The Role of Roughness in Earthquake Source Physics

by

Yuval Tal

Submitted to the Department of Earth, Atmospheric and Planetary Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Geophysics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

© Massachusetts Institute of Technology 2017. All rights reserved.
The role of roughness in earthquake source physics

by

Yuval Tal

Submitted to the Department of Earth, Atmospheric and Planetary Sciences
on July 31, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Geophysics

Abstract

Faults are rough at all scales and can be described as self-affine fractals. This deviation from planarity results in geometric asperities and a locally heterogeneous stress field, which affect the nucleation and propagation of shear rupture. I study this effect numerically at the scale of small earthquakes, in which realistic geometry and friction law parameters can be incorporated in the model. The numerical approach developed in this thesis includes three main features. First, to enable slip that is large relative to the size of the elements near the fault, as well to capture accurately of the variation of normal stress during slip, I implement slip-weakening and rate and state friction laws into the Mortar Finite Element Method, in which non-matching meshes are allowed across the fault and the contacts are continuously updated. Second, the mesh near the fault is refined using hanging nodes to enable accurate representation of the fault geometry. Finally, to model the whole seismic cycle, including a completely spontaneous nucleation process, the method uses variable time stepping with quasi-static and fully dynamic implicit schemes. The developed methodology is used to study the response of rough faults governed by rate and state friction to slow tectonic loading, where, in each simulation, the earthquake sequence includes at least two seismic cycles. With increasing roughness, there is a transition from seismic to aseismic slip behavior, in which the load on the fault is released by more slip events but with lower slip rate, seismic moment, and average static stress drop. We analyze the nucleation process in the fast slip events and show that the roughness introduces local barriers that complicate the nucleation process and result in asymmetric expansions of the rupture, non-monotonic increases in the slip rates on the fault, and the generation of multiple slip pulses. In general, the nucleation length increases with increasing roughness amplitude. However, there are large differences between first slip events in the sequences, where the initial conditions are homogenous, and later events, where the initial stress field and friction conditions are determined by the rupture growth and arrest in previous slip events.

Thesis Supervisor: Bradford H. Hager
Title: Cecil and Ida Green Professor of Earth Sciences
Acknowledgments

First, I would like to thank my advisor, Bradford Hager, for the opportunity to study at MIT and for the freedom and trust he gave me, while always leaving his door open for questions and discussions. I would like to thank Brian Evans for supervising my second generals project, where I gained important experience in experimental work, and for many discussions regarding the structure and physics of fault zones. I have benefited greatly from the advice and help of my thesis committee and my general exam committee: Robert Viesca, Matej Pec, Thomas Herring, and Sam Bowring. Dale Morgan provided a different and enriching perspective on geophysics, which I will take with me further. William Rodi provided fruitful discussions about how to insert the hypocenter locations error in my first general project. Zhulin Yu generated the maps in the paper that came out of my first general project. Haoyue Wang helped me a lot with running the simulations on the cluster. My experimental work would not have been possible without the help of Ulrich Mok, it was a great pleasure to work with him in the Lab.

The administrators at the EAPS Education Office and the ERL definitely make a graduate student's life easier, and I want to particularly thank Anna Shaughnessy, Joshua Kastorf, Theresa Macloon, Vicki McKenna, Roberta Allard, and Linda Meinke.

The support from and interaction with the fellow students, post-docs, and researchers at EAPS is an indispensable component of the journey here. The few I feel most indebted to are Yossi Cohen, Noa Bechor Ben-Dov, Andrey Shabelansky, Lucas Willemsen, Alan Richardson, Ali Fuad AlJishi, Saleh Al Nasser, Jamie Potter, Nasruddin Nazerali, Josimar Alves da Silva, Manuel Florez Torres, Alejandra Quintanilla Terminel, Eva Golos, Stephen Brown, Oleg Poliannikov, Chen Gu, Chunquan Yu, Farrokh Sheibani, Chunfang Meng, Saied Mighani, Yusuke Mukuhira, Hua Wang, Diming Yu, and Hamed Ibrahim.

Throughout my PhD I received financial support from the Aramco grant 6500009957, DOE grant DE-FE0009738, and the Southern California Earthquake Center (SCEC) grant 16108. SCEC is funded by NSF Cooperative Agreement EAR-1033462 and USGS Cooperative Agreement G12AC20038.

