Rate and State Friction as a Spatially Regularized Transient Viscous Flow Law

Casper Pranger, Patrick Sanan, Dave May, Laetitia Le Pourhiet, Alice-Agnes Gabriel

Rate and State Friction as a Spatially Regularized Transient Viscous Flow Law

Casper Pranger[1,2], Patrick Sanan[2], Dave A. May[3], Laetitia Le Pourhiet[4], Alice-Agnes Gabriel[1,3]

[1]Ludwig-Maximilians Universität München (LMU), Theresienstrasse 41, 80333 München, Germany.

[2]Institute of Geophysics, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland.

[3]Scripps Institution of Oceanography, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA.

[4]Sorbonne Université, CNRS-INSU, Institut des Sciences de la Terre Paris, ISTeP UMR 7193, F-75005, Paris, France

Key Points:

- We reformulate the empirical rate and state friction law as a bulk viscous flow law in terms of anelastic shear strain rate.
- We show how mesh independence is achieved by including a gradient-like non-local anelastic shear strain rate equivalent.
- We show analytically and numerically that the proposed continuum model closely reproduces existing results of rate and state friction.

Corresponding author: Casper Pranger, casper.pranger@geophysik.uni-muenchen.de

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2021JB023511.

Abstract

The theory of rate and state friction unifies field, laboratory, and theoretical analysis of the evolution of slip on natural faults. While the observational study of earthquakes and aseismic fault slip is hampered by its strong multi-scale character in space and time, numerical simulations are well-positioned to link the laboratory study of grain-scale processes to the scale at which rock masses move. However, challenges remain in accurately representing the complex and permanently evolving sub-surface fault networks that exist in nature. Additionally, the common representation of faults as interfaces may miss important physical aspects governing volumetric fault system behavior. In response, we propose a transient viscous rheology that produces shear bands that closely mimic the rate- and state-dependent sliding behavior of equivalent fault interfaces. Critically, we show that the expected tendency of the continuum rheology for runaway localization and mesh dependence can be halted by including an artificial diffusion-type regularization of anelastic strain rate in the softening law. We demonstrate analytically and numerically using a simplified fault transect that important aspects of the frictional behavior are not significantly affected by the introduced regularization. Any discrepancies with respect to the interfacial description of fault behavior are critically evaluated using 1D numerical velocity stepping and spring-slider experiments. Since no new physical parameters are introduced, our model may be straightforwardly used to extend the existing modeling techniques. The model predicts the emergence of complex patterns of shear localization and delocalization that may inform the interpretation of complex damage distributions observed around faults in nature.

Plain Language Summary

How, where, and when earthquakes nucleate is one of the great questions in science and society that, despite steady progress, has hardly been answered to any practical degree. Based on field observations, laboratory experiments, and theoretical work it is believed that a cocktail of escalating mechanical, chemical, and thermal grain-scale processes cause the sudden and rapid onset of earthquakes. The net effect of these processes are characterized by an immediate strengthening and a gradual weakening response to deformation and are unified in simplified form in the theory of 'rate and state friction'. This theory is commonly used in computer simulations of earthquake sequences. We point out that rate and state friction, unlike some physical theories of earthquake

rupture, does not incorporate a diffusion process such as for example heat conduction. We show the introduction of an artificial diffusion process can prevent the mathematical reduction of a fault zone to a two-dimensional interface while retaining the properties of the original friction law. This in turn enables simulation techniques that rely on an interface-free description of the earth and promise to provide new insights into the spontaneous organization of seismic and aseismic phenomena in developing fault zones.

1 Introduction

Over the second half of the last century the study of the sliding behavior of frictional surfaces, such as those believed to occur in the earth's crust, has led to a general understanding that it is governed by competition of stabilizing viscous-like effects and potentially destabilizing processes that affect the texture of the interface in a time-dependent manner and self-organize into periods during which the interface is arbitrarily close to elastic stick punctuated with periods of anelastic slip (Rabinowicz, 1958; Brace & Byerlee, 1966; Bowden & Tabor, 1966; Scholz et al., 1972; Ida, 1972; Dieterich, 1978; Ruina, 1980, 1983; Dieterich & Kilgore, 1994).

1.1 Rate- and State-dependent Friction

This has led to the proposition of the phenomenological *rate- and state-dependent friction law* by Dieterich (1978, 1979a); Ruina (1980, 1983), which reads in its general form as

$$f = f_0 + a \log V + b \log \Theta. \tag{1}$$

In this law, the friction coefficient f is given by the linear combination of a reference friction coefficient f_0, a logarithmic contribution from dimensionless slip rate V multiplied by a coefficient a, and another logarithmic contribution from a dimensionless *state* variable Θ multiplied by a coefficient b. All quantities are taken to be positive.

The rate effect or *direct effect* is to a variable degree strengthening with increasing slip rate and thus exerts a stabilizing influence on the frictional interface for any strictly positive a. The state effect or *evolution effect* is to be governed by an evolution law that has the evolving steady state $\Theta_{\mathrm{ss}} V(t) = 1$. Then, if $b > a$, this allows an externally loaded system with a frictional interface to move to a configuration with lower elastic strain energy and thus a frictional-mechanical instability may occur.

One of the most widely used state evolution laws is the *aging law* of Ruina (1980, 1983), given by

$$\dot{\Theta} = r_0 \left(1 - V\Theta\right), \tag{2}$$

with r_0 a rate constant that is commonly expressed by dividing the reference velocity v_0 by a critical slip distance d_c. It should be noted that several other evolution laws that govern the state variable have been proposed. The most commonly cited contender is the *slip law* (Dieterich, 1979a; Ruina, 1980, 1983). In this work we do not consider it because its functional form ($\dot{\Theta} \propto \Theta\, V \log \Theta\, V$) is unfortunately not amenable to the type of analytical treatment that is performed. We note that rate and state friction captures only the first-order behavior of the stick-slip cycle and that laboratory experiments have brought to light many secondary effects (e.g. Ruina, 1983; Marone, 1998; Mair & Marone, 1999; Chester, 1994; Passelègue et al., 2020).

Following Amonton's law the friction coefficient is expressed as the ratio of shear stress τ to normal stress σ, implying a cohesionless fault. It is further assumed that the fault is always critically loaded, i.e. shear stress is equal to the shear strength and $V > 0$. The assumption of criticality is necessary because (1) degenerates at $V = 0$, but has also been suggested to be realistic for faults in nature (Bak & Tang, 1989).

The rate and state friction law is frequently applied in numerical studies of fault slip, whether that be a study over the course of a single earthquake or a complex sequence of slip transients (e.g. Dieterich, 1979b; P. G. Okubo, 1989; Rice, 1993; Rice & Ben-Zion, 1996; Ben-Zion & Rice, 1997; Lapusta et al., 2000; Lapusta & Liu, 2009; Rubin & Ampuero, 2005; Gabriel et al., 2012; Jiang & Lapusta, 2016; D. Li & Liu, 2017; Erickson et al., 2020). The use of interfacial friction laws in these models necessitates the treatment of faults as mesh features on which internal boundary conditions can be applied. It can be laborious to construct such meshes for non-trivial fault geometries, and the procedure does not scale well with increasing fault network complexity. There is currently no established method to construct meshes with time-dependent geometry and topology reflecting an evolving fault system, but we refer to K. Okubo et al. (2019) for work in this direction. With some exceptions (e.g. D. Li & Liu, 2017; Sathiakumar et al., 2020; Perez-Silva et al., 2021) many studies focus on the already complex behavior of long-term sequences of slip transients on linear or planar faults in simple domains.

1.2 Continuum Fault Rheology

There is strong field and experimental evidence that faults in nature are not infinitely thin planes but consist of complex evolving networks of strongly localized shear zones within a wider region of damaged host rock (Tchalenko, 1970; Granier, 1985; Chester & Chester, 1998; Katz et al., 2004; Savage & Brodsky, 2011; Faulkner et al., 2011; Barth et al., 2013; Passelègue et al., 2016; Perrin et al., 2016; Locatelli et al., 2018, 2019; Pozzi et al., 2018, 2019, 2021; Ritter, Rosenau, & Oncken, 2018; Ritter, Santimano, et al., 2018; Ross et al., 2019). This knowledge has spurred a search for intermediate-scale continuum homogenizations of distributed micro-scale processes that obey the rate and state framework in a somewhat generalized sense. Currently proposed models focus on rheological feedback mechanisms involving temperature (Rice, 2006; Braeck & Podladchikov, 2007; John et al., 2009; Garagash, 2012; Roubek, 2014; Rice et al., 2014; Platt et al., 2014; Thielmann et al., 2015; Thielmann, 2018; Pozzi et al., 2021), grain size (Rozel et al., 2011; Thielmann et al., 2015; Thielmann, 2018; Barbot, 2019; Pozzi et al., 2021), porosity (Sleep, 1997; Niemeijer & Spiers, 2007; Chen & Spiers, 2016; Van den Ende et al., 2018), fluid pressure (Barbot & Fialko, 2010; Garagash, 2012; Rice et al., 2014; Platt et al., 2014; Poulet et al., 2014; Veveakis et al., 2014; Alevizos et al., 2014; Rattez, Stefanou, & Sulem, 2018; Rattez, Stefanou, Sulem, Veveakis, & Poulet, 2018), damage (Lyakhovsky et al., 2011; Lyakhovsky & Ben-Zion, 2014b, 2014a; Lyakhovsky et al., 2016; Kurzon et al., 2019, 2020), granular physics (Daub et al., 2008; Daub & Carlson, 2008, 2009; Daub et al., 2010; Hermundstad et al., 2010; Elbanna & Carlson, 2014) or combinations thereof. Continuum models employing a rate and state formulation with the slip rate and tractions replaced by strain rate and stress invariants have been developed by Herrendörfer et al. (2018) and applied to self-organizing fault systems in Preuss et al. (2019, 2020); DalZilio et al. (2021); Behr et al. (2021). Since the state of the subsurface is typically difficult to accurately characterize on a regional basis, these models are important to develop a generic understanding of the long-term evolution of seismogenic fault systems.

A problem commonly associated with strain softening rheologies without internal length scale is unconstrained localization (Hobbs et al., 1990). As a shear zone localizes to the scale of a mesh element or grid cell they become poorly resolved and cause mesh dependence of the simulation. Models that possess an internal length scale perpendicular to the direction of shear may not suffer from this issue, provided that this length scale is resolved by the discretization. For example, as thermo-rheological feedback mech-

anism causes localization the heat diffusion term may gain prominence until localizing and delocalizing (e.g. diffusing) influences find a balance. Unfortunately some of the proposed fault rheologies don't possess an internal length scale, and some possess one that has physical meaning and can't be changed without changing the behavior of interest. If such a physical length scale is very small in nature, then that may place insurmountable constraints on the mesh resolution in two- or three-dimensional regional simulations (e.g. Rice et al., 2014; Platt et al., 2014). A few models *do* contain a controllable internal length scale; these are the unified rate and state friction theory of Sleep (1997) and the damage-breakage rheology of Lyakhovsky and Ben-Zion (2014b, 2014a); Lyakhovsky et al. (2016); Kurzon et al. (2019, 2020).

1.3 Objective

In summary, we state the need for a rigorous treatment of a continuum framework in which Dieterich–Ruina-type rate and state friction can be embedded based on the following observations:

1. Rate and state friction is a useful and powerful homogenization of the complex micro-scale processes that give rise to slip transients in nature.
2. Numerical modeling exploration of the complex and evolving relation that fault networks have with their tectonic environment is hampered by the long-standing challenges with mesh adaptive generation as well as lack of detailed knowledge of the structure of the subsurface.
3. The description of fault friction as a continuum process can be an elegant and practical way to avoid these problems provided they possess a controllable internal length scale in order to produce sensible results given reasonable computational resources.
4. A continuum faulting rheology may readily model important transient effects that might be missed in an equivalent interfacial description.
5. Much effort has been dedicated to understanding rate and state friction as an ingredient in numerical models – not all proposed continuum rheologies can equally benefit from this.

In response, we develop a Dieterich–Ruina-type continuum reformulation of rate and state friction that

1. possesses a controllable internal length scale λ_0 that enforces a constraint on localization,
2. yields mesh-independent results once λ_0 is sufficiently well resolved,
3. retains the parameters to the interfacial friction (1) and evolution (2) laws and yields similar trajectories of V and f for the same choice of parameter values, but that also
4. has interesting implications for fault zone behavior during and between earthquakes.

The model we propose here bears close relation to the earlier works of Sleep (1997) and Herrendörfer et al. (2018). Contrary to Sleep (1997) we refrain from in-depth discussion of fault physics but concentrate on the introduced spatial regularization, which is also a necessary extension of Herrendörfer et al. (2018).

In Section 2 we will motivate our choice of constitutive model, focusing on the formal links between interfacial and continuum models. In Section 3 we will present our equivalent continuum rate and state rheology. In Section 4 we build our rheology into a numerical model of a simplified one-dimensional fault transect to demonstrate both its localization and delocalization behavior as well as its similarities and differences with respect to the original laws of Dieterich and Ruina. We close with a discussion – in which we will comment on possible links to existing theories of the physics of faulting – and a summary in Sections 5 and 6.

2 Material Model and Continuum Mechanics

We will first describe how the existing interfacial description of rate and state friction can be embedded in a continuum mechanical framework with internal discontinuities, and then generalize this to internal friction and distributed anelastic deformation in an interface-free model.

2.1 A Fault Reference Frame

We consider the immediate neighborhood $\Omega \subset \mathbb{R}^3$ of a section of fault $\Gamma \subset \Omega$ (Figure 1) that is described as the set of points belonging to a differentiable 2D manifold. We invoke the existence of a surjective function $\vec{x}_f(\vec{x})$ that maps a coordinate $\vec{x} \in \Omega$ to its nearest Euclidian neighbor $\vec{x}_f \in \Gamma$. Next, we define the fault transect $T(\vec{x}_f) \subset \Omega$ as the 1D curve that includes all the points $\vec{x} \in \Omega$ that map to a particular fault co-

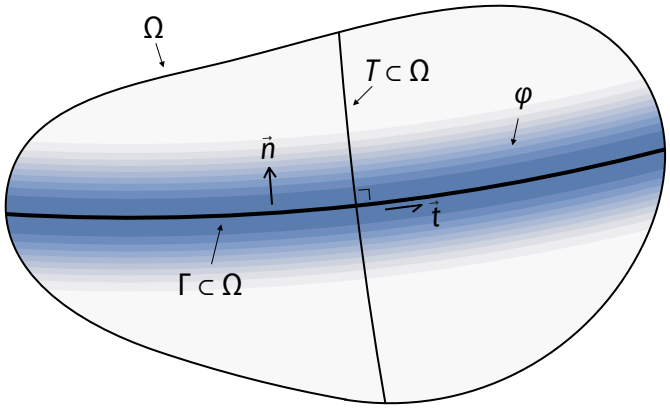

Figure 1. A region Ω in the enclosing a portion of a fault interface Γ, with normal and tangent vectors $\vec{n}$ and $\vec{t}$ indicated. A transect T is drawn perpendicular to the fault. The fault interface may be substituted by a shear band on which anelastic shear strain is distributed according to $\varphi(\vec{x})$, as indicated by the colored contours.

ordinate $\vec{x}_f \in \Gamma$, in short $T = \vec{x}_f^{-1}(\{\vec{x}\}) \subset \Omega$. Finally, we let $\hat{n}(\vec{x}) \in \mathbb{R}^3,\ \vec{x} \in \Omega$ be a unit vector field locally tangent to the transecting curve $T(\vec{x}_f(\vec{x}))$ and therefore normal to the fault.

The *state* of the body Ω is described by a symmetric Cauchy stress tensor field $\boldsymbol{\sigma}(t, \vec{x}) \in \mathbb{R}^3 \otimes \mathbb{R}^3$ and a velocity vector field $v_0\, \vec{v}(t, \vec{x}) \in \mathbb{R}^3$. Their governing equations are stated after the introduction of some further scaffolding. Let t denote time.

We define the normal traction $\sigma = -\hat{n} \cdot \boldsymbol{\sigma} \cdot \hat{n}$ (positive in compression), the shear traction vector $\vec{\tau} = \boldsymbol{\sigma} \cdot \hat{n} + \sigma\, \hat{n}$, and the shear traction magnitude $\tau = \sqrt{\vec{\tau} \cdot \vec{\tau}}$. The rupture process is most efficient when the slip rate vector $\vec{V} \in \mathbb{R}^3$ is parallel to the in-plane shear traction vector $\vec{\tau}$, and so alongside the full velocity field $\vec{v}$ we define a purely slip-induced velocity double couple field $\vec{v}_{\rightleftharpoons}$ around the coordinate $\vec{x}$ as

$$\vec{v}_{\rightleftharpoons}(t, \vec{x}) := \tfrac{1}{2} V(t, \vec{x}_f)\, \hat{t}(\vec{x}_f)\, \mathrm{sgn}(\hat{n}(\vec{x}_f) \cdot [\vec{x} - \vec{x}_f]), \tag{3}$$

with $\hat{t} \in \mathbb{R}^3 := \vec{\tau}/\tau$ the unit tangent vector and sgn the sign function, which has $\mathrm{sgn}(0) = 0$.

2.2 Assumptions

In the following, we will assume that continuum processes that occur around the fault are predominantly confined to the small neighborhood Ω, which itself includes only a small section of a whole fault. The neighborhood Ω is assumed to be sufficiently small that in-plane variations of fault properties such as state, slip rate, and curvature are negligible. This assumption comes at a loss of generality, especially around fault kinks, branches, and tips, but does not neccesarily limit the applicability of our formulation. We will revisit this assumption in the Section 5.5 of the Discussion.

We also assume that the material has uniform static properties throughout the domain. This means that the gouge layer is effectively assumed to be infinitely wide, or alternatively that the only quality that distinguishes the fault gouge from the host rock is it's 'state'. Undoubtedly there is a collection of state variables that govern the long-term structure of fault zones in nature that are not modeled here. We will revisit this assumption too in Section 5.5, but in the mean time note that its impact is largely confined to inferences made about the de-localization behavior of rate strengthening faults in Section 3

2.3 Generalization of a Fault to a Shear Band

We now generalize the fault interface to a distributed shear band by substituting a smoothed step function $\Phi = \Phi(t, x; \lambda_0) \in [-1, +1]$, $\lambda_0 > 0$ (e.g. $\Phi(t, x; \lambda_0) := \tanh(x/\lambda_0)$) for the sign function in (3). We require that the chosen function converges in a pointwise manner to the sign function as $\lambda_0 \to 0^+$, such that (3) can be seen as the result of taking the limit

$$\begin{aligned} \vec{v}_{\rightleftharpoons}(t, \vec{x}; \lambda_0) &:= \tfrac{1}{2} V(t, \vec{x}_f)\, \hat{t}(\vec{x})\, \Phi\left(t, \hat{n} \cdot \left[\frac{\vec{x} - \vec{x}_f}{\lambda_0}\right]\right) \\ \vec{v}_{\rightleftharpoons}(t, \vec{x}) &:= \lim_{\lambda_0 \to 0^+} \vec{v}_{\rightleftharpoons}(t, \vec{x}; \lambda_0)\,. \end{aligned} \tag{4}$$

Recognizing that fault slip is an anelastic process, the anelastic strain rate tensor $\dot{\boldsymbol{\varepsilon}}_{\neg \mathrm{e}}$ is defined to be the symmetric gradient ∇^s of the (continuous) slip-induced velocity field $\vec{v}_{\rightleftharpoons}(t, \vec{x}; \lambda_0)$:

$$\begin{aligned} \dot{\boldsymbol{\varepsilon}}_{\neg \mathrm{e}}(t, \vec{x}) &:= v_0 \nabla^s \vec{v}_{\rightleftharpoons}(t, \vec{x}; \lambda_0) := \tfrac{1}{2} v_0 \left[(\nabla \vec{v}_{\rightleftharpoons})^{\mathrm{T}} + (\nabla \vec{v}_{\rightleftharpoons})\right](t, \vec{x}; \lambda_0) \\ &\approx \frac{1}{2}\frac{v_0}{\lambda_0} V(t, \vec{x}_f) \varphi\left(t, \hat{n} \cdot \left[\frac{\vec{x} - \vec{x}_f}{\lambda_0}\right]\right) \left[\hat{t} \otimes \hat{n} + \hat{n} \otimes \hat{t}\right](\vec{x}) =: \gamma_0 \gamma(t, \vec{x}) \hat{\boldsymbol{s}}(\vec{x}). \end{aligned} \tag{5}$$

Here the *equivalent anelastic shear strain rate* (a scalar value) has been introduced as

$$\gamma_0\dot{\gamma}(t,x) = v_0 V(t)\varphi(t,x), \tag{6}$$

with reference value $\gamma_0 = v_0/\lambda_0$ and *strain rate distribution* $\varphi(t,x;\lambda_0) = \partial_x\Phi(t,x;\lambda_0)$ (illustrated in blue shading in Figure 1). The symbol $\hat{\boldsymbol{s}}$ denotes the Schmidt tensor $\hat{\boldsymbol{s}} = \frac{1}{2}\left[\hat{t}\otimes\hat{n} + \hat{n}\otimes\hat{t}\right]$, and its *Frobenius norm* $\|\hat{\boldsymbol{s}}\| = \sqrt{\mathrm{tr}\,\hat{\boldsymbol{s}}^{\mathrm{T}}\hat{\boldsymbol{s}}} = 1/\sqrt{2}$, which is an important property to maintain when generalizing $\hat{\boldsymbol{s}}$ later on because it is compatible with the interpretation of $\gamma_0\dot{\gamma}$ as the fault-perpendicular derivative of fault-parallel velocity, i.e. as an accurate measure of simple shear. In the derivation of (5) the gradients of V and the orthonormal bases $\hat{t}$ and $\hat{n}$ have been dropped under the assumption set out in Section 2.2 that they are small compared to the gradient of Φ. For this assumption to be met, the across-fault length scale λ_0 must be small compared to the along-fault length scales that exists in the interfacial rate and state formulation.

