same linear stability shift as the discrete delay: at $t_c = 0.12$ s, L_c changes from 2.300323 km to 2.293780 km, compared with 2.293771 km for Eq. (4), and the nonlinear near-critical turnaround differs by only about 2 ms. A matched slip-

controlled relaxation behaves differently and remains close to the classical trajectory. Velocity-only and state-only decompositions further show that the destabilizing contribution is associated primarily with retarded state adjustment rather than with a generic delay applied to any variable.

Reviewer comment: The spatial model also needs a stronger physical basis. The effective thickness of the model is set equal to the grid spacing, meaning that the physical mass, acceleration, and damping change with the numerical discretization. This is particularly important because the reported core widths correspond to relatively small numbers of grid cells. The authors should either assign the thickness an independently defined physical value or clearly derive the formulation as a discrete block-chain model. There is also an inconsistency between the stated Mode II rupture configuration and the scalar wave equation used in the manuscript. The elastic formulation could be derived carefully, including an explanation of why the shear-wave operator and radiation damping are both required.

Response: We accept this suggestion. We have provided an explicit mechanical derivation and separated the reduced-layer thickness from the numerical grid. Equation (6) is derived as the continuum limit of a reduced 1-D scalar shear layer / nearest-neighbour block chain. For a cell of length Δx , $m = \rho H \Delta x$ and $K_s = \mu H / \Delta x$, giving $u_{tt} = c_s^2 u_{xx} + [\tau_{load} - \tau_f - \eta(V - V_{init})] / (\rho H)$. The formulation is described consistently as a reduced scalar shear-layer model. H is defined independently of Δx , with $H = 20$ m used as the reference reduced-layer thickness and $H = 5$ -160 m explored in sensitivity tests. At fixed H , the $N = 2048$ solution differs from $N = 4096$ by about 0.00036% in V_{max} and about 0.03% in the continuous RMS width. We also state explicitly that $c_s^2 u_{xx}$ represents elastic redistribution along the modeled layer, whereas $\eta(V - V_{init})$ represents local radiation loss to the unresolved surrounding medium.

Reviewer comment: I am not convinced that the high-velocity core represents a new physical regime. The instantaneous model already reaches a velocity close to the selected high-velocity threshold, and the smallest delay moves it only slightly above that threshold. The apparent transition may therefore reflect a smooth increase passing through a chosen diagnostic value rather than a genuine change in fault dynamics. The authors should use a physically based instability measure, such as acceleration rate, energy release, or a change in the stability of the governing system. I recommend that the authors also clarify whether the reported peak velocities are actual maxima or simply the values reached when the simulations end. If the fault is still accelerating at that point, the reported amplification will depend on the selected twelve-second analysis window.

Response: We accept this suggestion. We therefore evaluate the dynamics with physically based instability measures in addition to the descriptive velocity threshold. The 10^{-2} m s $^{-1}$ level is treated only as a diagnostic velocity level, not as a stability boundary. In the 12 s finite-window runs, the reported values are maxima within that analysis window and the trajectories are still accelerating at its endpoint. In the converged 2 km Master calculation, the true peak occurs during the initial direct-effect transient at about 2.5-2.6 ms, with $V_{max} \approx 4.744 \times 10^{-4}$ m s $^{-1}$, after which the center velocity decreases through 16 s. We therefore use $d \ln V / dt$, characteristic-root growth rates, post-transient turnaround time, and a mechanical-energy balance as the primary dynamical diagnostics. The numerical energy balance closes to an RMS relative error of approximately 2.4×10^{-4} . Near the intrinsic stability boundary, these measures show a genuine delay-dependent increase in growth tendency even though the subcritical Master-case V_{max} is nearly unchanged after convergence.

Reviewer comment: The relationship between the imposed perturbation and the inferred nucleation scales needs to be separated more carefully. The simulations begin with a prescribed two-kilometre stress perturbation, and the broader accelerated region remains largely confined within this imposed footprint. This makes it difficult to determine how much of the spatial structure is generated by delayed friction and how much is inherited from the initial forcing. The authors should calculate the expected nucleation length for the adopted friction, α and elastic properties and compare it with the imposed perturbation size. It would also be

useful to compare the delayed and instantaneous cases at equivalent stages of evolution, such as the same central velocity or accumulated slip, rather than only at the same time. the sharper delayed profiles could partly reflect comparisons between different stages of the same accelerating process.