I am very grateful to my parents and brothers for their love and encouragement.
My deepest gratitude goes to my wife Irit, for her love and support and for walking with me on this journey, and to my daughters Adva and Netta.
# Contents

1 Introduction ........................................... 21  
  1.1 Background: Fault zones structure .................. 21  
  1.2 Previous studies on the role of roughness in earthquake mechanics ........ 22  
  1.3 Overview of this dissertation ....................... 24  
  1.4 Modeling considerations .............................. 26  

2 Dynamic mortar finite element method for modeling of shear rupture on frictional rough surfaces ............. 33  
  2.1 Introduction ........................................ 33  
  2.2 Problem definition ................................... 36  
    2.2.1 Slip-weakening friction .......................... 37  
    2.2.2 Rate and state friction ......................... 38  
  2.3 Weak form .......................................... 39  
  2.4 Finite element spatial discretization ................. 39  
  2.5 Time discretization ................................ 42  
  2.6 Solution with a semi-smooth Newton method ........... 43  
    2.6.1 Non-smooth complementarity functions ............ 43  
    2.6.1 Consistent linearization within the semi-smooth Newton method ....... 44  
      2.6.2.1 Slip-weakening friction discretization ........ 47  
      2.6.2.2 Rate and state friction discretization ........ 48  
    2.6.3 Algebraic representation ........................ 50  
    2.6.4 Primal–dual active set algorithm ................ 51
2.7 Examples ......................................................... 52
2.7.1 Quasi-static benchmark for rate and state friction ........ 52
2.7.2 Dynamic rupture with slip-weakening friction ............. 53
2.7.3 Dynamic Rupture with rate and state friction .......... 55
2.8 Conclusions .................................................. 57

3 The slip behavior and source parameters of spontaneous slip events on rough faults subjected to slow tectonic loading 71
3.1 Introduction .................................................. 71
3.2 Model description .......................................... 74
3.3 Results ..................................................... 76
  3.3.1 The effect of roughness on the earthquake sequence .... 76
  3.3.2 Seismic moment ...................................... 77
  3.3.3 Static stress drop ................................... 80
3.4 Discussion .................................................. 82
  3.4.1 Scaling relations .................................... 82
  3.4.2 Dimensionality ..................................... 84
3.5 Conclusions ................................................ 85

4 The effects of roughness on the nucleation process .......... 97
4.1 Introduction ................................................ 97
4.2 Model description ......................................... 99
4.3 Results ................................................... 101
  4.3.1 The nucleation process .............................. 101
  4.3.2 Nucleation length .................................. 105
4.3.3 Initiation of the rupture ........................................ 106
   4.3.3.1 Location on the fault .................................. 106
   4.3.3.2 Loading stage ........................................... 107

4.4 Discussion .......................................................... 108
   4.4.1 The effect of the initial conditions and fault geometry on the nucleation process ........................................... 108
   4.4.2 Detection of the nucleation stage at its relationship to the final size of the event ........................................... 110

4.5 Conclusions .......................................................... 112

5 Conclusion and future directions 123
   4.1 Conclusion ....................................................... 123
   4.2 Future directions ................................................ 125
## List of Figures

1-1  (a) A Schematic structure of a brittle fault zone in the upper half of the seismogenic layer (modified after Sibson [2003]). (b) The structure of the Gole Larghe fault zone, northern Italy, exhumed from a depth of 9-11 km (modified after Di Toro and Pennacchion [2005]). ........................................... 28

1-2  (a) Map of the main faults in southern California (after Plesch et al. [2007]). (b) Surface rupture map of the 1999, Hector Mine earthquake (after Klinger [2010]). ................................................................. 29

1-3  Roughness power spectral density over nine decades of wavelengths. The data include five parallel to slip direction exhumed faults and eight parallel and five normal to slip direction continental earthquake surface rupture traces (modified after Candela et al. [2012]). ............................. 30

1-4  The geometry of a 20 m rough fault with \( b_r = 0.005 \), \( \lambda_{min} = 20 \text{ cm} \), and \( H = 0.8 \). The scaled thickness of the black line that represents the fault is 4 cm, while the surrounding medium shown is 22 m x 4 m. ................................. 31

1-5  The relation between thickness and roughness of the PSZ and the minimum wavelength that should be considered in the mathematical problem for profiles with a thickness of 5 mm and different amplitude of roughness. Each profile is a superposition of sine functions with 10 m and 30 cm wavelengths. The black lines represent profiles with only the larger wavelength. In the case of the smoothest profile, the black line is completely within the PSZ, so the effect of the smaller wavelength is negligible. With increasing roughness, the black line is not restricted only to the PSZ, and the smaller wavelength and its corresponding stress variations also affect the deformation of the PSZ. ........................................... 32

2-1  Notation for finite deformation contact problem. ........................................... 59
2-2 (a) A linear SW friction law. (b) The change in friction in response to sudden increase in sliding velocity as predicted by the RS friction law.