Since the derivative of half a smoothed step function $\frac{1}{2}\Phi(x) \in [-\frac{1}{2},+\frac{1}{2}]$ has the properties of a distribution – is a non-negative and importantly integrates to unity over its domain – we may express the integral relation

$$\mathcal{V}(\dot{\gamma}) := \gamma_0 \int_T \dot{\gamma}(t,\vec{x})\mathrm{d}x = v_0 V(t,\vec{x}_f) \tag{7}$$

over the fault transect T, which is again compatible with the notion that $\gamma_0\dot{\gamma}$ is a measure of simple shear in the fault reference frame. We note that the above integral strictly requires that all anelastic strain occurs over the length of the fault transect T, implying that the anelastic strain rate distribution has compact support in space. In practice we can relax that requirement if said distribution is narrow and thin-tailed.

Subtracting the anelastic shear strain rate tensor $\dot{\boldsymbol{\varepsilon}}_{\neg\mathrm{e}}$ from the full strain rate tensor $\dot{\boldsymbol{\varepsilon}} = \nabla^s\vec{v}$ results in the elastic strain rate tensor. This additive decomposition of elastic and anelastic strains is known as the Maxwell model. Application of Hooke's law of linear elasticity to the elastic strain rate tensor leads to

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{S}\,[\dot{\boldsymbol{\varepsilon}} - \gamma_0\dot{\gamma}\hat{\boldsymbol{s}}] \tag{8}$$

with fourth-order elastic stiffness tensor $\boldsymbol{S}$ and the dot over a symbol denoting the time derivative. The relation above is standard in the context of elasto-plasticity (e.g. Mühlhaus & Aifantis, 1991), in which $\gamma_0\dot{\gamma}$ is known as the plastic multiplier and $\hat{\boldsymbol{s}}$ coincides with the derivative of the magnitude of shear traction τ with respect to the full stress ten-

sor $\boldsymbol{\sigma}$. In this case τ is to be interpreted as the *non-associated* plastic potential of a material that does not undergo anelastic volume change.

For completeness we list ordinary differential equation that governs $\vec{v}$, the *momentum balance equation*:

$$v_0\dot{\vec{v}} - \vec{g} = \rho^{-1}\nabla\cdot\boldsymbol{\sigma}. \tag{9}$$

Here $\vec{g}$ is the acceleration due to gravity and ρ the mass density. Gravity may be ignored in the remainder of this work at no loss of generality and the inertial term $\dot{\vec{v}}$ may be ignored at *some* loss of generality under the assumption of static momentum balance.

2.4 Plasticity and Coordinate Invariance

For a spontaneously developing fault zone the shear and normal stress can not be defined in relation to a known plane. Plasticity models avoid this problem through the use of stress tensor invariants and scalar parameters. The Mohr-Coulomb plasticity model generates anelastic shear strain parallel to the Schmidt tensor $\hat{\boldsymbol{s}} = \frac{1}{2}(\hat{n}\otimes\hat{t} + \hat{t}\otimes\hat{n})$, in which $\hat{t}$ makes an angle $\phi = \tan^{-1} f$ with respect to the largest principal stress axis $\vec{\sigma}_1$, and $\hat{t}$ and $\hat{n}$ form an orthogonal triad of vectors together with the intermediate principal stress axis $\vec{\sigma}_2$. In other words, anelastic shear strain occurs in a plane perpendicular to $\vec{\sigma}_2$. The corresponding cohesionless yield criterion is written as

$$\frac{1}{2}\left(\sigma_1 - \sigma_3\right) = -\sin(\phi)\,\frac{1}{2}\left(\sigma_1 + \sigma_3\right), \tag{10}$$

with σ_1 and σ_3 the magnitudes of the largest and smallest principal stresses $\vec{\sigma}_1$ and $\vec{\sigma}_3$. The quantity $-\frac{1}{2}\left(\sigma_1 + \sigma_3\right) = \sigma_m$ is called the *mean stress* and $\frac{1}{2}\left(\sigma_1 - \sigma_3\right) = \sigma_d$ the *effective deviatoric stress*. We note that the angle of a shear band that is generated does not need to be parallel to the direction of shear strain (Vardoulakis, 1980; Vermeer, 1990; Marone et al., 1992; Kaus, 2010; Le Pourhiet, 2013), and moreover that a local change in friction is not expected to immediately alter the macroscopic fault angle (e.g. Preuss et al., 2019).

In the Mohr-Coulomb model, the friction coefficient $f = \tan(\phi)$ of a yielding or even of a non-yielding material may be expressed from (10) as a function of the stress tensor $\boldsymbol{\sigma}$ as

$$f(\boldsymbol{\sigma}) = \sigma_{\mathrm{d}}(\boldsymbol{\sigma})\left[\sigma_{\mathrm{m}}(\boldsymbol{\sigma})^2 - \sigma_{\mathrm{d}}(\boldsymbol{\sigma})^2\right]^{-\frac{1}{2}}. \tag{11}$$

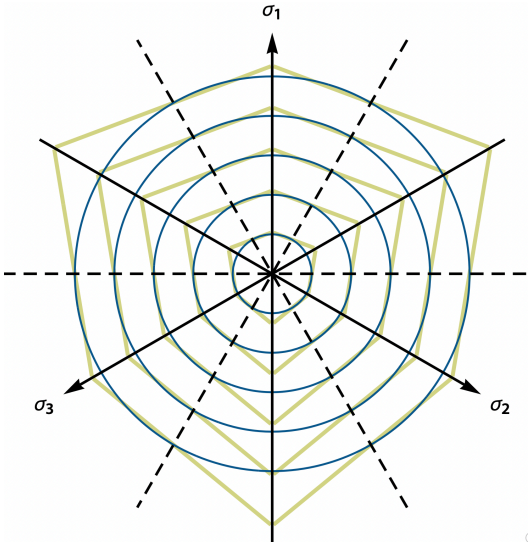

Figure 2. Comparison of Mohr-Coulomb (thick yellow lines) and Drucker-Prager yield envelopes (thin blue lines) at increments of pressure. The field of view is the octahedral plane of principal stress space. Projections of principal stress axes are indicated with arrows, and projections of planes of simple shear are indicated with dashed lines. A friction coefficient of 0.3 was used to generate this figure.

Herrendörfer et al. (2018) used the Drucker-Prager model (Drucker & Prager, 1952) as a simple and smooth approximation to Mohr-Coulomb plasticity. The model is defined in terms of the straightforwardly computable invariants

1. pressure $p = -\frac{1}{3}\operatorname{tr}\boldsymbol{\sigma}$, and
2. effective shear stress τ_e proportional to the *Frobenius norm* of the deviatoric stress tensor $\boldsymbol{\tau} := \boldsymbol{\sigma} + p\,\boldsymbol{\delta}$, with $\boldsymbol{\delta}$ the Kronecker delta:

$$\tau_\mathrm{e} := \frac{\|\boldsymbol{\tau}\|}{\sqrt{2}} = \sqrt{\tfrac{1}{2}\mathrm{tr}\,(\boldsymbol{\tau}^\mathrm{T}\cdot\boldsymbol{\tau})}. \tag{12}$$

In this model the existing definition of the Schmidt tensor $\hat{\boldsymbol{s}}$ is no longer applicable because the unit vectors $\hat{t}$ and $\hat{n}$ are not known *a priori*, and are not even uniquely constrained *a posteriori*. Keeping in line with plasticity theory, $\hat{\boldsymbol{s}}$ is redefined as the stress derivative of the non-associated plastic potential τ_e:

$$\hat{\boldsymbol{s}} := \frac{\partial\tau_\mathrm{e}}{\partial\boldsymbol{\sigma}} = \frac{\boldsymbol{\tau}}{\tau_\mathrm{e}}.$$

This definition still satisfies the criterion set out in the preceding section that $\|\hat{\boldsymbol{s}}\| = 1/\sqrt{2}$.

We calibrate the Drucker-Prager model to the Mohr-Coulomb model around a reference stress state $\boldsymbol{\sigma}_0$ of simple shear combined with isotropic compression, which encompasses all stress states possible in 2D plane strain and may be considered the most relevant stress state even in 3D tectonic settings. Let $\vec{\sigma}_0$ be given in ordered principal stress space by

$$\vec{\sigma}_0 := -p_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \tau_{\mathrm{e},0} \begin{bmatrix} +1 \\ 0 \\ -1 \end{bmatrix}.$$

At $\vec{\sigma} = \vec{\sigma}_0$, Mohr-Coulomb and Drucker-Prager measures of effective shear and normal stress coincide (up to a sign) and thus the Drucker-Prager model may be written as

$$\tau_{\mathrm{e}}(\boldsymbol{\sigma}) = \sin(\phi)\, p(\boldsymbol{\sigma}). \tag{13}$$

Away from $\vec{\sigma}_0$ (and towards a uniaxial stress state) the Drucker-Prager yield surface becomes an increasingly worse approximation to the Mohr-Coulomb yield surface (Figure 2). Higher-order approximations like the Willam-Warnke yield envelope (Ulm et al., 1999) may be considered too. The yield surfaces discussed in this section serve as part of a general model proposition – in the one-dimensional numerical experiments considered in this work only the two effective stresses of Amonton's friction law are defined.

Similar to the Mohr-Coulomb model (11), the friction coefficient $f = \tan\phi$ of the Drucker-Prager model may also be expressed from (13) as a function of the stress tensor $\boldsymbol{\sigma}$ as

$$f(\boldsymbol{\sigma}) = \tau_{\mathrm{e}}(\boldsymbol{\sigma}) \left[p(\boldsymbol{\sigma})^2 - \tau_{\mathrm{e}}(\boldsymbol{\sigma})^2 \right]^{-\frac{1}{2}}. \tag{14}$$

In both the Mohr-Coulomb model and the Drucker-Prager model, the friction coefficient becomes ill-posed whenever the effective shear stresses σ_{d} or τ_{e} exceed the effective normal stresses σ_m or p. We have found a practical solution to be to add a constant value to the effective normal stresses, which for steady-state friction behaves as a cohesive strength.

Despite adopting a plasticity framework to enforce the frictional yield constraint we emphasise that the resulting model is still best regarded as 'frictional viscous' rather than frictional plastic because it lacks a distinct boundary between yielding and non-yielding states in space and time – the yield constraint (an equality, not an inequality) is enforced everywhere and anytime, and at any stress.

3 Continuum Evolution Law

Our goal is to find a continuum state evolution law that produces nearly the same history of loading and sliding as the original formulation of Dieterich and Ruina on a discrete fault. Like $V(t)$ is a 'global' measure of $\gamma(t,x)$ over a fictitious fault transect T, we introduce a local variable $\theta(t,x)$ of which $\Theta(t)$ is a global measure. We target continuum friction and evolution laws expressed in terms of γ and θ of the form

$$f = f_0 + a\log\gamma + b\log\theta, \tag{15}$$

$$\dot{\theta} = r_0\left(c_2 - c_1\mathcal{M}(\gamma)\,\theta\right), \tag{16}$$

with $\mathcal{M}$ a non-local interaction operator that will be further elaborated in Section 3.3 and c_1 and c_2 additional coefficients that are required to calibrate (15) and (16) to their respective interfacial equivalent. The local friction and evolution laws (15) and (16) retain the structure of their interfacial counterparts (1) and (2) and generalize the continuum formulations of Sleep (1997) and Herrendörfer et al. (2018). We will revisit this connection in the Section 5.4 of the Discussion. The targeted form of the continuum equations does not come out of the blue. In coming sections we hope to make clear how it arises.

3.1 Analytical Framework

By

1. reorganizing the rate and state friction law (1) into a definition of $\Theta(V,f)$,
2. differentiating (1) with respect to time and reorganizing it into a definition of $\dot{V}(t)$, and
3. substituting the definition of $\Theta(V,f)$ and the definition of $\dot{\Theta}$ that follows from the aging law (2) into the definition of $\dot{V}(t)$,

we arrive at the ordinary differential equation

$$a\,\dot{V} = b\,r_0\left[V^2 - V\,\Theta(V,f)^{-1}\right] + V\dot{f}, \tag{17a}$$

$$\Theta(V,f)^{-1} := V^{\frac{a}{b}}\exp\left(-[f-f_0]/b\right) \tag{17b}$$

in which the friction coefficient $f = f(\boldsymbol{\sigma};\hat{t},\hat{n})$ is treated as a property of – rather than a constraint on – the stress tensor $\boldsymbol{\sigma}$.

By applying the same procedure to the proposed continuum friction and aging laws (15) and (16) we obtain the analogous ODE for γ:

$$a\,\dot{\gamma} = b\,r_0\left[c_1\gamma\mathcal{M}(\gamma) - c_2\gamma\,\theta(\gamma,f)^{-1}\right] + \gamma\dot{f}, \tag{18a}$$

$$\theta(\gamma,f)^{-1} := \gamma^{\frac{a}{b}}\exp\left(-[f-f_0]/b\right), \tag{18b}$$

in which $f = f(\boldsymbol{\sigma})$ is given by the yield criterion (14) and $\dot{f} = \dot{f}(\boldsymbol{\sigma},\dot{\boldsymbol{\sigma}})$ is expressible using the same yield criterion and the elastic constitutive equation (8).

Equations (17a,b) and (18a,b) are useful in mathematical and numerical analysis of the problem because they eliminate a variable and an algebraic constraint, and at the same time provide an ideal reference frame for establishing, and if necessary influencing, the ability of (18a,b) to reproduce the predictions made by (17a,b) in the sense of the integral relation (7).

In the following derivation we will assume that the field of effective friction $f(\vec{x}), \vec{x} \in \Omega$ in the neighborhood of a point $\vec{x}_f$ on the fault is the same whether it is generated by a hard or a soft discontinuity. This assumption ties into the principal assumption set out in Section 2.2, which is revisited in Discussion section 5.5. We will also assume that $\mathcal{M}(\square(t)\lozenge(x,\ldots)) = \square(t)\mathcal{M}(\lozenge(x,\ldots))$. This assumption is satisfied in Section 3.3. Combining (17a,b), (18a,b), (6), and (7) leads to the following useful relation that describes the evolution of the anelastic shear strain rate distribution over time:

$$\dot{\varphi} \propto V\left[\lambda_0\,c_1\,\varphi\,\mathcal{M}(\varphi) - \varphi\right] - \Theta(V,f)^{-1}\left[\lambda_0^{\frac{a}{b}}\,c_2\,\varphi^{1+\frac{a}{b}} - \varphi\right]. \tag{19}$$

This differential equation tells us several things. Most importantly, for γ to remain in accordance with V through the integral relation (7) over the transect T, the same integral of the left hand side of (19) must be zero. This requirement can only be met at abitrary (V,f) if both the term multiplied by V and the term multiplied by $\Theta(V,f)^{-1}$ integrate to zero over T. This in turn can not in general be guaranteed unless

$$c_1 := \mathcal{C}_1(\varphi) := \lambda_0^{-1}\left[\int \varphi(x)\mathcal{M}(\varphi)(x)\,\mathrm{d}x\right]^{-1}, \tag{20a}$$

$$c_2 := \mathcal{C}_2(\varphi) := \lambda_0^{-\frac{a}{b}}\left[\int \varphi(x)^{1+\frac{a}{b}}\,\mathrm{d}x\right]^{-1}. \tag{20b}$$

We consider it undesirable for coefficients to depend in a time-dependent way on integrals of the modeled quantities and therefore will be restricting our attention to specific regimes of interest during which the values of the coefficients $c_{1,2}$ can be predicted analytically. We then employ those predicted values as model-specific constants in time and

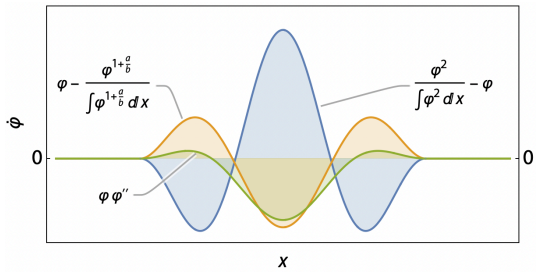

Figure 3. Influence of the different terms in Eq. (19) on the time derivative $\dot{\varphi}$ under the assumption that $\mathcal{M}(\varphi) = \varphi$. The first term (blue) can be seen to promote localization, while the second term (orange) favors the opposite. Also shown in this figure in green is the delocalizing influence of the nonlinear Laplacian term $\varphi(x)\varphi''(x)$, which is introduced in Sections 3.3 and 3.4. The input function φ follows a cosine-squared distribution.

space. We accept that this incurs a potential error whenever the state of the model is outside the selected regime, and will critically evaluate this error using numerical models in Section 4.

Assuming the coefficients $c_{1,2}$ have been chosen appropriately and for simplicity that $\mathcal{M}(\varphi) = \varphi$, two additional observations can be made on the basis of (19):

1. The term that is multiplied by V promotes localization if $\mathcal{M}(\varphi) = \varphi$ because φ^2 is a narrower distribution than φ (Figure 3).
2. The opposite (delocalization; Figure 3) holds true for the term that is multiplied by $\Theta(V, f)^{-1}$ on account of the corresponding minus sign and the requirement that $a, b > 0$.

This localizing and delocalizing behavior is of great interest and will be treated in detail in the following sections.

In the sections that follow we will assume to be operating on the fault-perpendicular transect T and will denote with the scalar coordinate x the distance with respect to the fault core, i.e. $x = \hat{n}(\vec{x}) \cdot [\vec{x} - \vec{x}_f(\vec{x})]$. This setting is in line with our assumption set out in Section 2.2 that across-fault variations in anelastic strain rate are more compact than its along-fault variations.

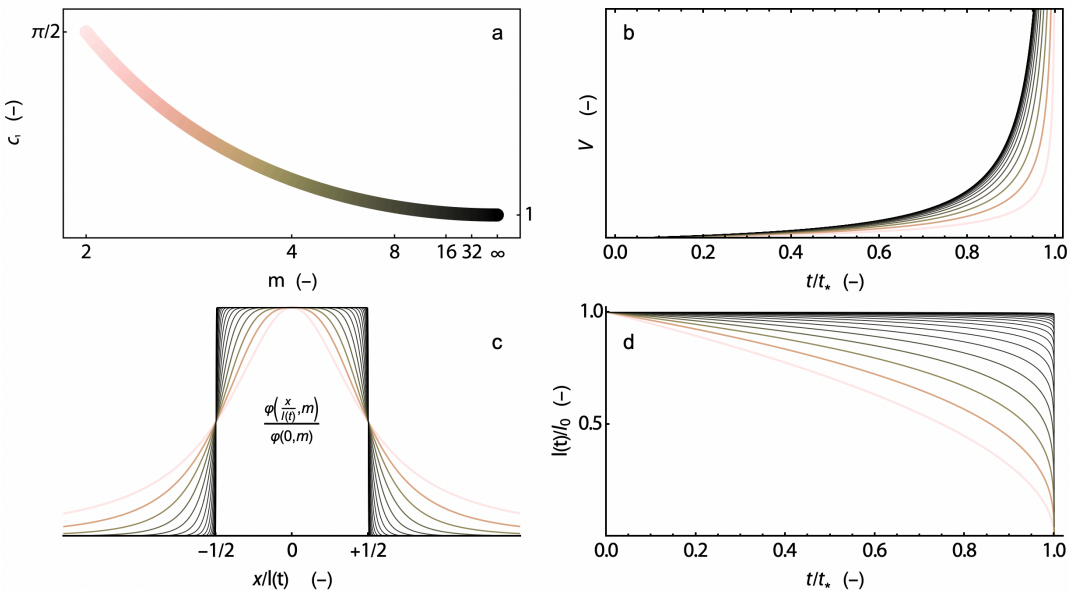

Figure 4. Asymptotic solutions during earthquake nucleation. a) Value of the dimensionless coefficient $c_1(m)$ versus distribution exponent m (see main text) represented on a reciprocal axis that ranges between 2 and ∞. The thick line acts also as a color bar approximately indicating the value of m at which the individual lines in the remaining figures are plotted. b) Dimensionless slip rate V versus dimensionless time t/t_* to the slip rate asymptote for different values $m \in \{2^{1.0}, 2^{1.5}, 2^{2.0} \ldots\}$. c) Nondimensionalized distributions $\varphi(x/l(t); m)/\varphi(0; m)$ versus dimensionless coordinate $x/l(t)$. d) Dynamic dimensionless length scale $l(t/t_*)/l_0$ versus dimensionless time t/t_*.

