

Adjoint-based inversion for stress and frictional parameters in earthquake modeling

Vidar Stiernström*, Martin Almquist*, Eric M. Dunham†

November 3, 2024

Abstract

We present an adjoint-based optimization method to invert for stress and frictional parameters used in earthquake modeling. The forward problem is linear elastodynamics with nonlinear rate-and-state frictional faults. The misfit functional quantifies the difference between simulated and measured particle displacements or velocities at receiver locations. The misfit may include windowing or filtering operators. We derive the corresponding adjoint problem, which is linear elasticity with linearized rate-and-state friction and, for forward problems involving fault normal stress changes, nonzero fault opening, with time-dependent coefficients derived from the forward solution. The gradient of the misfit is efficiently computed by convolving forward and adjoint variables on the fault. The method thus extends the framework of full-waveform inversion to include frictional faults with rate-and-state friction. In addition, we present a space-time dual-consistent discretization of a dynamic rupture problem with a rough fault in antiplane shear, using high-order accurate summation-by-parts finite differences in combination with explicit Runge–Kutta time integration. The dual consistency of the discretization ensures that the discrete adjoint-based gradient is the exact gradient of the discrete misfit functional as well as a consistent approximation of the continuous gradient. Our theoretical results are corroborated by inversions with synthetic data. We anticipate that adjoint-based inversion of seismic and/or geodetic data will be a powerful tool for studying earthquake source processes; it can also be used to interpret laboratory friction experiments.

1 Introduction

Earthquake cycle and dynamic rupture simulations with fault dynamics governed by rate-and-state friction have emerged as promising tools to better understand the processes governing earthquake nucleation and seismicity. Dynamic

*Department of Information Technology, Uppsala University, SE-751 05 Uppsala, Sweden, cstierns@stanford.edu, martin.almquist@it.uu.se

†Department of Geophysics, Stanford University, Stanford CA 94305, USA, edunham@stanford.edu

source models are complementary to kinematic source models that express the seismic wavefield and solid displacements through a convolution of the specified space-time history of slip with elastic Green’s functions. Kinematic models are widely used in slip inversions, which can be set up as linear least squares problems due to the linear relation between slip and the wavefield. However, kinematic inversions only reveal how faults slip, but not why they slip in a certain way. Furthermore, the solutions to the inverse problem can violate certain physical constraints, such as dissipation rather than creation of mechanical energy during frictional sliding. Dynamic source models go beyond kinematic source models through the introduction of friction laws that are based on laboratory experiments. These friction laws involve nonlinear relations between fault shear and normal tractions, slip velocity, and one or more state variables that capture the dependence of frictional strength on the slip history of the interface. After specification of the initial stresses and frictional parameters, dynamic rupture and earthquake cycle models provide both the slip history and wavefield (or quasi-static solid displacement field) as part of the solution to the nonlinear problem. Connecting cycle and dynamic rupture simulations to real data permits determination of the stresses and frictional parameters, which are otherwise difficult or impossible to constrain. This requires augmenting dynamic source models with inversion capabilities. The inversion is a PDE-constrained optimization problem where we seek parameter values that minimize the misfit between model output and data. We propose a gradient-based optimization where the gradient of the misfit is computed with only two simulations: one of the forward problem and one of the adjoint problem. Gradient-based optimization with the adjoint method is the preferred approach in full waveform inversions (FWI) in seismic tomography and reflection seismology [Tromp et al. \[2005\]](#), [Fichtner \[2011\]](#), [Virieux and Operto \[2009\]](#).

To date, dynamic rupture inversions have been done either by grid-search algorithms or Bayesian methods that require running thousands to order of a million forward models [Peyrat and Olsen \[2004\]](#), [Gallovič et al. \[2019a,b\]](#), [Premus et al. \[2022\]](#). Friction inversion for earthquake cycle models has been posed as a sequential data assimilation problem that is solved, for example, using the ensemble Kalman filter [van Dinther et al. \[2019\]](#), [Hirahara and Nishikiori \[2019\]](#). Frictional properties are treated as part of the state vector and updated in time. However, these parameters should not be time-dependent and we seek an alternative formulation that respects this.

In this work, we present an adjoint-based optimization framework for inverting rate-and-state frictional parameters and initial stresses from geodetic and seismic observations. The method presented in this work extends FWI to include nonlinear rate-and-state friction laws. Similar to FWI, the adjoint equations are the same linear elasticity equations as the forward equations, but with fault dynamics governed by what resembles linearized rate-and-state friction and, for forward problems involving fault normal stress changes, a nonzero fault opening condition. The adjoint friction law, state evolution equation, and the opening condition include time-dependent coefficients that are functions of slip velocity, state, and normal stress from the forward problem. The frame-

work is presented for dynamic elasticity but may be straightforwardly adapted to quasi-static or quasi-dynamic settings, to be used in earthquake cycle or aseismic slip simulations. We note that a special case of this method has been derived for quasi-dynamic cycle models, for the spatially discretized problem with a boundary element discretization of the linear elastic slip-stress change relation, by Kano et al. [Kano et al. \[2013, 2015, 2020\]](#). We also draw attention to new work exploring the use of physics-informed neural networks (PINNs) for inversion of rate-and-state friction parameters [Rucker and Erickson \[2024\]](#). Additionally, FWI has been applied to seismic source inversion in, e.g., [Sjögreen and Petersson \[2014\]](#), with seismic sources modeled as point moment tensor forcings in the elastic wave equation.

To demonstrate the adjoint-based optimization framework we consider a 2D dynamic rupture problem with a rough fault in antiplane shear, discretized using finite difference methods and explicit Runge–Kutta time stepping methods. High-order finite difference methods satisfying a summation-by-parts (SBP) property have successfully been applied to linear elasticity with rate-and-state friction in e.g., [Kozdon et al. \[2013\]](#), [Erickson and Dunham \[2014\]](#), [O’Reilly et al. \[2015\]](#), [Duru and Dunham \[2016\]](#), [Duru et al. \[2019\]](#), [Harvey et al. \[2023\]](#), and extended to poro- and viscoelasticity in [Torberntsson et al. \[2018\]](#), [Allison and Dunham \[2018\]](#), [Heimisson et al. \[2019\]](#). In this work, we utilize the recently developed boundary-optimized second derivative operators of [Stiernström et al. \[2023\]](#), providing increased accuracy for problems where boundary or interface effects are of particular importance. The operators are combined with the non-stiff interface treatment based on characteristic variables presented in [Erickson et al. \[2022\]](#). The fault interface conditions are imposed weakly, using simultaneous approximation terms (SAT). The resulting spatial SBP-SAT discretization allows for efficient time integration by means of explicit methods. We prove that the SBP-SAT discretization of the misfit functional and forward problem is dual consistent [Pierce and Giles \[2000\]](#), [Hartmann \[2007\]](#), [Berg and Nordström \[2012\]](#), [Hicken and Zingg \[2014\]](#), [Ghasemi Zinatabadi \[2019\]](#), meaning that in deriving the semi-discrete adjoint equations one obtains a consistent discretization of the continuous adjoint problem. As a result, the gradient of the semi-discrete problem is the exact gradient (to machine precision) of the semi-discrete misfit functional as well as a consistent approximation of the continuous gradient. Obtaining an exact gradient to the discrete misfit is beneficial from a numerical point of view, since discretization errors present in the gradient otherwise may negatively affect the convergence of the optimization algorithm [Giles and Pierce \[2000\]](#), especially when simulations are carried out on marginally resolved grids. That the characteristics-based SAT of [Erickson et al. \[2022\]](#), when paired with an SBP discretization, yields a dual-consistent scheme is to our knowledge a new result. We extend dual consistency to the fully discrete problem by employing the standard fourth-order Runge–Kutta time stepping and using the associated Runge–Kutta quadrature to discretize the time integral in the misfit functional [Sanz-Serna \[2016\]](#). While our discretization is based on SBP finite differences, dual consistency should straightforwardly extend to other methods satisfying SBP or corresponding discrete integration-by-parts properties. This includes

discontinuous- and continuous finite element methods as well as finite volume methods.

Our theoretical findings are corroborated by a series of numerical experiments on model problems using synthetic data. First, the discrete adjoint-based gradient is compared to a brute-force finite-difference approximation of the gradient. Second, we perform inversions for frictional parameters in an inverse crime setting, as well as using high-resolution synthetic data, demonstrating the capabilities of the method.

The article is developed as follows: In Section 2 we establish the notation used in the continuous and discrete analysis. In Section 3 the equations of linear elasticity and rate-and-state friction are presented. In Section 4 the inverse problem is formulated as a PDE-constrained optimization problem. The adjoint-based gradient of the misfit functional is derived in Section 5. We demonstrate how filtered and windowed residuals in particle displacement and velocity enter the adjoint equations and discuss well-posedness of the forward and adjoint problem. In Section 6 we present the dual-consistent SBP-SAT discretization for dynamic rupture simulations in antiplane shear and use it to invert synthetic data, demonstrating that the method indeed is capable of reconstructing frictional parameters. Finally, the article is concluded in Section 7.

2 Notation

2.1 Integrals, inner products and index notation

The L^2 -inner product for scalar functions u, v on Ω is denoted

$$(u, v)_\Omega := \int_\Omega u v d\Omega. \quad (2.1)$$

Similarly, define the bilinear form for integrating over a surface $\Gamma \in \partial\Omega$ as

$$(u, v)_\Gamma := \int_\Gamma u v d\Gamma. \quad (2.2)$$

The same notation is used when integrating over a space-time domain, i.e., for the domain $\Omega \times \mathcal{T}$ with $\mathcal{T} = [0, T]$, we write

$$(u, v)_{\Omega \times \mathcal{T}} := \int_0^T (u, v)_\Omega dt, \quad (2.3)$$

and correspondingly for $(u, v)_{\Gamma \times \mathcal{T}}$.

Vector-valued functions are denoted with a bar, e.g., $\bar{u} = [u_I]_{I=1}^d$ where $d = \dim(\Omega)$. Throughout the article, we will make use of index notation. Here the Einstein summation convention is applied to sum over repeated indices I, J, K, L used to denote spatial components. For instance, (2.3) is extended to vector-valued functions \bar{u} and \bar{v} as

$$(u_I, v_I)_{\Omega \times T} := \sum_{I=1}^d (u_I, v_I)_{\Omega \times \mathcal{T}} = (\bar{u}, \bar{v})_{\Omega \times T}. \quad (2.4)$$

Furthermore, we will use the notation $\partial_I := \frac{\partial}{\partial x_I}$, and the summation convention then applies to derivatives. For instance, the variable coefficient Laplace operator is $\partial_I \mu \partial_I$. In a few places we make use of mixed index and vector notation. For instance, $f_I(\bar{u})$, means that the I th component of the vector-valued function \bar{f} is a function of the components of \bar{u} .

2.2 Grid functions and discrete operators

Boldface symbols will be used to denote discrete quantities, e.g., the discrete grid function approximating the function $u(\bar{x})$, $\bar{x} \in \Omega$ is denoted \mathbf{u} and is defined on the grid Ω . Similarly, vector-valued grid functions are denoted with a bar, e.g., the grid function restriction of \bar{u} is denoted $\bar{\mathbf{u}}$. Scalar functions acting as diagonal operators are denoted with double bars, e.g., the operator approximating $f(\bar{x})$ is $\bar{\mathbf{f}} := \text{diag}(\mathbf{f})$.

\mathbf{H} with appropriate subscripts will be used to denote quadrature rules. For instance, a quadrature on Ω is denoted \mathbf{H}_Ω . The notation for inner products is analogous to the continuous setting. For instance

$$(\mathbf{u}, \mathbf{v})_\Omega := \mathbf{u}^T \mathbf{H}_\Omega \mathbf{v}. \quad (2.5)$$

For numerical integration along a boundary grid segment Γ , we first introduce the boundary restriction operators \mathbf{e}_Γ^T , where $\mathbf{e}_\Gamma^T \mathbf{u} \approx u(\bar{x})$, $\bar{x} \in \Gamma$. A discrete bilinear form approximating (2.2) is then defined as

$$(\mathbf{u}, \mathbf{v})_\Gamma := (\mathbf{e}_\Gamma^T \mathbf{u})^T \mathbf{H}_\Gamma (\mathbf{e}_\Gamma^T \mathbf{v}), \quad (2.6)$$

where \mathbf{H}_Γ is a quadrature rule on Γ . A bilinear form over the entire boundary grid $\partial\Omega$ is defined by summing over all boundary grid segments, i.e.,

$$(\mathbf{u}, \mathbf{v})_{\partial\Omega} := \sum_{\Gamma \subset \partial\Omega} (\mathbf{u}, \mathbf{v})_\Gamma. \quad (2.7)$$

In a semi-discrete setting, where time is left continuous, we will also use semi-discrete inner products, e.g., $(\mathbf{u}, \mathbf{v})_{\Omega \times \mathcal{T}} := \int_0^T \mathbf{u}^T \mathbf{H}_\Omega \mathbf{v} dt$. The summation convention for I, J, K, L also applies to discrete relations. For instance, for vector-valued grid functions $\bar{\mathbf{v}}$ and $\bar{\mathbf{u}}$ the extension of (2.5) is

$$(\mathbf{u}_I, \mathbf{v}_I)_\Omega := \sum_{I=1}^d (\mathbf{u}_I, \mathbf{v}_I)_\Omega. \quad (2.8)$$

3 Linear elasticity with rate-and-state friction

Consider a linear elastic medium $\Omega = \Omega_- \cup \Omega_+$, separated by a frictional fault $\Gamma = \Omega_- \cap \Omega_+$. Let \bar{u} denote the displacement vector. From Hooke's law, the changes in the stress tensor σ and traction vector \bar{T} for elastic deformations about a prestressed reference configuration are

$$\sigma_{IJ} := C_{IJKL} \partial_K u_L, \quad (3.1)$$

and

$$T_J := n_I \sigma_{IJ}, \quad (3.2)$$

where \bar{n} denotes the outward unit normal and C_{IJKL} is the elastic stiffness tensor. (We note that these stress changes are sometimes denoted as $\Delta\sigma_{IJ}$.) Henceforth, superscripts $+$ and $-$ will be used to distinguish between fields on the $+$ and $-$ sides of the fault. The change in shear traction on the $-$ side of the fault is the projection of \bar{T}^- onto the fault plane, given by

$$\bar{\tau} := \bar{T}^- - (\bar{n}^- \cdot \bar{T}^-) \bar{n}^-, \quad (3.3)$$

while the change in compressive normal stress is

$$\sigma_n := -\bar{n}^- \cdot \bar{T}^-. \quad (3.4)$$

We denote the the jump in particle velocity across the fault by

$$[\dot{\bar{u}}] := \dot{\bar{u}}^+ - \dot{\bar{u}}^- = V_n \bar{n}^- + \bar{V}, \quad (3.5)$$

where the opening velocity V_n and the slip velocity \bar{V} are defined as the normal component of $[\dot{\bar{u}}]$ and the projection of $[\dot{\bar{u}}]$ onto the fault plane, respectively, i.e.,

$$V_n := \bar{n}^- \cdot [\dot{\bar{u}}], \quad (3.6)$$

$$\bar{V} := [\dot{\bar{u}}] - V_n \bar{n}^-. \quad (3.7)$$

Exploiting linearity of the elasticity equations, the total stress on the fault is the sum of prestress, assumed to be in equilibrium with any external loading, and changes in stress due to the displacement \bar{u} . Denoting the initial compressive normal stress as σ_n^0 , the total compressive normal stress is

$$\sigma_n^{tot} := \sigma_n^0 + \sigma_n. \quad (3.8)$$

The total shear traction vector is

$$\bar{\tau}^{tot} := \bar{\tau}^0 + \bar{\tau} - \eta \bar{V}, \quad (3.9)$$

where $\bar{\tau}^0$ is the initial shear traction and $\eta \bar{V}$ is the radiation damping term. The radiation damping term is only used in quasi-dynamic models, where the coefficient η is the impedance of shear waves radiating away from the fault [Rice \[1993\]](#). In the fully dynamic setting, η is set to zero. Although we only perform experiments with $\eta = 0$ in this paper, we include the radiation damping term in the derivation for completeness.

Force balance across the fault and the condition of no opening or interpenetration of the fault walls are stated as

$$\bar{T}^+ = -\bar{T}^-, \quad (3.10)$$

$$V_n = 0. \quad (3.11)$$

The force balance condition (3.10) is what permits the simplified notation in (3.8) - (3.9), where $\bar{\tau}^{tot}$ and σ_n^{tot} , without superscripts $+$ or $-$, are used to represent values for shear and normal tractions on the two sides of the fault. We also restrict attention to problems where σ_n^{tot} remains compressive. A more general problem formulation would allow for fault opening and the transition to traction-free fault walls when constraining the fault against opening would lead to tensile σ_n^{tot} . This requires enforcing inequality constraints Day et al. [2005]. It is more common in earthquake modeling to always enforce the no opening condition, but then to handle tensile normal stresses by setting σ_n^{tot} to zero when evaluating frictional strength Harris et al. [2009, 2018], Erickson et al. [2023].

Equating shear traction with frictional shear strength yields

$$\bar{\tau}^{tot} = \sigma_n^{tot} f \frac{\bar{V}}{|\bar{V}|}, \quad (3.12)$$

where f is the friction coefficient. In this work, we consider rate- and state-dependent friction coefficients $f(|\bar{V}|, \Psi)$, where the dimensionless state variable Ψ is a measure of the interface contact strength related to the past history of sliding Dieterich [1979], Rice and Ruina [1983], Ruina [1983]. Solving (3.9) and (3.12) for $\bar{\tau}$ and using tangential components of (3.10) leads to the rate-and-state friction law, given by

$$\bar{\tau} = \bar{F}(\bar{V}, \Psi, \sigma_n) := (\sigma_n^0 + \sigma_n) f(|\bar{V}|, \Psi) \frac{\bar{V}}{|\bar{V}|} - \bar{\tau}^0 + \eta \bar{V}. \quad (3.13)$$

Note that $\bar{\tau}$, \bar{V} and \bar{F} are three-component vectors. However, since $\bar{\tau}$ and \bar{V} are coplanar, one of the three equations implied by $\bar{\tau} = \bar{F}$ is $0 = 0$, and the third interface condition is given by the no opening condition (3.11).

The state variable Ψ is in turn governed by a state evolution equation

$$\dot{\Psi} = G(\bar{V}, \Psi, \sigma_n, \dot{\sigma}_n). \quad (3.14)$$

The rate-and-state friction coefficient $f(|\bar{V}|, \Psi)$ and state evolution equation $G(\bar{V}, \Psi, \sigma_n, \dot{\sigma}_n)$ are empirical laws obtained from laboratory experiments Dieterich [1979], Linker and Dieterich [1992], Rice and Ruina [1983], Marone [1998]. They contain parameters which can be difficult to constrain at depths that are inaccessible to drilling. Even if laboratory friction experiments are performed on core samples from drilling, the relation between experimental parameter values and parameter values relevant for large-scale slip remains unclear. Examples of such parameters are the direct effect parameter a and the state evolution parameter b , which we define in a later section. Their difference, $a - b$, determines if friction increases or decreases with slip velocity, which controls the stability of sliding and the possibility of unstable rupture. These parameters may be spatially variable along the fault, but are time-independent.

We are now ready to state the forward problem. Let $\bar{Q}(\bar{x}, t)$ denote external

forcings. The governing equations of the forward problem are given by

$$\begin{aligned}
\rho \ddot{u}_J &= \partial_I C_{IJKL} \partial_K u_L + Q_J, & \bar{x} \in \Omega, & \quad t \in \mathcal{T}, \\
\bar{u} &= \bar{u}_0, \quad \dot{\bar{u}} = \bar{v}_0, & \bar{x} \in \Omega, & \quad t = 0, \\
L\bar{u} &= \bar{g}, & \bar{x} \in \partial\Omega, & \quad t \in \mathcal{T}, \\
V_n &= 0, & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\
\bar{T}^+ &= -\bar{T}^-, & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\
\bar{\tau} &= \bar{F}(\bar{V}, \Psi, \sigma_n), & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\
\dot{\Psi} &= G(\bar{V}, \Psi, \sigma_n, \dot{\sigma}_n), & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\
\Psi &= \Psi_0, & \bar{x} \in \Gamma, & \quad t = 0,
\end{aligned} \tag{3.15}$$

where the first equation is the momentum balance equation with Hooke's law used to replace stress with spatial derivatives of displacement, \bar{u}_0 and \bar{v}_0 are initial data, and $L\bar{u} = \bar{g}(\bar{x}, t)$ denotes boundary conditions on exterior boundaries, for a boundary operator L . The data \bar{u}_0 , \bar{v}_0 , and \bar{g} are assumed to be independent of frictional parameters or initial stress that we later invert for. We will consider L such that for homogeneous boundary data $\bar{g} = \bar{0}$, the boundary conditions are of the form

$$T_J = -\alpha_{JL} \dot{u}_L, \tag{3.16}$$

where $\alpha_{JL} = \alpha_{LJ}$ is positive semi-definite. Note that this includes traction-free conditions ($\alpha_{JL} = 0$), rigid-wall conditions ($\alpha_{JL} \rightarrow \infty$), and characteristic non-reflecting conditions [Petersson and Sjögreen \[2009\]](#). Again, we emphasize that the friction law $\bar{\tau} = \bar{F}$ constitutes two equations, such that there is no overspecification of the fault interface conditions.

4 The inverse problem

Now consider N_{rec} receivers positioned at $\bar{x} = \bar{x}_r^{(k)}$, $k = 1, \dots, N_{rec}$, each with a time series of measurements $\bar{m}_{data}^{(k)}(t)$ of either particle displacements ($\bar{m} = \bar{u}$) or velocities ($\bar{m} = \dot{\bar{u}}$). We seek parameter values that minimize the residual \bar{r} , defined as the difference between model predictions at the receiver locations (obtained by solving (3.15)) and the measured data, i.e.,

$$\bar{r}^{(k)}(t) = \bar{m}(\bar{x}_r^{(k)}, t) - \bar{m}_{data}^{(k)}(t). \tag{4.1}$$

In practice, the residual is often filtered or windowed. We therefore consider applying linear operators $W^{(k)}$ to obtain the adjusted residuals

$$\bar{R}^{(k)}(t) = W^{(k)} \left[\bar{r}^{(k)} \right] (t). \tag{4.2}$$

The notation $W^{(k)} \left[\bar{r}^{(k)} \right]$ means that the linear operator $W^{(k)}$ acts on $\bar{r}^{(k)}$. The result is the time-dependent adjusted residual. We define the misfit functional

\mathcal{F} as the sum of the L^2 -norms of the adjusted residuals:

$$\mathcal{F} = \frac{1}{2} \sum_{k=1}^{N_{rec}} \int_0^T |\bar{R}^{(k)}(t)|^2 dt. \quad (4.3)$$

Weighting of the different terms can be included in the $W^{(k)}$ operator.