2-3 Benchmark problem for RS friction. (a) Geometry and boundary conditions. (b) Friction coefficient versus relative slip at the central node of the upper side of the fault in response to a change in the velocity of the upper edge and consequently the slip rate on the fault. The red circles represent the numerical solution, while the black line is the analytic solution. Note that we begin with a non-matching grid on the fault.

2-4 (a) The problem set up: a 60 x 30 m elastic medium with a 20 m long fault is subjected to a prescribed slow horizontal velocity \( \pm V_0 \) at the top and bottom and zero vertical displacements on all boundaries, resulting in a simple shear loading conditions. (b) The geometry of the fault: self-affine fractal roughness geometry with the minimum wavelength of roughness set to 20 cm.

2-5 The distribution of shear stress around the fault at four different stages of a simulation with time discretization parameters of \( \alpha = 1, \beta = 0.3, \) and \( \gamma = 0.6. \) The reference for the time shown is the end the quasi-static loading stage (stage 1). The black circles show the locations where the sliding velocity and shear traction are measured in Figure 6.

2-6 The effect of the time discretization parameters on the time evolution of sliding velocity and shear traction at nodes A and B for a fault governed by SW friction. The nodes are located inside and outside the nucleation region, respectively (see Figure 2-5). Note that in all schemes we choose \( \gamma = 0.5\beta \) and the time is calculated from the end of the quasi-static loading stage.

2-7 Potential (a), kinetic (b), contact (c), and the energy balance (d) vs. time for the four time discretization schemes. Note that in all schemes we choose \( \gamma = 0.5\beta \) and the time is calculated from the end of the quasi-static loading stage.
2-8 Number of slipping nodes (black) and iterations (red) during the simulation for scheme #1 (solid) and scheme #4 (dashed).

2-9 (a) The evolution of slip along the fault with time for a simulation with time discretization scheme #2. The initial quasi-static stage is represented by red contours, with decreasing time intervals between the contours, while the dynamic stage is represented by black contours with the time interval between the contours set to a value of $5 \times 10^{-4}$ s. (b) The rupture velocity $V_r$ along the fault.

2-10 The effect of the time discretization parameters on the time evolution of sliding velocity and shear traction at nodes A and B for fault governed by RS friction. Note that in all schemes we choose $\gamma = 0.5\beta$ and the time is calculated from the end of the quasi-static stage.

2-11 Potential (a), kinetic (b), contact (c), and the energy balance (d) vs. time for the four time discretization schemes. Note that in all schemes we choose $\gamma = 0.5\beta$ and the time is calculated from the end of the quasi-static stage.

2-12 Number of slipping (black) and inactive (blue) nodes and number of iterations (red) during the simulation for scheme #1 (solid) and scheme #4 (dashed).

3-1 (a) The problem set up: a finite fault with length $L_f$ is embedded in a 2D elastic medium with dimensions $3L_f \times 1.5L_b$, which is subjected to a prescribed slow horizontal velocity $V_b = \pm \frac{L_f}{20} \times 10^{-10}$ m/s at the top and bottom, and initial stresses $\sigma_{xx0}$, $\sigma_{yy0}$, and $\sigma_{xy0}$. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. The model includes an absorbing layer with gradual Rayleigh damping. (b) The mesh structure around the left side portion of a rough fault with length of 30 m.
The mesh is refined with hanging nodes to give elements on the fault with dimensions of about 1.56 x 1.56 cm.

3-2 The fault profiles examined in this study: We consider a total of nine different general geometries for fault lengths of \( L_f = 20 \text{m} \) (a), 30 m (b), and 40 m (c). For each geometry eight profiles are generated with roughness pre-factor values of \( b_r = 0.001, 0.002, 0.005 \) and 0.01 and minimum wavelength of \( \lambda_{min} = 0.2 \text{m} \) and 1 m. For reference, we also run a simulation with a smooth fault for each length.

3-3 The evolution of maximum slip rate (a, c, and e) and average shear traction (b, d, and f) obtained for the fault geometries Geo-1 (20m), Geo-4 (30m), and Geo-7 (40m) shown in Figure 3-2, as well as for smooth faults.

3-4 The number of events vs. \( b_r \) for different values of \( L_f \) and \( L_{min} \). For rough faults, there are three different sequences for each of \( b_r, L_f \), and \( L_{min} \), and the number of event is the sum of all the events in these sequences divided by 3.