3.2 Runaway Sliding and Localization

Under rate-weakening conditions $a < b$, earthquakes occur as superexponential solutions to (17a,b). Prior to inertial damping the seismic slip rate V behaves as

$$\dot{V} \propto V^2.$$

This ODE has the solution

$$V(t) = V(0)\left(1 - \frac{t}{t_*}\right)^{-1}, \tag{21}$$

in which $t_* \propto V(0)^{-1}$ is the time of the (hypothetical) singularity.

Assuming $\mathcal{M}(\square) = \square$, the same behavior occurs in (18a,b) in the limit

$$\dot{\gamma} \propto c_1 \gamma^2, \tag{22}$$

which is similarly solved by

$$\gamma(t, x) = \gamma(0, x)\left(1 - \frac{t}{t_*(x)}\right)^{-1},$$

in which $t_*(x) \propto (c_1 \gamma(0, x))^{-1}$.

We can make the following qualitative observations:

1. small spatial variations in initial conditions $\gamma(0, x)$ can be amplified to infinity due to the fact that points with larger initial conditions on γ are closer to the strain rate asymptote at $t = t_*(x)$ than points with smaller initial conditions on γ, and
2. $\gamma(t, x)$ and $V(t)$ do not generally satisfy the integral relation (7) for all time $t \in [0, t_*)$.

Noting that this regime of runaway sliding corresponds in (19) to the limit

$$\dot{\varphi} \propto V\big[\lambda_0\, c_1\, \varphi\, \mathcal{M}(\varphi) - \varphi\big],$$

we can further illustrate the observations made above by choosing the class of solutions (Appendix A)

$$V(t) \propto \left(1 - \frac{t}{t_*}\right)^{m^{-1}-1}, \tag{23a}$$

$$\varphi(t,x;m) := \left[l(t)c_1(m)\left(1 + \left|\frac{2x}{l(t)}\right|^m\right)\right]^{-1}, \qquad m \in \mathbb{R} \geq 2 \tag{23b}$$

$$l(t) \propto V(t)^{(1-m)^{-1}}, \tag{23c}$$

$$\gamma_0\gamma(t,x;m) := v_0 V(t)\varphi(t,x;m), \tag{23d}$$

$$c_1 = c_1(m) := \frac{\pi}{m}\csc\frac{\pi}{m}, \tag{23e}$$

in which the family of strain rate distributions $\varphi(t,x;m)$ generated by the exponent $m \in \mathbb{R} \geq 2$ (Figure 4c) are of generalized Cauchy type and evolve in accordance with a time-variable characteristic width $l(t)$ (Figure 4d). We emphasize that, though illustrative, this class of solutions is by no means unique because (18a,b) contains no constraint on the spatial distribution of strain rate. Nevertheless, we note the following:

1. In the limit $m \to \infty$, the distribution $\varphi(t,x;m \to \infty)$ reduces to a uniform distribution on the constant interval $x \in [-l(0)/2, +l(0)/2]$. The parameter $c_1(m \to \infty) = 1$ (Figure 4a) and the solution (23a) reduces to the interfacial solution (21). However this limit also requires infinite mesh resolution or alternatively the definition of two mesh discontinuities, which defeats our intent.
2. In the other end-member case where $m = 2$, the distribution $\varphi(t,x;m = 2)$ is known as the Cauchy distribution, and its characteristic width $l(t) \propto V(t)^{-1}$ for a given initial condition $l(0) = l_0$. The parameter c_1 is chosen in such a way that the distribution $\varphi(t,x;m)$ integrates to one and the time of the singularity matches the prediction of the interfacial model. However, the trajectory of V *towards* the asymptote is unavoidably different (Figure 4b).

We therefore come to the qualitative conclusion that for initially smooth and numerically resolvable strain rate distributions, orders-of-magnitude increases of slip rate readily correspond to orders-of-magnitude increase of strain rate locality, placing insurmountable demands on mesh resolution. If left unconstrained, this process culminates in finite

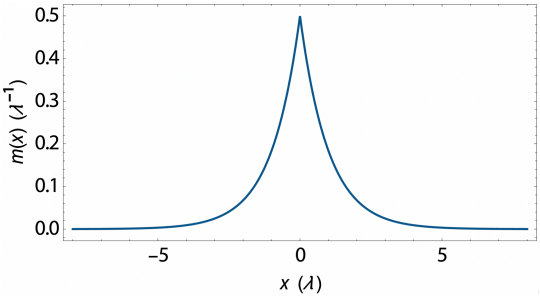

Figure 5. Plot of the mollifier $m(x) = \frac{1}{2\lambda} \exp\left(-\left|\frac{x}{\lambda}\right|\right)$.

time to an infinite slip rate concentrated in an infinitesimally thin region of space – a plane.

3.3 Non-local Equivalent Strain Rate

In order to counter this tendency towards complete localization to a plane we introduce the *non-local* interaction operator $\mathcal{M}$, which *mollifies* its operand by means of convolution with the symmetric distribution $m(x)$:

$$\mathcal{M}(\gamma)(t,x) := \int_{\mathrm{supp}(m)} m(\xi)\gamma(t, x-\xi)\mathrm{d}\xi. \tag{24}$$

We note that $\mathcal{M}(\gamma)(t,x)$ still satisfies the relation (7) given that the distribution m integrates to unity over its domain.

Evaluation of $\mathcal{M}(\gamma)$ requires the evaluation of the integral of γ over the support of $m(\xi)$ and therefore lacks the *mathematical* (as opposed to physical) locality that is desirable for continuum models. We follow Peerlings et al. (1996) in constructing and truncating a Taylor series linearization of (24). We find that a mollifier $m(x) = \frac{1}{2\lambda_0} \exp\left(-\left|\frac{x}{\lambda_0}\right|\right)$ (Figure 5) leads to series coefficients that are even powers of the length scale λ_0:

$$\mathcal{M}(\gamma)(t,x) = \sum_{n=0}^{\infty} \lambda_0^{2n} \partial_x^{2n} \gamma$$

This infinite sum can be rolled up into a recursive or *implicit* definition of $\mathcal{M} = \mathcal{M}_{\mathrm{im}}$

$$\mathcal{M}_{\mathrm{im}}^{-1}(\bar{\gamma}) := \bar{\gamma} - \lambda_0^2 \partial_x^2 \bar{\gamma}, \tag{25}$$

which is to be solved for the auxiliary variable $\bar{\gamma} = \mathcal{M}_{\mathrm{im}}(\gamma)$. Alternatively, the terms beyond $n = 1$ can be dropped, leading to the *explicit* definition of $\mathcal{M} = \mathcal{M}_{\mathrm{ex}}$

$$\mathcal{M}_{\mathrm{ex}}(\gamma) := \gamma + \lambda_0^2 \partial_x^2 \gamma. \tag{26}$$

We note that any truncation of the Taylor series into a linear combination of even derivatives of γ adheres to the integral property expressed in (7), even if these truncations lose correspondence to a concrete mollifier. The implicit definition (25) has numerically advantageous properties (Peerlings et al., 1996), but requires the potentially costly solution of an elliptic PDE. For this reason we will use the explicit definition (26) in this work and discuss the implicit definition (25) as an option for future implementations.

We generalize the second partial derivatives ∂_x^2 with respect to the x coordinate that appear in (25) and (26) to Laplacians ∇^2 in 3D, again assuming that variations of strain rate across the fault are much bigger than variations along it.

A convolution over slip history $d(t)$ with a kernel $w(d) \propto \exp -d/d_c$, which is identical to the kernel discussed above, was also employed by Ruina (1980, eq. 15, 16a, 17a) and Ruina (1983, eq. 13a,b) as a 'generator' of evolution laws. There, it represented the 'fading memory' of slip history on the fault. Here, we have consistently extended that notion with a non-local interaction of deformation that fades with distance away from a point in the shear band. Although not primarily intended to have physical meaning in this work, we note that the Laplacian of effective strain rate plays a role in granular physics (e.g. Bouzid et al., 2013, 2015). There, the length scale λ_0 is proportional to the grain size (Mühlhaus & Vardoulakis, 1987; Bouzid et al., 2013, 2015; Rice, 2006, and references in the latter).

3.4 Non-locality as a Localization Limiter

We repeat the analysis of Section 3.2, but now use the explicit non-local operator $\mathcal{M}_{\text{ex}}$ in

$$\dot{\gamma} \propto c_1 \gamma \, \mathcal{M}_{\text{ex}}(\gamma)$$

and

$$\dot{\varphi} \propto V\big[\lambda_0 \, c_1 \, \varphi \, \mathcal{M}_{\text{ex}}(\varphi) - \varphi\big]. \tag{27}$$

As shown in Appendix C, we find that it is solved by

$$V(t) \propto \left(1 - \frac{t}{t_*}\right)^{-1}, \tag{28a}$$

$$\varphi(x) := \begin{cases} \frac{1}{\pi\,\lambda_0} \cos^2\left(\frac{1}{2}\,\frac{x}{\lambda_0}\right) & \forall\, x \in [-\pi\lambda_0, +\pi\lambda_0] \\ 0 & \forall\, x \notin [-\pi\lambda_0, +\pi\lambda_0] \end{cases}, \tag{28b}$$

$$\gamma_0 \gamma(t, x) := v_0 V(t) \varphi(x), \tag{28c}$$

$$c_1 := 2\pi, \tag{28d}$$

and observe that

1. The trajectory of V towards its asymptote can be made to exactly match the prediction (21) of the interfacial rate and state friction formulation (17a,b) by choosing $c_1 = 2\pi$.
2. The cosine-squared distribution with prescribed size λ_0 is an attractive steady state during the process of earthquake nucleation. Further localization will not spontaneously occur. We interpret this distribution as striking a balance between the localizing tendency of "$\varphi^2 - \varphi$" and the delocalizing tendency of "$\varphi\nabla^2\varphi - \varphi$" (Equation (19) and Figure 3).

The implicit non-local operator $\mathcal{M}_{\text{im}}$ in the same context does not have analytical solutions that simultaneously satisfy $\dot{V} \propto V^2$ and $\dot{\varphi} = 0$ with non-degenerate distribution $\varphi = \varphi(x/\lambda_0)$.

3.5 Steady-state Friction and Deformation

As mentioned in Section 1.1, an important property of the state evolution law is the steady state $\Theta\, V = 1$. We wish to retain this steady state in the continuum equivalent (16) of the state evolution law, ideally at a steady strain rate distribution $\varphi(x)$, cf. (19). We are thus interested in a steady-state solution to

$$\dot{\varphi} \propto \lambda_0\, c_1\, \varphi\, \mathcal{M}_{\text{ex}}(\varphi) - \lambda_0^{\frac{a}{b}}\, c_2\, \varphi^{1+\frac{a}{b}}, \tag{29}$$

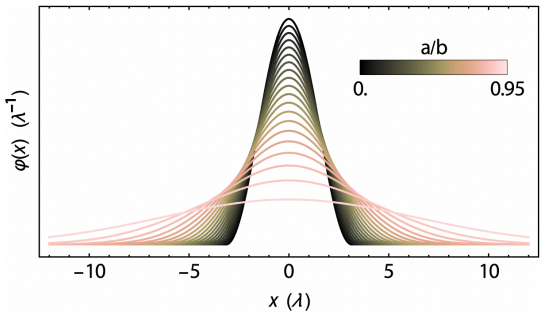

Figure 6. Steady-state distributions φ [cf. Eq. (30a)] of strain rate at various a/b versus distance x with respect to the fault core. The curve plotted for $a/b = 0$ also corresponds to the most extreme strain rate localization at any value of a/b (see main text).

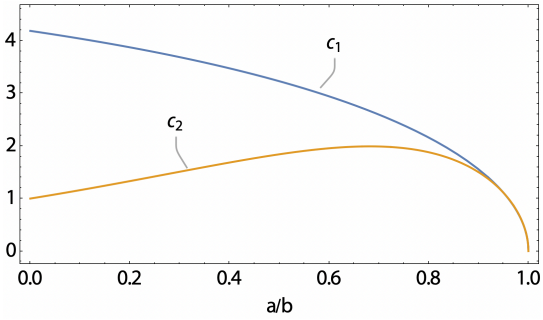

Figure 7. The coefficients $c_1\left(\frac{a}{b}\right)$ and $c_2\left(\frac{a}{b}\right)$, given by Eq. (31a) and (31b), are plotted versus the acceptable range of $\frac{a}{b}$.

which exists as long as $a, b > 0$ (standard requirements) and $a < b$ (a net weakening fault). As detailed in Appendix B, this solution is given by

$$\varphi_{\rm ss}(x) = \begin{cases} \varphi_0 \left[\cos\left(\frac{1}{2}\frac{x}{\lambda_1}\right)\right]^{2\left[1-\frac{a}{b}\right]^{-1}} & \forall\, x \in [-\pi\lambda_1, +\pi\lambda_1] \\ 0 & \forall\, x \notin [-\pi\lambda_1, +\pi\lambda_1] \end{cases} \tag{30a}$$

$$\lambda_1 := \lambda_0 \left[1 - \frac{a}{b}\right]^{-1}. \tag{30b}$$

Faults that are net strengthening or neutral ($a > b$) only possess a uniform steady state strain rate distribution that is bounded by the (fictitious) walls of the gouge layer (Sections 2.2 and 5.5).

The coefficients c_1 and c_2, and the newly introduced φ_0, are given by

$$c_1 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(1 + 2\left[1 - \frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2} + 2\left[1 - \frac{a}{b}\right]^{-1}\right)} \left[\lambda_1 \varphi_0\right]^{-2} \tag{31a}$$

$$c_2 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(2\left[1 - \frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2}\left[3 + \frac{a}{b}\right]\left[1 - \frac{a}{b}\right]^{-1}\right)} \left[\lambda_1 \varphi_0\right]^{-1-\frac{a}{b}} \tag{31b}$$

$$\varphi_0 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(1 + \left[1 - \frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2} + \left[1 - \frac{a}{b}\right]^{-1}\right)} \lambda_1^{-1}, \tag{31c}$$

with Γ the gamma function that has the property $\Gamma(n) = (n-1)! \forall n \in \mathbb{N}^+$. The coefficients c_1 and c_2 are plotted as functions of a/b in Figure 7. The solutions (30a) are plotted in Figure 6, in which we can see that steady-state anelastic strain rate distributions have finite width and are therefore numerically resolvable as long as the aforementioned requirements on a, b, and a/b are met.

3.6 Damping

Finite time blow-up of the solution (Section 3.5) under quasi-static loading conditions is understood to be the consequence of the model's inability to radiate or dissipate liberated potential strain energy away from the fault. It appears impossible to directly counteract this problem by implementing inertia with absorbing boundary conditions in the 1D models we present here, due to the resulting model's tendency to quickly establish a global steady state at an unphysically high strain rate. Therefore, inspired by the *radiation damping approximation* of Rice (1993) (see also Ben-Zion & Rice, 1995; Rice & Ben-Zion, 1996; Lapusta et al., 2000; Thomas et al., 2014) and the Kelvin-Voigt visco-plastic regularization of Needleman (1988) (see also F. Wu & Freund, 1984; Peirce et al., 1983; Duretz et al., 2019, 2020, 2021; de Borst & Duretz, 2020; Stathas & Stefanou, 2022), we add a linear viscous damping term $\eta\gamma$ in a parallel configuration to the continuum friction law:

$$f = f_0 + a \log \gamma + b \log \theta + \eta\gamma. \tag{32}$$

Informally, the action of this term may be understood to truncate the anelastic shear strain rate, which then causes the distribution of said strain to widen under continued loading. We formalize this claim by repeating the derivation in Section 3.1 of the anelastic shear strain rate ODE (18a,b), now incorporating the damping term in (32). We arrive at the following result, written in terms of the newly introduced auxiliary field ζ:

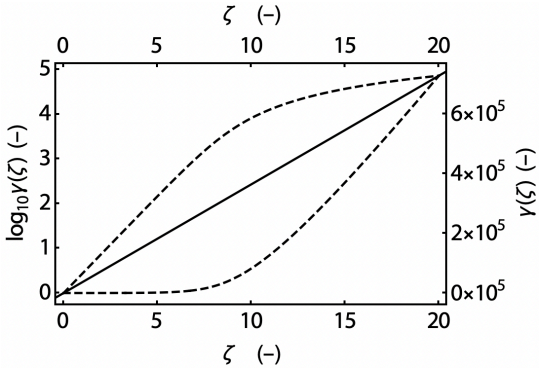

Figure 8. The function $\gamma(\zeta)$ defined in Eq. (33c) plotted on a logarithmic scale in the top left diagonal portion of the figure and on a linear scale in the opposing portion, highlighting the exponential behavior of $\gamma(\zeta)$ at low ζ and the linear behavior at high ζ.

$$a\,\dot{\zeta}(t,x) = b\,r_0\left(c_1\,\mathcal{M}_{\text{ex}}\circ\gamma\circ\zeta(t,x) - c_2\,\theta(\zeta,f)^{-1}\right) + \dot{f}(t), \tag{33a}$$

$$\theta(\zeta,f)^{-1} := \left(\gamma\circ\zeta(t,x)\right)^{\frac{a}{b}}\exp\left(-\tfrac{1}{b}\left[f - f_0 - \eta\,\gamma\circ\zeta(t,x)\right]\right), \tag{33b}$$

$$\gamma(\zeta) := \frac{a}{\eta}W_0\left(\frac{\eta}{a}\exp\zeta\right), \tag{33c}$$

with the principal branch $W_0(\Diamond)$ of the Lambert W function, which is an implicit transcendental function that is defined to be the solution to the equation $W_0\exp W_0 = \Diamond$. While we acknowledge that the use of implicit functions in physical descriptions is not ideal, from a practical standpoint there is little problem because reliable and fast algoritms for computing the common Lambert W function are included in many programming languages (Fritsch et al., 1973; Barry, Barry, & Culligan-Hensley, 1995; Barry, Culligan-Hensley, & Barry, 1995; Johansson, 2020).

As indicated in Figure 8, the function $\gamma(\zeta)$ is approximately linear at high ζ (and correspondingly high γ). This means that in this limit, (33) simplifies to

$$\eta\,\dot{\gamma}(t,x) = b\,r_0\left(c_1\,\mathcal{M}_{\text{ex}}\circ\gamma(t,x) - c_2\,\theta(\gamma,f)^{-1}\right) + \dot{f}(t), \tag{34a}$$

$$\theta(\gamma,f)^{-1} := \exp\left(-\tfrac{1}{b}\left[f - f_0 - \eta\,\gamma\right]\right), \tag{34b}$$

We can now compare reaction terms of (18) and (34). Ignoring spatial derivatives and the temporal derivative $\dot{f}$ of friction in both, the reaction term of (18) that is active at

low strain rate looks like $\dot{\gamma} \propto \gamma^2 - \alpha(f)\gamma^{1+\frac{a}{b}}$ while the reaction term of (34) that is active at high strain rate looks like $\dot{\gamma} \propto \gamma - \alpha(f)\exp(\frac{\eta}{b}\gamma)$. The former is a convex function of γ, the latter is a concave function of γ. We can thus see that the same conditions (large f, small $\alpha(f)$) that readily trigger runaway sliding in (18) also introduce a stable steady state at high γ in (34) and therefore also in (33). We note that this quality of having a stable steady state at non-zero γ means that (33) can be classified as a reaction-diffusion equation of generalized Fisher or KolmogorovPetrovskyPiskunov (KPP) type (Fisher, 1937; Kolmogorov et al., 1937). This type of equation is known to exhibit outward propagating (strain rate) fronts, and consequently does not admit a steady strain rate distribution. This markedly limits the validity of the analytical values derived for $c_{1,2}$ during the coseismic phase. We will revisit this issue in detail throughout the results section.