For a model parameter p in (3.15), the inverse problem is the PDE-constrained optimization problem given by

$$\min_p \mathcal{F} \text{ subject to (3.15)}. \quad (4.4)$$

In this work, we particularly consider the case where $p(\bar{x})$, $\bar{x} \in \Gamma$ is a spatially variable parameter in the rate-and-state friction coefficient f in (3.13) and/or state evolution equation G in (3.14) (e.g., the direct effect parameter a). Additionally, p may also represent initial stresses $\bar{\tau}^0$ and σ_n^0 in (3.13) or initial state Ψ_0 . To solve (4.4) using gradient-based optimization algorithms, the misfit gradient $\frac{\delta \mathcal{F}}{\delta p}$ is required, where $\frac{\delta}{\delta p}$ denotes the functional derivative with respect to p . However, taking the functional derivative of (4.3) directly requires computing $\frac{\delta \bar{m}}{\delta p}$ (see Section 5.2). In a discrete setting, where \mathbf{p} is represented by N degrees of freedom, even the simplest first-order difference approximation of $\frac{\delta \mathcal{F}}{\delta p}$ would require (3.15) to be solved numerically $N + 1$ times. This is of course not feasible for large-scale 3D computations. We therefore seek an alternative expression through the adjoint-state framework.

5 Adjoint equations and misfit gradient

We begin this section by presenting the adjoint equations and the adjoint-based gradient of the misfit (4.3) in the form of a theorem. To this end, introduce the following adjoint variables:

$$\begin{aligned} \bar{u}^\dagger & \quad (\text{adjoint displacement}), \\ \Psi^\dagger & \quad (\text{adjoint state variable}), \\ V_n^\dagger & := -\bar{n}^- \cdot \llbracket \dot{u}^\dagger \rrbracket \quad (\text{adjoint opening velocity}), \\ \bar{V}^\dagger & := -\llbracket \dot{u}^\dagger \rrbracket - V_n^\dagger \bar{n}^- \quad (\text{adjoint slip velocity}), \\ T_J^\dagger & := n_I C_{IJKL} \partial_K u_L^\dagger \quad (\text{adjoint traction}), \\ \bar{\tau}^\dagger & := \bar{T}^{\dagger-} - (\bar{n}^- \cdot \bar{T}^{\dagger-}) \bar{n}^- \quad (\text{adjoint shear traction}), \\ \sigma_n^\dagger & := -\bar{n}^- \cdot \bar{T}^{\dagger-} \quad (\text{adjoint normal stress}). \end{aligned} \quad (5.1)$$

The negative signs in V_n^\dagger and \bar{V}^\dagger are for notational convenience. As shortly explained, the adjoint problem will later be phrased in reversed time in which case V_n^\dagger and \bar{V}^\dagger are defined analogously to (3.6)-(3.7). Furthermore, let $\bar{Q}^\dagger(\bar{x}, t)$ be the adjoint source term. It consists of the residuals (4.2) acting as point

forces at the receiver locations. For a complete description, see Section 5.2. For now, it is sufficient that \bar{Q}^\dagger satisfies the relation

$$\left(Q_J^\dagger, \frac{\delta \dot{u}_J}{\delta p} \right)_{\Omega \times \mathcal{T}} = \frac{\delta \mathcal{F}}{\delta p}. \quad (5.2)$$

We now define the adjoint equations to the inverse problem (4.4) as

$$\begin{aligned} \rho \ddot{u}_J^\dagger &= \partial_I C_{IJKL} \partial_K u_L^\dagger + Q_J^\dagger, & \bar{x} \in \Omega, & \quad t \in \mathcal{T}, \\ \bar{u}^\dagger &= \bar{0}, \quad \dot{\bar{u}}^\dagger = \bar{0}, & \bar{x} \in \Omega, & \quad t = T, \\ L^\dagger \bar{u}^\dagger &= 0, & \bar{x} \in \partial\Omega, & \quad t \in \mathcal{T}, \\ V_n^\dagger &= H^\dagger(\bar{V}^\dagger, \Psi^\dagger), & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\ \bar{T}^{\dagger+} &= -\bar{T}^{\dagger-}, & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\ \bar{\tau}^\dagger &= \bar{F}^\dagger(\bar{V}^\dagger, \Psi^\dagger), & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\ -\dot{\Psi}^\dagger &= G^\dagger(\bar{V}^\dagger, \Psi^\dagger), & \bar{x} \in \Gamma, & \quad t \in \mathcal{T}, \\ \Psi^\dagger &= 0, & \bar{x} \in \Gamma, & \quad t = T. \end{aligned} \quad (5.3)$$

The functions F^\dagger , G^\dagger , and H^\dagger are linear in their arguments. They govern the fault dynamics of the adjoint problem and are discussed in more detail below. The adjoint problem is equipped with terminal conditions (imposed at time $t = T$) and is naturally solved in reversed time. It is common to introduce the change of variables $t^\dagger = T - t$. We utilize this when discretizing (5.3) (see Section 6.3). Here, for notational convenience, we keep the original time t in the derivation of the adjoint-based gradient. Note that $\frac{\partial}{\partial t} = -\frac{\partial}{\partial t^\dagger}$, and therefore

$$-\dot{\Psi}^\dagger = \frac{\partial \Psi^\dagger}{\partial t^\dagger}, \quad -[\dot{\bar{u}}^\dagger] = \frac{\partial}{\partial t^\dagger} [\bar{u}^\dagger]. \quad (5.4)$$

Notably, the boundary operator is self-adjoint in reversed time, i.e., $L^\dagger = L$, and the adjoint boundary conditions can be formulated as

$$T_J^\dagger = \alpha_{JL} \dot{u}_L^\dagger = -\alpha_{JL} \frac{\partial u_L^\dagger}{\partial t^\dagger}, \quad (5.5)$$

(cf. (3.16)). As mentioned in Section 3, (5.5) includes rigid-wall, traction-free, and characteristic non-reflecting conditions. The relations (5.4) - (5.5) show that the adjoint problem (5.3), when considered in reversed time, is of the same form as the forward problem (3.15).

The adjoint friction law and state evolution equation in (5.3) are given by

$$\tau_J^\dagger = F_J^\dagger(\bar{V}^\dagger, \Psi^\dagger) := \frac{\partial F_I}{\partial V_J} V_I^\dagger + \frac{\partial G}{\partial V_J} \Psi^\dagger, \quad (5.6)$$

and

$$-\dot{\Psi}^\dagger = G^\dagger(\bar{V}^\dagger, \Psi^\dagger) := \frac{\partial F_J}{\partial \Psi} V_J^\dagger + \frac{\partial G}{\partial \Psi} \Psi^\dagger. \quad (5.7)$$

Note that (5.6) - (5.7) resemble the equations of linearized rate-and-state friction. However, the coefficients are time-dependent and are actually functions of the forward variables, e.g. $\frac{\partial F_L}{\partial V_J} = \frac{\partial F_L}{\partial V_J}(\bar{V}, \Psi, \sigma_n)$. Instead of the no opening or interpenetration condition (3.11) in the forward problem, the adjoint problem satisfies an equation for nonzero adjoint opening velocity if the forward problem involves normal stress changes, given by

$$\begin{aligned} V_n^\dagger &= H^\dagger(\bar{V}^\dagger, \Psi^\dagger) \\ &:= \left[\frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial F_J}{\partial \Psi} + \frac{\partial F_J}{\partial \sigma_n} \right] V_J^\dagger + \left[\frac{\partial G}{\partial \sigma_n} + \frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial G}{\partial \Psi} - \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{\sigma}_n} \right) \right] \Psi^\dagger. \end{aligned} \quad (5.8)$$

With the adjoint problem defined, we are ready to state the first major result of this paper.

Theorem 5.1. *Let $\bar{V}^\dagger, \Psi^\dagger$, satisfy (5.3). Further, let p be a parameter in \bar{F} or G , and let Ψ_0 be the initial state. Then, the gradient of the misfit in the inverse problem (4.4) is given by*

$$\begin{aligned} \frac{\delta \mathcal{F}}{\delta p} &= - \left(V_J^\dagger, \frac{\partial F_J}{\partial p} \right)_\tau - \left(\Psi^\dagger, \frac{\partial G}{\partial p} \right)_\tau, \\ \frac{\delta \mathcal{F}}{\delta \Psi_0} &= \Psi_0^\dagger, \end{aligned} \quad (5.9)$$

where $\Psi_0^\dagger = \Psi^\dagger(t=0)$.

Proof. See section 5.1. □

5.1 Derivation of the adjoint-based gradient

To prove Theorem 5.1, we formulate the Lagrangian cost functional

$$\begin{aligned} \mathcal{L} &= \mathcal{F} + \left(\dot{u}_J^\dagger, \rho \ddot{u}_J - \partial_I C_{IJKL} \partial_K u_L - Q_J \right)_{\Omega \times \mathcal{T}} \\ &\quad + \left(\Psi^\dagger, \dot{\Psi} - G \right)_{\Gamma \times \mathcal{T}}, \end{aligned} \quad (5.10)$$

where \bar{u}^\dagger and Ψ^\dagger are solutions to (5.3). Note that for any \bar{u} and Ψ satisfying (3.15), $\mathcal{L} = \mathcal{F}$ and further $\frac{\delta \mathcal{L}}{\delta p} = \frac{\delta \mathcal{F}}{\delta p}$ by the Lagrange multiplier theorem Tröltzsch [2010]. Also note that we are using adjoint velocity \dot{u}^\dagger as a Lagrange multiplier when forming the cost functional. This choice is explained in a remark at the end of this section. Moreover, since initial- and boundary data in (3.15) are independent of p we may set $\bar{u}_0 = 0$, $\bar{v}_0 = 0$ and $\bar{g} = 0$ for the remaining part of the derivation.

Before considering the first variation $\delta \mathcal{L}$, we rewrite the cost functional by performing partial integrations to shift derivatives from \bar{u} to \bar{u}^\dagger . Integrating by parts in time yields

$$\left(\dot{u}_J^\dagger, \rho \ddot{u}_J \right)_{\Omega \times \mathcal{T}} = - \left(\rho \ddot{u}_J^\dagger, \dot{u}_J \right)_{\Omega \times \mathcal{T}}, \quad (5.11)$$

where terms at $t = 0$ and $t = T$ vanish due to the initial and terminal conditions $\dot{u} = 0$ and $\dot{u}^\dagger = 0$ in (3.15) and (5.3). Next, consider the spatial elastic operator and integrate by parts in space to obtain

$$\begin{aligned}
-\left(\dot{u}_J^\dagger, \partial_I C_{IJKL} \partial_K u_L\right)_{\Omega \times \mathcal{T}} &= -\left(\dot{u}_J^\dagger, T_J\right)_{\partial\Omega \times \mathcal{T}} + \left(\dot{T}_J^\dagger, u_J\right)_{\partial\Omega \times \mathcal{T}} \\
&\quad - \left(\dot{u}_J^{\dagger+}, T_J^+\right)_{\Gamma \times \mathcal{T}} + \left(\dot{T}_J^{\dagger+}, u_J^+\right)_{\Gamma \times \mathcal{T}} \\
&\quad - \left(\dot{u}_J^{\dagger-}, T_J^-\right)_{\Gamma \times \mathcal{T}} + \left(\dot{T}_J^{\dagger-}, u_J^-\right)_{\Gamma \times \mathcal{T}} \\
&\quad - \left(\partial_I C_{IJKL} \partial_K \dot{u}_L^\dagger, u_J\right)_{\Omega \times \mathcal{T}},
\end{aligned} \tag{5.12}$$

where we used the major symmetry of the stiffness tensor: $C_{IJKL} = C_{KLIJ}$. Integrating by parts in time leads to

$$\begin{aligned}
\left(\dot{T}_J^\dagger, u_J\right)_{\partial\Omega \times \mathcal{T}} &= -\left(T_J^\dagger, \dot{u}_J\right)_{\partial\Omega \times \mathcal{T}} \\
\left(\dot{T}_J^{\dagger\pm}, u_J^\pm\right)_{\Gamma \times \mathcal{T}} &= -\left(T_J^{\dagger\pm}, \dot{u}_J^\pm\right)_{\Gamma \times \mathcal{T}}
\end{aligned} \tag{5.13}$$

and

$$-\left(\partial_I C_{IJKL} \partial_K \dot{u}_L^\dagger, u_J\right)_{\Omega \times \mathcal{T}} = \left(\partial_I C_{IJKL} \partial_K u_L^\dagger, \dot{u}_J\right)_{\Omega \times \mathcal{T}}, \tag{5.14}$$

where, again, terms at the initial and final times vanish due to the initial and terminal conditions $\bar{u} = 0$, $\bar{u}^\dagger = 0$. We have now derived

$$\begin{aligned}
\left(\dot{u}_J^\dagger, \rho \ddot{u}_J - \partial_I C_{IJKL} \partial_K u_L\right)_{\Omega \times \mathcal{T}} &= -\left(\rho \ddot{u}_J^\dagger - \partial_I C_{IJKL} \partial_K u_L^\dagger, \dot{u}_J\right)_{\Omega \times \mathcal{T}} \\
&\quad - \underbrace{\left(\dot{u}_J^\dagger, T_J\right)_{\partial\Omega \times \mathcal{T}} - \left(T_J^\dagger, \dot{u}_J\right)_{\partial\Omega \times \mathcal{T}}}_{BT_{ext}} \\
&\quad - \underbrace{\left(\dot{u}_J^{\dagger+}, T_J^+\right)_{\Gamma \times \mathcal{T}} - \left(T_J^{\dagger+}, \dot{u}_J^+\right)_{\Gamma \times \mathcal{T}}}_{BT_{falt}^+} \\
&\quad - \underbrace{\left(\dot{u}_J^{\dagger-}, T_J^-\right)_{\Gamma \times \mathcal{T}} - \left(T_J^{\dagger-}, \dot{u}_J^-\right)_{\Gamma \times \mathcal{T}}}_{BT_{falt}^-}.
\end{aligned} \tag{5.15}$$

To simplify this expression, we first note that the exterior boundary terms BT_{ext} vanish. To see this, we use that the forward and adjoint boundary conditions satisfy (3.16) and (5.5), and $\alpha_{JL} = \alpha_{LJ}$, such that

$$\begin{aligned}
BT_{ext} &= \left(\dot{u}_J^\dagger, \alpha_{JL} \dot{u}_L\right)_{\partial\Omega \times \mathcal{T}} - \left(T_J^\dagger, \dot{u}_J\right)_{\partial\Omega \times \mathcal{T}} \\
&= \left(\alpha_{JL} \dot{u}_L^\dagger - T_J^\dagger, \dot{u}_J\right)_{\partial\Omega \times \mathcal{T}} = 0.
\end{aligned} \tag{5.16}$$

This naturally holds for self-adjoint boundary operators $L = L^\dagger$. Second, for the fault terms BT_{fault}^\pm , we use the forward and adjoint force balances, $\bar{T}^+ = -\bar{T}^-$ and $\bar{T}^{\dagger+} = -\bar{T}^{\dagger-}$, to obtain

$$BT_{fault}^+ + BT_{fault}^- = \left(\llbracket \dot{u}_J^\dagger \rrbracket, T_J^- \right)_{\Gamma \times \mathcal{T}} + \left(T_J^{\dagger-}, \llbracket \dot{u}_J \rrbracket \right)_{\Gamma \times \mathcal{T}}. \quad (5.17)$$

Next, we decompose tractions and jumps in particle velocity into fault-parallel and fault-normal components:

$$\begin{aligned} \llbracket \dot{u}_J \rrbracket &= V_J + V_n n_J^-, & T_J^- &= \tau_J - \sigma_n n_J^-, \\ -\llbracket \dot{u}_J^\dagger \rrbracket &= V_J^\dagger + V_n^\dagger n_J^-, & T_J^{\dagger-} &= \tau_J^\dagger - \sigma_n^\dagger n_J^-. \end{aligned} \quad (5.18)$$

Using the decomposition, together with the forward and adjoint friction laws $\tau_J = F_J$ and $\tau_J^\dagger = F_J^\dagger$ and the no opening condition $V_n = 0$, allows us to write

$$BT_{fault}^+ + BT_{fault}^- = - \left(V_J^\dagger, F_J \right)_{\Gamma \times \mathcal{T}} + \left(F_J^\dagger, V_J \right)_{\Gamma \times \mathcal{T}} + (V_n^\dagger, \sigma_n)_{\Gamma \times \mathcal{T}}. \quad (5.19)$$

The last term in the Lagrangian cost functional (5.10) containing the state evolution equation can be rewritten, using integration by parts in time,

$$\left(\Psi^\dagger, \dot{\Psi} - G \right)_{\Gamma \times \mathcal{T}} = - \left(\dot{\Psi}^\dagger, \Psi \right)_{\Gamma \times \mathcal{T}} - \left(\Psi^\dagger, G \right)_{\Gamma \times \mathcal{T}} + \left(\Psi_0^\dagger, \Psi_0 \right)_\Gamma, \quad (5.20)$$

where Ψ_0^\dagger, Ψ_0 are the state variables at $t = 0$. The corresponding term at $t = T$ vanishes due to the terminal condition $\Psi(t = T) = 0$ in (5.3). We have derived

$$\begin{aligned} \mathcal{L} &= \mathcal{F} - \left(\rho \ddot{u}_J^\dagger - \partial_I C_{IJKL} \partial_K u_L^\dagger, \dot{u}_J \right)_{\Omega \times \mathcal{T}} - \left(\dot{u}_J^\dagger, Q_J \right)_{\Omega \times \mathcal{T}} \\ &\quad - \left(V_J^\dagger, F_J \right)_{\Gamma \times \mathcal{T}} + \left(F_J^\dagger, V_J \right)_{\Gamma \times \mathcal{T}} + (V_n^\dagger, \sigma_n)_{\Gamma \times \mathcal{T}} \\ &\quad - \left(\dot{\Psi}^\dagger, \Psi \right)_{\Gamma \times \mathcal{T}} - \left(\Psi^\dagger, G \right)_{\Gamma \times \mathcal{T}} + \left(\Psi_0^\dagger, \Psi_0 \right)_\Gamma. \end{aligned} \quad (5.21)$$

We are now ready to derive the first variation of \mathcal{L} . First, $\delta Q_J = 0$ since external forces do not depend on frictional parameters or initial stresses. Second, by the chain rule,

$$\delta F_J = \frac{\partial F_J}{\partial V_I} \delta V_I + \frac{\partial F_J}{\partial \Psi} \delta \Psi + \frac{\partial F_J}{\partial \sigma_n} \delta \sigma_n + \frac{\partial F_J}{\partial p} \delta p, \quad (5.22)$$

$$\delta G = \frac{\partial G}{\partial V_I} \delta V_I + \frac{\partial G}{\partial \Psi} \delta \Psi + \frac{\partial G}{\partial \sigma_n} \delta \sigma_n + \frac{\partial G}{\partial \dot{\sigma}_n} \delta \dot{\sigma}_n + \frac{\partial G}{\partial p} \delta p. \quad (5.23)$$

The above considerations together with the relation (5.2) lead to

$$\begin{aligned}
\delta\mathcal{L} = & - \left(\rho \ddot{u}_J^\dagger - \partial_I C_{IJKL} \partial_K u_L^\dagger - Q_J^\dagger, \delta \dot{u}_J \right)_{\Omega \times \mathcal{T}} \\
& - \left(V_J^\dagger, \frac{\partial F_J}{\partial V_I} \delta V_I + \frac{\partial F_J}{\partial \Psi} \delta \Psi + \frac{\partial F_J}{\partial \sigma_n} \delta \sigma_n + \frac{\partial F_J}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} \\
& + \left(F_J^\dagger, \delta V_J \right)_{\Gamma \times \mathcal{T}} + \left(V_n^\dagger, \delta \sigma_n \right)_{\Gamma \times \mathcal{T}} \\
& - \left(\dot{\Psi}^\dagger, \delta \Psi \right)_{\Gamma \times \mathcal{T}} - \left(\Psi^\dagger, \frac{\partial G}{\partial V_I} \delta V_I + \frac{\partial G}{\partial \Psi} \delta \Psi + \frac{\partial G}{\partial \sigma_n} \delta \sigma_n + \frac{\partial G}{\partial \dot{\sigma}_n} \delta \dot{\sigma}_n + \frac{\partial G}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} \\
& + \left(\Psi_0^\dagger, \delta \Psi_0 \right)_\Gamma.
\end{aligned} \tag{5.24}$$

For the term involving $\delta \dot{\sigma}_n$, integrating by parts in time yields

$$- \left(\Psi^\dagger, \frac{\partial G}{\partial \dot{\sigma}_n} \delta \dot{\sigma}_n \right)_{\Gamma \times \mathcal{T}} = \left(\frac{\partial G}{\partial \dot{\sigma}_n} \dot{\Psi}^\dagger + \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{\sigma}_n} \right) \Psi^\dagger, \delta \sigma_n \right)_{\Gamma \times \mathcal{T}}, \tag{5.25}$$

where we used that $\Psi^\dagger(t=T)=0$ and $\delta \sigma_n(t=0)=0$ due to the terminal and initial conditions. Using the adjoint state evolution equation (5.7) to substitute $\dot{\Psi}^\dagger$ yields

$$\begin{aligned}
& - \left(\Psi^\dagger, \frac{\partial G}{\partial \dot{\sigma}_n} \delta \dot{\sigma}_n \right)_{\Gamma \times \mathcal{T}} = \\
& - \left(\frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial F_J}{\partial \Psi} V_J^\dagger + \left[\frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial G}{\partial \Psi} - \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{\sigma}_n} \right) \right] \Psi^\dagger, \delta \sigma_n \right)_{\Gamma \times \mathcal{T}}.
\end{aligned} \tag{5.26}$$

The volume term in (5.24) vanishes due to the adjoint momentum balance in (5.3). Gathering the remaining terms leads to

$$\begin{aligned}
\delta\mathcal{L} = & \left(F_J^\dagger - \frac{\partial F_I}{\partial V_J} V_I^\dagger - \frac{\partial G}{\partial V_J} \Psi^\dagger, \delta V_J \right)_{\Gamma \times \mathcal{T}} \\
& - \left(\dot{\Psi}^\dagger + \frac{\partial F_J}{\partial \Psi} V_J^\dagger + \frac{\partial G}{\partial \Psi} \Psi^\dagger, \delta \Psi \right)_{\Gamma \times \mathcal{T}} \\
& + \left(V_n^\dagger - \left[\frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial F_J}{\partial \Psi} + \frac{\partial F_J}{\partial \sigma_n} \right] V_J^\dagger - \left[\frac{\partial G}{\partial \sigma_n} + \frac{\partial G}{\partial \dot{\sigma}_n} \frac{\partial G}{\partial \Psi} - \frac{d}{dt} \left(\frac{\partial G}{\partial \dot{\sigma}_n} \right) \right] \Psi^\dagger, \delta \sigma_n \right)_{\Gamma \times \mathcal{T}} \\
& - \left(\Psi^\dagger, \frac{\partial G}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} - \left(V_J^\dagger, \frac{\partial F_J}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} \\
& + \left(\Psi_0^\dagger, \delta \Psi_0 \right)_\Gamma.
\end{aligned} \tag{5.27}$$

The first two terms vanish due to the adjoint friction law (5.6) and state evolution equation (5.7), while the third term vanishes due to the adjoint opening equation (5.8). We have arrived at

$$\delta\mathcal{L} = - \left(\Psi^\dagger, \frac{\partial G}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} - \left(V_J^\dagger, \frac{\partial F_J}{\partial p} \delta p \right)_{\Gamma \times \mathcal{T}} + \left(\Psi_0^\dagger, \delta \Psi_0 \right)_\Gamma. \tag{5.28}$$

Since $\frac{\delta \mathcal{F}}{\delta p} = \frac{\delta \mathcal{L}}{\delta p}$ and $\frac{\delta}{\delta p} (v, w \delta p)_{\Gamma \times \mathcal{T}} = (v, w)_{\mathcal{T}}$, (5.9) follows. This proves Theorem 5.1.