3-5 (a) The evolution of \( M_{0,1d} \) (black), \( M_{0,dynd,1d} \) (blue), and \( \dot{M}_{0,1d} \) (red) during 0.1 s of the first slip event of the sequence obtained for fault with Geo-4 (30 m), \( b_r = 0.001 \), and \( \lambda_{min} = 0.2 \text{m} \). We define the beginning of the dynamic stage of the rupture at the stage when the moment rate exceeds a threshold value of \( \dot{M}_{th,ini} = 5 \times 10^9 \text{N/s} \) and calculate \( M_{0,dynd,1d}(t) \) with the slip accumulated from that stage. (b) The evolution of shear traction at six locations along the fault. (c) Profiles of slip rate along the fault. (d) Profiles of slip. The stage where \( \dot{M}_{0,1d} < 5 \times 10^9 \) is represented by red contours at decreasing time intervals. The dynamic stage is represented by black contours for the propagation phase and gray contours for the arrest phase, both with time intervals of \( 1 \times 10^{-3} \text{s} \).

3-6 (a) The evolution of \( M_{0,1d}, M_{0,dynd,1d}, \) and \( \dot{M}_{0,1d} \) during the third slip event of the sequence obtained for fault with Geo-4 (30 m), \( b_r = 0.01 \), and \( \lambda_{min} = 1 \text{m} \). Most of the event is aseismic, but the figure represents a time range of
85 seconds, which includes three short sub-events where the rupture propagates dynamically ($\dot{M}_{0,1d} > 5 \times 10^9$). (b) Contours of the evolution of slip along the fault. The stage where $\dot{M}_{0,1d} < 5 \times 10^9$ is represented by red contours at decreasing time intervals. The dynamic stage is represented by black contours with time intervals of $1 \times 10^{-3}$ s. (c) The geometry of the fault.

3-7
(a) $M_{0,1d}$ vs. $b_r$ for $L_f = 20$ m (blue), $L_f = 30$ m (green), and $L_f = 40$ m (red) in all the slow (‘x’ symbols) and fast (‘o’ symbols) slip events in the 75 simulated earthquake sequences, as well as the average values (solid curves). The events on faults with $\lambda_{min} = 0.2$ are shown on the left and on those with $\lambda_{min} = 1$ are shown on the right. (b) The moment ratio $M_{0,dynd}/M_{0,1d}$ vs. $b_r$ for $\lambda_{min} = 0.2$ and $\lambda_{min} = 1$ for all the fast slip events, as well as the average values. Note that in the case of $b_r = 0.01$, $\lambda_{min} = 1$ m, and $L_f = 20$ m, there is only one fast slip event, so the average value is not representative.

3-8
The evolution of $\tau_e$ vs. normalized slip ($u/L$) at six locations on the fault, as well as the average behavior, during the first slip event in the sequence obtained for fault with $L_f = 30$ m, $b_r = 0.001$, and $\lambda_{min} = 1$ m (Geo-4). The slip and the shear traction are plotted from the beginning of the simulation, and circles denote the beginning of the slip event (black) and the dynamic stage of the rupture (blue) and end of the slip event (red). (b) The spatial distributions of shear traction along the fault at these stages.

3-9
(a) $\Delta \tau_e$ vs. $b_r$ for $L_f = 20$ m (blue), $L_f = 30$ m (green), and $L_f = 40$ m (red) in all the slow (‘x’ symbols) and fast (‘o’ symbols) slip events in the 75 simulated earthquake sequences, as well as the average values (solid curves). The events on faults with $\lambda_{min} = 0.2$ are shown on the left and on those with $\lambda_{min} = 1$ are shown on the right. (b) $\Delta \tau_{t(dyn)}$ vs. $b_r$ for $\lambda_{min} = 0.2$ and $\lambda_{min} = 1$ for all the fast slip events, as well as the average values. Note that in the case of $b_r = 0.01$, $\lambda_{min} = 1$ m, and $L_f = 30$ m, there is only one fast slip event, so the average value is not representative.
3-10 $\Delta \tau_t$ vs. $\overline{\Delta \tau_t}/L_{rup}$ (a) and $\Delta \tau_t(dyn)$ vs. $\overline{\Delta \tau_t(dyn)}/L_{rup}$ (b) for all the slow (filled symbols) and fast (open symbols) slip events in the 75 simulated earthquake. The color of the symbols denotes different values of $b_r$, while the type denotes different fault lengths. Slip events on fault with $\lambda_{min} = 0.2$ and $\lambda_{min} = 1$ are shown with thin and thick symbols, respectively. The static solution for the case a uniform shear stress change on a smooth 1-D finite fault embedded in a 2-D domain is also shown (solid line).