In the interfacial representation of a fault in a homogeneous half-space, the shear traction $\tau(t,z)$ at a time t and location z along the fault may be decomposed into

$$f(t,z)\sigma = \tau(t,z) = \tau_0(t,z) + \tau_{\mathrm{d}}(t,z) - \eta_* V(t,z),$$

where $\tau_0(t,z)$ is the loading stress, $\tau_{\mathrm{d}}(t,z)$ is the convolution that expresses the dynamic stress transferred to the point z at time t by slip at all points within the causality cone around (t,z) (Rice, 1993; Cochard & Madariaga, 1994, 1996; Lapusta et al., 2000). The term $\eta_* V$ exists to compensate for the removal of a singularity in aforementioned convolution, and this sets the viscosity $\eta_* = \frac{1}{2}\frac{\mu}{c}$ equal to half the material's shear impedance, with $c_s = \sqrt{\mu/\rho}$ the shear wave speed and in turn μ the shear modulus and ρ the mass density. In the *quasi-dynamic* or *radiation damping* approximation (Rice, 1993), the dynamic stress transfer term $\tau_{\mathrm{d}}(t,z)$ is replaced by an easier to compute static stress transfer term $\tau_{\mathrm{s}}(t,z)$, but the damping term $\eta_* V$ is retained. Here we take the traction $\tau_0 + \tau_{\mathrm{s}}$ to be the one produced by our static momentum balance equation (9) and elastic constitutive equation (8), and implicitly move over the damping term to the right-hand side of the equation. Calibration of (32) to the result gives our non-dimensional damping viscosity as

$$\eta = c_3 \frac{1}{2}\frac{\mu}{\sigma}\frac{v_0}{c_s},$$

with $c_3 \sim 1$ an additional dimensionless calibration constant that is chosen somewhat arbitrarily to be equal to c_1, which seems to give maximum macroscopic slip rates that are similar to those produced by the interfacial friction law with radiation damping.

4 Numerical Models

Analytical predictions made in preceding sections are complemented here with numerical simulation results that better illustrate the complex time-dependent behavior of the system of equations, and allow us to establish the consequences of approximations and other model choices made in the process of formulating a local equivalent continuum rate and state friction law. As before we restrict the scope to 1D models across the shear band.

4.1 Equations

On the one-dimensional line $x \in [-L/2, +L/2]$ that deforms under simple shear, analogous to the transect T defined in Figure 1, we model the compact equation (33a,b) together with the one-dimensional static momentum balance equation $\partial\tau/\partial x = 0$ cf. (9), elastic constitutive equation $\dot{\tau}(t,x) = \mu\,[\partial v(t,x)/\partial x - 2\gamma_0\gamma(t,x)]$ cf. (8) (with μ the shear modulus as before), yield equality $\tau(t) = f(t)\sigma$, and finally the boundary conditions $v(t, x = \mp L/2) = \mp V_p/2$ combined into the single ODE

$$\dot{f}(t) = \frac{\mu}{\sigma}\left[\frac{V_p}{L} - \frac{\mathcal{V}(\gamma \circ \zeta)}{L}\right], \tag{33d}$$

with V_p the 'plate' driving velocity that is applied antisymmetrically on both ends of the domain. This equation uses (7) that relates slip velocity to the integral (from $-L/2$ to $+L/2$) of anelastic shear strain rate γ, as well as (33c) for the relation between γ and the auxiliary field ζ.

The system is closed with natural boundary conditions $\partial\gamma/\partial x = 0$ on $\gamma(t, x = \mp L/2)$, and initial conditions $f(t=0) = f_0 + \Delta f$ and $\gamma(t=0,x) \propto \varphi_{\mathrm{ss}}(x) + \epsilon$, with $\varphi_{\mathrm{ss}}(x)$ deriving from the analytical prediction (30a) and ϵ a necessary but small homogeneous background value of anelastic shear strain rate. As long as this value is small enough ($\epsilon \ll V_p/L \ll v_0/L$) and the 'tails' of the strain rate distribution can be ignored, the required computational domain size may be based on the length scale λ_0 and therefore cover only a potentially small fraction of the line $[-L/2, +L/2]$, greatly reducing the computational cost.

The solution procedure is outlined in Appendix D.

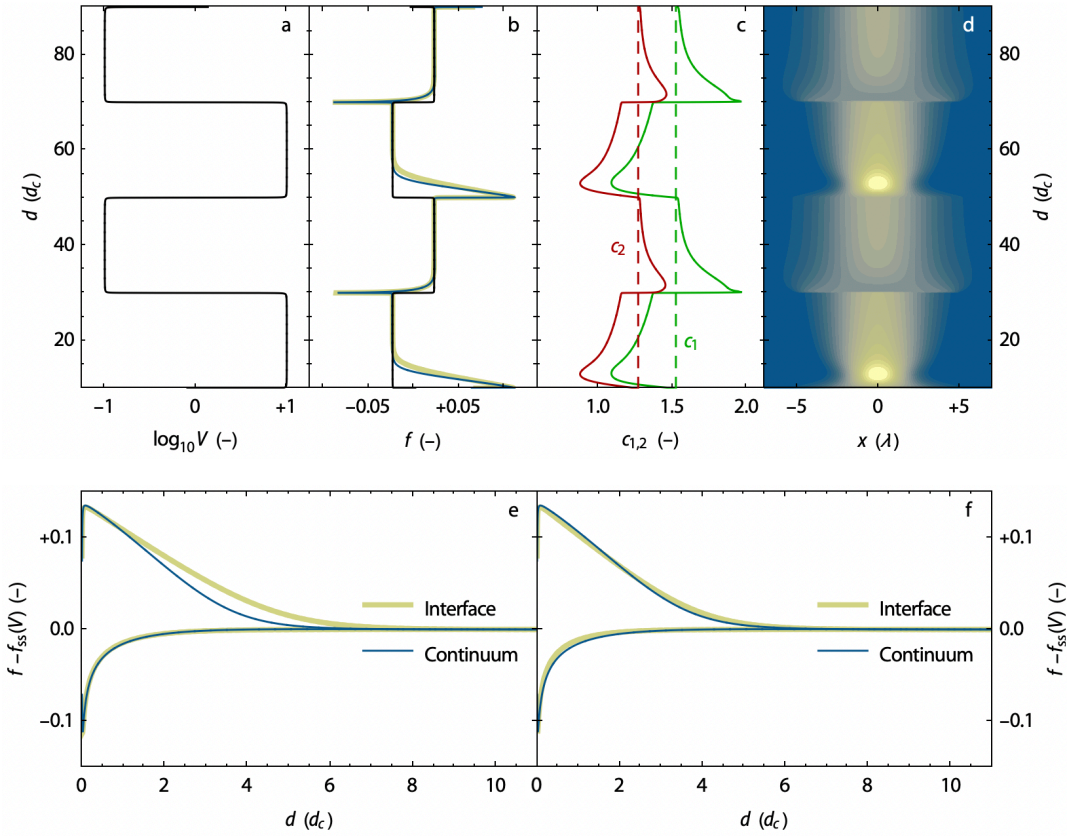

Figure 9. Results of the numerical velocity stepping experiments: a) dimensionless sliding velocity V closely tracking a smooth square wave signal (Eq. (36)); b) frictional response of the interfacial reference model (thick yellow line) and continuum model (dark blue line) with for reference the theoretical steady-state friction $f_{ss}(V)$ (black line); c) measured (solid lines; cf. Eq. (20a,b)) vs. used (dashed lines; cf. Eq. (31a,b)) values of the dimensionless coefficients $c_{1,2}$; d) distribution $\varphi = \gamma/V$ of anelastic shear strain rate with darker blue colors reflecting very low values and bright yellow colors reflecting high values, and contours distributed evenly on a linear scale. Reflecting the ultimate slip dependence (rather than direct time dependence) of the rate and state friction laws, curves in panels a)–d) are plotted against sliding distance d (measured in critical slip distances d_c) on the vertical axis. Panel e) shows an enhanced view of the evolution of the instantaneous friction coefficient towards the steady-state value with slip distance d, simultaneously showing a positive and negative step for both the interface and continuum models. Panel f) displays the same, but artificially increases the critical slip weakening distance of the interfacial friction law by 20%.

Table 1. Parameter values used in the velocity stepping experiment

parameter	value	unit
a	$2\ 10^{-2}$	1
b	$3\ 10^{-2}$	1
d_c	10^{-5}	m
v_0	10^{-6}	m s^{-1}
ρ	10^{+3}	kg m^{-3}
μ	10^{+10}	Pa
c_s	$3.162...\ 10^{+3}$	m s^{-1}
η	$1.581...\ 10^{-6}$	1
σ	10^{+6}	Pa
d_c	$5.00\ 10^{-4}$	m
λ_0	$\sim 1.4\ 10^{-4}$	m
c_1	1.531...[a]	1
c_2	1.276...[a]	1
r_0	v_0/d_c	s^{-1}
ν	10^{-3}	1
n	20	1

Some rate and state parameter values are roughly based on Erickson et al. (2020).
[a]Using (31a) and (31b) with stated values of a and b.

4.2 Continuum Velocity Stepping Friction Experiments

Velocity stepping laboratory experiments form the foundation on which rate and state friction was proposed (Dieterich, 1978). We compare numerical velocity stepping experiments of a traditional rate and state friction-governed interface with our proposed continuum equivalent under the same conditions. We impose a driving velocity that follows a smooth square wave function with respect to sliding distance d:

$$\log_{10}\left(V_p(d)/v_0\right) = \frac{2}{\pi}\tan^{-1}\left(-\frac{1}{\nu}\cos\left(\frac{\pi\, d}{n\, d_c}\right)\right), \tag{36}$$

with $n = 20$ the amount of slip weakening distances d_c between velocity steps, and $\nu = 10^{-3}$ a parameter that controls the smoothness of the smooth square wave, with small numbers giving the squarest result (see Figure 9a). Some smoothness is important to maintain numerically stable results. The low and high driving velocities are $10^{-1}v_0$ and $10^{+1}v_0$ respectively. Other model parameters are given in Table 4.2. Inertial effects may be ignored (both in the original lab experiments and in our simulations) due to the low driving velocity. Thus, the quasi-radiation damping viscosity η is set to zero and its effects not discussed here. In the limit $\eta \to 0^+$, (33c) simply reduces to $\gamma(\zeta) = \exp\zeta$ or equivalently $\zeta(\gamma) = \ln\gamma$, which reduces (33a,b) back to (18a,b).

Because the domain size L is very small and the system (33a–33d) therefore stiff, the friction tends to a stable steady state after being perturbed by a velocity step rather than developing a limit cycle. This also means that $\mathcal{V}(\gamma) \approx V_p(d)$. In fact, $\mathcal{V}(\gamma)$ and $V_p(d)$ are so close that their difference would not register in Figure 9a.

In the following, we use the terms time dependence and slip dependence interchangeably but note that the rate of change of either is not constant from the perspective of the other. A slip-centric presentation of velocity-stepping results is commonplace in the earlier cited literature.

In Figure 9b (and its detailed view 9e) we explore the similarities and differences in slip dependence of the friction coefficient f between the interfacial and continuum velocity stepping experiments. Based on (19), we expect that a velocity step perturbs not just the magnitude of anelastic shear strain, but also its distribution (Figure 9d) and therefore the theoretical value of the dimensionless coefficients $c_{1,2}$ (Section 3.3; Equations (20a) and (20b); Figure 9c, solid lines) away from steady state. Since these coefficients are actually set to their constant steady-state predictions (Section 3.5; Equations (31a)

and (31b); dashed lines in Figure 9c), we expect some differences in time or slip dependence to occur. Notably, a positive velocity step leads to a temporary localization of anelastic deformation and a corresponding decrease in the theoretical values of $c_{1,2}$ with respect to steady state. This leads to an overestimation of corresponding terms in the right-hand-side of (33a) and a faster decay towards the steady-state friction coefficient (Figure 9b,e) of the continuum model with respect to the interfacial model. A negative velocity step causes temporary delocalization (Figure 9d) and under-estimation of terms involving $c_{1,2}$ in (33a), but does not lead to any noticeable difference in friction coefficient evolution towards the steady state between the two models (Figure 9b,e). We conclude that those terms involving $c_{1,2}$ are negligible in this situation.

In Figure 9c and 9d it can also be seen that the *distribution* of anelastic shear strain rate tends more slowly to a steady state than the friction coefficient itself, and moreover that this decay is slower after a positive velocity step than after a negative one. In fact, the 40 d_c wavelength we use for the intput signal (36) is too small to enable a somewhat complete evolution towards steady-state of the anelastic strain rate distribition during the high-velocity regime, but we have taken care that this does not meaningfully impact the results during the subsequent low-velocity regime.

Most of the difference between the interfacial and continuum velocity stepping results disappears if the critical slip weakening distance d_c is seen as an additional calibration parameter. Continuing the analogy to the classical experimental setup, when we regard the continuum formulation proposed here to be the 'generator' of experimental data, and the result of the existing interfacial rate and state friction and aging laws as a data fitting curve, we observe a critical slip weakening distance d_c that is about 20% larger than the value that was used to generate the data. The result is plotted in Figure 9f. There it can be seen that the discrepancies remaining after the calibration of d_c are small compared to the expected noise level of experimental results.

4.3 Continuum Spring-slider Experiments

Whereas velocity stepping experiments are useful for studying the way in which friction can be attracted to a steady state, permanently out-of-equilibrium 'limit cycle' behavior is more interesting in the study of earthquake dynamics. We achieve this behavior by increasing the size of the domain and the distance of the boundary conditions

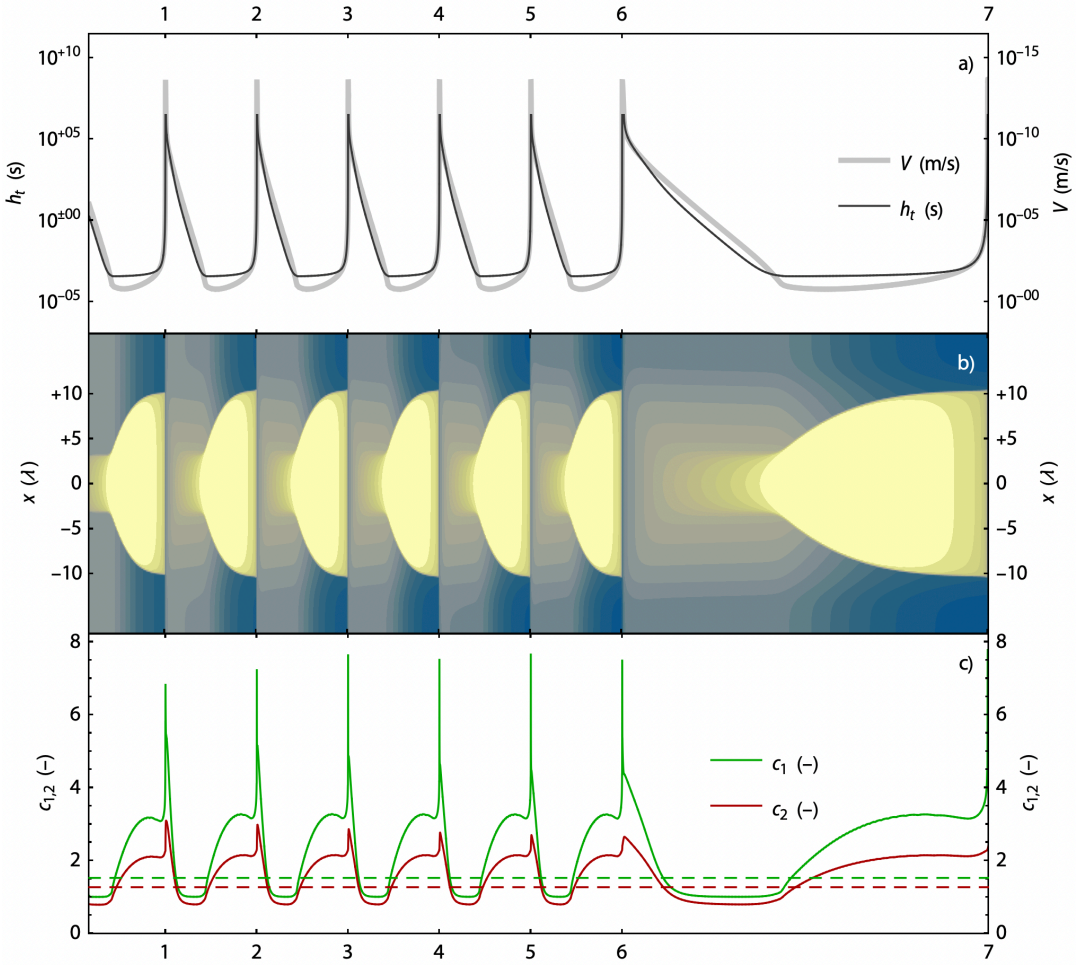

Figure 10. Evolution of the continuum seismic cycle simulation demonstrated by a succession of six events plotted against time step count on the horizontal axis, with a seventh event stretched by a factor four in order to better show some features of an individual event. The end of each event is punctuated by a slip velocity minimum and numbered 1 to 7 on the horizontal axes. Each revolution around the limit cycle takes 83.00 years and 107,500 time steps to complete. Panel a: time step size h_t (thin dark line; left axis; logarithmic scale) and slip velocity V (thick bright line, right axis; logarithmic scale). The former serves to aid the interpretation of the time-dependent progression of each event while the latter gives a sense of magnitude to the color scale of panel b. Axes are chosen so that together these curves express the imperfect reciprocal relation between time step and slip velocity. Panel b: logarithm of dimensionless anelastic shear strain rate versus signed distance x measured in number of length scales λ_0 away from the shear zone center. Bright yellow colors indicate high strain rate; dark blue colors indicate low strain rate. Contours are drawn at equal intervals in log space. No explicit color scale is given because the magnitude of anelastic shear strain rate γ is tied to the value of λ_0 in order to produce a length-scale independent history of sliding velocity V (panel a). Panel c: measured (solid lines; cf. Eq. (20a,b)) vs. used (dashed lines; cf. Eq. (31a,b)) values of the dimensionless parameters $c_{1,2}$.

by many orders of magnitude ($L = 80$ km) with respect to the velocity stepping experiments. Thereby, we greatly reduce the effective stiffness of the medium as can be seen in (33d). This is analogous to performing the so-called spring-slider experiment in which a mass is dragged over a surface by a spring that is tensed at a constant rate to give an educational example of the apparent stick-slip behavior of seismogenic faults. Our goal in performing this experiment is to determine the degree to which the here proposed continuum friction laws match the predictions made by Dieterich and Ruina's interfacial laws and to describe the transient behavior of anelastic strain rate in the added dimension. Damping plays an essential role in these experiments to close the limit cycle that otherwise extends to $V \to \infty$ (Section 3.6). As discussed in the same section, our implementation of damping as a linear bulk viscosity precludes a steady-state anelastic strain rate distribution at high slip velocity, a fact that directly contradicts one of our precepts (Section 3.1). Therefore we anticipate a larger discrepancy between interface and continuum models here than was observed during the velocity stepping experiments.

Figure 10 depicts the evolution of quantities of interest over the course of multiple orbits of the solution as it converges to the limit cycle. The first important observation is that the continuum model still appears to exhibit spontaneous limit cycle behavior. Due to the appropriately chosen initial condition, the wind-up period is short and there is little variability between successive revolutions. As predicted (Section 3.5), the nucleation phase is marked by runaway amplitude increase of a quasi-steady strain rate distribution (panel b). As before, we measure the dynamic value of the coefficients $c_{1,2}$ from the strain rate distribution following (20a,b), and compare to the analytically derived steady-state values (31a,b). In line with analytical predictions this phase is also accompanied by measured values of $c_{1,2}$ (panel c) that are slightly lower than the constant values that are prescribed (dashed lines in panel c). The opposite situation occurs throughout the remainder of each cycle where anelastic strain is more widely distributed. We can attribute this directly to the effect of damping at high strain rate, which exerts a delocalizing influence. Figure 11 provides a more intuitive view of the short-lived delocalizing behavior of a strain pulse using linear rather than logarithmic scales. We emphasize that the ultimate extent of coseismic delocalization compared to the length scale λ_0 is dependent on model and material parameters, among which the effective stiffness of the medium.