Remark. Note that in the Lagrangian cost functional (5.10) we have taken the inner product with adjoint particle velocity $\dot{\bar{u}}^\dagger$, rather than particle displacement \bar{u}^\dagger . This choice significantly simplifies the adjoint friction and state evolution equations. If one instead were to use adjoint displacement, the equations would, in addition to the current terms, depend also on forward particle accelerations, adjoint slip and adjoint state rate.

Remark. Interestingly, the adjoint rate-and-state equations and the adjoint-based gradient do not depend on the adjoint normal stress σ_n^\dagger . This is due to the no opening condition (3.11). If the fault in the forward problem instead satisfied an opening condition $V_n = H$, with H a function of a forward variable and/or a frictional parameter, then σ_n^\dagger would appear in the adjoint rate-and-state equations and/or the gradient.

5.2 Misfits and adjoint sources

This section details how misfits of particle displacement and velocity enter the adjoint problem (5.3), i.e., how the different types of misfits translate to the adjoint source term \bar{Q}^\dagger .

Consider the adjusted residuals $\bar{R}^{(k)}$ in (4.2). We will refer to the linear operators $W^{(k)}$ as filters, although they could equally well be windowing operators. It is understood that the $W^{(k)}$ may be different for each component of $\bar{r}^{(k)}$, but to simplify the presentation we avoid explicit notation for this. Furthermore, we first consider a single receiver at $\bar{x} = \bar{x}_r$ and will temporarily drop the (k) superscript. Since W is linear, the first variation of R_J can be expressed as

$$\delta R_J = W[\delta r_J] = \int_{\Omega} W[\delta m_J] \hat{\delta}_{\bar{x}_r} d\bar{x}, \quad (5.29)$$

where $\hat{\delta}_{\bar{x}_r}(\bar{x}) := \hat{\delta}(\bar{x} - \bar{x}_r)$ is the shifted Dirac delta function. Let W^\dagger denote the Hilbert adjoint of W with respect to the L^2 -inner product, such that

$$\int_0^T q(t) W[s](t) dt = \int_0^T W^\dagger[q](t) s(t) dt \quad \forall q, s \in L^2(\mathcal{T}). \quad (5.30)$$

Note that windowing operators of the form $W[r](t) = w(t)r(t)$, for some function w , are self-adjoint ($W^\dagger = W$), while a general filter might not be self-adjoint.

Now consider the first variation δF of the misfit functional (4.3) with a single

receiver. Carrying out the chain rule for the first variation yields

$$\begin{aligned}\delta\mathcal{F} &= \int_0^T R_J \delta R_J dt = \int_0^T \int_{\Omega} R_J W[\delta m_J] \hat{\delta}_{\bar{x}_r} d\bar{x} dt \\ &= \int_0^T \int_{\Omega} W^\dagger[R_J] \hat{\delta}_{\bar{x}_r} \delta m_J d\bar{x} dt = \left(W^\dagger[R_J] \hat{\delta}_{\bar{x}_r}, \delta m_J \right)_{\Omega \times \mathcal{T}}.\end{aligned}\tag{5.31}$$

For a velocity misfit where $\bar{m} = \dot{u}$, we note that $Q_J^\dagger = W^\dagger[R_J] \hat{\delta}_{\bar{x}_r}$ satisfies (5.2), which was used to derive the adjoint-based gradient. In the case of a displacement misfit, i.e., $\bar{m} = \bar{u}$, introduce the time-integrated quantity

$$\hat{R}_J(t) = \int_0^t W^\dagger[R_J](t') dt' + \hat{R}_{J,0},\tag{5.32}$$

where the constant $\hat{R}_{J,0}$ is selected such that $\hat{R}_J(T) = 0$. By (5.32), \hat{R}_J satisfies the ODE

$$\dot{\hat{R}}_J = W^\dagger[R_J].\tag{5.33}$$

Using integration by parts in time, it therefore follows that

$$\delta\mathcal{F} = \left(W^\dagger[R_J] \hat{\delta}_{\bar{x}_r}, \delta u_J \right)_{\Omega \times \mathcal{T}} = - \left(\hat{R}_J \hat{\delta}_{\bar{x}_r}, \delta \dot{u}_J \right)_{\Omega \times \mathcal{T}},\tag{5.34}$$

where the term at $t = 0$ vanishes because of the initial condition $u_J = 0$ (or in general the p -independence of the initial data) and the term at $t = T$ vanishes because $\hat{R}_J(T) = 0$. Considering multiple receivers again, the adjoint source function at receiver k is thus

$$S_J^{(k)}(t) = \begin{cases} -\hat{R}_J^{(k)}(t) & m_J = u_J \quad (\text{displacement}), \\ W^{(k)\dagger}[R_J^{(k)}](t) & m_J = \dot{u}_J \quad (\text{velocity}). \end{cases}\tag{5.35}$$

Note that both types of signals involve the application of an adjoint filter operator. The complete adjoint source term is obtained by summing over all receivers,

$$Q_J^\dagger(\bar{x}, t) = \sum_{k=1}^{N_{rec}} S_J^{(k)}(t) \hat{\delta}(\bar{x} - \bar{x}_r^{(k)}).\tag{5.36}$$

In this work, the misfit functionals are considered to measure either displacement or velocity exclusively. It is straightforward to extend the derivation to a linear combination of displacement and velocity misfits. We briefly comment on extending the misfit to other types of measurements. Measurements to consider include strain rate from fiber optics cables, pressure from hydrophones, or snapshot measurements of the displacement of Earth's surface through InSAR. If the resulting misfit can be cast in the form of (5.2), the expression for the adjoint-based gradient in Theorem 5.1 follows directly. The difference will only be in the adjoint source term. In the case of an InSAR measurement, we anticipate that the adjoint source term consists of a surface forcing in space, multiplied by a Heaviside function in time activating at the snapshot time.

5.3 Well-posedness

For Theorem 5.1 to be valid, both the forward and adjoint problems need to be well-posed. A problem is well-posed if a solution a) exists, b) is unique, and c) depends continuously on boundary and initial data (also referred to as stability) Gustafsson et al. [2013]. Well-posedness results for linear elastodynamics with rate-and-state friction are incomplete, and we provide a brief review of known results below. For the adjoint problem, well-posedness is tied to that of the linearized forward problem, i.e., the adjoint problem is well-posed if and only if the linearized forward problem is Bui-Thanh [2023].

For linear initial-boundary value problems, conditions for well-posedness may be established using the energy method, in which an energy estimate for the problem is derived. By prescribing appropriate boundary and interface conditions, the energy estimate shows that the solution is bounded in terms of data, ensuring stability Gustafsson et al. [2013]. Uniqueness can then be established by applying the the energy method to the difference of two solutions. Existence further requires that a minimal number of initial, boundary, and interface conditions are specified Nordström [2013]. For the forward problem (3.15), where non-linearity is limited to the friction law, the energy method leads to conditions on the friction law, required for stability and uniqueness, as discussed below.

Energy estimates for problems with general friction laws were established decades ago Kostrov [1974], Rudnicki and Freund [1981]. Dissipation (rather than creation) of mechanical energy during slip is guaranteed when the frictional shear traction is coplanar with and opposite to slip velocity and the normal stress remains compressive. In Kozdon et al. [2012], the energy method was applied to 2D antiplane shear problems with purely velocity-dependent nonlinear friction and no prestress. For scalar friction laws of the form $\tau = F(V)$, the stability condition of frictional energy dissipation reads $F(V)V \geq 0$. Additionally, the authors showed that a sufficient condition for uniqueness is $\frac{dF}{dV} \geq 0$. In Kozdon et al. [2010], the results were extended to full 3D elastodynamics assuming that fault normal stress remains compressive. Here, the stability condition reads $F_J V_J \geq 0$ and the sufficient uniqueness condition is that the Jacobian of the friction law is symmetric-positive definite, i.e., $\frac{\partial F_I}{\partial V_J} = \frac{\partial F_J}{\partial V_I}$, and $V_I \frac{\partial F_I}{\partial V_I} V_J \geq 0$. Moreover, existence was shown using the method of characteristics, thereby proving well-posedness for this particular class of problems with purely velocity-dependent friction and no prestress.

In Duru et al. [2019], the analysis of 2D antiplane shear problems was extended to rate-and-state friction with prestress. In addition to the previously mentioned constraints on F , it was further established that well-posedness requires that the state evolution law $G(V, \Psi)$ is such that $\frac{\partial G}{\partial \Psi}$ is bounded and such evolution laws were termed admissible¹. Well-posedness results for general 3D elastodynamics with rate-and-state friction are to the best of our knowledge currently lacking.

There have been many studies of linearized rate-and-state friction, as sum-

¹It suffices that $\frac{\partial G}{\partial \Psi}$ is bounded almost everywhere, i.e., $G(V, \Psi)$ is Lipschitz continuous in Ψ .

marized by [Rice et al. \[2001\]](#). With a few exceptions [Ray and Viesca \[2017\]](#), [Viesca \[2023\]](#), these studies have been limited to linearization about a state of steady sliding on a planar fault with constant frictional parameters (making this a constant coefficient problem). Sliding occurs between two homogeneous elastic half-spaces though some studies have considered other (e.g., layered) geometries [Ranjith \[2009, 2014\]](#), [Aldam et al. \[2016\]](#). Of particular note are so-called bimaterial problems, where sliding occurs between dissimilar elastic solids. Spatially nonuniform sliding alters the fault normal stress and therefore frictional strength. This feedback can destabilize steady sliding. Compromises to the full rate-and-state formulation, such as removing the direct effect so that the friction coefficient satisfies $\frac{\partial f}{\partial |V|} = 0$, are known to render the problem ill-posed [Ranjith and Rice \[2001\]](#), [Rice et al. \[2001\]](#).

To summarize, the adjoint problem (5.3) is well-posed if and only if the forward problem (3.15) with linearized rate-and-state friction is well-posed. The latter is determined by the stability of the linearized rate-and-state friction laws, for which well-posedness results are currently incomplete. Studying the stability of linearized rate-and-state friction is thus of interest also for adjoint-based optimization. Finally, we note that the 2D antiplane shear numerical studies presented herein are performed using the typical rate-and-state friction coefficient (6.2) together with the slip law (6.4). No blow-ups of the adjoint solution have been observed, and the results therefore indicate that the adjoint equations are well-posed, at least in this setting.

6 Dynamic rupture in antiplane shear

We now proceed to demonstrate the method applied to dynamic rupture simulations in 2D antiplane shear. Consider the domain Ω with a frictional fault along Γ , as illustrated in Figure 6.1. In the semi-discrete setting, the discretizations will be carried out blockwise in Ω_{\pm} . For this reason, quantities on the respective domain will be denoted with the subscripts $+$ and $-$, or compactly when the equations apply to both domains individually by \pm . Note that this use of $+$ and $-$ differs slightly from the continuous setting in Sections 3 and 5 where it was used only when referring to quantities on the fault; here it is used for both the entire volumes including boundaries and the fault. The notation should be clear from context. Letting $(x_1, x_2) = (x, y) \in \Omega$ be the in-plane coordinates, the only nonzero component of particle displacement, $u = u_z$, is in the antiplane direction $x_3 = z$, with all fields being functions only of x and y . Slip, slip velocity, and shear traction are scalars, and the fault normal stress is

unaltered by slip. In this setting, the governing equations (3.15) reduce to

$$\begin{aligned}
\rho_{\pm} \ddot{u}_{\pm} &= \partial_I \mu_{\pm} \partial_I u_{\pm}, & \bar{x} \in \Omega_{\pm}, & & t \in \mathcal{T}, \\
u_{\pm} &= u_{0\pm}, \quad \dot{u}_{\pm} = v_{0\pm}, & \bar{x} \in \Omega_{\pm}, & & t = 0, \\
\tau_{\pm} + Z \dot{u}_{\pm} &= 0, & \bar{x} \in \partial\Omega_{\pm} \setminus \Gamma, & & t \in \mathcal{T}, \\
\tau_{\pm} &= \mp F(V, \Psi), & \bar{x} \in \Gamma, & & t \in \mathcal{T}, \\
\dot{\Psi} &= G(V, \Psi), & \bar{x} \in \Gamma, & & t \in \mathcal{T}, \\
\Psi &= \Psi_0, & \bar{x} \in \Gamma, & & t = 0.
\end{aligned} \tag{6.1}$$

Here, we use $\tau_{\pm} = \mu_{\pm} \partial_{n_{\pm}} u_{\pm}$ to denote shear traction on the boundaries of Ω_{\pm} as well as the two sides of the fault, while in Section 3 $\bar{\tau}$ was only defined for the $-$ side of the fault. The equation $\tau_{\pm} = \mp F(V, \Psi)$ therefore specifies both the friction law and force balance on the fault. On the exterior boundaries $\partial\Omega_{\pm} \setminus \Gamma$ characteristic non-reflecting conditions Engquist and Majda [1977] are imposed. On the fault, we consider a typical rate-and-state friction coefficient Rice et al. [2001]

$$f(|V|, \Psi) = a \sinh^{-1} \left(\frac{|V|}{2V_0} e^{\frac{\Psi}{a}} \right), \tag{6.2}$$

where a is the dimensionless direct effect parameter and V_0 an arbitrarily chosen reference velocity. The direct effect parameter is positive ($a > 0$), which ensures that the instantaneous response of the fault to perturbations is of velocity-strengthening character ($\partial F / \partial V > 0$). This is well supported by experiments and also expected on a theoretical basis Rice et al. [2001]. Furthermore, some studies referenced in Section 5.3 suggest it to be a requirement for well-posedness. In addition, an external loading $\tau_L(\bar{x})$, $\bar{x} \in \Gamma$ is added to the fault to initiate the rupture at a pre-specified location, such that the friction law reads

$$F(V, \Psi) = \sigma_n^0 f(|V|, \Psi) \frac{V}{|V|} - \tau^0 - \tau_L. \tag{6.3}$$

The external loading could of course be included directly in τ^0 , but is kept separate here since we later will perform inversions for τ^0 with a known τ_L . State evolution is governed by the slip law Marone [1998], Lapusta et al. [2000]

$$G(V, \Psi) = -\frac{|V|}{D_c} (f(|V|, \Psi) - f_{ss}(|V|)), \tag{6.4}$$

where $f_{ss}(|V|) = f_0 + (a - b) \ln(|V|/V_0)$ is the steady state friction coefficient Rice et al. [2001]. Here b is the state evolution parameter, D_c the state evolution distance and f_0 is the reference coefficient for steady sliding at velocity V_0 . The fault is velocity-strengthening where $a - b > 0$ and velocity-weakening where $a - b < 0$; the latter is required for unstable slip and sustained rupture propagation Marone [1998].

In the following sections, we introduce a dual consistent Hicken and Zingg [2014] space-time discretization based on SBP difference operators combined with weakly enforced boundary and interface conditions through characteristics-based SAT and explicit Runge–Kutta time integration. Based on this discretization we present a discrete counterpart to Theorem 5.1 for dynamic rupture in

antiplane shear. While the discretization of the forward problem was developed previously in [Duru et al. \[2019\]](#), [Erickson et al. \[2022\]](#), [Harvey et al. \[2023\]](#), [Stiernström et al. \[2023\]](#), its space-time dual consistency is to the best of our knowledge a novel result.

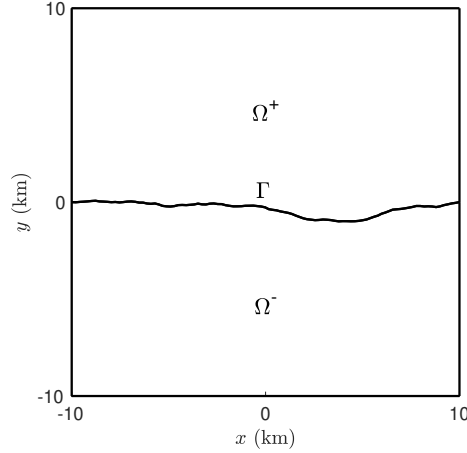


Figure 6.1: Domain with the fault Γ separating Ω^+ and Ω^-

6.1 SBP finite difference discretization

Discretizations of (6.1) using SBP finite differences, with weak enforcement of fault and boundary conditions through the simultaneous-approximation-term (SAT) method have been presented in e.g. [Erickson and Dunham \[2014\]](#), [Duru et al. \[2019\]](#), [Erickson et al. \[2022\]](#), [Harvey et al. \[2023\]](#). Here, we utilize the boundary-optimized SBP difference operators presented in [Mattsson et al. \[2018\]](#), [Stiernström et al. \[2023\]](#) to discretize the scalar wave operator on the physical grid Ω . The physical grid is obtained through transfinite interpolation of the coordinate mapping between Ω and a two-block Cartesian reference domain. See e.g. [Almquist and Dunham \[2020, 2021\]](#), [Stiernström et al. \[2023\]](#) for further details on the finite difference discretization of the scalar wave operator and the discretization of Ω .

Let $\mathbf{D}_{II}(\mu_{\pm}) \approx \partial_I \mu_{\pm} \partial_I$ be the SBP finite difference approximation of the variable coefficient Laplace operator on the physical domains Ω_{\pm} . Then, the discretization of the scalar wave equation in (6.1) is given by

$$\begin{aligned} \bar{\rho}_{\pm} \ddot{\mathbf{u}}_{\pm} &= \mathbf{D}_{II}(\mu_{\pm}) \mathbf{u}_{\pm} + \text{SAT}_{\pm}, & t \in \mathcal{T}, \\ \mathbf{u}_{\pm} &= \mathbf{u}_{0\pm}, \quad \dot{\mathbf{u}}_{\pm} = \mathbf{v}_{0\pm}, & t = 0, \end{aligned} \quad (6.5)$$

where SAT_{\pm} weakly enforces boundary and interface conditions. Imposing the non-linear rate-and-state friction law through a standard traction SAT (e.g. as in [Almquist and Dunham \[2020\]](#)) may lead to significant stiffness in dynamic rupture problems. To resolve the issue, we use the non-stiff SAT presented

in [Erickson et al. \[2022\]](#), [Harvey et al. \[2023\]](#), enforcing boundary and interface conditions through characteristic variables. Since the characteristics-based non-stiff SAT is a recent development for second-order wave equations it is described in detail below, for the sake of completeness.

Consider first a single block Ω (either Ω_+ or Ω_-). On a boundary segment ζ (either an exterior boundary or the fault interface), the non-stiff SAT is given by

$$\text{SAT}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*) = \mathbf{H}_\Omega^{-1} (\mathbf{e}\mathbf{H}_\zeta(\boldsymbol{\tau}^* - \boldsymbol{\tau}) - \mathbf{T}^T \mathbf{H}_\zeta(\mathbf{u}^* - \mathbf{u})), \quad (6.6)$$

where \mathbf{u} , \mathbf{u}^* , and $\boldsymbol{\tau}^*$ all are evaluated on ζ . Here, \mathbf{H}_Ω and \mathbf{H}_Γ are the SBP quadratures on Ω and ζ , while \mathbf{e}^T is the boundary restriction operator. Moreover, \mathbf{T} is the boundary traction operator such that $\boldsymbol{\tau} := \mathbf{T}\mathbf{u} \approx \mu n_I \partial_I u(\bar{x})$, $\bar{x} \in \zeta$, while \mathbf{u}^* and $\boldsymbol{\tau}^*$ are target values (or fluxes) for displacement and traction. The target values are required to preserve the outgoing grid-based characteristic variables,

$$\mathbf{w}^* := \mathbf{Z}\dot{\mathbf{u}}^* - \boldsymbol{\tau}^* = \mathbf{Z}\dot{\mathbf{u}} - \tilde{\boldsymbol{\tau}} =: \mathbf{w}, \quad (6.7)$$

where $\mathbf{Z} = \sqrt{\rho\mu}$ is the shear impedance and $\tilde{\boldsymbol{\tau}}$ is the modified traction

$$\tilde{\boldsymbol{\tau}} = \boldsymbol{\tau} + \gamma(\mathbf{u}^* - \mathbf{u}). \quad (6.8)$$

The parameter γ is the semi-bounded Dirichlet penalty parameter of [Almqvist and Dunham \[2020\]](#). Note that the units of γ are such that $\gamma(\mathbf{u}^* - \mathbf{u})$ has the units of traction. Moreover, \mathbf{u}^* and $\boldsymbol{\tau}^*$ should satisfy the boundary or fault conditions.

On an exterior boundary, the boundary conditions are conveniently expressed by introducing a reflection coefficient \mathbf{R} , such that we may write the ingoing characteristic as $\mathbf{Z}\dot{\mathbf{u}}^* + \boldsymbol{\tau}^* = \mathbf{R}\mathbf{w}^*$. This together with (6.7) results in

$$\boldsymbol{\tau}^* = \frac{\mathbf{R} - \mathbf{I}}{2}(\mathbf{Z}\dot{\mathbf{u}} - \tilde{\boldsymbol{\tau}}), \quad (6.9)$$

$$\dot{\mathbf{u}}^* = \frac{\mathbf{R} + \mathbf{I}}{2}(\dot{\mathbf{u}} - \mathbf{Z}^{-1}\tilde{\boldsymbol{\tau}}). \quad (6.10)$$

Here \mathbf{I} is an identity matrix with dimensions matching the number of boundary points. Then, rigid-wall (Dirichlet), traction-free (Neumann) or characteristic non-reflecting boundary conditions are imposed by setting

$$\begin{aligned} \mathbf{R} &= -\mathbf{I} & (\text{rigid-wall}), \\ \mathbf{R} &= \mathbf{I} & (\text{traction-free}), \\ \mathbf{R} &= \mathbf{0} & (\text{non-reflecting}), \end{aligned} \quad (6.11)$$

along the corresponding boundary segment. On the outer boundaries, characteristic non-reflecting conditions are specified, and we therefore set $\mathbf{R} = \mathbf{0}$.

On the fault, the target tractions instead satisfy the fault condition

$$\boldsymbol{\tau}_\pm^* = \mp \mathbf{F}(\mathbf{V}^*, \boldsymbol{\Psi}), \quad (6.12)$$

where \mathbf{F} is considered to apply point-wise. From (6.7) and (6.12) the target slip velocity \mathbf{V}^* can be shown to satisfy

$$\kappa \mathbf{V}^* + \mathbf{F}(\mathbf{V}^*, \Psi) = -\tau_\ell, \quad (6.13)$$

where $\kappa = \frac{\mathbf{Z}_+ \mathbf{Z}_-}{\mathbf{Z}_+ + \mathbf{Z}_-}$ and $\tau_\ell = \frac{\mathbf{Z}_+ \mathbf{w}_- - \mathbf{Z}_- \mathbf{w}_+}{\mathbf{Z}_+ + \mathbf{Z}_-}$. For non-linear F , solving (6.13) requires employing a root-finding algorithm. Note that Ψ is held fixed while solving (6.13). Let $\tilde{\mathbf{V}}^*(\tau_\ell)$ denote the solution to

$$\mathbf{F}(\tilde{\mathbf{V}}^*, \Psi) = -\tau_\ell, \quad (6.14)$$

which can be solved analytically with our friction coefficient (6.2). As noted in Erickson et al. [2022] one can show that $\mathbf{V}^* \in [\tilde{\mathbf{V}}^*(\tau_\ell), 0]$ for $\tau_\ell \geq \tau^0 + \tau_L$ and $\mathbf{V}^* \in [0, \tilde{\mathbf{V}}^*(\tau_\ell)]$ for $\tau_\ell < \tau^0 + \tau_L$. Since $F_V > 0$ for admissible friction laws, there exists a unique solution \mathbf{V}^* .