4-1 (a) The problem setup: a 40 m long finite fault is embedded in a 2D elastic medium with dimensions 120 x 80, which is subjected a prescribed slow horizontal velocity $V_b = \pm 10^{-9}$ at the top and bottom, and initial stresses $\sigma_{x0}$, $\sigma_{y0}$, and $\sigma_{xy0}$. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. The model includes an absorbing layer with gradual Rayleigh damping. (b) The fault profiles examined in this study. We consider a total of three different general geometries, and for each geometry, eight profiles are generated with roughness pre-factor values of $b_r = 0.001$, 0.002, 0.005 and 0.01 and minimum wavelength of $\lambda_{min} = 0.2$ m and 1 m. For reference, we also run a simulation with a smooth fault.

4-2 (a) The evolution of the $\tau_{xy}$ vs. time for the eight profiles of Geo-1, as well as for a smooth fault. (b) The evolution of the $\tau_{xy}$ vs. time during the first slip events in the sequences obtained for faults with $\lambda_{min} = 0.2$ m, and $b_r = 0.001$ (right) and 0.002 (left). (c) The evolution of the maximum slip rate on the fault vs. time for the eight profiles of Geo-1, as well as for smooth faults.

4-3 Profiles of $u$, $v$, and $\mu$ along the fault for a smooth fault (a, c, and e) and a rough fault with Geo-1, $b_r = 0.001$, and $\lambda_{min} = 0.2$ m (b, d, and f). The contours of $\mu$ are shown only for the first slip event in each sequence.
contours of $u$ are plotted for all the stages in the seismic cycle, i.e. the loading (dashed purple), nucleation (red), and dynamic (black) stages, where the dynamic stage includes both the propagation and arrest of the rupture. The final stage of each slip event is shown in blue. In the case of the first event of the sequence obtained for a smooth fault, the contours of $v$ and $\mu$ are plotted for the whole dynamic stage, with gray contours for the arrest phase. However, for better visualization of the nucleation stage, the contours of $v$ and $\mu$ are shown only for the nucleation stage and the beginning of the dynamic stage in other slip events. Moreover, the contours of $\mu$ are shown only for the first event in each sequence and we plot only six contours (with different colors) for the nucleation stage.

Profiles of $u$, $v$, and $\mu$ along rough faults with Geo-1, $\lambda_{\text{min}} = 0.2$ m, and $b_r = 0.002$ (a, c, and e) and $0.005$ (b, d, and f). The contours of $u$ are plotted for all the stages in the seismic cycle, i.e. the loading (dashed purple), nucleation (red), and dynamic (black) stages, where the dynamic stage includes both the propagation and arrest of the rupture. The final stage of each of the slip events is shown in blue. For $b_r = 0.005$, the contours of $v$ are shown only for the first and second slip events of the sequence. For better visualization of the nucleation stage, the contours of $v$ and $\mu$ are shown only for the nucleation stage and the beginning of the dynamic stage. Moreover, the contours of $\mu$ are shown only for the first event in each sequence and we plot only six contours (with different colors) for the nucleation stage.

The nucleation lengths of the fast slip events in the 25 simulated earthquake sequences. Only slip events where the slip accumulated during the dynamic stage of the rupture is more than half of the total slip are considered. Slip events that are the first in the sequence are shown in blue, while the later events are shown in red, both with open circles for faults with $\lambda_{\text{min}} = 0.2$ m and ‘+’ symbols for faults with $\lambda_{\text{min}} = 1$ m. The solid curves represent the average values for the first (blue) and later (red) slip events on faults.
with $\lambda_{\text{min}} = 0.2$ m (solid) and $\lambda_{\text{min}} = 0.2$ m (dashed).

4-6 Profiles of $u$ and $v$ for the loading (purple) and the beginning of the nucleation stage (red) during the first slip events in the sequences obtained for faults with Geo-2, $\lambda_{\text{min}} = 0.2$ m, and $b_r = 0.001$ (a and c) and 0.002 (b and d).

4-7 The locations where the fast slip events in the 25 simulated earthquake sequences initiate at the beginning of the nucleation stage. The plot shows the absolute distance of the locations from the center of the fault. Slip events that are first in the sequence are shown in blue, while the later events are shown in red, both with open circles for faults with $\lambda_{\text{min}} = 0.2$ m and '+' symbols for faults with $\lambda_{\text{min}} = 1$ m.