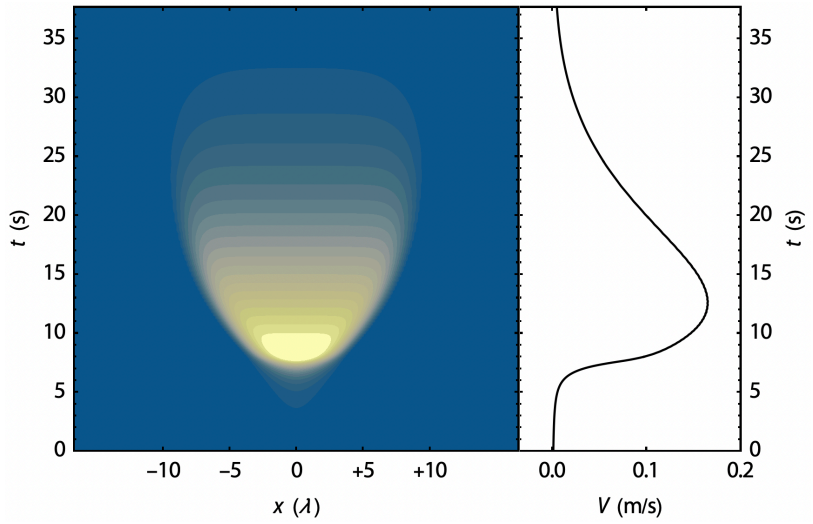

Figure 11. Focus on the last event of Figure 10, here visualized in panel a by strain rate on a linear color and contour scale and in panel b by slip rate V, both versus a linear time scale in seconds.

Large peaks in the measured values of $c_{1,2}$ shown in panel c of Figure 10 occur during the post- and interseismic phase, where anelastic shear strain rate is broadly and quite uniformly distributed. The values that are actually used (dashed lines) are way too small here, however, terms that do not involve these coefficients dominate the time-dependent behavior in this regime and the discrepancy's consequences are limited. This is largely confirmed in Figure 12, which tabulates the evolving scalar properties of a limit cycle of the interfacial and continuum models of rate and state friction. In particular, the duration of the limit cycle (first column, panels a–c, time in years on the horizontal axis) is very similar even though the timing of the event in relation to its enclosing slip velocity minima is somewhat different.

Markedly different are the coseismic aspects of the limit cycle. Panel e of the middle column and panels h–j of the right column of Figure 12 clearly show this. The continuum model has a faster stress drop (panel h, $U' \propto f$) and a more symmetric slip rate response with time compared to the interfacial model. The limit cycle depicted in linear phase space (U', V) in Figure 12e shows a skewed triangular trajectory of the interfacial model and a more parabolic trajectory for the continuum model. It appears that the amount of stress drop that happens *before* peak slip rate in the interfacial model is small – about 10% of the total stress drop. At the same point in the continuum model

the stress drop is already about one third of the total. The same limit cycle shown in Figure 12f in log-linear phase space $(U', log_{10}V)$ allows us to see that there are subtle differences between the limit cycles of the interfacial and continuum models over the whole range of slip velocities.

A further result of practical interest is the difference in adaptively chosen time step between the interfacial and continuum models. Figure 10a shows an approximate inverse relation between the time step of continuum model to the aggregate slip velocity $v_0 V$, which is unsurprising seeing how $d_c/[v_0 V(t)]$ gives a sensible local time scale for both interfacial and continuum rate and state friction models. The approximate inverse relation breaks down at large time scales where the corresponding time steps are truncated to a fixed value of 0.1 yr, and at large slip velocity and small time step, where the time scale of the continuum model becomes dominated by large spatial gradients and is generally much smaller than that of the interfacial model (Figures 12d,g,k). From these same plots it can be seen that, just like stress begins to drop in significant advance of an event, so too does the time step of the continuum model.

Finally, we remark that spatial resolution tests indicate that the quality of the solution is only influenced by the ratio of cell size h_x to λ_0, with acceptable results achieved when $h_x/\lambda_0 \geq 10$. For the results presented here we have used $h_x/\lambda_0 = 20$. Changing λ_0 in proportion to h_x does nothing except to cause a wider or narrower but equally well resolved strain distribution and a virtually identical limit cycle.

5 Discussion

So far we have argued for a visco-elastic continuum rheology that resembles both the mathematical presentation as well as the resulting behavior of traditional interfacial rate and state friction as described by Dieterich and Ruina in their respective seminal puplications. Before addressing this work in the context of a much broader body of existing research and reflecting on critical assumptions made in this work and its extension towards the future, we briefly touch upon some alternate branches of our proposed theory that have been left out so far for the sake of clarity.

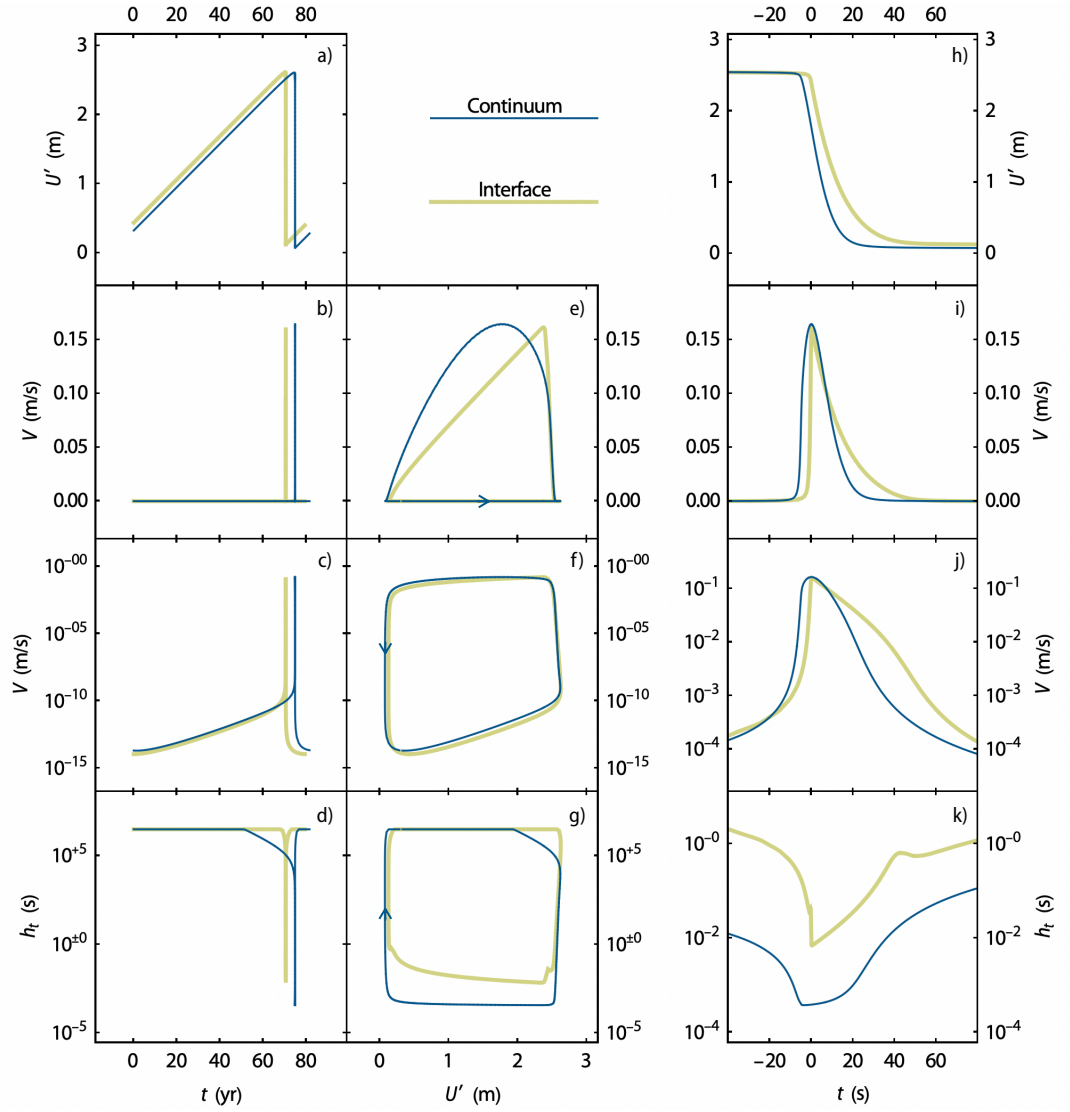

Figure 12. Panels a–g: behavior of the interfacial (bright yellow lines) and continuum (dark blue lines) models for a full limit cycle. Panels h–k: focus on the 100 seconds around an event. From left to right, horizontal axes are: time t in years, slip deficit U' in meters (linearly related to shear stress τ and friction coefficient f), and time t in seconds. Vertical axes from top to bottom are slip deficit U' in meters, slip velocity V in meters per second on a linear scale, the same on a logarithmic scale, and finally the time step size h_t in seconds on a logarithmic scale. Where limit cycles are shown (middle column, panels e–g), the cycle sense is indicated with small arrows.

5.1 A Primitive Reformulation

In contrast to traditional models of plasticity, the continuum rate and state friction laws presented here possess a continuous time dependence that makes the anelastic shear strain rate γ a predictable quantity rather than one that is to be solved by a constrained optimization algorithm (e.g. Simo & Taylor, 1985; Duretz et al., 2018). This is apparent from the ODE (18a,b) for γ, but is hidden in the Dieterich-Ruina form (15) and (16) of the continuum rate and state friction rheology. However, (18a,b) lacks the simplicity and elegance of the Dieterich-Ruina system and is difficult to interpret as the resultant of physical phenomena. By defining a set of primitive variables κ and ψ as

$$\begin{aligned} \dot{\kappa} &= r_\kappa \gamma \\ \dot{\psi} &= r_\psi \gamma^{a/b} \exp\left(-[f(\boldsymbol{\sigma}) - f_0]/b\right), \end{aligned} \tag{37}$$

with the respective reference rates $r_\kappa = r_0 c_1$ and $r_\psi = r_0 c_2$, we are able to write a more expressive formulation of (18a,b) as

$$\dot{\kappa} = r_\kappa \gamma = r_\kappa \left(\dot{\psi}/r_\psi\right)^{\frac{b}{a}} \exp\left([f(\boldsymbol{\sigma}) - f_0]/a\right) \tag{38}$$

$$\dot{\psi} = r_\psi \exp\left(\mathcal{M}_{\text{ex}}(\kappa) - \psi\right), \tag{39}$$

that still closes the momentum balance equation (9) and elasto-plastic constitutive equation (8) without algebraic constraints and thus without requiring a constrained optimization algorithm. Equation (39) can optionally accommodate damping using the Lambert W function analogous to the way it was discussed in Section 3.6.

We recognize $\kappa(t, x)$ as a dimensionless measure of accumulated anelastic shear strain and therefore assign to ψ the meaning of a 'virtual' or 'unrealized' strain. In spite of the conceptual appeal of (38) and (39), it can be difficult to choose appropriate initial conditions for ψ.

5.2 Quadratic Diffusion

Instead of the mollified term $\gamma\,\mathcal{M}(\gamma)$ that was introduced in (16) and used throughout the remainder of this work, we could have chosen to use $\mathcal{M}(\gamma^2)$:

$$a\,\dot{\gamma} = b\,r_0\left[c_1 \mathcal{M}(\gamma^2) - c_2 \gamma\,\theta(\gamma, f)^{-1}\right] + \gamma \dot{f}. \tag{40}$$

and the definition of c_1 is adapted to the new way of mollifying according to the same principles as laid out in Section 3.

This choice has advantages and disadvantages. A major disadvantage is that it does not seem to allow a return from (40) to the more elegant Dieterich-Ruina form (15) and (16), nor to the compact form (38)–(39) proposed in the preceding section. A clear advantage is that, taken together with the explicit non-local operator $\mathcal{M}_{\text{ex}}$, (40) produces a *degenerate* reaction-diffusion equation of generalized Fisher or KolmogorovPetrovskyPiskunov (KPP) type (Fisher, 1937; Kolmogorov et al., 1937) that remains well-posed at and around $\gamma = 0$. This leads to shear zone solutions that can propagate into perfectly intact rock without taking recourse to an arbitrary small initial condition on γ. Furthermore, under quadratic diffusion, the function $\gamma(\zeta)$ used in the ODE form (33a–d) in Section 3.6 becomes algebraic instead of transcendentally implicit.

We have reproduced all results presented in Section 4 with this quadratic diffusion term and zero background anelastic strain rate too, and observe numerically smoother and better resolved shear fronts during the coseismic delocalization phase. These fronts also travel faster and further than in the linear diffusion case leading to a larger discrepancy between measured and used values of the coefficients $c_{1,2}$ and consequently larger deviations of the limit cycle with respect to that of the interfacial rate and state friction model. The generalized Fisher–KPP type equation has merited a large body of theoretical research (e.g. Sánchez-Garduño & Maini, 1994, 1995, 1997; Sherratt & Marchant, 1996; Petrovskii & Li, 2003, 2006; Gilding & Kersner, 2005; Y. Wu et al., 2006; Y. Li & Wu, 2008; Malaguti & Ruggerini, 2010; Broadbridge & Bradshaw-Hajek, 2016) and therefore some of the above mentioned empirical findings may be given a theoretical underpinning in future work.

5.3 Relation to Regularized Damage or Plasticity Models

As shown in Section 4.3 there are circumstances under which the transient viscous rheology proposed in this work promotes a spontaneous organization of periods and regions of negligible anelastic strain rate and those of significant anelastic strain rate. Even though a critical yield stress seems to be an emergent rather than an inherent property of the system of equations and initial and boundary conditions, it is tempting to think of this rheology as a smooth plasticity model.

We have shown (Section 3.5) that our model requires regularization by a non-local strain rate measure to avoid spurious mesh dependence. Non-locality in one form or an-

other has been applied to combat ill-posedness in e.g. Bažant et al. (1984); Triantafyllidis and Aifantis (1986); Schreyer and Chen (1986); Pijaudier-Cabot and Bažant (1987); de Borst and Mhlhaus (1992); Peerlings et al. (1996); Jirásek (1998); Bažant and Jirásek (2002); Engelen et al. (2003); Jirásek and Rolshoven (2009b, 2009a); Burghardt et al. (2012); Lyakhovsky et al. (2011); Lyakhovsky and Ben-Zion (2014b, 2014a); Lyakhovsky et al. (2016); Kurzon et al. (2019, 2020); Kiefer et al. (2018); Abdallah et al. (2020).

The comprehensive non-local damage-breakage rheology of Lyakhovsky et al. (2011); Lyakhovsky and Ben-Zion (2014b, 2014a); Lyakhovsky et al. (2016), firmly rooted in thermodynamic theory and well-calibrated to match observations, is one of few damage theories developed to cover the full process of earthquake generation and healing (Lyakhovsky et al., 2016; Kurzon et al., 2019, 2020). As in our model, the nonlocality in theirs is of Gradient type, but differs subtly in that it acts on the damage parameter rather than on the anelastic shear strain rate. Another example is the Godunov-Peshkov-Romenski model (S. Godunov & Romenskii, 1972; S. K. Godunov & Romenskii, 2003; Resnyansky et al., 2003; Romenskii, 2007; Romenski et al., 2020), which was first used in Gabriel et al. (2021) to simulate dynamic rupture and off-fault damage generation. This model differs from our model and the aforementioned damage-breakage rheology in a fundamental way in that shear bands are produced by damage waves described by hyperbolic equations rather than parabolic equations of reaction-diffusion type.

Our model differs from phase-field models of fracture in some respects. We have used the distribution φ of anelastic shear strain rate as an analytical tool during the derivation of the continuum rheology proposed in this work. While this distribution could be renormalized to form a phase field $d \in (0, 1)$, the necessary inclusion of a damping viscosity in Section 3.6 causes temporal variations in the anelastic shear strain rate distribution that transcend the modeled distributions, and make renormalization to a phase field possible at best as a post-processing step in a simulation. We recognize that the inclusion of a damping viscosity constitutes the use of a 'double-well potential' in phase field terminology, which multiple authors have commented on critically in this context (e.g. Kuhn et al., 2015; J.-Y. Wu, 2017).

Our linear viscous damping appears as a regularization technique in some plasticity models (Needleman, 1988; Peirce et al., 1983; F. Wu & Freund, 1984; Duretz et al., 2019, 2020, 2021; de Borst & Duretz, 2020; Stathas & Stefanou, 2022). In these mod-

els the Kelvin–Voigt arrangement of yield strength and Newtonian viscosity truncates the steady-state anelastic strain rate that may be achieved, forcing a shear zone to have a finite width in order to slide at a certain macroscopic rate. In our model, the same viscosity also introduces an effective upper bound on anelastic strain rate attained during the limit cycle and causes subsequent delocalization, but we must still rely on the gradient regularization to combat the unbounded localization that would otherwise happen even at low strain rate. A detailed analysis of the efficacy of this viscous regularization in dynamical problems is made in Stathas and Stefanou (2022).

5.4 Relation to Other Transient Continuum Rheologies

The framework here proposed can be seen as a generalization of the work of Herrendörfer et al. (2018); Preuss et al. (2019, 2020), who made the purely local substitution $v_0 V \rightarrow h_x \gamma_0 \gamma$ and used a Drucker-Prager elastoplastic model similar to the one set out in Section 2.4. Setting $\lambda_0, \eta = 0$ and assuming that strain rate fully localizes into a discrete Dirac function sampled every h_x, we find that the coefficients c_1 and c_2 become h_x/d_c and 1 respectively, substitution of which into (16) yields Herrendörfer's version of the aging law. Their model was first applied along a predefined staggered grid line in Herrendörfer et al. (2018), obeying the discrete Dirac distribution of anelastic shear strain automatically and effectively yielding a numerical method analogous to the stress glut method of Andrews (1999). Herrendörfer's model was subsequently applied in an unconstrained evolving continuum model in Preuss et al. (2019, 2020), but notwithstanding measures put in place that acknowledge the changing distribution of shear strain rate within a shear zone, their model ultimately lacks regularizations that remove mesh dependence. In the chapters titled 'Localization of Deformation' and 'Relationship of Localization to Instability' of his PhD thesis, Ruina (1980) gives a thoughtful take on aspects of the localization behavior of a strain rate formulation of rate and state friction (without spatial regularization), which is in some aspects in line with findings reported in this work, and complementary in others.

As noted in Section 1.2 of the Introduction, a variety of continuum theories have been developed to explain the general rate- and state-like behavior of deformation in faults and shear zones (e.g. Sleep, 1997; Braeck & Podladchikov, 2007; Daub et al., 2008; Daub & Carlson, 2008, 2009; Daub et al., 2010; Elbanna & Carlson, 2014; John et al., 2009; Thielmann et al., 2015; Thielmann, 2018; Rozel et al., 2011; Barbot, 2019; Niemeijer &

Spiers, 2007; Chen & Spiers, 2016; Van den Ende et al., 2018; Roubek, 2014; Lyakhovsky et al., 2011; Lyakhovsky & Ben-Zion, 2014b, 2014a; Lyakhovsky et al., 2016; Kurzon et al., 2019, 2020; Poulet et al., 2014; Veveakis et al., 2014; Alevizos et al., 2014; Rattez, Stefanou, & Sulem, 2018; Rattez, Stefanou, Sulem, Veveakis, & Poulet, 2018; Pozzi et al., 2021). We note again that an internal length scale in these models typically arises from the inclusion of a diffusion process (of e.g. temperature, pore pressure) but that may not always be adjusted to meet the constraints imposed by scale and computational power without changing the outcome of the model. In our proposed formulation an artificial diffusion process acts directly on the anelastic shear strain rate, resulting in a robust and controllable internal length scale.

In this last respect our work is preceded by nearly 25 years by Sleep (1997). In this work Sleep combined and extended earlier works (Sleep, 1995; Segall & Rice, 1995; Linker & Dieterich, 1992; Chester, 1994, 1995) in which rate and state friction was interpreted as the product of crack generation and healing, associated rheological weakening, and dissipative heating. This physical reasoning resulted in a model that contains only quantities that are either directly measurable or can be modeled by independent methods. This contrasts with our purely mathematical argumentation that serves to retain close correspondence to the original phenomenological description of rate and state dependent sliding on a frictional interface. Sleep (1997) neglects the fluxes and associated spatial gradients of the pore fluid but does include heat diffusion, which they note does however not play a significant role at the scale of their numerical experiments. Instead, resembling our approach, they impose an artificial length scale and forced strain distribution by explicit mollification of the anelastic shear strain rate with a Gaussian kernel. This leads to an aging law (Equation 53 in Sleep (1997)) that is structurally identical to our result (16). Like us, they find that strain localization can only occur when $a < b$, and that a rate-strengthening effect that activates at high strain rate leads to strain delocalization. With respect to Sleep's valuable contribution, in this work we provide a more complete argumentation for this type of spatial regularization and analysis of the resulting patterns of strain localization and delocalization over the seismic cycle.