Response: We accept this suggestion. We quantified the relationship between the imposed forcing scale and the intrinsic stability scale and added the requested equal-stage comparisons. For $H = 20$ m and the input-derived reference speed $V_\Delta = 4.7463 \times 10^{-4} \text{ m s}^{-1}$, the instantaneous critical wavelength is $L_c = 2.300323$ km, so the 2 km Master perturbation is subcritical, corresponding to $0.869 L_c$. At $t_c = 0.12$ s, L_c decreases to 2.293771 km. A finite-amplitude perturbation-size sweep gives a grid-converged end-member shift of the 8-16 s crossover diameter of about -4.17 m, consistent in sign and order of magnitude with the -6.55 m linear-stability shift. For a near-critical 2.542 km perturbation, $t_c = 0.12$ s advances the post-transient velocity turnaround by approximately 0.07 s. At equal accumulated center slip, the normalized spatial profiles nearly collapse; at equal center velocity, the differences are sub-percent. We therefore characterize localization through the relation between imposed perturbation size, intrinsic stability length, and matched-stage spatial structure, while common-time sharpening is treated as sensitive to evolutionary phase.

Reviewer comment: The proposed memory-strength scaling is not still independent or predictive because it contains the maximum velocity, which is itself one of the main model outputs. Plotting maximum velocity against a quantity that already includes maximum velocity naturally produces a relationship between the axes. Since the characteristic slip distance and all other model properties remain fixed, the simulations also do not demonstrate a broader scaling law. I suggest the authors to develop a dimensionless measure using only prescribed input quantities and support it with a stability or timescale analysis. More generally, a linear stability analysis of the delayed system would greatly strengthen the manuscript by showing whether the delay changes the growth rate, creates a new unstable mode, or simply amplifies a transient response. This would provide a much firmer technical basis for the novelty than the current threshold-based classification, acc to me.

Response: We accept this suggestion. We use the input-only timescale ratio $\chi = t_c V_s / D_c$ for prescribed steady sliding and $\chi_\Delta = t_c V_\Delta / D_c$ for the Master forcing, where V_Δ is calculated entirely from prescribed stress and friction inputs. The quantity $\Pi_d = t_c V_{\max} / D_c$ is retained only as a post-simulation descriptive coordinate and is not used as an independent predictor. We also performed the requested linear-stability analysis. Linearization gives $[\lambda^2 + A\lambda + c_s^2 k^2][\lambda + \omega \exp(-\lambda t_c)] - B\lambda \omega \exp(-\lambda t_c) = 0$. At V_Δ and $H = 20$ m, $t_c = 0.12$ s shifts L_c by -0.285% . A 2.270 km mode changes from $\text{Re}(\lambda) = -3.14 \times 10^{-5} \text{ s}^{-1}$ at $t_c = 0$ to $+3.19 \times 10^{-5} \text{ s}^{-1}$ at $t_c = 0.12$ s, with the stability crossing near $t_c = 0.0594$ s. Thus finite response time can change the sign of a near-critical growth rate, while long already-unstable modes need not be amplified. The fractional stability-boundary shift grows systematically with χ , providing the requested input-based scaling and timescale interpretation.

Reviewer comment: Section 3.3 refers to Figures 5 and 6 when the discussion appears to concern Figures 3 and 4. Several sources cited in the Supplement, including Bizzarri and colleagues, Guatteri and colleagues, and Wei and Chen, are missing from the main reference list. The displacement integral described as a moment proxy should also be reported with its units and should not be interpreted as seismic moment without a defined rupture area and shear modulus.

Response: We accept this suggestion. We corrected the Section 3.3 figure cross-references and reconciled the Supplement citations with the main reference list, including the cited works by Bizzarri and colleagues, Guatteri and colleagues, and Wei and Chen. We also report $\int u \, dx$ explicitly as an integrated-displacement proxy with units of m^2 . It is not interpreted as seismic moment because the reduced model does not prescribe the physical rupture area and complete geometric information required for that conversion.