4-8 The evolution of the shear tractions $\tau$, $\nu$, $\theta$, and $\mu$ at six loading stages between the end of the first slip event and the beginning of the second slip event for a smooth fault and a fault with Geo-1, $b_r = 0.002$ and $\lambda_{\text{min}} = 1$ m. The six loading stages are shown on curves of the average shear stress on the fault vs. time.

4-9 (a) Curves of the evolution of $\dot{M}_{0,1d}$ during 18 s near their maximum values for the first (solid) and later (dashed) slip events in the earthquake sequences shown in Figures 4-3 and 4-4, except for the first slip event for $b_r = 0.005$. To show all slip events in a single plot, the curves begin at the time when the moment rate exceeds a value of $\dot{M}_{0,1d} = 10^6$ N/s for the last time before it approaches its maximum value. The black circles denote the transition between the nucleation and the dynamic stages, as is defined in this study. (b) The evolution of $\dot{M}_{0,1d}$ during 0.05 s near its maximum value for the same slip events. The curves are aligned such that they all begin when $\dot{M}_{0,1d}$ exceeds a value of $2 \times 10^6$ N/s for the last time before it approaches the maximum value. The black circles denote the transition between the nucleation and the dynamic stages.
List of Tables

3-1  Model Parameter Values. ..................................................... 96
Chapter 1

Introduction

1.1 Background: Fault zone structure

Fault zones are geometrically complex and their characterization depends on the scale of observation [Ben-Zion and Sammis, 2003]. Considering the structure of an individual fault, it is generally composed of one or more principal slip zones (PSZ), a few mm to several cm thick, within a fault core (FC), that may be surrounded by an associated zone of fractures known as the damage zone (DZ) [Sibson, 2003]. A schematic representation of a brittle fault zone for faults exhumed from shallow depth is shown in Figure 1-1a. Although the fault core thickness may vary considerably (Figure 1-1b), a range of studies on active and exhumed faults from different depths, with different host rock lithologies, and different fault core compositions and thicknesses suggests that the deformation during an individual slip event is mostly accommodated within the principal slip zone and may even be localized further on slip surfaces tens of micron wide and less. The slip zones are often localized within a fault core [Chester and Chester, 1998; Katz et al., 2003; Sibson, 2003; Wibberley and Shimamoto, 2005; De Paola et al., 2008 (faults with total slip > 100 m); Boullier et al., 2009; Smith et al., 2011], may be located at the boundary between the damaged zone and the fault core [Sagy and Brodsky, 2009; Barth et al., 2013; Kirkpatrick and Brodsky, 2014], and may be developed in intact rock [Di Toro and Pennacchioni, 2005; De Paola et al., 2008 (total slip < 10 m)].

While, in terms of modeling, the observations that most of the deformation during a slip event is restricted to a very thin PSZ somewhat simplify the problem, a source of complexity arises from the deviation of the PSZ from planarity. At large scales, map traces of the major fault systems in southern California (Figure 1-2a) and high-resolution maps of large continental strike-slip earthquake surface ruptures [Klinger, 2010] clearly show that faults, and consequently their PSZ, deviate from planarity. In the latter, the resolution is higher, with minimum roughness wavelength of about 200 m.
At scales between 10 μm and 20 m, measurements of roughness on exhumed faults suggest that faults have fractal geometry with a self-similar [Brown and Scholz, 1985; Power et al., 1987; Power and Tullis, 1991] or self-affine [Renard et al., 2006; Sagy et al., 2007; Candela et al., 2009, 2012; Bistacchi et al., 2011; Brodsky et al., 2011] behavior. In the spatial domain, the roughness can be measured by the average deviation of the profile from planarity (RMS height), which for a profile with length $L_f$ is expressed as

$$h(L_f) = b_r L_f^H,$$

where $b_r$ is the RMS pre-factor and $H$ is the Hurst exponent, which is equal to 1 in the case of a self-similar fractal and lower than 1 in the case of a self-affine fractal. The spectral density quantifies the roughness as a power law in the form of

$$P(k) = C k^{-(2H+1)},$$

where $k$ is the spatial frequency and $C = 2H b_r^2$ is a pre-factor.

Candela et al. [2012] compiled rupture trace data from Klinger [2010] with measurements from exhumed faults to generate a description of roughness over nine decades of length scales (Figure 1-3). They suggested a self-affine geometry with $H = 0.6 - 0.8$ and $b_r$ ranging from 0.001 to 0.01 in the slip direction. However, while measurements at a particular wavelength bandwidth seems to fit self-affine surfaces, the data as a whole may be better fit by a single self similar fractal, as pointed out by Shi and Day [2013]. It is important to note that earthquake surface ruptures include also traces perpendicular to the direction of slip (thrust and normal fault), and that even in the case of strike-slip faults, the traces may not be completely aligned with the direction of slip, as discussed in Candela et al. [2012].