5.5 Assumptions and Future Work

We proposed our model in a very general three-dimensional continuum mechanics framework, but for simplicity have considered only a small fault neighborhood in which

in-plane variations of fault properties can be neglected (Section 2.2) so that the model becomes effectively one-dimensional. This same assumption was also taken in the numerical models that we have used in our analysis. The assumption is clearly violated around fault branches, at fault kinks or on rough faults, and near the fault or rupture tip. We note however that rate and state friction was proposed based on laboratory studies that also neglect these geometrical complexities. Although rate and state faults with branches and kinks are still largely non-standard in present-day numerical modeling studies, plenty of attention has been given to the critical nucleation patch and the structure of the cohesive zone near the rupture tip (e.g. Rice, 1993; Cocco & Bizzarri, 2002; Cocco et al., 2004; Rubin & Ampuero, 2005; Day et al., 2005; Lapusta & Liu, 2009; Viesca, 2016b, 2016a; Putelat et al., 2017). Given our main assumption, these features can only be accurately reproduced with our continuum formulation in higher-dimensional numerical models if the regularization length scale λ_0 is significantly smaller than the length scales associated with the critical nucleation patch and the cohesive zone. In turn the cell size must be sufficient to resolve λ_0, and so we expect to need a grid resolution that is significantly higher than that of existing methods to simulate rate and state frictional interfaces. We have also seen that the diffusion process is associated with a smaller time scale, and thus, stricter time step constraints than the interfacial model. We note that both spatial and temporal resolution requirements already place challenging constraints on simulations of seismic and aseismic slip sequences (Erickson et al., 2020; Jiang et al., 2022).

It seems prudent to first make a more detailed assessment of the computational demands and the techniques that may be required to meet the resolution requirements (e.g. adaptive mesh refinement, local adaptive time stepping). In the process of constructing higher-dimensional models one may first concentrate on the friendliest regions of parameter space, for example those that promote stable sliding, or undamped runaway localization as in Viesca (2016b, 2016a, 2020).

The temporal patterns of localization and delocalization that occur in our model yield testable predictions that tie in to a recent surge in interest in similar patterns observed in the lab and in nature (e.g. Ben-Zion & Zaliapin, 2020; McBeck et al., 2018; McBeck, Aiken, et al., 2020; McBeck, Ben-Zion, & Renard, 2020; McBeck et al., 2021). It could be the scope of future research to reinterpret anelastic strain rate in our model as a measure of the activity of a statistical distribution of cracks of various properties and com-

pare to aforementioned lab and field observations. In this context, the delocalization that is in our models induced by a rate-limiting Kelvin viscosity is reminiscent of the growth of off-fault fracture networks during dynamic rupture (e.g. Templeton & Rice, 2008; K. Okubo et al., 2019; Gabriel et al., 2021).

We have assumed an infinite gauge layer and observed that materials that are rate-strengthening at steady state ($a > b$) can only feature delocalization of strain. This appears inconsistent with the widespread observation of faults and fault gauges that are more or less stable (e.g. Ikari et al., 2011; Carpenter et al., 2012; Coble et al., 2014; Bedford et al., 2021). A possible explanation is that a gauge layer is initially formed by a set of strain-softening processes, subsequently chemically, petrologically, texturally or geometrically matured over time to become rate-strengthening, but generally weaker than host rock. The gauge layer walls may then act as a barrier to further delocalization. While this situation could be simulated in our approach by varying the frictional properties accordingly, the consistent way to model the long term evolution of faults likely requires at least two internal state variables. The damage-breakage rheology of Lyakhovsky and Ben-Zion (2014b, 2014a) has this feature.

In this work, we have restricted ourselves to classical Dieterich-Ruina rate and state friction with aging law. In Section 1.2, we have hinted at a sizable number of physical interpretations of the phenomenological rate and state friction problem. This concerns notably the behavior of weakening and the interpretation given to the 'state' variable. At the same time, a number of alternative phenomenological evolution laws have been proposed, among which the slip law (Dieterich, 1979a). It would be worthwhile to investigate the possibility of applying the mathematical techniques developed here to a wider range of friction and evolution laws, especially those involving multiple states, such as for example temperature, pore pressure, or grain size. However, this might not be straightforward because we have relied heavily on analytical solutions to calibrate the interfacial and continuum descriptions of friction, and finding these solutions is frequently a time-consuming task with uncertain prospects for success.

6 Conclusions

In this work we have carefully constructed a coordinate-invariant and mesh-independent transient visco-elastic continuum rheology that behaves in a way that is consistent with

rate and state friction on an interface. We have shown that inclusion of a diffusion-like spatial regularization ensures a limit to strain localization and thus guarantees mesh convergence. In a simplified 1D fault transect, important metrics of the seismic cycle – such as slip rate and friction – are independent of the diffusion length scale associated with the regularization. However, throughout this work we have assumed the regularization length scale to be small compared to the length scales associated with other features of interest, such as fault curvature or along-strike variations of slip rate and stress. Therefore, going forward, high-resolution 2D or 3D numerical models are required to apply this model to the study of the seismogenic behavior of emerging and evolving fault zone networks. Our continuum rheology resembles a reaction-diffusion equation for anelastic strain rate. Processes described by such equations are ubiquitous in nature, and it is tempting to compare temporal patterns of localization and delocalization produced by our model with natural observations.

Acknowledgments

We are grateful to Yehuda Ben-Zion for his recommendation to look into damage diffusion that ultimately spurred us to pursue this work. We also gratefully acknowledge fruitful discussions with Robert Herrendörfer, Robert Viesca, Vladimir Lyakhovsky, Dmitry Garagash, Sylvain Barbot, and René de Borst. We greatly appreciate the thoughtful reviews of Norman Sleep and two anonymous reviewers, which encouraged us to improve the work. This research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under the ERC StG TEAR, grant no. 852992. CP acknowledges additional financial support by Swiss National Science Foundation grant 200021-169880. PS acknowledges financial support from the Swiss University Conference and the Swiss Council of Federal Institutes of Technology through the Platform for Advanced Scientific Computing (PASC) program. A.-A.G. acknowledges additional support by the European Union's Horizon 2020 Research and Innovation Programme under ChEESE, grant no. 823844 and the German Research Foundation (DFG projects grants no. GA 2465/2-1, GA 2465/3-1). None of the authors declare any competing interests or financial conflicts. All scripts used to generate the figures and their underlying data are available in the supplementary materials. We depend on the MIT-licenced 'Scientific Colour Maps' package (Crameri, 2021) for distortion-free representation of the model results – also for readers with color vision deficiencies (Crameri et al., 2020).

Data Availability Statement

All codes and algorithms to generate and visualize the results discussed in this work can be found as Wolfram Mathematica (Wolfram Research, Inc., 2017) notebooks in the supplementary information.

Appendix A Strain Localization in the Runaway Slip Regime

In this appendix we derive a generic family of solutions to the runaway slip regime discussed in Section 3.2. We consider the ODE

$$\dot{\gamma}(t,x) = \gamma(t,x)^2, \tag{A1}$$

in which t and γ are taken to be non-dimensionalized in such a way that the rate constant equals one and is thus dropped. We emphasize that since no derivatives with respect to x are included, the ultimate collective behavior of $\gamma(t,x)$ is not uniquely defined. Even so, our Ansatz is a product-wise decomposition of the solution into a purely time-dependent term Γ and a self-similar contribution g:

$$\gamma(t,x) =: \Gamma(t) g\left(\frac{x}{l(t)}\right), \qquad g(0) = 1. \tag{A2}$$

We then take the time derivative of the Ansatz (A2) and equate the result to (A1), yielding

$$\dot{\gamma}(t,x) = \gamma(t,x)^2 = \dot{\Gamma}(t) g\left(\frac{x}{l(t)}\right) - \Gamma(t)\frac{\dot{l}(t)}{l(t)}\frac{x}{l(t)} g'\left(\frac{x}{l(t)}\right) = \Gamma(t)^2 g\left(\frac{x}{l(t)}\right)^2. \tag{A3}$$

At $x = 0$, we have:

$$\dot{\Gamma}(t) = \Gamma(t)^2. \tag{A4}$$

The solution to (A4) is

$$\Gamma(t) = \Gamma_0 \left(1 - \Gamma_0 t\right)^{-1}, \qquad \Gamma_0 = \Gamma(t=0). \tag{A5}$$

When we substitute this solution into (A3), we get

$$\frac{\dot{l}(t)}{l(t)}\frac{x}{l(t)} g'\left(\frac{x}{l(t)}\right) = -\Gamma(t)\left[g\left(\frac{x}{l(t)}\right)^2 - g\left(\frac{x}{l(t)}\right)\right]. \tag{A6}$$

If we now take the additional Ansatz that

$$\frac{x}{l(t)} g'\left(\frac{x}{l(t)}\right) = m\left[g\left(\frac{x}{l(t)}\right)^2 - g\left(\frac{x}{l(t)}\right)\right], \qquad m \in \mathbb{R} \geq 2 \tag{A7}$$

then we obtain

$$g(t,x) = \left(1 + 2\left|\frac{x}{l(t)}\right|^m\right)^{-1}, \tag{A8}$$

and

$$l(t) = l_0\,(1 - l_0 t)^{\frac{1}{m}}\,, \qquad l_0 = l(t=0), \tag{A9}$$

The function $g(t, x/l(t))$ integrates to

$$\Lambda(t) := \int_{-\infty}^{+\infty} g\left(t, \frac{\xi}{l(t)}\right) \mathrm{d}\xi = l(t)\frac{\pi}{m}\csc\frac{\pi}{m}, \tag{A10}$$

and thus the *distribution* $\varphi(t,x)$ corresponding to $g(t, x/l(t))$ is given by

$$\varphi(t,x) = \Lambda(t)^{-1}\left(1 + 2\left|\frac{x}{l(t)}\right|^m\right)^{-1}, \tag{A11}$$

Finally, we use (7) and (A11) to write (A5), (A8), and (A9) in terms of $V(t)$, and obtain:

$$V(t) \propto \left(1 - \frac{t}{t_*}\right)^{m^{-1}-1},$$
$$l(t) \propto V(t)^{(1-m)^{-1}}. \tag{A12a}$$

Appendix B Derivation of the Steady-state Strain Rate Distribution

Here we derive the steady-state solution presented in Section 3.4 to the time-dependent PDE

$$\dot{\varphi} \propto \lambda_0\, c_1\, \varphi\, \mathcal{M}_{\mathrm{ex}}(\varphi) - \lambda_0^{\frac{a}{b}}\, c_2\, \varphi^{1+\frac{a}{b}}, \tag{B1}$$

subject to $a, b > 0$ and $a < b$, with $\mathcal{M}_{\mathrm{ex}}$ given by (26) and the coefficients c_1 and c_2 measured from the sought solution $\varphi(\infty, x)$ by means of (20a) and (20b). We summarize the PDE that is to be solved as

$$\varphi^2 + \varphi\lambda_0^2\partial_x^2\varphi - c\,\varphi^{1+\frac{a}{b}} = 0, \tag{B2}$$

with c a parameter to be determined later. Dividing the whole equation by φ, and making the substitution $\lambda_0 \to \left[1 - \frac{a}{b}\right]\lambda_1$, (B2) becomes

$$\varphi + \left[1 - \frac{a}{b}\right]^2 \lambda_1^2\partial_x^2\varphi - \xi\,\varphi^{\frac{a}{b}} = 0. \tag{B3}$$

We now introduce $f(\varphi) = \left[1 - \frac{a}{b}\right]\lambda_1\partial_x\varphi(x)$, which allows us to write

$$\left[1 - \frac{a}{b}\right]^2 \lambda_1^2\partial_x^2\varphi = \left[1 - \frac{a}{b}\right]\lambda_1\partial_x f(\varphi) = \left[1 - \frac{a}{b}\right]\lambda_1 f'(\varphi)\partial_x\varphi = f'(\varphi)f(\varphi). \tag{B4}$$

Applying this reasoning to (B3) and reorganizing slightly gives

$$f'(\varphi)f(\varphi) = c\,\varphi^{\frac{a}{b}} - \varphi. \tag{B5}$$

We now integrate (B5) by parts as follows:

$$\int_0^{\varphi} f'(\varepsilon)f(\varepsilon)\mathrm{d}\varepsilon = \tfrac{1}{2}\left(f(\varphi)^2 - f(0)^2\right) = c\int_0^{\varphi}\varepsilon^{\frac{a}{b}}\mathrm{d}\varepsilon - \int_0^{\varphi}\varepsilon\,\mathrm{d}\varepsilon = c\,\left[1+\tfrac{a}{b}\right]^{-1}\varphi^{1+\frac{a}{b}} - 2^{-1}\varphi^2. \tag{B6}$$

We arbitrarily take the distribution $\varphi(\infty, x)$ to be symmetric around $x = 0$, where it reaches its maximum value $\varphi_0 = \varphi(\infty, 0)$. At this point, as well as at the extremities of the distribution where $\varphi = 0$, we set the gradient, i.e. $f(\varphi)$, to zero. This eliminates the term $f(0)^2$ in (B6) and constrains the free parameter c to

$$c = \tfrac{1}{2}\left[1+\tfrac{a}{b}\right]\varphi_0^{1-\frac{a}{b}}. \tag{B7}$$

We now multiply both sides of (B6) by 2 and take the square root to obtain

$$f(\varphi) = \sqrt{\varphi_0^{1-\frac{a}{b}}\varphi^{1+\frac{a}{b}} - \varphi^2}. \tag{B8}$$

Applying the inverse function theorem to $f(\varphi)^{-1} = \left(\left[1-\frac{a}{b}\right]\lambda_1^{-1}\varphi'(x)\right)^{-1} = \left(\left[1-\frac{a}{b}\right]\lambda_1\right)^{-1}x'(\varphi)$, we obtain from (B8)

$$\frac{x'(\varphi)}{\lambda_1} = \left[1-\tfrac{a}{b}\right]\left(\varphi_0^{1-\frac{a}{b}}\varphi^{1+\frac{a}{b}} - \varphi^2\right)^{-\frac{1}{2}}. \tag{B9}$$

We seek to integrate this relation once more over the region $\varepsilon \in [0, \varphi)$ to obtain a solution for $x(\varphi)$. The integral of the right hand side of (B8) can be reverse-engineerded from the known derivatives

$$\cos^{-1\prime}(f) = -(1-f^2)^{-\frac{1}{2}}.$$

and $\mathrm{d}_f f^\alpha = \alpha f^{\alpha-1}$, noting also that $\varphi/\varphi_0 \leq 1$, to give

$$\left[1-\tfrac{a}{b}\right]\left(\varphi_0^{1-\frac{a}{b}}\varphi^{1+\frac{a}{b}} - \varphi^2\right)^{-\frac{1}{2}} = \partial_\varphi\left[-2\cos^{-1}\left(\left[\frac{\varphi}{\varphi_0}\right]^{\frac{1}{2}\left[1-\frac{a}{b}\right]}\right)\right]. \tag{B10}$$

This yields

$$\frac{x(\varphi)}{\lambda_1} - \frac{x(0)}{\lambda_1} = -2\cos^{-1}\left(\left[\frac{\varphi}{\varphi_0}\right]^{\frac{1}{2}\left[1-\frac{a}{b}\right]}\right) + \pi. \tag{B11}$$

Setting $\frac{x(0)}{\lambda_1} = -\pi$ and inverting for $\varphi(\infty, x)$ gives

$$\varphi(\infty, x) = \varphi_0\cos\left(\frac{1}{2}\frac{x}{\lambda_1}\right)^{2\left[1-\frac{a}{b}\right]^{-1}}. \tag{B12}$$

It is possible and in fact desirable to isolate a single strain pulse in the range $x \in [-\pi\lambda_1, +\pi\lambda_1]$:

$$\varphi(\infty, x) = \begin{cases} \varphi_0 \cos\left(\frac{1}{2}\frac{x}{\lambda_1}\right)^{2\left[1-\frac{a}{b}\right]^{-1}} & \forall\, x \in [-\pi\lambda_1, +\pi\lambda_1] \\ 0 & \forall\, x \notin [-\pi\lambda_1, +\pi\lambda_1]. \end{cases} \tag{B13}$$

Finally, we determine the coefficient φ_0 that ensures that the distribution φ integrates to one over its domain, and the coefficients c_1 and c_2 following the expressions (20a) and (20b). Definite integrals of (B13) are evaluated using Wolfram Mathematica (Wolfram Research, Inc., 2017), and we find

$$\varphi_0 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(1 + \left[1-\frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2} + \left[1-\frac{a}{b}\right]^{-1}\right)} \lambda_1^{-1} \tag{B14a}$$

$$c_1 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(1 + 2\left[1-\frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2} + 2\left[1-\frac{a}{b}\right]^{-1}\right)} \left[\lambda_1\varphi_0\right]^{-2} \tag{B14b}$$

$$c_2 = \frac{1}{2}\pi^{-\frac{1}{2}} \frac{\Gamma\left(2\left[1-\frac{a}{b}\right]^{-1}\right)}{\Gamma\left(\frac{1}{2}\left[3+\frac{a}{b}\right]\left[1-\frac{a}{b}\right]^{-1}\right)} \left[\lambda_1\varphi_0\right]^{-1-\frac{a}{b}}. \tag{B14c}$$

We have verified that (B13) is a solution to (B2) given (B7) using automated symbolic manipulation in Wolfram Mathematica.