From (6.7) it directly follows that $\dot{\mathbf{u}}_\pm^*$ on the fault is given by

$$\dot{\mathbf{u}}_\pm^* = \dot{\mathbf{u}}_\pm - \mathbf{Z}_\pm^{-1}(\tilde{\tau}_\pm - \tau_\pm^*). \quad (6.15)$$

Note that the characteristic boundary and fault conditions require the target displacements \mathbf{u}_\pm^* to be regarded as additional unknown grid functions, evolved together with \mathbf{u}_\pm and Ψ through (6.10) and (6.15). For notational purposes we introduce

$$\dot{\mathbf{u}}_\pm^* = \mathbf{L}^*(\dot{\mathbf{u}}_\pm, \tilde{\tau}_\pm) = \begin{cases} \frac{\mathbf{R} + \mathbf{I}}{2}(\dot{\mathbf{u}}_\pm - \mathbf{Z}_\pm^{-1} \tilde{\tau}_\pm) & \text{(exterior boundaries),} \\ \dot{\mathbf{u}}_\pm - \mathbf{Z}_\pm^{-1}(\tilde{\tau}_\pm - \tau_\pm^*) & \text{(fault),} \end{cases} \quad (6.16)$$

with τ_\pm^* given by (6.12). Similarly, the linear operator for the right-hand-side of (6.5) is denoted

$$\mathbf{L}(\mathbf{u}_\pm, \mathbf{u}_\pm^*, \tau_\pm^*) := \mathbf{D}_{II}(\mu_\pm) \mathbf{u}_\pm + \sum_\zeta \text{SAT}(\mathbf{u}_\pm, \mathbf{u}_\pm^*, \tau_\pm^*), \quad (6.17)$$

The semi-discretization of the state evolution equation (6.4) is

$$\begin{aligned} \dot{\Psi} &= \mathbf{G}(\mathbf{V}^*, \Psi), & t \in \mathcal{T}, \\ \Psi &= \Psi_0, & t = 0, \end{aligned} \quad (6.18)$$

where \mathbf{G} is evaluated pointwise. Note that \mathbf{V}^* is used in evolving Ψ . This will be of importance when deriving the gradient to the discrete optimization problem in Section 6.2.

Finally, the SBP-SAT semi-discretization of (6.1) is given by

$$\begin{aligned} \bar{\rho}_\pm \ddot{\mathbf{u}}_\pm &= \mathbf{L}(\mathbf{u}_\pm, \mathbf{u}_\pm^*, \tau_\pm^*), & t \in \mathcal{T}, \\ \dot{\mathbf{u}}_\pm^* &= \mathbf{L}^*(\dot{\mathbf{u}}_\pm, \tilde{\tau}_\pm) & t \in \mathcal{T}, \\ \mathbf{u}_\pm &= \mathbf{u}_{0\pm}, \quad \dot{\mathbf{u}}_\pm = \mathbf{v}_{0\pm}, & t = 0, \\ \mathbf{u}_\pm^* &= \mathbf{e}_{\partial\Omega_\pm}^T \mathbf{u}_{0\pm}, & t = 0, \\ \dot{\Psi} &= \mathbf{G}(\mathbf{V}^*, \Psi), & t \in \mathcal{T}, \\ \Psi &= \Psi_0, & t = 0. \end{aligned} \quad (6.19)$$

In [Duru et al. \[2019\]](#), [Erickson et al. \[2022\]](#) it is shown (by using the SBP properties in [Appendix A](#)) that the scheme is energy stable and that the semi-discrete energy rate is a consistent approximation of the continuous energy rate.

With the forward problem discretized in space, the semi-discretization of the optimization problem [\(4.4\)](#) reads

$$\min_{\mathbf{p}} \mathcal{F} \quad \text{subject to } \text{[\(6.19\)](#)}, \quad (6.20)$$

where the semi-discrete misfit is given by

$$\mathcal{F} = \frac{1}{2} \sum_{k=1}^{N_{rec}} \int_0^T |\mathbf{r}^{(k)}(t)|^2 dt, \quad (6.21)$$

and the semi-discrete residuals are

$$\mathbf{r}^{(k)}(t) = \left(\left(\hat{\boldsymbol{\delta}}_{\bar{x}_r^{(k)}}, \mathbf{m}(t) \right)_{\Omega} - m_{data}(t) \right), \quad (6.22)$$

for a measured field \mathbf{m} . We refrain from including filtering operators W in the semi-discrete presentation since no filtering is used in the numerical experiments performed in [Section 6.5](#). The discrete Dirac delta $\hat{\boldsymbol{\delta}}_{\bar{x}_r^{(k)}} \approx \hat{\delta}(\bar{x} - \bar{x}_r^{(k)})$ presented in [Petersson et al. \[2016\]](#) is used to restrict the measured field to the receiver location $\bar{x}_r^{(k)}$. The number of moment conditions used when constructing $\hat{\boldsymbol{\delta}}$ is the same as the order of the SBP difference operators. No smoothness conditions are used.

6.2 Adjoint scheme and semi-discrete gradient

Analogously to the forward problem, the adjoint equations [\(5.3\)](#) reduce to

$$\begin{aligned} \rho_{\pm} \ddot{u}_{\pm}^{\dagger} &= \partial_I \mu_{\pm} \partial_I u_{\pm}^{\dagger} + Q_{\pm}^{\dagger}, & \bar{x} \in \Omega_{\pm}, & t \in \mathcal{T}, \\ u_{\pm}^{\dagger} &= 0, \quad \dot{u}_{\pm}^{\dagger} = 0, & \bar{x} \in \Omega_{\pm}, & t = T, \\ \tau_{\pm}^{\dagger} - Z \dot{u}_{\pm}^{\dagger} &= 0, & \bar{x} \in \partial\Omega_{\pm} \setminus \Gamma, & t \in \mathcal{T}, \\ \tau_{\pm}^{\dagger} &= \mp F^{\dagger}(V^{\dagger}, \Psi^{\dagger}), & \bar{x} \in \Gamma, & t \in \mathcal{T}, \\ -\dot{\Psi}^{\dagger} &= G^{\dagger}(V^{\dagger}, \Psi^{\dagger}), & \bar{x} \in \Gamma, & t \in \mathcal{T}, \\ \Psi^{\dagger} &= 0, & \bar{x} \in \Gamma, & t = T. \end{aligned} \quad (6.23)$$

In the semi-discrete setting, we start by defining

$$\mathbf{w}^{\dagger*} := -Z \dot{\mathbf{u}}^{\dagger*} - \boldsymbol{\tau}^{\dagger*} = -Z \dot{\mathbf{u}}^{\dagger} - \tilde{\boldsymbol{\tau}}^{\dagger} =: \mathbf{w}^{\dagger}, \quad (6.24)$$

which are the outgoing characteristic adjoint variables on a boundary segment after time reversal, where $\tilde{\boldsymbol{\tau}}^{\dagger}$ is defined analogously to [\(6.8\)](#). On an exterior boundary, $\boldsymbol{\tau}^{\dagger*}$ and $\dot{\mathbf{u}}^{\dagger*}$ then satisfies

$$\boldsymbol{\tau}^{\dagger*} = \frac{\mathbf{R} - \mathbf{I}}{2} (-Z \dot{\mathbf{u}}^{\dagger} - \tilde{\boldsymbol{\tau}}^{\dagger}), \quad (6.25)$$

$$-\dot{\mathbf{u}}^{\dagger*} = \frac{\mathbf{R} + \mathbf{I}}{2} (-\dot{\mathbf{u}}^{\dagger} - Z^{-1} \tilde{\boldsymbol{\tau}}^{\dagger}). \quad (6.26)$$

Similarly, on the fault

$$\tau_{\pm}^{\dagger*} = \mp \mathbf{F}^{\dagger}(\mathbf{V}^{\dagger*}, \Psi^{\dagger}) = \mp(\bar{\mathbf{F}}_V \mathbf{V}^{\dagger*} + \bar{\mathbf{G}}_V \Psi^{\dagger}), \quad (6.27)$$

$$-\dot{\mathbf{u}}_{\pm}^{\dagger*} = -\dot{\mathbf{u}}_{\pm}^{\dagger} - \mathbf{Z}_{\pm}^{-1}(\tau_{\pm}^{\dagger} - \tau_{\pm}^{\dagger*}), \quad (6.28)$$

where $\mathbf{F}_V(\mathbf{V}^*, \Psi) \approx \frac{\partial F}{\partial \mathbf{V}}$ and $\mathbf{G}_V(\mathbf{V}^*, \Psi) \approx \frac{\partial G}{\partial \mathbf{V}}$ are evaluated point-wise. The adjoint target slip velocity now satisfies

$$\kappa \mathbf{V}^{\dagger*} + \mathbf{F}^{\dagger}(\mathbf{V}^{\dagger*}, \Psi^{\dagger}) = -\tau_{\ell}^{\dagger} = -\frac{\mathbf{Z}_+ \mathbf{w}_{-}^{\dagger} - \mathbf{Z}_- \mathbf{w}_{+}^{\dagger}}{\mathbf{Z}_+ + \mathbf{Z}_-}. \quad (6.29)$$

Note that F^{\dagger} is linear in V^{\dagger} , such that (6.29) reduces to a time-dependent linear equation in $\mathbf{V}^{\dagger*}$. The semi-discrete adjoint state evolution equation is simply

$$\begin{aligned} -\dot{\Psi}^{\dagger} &= \mathbf{G}^{\dagger}(\mathbf{V}^{\dagger*}, \Psi^{\dagger}) = (\bar{\mathbf{G}}_{\Psi} \Psi^{\dagger} + \bar{\mathbf{F}}_{\Psi} \mathbf{V}^{\dagger*}), & t \in \mathcal{T}, \\ \Psi^{\dagger} &= 0, & t = T. \end{aligned} \quad (6.30)$$

Again, $\mathbf{G}_{\Psi}(\mathbf{V}^*, \Psi) \approx \frac{\partial G}{\partial \Psi}$ and $\mathbf{F}_{\Psi}(\mathbf{V}^*, \Psi) \approx \frac{\partial F}{\partial \Psi}$ are evaluated point-wise.

Adopting the notation for the operators \mathbf{L} and \mathbf{L}^* in (6.17) and (6.16), the adjoint equations are then discretized according to (6.19), resulting in

$$\begin{aligned} \bar{\rho}_{\pm} \ddot{\mathbf{u}}_{\pm}^{\dagger} &= \mathbf{L}(\mathbf{u}_{\pm}^{\dagger}, \mathbf{u}_{\pm}^{\dagger*}, \tau_{\pm}^{\dagger*}) + \mathbf{Q}_{\pm}^{\dagger}, & t \in \mathcal{T}, \\ -\dot{\mathbf{u}}_{\pm}^{\dagger*} &= \mathbf{L}^*(-\dot{\mathbf{u}}_{\pm}^{\dagger}, \hat{\tau}_{\pm}^{\dagger}) & t \in \mathcal{T}, \\ \mathbf{u}_{\pm}^{\dagger} &= 0, \quad \dot{\mathbf{u}}_{\pm}^{\dagger} = 0, & t = T, \\ \mathbf{u}_{\pm}^{\dagger*} &= 0, & t = T, \\ -\dot{\Psi}^{\dagger} &= \mathbf{G}^{\dagger}(\mathbf{V}^{\dagger*}, \Psi^{\dagger}), & t \in \mathcal{T}, \\ \Psi^{\dagger} &= 0, & t = T. \end{aligned} \quad (6.31)$$

The adjoint residual source term is given by $\mathbf{Q}_{\pm}^{\dagger} = \sum_{k=1}^{N_{rec}} \mathbf{S}^{(k)} \hat{\delta}_{\bar{x}_r^{(k)}}$ where

$$\mathbf{S}^{(k)}(t) = \begin{cases} -\hat{\mathbf{r}}^{(k)}(t), & \mathbf{m} = \mathbf{u}, \\ \mathbf{r}^{(k)}(t), & \mathbf{m} = \dot{\mathbf{u}}, \end{cases} \quad (6.32)$$

with $\hat{\mathbf{r}}^{(k)}(t) = \int_0^t \mathbf{r}^{(k)}(t') dt' + \hat{\mathbf{r}}_0^{(k)}$, with $\hat{\mathbf{r}}_0^{(k)}$ chosen such that $\hat{\mathbf{r}}^{(k)}(T) = 0$. See Appendix B for a derivation of the residual source term. As mentioned in Section 6.1 energy stability of the forward scheme (6.19) was established in Duru et al. [2019], Erickson et al. [2022]. Since the adjoint scheme (6.31) only differs in the source term and the rate-and-state equations, the stability of the adjoint scheme follows directly from the stability of the continuous adjoint rate-and-state equations.

We may now state the gradient to (6.20). It is given by the following Lemma which is the semi-discrete counterpart to Theorem 5.1 for antiplane shear.

Lemma 6.1. *Let $\mathbf{V}^{\dagger*}$, Ψ^{\dagger} , satisfy (6.31). Further, let \mathbf{p} be the fault grid function of a parameter in F or G , and let Ψ_0 be the initial state. If \mathbf{V}^* is*

computed exactly from (6.13), then the gradient of the misfit in the semi-discrete inverse problem (6.20) is given by

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial \mathbf{p}} &= - \int_0^T \mathbf{H}_\Gamma \left(\bar{\mathbf{F}}_p \mathbf{V}^{\dagger*} + \bar{\mathbf{G}}_p \Psi^\dagger \right) dt, \\ \frac{\partial \mathcal{F}}{\partial \Psi_0} &= \mathbf{H}_\Gamma \Psi_0^\dagger,\end{aligned}\tag{6.33}$$

where $\mathbf{F}_p(\mathbf{V}^*, \Psi) \approx \frac{\partial F}{\partial p}$ and $\mathbf{G}_p(\mathbf{V}^*, \Psi) \approx \frac{\partial G}{\partial p}$ are evaluated point-wise and $\Psi_0^\dagger = \Psi^\dagger(t=0)$.

Proof. See Appendix C. □

Note that (6.33) is a consistent approximation of (5.9). In proving Lemma 6.1 we further show that the forward scheme (6.19) is dual consistent Hicken and Zingg [2011], Berg and Nordström [2012], Hicken and Zingg [2014], Ghasemi Zinatabadi [2019], i.e., that the adjoint scheme (6.31) is the adjoint of (6.19). This implies that the semi-discrete gradient (6.33) is the exact (to machine precision) gradient to (6.21), and that it is a consistent approximation of the continuous gradient. The condition on the accuracy of \mathbf{V}^* is discussed in the remark at the end of Appendix C. In summary, errors in \mathbf{V}^* translate to an error in the gradient.

6.3 Time integration and the discrete gradient

To integrate the forward and adjoint schemes (6.19), (6.31) in time, the standard fourth-order accurate Runge–Kutta method (RK4) will be used. To this end, introduce $\mathbf{v}_\pm = \dot{\mathbf{u}}_\pm$ such that (6.19) may be rewritten to first-order in time as

$$\begin{aligned}\frac{d\mathbf{U}(t)}{dt} &= \mathbf{A}(t, \mathbf{U}), \quad t \in \mathcal{T}, \\ \mathbf{U} &= \mathbf{U}_0, \quad t = 0,\end{aligned}\tag{6.34}$$

where

$$\mathbf{A}(t, \mathbf{U}) = \begin{bmatrix} \mathbf{A}_-(t, \mathbf{U}_-) \\ \mathbf{A}_+(t, \mathbf{U}_+) \\ \mathbf{G}(\mathbf{V}^*, \Psi) \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{U}_- \\ \mathbf{U}_+ \\ \Psi \end{bmatrix},\tag{6.35}$$

and

$$\mathbf{A}_\pm(t, \mathbf{U}_\pm) = \begin{bmatrix} \mathbf{v}_\pm \\ \bar{\rho}_\pm^{-1}(\mathbf{L}(\mathbf{u}_\pm, \mathbf{u}_\pm^*, \tau_\pm^*) + \mathbf{Q}_\pm(t)) \\ \mathbf{L}^*(\mathbf{v}_\pm, \tilde{\tau}_\pm) \end{bmatrix}, \quad \mathbf{U}_\pm = \begin{bmatrix} \mathbf{u}_\pm \\ \mathbf{v}_\pm \\ \mathbf{u}_\pm^* \end{bmatrix}.\tag{6.36}$$

The ODE system (6.34) is then integrated in time using RK4, where in each step (6.13) is solved for \mathbf{V}^* using bisection.

To integrate the adjoint scheme (6.19) the reversed time $t^\dagger = T - t$ is introduced. Then, the adjoint ODE system is

$$\begin{aligned} \frac{d\mathbf{U}^\dagger(t^\dagger)}{dt^\dagger} &= \mathbf{A}^\dagger(t^\dagger, \mathbf{U}^\dagger), & t^\dagger &\in \mathcal{T}, \\ \mathbf{U}^\dagger &= \mathbf{U}_0^\dagger, & t^\dagger &= 0. \end{aligned} \quad (6.37)$$

Here

$$\mathbf{A}^\dagger(t^\dagger, \mathbf{U}^\dagger) = \begin{bmatrix} \mathbf{A}_-^\dagger(t^\dagger, \mathbf{U}_-^\dagger) \\ \mathbf{A}_+^\dagger(t^\dagger, \mathbf{U}_+^\dagger) \\ \mathbf{G}^\dagger(\mathbf{V}^{\dagger*}, \mathbf{\Psi}^\dagger) \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \mathbf{U}_-^\dagger \\ \mathbf{U}_+^\dagger \\ \mathbf{\Psi}^\dagger \end{bmatrix}, \quad (6.38)$$

$$\mathbf{A}_\pm^\dagger(t^\dagger, \mathbf{U}_\pm^\dagger) = \begin{bmatrix} \bar{\rho}_\pm^{-1}(\mathbf{L}(\mathbf{u}_\pm^\dagger, \mathbf{u}_\pm^{\dagger*}, \mathbf{v}_\pm^\dagger, \mathbf{v}_\pm^{\dagger*}) + \mathbf{Q}_\pm^\dagger(T - t^\dagger)) \\ \mathbf{L}^*(\mathbf{v}_\pm^\dagger, \mathbf{v}_\pm^{\dagger*}) \end{bmatrix}, \quad \mathbf{U}_\pm^\dagger = \begin{bmatrix} \mathbf{u}_\pm^\dagger \\ \mathbf{v}_\pm^\dagger \\ \mathbf{u}_\pm^{\dagger*} \end{bmatrix}, \quad (6.39)$$

where for each t^\dagger (6.29) is solved for $\mathbf{V}^{\dagger*}$. Note that the forward variables \mathbf{V}^* and $\mathbf{\Psi}$ entering as variable coefficients in \mathbf{F}^\dagger and \mathbf{G}^\dagger , also are evaluated at time $T - t^\dagger$.

To evaluate the integrals over time in (6.21) and (6.33), the Runge–Kutta quadrature is utilized. That is, for time interval n with step size Δt_n , the Runge–Kutta quadrature weight at substage $s \in [1, 4]$ is given by $\mathbf{H}_{\mathcal{T}, 4n+s} = \Delta t_n b_s$, where $b = [1/6, 1/3, 1/3, 1/6]$ Butcher [2016]. Then for N time steps, the $4N \times 4N$ diagonal matrix $\mathbf{H}_{\mathcal{T}}$ is a quadrature on the temporal grid \mathcal{T} . The fully discrete optimization problem reads

$$\min_{\mathbf{p}} \mathcal{F} \quad \text{subject to (6.34) solved using RK4}, \quad (6.40)$$

where

$$\mathcal{F} = \frac{1}{2} \sum_{k=1}^{N_{rec}} \sum_{n=0}^{N-1} \sum_{s=1}^4 \mathbf{H}_{\mathcal{T}, 4n+s} |\hat{\mathbf{r}}_{4n+s}^{(k)}|^2. \quad (6.41)$$

with the discrete residual given by

$$\hat{\mathbf{r}}_{4n+s}^{(k)} = \left(\hat{\delta}_{\bar{x}_r}^{(k)}, \hat{\mathbf{m}}_{4n+s}^{(k)} \right)_{\Omega} - m_{data}^{(k)}(t_n + c_s \Delta t_n). \quad (6.42)$$

Here $\hat{\mathbf{m}}^{(k)}$ is the measured field at receiver k for all Runge–Kutta stages between $t = 0$ and $t = T$, $c = [0, 1/2, 1/2, 1]$ are the RK4 stage nodes, and $t_0 = 0$. Note that for each index i , $\hat{\mathbf{m}}_i^{(k)}$ is a grid function.

We now arrive at one of the major results of this work, namely the discrete counterpart to Theorem 5.1 for dynamic rupture inversions in antiplane shear.

Theorem 6.2. *Let $\hat{\mathbf{V}}^{\dagger*}$, $\hat{\mathbf{\Psi}}^\dagger$, be the space-time solutions (including Runge–Kutta substages) to (6.37) solved using RK4. Further, let \mathbf{p} be the fault grid function of a parameter in F or G , and let $\mathbf{\Psi}_0$ be the initial state. If in each*

time step \mathbf{V}^* is computed exactly from (6.13), then the gradient of the misfit in the discrete inverse problem (6.40) is given by

$$\begin{aligned}\frac{\partial \mathcal{F}}{\partial \mathbf{p}} &= - \sum_{n=0}^{N-1} \sum_{s=1}^4 \mathbf{H}_{\mathcal{T}, 4n+s} \mathbf{H}_{\Gamma} \left(\bar{\mathbf{F}}_{p, 4n+s} \hat{\mathbf{V}}_{4n+s}^{\dagger*} + \bar{\mathbf{G}}_{p, 4n+s} \hat{\mathbf{\Psi}}_{4n+s}^{\dagger} \right), \\ \frac{\partial \mathcal{F}}{\partial \mathbf{\Psi}_0} &= \mathbf{H}_{\Gamma} \hat{\mathbf{\Psi}}_0^{\dagger},\end{aligned}\tag{6.43}$$

where $\mathbf{F}_p(\hat{\mathbf{V}}^*, \hat{\mathbf{\Psi}}) \approx \frac{\partial F}{\partial p}$, $\mathbf{G}_p(\hat{\mathbf{V}}^*, \hat{\mathbf{\Psi}}) \approx \frac{\partial G}{\partial p}$ are evaluated point-wise using the Runge–Kutta stage approximations.

Proof. RK4 is self-adjoint with respect to $\mathbf{H}_{\mathcal{T}}$ under time reversal, see Sanz-Serna [2016], Matsuda and Miyatake [2021]. Therefore, by Lemma 6.1, the result follows. \square

The combination of RK4 with the SBP-SAT discretization thus results in a space-time dual-consistent discretization. Specifically, this means that (6.43) is the exact gradient to (6.41) and a consistent approximation of the continuous gradient. Again, errors in the non-linear solve of \mathbf{V}^* will result in an error in $\frac{\partial \mathcal{F}}{\partial \mathbf{p}}$, see the remark at the end of Appendix C.