### 1.2 Previous studies on the role of roughness in earthquake mechanics

Experimental observations show that, on the lab-scale, roughness affects frictional constitutive parameters such as stress drop and critical slip weakening distance and results in larger nucleation zones [Okubo and Dieterich, 1984; Ohnaka and Shen, 1999; Ohnaka, 2003]. Moreover, Ohnaka [2003] compared laboratory measurements of slip
with earthquake observations over a wide range of scales and argued that roughness properties in the fault zone govern the final size of earthquakes.

Numerical static analyses of the response of rough faults to an external increase in shear load [Dieterich and Smith, 2009] or a prescribed uniform shear stress reduction on the fault [Zielke et al., 2017] show that the slip on the fault (as well as the area of the rupture in the later study) decreases with increasing roughness amplitude. Using a quasi-static cascade rupture model for an earthquake, Candela et al., [2011] suggest that the stress drop decreases with increasing fault dimension because of the self-affine fractal geometry of faults. However, although these studies provide important first order information regarding the effects of roughness on the source parameters, they do not take into accounts the dynamics of the rupture and the effect of friction laws and may be oversimplified. Moreover, they do not provide any information regarding the maximum slip rate on the fault during the events, thus whether they happen seismically or aseismically.

Several numerical studies of dynamic rupture on rough faults governed by frictional laws have been performed, all of them with strongly rate-weakening friction laws and visco-plasticity for the bulk material [Dunham et al., 2011; Fang and Dunham, 2013; Shi and Day, 2013; Bruhat et al., 2016]. These studies show that roughness promotes the development of self-healing rupture pulses, as well as secondary slip pulses that rerupture previously slipped fault sections, substantial fluctuations in rupture velocity with supershear transitions, heterogeneous slip distribution and inelastic deformation. Moreover, it provides a realistic mechanism for rupture arrest.

The dynamic simulations above are limited to a single rupture and adopt a constant time step size and an artificial nucleation procedure. The question then arises: would the modeled ruptures nucleate spontaneously into fast seismic events when the faults are only loaded by tectonic stress. Moreover, these models assume a spatially uniform initial stress field and friction parameters, but the slip and stress heterogeneities at the end of slip events affect the rupture process in subsequent events. In addition, the contact formulation assumes that the grid points are collocated on either side of the fault during all stages of the simulation, thus the range of roughness wavelengths is limited and the variations of the normal stresses on the fault during slip may be underestimated.
1.3 Overview of this dissertation

So far, numerical studies of rough faults have considered only quite a limited roughness bandwidth and modeled only the dynamic stage of the shear rupture. In this thesis, I aim to study the role of roughness in all the stages of the seismic cycle, i.e. slow tectonic loading, earthquake nucleation, and dynamic propagation and arrest of the shear rupture. Moreover, by modeling sequences of at least two earthquakes I study how the stress heterogeneities and friction properties at the end of slip events affect the rupture process at subsequent events.

Any modeling study involves the idealization of the physical problem to a mathematical model. In the remainder of this chapter, I discuss the approximation of the principal slip zones as a discrete surface.

In chapter 2, I present a mortar-based finite element formulation for modeling the dynamics of shear rupture on rough interfaces governed by slip-weakening and rate and state friction laws. To model faults with roughness extending over a large range in wavelengths, the contact formulation should allow slip that is large relative to the element size near the fault and should accurately capture the variation of normal stress during slip. In the mortar method, non-matching meshes are allowed across the fault, the contacts are continuously updated, and Lagrange multipliers are used to enforce the continuity of stress, the non-penetration condition, and the frictional resistance on the fault in a weak integral sense. I extend the 2-D large sliding mortar formulation to dynamic problems and implement slip-weakening and rate and state friction laws into the method. The friction laws are discretized and linearized such that the efficient semi-smooth Newton algorithm for the solution of the nonlinear system of equations is preserved. Moreover, the discretization of the RS friction law involves a procedure to condense out the state variables, thus eliminating the addition of another set of unknowns into the system. Because I use a finite element based method, the meshing structure is quite flexible. In this study, I refine the mesh near the fault using a hanging nodes technique, thereby enabling accurate representation of the fault geometry.