Appendix C Derivation of the Runaway Strain Rate Distribution

We simply note that the time-depedent PDE

$$\dot{\varphi} \propto V\left[\lambda_0\, c_1\, \varphi\, \mathcal{M}_{\text{ex}}(\varphi) - \varphi\right] \tag{C1}$$

to which a steady-state solution is sought in Section 3.5, closely relates to (B1) if in that equation the substitution $\frac{a}{b} \to 0$ is made. Making the same substitution in the result of Appendix B, we obtain:

$$\varphi(\infty, x) = \begin{cases} \frac{1}{\pi\lambda}\cos\left(\frac{1}{2}\frac{x}{\lambda_0}\right)^2 & \forall\, x \in [-\pi\lambda_0, +\pi\lambda_0] \\ 0 & \forall\, x \notin [-\pi\lambda_0, +\pi\lambda_0]. \end{cases} \tag{C2}$$

Appendix D Numerical Solution Procedure

We discretize the Laplacian using a standard second-order accurate central difference stencil, with natural boundary conditions of the same accuracy implemented by staggering the fields with respect to the physical domain walls. The problem size is halved

by exploiting symmetry across the shear zone. Integrals are evaluated using a midpoint rule. This gives the discrete system

$$\dot{\tilde{\boldsymbol{\zeta}}} = \mathbf{g}(\tilde{\boldsymbol{\zeta}}, \tilde{f})$$
$$\dot{\tilde{f}} = \mathrm{h}(\tilde{\boldsymbol{\zeta}}, \tilde{f}),$$

where tildes indicate approximate space-discrete quantities and numerical arrays are indicated in bold face. We form a symbolic rules for computing the Jacobian matrix $\mathbf{J}$ of this system of non-linear space-discrete equations,

$$\mathbf{J}(\tilde{\boldsymbol{\zeta}}, \tilde{f}) = \begin{bmatrix} \nabla_{\tilde{\boldsymbol{\zeta}}}\, \mathbf{g}(\tilde{\boldsymbol{\zeta}}, \tilde{f}) & \nabla_{\tilde{f}}\, \mathbf{g}(\tilde{\boldsymbol{\zeta}}, \tilde{f}) \\ \nabla_{\tilde{\boldsymbol{\zeta}}}\, \mathrm{h}(\tilde{\boldsymbol{\zeta}}, \tilde{f}) & \nabla_{\tilde{f}}\, \mathrm{h}(\tilde{\boldsymbol{\zeta}}, \tilde{f}) \end{bmatrix},$$

with the upper left block a dense $N \times N$ matrix with a dominant sparse band structure (N being the problem size), the lower right block a 1×1 empty matrix, and the off-diagonal blocks densely populated vectors of compatible shape. The system is then linearized as

$$\mathrm{d} \begin{bmatrix} \dot{\tilde{\boldsymbol{\zeta}}} \\ \dot{\tilde{f}} \end{bmatrix} = \mathbf{J}(\tilde{\boldsymbol{\zeta}}, \tilde{f})\, \mathrm{d} \begin{bmatrix} \tilde{\boldsymbol{\zeta}} \\ \tilde{f} \end{bmatrix}$$

and the smallest time scale $S_t(\tilde{\boldsymbol{\zeta}}, \tilde{f})$ to be resolved is computed as the reciprocal of the largest eigenvalue of the Jacobian matrix $\mathbf{J}$. The actual time step Δ_t is computed as some fixed fraction of S_t, optionally bounded by a maximum value and/or maximum growth rate to prevent time step overestimation as the time scale increases. The equations are discretized in time with forward and backward Euler schemes, both first-order accurate, respectively as

$$\begin{bmatrix} \tilde{\boldsymbol{\zeta}}^{k+1} \\ \tilde{f}^{k+1} \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{\zeta}}^{k} \\ \tilde{f}^{k} \end{bmatrix} + \Delta_t \begin{bmatrix} \mathbf{g}(\tilde{\boldsymbol{\zeta}}^{k}, \tilde{f}^{k}) \\ \mathrm{h}(\tilde{\boldsymbol{\zeta}}^{k}, \tilde{f}^{k}) \end{bmatrix} \tag{D1}$$

$$\mathbf{f} = \begin{bmatrix} \tilde{\boldsymbol{\zeta}}^{k+1} \\ \tilde{f}^{k+1} \end{bmatrix} - \begin{bmatrix} \tilde{\boldsymbol{\zeta}}^{k} \\ \tilde{f}^{k} \end{bmatrix} - \Delta_t \begin{bmatrix} \mathbf{g}(\tilde{\boldsymbol{\zeta}}^{k+1}, \tilde{f}^{k+1}) \\ \mathrm{h}(\tilde{\boldsymbol{\zeta}}^{k+1}, \tilde{f}^{k+1}) \end{bmatrix} = \mathbf{0}. \tag{D2}$$

Our algorithm makes an explicit-in-time prediction using (D1) and evaluates the l_2 norm of the implicit-in-time residual $\mathbf{f}$. It includes the possibility to perform Newton-Raphson iterations using the Jacobian $\mathbf{J}$ to keep the residual bounded, although we find it to be more efficient to experimentally set the dimensionless time step Δ_t / S_t sufficiently small (e.g. 0.1) to never cause tolerances to be violated. This value is then taken to correspond to a stable time step size of the explicit problem.

We make our algorithms and scripts to produce figures available as Wolfram Mathematica (Wolfram Research, Inc., 2017) notebooks in the Supplementary Materials.

References

Abdallah, Y., Sulem, J., & Stefanou, I. (2020). Compaction banding in high-porosity carbonate rocks: 2. a gradient-dependent plasticity model. *Journal of Geophysical Research: Solid Earth*, *125*(12), e2020JB020610.

Alevizos, S., Poulet, T., & Veveakis, E. (2014). Thermo-poro-mechanics of chemically active creeping faults. 1: Theory and steady state considerations. *Journal of Geophysical Research: Solid Earth*, *119*(6), 4558–4582.

Andrews, D. (1999). Test of two methods for faulting in finite-difference calculations. *Bulletin of the Seismological Society of America*, *89*(4), 931–937.

Bak, P., & Tang, C. (1989). Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research: Solid Earth*, *94*(B11), 15635-15637. doi: 10.1029/JB094iB11p15635

Barbot, S. (2019). Modulation of fault strength during the seismic cycle by grain-size evolution around contact junctions. *Tectonophysics*, *765*, 129 - 145. doi: 10.1016/j.tecto.2019.05.004

Barbot, S., & Fialko, Y. (2010). A unified continuum representation of post-seismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow. *Geophysical Journal International*, *182*, 1124-1140.

Barry, D., Barry, S., & Culligan-Hensley, P. (1995). Algorithm 743: Wapr–a fortran routine for calculating real values of the w-function. *ACM Transactions on Mathematical Software (TOMS)*, *21*(2), 172–181.

Barry, D., Culligan-Hensley, P., & Barry, S. (1995). Real values of the w-function. *ACM Transactions on Mathematical Software (TOMS)*, *21*(2), 161–171.

Barth, N. C., Boulton, C., Carpenter, B. M., Batt, G. E., & Toy, V. G. (2013). Slip localization on the southern alpine fault, new zealand. *Tectonics*, *32*(3), 620-640. doi: 10.1002/tect.20041

Bažant, Z. P., Belytschko, T. B., & Chang, T.-P. (1984). Continuum theory for strain-softening. *Journal of Engineering Mechanics*, *110*(12), 1666–1692.

Bažant, Z. P., & Jirásek, M. (2002). Nonlocal integral formulations of plasticity and damage: survey of progress. *Journal of engineering mechanics*, *128*(11), 1119–

1149.

Bedford, J. D., Faulkner, D. R., Allen, M. J., & Hirose, T. (2021). The stabilizing effect of high pore-fluid pressure along subduction megathrust faults: Evidence from friction experiments on accretionary sediments from the nankai trough. *Earth and Planetary Science Letters*, *574*, 117161.

Behr, W. M., Gerya, T. V., Cannizzaro, C., & Blass, R. (2021). Transient slow slip characteristics of frictional-viscous subduction megathrust shear zones. *AGU Advances*, *2*(3), e2021AV000416.

Ben-Zion, Y., & Rice, J. R. (1995). Slip patterns and earthquake populations along different classes of faults in elastic solids. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 12959–12983.

Ben-Zion, Y., & Rice, J. R. (1997). Dynamic simulations of slip on a smooth fault in an elastic solid. *Journal of Geophysical Research: Solid Earth*, *102*(B8), 17771-17784. doi: 10.1029/97JB01341

Ben-Zion, Y., & Zaliapin, I. (2020). Localization and coalescence of seismicity before large earthquakes. *Geophysical Journal International*, *223*(1), 561-583. doi: 10.1093/gji/ggaa315

Bouzid, M., Trulsson, M., Claudin, P., Clément, E., & Andreotti, B. (2013). Nonlocal rheology of granular flows across yield conditions. *Physical review letters*, *111*(23), 238301.

Bouzid, M., Trulsson, M., Claudin, P., Clément, E., & Andreotti, B. (2015). Microrheology to probe non-local effects in dense granular flows. *EPL (Europhysics Letters)*, *109*(2), 24002.

Bowden, F. P., & Tabor, D. (1966). Friction, lubrication and wear: a survey of work during the last decade. *British Journal of Applied Physics*, *17*(12), 1521–1544. doi: 10.1088/0508-3443/17/12/301

Brace, W., & Byerlee, J. (1966). Stick-slip as a mechanism for earthquakes. *Science*, *153*(3739), 990–992.

Braeck, S., & Podladchikov, Y. Y. (2007). Spontaneous thermal runaway as an ultimate failure mechanism of materials. *Phys. Rev. Lett.*, *98*, 095504. doi: 10.1103/PhysRevLett.98.095504

Broadbridge, P., & Bradshaw-Hajek, B. H. (2016). Exact solutions for logistic reaction–diffusion equations in biology. *Zeitschrift für angewandte Mathematik*

und Physik, *67*(4), 93. doi: 10.1007/s00033-016-0686-3

Burghardt, J., Brannon, R., & Guilkey, J. (2012). A nonlocal plasticity formulation for the material point method. *Computer methods in applied mechanics and engineering*, *225*, 55–64.

Carpenter, B., Saffer, D., & Marone, C. (2012). Frictional properties and sliding stability of the san andreas fault from deep drill core. *Geology*, *40*(8), 759–762.

Chen, J., & Spiers, C. J. (2016). Rate and state frictional and healing behavior of carbonate fault gouge explained using microphysical model. *Journal of Geophysical Research: Solid Earth*, *121*(12), 8642–8665.

Chester, F. M. (1994). Effects of temperature on friction: Constitutive equations and experiments with quartz gouge. *Journal of Geophysical Research: Solid Earth*, *99*(B4), 7247–7261.

Chester, F. M. (1995). A rheologic model for wet crust applied to strike-slip faults. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 13033–13044.

Chester, F. M., & Chester, J. S. (1998). Ultracataclasite structure and friction processes of the punchbowl fault, san andreas system, california. *Tectonophysics*, *295*(1-2), 199–221.

Coble, C. G., French, M. E., Chester, F. M., Chester, J. S., & Kitajima, H. (2014). In situ frictional properties of San Andreas Fault gouge at SAFOD. *Geophysical Journal International*, *199*(2), 956-967. Retrieved from `https://doi.org/10.1093/gji/ggu306` doi: 10.1093/gji/ggu306

Cocco, M., & Bizzarri, A. (2002). On the slip-weakening behavior of rate- and state dependent constitutive laws. *Geophysical Research Letters*, *29*(11), 11-1-11-4. doi: 10.1029/2001GL013999

Cocco, M., Bizzarri, A., & Tinti, E. (2004). Physical interpretation of the breakdown process using a rate- and state-dependent friction law. *Tectonophysics*, *378*(3), 241 - 262. (Physics of Active Faults - Theory, Observation and Experiments) doi: 10.1016/j.tecto.2003.09.015

Cochard, A., & Madariaga, R. (1994). Dynamic faulting under rate-dependent friction. *Pure and Applied Geophysics*, *142*, 419-445. doi: 10.1007/BF00876049

Cochard, A., & Madariaga, R. (1996). Complexity of seismicity due to highly rate-dependent friction. *Journal of Geophysical Research: Solid Earth*, *101*(B11), 25321–25336.

Crameri, F. (2021). *Scientific colour maps.* Zenodo. Retrieved from `https://doi.org/10.5281/zenodo.5501399` doi: 10.5281/zenodo.5501399

Crameri, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. *Nature communications*, *11*(1), 1–10.

DalZilio, L., Lapusta, N., Avouac, J.-P., & Gerya, T. (2021). Subduction earthquake sequences in a non-linear visco-elasto-plastic megathrust. *Geophysical Journal International*, *229*(2), 1098-1121. Retrieved from `https://doi.org/10.1093/gji/ggab521` doi: 10.1093/gji/ggab521

Daub, E. G., & Carlson, J. M. (2008). A constitutive model for fault gouge deformation in dynamic rupture simulations. *Journal of Geophysical Research: Solid Earth*, *113*(B12).

Daub, E. G., & Carlson, J. M. (2009). Stick-slip instabilities and shear strain localization in amorphous materials. *Physical Review E*, *80*(6), 066113.

Daub, E. G., Manning, M. L., & Carlson, J. M. (2008). Shear strain localization in elastodynamic rupture simulations. *Geophysical research letters*, *35*(12).

Daub, E. G., Manning, M. L., & Carlson, J. M. (2010). Pulse-like, crack-like, and supershear earthquake ruptures with shear strain localization. *Journal of Geophysical Research: Solid Earth*, *115*(B5).

Day, S. M., Dalguer, L. A., Lapusta, N., & Liu, Y. (2005). Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *Journal of Geophysical Research: Solid Earth*, *110*(B12). doi: 10.1029/2005JB003813

de Borst, R., & Duretz, T. (2020). On viscoplastic regularisation of strain-softening rocks and soils. *International Journal for Numerical and Analytical Methods in Geomechanics*, *44*(6), 890-903. doi: 10.1002/nag.3046

de Borst, R., & Mhlhaus, H.-B. (1992). Gradient-dependent plasticity: Formulation and algorithmic aspects. *International Journal for Numerical Methods in Engineering*, *35*(3), 521-539. doi: 10.1002/nme.1620350307

Dieterich, J. H. (1978). Time-dependent friction and the mechanics of stick-slip. *pure and applied geophysics*, *116*(4), 790–806. doi: 10.1007/BF00876539

Dieterich, J. H. (1979a). Modeling of rock friction: 1. experimental results and constitutive equations. *Journal of Geophysical Research: Solid Earth*, *84*(B5), 2161–2168. doi: 10.1029/JB084iB05p02161

Dieterich, J. H. (1979b). Modeling of rock friction: 2. simulation of preseismic slip. *Journal of Geophysical Research: Solid Earth*, *84*(B5), 2169–2175. doi: 10.1029/JB084iB05p02169

Dieterich, J. H., & Kilgore, B. D. (1994). Direct observation of frictional contacts: New insights for state-dependent properties. *pure and applied geophysics*, *143*(1), 283–302. doi: 10.1007/BF00874332

Drucker, D. C., & Prager, W. (1952). Soil mechanics and plastic analysis or limit design. *Quarterly of applied mathematics*, *10*(2), 157–165.

Duretz, T., de Borst, R., & Le Pourhiet, L. (2019). Finite thickness of shear bands in frictional viscoplasticity and implications for lithosphere dynamics. *Geochemistry, Geophysics, Geosystems*. doi: 10.1029/2019GC008531

Duretz, T., de Borst, R., & Yamato, P. (2021). Modeling lithospheric deformation using a compressible visco-elasto-viscoplastic rheology and the effective viscosity approach. *Geochemistry, Geophysics, Geosystems*, *22*(8). doi: 10.1029/2021GC009675

Duretz, T., de Borst, R., Yamato, P., & Le Pourhiet, L. (2020). Toward robust and predictive geodynamic modeling: The way forward in frictional plasticity. *Geophysical Research Letters*, *47*(5), e2019GL086027. doi: 10.1029/2019GL086027

Duretz, T., Souche, A., de Borst, R., & Le Pourhiet, L. (2018). The benefits of using a consistent tangent operator for viscoelastoplastic computations in geodynamics. *Geochemistry, Geophysics, Geosystems*, *19*(12), 4904–4924.

Elbanna, A. E., & Carlson, J. M. (2014). A two-scale model for sheared fault gouge: Competition between macroscopic disorder and local viscoplasticity. *Journal of Geophysical Research: Solid Earth*, *119*(6), 4841–4859.

Engelen, R. A., Geers, M. G., & Baaijens, F. P. (2003). Nonlocal implicit gradient-enhanced elasto-plasticity for the modelling of softening behaviour. *International Journal of Plasticity*, *19*(4), 403 - 433. doi: 10.1016/S0749-6419(01)00042-0

Erickson, B. A., Jiang, J., Barall, M., Lapusta, N., Dunham, E. M., Harris, R., ... Wei, M. (2020). The Community Code Verification Exercise for Simulating Sequences of Earthquakes and Aseismic Slip (SEAS). *Seismological Research Letters*. doi: 10.1785/0220190248

Faulkner, D., Mitchell, T., Jensen, E., & Cembrano, J. (2011). Scaling of fault damage zones with displacement and the implications for fault growth processes. *Journal of Geophysical Research: Solid Earth*, *116*(B5).

Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of Eugenics*, *7*(4), 355-369. doi: 10.1111/j.1469-1809.1937.tb02153.x

Fritsch, F. N., Shafer, R., & Crowley, W. (1973). Algorithm 443: Solution of the transcendental equation $we^w = x$. *Communications of the ACM*, *16*(2), 123–124.

Gabriel, A.-A., Ampuero, J.-P., Dalguer, L. A., & Mai, P. M. (2012). The transition of dynamic rupture styles in elastic media under velocity-weakening friction. *Journal of Geophysical Research: Solid Earth*, *117*(B9).

Gabriel, A.-A., Li, D., Chiocchetti, S., Tavelli, M., Peshkov, I., Romenski, E., & Dumbser, M. (2021). A unified first-order hyperbolic model for nonlinear dynamic rupture processes in diffuse fracture zones. *Philosophical Transactions of the Royal Society A*, *379*(2196), 20200130.

Garagash, D. I. (2012). Seismic and aseismic slip pulses driven by thermal pressurization of pore fluid. *Journal of Geophysical Research: Solid Earth*, *117*(B4).

Gilding, B., & Kersner, R. (2005). A fisher/kpp-type equation with density-dependent diffusion and convection: travelling-wave solutions. *Journal of Physics A: Mathematical and General*, *38*, 3367. doi: 10.1088/0305-4470/38/15/009

Godunov, S., & Romenskii, E. (1972). Nonstationary equations of nonlinear elasticity theory in eulerian coordinates. *Journal of Applied Mechanics and Technical Physics*, *13*(6), 868–884.

Godunov, S. K., & Romenskii, E. (2003). *Elements of continuum mechanics and conservation laws*. Springer Science & Business Media.

Granier, T. (1985). Origin, damping, and pattern of development of faults in granite. *Tectonics*, *4*(7), 721-737. doi: 10.1029/TC004i007p00721

Hermundstad, A., Daub, E., & Carlson, J. (2010). Energetics of strain localization in a model of seismic slip. *Journal of Geophysical Research: Solid Earth*, *115*(B6).

Herrendörfer, R., Gerya, T., & Van Dinther, Y. (2018). An invariant rate-and state-dependent friction formulation for viscoeastoplastic earthquake cycle

simulations. *Journal of Geophysical Research: Solid Earth*, *123*(6), 5018–5051.

Hobbs, B. E., Mühlhaus, H.-B., & Ord, A. (1990). Instability, softening and localization of deformation. *Geological Society, London, Special Publications*, *54*(1), 143–165. doi: 10.1144/GSL.SP.1990.054.01.15

Ida, Y. (1972). Cohesive force across the tip of a longitudinal-shear crack and griffith's specific surface energy. *Journal of Geophysical Research*, *77*(20), 3796–3805.

Ikari, M. J., Marone, C., & Saffer, D. M. (2011). On the relation between fault strength and frictional stability. *Geology*, *39*(1), 83–86.

Jiang, J., Erickson, B. A., Lambert, V. R., Ampuero, J.-P., Ando, R., Barbot, S. D., ... van Dinther, Y. (2022). Community-driven code comparisons for three-dimensional dynamic modeling of sequences of earthquakes and aseismic slip. *Journal of Geophysical Research: Solid Earth*, *127*(3), e2021JB023519. doi: 10.1029/2021JB023519

Jiang, J., & Lapusta, N. (2016). Deeper penetration of large earthquakes on seismically quiescent faults. *Science*, *352*(6291), 1293–1297.

Jirásek, M. (1998). Nonlocal models for damage and fracture: comparison of approaches. *International Journal of Solids and Structures*, *35*(31-32), 4133–4145.

Jirásek, M., & Rolshoven, S. (2009a). Localization properties of strain-softening gradient plasticity models. part ii: Theories with gradients of internal variables. *International Journal of Solids and Structures*, *46*(11-12), 2239–2254.

Jirásek, M., & Rolshoven, S. (2009b). Localization properties of strain-softening gradient plasticity models. part i: Strain-gradient theories. *International Journal of Solids and Structures*, *46*(11-12), 2225–2238.

Johansson, F. (2020). Computing the lambert w function in arbitrary-precision complex interval arithmetic. *Numerical Algorithms*, *83*(1), 221–242.

John, T., Medvedev, S., Rüpke, L. H., Andersen, T. B., Podladchikov, Y. Y., & Austrheim, H. (2009). Generation of intermediate-depth earthquakes by self-localizing thermal runaway. *Nature Geoscience*, *2*, 137–140.

Katz, Y., Weinberger, R., & Aydin, A. (2004). Geometry and kinematic evolution of riedel shear structures, capitol reef national park, utah. *Journal of Structural Geology*, *26*(3), 491-501. doi: 10.1016/j.jsg.2003.08.003

Kaus, B. J. (2010). Factors that control the angle of shear bands in geodynamic numerical models of brittle deformation. *Tectonophysics*, *484*(1), 36 - 47. (Quantitative modelling of geological processes) doi: 10.1016/j.tecto.2009.08.042

Kiefer, B., Waffenschmidt, T., Sprave, L., & Menzel, A. (2018). A gradient-enhanced damage model coupled to plasticitymulti-surface formulation and algorithmic concepts. *International Journal of Damage Mechanics*, *27*(2), 253-295. doi: 10.1177/1056789516676306

Kolmogorov, A. N., Petrovskii, I. G., & Piskunov, N. S. (1937). A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. In V. Tikhomirov (Ed.), *Selected works of a.n. kolmogorov 1 (1991)* (pp. 242–270). Kluwer.

Kuhn, C., Schlüter, A., & Müller, R. (2015). On degradation functions in phase field fracture models. *Computational Materials Science*, *108*, 374–384.

Kurzon, I., Lyakhovsky, V., & Ben-Zion, Y. (2019). Dynamic rupture and seismic radiation in a damage–breakage rheology model. *Pure and Applied Geophysics*, *176*(3), 1003–1020.