To summarize, when solving the optimization problem (6.40), the gradient of (6.41) is obtained through the following procedure:

1. Solve (6.34) using RK4, storing the Runge–Kutta stage approximations $\hat{\mathbf{V}}^*$, $\hat{\mathbf{\Psi}}$ and $\hat{\mathbf{r}}$.
2. Solve (6.37) using RK4, with $\hat{\mathbf{V}}^*$, $\hat{\mathbf{\Psi}}$ and the residuals $\hat{\mathbf{r}}^{(k)}$ in \mathbf{Q}^{\dagger} reversed in time. Use the time steps from the forward solve in reverse. Store the Runge–Kutta stage approximations $\hat{\mathbf{V}}^{\dagger*}$, $\hat{\mathbf{\Psi}}^{\dagger}$.
3. Reverse $\hat{\mathbf{V}}^{\dagger*}$, $\hat{\mathbf{\Psi}}^{\dagger}$ in time and compute the gradient according to (6.43).

6.4 Lower-dimensional parameter representation

The misfit functional in the inverse problem (6.40) is typically non-convex with many local minima. As a way to regularize (6.40) by reducing the discrete solution space, a lower-dimensional parameter representation of \mathbf{p} may therefore be beneficial. Examples of such lower-dimensional representations are polynomials or splines Samareh [2001], Bader et al. [2023]. Here we make use of a coarse grid representation of the parameters and move between the lower-dimensional parameter grid and the higher-dimensional computational grid using interpolation operators.

Denote an inversion parameter on the coarse fault grid by \mathbf{p}^C and let \mathbf{I}^{C2F} be an interpolation operator from the coarse grid to the fine computational grid. Using superscript F for fault grid functions on the computational grid, \mathbf{p}^F is given by

$$\mathbf{p}^F = \mathbf{I}^{C2F} \mathbf{p}^C.\tag{6.44}$$

The gradient expressions in (6.43) are derived with respect to \mathbf{p}^F , and are of the form

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}^F} = \mathbf{H}_\Gamma^F \Phi^F, \quad (6.45)$$

where Φ^F is a combination of forward and adjoint fields. By the chain rule, the gradient with respect to \mathbf{p}^C is

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}^C} = \frac{\partial \mathcal{F}}{\partial \mathbf{p}^F} \frac{\partial \mathbf{p}^F}{\partial \mathbf{p}^C}. \quad (6.46)$$

Differentiating (6.44) stated in index notation (where i, j, k denote components in the matrices and vectors and the summation convention applies) leads to

$$\frac{\partial \mathbf{p}_i^F}{\partial \mathbf{p}_k^C} = \mathbf{I}_{ij}^{C2F} \delta_{jk} = \mathbf{I}_{ik}^{C2F} = (\mathbf{I}^{C2F})_{ki}^T. \quad (6.47)$$

Hence,

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}^C} = (\mathbf{I}^{C2F})^T \mathbf{H}_\Gamma^F \Phi^F. \quad (6.48)$$

If \mathbf{I}^{C2F} is any interpolation operator, it is not guaranteed that $(\mathbf{I}^{C2F})^T$ is a consistent approximation of interpolation of fields from the fine to the coarse grid. Thus, it is not guaranteed that $\frac{\partial \mathcal{F}}{\partial \mathbf{p}^C}$ is consistent with the continuous gradient. However, by requiring that \mathbf{I}^{C2F} is constructed such that

$$\mathbf{I}^{F2C} := (\mathbf{H}_\Gamma^C)^{-1} (\mathbf{I}^{C2F})^T \mathbf{H}_\Gamma^F \quad (6.49)$$

is a consistent interpolation operator from the fine grid to the coarse, then (6.48) reads

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}^C} = \mathbf{H}_\Gamma^C \mathbf{I}^{F2C} \Phi^F. \quad (6.50)$$

This corresponds to first interpolating Φ^F to the coarse grid and then applying the coarse fault quadrature. In [Almquist et al. \[2019\]](#) interpolation operators satisfying (6.49) are shown to exist, provided $\mathbf{H}_\Gamma^{C,F}$ are accurate quadratures. Here, we use intermediate glue grids to construct interpolation operators for the boundary-optimized SBP difference operators, see [Kozdon and Wilcox \[2016\]](#), [Stiernström and Almquist \[2022\]](#).

Remark. Replacing $(\mathbf{I}^{C2F})^T$ in (6.48) with any accurate interpolation operator \mathbf{I}^{F2C} would result in consistency with the continuous gradient and an approximation of (6.45) to the order of accuracy of \mathbf{I}^{F2C} . The resulting gradient computation would then correspond to a continuous-adjoint approach [Hicken and Zingg \[2014\]](#). The additional requirement (6.49) guarantees that the continuous-adjoint approach is equivalent to the discrete-adjoint approach, meaning that we obtain the exact gradient of (6.41) while remaining consistent with the continuous gradient. In other words, the interpolation preserves the dual consistency of the forward scheme (6.19). The property (6.49) states that $\mathbf{I}^{F2C} = (\mathbf{I}^{C2F})^\dagger$

in the \mathbf{H}_Γ norm. It is termed SBP or inner product preserving and is commonly used to couple SBP schemes across non-conforming interfaces [Mattsson and Carpenter \[2010\]](#), [Nissen et al. \[2015\]](#), [Lundquist et al. \[2018\]](#), [Kozdon and Wilcox \[2016\]](#), [Almquist et al. \[2019\]](#), [Lundquist et al. \[2020\]](#). Interestingly, the property is useful also in the present setting.

6.5 Numerical studies

In this section, we corroborate our theoretical results and present inversions for friction parameters and initial stresses in dynamic rupture simulations using synthetic data. The experiments are performed on the domain in [Figure 6.1](#), where the two elastic blocks are separated by a rough (band-limited self-similar fractal) fault [Dunham et al. \[2011\]](#), [Fang and Dunham \[2013\]](#). The parameter set used for synthetic data is listed in [Table 6.1](#). Here $\Gamma_{VW} = \{\bar{x} \in \Gamma, x \in [-5, 6] \text{ km}\}$

Table 6.1: Parameter values used in reference problem setup, divided into material parameters, frictional parameters, and initial conditions.

Parameter	Symbol	Value
Density	ρ_\pm	2.6700 g/cm ³
Shear modulus	μ_\pm	32.0381 GPa
Initial stress	σ_n^0	120 MPa
	σ_{yz}^0	72 MPa
Reference friction coefficient for steady sliding	f_0	0.6
Direct effect parameter	a	0.009 $\bar{x} \in \Gamma_{VW}$, 0.013 otherwise
State evolution effect parameter	b	0.011
State evolution distance	D_c	0.2 m $\bar{x} \in \Gamma_{VW}$, 1.0 m otherwise
Reference slip velocity	V_0	10 ⁻⁶ m/s
Initial displacement	$u_{0\pm}$	0
Initial velocity	$v_{0\pm}$	$\pm 5 \times 10^{-13}$ m/s
Initial state	Ψ_0	0.7243

denotes the central velocity-weakening (VW) region of the fault, i.e., the part of the fault where $a(\bar{x}) - b(\bar{x}) < 0$. Almost all slip occurs in this region. In contrast, the region where $a(\bar{x}) - b(\bar{x}) > 0$ is velocity-strengthening (VS), which serves to arrest the rupture. The initial stress tensor is such that the normal stress σ_n^0 is constant while the shear traction varies along the fault according to $\tau^0(\bar{x}) = \sigma_{yz}^0 \hat{n}_y^-(\bar{x})$, where $\hat{n}_y^-(\bar{x})$ is the (spatially variable) y -component of the fault normal on the $-$ side of the fault, and σ_{yz}^0 is a constant remote shear stress (and $\sigma_{xz}^0 = 0$). A one-dimensional Gaussian distribution centered about $x_c = 3$ km, with amplitude $A = 25$ MPa and standard deviation $d = 2$ km is added as loading to the fault, i.e., setting $\tau_L(\bar{x}) = Ae^{-\frac{(x-x_c)^2}{2d^2}}$ in [\(6.3\)](#). The problem setup is such that the external loading will cause an earthquake to nucleate at $\bar{x}_c = (x_c, y_c)$ where $y_c = -0.896$ km is the y -coordinate on the fault

at x_c . The stress and friction conditions are chosen so that rupture nucleates gradually over the first few seconds of the simulation, before transitioning to a propagating rupture. Left- and right-going rupture fronts propagate along the fault at a speed slightly lower than the shear wave speed until they reach the VS region where rupture arrests.

Using m grid points in the x -direction, the domain is discretized such that the grid spacing is approximately equal in both directions, resulting in $m \times (m + 1)$ grid points in total across the multiblock grid $\Omega = \Omega_- \cup \Omega_+$. The number of grid points used for lower-dimensional parameter representations is denoted m_p . 8th-order accurate boundary-optimized SBP operators are used for the spatial discretization. In [Stiernström et al. \[2023\]](#) it was shown that these operators provide an efficient alternative for wave propagation problems where boundary or interface effects are significant. The non-linear equation (6.13) for target slip velocity \mathbf{V}^* is solved using bisection, specifying an absolute tolerance 10^{-13} m/s.

Snapshots of the velocity field from a forward simulation using $m = 1001$ and $\Delta t = 0.0005$ s up to $T = 6$ s are presented in Figure 6.2, illustrating the nucleation, propagation, and arrest of the rupture. The velocity field is nonzero within a circular wavefront known in seismology as the starting phase. It has peaks at the two rupture fronts and is discontinuous across the parts of the fault that are actively slipping. Additional circular wavefronts emanate from the fault as the rupture accelerates or decelerates in response to the spatially variable initial stress and complex fault geometry. Additional larger amplitude waves are produced by rupture arrests that bring the particle velocity back toward zero; these are known as stopping phases. Finally, wave reflections from the right boundary are visible, as the outflow boundary condition is only effective at absorbing normally incident waves. In Figure 6.3 plots of slip velocity \mathbf{V}^* and fault slip $[\mathbf{u}]$ are presented. This setup will be used in Section 6.5.3 to generate high-resolution synthetic data used for inversion.

In the inversions, receivers are placed in a cut-out rectangle such that the region closest to the VW part of the fault is excluded. An example of $N_{rec} = 88$ receivers placed in the cut-out rectangle with outer boundary $-9 \leq x \leq 9$, $-9 \leq y \leq 9$, and inner boundary $-7 \leq x \leq 7$, $-3 \leq y \leq 3$, spaced 2 km apart, is shown in Figure 6.4a. In Figure 6.4b an example of recorded synthetic velocity data from the high-resolution simulation described above is shown, with receivers ordered according to the radial distance r to the hypocenter \bar{x}_c . From Figure 6.4b it is clear that the wavefield is densely sampled and that the receivers record the radial moveout of the high-amplitude waves generated as the rupture front expands.

6.5.1 Verification of discrete gradient

To verify Theorem 6.2 we compare with a first-order finite difference approximation of the gradient. Consider the parameter-normalized norm

$$\|\mathbf{v}\|_{\mathbf{p}} = \|\bar{\mathbf{p}}^{-1}\mathbf{v}\|_{\infty}, \quad (6.51)$$

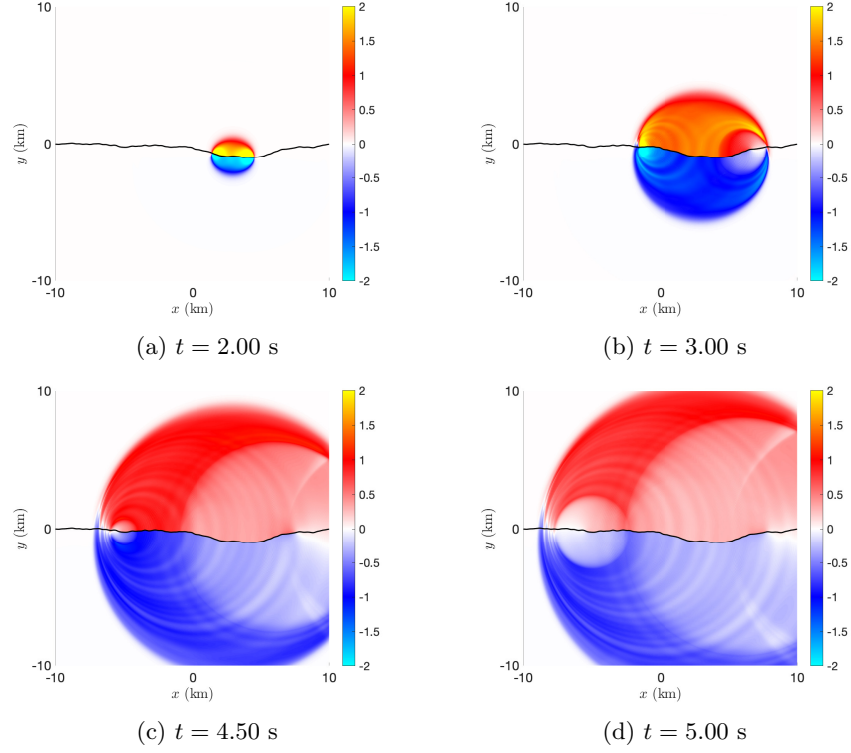


Figure 6.2: Velocity (m/s) at different points in time, showing rupture propagation and arrest.

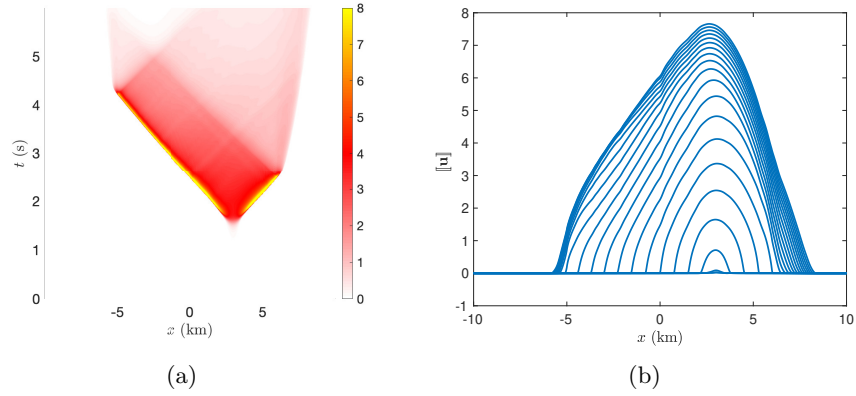


Figure 6.3: (a) Space-time plot of slip velocity \mathbf{V}^* (m/s). (b) Slip $[\mathbf{u}]$ (m) plotted every $t = 0.25$ s.

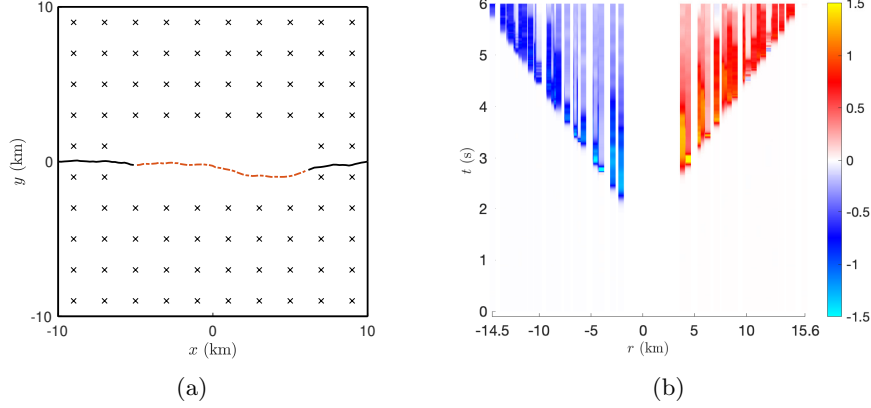


Figure 6.4: (a) Position of $N_{rec} = 88$ receivers spaced 2 km apart (\times). Velocity-weakening part of fault (---). (b) Synthetic velocity data (m/s) at receivers, ordered according to the radial distance r between the receiver and the hypocenter \bar{x}_c . $r < 0$ refers to receivers in Ω^- while $r > 0$ to receivers in Ω^+ .

where \mathbf{v} is a fault grid function on the parameter grid (i. e. , of size $m_p \times 1$). We then define the relative error in the gradient as

$$e(\Delta p) = \frac{\|\frac{\partial \mathcal{F}}{\partial \mathbf{p}} - \mathbf{D}_+ \mathcal{F}(\Delta p)\|_{\mathbf{p}}}{\|\frac{\partial \mathcal{F}}{\partial \mathbf{p}}\|_{\mathbf{p}}}, \quad (6.52)$$

where $\frac{\partial \mathcal{F}}{\partial \mathbf{p}}$ is the adjoint-based gradient and $\mathbf{D}_+ \mathcal{F}(\Delta p)$ is the first-order finite difference approximation, with the i th component given by

$$(\mathbf{D}_+ \mathcal{F}(\Delta p))_i = \frac{\mathcal{F}(\mathbf{p} + \Delta p \mathbf{e}_i) - \mathcal{F}(\mathbf{p})}{\Delta p}. \quad (6.53)$$

Here, \mathbf{e}_i is a vector with a 1 at the i th entry and zeros elsewhere, such that $\mathbf{D}_+ \mathcal{F}(\Delta p)$ is computed by individually perturbing each element in \mathbf{p} by Δp . Synthetic data is generated up to $T = 6$ s using the parameter values in Table 6.1, and discretization parameters $m = 101$, $m_p = 11$ and $\Delta t = 0.005$ s. Note that the grid is too coarse for accurate dynamic rupture simulations but the setup is still useful in verifying that the correct discrete gradient is obtained since any approximation errors in the gradient should be clearly visible. To interpolate between the computational and the parameter grid we construct interpolation operators based on the 8th-order accurate boundary-optimized norm, and the second-order accurate traditional SBP norm on an equispaced grid. This means that the interpolation is second-order accurate. The reason for using a lower-order accurate interpolation is that it reduces oscillations when interpolating non-smooth parameters. $N_{rec} = 88$ receivers are used, positioned as illustrated in Figure 6.4a.

Initializing \mathbf{a} from $1.1a(\bar{x})$ as given by Table 6.1, the error $e(\Delta a)$ is then computed for Δa decreasing from 10^{-5} to 10^{-12} . The results are presented in Figure 6.5 where it can be observed that $e(\Delta a)$ decreases with a first-order rate with decreasing Δa until $e(\Delta a) \approx 2.3 \times 10^{-6}$, at $\Delta a = 10^{-8.3}$ with a displacement misfit and $e(\Delta a) \approx 2.2 \times 10^{-5}$, at $\Delta a = 10^{-9}$ with a velocity misfit. At these points, cancellation errors in the first-order approximation $\mathbf{D}_+\mathcal{F}(\Delta a)$ take effect after which $e(\Delta a)$ increases linearly (although oscillatory) as Δa is further decreased. Note that if $\frac{\partial \mathcal{F}}{\partial \mathbf{a}}$ included an approximation error ε then the error curve would become horizontal at $e(\Delta a) = \varepsilon$, provided ε is larger than cancellation errors in $\mathbf{D}_+\mathcal{F}(\Delta a)$. Since no such piecewise constant parts of the error curves are observed, we conclude that (6.43) indeed is the gradient of \mathcal{F} , and that errors in computing \mathbf{V}^* are negligible. A Matlab script and related code for reproducing Figure 6.5 is available in [Stiernström and Almquist \[2024\]](#).

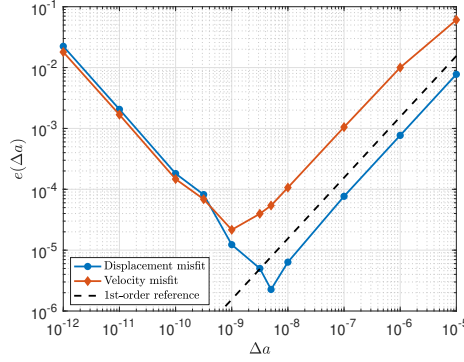


Figure 6.5: Relative error $e(\Delta a)$ between the adjoint-based gradient and the first-order finite difference gradient, as a function of Δa .

6.5.2 Inverse crimes

We now proceed to perform inversions in an inverse crime setting, i.e., we aim to reconstruct parameters based on synthetic velocity data generated from a forward simulation using the same spatio-temporal resolution. There are no model errors or differences in the fault geometry, boundary conditions, or problem parameters (except for the parameters selected as model parameters in the inversions). The wavefield is densely sampled. Thus these inversions are not representative of conditions that might be encountered in real-world applications; instead they serve to test our methodology and to demonstrate the possibility of adjoint-enabled gradient-based optimization for this class of problems. We also perform inversions for a single parameter (i.e., the direct effect parameter $a(\bar{x})$ or the initial shear stress $\tau^0(\bar{x})$) rather than doing inversions for multiple parameters simultaneously. This avoids having to develop regularization schemes to handle parameter trade-offs that will inevitably arise in real-world applications.

The optimization problem (6.40) is solved using the quasi-Newton limited-

memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm provided by the Matlab function `fmincon`. The forward and adjoint solves are performed using $m = 251$ and $\Delta t = 0.003$ s, with m_p varying depending on the parameter considered for inversion. As before, synthetic data are generated using the parameters in Table 6.1 as the true values.

Inverting for the direct effect parameter $a(\bar{x})$ we set $m_p = 26$, and initialize \mathbf{a} identically to 0.0135, which corresponds to a 50% error in the VW region. This choice of initial \mathbf{a} makes the entire fault velocity-strengthening, so in the initial iteration, rupture immediately arrests. Again, the $N_{rec} = 88$ receiver setup illustrated in Figure 6.4a is used. Figure 6.6 presents \mathbf{a} plotted for different iterations; the value of \mathbf{b} is also shown to help delimit VW and VS parameter values. After the first 5 iterations, we observe that \mathbf{a} has been decreased close to its true value at the hypocenter $x_c = 3$ km. This allows for the earthquake to nucleate, radiating waves and thereby reducing the misfit for first-arrival waves. However, since $\mathbf{a} - \mathbf{b} > 0$ outside of the hypocenter, the rupture will not propagate. In subsequent iterations, \mathbf{a} is adjusted to account for the spatial extent of the rupture. After 200 iterations, \mathbf{a} has decreased close to its true value, such that the fault becomes VW within the central region where slip is nonzero (Figure 6.3). Note that the difference in \mathbf{a} in the first and final iterations is negligible in the VS parts of the fault. This part of the fault started VS, so that the rupture arrests there, and no adjustments are required for consistency with the seismograms at the receivers because without slip this part of the fault does not radiate. Thus there is a very low sensitivity (small value of $\frac{\partial \mathcal{F}}{\partial \mathbf{a}}$) in the VS region. In contrast, matching the timing and amplitude of wave arrivals from the interior slipped part of the fault is only possible when \mathbf{a} is close to its true value. A Matlab script and related code for performing the inversion, reproducing Figure 6.6d, is available in [Stiernström and Almquist \[2024\]](#).

To illustrate the effect of the parameter resolution m_p on the optimization procedure, we perform the same inversion using $m_p = 51$ and $m_p = 101$. The computed values of \mathbf{a} after 200 iterations, and the misfit histories are presented in Figures 6.7. Increasing the resolution on the parameter grid, setting $m_p = 51$, we see larger oscillations about the hypocenter x_c . This shows that reducing the discrete parameter space may be beneficial for faster convergence of the misfit, and for reducing the number of local minima. The oscillations around the hypocenter observed for higher m_p could likely be remedied by introducing regularization to the misfit, for instance by penalizing large gradients in \mathbf{a} . Finding a suitable regularization is however out of the scope of this paper.