Kurzon, I., Lyakhovsky, V., & Ben-Zion, Y. (2020). Earthquake source properties from analysis of dynamic ruptures and far-field seismic waves in a damage-breakage model. *Geophysical Journal International*, *224*(3), 1793-1810. doi: 10.1093/gji/ggaa509

Lapusta, N., & Liu, Y. (2009). Three-dimensional boundary integral modeling of spontaneous earthquake sequences and aseismic slip. *Journal of Geophysical Research: Solid Earth*, *114*(B9). doi: 10.1029/2008JB005934

Lapusta, N., Rice, J. R., Ben-Zion, Y., & Zheng, G. (2000). Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research: Solid Earth*, *105*(B10), 23765-23789. doi: 10.1029/2000JB900250

Le Pourhiet, L. (2013). Strain localization due to structural softening during pressure sensitive rate independent yielding. *Bulletin de la Société géologique de France*, *184*(4-5), 357–371.

Li, D., & Liu, Y. (2017). Modeling slow-slip segmentation in cascadia subduction zone constrained by tremor locations and gravity anomalies. *Journal of Geophysical Research: Solid Earth*, *122*(4), 3138–3157.

Li, Y., & Wu, Y. (2008). Stability of travelling waves with noncritical speeds for double degenerate fisher-type equations. *Discrete and Continuous Dynamical Systems - Series B*, *10*, 149-170. doi: 10.3934/dcdsb.2008.10.149

Linker, M., & Dieterich, J. H. (1992). Effects of variable normal stress on rock friction: Observations and constitutive equations. *Journal of Geophysical Research: Solid Earth*, *97*(B4), 4923–4940.

Locatelli, M., Verlaguet, A., Agard, P., Federico, L., & Angiboust, S. (2018). Intermediate-depth brecciation along the subduction plate interface (monviso eclogite, w. alps). *Lithos*, *320*, 378–402.

Locatelli, M., Verlaguet, A., Agard, P., Pettke, T., & Federico, L. (2019). Fluid pulses during stepwise brecciation at intermediate subduction depths (monviso eclogites, w. alps): First internally then externally sourced. *Geochemistry, geophysics, geosystems*, *20*(11), 5285–5318.

Lyakhovsky, V., & Ben-Zion, Y. (2014a). A continuum damage–breakage faulting model and solid-granular transitions. *Pure and Applied Geophysics*, *171*(11), 3099–3123.

Lyakhovsky, V., & Ben-Zion, Y. (2014b). Damage–breakage rheology model and solid-granular transition near brittle instability. *Journal of the Mechanics and Physics of Solids*, *64*, 184–197.

Lyakhovsky, V., Ben-Zion, Y., Ilchev, A., & Mendecki, A. (2016). Dynamic rupture in a damage-breakage rheology model. *Geophysical Journal International*, *206*(2), 1126–1143.

Lyakhovsky, V., Hamiel, Y., & Ben-Zion, Y. (2011). A non-local visco-elastic damage model and dynamic fracturing. *Journal of the Mechanics and Physics of Solids*, *59*(9), 1752–1776.

Mair, K., & Marone, C. (1999). Friction of simulated fault gouge for a wide range of velocities and normal stresses. *Journal of Geophysical Research: Solid Earth*, *104*(B12), 28899-28914. doi: 10.1029/1999JB900279

Malaguti, L., & Ruggerini, S. (2010). Asymptotic speed of propagation for fisher-type degenerate reaction-diffusion-convection equations. *Advanced Nonlinear Studies*, *10*. doi: 10.1515/ans-2010-0306

Marone, C. (1998). The effect of loading rate on static friction and the rate of fault healing during the earthquake cycle. *Nature*, *391*(6662), 69-72. doi: 10.1038/

34157

Marone, C., Hobbs, B. E., & Ord, A. (1992). Coulomb constitutive laws for friction: Contrasts in frictional behavior for distributed and localized shear. *pure and applied geophysics*, *139*(2), 195–214. doi: 10.1007/BF00876327

McBeck, J. A., Aiken, J. M., Mathiesen, J., Ben-Zion, Y., & Renard, F. (2020). Deformation precursors to catastrophic failure in rocks. *Geophysical Research Letters*, *47*(24), e2020GL090255.

McBeck, J. A., Ben-Zion, Y., & Renard, F. (2020). The mixology of precursory strain partitioning approaching brittle failure in rocks. *Geophysical Journal International*, *221*(3), 1856–1872.

McBeck, J. A., Kobchenko, M., Hall, S. A., Tudisco, E., Cordonnier, B., Meakin, P., & Renard, F. (2018). Investigating the onset of strain localization within anisotropic shale using digital volume correlation of time-resolved x-ray microtomography images. *Journal of Geophysical Research: Solid Earth*, *123*(9), 7509–7528.

McBeck, J. A., Zhu, W., & Renard, F. (2021). The competition between fracture nucleation, propagation, and coalescence in dry and water-saturated crystalline rock. *Solid Earth*, *12*(2), 375–387.

Mühlhaus, H.-B., & Aifantis, E. (1991). A variational principle for gradient plasticity. *International Journal of Solids and Structures*, *28*(7), 845–857.

Mühlhaus, H.-B., & Vardoulakis, I. (1987). The thickness of shear bands in granular materials. *Geotechnique*, *37*(3), 271–283.

Needleman, A. (1988). Material rate dependence and mesh sensitivity in localization problems. *Computer Methods in Applied Mechanics and Engineering*, *67*(1), 69 - 85. doi: 10.1016/0045-7825(88)90069-2

Niemeijer, A., & Spiers, C. (2007). A microphysical model for strong velocity weakening in phyllosilicate-bearing fault gouges. *Journal of Geophysical Research: Solid Earth*, *112*(B10).

Okubo, K., Bhat, H. S., Rougier, E., Marty, S., Schubnel, A., Lei, Z., ... Klinger, Y. (2019). Dynamics, radiation, and overall energy budget of earthquake rupture with coseismic off-fault damage. *Journal of Geophysical Research: Solid Earth*, *124*(11), 11771-11801. doi: 10.1029/2019JB017304

Okubo, P. G. (1989). Dynamic rupture modeling with laboratory-derived consti-

tutive relations. *Journal of Geophysical Research: Solid Earth*, *94*(B9), 12321–12335.

Passelègue, F. X., Almakari, M., Dublanchet, P., Barras, F., Fortin, J., & Violay, M. (2020). Initial effective stress controls the nature of earthquakes. *Nature communications*, *11*(1), 1–8.

Passelègue, F. X., Spagnuolo, E., Violay, M., Nielsen, S., Di Toro, G., & Schubnel, A. (2016). Frictional evolution, acoustic emissions activity, and off-fault damage in simulated faults sheared at seismic slip rates. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7490–7513.

Peerlings, R. H. J., de Borst, R., Brekelmans, W. A. M., & De Vree, J. H. P. (1996). Gradient enhanced damage for quasi-brittle materials. *International Journal for Numerical Methods in Engineering*, *39*(19), 3391-3403. doi: 10.1002/(SICI)1097-0207(19961015)39:19⟨3391::AID-NME7⟩3.0.CO;2-D

Peirce, D., Asaro, R. J., & Needleman, A. (1983). Material rate dependence and localized deformation in crystalline solids. *Acta metallurgica*, *31*(12), 1951–1976.

Perez-Silva, A., Li, D., Gabriel, A.-A., & Kaneko, Y. (2021). 3d modeling of long-term slow slip events along the flat-slab segment in the guerrero seismic gap, mexico. *Geophysical Research Letters*, *48*(13), e2021GL092968.

Perrin, C., Manighetti, I., & Gaudemer, Y. (2016). Off-fault tip splay networks: A genetic and generic property of faults indicative of their long-term propagation. *Comptes Rendus Geoscience*, *348*(1), 52-60. (From past to current tectonics) doi: 10.1016/j.crte.2015.05.002

Petrovskii, S. V., & Li, B.-L. (2003). An exactly solvable model of population dynamics with density-dependent migrations and the allee effect. *Mathematical Biosciences*, *186*(1), 79 - 91. doi: 10.1016/S0025-5564(03)00098-1

Petrovskii, S. V., & Li, B.-L. (2006). *Exactly solvable models of biological invasion.* Chapman and Hall/CRC.

Pijaudier-Cabot, G., & Bažant, Z. P. (1987). Nonlocal damage theory. *Journal of engineering mechanics*, *113*(10), 1512–1533.

Platt, J. D., Rudnicki, J. W., & Rice, J. R. (2014). Stability and localization of rapid shear in fluid-saturated fault gouge: 2. localized zone width and strength evolution. *Journal of Geophysical Research: Solid Earth*, *119*(5), 4334–4359.

Poulet, T., Veveakis, E., Regenauer-Lieb, K., & Yuen, D. A. (2014). Thermo-poro-mechanics of chemically active creeping faults: 3. the role of serpentinite in episodic tremor and slip sequences, and transition to chaos. *Journal of Geophysical Research: Solid Earth*, *119*(6), 4606–4625.

Pozzi, G., De Paola, N., Nielsen, S. B., Holdsworth, R. E., & Bowen, L. (2018). A new interpretation for the nature and significance of mirror-like surfaces in experimental carbonate-hosted seismic faults. *Geology*, *46*(7), 583-586. doi: 10.1130/G40197.1

Pozzi, G., De Paola, N., Nielsen, S. B., Holdsworth, R. E., Tesei, T., Thieme, M., & Demouchy, S. (2021). Coseismic fault lubrication by viscous deformation. *Nature Geoscience*, *14*(6), 437–442.

Pozzi, G., Paola, N. D., Holdsworth, R. E., Bowen, L., Nielsen, S. B., & Dempsey, E. D. (2019). Coseismic ultramylonites: An investigation of nanoscale viscous flow and fault weakening during seismic slip. *Earth and Planetary Science Letters*, *516*, 164 - 175. doi: 10.1016/j.epsl.2019.03.042

Preuss, S., Ampuero, J. P., Gerya, T., & van Dinther, Y. (2020). Characteristics of earthquake ruptures and dynamic off-fault deformation on propagating faults. *Solid Earth*, *11*(4), 1333–1360.

Preuss, S., Herrendrfer, R., Gerya, T., Ampuero, J.-P., & van Dinther, Y. (2019). Seismic and aseismic fault growth lead to different fault orientations. *Journal of Geophysical Research: Solid Earth*, *124*(8), 8867-8889. doi: 10.1029/2019JB017324

Putelat, T., Dawes, J., & Champneys, A. (2017). A phase-plane analysis of localized frictional waves. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *473*(2203), 20160606.

Rabinowicz, E. (1958). The intrinsic variables affecting the stick-slip process. *Proceedings of the Physical Society (1958-1967)*, *71*(4), 668.

Rattez, H., Stefanou, I., & Sulem, J. (2018). The importance of thermo-hydro-mechanical couplings and microstructure to strain localization in 3d continua with application to seismic faults. part i: Theory and linear stability analysis. *Journal of the Mechanics and Physics of Solids*, *115*, 54 - 76.

Rattez, H., Stefanou, I., Sulem, J., Veveakis, M., & Poulet, T. (2018). The importance of thermo-hydro-mechanical couplings and microstructure to strain

localization in 3d continua with application to seismic faults. part ii: Numerical implementation and post-bifurcation analysis. *Journal of the Mechanics and Physics of Solids*, *115*, 1–29.

Resnyansky, A., Romensky, E., & Bourne, N. (2003). Constitutive modeling of fracture waves. *Journal of applied physics*, *93*(3), 1537–1545.

Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research: Solid Earth*, *98*(B6), 9885-9907. doi: 10.1029/93JB00191

Rice, J. R. (2006). Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth*, *111*(B5).

Rice, J. R., & Ben-Zion, Y. (1996). Slip complexity in earthquake fault models. *Proceedings of the National Academy of Sciences*, *93*(9), 3811–3818. doi: 10.1073/pnas.93.9.3811

Rice, J. R., Rudnicki, J. W., & Platt, J. D. (2014). Stability and localization of rapid shear in fluid-saturated fault gouge: 1. linearized stability analysis. *Journal of Geophysical Research: Solid Earth*, *119*(5), 4311–4333.

Ritter, M. C., Rosenau, M., & Oncken, O. (2018). Growing faults in the lab: Insights into the scale dependence of the fault zone evolution process. *Tectonics*, *37*(1), 140-153. doi: 10.1002/2017TC004787

Ritter, M. C., Santimano, T., Rosenau, M., Leever, K., & Oncken, O. (2018). Sandbox rheometry: Co-evolution of stress and strain in riedel and critical wedgeexperiments. *Tectonophysics*, *722*, 400-409. doi: 10.1016/j.tecto.2017.11.018

Romenski, E., Reshetova, G., Peshkov, I., & Dumbser, M. (2020). Modeling wavefields in saturated elastic porous media based on thermodynamically compatible system theory for two-phase solid-fluid mixtures. *Computers & Fluids*, *206*, 104587. doi: 10.1016/j.compfluid.2020.104587

Romenskii, E. (2007). Deformation model for brittle materials and the structure of failure waves. *Journal of applied mechanics and technical physics*, *48*(3), 437–444.

Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., ... others (2019). Hierarchical interlocked orthogonal faulting in the 2019 ridgecrest earthquake sequence. *Science*, *366*(6463), 346–351.

Roubek, T. (2014). A note about the rate-and-state-dependent friction model in a thermodynamic framework of the Biot-type equation . *Geophysical Journal In-*

ternational, *199*(1), 286-295. doi: 10.1093/gji/ggu248

Rozel, A., Ricard, Y., & Bercovici, D. (2011). A thermodynamically self-consistent damage equation for grain size evolution during dynamic recrystallization. *Geophysical Journal International*, *184*(2), 719–728.

Rubin, A. M., & Ampuero, J.-P. (2005). Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research: Solid Earth*, *110*(B11).

Ruina, A. L. (1980). *Friction laws and instabilities: A quasistatic analysis of some dry friction behaviour* (Unpublished doctoral dissertation). Ph.D. Thesis, Division of Engineering, Brown University, Providence RI.

Ruina, A. L. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research: Solid Earth*, *88*(B12), 10359-10370. doi: 10.1029/JB088iB12p10359

Sánchez-Garduño, F., & Maini, P. (1994). An approximation to a sharp type solution of a density-dependent reaction-diffusion equation. *Applied Mathematics Letters*, *7*(1), 47 - 51. doi: 10.1016/0893-9659(94)90051-5

Sánchez-Garduño, F., & Maini, P. (1995). Traveling wave phenomena in some degenerate reaction-diffusion equations. *Journal of Differential Equations*, *117*(2), 281 - 319. doi: 10.1006/jdeq.1995.1055

Sánchez-Garduño, F., & Maini, P. (1997). Travelling wave phenomena in non-linear diffusion degenerate nagumo equations. *Journal of Mathematical Biology*, *35*(6), 713–728. doi: 10.1007/s002850050073

Sathiakumar, S., Barbot, S., & Hubbard, J. (2020). Earthquake cycles in fault-bend folds. *Journal of Geophysical Research: Solid Earth*, *125*(8), e2019JB018557.

Savage, H. M., & Brodsky, E. E. (2011). Collateral damage: Evolution with displacement of fracture distribution and secondary fault strands in fault damage zones. *Journal of Geophysical Research: Solid Earth*, *116*(B3).

Scholz, C., Molnar, P., & Johnson, T. (1972). Detailed studies of frictional sliding of granite and implications for the earthquake mechanism. *Journal of geophysical research*, *77*(32), 6392–6406.

Schreyer, H. L., & Chen, Z. (1986). One-Dimensional Softening With Localization. *Journal of Applied Mechanics*, *53*(4), 791-797. doi: 10.1115/1.3171860

Segall, P., & Rice, J. R. (1995). Dilatancy, compaction, and slip instability of a fluid-infiltrated fault. *Journal of Geophysical Research: Solid Earth*, *100*(B11),

22155-22171. doi: 10.1029/95JB02403

Sherratt, J., & Marchant, B. (1996). Nonsharp travelling wave fronts in the fisher equation with degenerate nonlinear diffusion. *Applied Mathematics Letters*, *9*(5), 33 - 38. doi: 10.1016/0893-9659(96)00069-9

Simo, J., & Taylor, R. (1985). Consistent tangent operators for rate-independent elastoplasticity. *Computer Methods in Applied Mechanics and Engineering*, *48*(1), 101 - 118. doi: 10.1016/0045-7825(85)90070-2

Sleep, N. H. (1995). Ductile creep, compaction, and rate and state dependent friction within major fault zones. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 13065–13080.

Sleep, N. H. (1997). Application of a unified rate and state friction theory to the mechanics of fault zones with strain localization. *Journal of Geophysical Research: Solid Earth*, *102*(B2), 2875-2895. doi: 10.1029/96JB03410

Stathas, A., & Stefanou, I. (2022). The role of viscous regularization in dynamical problems, strain localization and mesh dependency. *Computer Methods in Applied Mechanics and Engineering*, *388*, 114185. doi: 10.1016/j.cma.2021.114185

Tchalenko, J. (1970). Similarities between shear zones of different magnitudes. *Geological Society of America Bulletin*, *81*(6), 1625–1640.

Templeton, E. L., & Rice, J. R. (2008). Off-fault plasticity and earthquake rupture dynamics: 1. dry materials or neglect of fluid pressure changes. *Journal of Geophysical Research: Solid Earth*, *113*(B9).

Thielmann, M. (2018). Grain size assisted thermal runaway as a nucleation mechanism for continental mantle earthquakes: Impact of complex rheologies. *Tectonophysics*, *746*, 611–623. doi: 10.1016/j.tecto.2017.08.038

Thielmann, M., Rozel, A., Kaus, B., & Ricard, Y. (2015). Intermediate-depth earthquake generation and shear zone formation caused by grain size reduction and shear heating. *Geology*, *43*(9), 791. doi: 10.1130/G36864.1

Thomas, M. Y., Lapusta, N., Noda, H., & Avouac, J.-P. (2014). Quasi-dynamic versus fully dynamic simulations of earthquakes and aseismic slip with and without enhanced coseismic weakening. *Journal of Geophysical Research: Solid Earth*, *119*(3), 1986-2004. doi: 10.1002/2013JB010615

Triantafyllidis, N., & Aifantis, E. C. (1986). A gradient approach to localization of

deformation. i. hyperelastic materials. *Journal of Elasticity*, *16*(3), 225–237.

Ulm, F.-J., Coussy, O., & Bažant, Z. P. (1999). The "chunnel" fire. i: Chemoplastic softening in rapidly heated concrete. *Journal of engineering mechanics*, *125*(3), 272–282.

Van den Ende, M., Chen, J., Ampuero, J.-P., & Niemeijer, A. (2018). A comparison between rate-and-state friction and microphysical models, based on numerical simulations of fault slip. *Tectonophysics*, *733*, 273–295.

Vardoulakis, I. (1980). Shear band inclination and shear modulus of sand in biaxial tests. *International journal for numerical and analytical methods in geomechanics*, *4*(2), 103–119.

Vermeer, P. (1990). The orientation of shear bands in biaxial tests. *Geotechnique*, *40*(2), 223–236.

Veveakis, E., Poulet, T., & Alevizos, S. (2014). Thermo-poro-mechanics of chemically active creeping faults: 2. transient considerations. *Journal of Geophysical Research: Solid Earth*, *119*(6), 4583–4605.

Viesca, R. C. (2016a). Self-similar slip instability on interfaces with rate-and state-dependent friction. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *472*(2192), 20160254.

Viesca, R. C. (2016b). Stable and unstable development of an interfacial sliding instability. *Physical Review E*, *93*(6), 060202.

Viesca, R. C. (2020). On the existence of a nucleation length for dynamic shear rupture. *arXiv preprint arXiv:2008.11854*.

Wolfram Research, Inc. (2017). *Mathematica, Version 11.1.1.0.* (Champaign, IL, 2017)

Wu, F., & Freund, L. (1984). Deformation trapping due to thermoplastic instability in one-dimensional wave propagation. *Journal of the Mechanics and Physics of Solids*, *32*(2), 119-132. doi: 10.1016/0022-5096(84)90014-0

Wu, J.-Y. (2017). A unified phase-field theory for the mechanics of damage and quasi-brittle failure. *Journal of the Mechanics and Physics of Solids*, *103*, 72–99.

Wu, Y., Xing, X., & Ye, Q. (2006). Stability of travelling waves with algebraic decay for n-degree fisher-type equations. *Discrete and Continuous Dynamical Systems*, *16*. doi: 10.3934/dcds.2006.16.47