Next, we invert for the shear traction $\tau^0(\bar{x}) = \sigma_{yz}^0 \hat{n}_y^-(\bar{x})$, setting $m_p = 51$, and starting from τ^0 given by $\sigma_{yz}^0 = 68$ MPa corresponding to approximately 5% initial error. The smaller error in the initial guess is due to inversions of $\tau^0(\bar{x})$ proving to be significantly more difficult for this problem setup, with a higher likelihood of incorrectly converging to a local minimum. This choice of initial guess reduces the initial stress below the true value, to the point where the rupture with the initial parameter values quickly arrests. The inversion must therefore increase τ^0 to produce a propagating rupture. Two inverse-crime inversions are performed, one with the receiver distribution in Figure 6.4a and

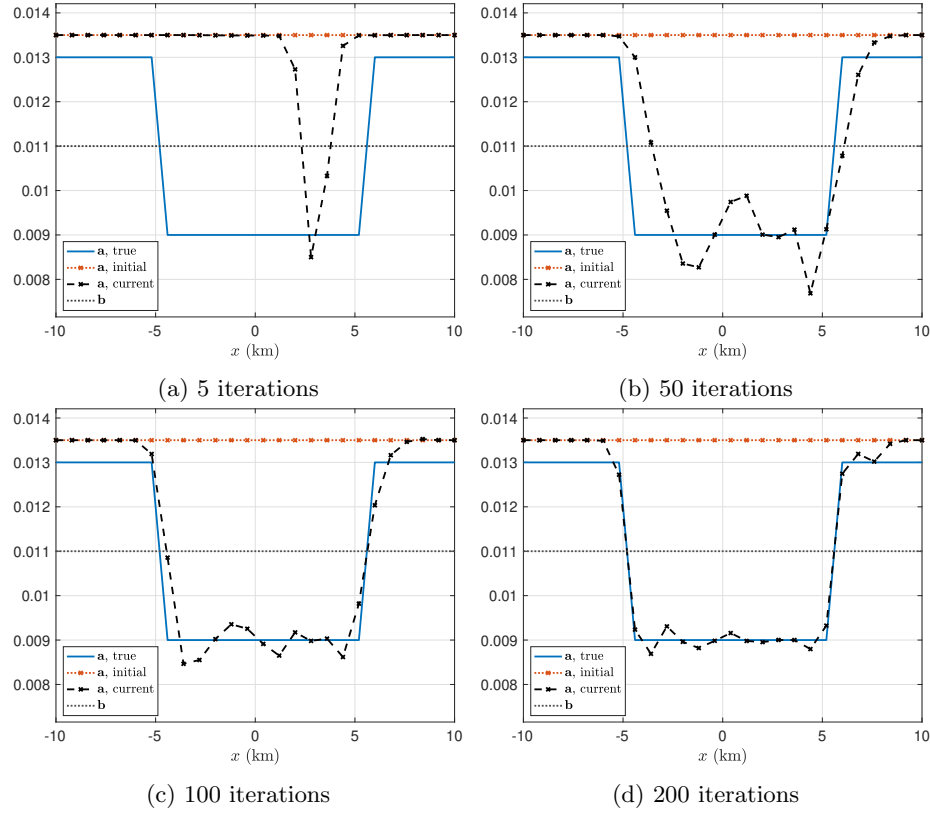


Figure 6.6: Direct effect parameter \mathbf{a} at different iterations for $m_p = 26$.

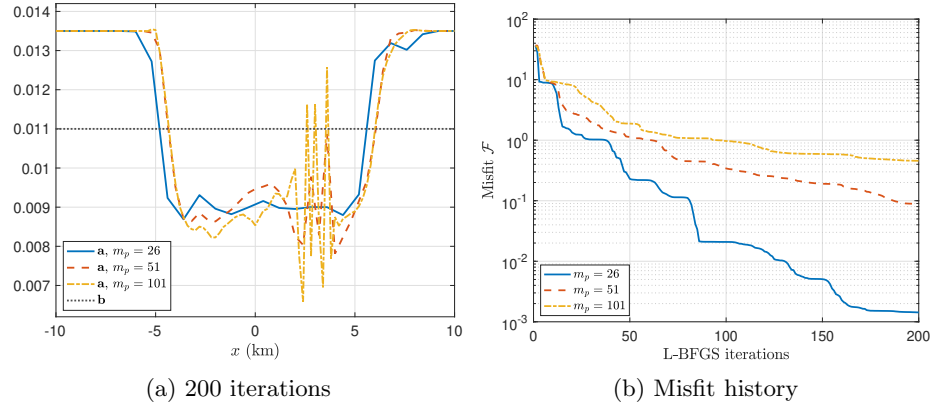


Figure 6.7: Direct effect parameter \mathbf{a} and misfit history for different m_p .

one where the number of receivers is increased by placing them 1 km apart and

decreasing the inner bounding box to $-6 \leq x \leq 7$, $-2 \leq y \leq 2$, for distances in km, totaling $N_{rec} = 325$. In Figure 6.8, τ^0 obtained with $N_{rec} = 88$ is plotted for different iterations. Again, the first 10 iterations adjust the parameter at the hypocenter, while subsequent iterations adjust for the spatial extent of the rupture. However, in this case, the optimization stagnates at a local minimum after about 100 iterations, as illustrated by the misfit history shown in Figure 6.9b. A comparison of the results at 200 iterations with $N_{rec} = 88$ and $N_{rec} = 325$ is shown in Figure 6.9. The result using $N_{rec} = 325$ matches the true value well in the VW region. The increased receiver density as well as having receivers closer to the fault provides more constraints on the shorter wavelength and shorter timescale parts of the rupture process, which manifests as improved resolution of τ^0 .

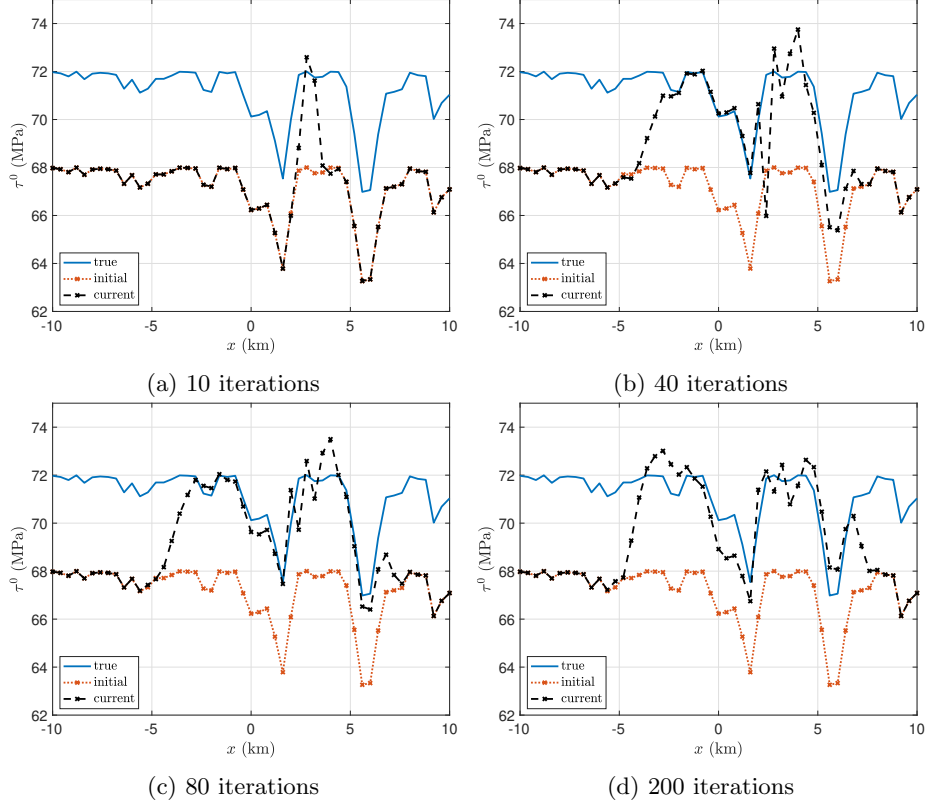


Figure 6.8: Shear traction τ^0 at different iterations using $N_{rec} = 88$.

Remark. Numerical experiments (not presented here) indicate that the inversion is more robust when initializing the inverted parameter such that no nucleation occurs. As observed, this causes the inversion to adjust the parameter to create radiation in order to reduce the misfit, which is done by first adjusting the parameter at the hypocenter and subsequently for the full spatial extent of the

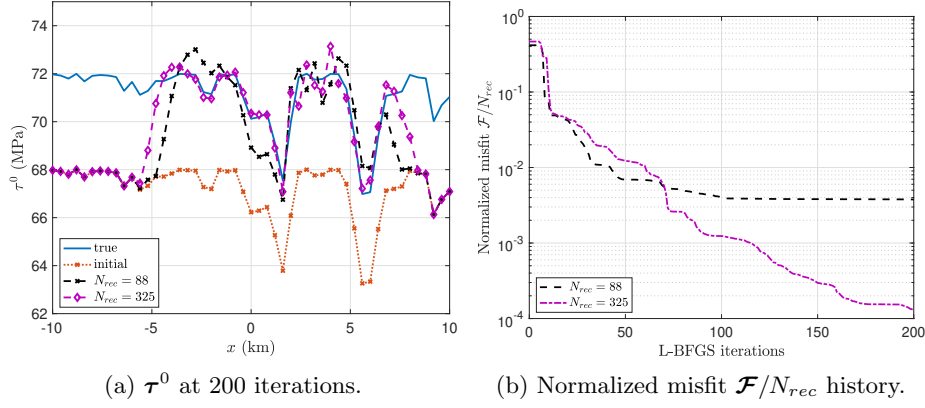


Figure 6.9: Comparison of shear traction τ^0 at 200 iterations for $N_{rec} = 88$ and $N_{rec} = 325$.

rupture. Conversely, starting from parameters causing truly stable parts of the fault to rupture, e. g., if initially $\mathbf{a} - \mathbf{b} < 0$ in parts of the fault that truly are VS, the inversion is very likely to converge to an incorrect local minimum.

6.5.3 Inversion using high-resolution synthetic data

Slightly increasing the difficulty of the optimization problem, we perform inversions using the same setups as in Section 6.5.2, but use high-resolution velocity data generated from the forward simulation presented at the beginning of this section (illustrated in Figures 6.2 - 6.4). In this setting, we effectively have no modeling errors (other than the perturbed parameters considered for inversion), but the data contain features that are not resolvable on the computational grid. When inverting for $a(\bar{x})$ we use $m_p = 26$ and $N_{rec} = 88$ receivers (i. e., data as shown in Figure 6.4b), while for $\tau^0(\bar{x})$, we set $m_p = 51$ and $N_{rec} = 325$. The results after 200 iterations are presented in Figures 6.10 - 6.11. In this setting, \mathbf{a} oscillates around the true value in the VW region, while τ^0 seems to interpolate the true value fairly well. From the misfit in Figure 6.10b it is clear that the optimization for \mathbf{a} has stagnated at a local minimum.

Figure 6.12 shows the velocity field at $t = 5$ corresponding to: the true parameter set on (a) the computational grid and (b) the high-resolution grid, (c, e) the inverted parameters, and (d, f) the initial parameters. Comparing Figures 6.12a and 6.12b it is clear that spurious oscillations due to lower spatial resolution are present in the velocity fields on the computational grid. The oscillations observed in \mathbf{a} in Figure 6.10a are therefore likely a result of numerical errors in the forward and adjoint fields being mapped into the parameter. Increasing the number of receivers to $N_{rec} = 325$ had no significant effect, and the results are therefore omitted. To further improve the results, a combination of regularization and filtering of the misfit residual is likely needed.

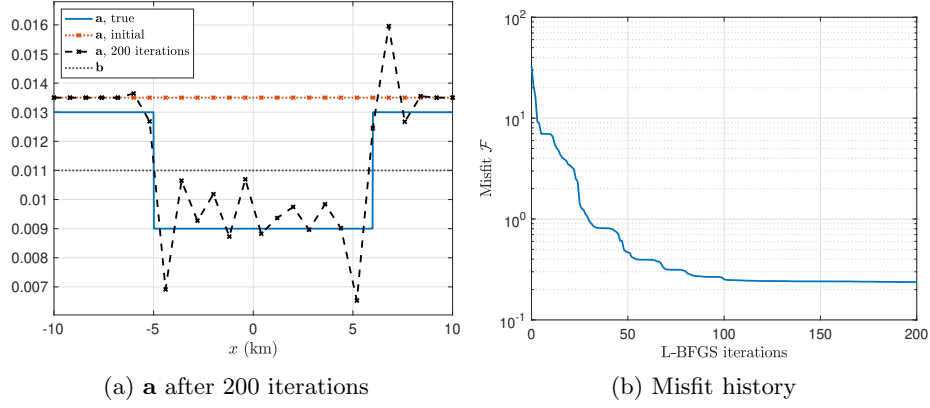


Figure 6.10: Inversion for direct effect parameter \mathbf{a} using high-resolution data and $N_{rec} = 88$.

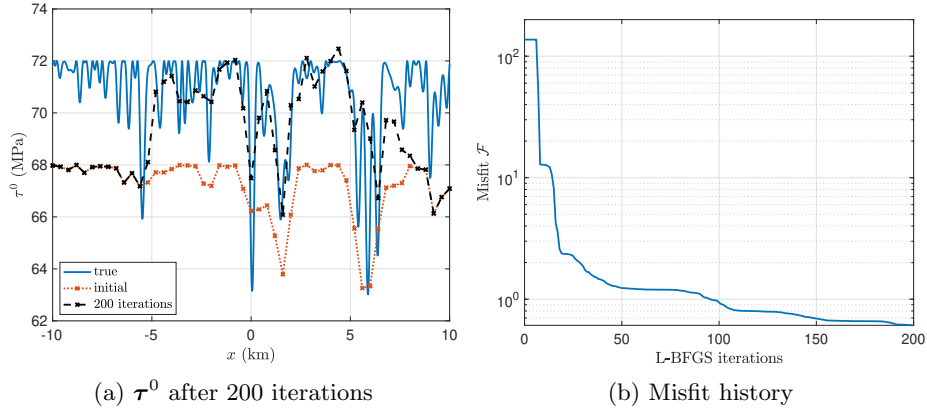
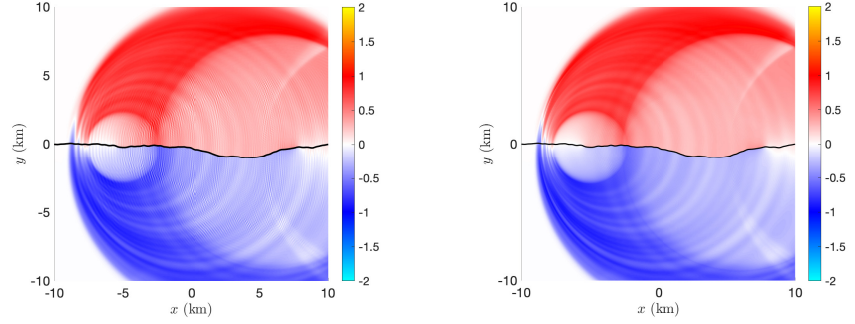


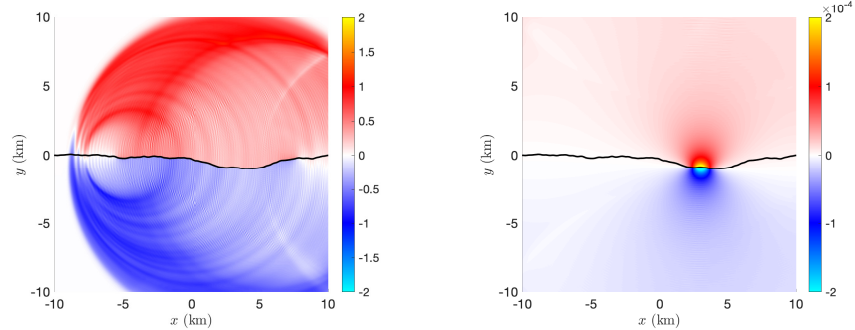
Figure 6.11: Inversion for shear traction τ^0 using high-resolution data and $N_{rec} = 325$.

7 Conclusion

This work presents an adjoint-based optimization framework for earthquake modeling with rate-and-state friction. The main contributions are two-fold. Firstly, the continuous adjoint equations to linear elasticity with rate-state frictional faults are derived, considering displacement and velocity misfit functionals. The adjoint equations satisfy the equations of linear elasticity in reversed time, together with an adjoint friction law and state evolution equation. The adjoint friction law and state evolution equation resemble those of linearized rate-and-state friction but with time-dependent variable coefficients from the forward problem. Additionally, for forward problems involving fault normal stress changes, the adjoint equations include a nonzero fault opening condition,

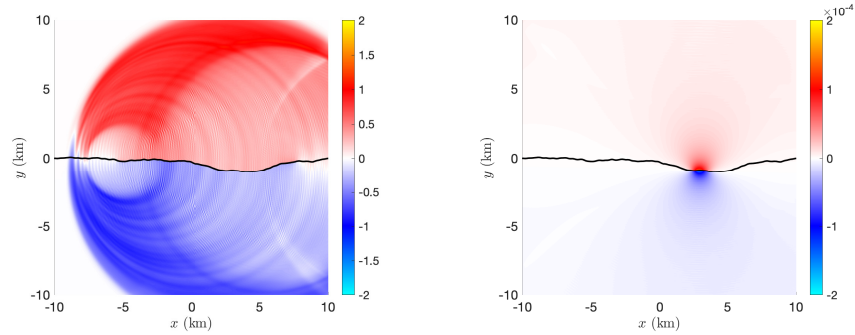


(a) True parameters, computational grid (b) True parameters, high-resolution grid



(c) \mathbf{a} 200 iterations

(d) \mathbf{a} initial parameters



(e) $\boldsymbol{\tau}^0$ 200 iterations

(f) $\boldsymbol{\tau}^0$ initial parameters

Figure 6.12: Comparison of velocity (m/s) at $t = 5$ using the true parameter sets on (a) the computational grid, (b) the high-resolution grid, (c, e) \mathbf{a} and $\boldsymbol{\tau}^0$ obtained after 200 L-BFGS iterations and (d, f) initial parameter values. Note the difference in scales using the initial parameter values.

also with time-dependent variable coefficients from the forward problem. By convolving the adjoint fields with derivatives of the friction law and state evolution equation, the gradient of the misfit with respect to the inversion parameter

is obtained.

Secondly, we present SBP-SAT finite difference discretizations of the forward and adjoint equations in antiplane shear, utilizing the non-stiff interface treatment in [Erickson et al. \[2022\]](#) and the high-order accurate boundary-optimized SBP operators of [Stiernström et al. \[2023\]](#). SBP-preserving interpolation allows the use of a lower-dimensional parameter representation while maintaining dual consistency. In combination with RK4 time integration, a fully discrete gradient expression is presented. Due to the self-adjointness of RK4 [Sanz-Serna \[2016\]](#), [Matsuda and Miyatake \[2021\]](#), the space-time discretization is dual consistent, meaning that the gradient expression is the true gradient of the discrete misfit and consistent with the continuous gradient. This claim is corroborated by numerical experiments in a dynamic rupture setting, presented in Section 6.5.1. Dual consistency has been shown to be beneficial in other applications, see e.g. [Hartmann \[2007\]](#), [Hicken and Zingg \[2014\]](#), where dual consistent discretizations leads to superconvergent approximations of integral functionals in fluid dynamics. The practical benefits of dual consistency for the type of inverse modeling considered herein are yet to be determined.

We then proceed with inversions of model problems in increasingly complex settings in Sections 6.5.2 - 6.5.3. To start with, inverse-crime inversions are performed. Here, we are able to reconstruct the direct effect parameter and fault shear traction fairly accurately in the velocity-weakening region of the fault with appreciable slip. In the velocity-strengthening region that has negligible slip, the parameter sensitivity is close to zero, due to a lack of radiation from this region. Next, inversions utilizing synthetic high-resolution data are performed. For the direct effect parameter, a local minimum is found, which exhibits oscillations in the parameter around the true value. Inversions for shear traction interpolate the true value fairly well. We conclude that regularization, likely in combination with filtering of the misfit and adjoint source signals, is required to further improve the results.

An additional issue which will arise in real-world applications is the likely trade-off between model parameters when doing inversions for both friction parameters and initial stresses. The wavefield that is measured by the receivers, which the inversion attempts to match, is controlled by the fault slip history. Slip is controlled by the reduction in stress from the initial shear stress to the residual strength, the product of dynamic friction and normal stress. Parameter trade-offs will likely exist between $a - b$, which determines dynamic friction, and the initial shear stress. Additional trade-offs can be anticipated by examination of the crack tip equation of motion, which balances the energy release rate with the fracture energy [Freund \[1998\]](#). [Garagash \[2021\]](#) provides expressions for fracture energy in terms of rate and state frictional parameters and normal stress, and the energy release rate is a functional of the shear stress drop history. Parameter combinations that produce similar rupture histories will not be uniquely constrained.

The primary future research direction is thus to investigate suitable regularizations of the misfit functional, which might be designed by anticipating trade-offs through knowledge of dynamic fracture mechanics. In addition, we

will extend the adjoint framework to include other types of measurements, e. g. , strain rate from fiber optics cables, pressure from hydrophones, and InSAR measurements of surface displacement.

Acknowledgments

We sincerely thank Stéphanie Chaillat (ENSTA Paris, Stanford visiting scholar), Laura Bagur and Alice Nassor (ENSTA Paris) for fruitful discussions during the fall of 2022, on the treatment of displacement misfits in Section 5.2.

V. Stiernström was supported by the Swedish Research Council (grant no. 2017-04626). The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at UPPMAX partially funded by the Swedish Research Council through grant agreement no. 2022-06725.

E. M. Dunham was supported by the U.S. National Science Foundation (EAR-1947448).

A Summation-by-parts properties

Integration by parts (IBP) is key in deriving energy balances and showing conservation properties of the continuous equations. SBP operators allow us to mimic the analysis in the discrete setting. It is shown in [Almquist and Dunham \[2020, 2021\]](#) that $\mathbf{D}_{II}(\boldsymbol{\mu})$ satisfies SBP properties on Ω , consistent with the continuous IBP property for $\partial_I \mu \partial_I$ on Ω . The results are summarized below.

For scalar functions u and v , the variable coefficient Laplace operator satisfies

$$(v, \partial_I \mu \partial_I u)_\Omega = (v, \tau_u)_{\partial\Omega} - (\partial_I v, \mu \partial_I u)_\Omega, \quad (\text{A.1})$$

where $\tau_u = \mu \partial_{\hat{n}} u$. Integrating by parts once more yields

$$(v, \partial_I \mu \partial_I u)_\Omega = (v, \tau_u)_{\partial\Omega} - (\tau_v, u)_{\partial\Omega} + (\partial_I \mu \partial_I v, u)_\Omega, \quad (\text{A.2})$$

In the discrete setting let $\mathbf{D}_I \approx \partial_I$ on Ω be a first-derivative SBP operator, based on the same SBP norm \mathbf{H} as \mathbf{D}_{II} . Then, $\mathbf{D}_{II}(\boldsymbol{\mu})$ satisfies the SBP property

$$(\mathbf{v}, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u})_\Omega = (\mathbf{v}, \boldsymbol{\tau}_u)_{\partial\Omega} - (\mathbf{D}_I \mathbf{v}, \bar{\boldsymbol{\mu}} \mathbf{D}_I \mathbf{u})_\Omega - \mathbf{v}^T \mathbf{R}_\Delta \mathbf{u}, \quad (\text{A.3})$$

where $\boldsymbol{\tau}_u = \mathbf{T} \mathbf{u}$, with $\mathbf{T} \approx \mu \partial_{\hat{n}}$ and $\mathbf{R}_\Delta = \mathbf{R}_\Delta^T \geq 0$, $\mathbf{R}_\Delta \approx 0$. Thus (A.3) is a discrete counterpart to the IBP property (A.1). Applying the SBP property once more yields

$$(\mathbf{v}, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u})_\Omega = (\mathbf{v}, \boldsymbol{\tau}_u)_{\partial\Omega} - (\boldsymbol{\tau}_v, \mathbf{u})_{\partial\Omega} + (\mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{v}, \mathbf{u})_\Omega, \quad (\text{A.4})$$

For details on how \mathbf{D}_I and \mathbf{T} are constructed see [Almquist and Dunham \[2020, 2021\]](#), [Stiernström et al. \[2023\]](#).

The SBP property (A.3) facilitates the derivation of a discrete energy rate for scalar wave equations. This is achieved by replacing \mathbf{v} by $\dot{\mathbf{u}}$, in which case the volume terms $(\cdot, \cdot)_\Omega$ contribute to the discrete potential energy rate. The discrete energy rates are then useful for guiding the enforcement of boundary or interface conditions, such that an energy-stable scheme may be obtained. See, e.g., [Duru et al. \[2019\]](#), [Wang and Petersson \[2019\]](#), [Almquist and Dunham \[2020, 2021\]](#), [Erickson et al. \[2022\]](#). The property (A.4) states that $\mathbf{D}_{II}(\boldsymbol{\mu})$ augmented with suitable boundary or interface treatment is self-adjoint with respect to \mathbf{H}_Ω .

B Semi-discrete adjoint sources

To derive the semi-discrete gradient we seek an adjoint source \mathbf{Q}^\dagger satisfying

$$\left(\mathbf{Q}^\dagger, \frac{\partial \dot{\mathbf{u}}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}} = \frac{\partial \mathcal{F}}{\partial \mathbf{p}}. \quad (\text{B.1})$$

Consider the gradient of the misfit (6.21) with a single receiver. For convenience, we drop the superscripts. By the chain rule, it follows that

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}} = \frac{\partial}{\partial \mathbf{p}} \left(\frac{1}{2} \int_{\mathcal{T}} \mathbf{r}^2 dt \right) = \int_{\mathcal{T}} \mathbf{r} \left(\hat{\delta}, \frac{\partial \mathbf{m}}{\partial \mathbf{p}} \right)_\Omega dt = \left(\mathbf{r} \hat{\delta}, \frac{\partial \mathbf{m}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}}. \quad (\text{B.2})$$

If $\mathbf{m} = \dot{\mathbf{u}}$, we have $\frac{\partial \mathcal{F}}{\partial \mathbf{p}} = \left(\mathbf{r} \hat{\delta}, \frac{\partial \dot{\mathbf{u}}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}}$, which is of the desired form (B.1). If $\mathbf{m} = \mathbf{u}$, introduce $\hat{\mathbf{r}}(t) = \int_0^t \mathbf{r}(t') dt' + \hat{\mathbf{r}}_0$ with $\hat{\mathbf{r}}_0 = -\int_0^T \mathbf{r}(t') dt'$ and integrate $\left(\mathbf{r} \hat{\delta}, \frac{\partial \dot{\mathbf{u}}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}}$ by parts in time to obtain $\frac{\partial \mathcal{F}}{\partial \mathbf{p}} = \left(-\hat{\mathbf{r}} \hat{\delta}, \frac{\partial \dot{\mathbf{u}}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}}$. Thus,

$$\frac{\partial \mathcal{F}}{\partial \mathbf{p}} = \left(\mathbf{S}, \frac{\partial \dot{\mathbf{u}}}{\partial \mathbf{p}} \right)_{\Omega \times \mathcal{T}}, \quad \mathbf{S}(t) = \begin{cases} -\hat{\mathbf{r}}(t) \hat{\delta}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_r), & \mathbf{m} = \mathbf{u}, \\ \mathbf{r}(t) \hat{\delta}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_r), & \mathbf{m} = \dot{\mathbf{u}}. \end{cases} \quad (\text{B.3})$$

Summing \mathbf{S} over all receivers results in \mathbf{Q}^\dagger in (6.31).

C Derivation of the semi-discrete gradient

To derive the gradient of (6.21), the procedure in Section 5 is followed. Omitting external forcing (since it does not contribute to the gradient) the Lagrangian misfit functional to (6.21) reads

$$\begin{aligned} \mathcal{L} = & \mathcal{F} + (\dot{\mathbf{u}}^\dagger, \bar{\rho} \ddot{\mathbf{u}} - \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}} \\ & + (\dot{\mathbf{u}}^\dagger, \bar{\rho} \ddot{\mathbf{u}} - \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_- \times \mathcal{T}} \\ & + (\Psi^\dagger, \dot{\Psi} - \mathbf{G})_{\Gamma \times \mathcal{T}}, \end{aligned} \quad (\text{C.1})$$

where we omit subscripts \pm of the grid functions since they are implied by the inner products. Consider first $(\dot{\mathbf{u}}^\dagger, \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}}$. By (6.6) it follows that

$$\begin{aligned} & (\dot{\mathbf{u}}^\dagger, \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}} = \\ & (\dot{\mathbf{u}}^\dagger, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u} + \text{SAT}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}} = \\ & (\dot{\mathbf{u}}^\dagger, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u})_{\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau}^* - \boldsymbol{\tau})_{\partial\Omega_+ \times \mathcal{T}} - (\dot{\boldsymbol{\tau}}^\dagger, \mathbf{u}^* - \mathbf{u})_{\partial\Omega_+ \times \mathcal{T}}. \end{aligned} \quad (\text{C.2})$$

Swapping forward and adjoint variables yields

$$\begin{aligned} & (\dot{\mathbf{u}}, \mathbf{L}(\mathbf{u}^\dagger, \mathbf{u}^{\dagger*}, \boldsymbol{\tau}^{\dagger*}))_{\Omega_+ \times \mathcal{T}} = \\ & (\dot{\mathbf{u}}, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u}^\dagger)_{\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}, \boldsymbol{\tau}^{\dagger*} - \boldsymbol{\tau}^\dagger)_{\partial\Omega_+ \times \mathcal{T}} - (\dot{\boldsymbol{\tau}}, \mathbf{u}^{\dagger*} - \mathbf{u}^\dagger)_{\partial\Omega_+ \times \mathcal{T}}. \end{aligned} \quad (\text{C.3})$$

Next, using the SBP property (A.4) and IBP in time it follows that

$$\begin{aligned} & (\dot{\mathbf{u}}^\dagger, \mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u})_{\Omega_+ \times \mathcal{T}} = \\ & -(\mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u}^\dagger, \dot{\mathbf{u}})_{\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau})_{\partial\Omega_+ \times \mathcal{T}} - (\dot{\boldsymbol{\tau}}^\dagger, \mathbf{u})_{\partial\Omega_+ \times \mathcal{T}}, \end{aligned} \quad (\text{C.4})$$

where terms evaluated at $t = 0$ and $t = T$ vanish due to the initial and terminal conditions (6.19) and (6.31). Due to the p -independence of data, we assumed homogeneous initial conditions for the forward problem, since such terms would

vanish in the gradient expression regardless. Using (C.4) in (C.2) together with (C.3) yields

$$\begin{aligned} & (\dot{\mathbf{u}}^\dagger, \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}} = \\ & -(\mathbf{D}_{II}(\boldsymbol{\mu})\mathbf{u}^\dagger, \dot{\mathbf{u}})_{\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau}^*)_{\partial\Omega_+ \times \mathcal{T}} - (\dot{\boldsymbol{\tau}}^\dagger, \mathbf{u}^*)_{\partial\Omega_+ \times \mathcal{T}} = \\ & -(\dot{\mathbf{u}}^\dagger, \mathbf{L}(\mathbf{u}, \mathbf{u}^*, \boldsymbol{\tau}^*))_{\Omega_+ \times \mathcal{T}} + BT \end{aligned} \quad (\text{C.5})$$

where

$$\begin{aligned} BT &= (\boldsymbol{\tau}^{\dagger*} - \boldsymbol{\tau}^\dagger, \dot{\mathbf{u}})_{\partial\Omega_+ \times \mathcal{T}} - (\mathbf{u}^{\dagger*} - \mathbf{u}^\dagger, \dot{\boldsymbol{\tau}})_{\partial\Omega_+ \times \mathcal{T}} \\ &+ (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau}^*)_{\partial\Omega_+ \times \mathcal{T}} - (\dot{\boldsymbol{\tau}}^\dagger, \mathbf{u}^*)_{\partial\Omega_+ \times \mathcal{T}} \\ &= (\boldsymbol{\tau}^{\dagger*} - \boldsymbol{\tau}^\dagger, \dot{\mathbf{u}})_{\partial\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \boldsymbol{\tau})_{\partial\Omega_+ \times \mathcal{T}} \\ &+ (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau}^*)_{\partial\Omega_+ \times \mathcal{T}} + (\boldsymbol{\tau}^\dagger, \dot{\mathbf{u}}^*)_{\partial\Omega_+ \times \mathcal{T}} \end{aligned} \quad (\text{C.6})$$

with the last equality obtained by IBP in time. Next we perform a change of variables in the term $(\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \boldsymbol{\tau})_{\partial\Omega_+ \times \mathcal{T}}$ from $\boldsymbol{\tau}$ to $\tilde{\boldsymbol{\tau}}$ using (6.8)

$$\begin{aligned} (\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \boldsymbol{\tau})_{\partial\Omega_+ \times \mathcal{T}} &= (\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \tilde{\boldsymbol{\tau}} - \gamma(\mathbf{u}^* - \mathbf{u}))_{\partial\Omega_+ \times \mathcal{T}} = \\ &(\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \tilde{\boldsymbol{\tau}})_{\partial\Omega_+ \times \mathcal{T}} + (\gamma(\mathbf{u}^{\dagger*} - \mathbf{u}^\dagger), \dot{\mathbf{u}}^*)_{\partial\Omega_+ \times \mathcal{T}} - (\gamma(\mathbf{u}^{\dagger*} - \mathbf{u}^\dagger), \dot{\mathbf{u}})_{\partial\Omega_+ \times \mathcal{T}}, \end{aligned} \quad (\text{C.7})$$

where IBP in time was used to obtain the last equality. Combining the result with (C.6) and substituting $\tilde{\boldsymbol{\tau}}^\dagger$ results in

$$\begin{aligned} BT &= (\boldsymbol{\tau}^{\dagger*} - \tilde{\boldsymbol{\tau}}^\dagger, \dot{\mathbf{u}})_{\partial\Omega_+ \times \mathcal{T}} + (\dot{\mathbf{u}}^{\dagger*} - \dot{\mathbf{u}}^\dagger, \tilde{\boldsymbol{\tau}})_{\partial\Omega_+ \times \mathcal{T}} \\ &+ (\dot{\mathbf{u}}^\dagger, \boldsymbol{\tau}^*)_{\partial\Omega_+ \times \mathcal{T}} + (\tilde{\boldsymbol{\tau}}^\dagger, \dot{\mathbf{u}}^*)_{\partial\Omega_+ \times \mathcal{T}}. \end{aligned} \quad (\text{C.8})$$

Next we use (6.7) and (6.24) to obtain

$$\begin{aligned} \tilde{\boldsymbol{\tau}} &= \boldsymbol{\tau}^* + Z(\dot{\mathbf{u}} - \dot{\mathbf{u}}^*), \\ \tilde{\boldsymbol{\tau}}^\dagger &= \boldsymbol{\tau}^{\dagger*} - Z(\dot{\mathbf{u}}^\dagger - \dot{\mathbf{u}}^{\dagger*}). \end{aligned} \quad (\text{C.9})$$

Substituting (C.9) into (C.8) and eliminating terms yields

$$BT = (\dot{\mathbf{u}}^{\dagger*}, \boldsymbol{\tau}^*)_{\partial\Omega_+ \times \mathcal{T}} + (\boldsymbol{\tau}^{\dagger*}, \dot{\mathbf{u}}^*)_{\partial\Omega_+ \times \mathcal{T}}. \quad (\text{C.10})$$

Since the target values satisfy the boundary conditions, BT vanishes on all boundaries but the fault. This is straightforwardly seen on rigid-wall or traction-free boundaries, where \mathbf{u}^* and $\mathbf{u}^{\dagger*}$, or $\boldsymbol{\tau}^*$ and $\boldsymbol{\tau}^{\dagger*}$, respectively, are zero. On non-reflecting boundaries, substituting (6.9), (6.10), (6.25), (6.26) with $\mathbf{R} = \mathbf{0}$ into (C.10) yields the result. Adding contributions of (C.10) from Ω_- yields

$$IT = BT_+ + BT_- = (\boldsymbol{\tau}_+^{\dagger*}, \mathbf{V}^*)_{\Gamma \times \mathcal{T}} - (\mathbf{V}^{\dagger*}, \boldsymbol{\tau}_+^*)_{\Gamma \times \mathcal{T}}. \quad (\text{C.11})$$

where it is assumed that $\mathbf{V}^* = \llbracket \dot{\mathbf{u}}^* \rrbracket$ and $\mathbf{V}^{\dagger*} = -\llbracket \dot{\mathbf{u}}^{\dagger*} \rrbracket$.

By (C.1) and (C.5) it then follows that

$$\begin{aligned}\mathcal{L} &= \mathcal{F} - (\bar{\rho}\ddot{\mathbf{u}}^\dagger - \mathbf{L}(\mathbf{u}^\dagger, \mathbf{u}^{\dagger*}, \boldsymbol{\tau}^{\dagger*}, \dot{\mathbf{u}})_{\Omega_+ \times \mathcal{T}} \\ &\quad - (\bar{\rho}\ddot{\mathbf{u}}^\dagger - \mathbf{L}(\mathbf{u}^\dagger, \mathbf{u}^{\dagger*}, \boldsymbol{\tau}^{\dagger*}, \dot{\mathbf{u}})_{\Omega_- \times \mathcal{T}} \\ &\quad - ST - IT,\end{aligned}\tag{C.12}$$

where

$$ST = (\dot{\Psi}^\dagger, \Psi)_{\Gamma \times \mathcal{T}} + (\Psi^\dagger, \mathbf{G})_{\Gamma \times \mathcal{T}} - (\Psi_0^\dagger, \Psi_0)_\Gamma,\tag{C.13}$$

is obtained from IBP in time, where the term at $t = T$ vanishes due to the terminal condition in (6.31).

We will now derive the gradient $\frac{\partial \mathcal{L}}{\partial \mathbf{p}}$. Combining (B.1) with the adjoint scheme (6.31), the gradient of the volume terms vanishes such that $\frac{\partial \mathcal{L}}{\partial \mathbf{p}} = -\frac{\partial ST}{\partial \mathbf{p}} - \frac{\partial IT}{\partial \mathbf{p}}$. First, consider the inversion for a parameter in either F or G . Since

$$\frac{\partial \tau_+^*}{\partial \mathbf{p}} = -\left(\bar{\mathbf{F}}_V \frac{\partial \mathbf{V}^*}{\partial \mathbf{p}} + \bar{\mathbf{F}}_\Psi \frac{\partial \Psi}{\partial \mathbf{p}} + \bar{\mathbf{F}}_p \frac{\partial \mathbf{p}}{\partial \mathbf{p}}\right),\tag{C.14}$$

and

$$\frac{\partial ST}{\partial \mathbf{p}} = \left(\dot{\Psi}^\dagger, \frac{\partial \Psi}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}} + \left(\Psi^\dagger, \bar{\mathbf{G}}_V \frac{\partial \mathbf{V}^*}{\partial \mathbf{p}} + \bar{\mathbf{G}}_\Psi \frac{\partial \Psi}{\partial \mathbf{p}} + \bar{\mathbf{G}}_p \frac{\partial \mathbf{p}}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}},\tag{C.15}$$

we obtain

$$\frac{\partial IT}{\partial \mathbf{p}} = \left(\tau_+^{\dagger*}, \frac{\partial \mathbf{V}^*}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}} + \left(\mathbf{V}^{\dagger*}, \bar{\mathbf{F}}_V \frac{\partial \mathbf{V}^*}{\partial \mathbf{p}} + \bar{\mathbf{F}}_\Psi \frac{\partial \Psi}{\partial \mathbf{p}} + \bar{\mathbf{F}}_p \frac{\partial \mathbf{p}}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}}\tag{C.16}$$

and

$$\begin{aligned}\frac{\partial ST}{\partial \mathbf{p}} + \frac{\partial IT}{\partial \mathbf{p}} &= \left(\tau_+^{\dagger*} + \bar{\mathbf{F}}_V \mathbf{V}^{\dagger*} + \bar{\mathbf{G}}_V \Psi^\dagger, \frac{\partial \mathbf{V}^*}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}} \\ &\quad + \left(\dot{\Psi}^\dagger + \bar{\mathbf{F}}_\Psi \mathbf{V}^{\dagger*} + \bar{\mathbf{G}}_\Psi \Psi^\dagger, \frac{\partial \Psi}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}} \\ &\quad + \left(\bar{\mathbf{G}}_p \Psi^\dagger + \bar{\mathbf{F}}_p \mathbf{V}^{\dagger*}, \frac{\partial \mathbf{p}}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}}.\end{aligned}\tag{C.17}$$

Using the discrete friction law (6.27) and state evolution equation (6.30), we arrive at

$$\frac{\partial ST}{\partial \mathbf{p}} + \frac{\partial IT}{\partial \mathbf{p}} = \left(\bar{\mathbf{G}}_p \Psi^\dagger + \bar{\mathbf{F}}_p \mathbf{V}^{\dagger*}, \frac{\partial \mathbf{p}}{\partial \mathbf{p}}\right)_{\Gamma \times \mathcal{T}}.\tag{C.18}$$

Similarly, if inverting for the initial state Ψ_0

$$\frac{\partial \tau_+^*}{\partial \Psi_0} = -\left(\bar{\mathbf{F}}_V \frac{\partial \mathbf{V}^*}{\partial \Psi_0} + \bar{\mathbf{F}}_\Psi \frac{\partial \Psi}{\partial \Psi_0}\right),\tag{C.19}$$

while

$$\begin{aligned} \frac{\partial ST}{\partial \Psi_0} = & \left(\dot{\Psi}^\dagger, \frac{\partial \Psi}{\partial \Psi_0} \right)_{\Gamma \times \mathcal{T}} + \left(\Psi^\dagger, \bar{\mathbf{G}}_V \frac{\partial \mathbf{V}^*}{\partial \Psi_0} + \bar{\mathbf{G}}_\Psi \frac{\partial \Psi}{\partial \Psi_0} \right)_{\Gamma \times \mathcal{T}} \\ & - \left(\Psi_0^\dagger, \frac{\partial \Psi_0}{\partial \Psi_0} \right)_\Gamma, \end{aligned} \quad (\text{C.20})$$

such that

$$\frac{\partial ST}{\partial \Psi_0} + \frac{\partial IT}{\partial \Psi_0} = - \left(\Psi_0^\dagger, \frac{\partial \Psi_0}{\partial \Psi_0} \right)_\Gamma. \quad (\text{C.21})$$

Since (in index notation) $\frac{\partial \mathbf{p}_j}{\partial \mathbf{p}_i} = \delta_{ij}$ (and similarly for Ψ_0) (6.33) follows.

Remark. To reach (C.18) and (C.21) we have assumed $\mathbf{V}^* = \llbracket \dot{\mathbf{u}}^* \rrbracket$ and $\mathbf{V}^{\dagger*} = -\llbracket \dot{\mathbf{u}}^{\dagger*} \rrbracket$. This is true for $\mathbf{V}^{\dagger*}$, since the adjoint scheme satisfies (6.29) exactly, due to F^\dagger being linear in V^\dagger such that the solve is trivial. However, due to the non-linear solve required in (6.13),

$$\llbracket \dot{\mathbf{u}}^* \rrbracket = \mathbf{V}^* + \varepsilon. \quad (\text{C.22})$$

When adding the contributions from the interface terms, this leads to

$$IT = \left(\tau_+^{\dagger*}, \mathbf{V}^* + \varepsilon \right)_{\Gamma \times \mathcal{T}} - \left(\mathbf{V}^{\dagger*}, \tau_+^* \right)_{\Gamma \times \mathcal{T}}, \quad (\text{C.23})$$

such that

$$\frac{\partial \mathcal{F}^{(\varepsilon)}}{\partial \mathbf{p}_i} = \frac{\partial \mathcal{F}}{\partial \mathbf{p}_i} + \left(\tau_+^{\dagger*}, \frac{\partial \varepsilon}{\partial \mathbf{p}_i} \right)_{\Gamma \times \mathcal{T}} \quad (\text{C.24})$$

is the resulting gradient. Thus, the error in the non-linear solve results in an error in the semi-discrete gradient. However, since F is invertible and monotone on the interval $[V_{min}, V_{max}]$, the root can be bracketed, i.e., $\mathbf{V}^* \in [\mathbf{V}_{min}, \mathbf{V}_{max}]$ Erickson et al. [2022]. Therefore we can control ε arbitrarily through the tolerance in a bracketed root finding algorithm, e.g. bisection.

References

- M. Aldam, Y. Bar-Sinai, I. Svetlizky, E. A. Brener, J. Fineberg, and E. Bouchbinder. Frictional sliding without geometrical reflection symmetry. *Phys. Rev. X*, 6:041023, Oct 2016. doi:[10.1103/PhysRevX.6.041023](https://doi.org/10.1103/PhysRevX.6.041023).
- K. L. Allison and E. M. Dunham. Earthquake cycle simulations with rate-and-state friction and power-law viscoelasticity. *Tectonophysics*, 733:232–256, 2018. doi:<https://doi.org/10.1016/j.tecto.2017.10.021>. Physics of Earthquake Rupture Propagation.
- M. Almquist and E. M. Dunham. Non-stiff boundary and interface penalties for narrow-stencil finite difference approximations of the Laplacian on curvilinear multiblock grids. *J. Comput. Phys.*, 408:109294, 33, 2020. doi:[10.1016/j.jcp.2020.109294](https://doi.org/10.1016/j.jcp.2020.109294).

- M. Almquist and E. M. Dunham. Elastic wave propagation in anisotropic solids using energy-stable finite differences with weakly enforced boundary and interface conditions. *J. Comput. Phys.*, 424:109842, 33, 2021. doi:[10.1016/j.jcp.2020.109842](https://doi.org/10.1016/j.jcp.2020.109842).
- M. Almquist, S. Wang, and J. Werpers. Order-preserving interpolation for summation-by-parts operators at nonconforming grid interfaces. *SIAM J. Sci. Comput.*, 41(2):A1201–A1227, 2019. doi:[10.1137/18M1191609](https://doi.org/10.1137/18M1191609).
- M. Bader, M. Almquist, and E. M. Dunham. Modeling and inversion in acoustic-elastic coupled media using energy-stable summation-by-parts operators. *Geophysics*, 88(3):T137–T150, 04 2023. doi:[10.1190/geo2022-0195.1](https://doi.org/10.1190/geo2022-0195.1).
- J. Berg and J. Nordström. Superconvergent functional output for time-dependent problems using finite differences on summation-by-parts form. *J. Comput. Phys.*, 231(20):6846–6860, 2012. doi:[10.1016/j.jcp.2012.06.032](https://doi.org/10.1016/j.jcp.2012.06.032).
- T. Bui-Thanh. Adjoint and its roles in sciences, engineering, and mathematics: A tutorial, 2023.
- J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, Ltd., Chichester, third edition, 2016. ISBN 978-1-119-12150-3. doi:[10.1002/9781119121534](https://doi.org/10.1002/9781119121534). With a foreword by J. M. Sanz-Serna.
- S. M. Day, L. A. Dalgner, N. Lapusta, and Y. Liu. Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *J. Geophys. Res. Solid Earth*, 110(B12), 2005. doi:<https://doi.org/10.1029/2005JB003813>.
- J. H. Dieterich. Modeling of rock friction: 1. experimental results and constitutive equations. *J. Geophys. Res. Solid Earth*, 84(B5):2161–2168, 1979. doi:<https://doi.org/10.1029/JB084iB05p02161>.
- E. M. Dunham, D. Belanger, L. Cong, and J. E. Kozdon. Earthquake Ruptures with Strongly Rate-Weakening Friction and Off-Fault Plasticity, Part 2: Nonplanar Faults. *Bull. Seismol. Soc. Am.*, 101(5):2308–2322, 10 2011. ISSN 0037-1106. doi:[10.1785/0120100076](https://doi.org/10.1785/0120100076).
- K. Duru and E. M. Dunham. Dynamic earthquake rupture simulations on non-planar faults embedded in 3D geometrically complex, heterogeneous elastic solids. *J. Comput. Phys.*, 305:185–207, 2016. doi:[10.1016/j.jcp.2015.10.021](https://doi.org/10.1016/j.jcp.2015.10.021).
- K. Duru, K. L. Allison, M. Rivet, and E. M. Dunham. Dynamic rupture and earthquake sequence simulations using the wave equation in second-order form. *Geophys. J. Int.*, 219:796–815, 2019. doi:[10.1093/gji/ggz319](https://doi.org/10.1093/gji/ggz319).
- B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.*, 31(139):629–651, 1977. doi:[10.2307/2005997](https://doi.org/10.2307/2005997).

- B. A. Erickson and E. M. Dunham. An efficient numerical method for earthquake cycles in heterogeneous media: Alternating subbasin and surface-rupturing events on faults crossing a sedimentary basin. *J. Geophys. Res. Solid Earth*, 119(4):3290–3316, 2014. doi:<https://doi.org/10.1002/2013JB010614>.
- B. A. Erickson, J. E. Kozdon, and T. Harvey. A non-stiff summation-by-parts finite difference method for the scalar wave equation in second order form: characteristic boundary conditions and nonlinear interfaces. *J. Sci. Comput.*, 93(1):Paper No. 17, 39, 2022. doi:[10.1007/s10915-022-01961-1](https://doi.org/10.1007/s10915-022-01961-1).
- B. A. Erickson, J. Jiang, V. Lambert, S. D. Barbot, M. Abdelmeguid, M. Almquist, J. Ampuero, R. Ando, C. Cattania, A. Chen, L. Dal Zilio, S. Deng, E. M. Dunham, A. E. Elbanna, A. Gabriel, T. W. Harvey, Y. Huang, Y. Kaneko, J. E. Kozdon, N. Lapusta, D. Li, M. Li, C. Liang, Y. Liu, S. Ozawa, A. Perez-Silva, C. Pranger, P. Segall, Y. Sun, P. Thakur, C. Uphoff, Y. van Dinther, and Y. Yang. Incorporating full elastodynamic effects and dipping fault geometries in community code verification exercises for simulations of earthquake sequences and aseismic slip (seas). *Bull. Seismol. Soc. Am.*, 113(2):499–523, 01 2023. doi:[10.1785/0120220066](https://doi.org/10.1785/0120220066).
- Z. Fang and E. M. Dunham. Additional shear resistance from fault roughness and stress levels on geometrically complex faults. *J. Geophys. Res. Solid Earth*, 118(7):3642–3654, 2013. doi:<https://doi.org/10.1002/jgrb.50262>.
- A. Fichtner. *Full Seismic Waveform Modelling and Inversion*. Springer Berlin, Heidelberg, 2011. doi:[10.1007/978-3-642-15807-0](https://doi.org/10.1007/978-3-642-15807-0).
- L. B. Freund. *Dynamic Fracture Mechanics*. Cambridge University Press, 1998.
- F. Gallovič, L. Valentová, J.-P. Ampuero, and A.-A. Gabriel. Bayesian dynamic finite-fault inversion: 1. method and synthetic test. *J. Geophys. Res.-Sol. Ea.*, 124(7):6949–6969, 2019a. doi:[10.1029/2019JB017510](https://doi.org/10.1029/2019JB017510).
- F. Gallovič, L. Valentová, J.-P. Ampuero, and A.-A. Gabriel. Bayesian dynamic finite-fault inversion: 2. application to the 2016 Mw 6.2 Amatrice, Italy, earthquake. *J. Geophys. Res.-Sol. Ea.*, 124(7):6970–6988, 2019b. doi:[10.1029/2019JB017512](https://doi.org/10.1029/2019JB017512).
- D. I. Garagash. Fracture mechanics of rate-and-state faults and fluid injection induced slip. *Philos. Trans. R. Soc. A*, 379(2196):20200129, 2021. doi:[10.1098/rsta.2020.0129](https://doi.org/10.1098/rsta.2020.0129).
- F. Ghasemi Zinatabadi. *Stability, dual consistency and conservation of summation-by-parts formulations for multiphysics problems*. Number 1998 in Linköping Studies in Science and Technology. Dissertations. 2019. ISBN 9789176850312. doi:[10.3384/diss.diva-159189](https://doi.org/10.3384/diss.diva-159189).
- M. B. Giles and N. A. Pierce. An introduction to the adjoint approach to design. *Flow, Turbulence and Combustion*, 65(3):393–415, 2000. doi:[10.1023/A:1011430410075](https://doi.org/10.1023/A:1011430410075).

- B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time-dependent problems and difference methods*. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2013. ISBN 978-0-470-90056-7. doi:[10.1002/9781118548448](https://doi.org/10.1002/9781118548448).
- R. A. Harris, M. Barall, R. Archuleta, E. Dunham, B. Aagaard, J. P. Ampuero, H. Bhat, V. Cruz-Atienza, L. Dalguer, P. Dawson, S. Day, B. Duan, G. Ely, Y. Kaneko, Y. Kase, N. Lapusta, Y. Liu, S. Ma, D. Oglesby, K. Olsen, A. Pitarka, S. Song, and E. Templeton. The scec/usgs dynamic earthquake rupture code verification exercise. *Seismol. Res. Lett.*, 80(1):119–126, 01 2009. doi:[10.1785/gssrl.80.1.119](https://doi.org/10.1785/gssrl.80.1.119).
- R. A. Harris, M. Barall, B. Aagaard, S. Ma, D. Roten, K. Olsen, B. Duan, D. Liu, B. Luo, K. Bai, J.-P. Ampuero, Y. Kaneko, A. Gabriel, K. Duru, T. Ulrich, S. Wollherr, Z. Shi, E. Dunham, S. Bydlon, Z. Zhang, X. Chen, S. N. Somala, C. Pelties, J. Tago, V. M. Cruz-Atienza, J. Kozdon, E. Daub, K. Aslam, Y. Kase, K. Withers, and L. Dalguer. A suite of exercises for verifying dynamic earthquake rupture codes. *Seismol. Res. Lett.*, 89(3):1146–1162, 04 2018. doi:[10.1785/0220170222](https://doi.org/10.1785/0220170222).
- R. Hartmann. Adjoint consistency analysis of discontinuous galerkin discretizations. *SIAM J. Numer. Anal.*, 45(6):2671–2696, 2007. doi:[10.1137/060665117](https://doi.org/10.1137/060665117).
- T. W. Harvey, B. A. Erickson, and J. E. Kozdon. A high-order accurate summation-by-parts finite difference method for fully-dynamic earthquake sequence simulations within sedimentary basins. *J. Geophys. Res. Solid Earth*, 128(4):e2022JB025357, 2023. doi:[10.1029/2022JB025357](https://doi.org/10.1029/2022JB025357).
- E. R. Heimisson, E. M. Dunham, and M. Almquist. Poroelastic effects destabilize mildly rate-strengthening friction to generate stable slow slip pulses. *J. Mech. Phys. Solids*, 130:262–279, 2019. doi:[10.1016/j.jmps.2019.06.007](https://doi.org/10.1016/j.jmps.2019.06.007).
- J. E. Hicken and D. W. Zingg. Superconvergent functional estimates from summation-by-parts finite-difference discretizations. *SIAM J. Sci. Comput.*, 33(2):893–922, 2011. doi:[10.1137/100790987](https://doi.org/10.1137/100790987).
- J. E. Hicken and D. W. Zingg. Dual consistency and functional accuracy: a finite-difference perspective. *J. Comput. Phys.*, 256:161–182, 2014. doi:[10.1016/j.jcp.2013.08.014](https://doi.org/10.1016/j.jcp.2013.08.014).
- K. Hirahara and K. Nishikiori. Estimation of frictional properties and slip evolution on a long-term slow slip event fault with the ensemble Kalman filter: Numerical experiments. *Geophys. J. Int.*, 219(3):2074–2096, 2019. doi:[10.1093/gji/ggz415](https://doi.org/10.1093/gji/ggz415).
- M. Kano, S. Miyazaki, K. Ito, and K. Hirahara. An adjoint data assimilation method for optimizing frictional parameters on the afterslip area. *Earth, Planets and Space*, 65(12):1575–1580, 2013. doi:[10.5047/eps.2013.08.002](https://doi.org/10.5047/eps.2013.08.002).

- M. Kano, S. Miyazaki, Y. Ishikawa, Y. Hiyoshi, K. Ito, and K. Hirahara. Real data assimilation for optimization of frictional parameters and prediction of afterslip in the 2003 Tokachi-oki earthquake inferred from slip velocity by an adjoint method. *Geophys. J. Int.*, 203(1):646–663, 2015. doi:[10.1093/gji/ggv289](https://doi.org/10.1093/gji/ggv289).
- M. Kano, S. Miyazaki, Y. Ishikawa, and K. Hirahara. Adjoint-based direct data assimilation of GNSS time series for optimizing frictional parameters and predicting postseismic deformation following the 2003 Tokachi-oki earthquake. *Earth, Planets and Space*, 72(1), 2020. doi:[10.1186/s40623-020-01293-0](https://doi.org/10.1186/s40623-020-01293-0).
- B. V. Kostrov. Seismic moment and energy of earthquakes, and seismic flow of rock. *Izv. Earth Phys.*, 1:23–40, 1974.
- J. E. Kozdon and L. C. Wilcox. Stable coupling of nonconforming, high-order finite difference methods. *SIAM J. Sci. Comput.*, 38(2):A923–A952, 2016. doi:[10.1137/15M1022823](https://doi.org/10.1137/15M1022823).
- J. E. Kozdon, E. M. Dunham, and J. Nordström. Interaction of waves with frictional interfaces using summation-by-parts difference operators, 2. extension to full elastodynamics. Technical Report 2010-018, Department of Information Technology, Uppsala University, June 2010.
- J. E. Kozdon, E. M. Dunham, and J. Nordström. Interaction of waves with frictional interfaces using summation-by-parts difference operators: weak enforcement of nonlinear boundary conditions. *J. Sci. Comput.*, 50(2):341–367, 2012. ISSN 0885-7474,1573-7691. doi:[10.1007/s10915-011-9485-3](https://doi.org/10.1007/s10915-011-9485-3).
- J. E. Kozdon, E. M. Dunham, and J. Nordström. Simulation of dynamic earthquake ruptures in complex geometries using high-order finite difference methods. *J. Sci. Comput.*, 55(1):92–124, 2013. doi:[10.1007/s10915-012-9624-5](https://doi.org/10.1007/s10915-012-9624-5).
- N. Lapusta, J. R. Rice, Y. Ben-Zion, and G. Zheng. Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *J. Geophys. Res. Solid Earth*, 105(B10):23765–23789, 2000. doi:<https://doi.org/10.1029/2000JB900250>.
- M. F. Linker and J. H. Dieterich. Effects of variable normal stress on rock friction: Observations and constitutive equations. *J. Geophys. Res. Solid Earth*, 97(B4):4923–4940, 1992. doi:<https://doi.org/10.1029/92JB00017>.
- T. Lundquist, A. Malan, and J. Nordström. A hybrid framework for coupling arbitrary summation-by-parts schemes on general meshes. *J. Comput. Phys.*, 362:49–68, 2018. doi:[10.1016/j.jcp.2018.02.018](https://doi.org/10.1016/j.jcp.2018.02.018).
- T. Lundquist, A. G. Malan, and J. Nordström. Efficient and error minimized coupling procedures for unstructured and moving meshes. *J. Comput. Phys.*, 406:109158, 21, 2020. doi:[10.1016/j.jcp.2019.109158](https://doi.org/10.1016/j.jcp.2019.109158).

- C. Marone. Laboratory-derived friction laws and their application to seismic faulting. *Annu. Rev. Earth Planet. Sci.*, 26(1):643–696, 1998. doi:[10.1146/annurev.earth.26.1.643](https://doi.org/10.1146/annurev.earth.26.1.643).
- T. Matsuda and Y. Miyatake. Generalization of partitioned Runge-Kutta methods for adjoint systems. *J. Comput. Appl. Math.*, 388:Paper No. 113308, 7, 2021. doi:[10.1016/j.cam.2020.113308](https://doi.org/10.1016/j.cam.2020.113308).
- K. Mattsson and M. H. Carpenter. Stable and accurate interpolation operators for high-order multiblock finite difference methods. *SIAM J. Sci. Comput.*, 32(4):2298–2320, 2010. doi:[10.1137/090750068](https://doi.org/10.1137/090750068).
- K. Mattsson, M. Almquist, and E. van der Weide. Boundary optimized diagonal-norm SBP operators. *J. Comput. Phys.*, 374:1261–1266, 2018. doi:[10.1016/j.jcp.2018.06.010](https://doi.org/10.1016/j.jcp.2018.06.010).
- A. Nissen, K. Kormann, M. Grandin, and K. Virta. Stable difference methods for block-oriented adaptive grids. *J. Sci. Comput.*, 65(2):486–511, 2015. doi:[10.1007/s10915-014-9969-z](https://doi.org/10.1007/s10915-014-9969-z).
- J. Nordström. Linear and nonlinear boundary conditions for wave propagation problems. In *Recent developments in the numerics of nonlinear hyperbolic conservation laws*, volume 120 of *Notes Numer. Fluid Mech. Multidiscip. Des.*, pages 283–299. Springer, Heidelberg, 2013. ISBN 978-3-642-33220-3; 978-3-642-33221-0. doi:[10.1007/978-3-642-33221-0_17](https://doi.org/10.1007/978-3-642-33221-0_17).
- O. O’Reilly, J. Nordström, J. E. Kozdon, and E. M. Dunham. Simulation of earthquake rupture dynamics in complex geometries using coupled finite difference and finite volume methods. *Commun. Comput. Phys.*, 17(2):337–370, 2015. doi:[10.4208/cicp.111013.120914a](https://doi.org/10.4208/cicp.111013.120914a).
- N. A. Petersson and B. Sjögreen. An energy absorbing far-field boundary condition for the elastic wave equation. *Commun. Comput. Phys.*, 6(3):483–508, 2009. doi:[10.4208/cicp.2009.v6.p483](https://doi.org/10.4208/cicp.2009.v6.p483).
- N. A. Petersson, O. O’Reilly, B. Sjögreen, and S. Bydlon. Discretizing singular point sources in hyperbolic wave propagation problems. *J. Comput. Phys.*, 321:532–555, 2016. doi:[10.1016/j.jcp.2016.05.060](https://doi.org/10.1016/j.jcp.2016.05.060).
- S. Peyrat and K. B. Olsen. Nonlinear dynamic rupture inversion of the 2000 Western Tottori, Japan, earthquake. *Geophys. Res. Lett.*, 31(5), 2004. doi:[10.1029/2003GL019058](https://doi.org/10.1029/2003GL019058).
- N. A. Pierce and M. B. Giles. Adjoint recovery of superconvergent functionals from pde approximations. *SIAM Rev.*, 42(2):247–264, 2000. doi:[10.1137/S0036144598349423](https://doi.org/10.1137/S0036144598349423).
- J. Premus, F. Gallovič, and J.-P. Ampuero. Bridging time scales of faulting: From coseismic to postseismic slip of the Mw 6.0 2014 South Napa, California earthquake. *Sci. Adv.*, 8(38):eabq2536, 2022. doi:[10.1126/sciadv.abq2536](https://doi.org/10.1126/sciadv.abq2536).

- K. Ranjith. Destabilization of long-wavelength love and stoneley waves in slow sliding. *Int. J. Solids Struct.*, 46(16):3086–3092, 2009. ISSN 0020-7683. doi:<https://doi.org/10.1016/j.ijsolstr.2009.04.010>.
- K. Ranjith. Instabilities in dynamic anti-plane sliding of an elastic layer on a dissimilar elastic half-space. *J. Elast.*, 115(1):47–59, Mar 2014. ISSN 1573-2681. doi:[10.1007/s10659-013-9446-1](https://doi.org/10.1007/s10659-013-9446-1).
- K. Ranjith and J. Rice. Slip dynamics at an interface between dissimilar materials. *J. Mech. Phys. Solids*, 49(2):341–361, 2001. ISSN 0022-5096. doi:[https://doi.org/10.1016/S0022-5096\(00\)00029-6](https://doi.org/10.1016/S0022-5096(00)00029-6).
- S. Ray and R. C. Viesca. Earthquake nucleation on faults with heterogeneous frictional properties, normal stress. *J. Geophys. Res. Solid Earth*, 122(10): 8214–8240, 2017. doi:<https://doi.org/10.1002/2017JB014521>.
- J. R. Rice. Spatio-temporal complexity of slip on a fault. *J. Geophys. Res. Solid Earth*, 98(B6):9885–9907, 1993. doi:<https://doi.org/10.1029/93JB00191>.
- J. R. Rice and A. L. Ruina. Stability of Steady Frictional Slipping. *J. Appl. Mech.*, 50(2):343–349, 06 1983. ISSN 0021-8936. doi:[10.1115/1.3167042](https://doi.org/10.1115/1.3167042).
- J. R. Rice, N. Lapusta, and K. Ranjith. Rate and state dependent friction and the stability of sliding between elastically deformable solids. *J Mech Phys Solids*, 49(9):1865–1898, 2001. doi:[10.1016/S0022-5096\(01\)00042-4](https://doi.org/10.1016/S0022-5096(01)00042-4).
- C. Rucker and B. A. Erickson. Physics-informed deep learning of rate-and-state fault friction. *Comput. Methods Appl. Mech. Eng.*, 430:117211, 2024. ISSN 0045-7825. doi:<https://doi.org/10.1016/j.cma.2024.117211>.
- J. W. Rudnicki and L. B. Freund. On energy radiation from seismic sources. *Bull. Seismol. Soc. Am.*, 71(3):583–595, 06 1981. ISSN 0037-1106. doi:[10.1785/BSSA0710030583](https://doi.org/10.1785/BSSA0710030583).
- A. Ruina. Slip instability and state variable friction laws. *J. Geophys. Res. Solid Earth*, 88(B12):10359–10370, 1983. doi:<https://doi.org/10.1029/JB088iB12p10359>.
- J. A. Samareh. Survey of shape parameterization techniques for high-fidelity multidisciplinary shape optimization. *AIAA Journal*, 39(5):877–884, 2001. doi:[10.2514/2.1391](https://doi.org/10.2514/2.1391).
- J. M. Sanz-Serna. Symplectic Runge-Kutta schemes for adjoint equations, automatic differentiation, optimal control, and more. *SIAM Rev.*, 58(1):3–33, 2016. doi:[10.1137/151002769](https://doi.org/10.1137/151002769).
- B. Sjögreen and N. A. Petersson. Source estimation by full wave form inversion. *J. Sci. Comput.*, 59(1):247–276, 2014. doi:[10.1007/s10915-013-9760-6](https://doi.org/10.1007/s10915-013-9760-6).

- V. Stiernström and M. Almquist. Order-preserving non-conforming grid interfaces for boundary-optimized summation-by-parts operators. In *The 15th International Conference on Mathematical and Numerical Aspects of Wave Propagation, Book of Abstracts, Version: July 17th, 2022*.
- V. Stiernström and M. Almquist. Testing and verification of numerical methods in Adjoint-based inversion for stress and frictional parameters in earthquake modeling, Sept. 2024. URL <https://doi.org/10.5281/zenodo.13776965>.
- V. Stiernström, M. Almquist, and K. Mattsson. Boundary-optimized summation-by-parts operators for finite difference approximations of second derivatives with variable coefficients. *J. Comput. Phys.*, 491:Paper No. 112376, 2023. doi:[10.1016/j.jcp.2023.112376](https://doi.org/10.1016/j.jcp.2023.112376).
- K. Torberntsson, V. Stiernström, K. Mattsson, and E. M. Dunham. A finite difference method for earthquake sequences in poroelastic solids. *Comput. Geosci.*, 22(5):1351–1370, 2018. doi:[10.1007/s10596-018-9757-1](https://doi.org/10.1007/s10596-018-9757-1).
- F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. ISBN 978-0-8218-4904-0. doi:[10.1090/gsm/112](https://doi.org/10.1090/gsm/112). Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.
- J. Tromp, C. Tape, and Q. Liu. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophys. J. Int.*, 160(1):195–216, 2005. doi:[10.1111/j.1365-246X.2004.02453.x](https://doi.org/10.1111/j.1365-246X.2004.02453.x).
- Y. van Dinther, H. R. Künsch, and A. Fichtner. Ensemble data assimilation for earthquake sequences: probabilistic estimation and forecasting of fault stresses. *Geophys. J. Int.*, 217(3):1453–1478, 2019. doi:[10.1093/gji/ggz063](https://doi.org/10.1093/gji/ggz063).
- R. C. Viesca. Frictional state evolution laws and the non-linear nucleation of dynamic shear rupture. *J. Mech. Phys. Solids*, 173:105221, 2023. ISSN 0022-5096. doi:<https://doi.org/10.1016/j.jmps.2023.105221>.
- J. Virieux and S. Operto. An overview of full-waveform inversion in exploration geophysics. *GEOPHYSICS*, 74(6):WCC1–WCC26, 2009. doi:[10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- S. Wang and N. A. Petersson. Fourth order finite difference methods for the wave equation with mesh refinement interfaces. *SIAM J. Sci. Comput.*, 41(5):A3246–A3275, 2019. doi:[10.1137/18M1211465](https://doi.org/10.1137/18M1211465).