The method developed in chapter 2 uses a quasi-static backward Euler time discretization scheme when inertial effects are negligible and implicit Newmark or mid-
point schemes for dynamic analysis. Because all these schemes are implicit, the implementation of variable time stepping is straightforward, and the whole seismic cycle can be modeled, including a completely spontaneous nucleation process. However, it is important to note that, although from the contact perspective the slip on the fault can be large, the amount of slip and consequently the number of cycles are limited because the model accounts currently only for an elastic medium. Otherwise, unphysical stresses may be developed in the medium surrounding the fault.

In chapter 3, I use the developed methodology to study the response of rough faults governed by rate and state friction to slow tectonic loading, where, in each simulation, the earthquake sequence includes at least two seismic cycles. This chapter focuses on the global behavior of the faults and examines the effect of the roughness amplitude and minimum wavelength and the fault length on the stress drop and seismic moment of the slip events, and on whether the fault slips seismically or aseismically. I study the scale of small earthquakes (faults with length of $L_f = 20 - 40$ m), and choose the minimum roughness wavelength to be at a size close to the lab samples (20 cm) and thus use the observed lab-scale rate and state friction laws without up scaling the constitutive parameters, assuming that the experimental data already include the effects of smaller wavelengths of roughness.

In chapter 4, I analyze further the simulations performed in chapter 3 with a fault length of 40 m and examine the effects of roughness on the nucleation process in the fast slip events. Because the simulations include sequences of at least two seismic cycles, they enable me to examine the effect of the stress state and frictional conditions resulting from the rupture growth and arrest in a given slip event on the nucleation process for subsequent events.

Finally, in chapter 5, I summarize the major contributions of this thesis and suggest possible future research directions.
1.4. Modeling considerations

Because of incompleteness of data, limited computational power, and mathematical challenges, only some aspects of the physical complexities in the fault zone can be addressed in a single numerical method. The relative importances of different aspects of earthquake mechanics are quite subjective, as well as the assumptions and simplifications that can be made. In this section, I discuss the approximation of the principal slip zones as a discrete surface.

Figure 1-4 shows the geometry of a 20 m rough fault with PSZ thickness of 4 cm. Whether the wall rocks of the PSZ are intact rocks, damaged rocks, fault rocks, or a combination of them, because of its small thickness one cannot model both the internal physical processes in the PSZ [e.g. Lyakhovsky and Ben-Zion, 2014; Platt et al., 2014; Rice et al., 2014] and a complex interaction of the PSZ with the surrounding medium. Therefore, to model the effects of geometrical irregularities and spatial variations of mechanical properties in the medium, the PSZ has to be approximated as a discrete surface. Moreover, striated fault surfaces suggest that in some cases a description of a slip on a discrete surface is more appropriate.

Constitutive laws for the shear resistance of the PSZ, whether it is a very thin layer or a discrete surface, can be provided by experimental data. In the case of slow slip rate, experiment on bare surfaces [Dieterich, 1979] and gouge layers [Karner et al., 1998; Mair and Marone, 1999] can be generalized into a rate and state friction description. At seismic slip rates, a large set of experiments performed on cohesive and non-cohesive rocks shows that the friction coefficient decays exponentially with slip, independent of the rock and weakening mechanism involved [Di Toro et al., 2011]. In cases where the experimental data are insufficient, studies of the internal physical processes in the PSZ, as these mentioned in the previous paragraph, can provide the constitutive law for models of faults with a complex structure.

In the context of rough faults, a PSZ with a finite thickness can be approximated as a mathematical surface as long as only geometric irregularities that are not masked by the thickness of PSZ are considered. In general, a wavelength should be included if its corresponding amplitude is larger than half the thickness of the PSZ, as demonstrated in
Figure 1-5. The figure shows three profiles with thickness of 5 mm and different amplitudes of roughness. Each profile is a superposition of sine functions with 10 m and 30 cm wavelengths. The amplitude to wavelength ratio of the two wavelengths is the same (self similar description). In the case of the smoothest profile, all of the shear deformation can localize onto a surface that is completely within the PSZ but has no 30 cm wavelength component, so the effect of the smaller wavelength is negligible. With increasing roughness a surface with no component of the smaller wavelength cannot be restricted only to the PSZ. Therefore, to allow deformation only in the PSZ, the surface must also have a component of the smaller wavelength. In the following, a simple calculation of the minimum wavelength for a given roughness and PSZ thickness is shown.

Assuming that a sine function can represent the total contributions of all wavelengths smaller than or equal to a wavelength, \( \lambda \), the average deviation of the profile from planarity is

\[
h \approx \frac{1}{\lambda/2} \int_{0}^{\lambda} \sin \left( \frac{2\pi}{\lambda} x \right) A dx = \frac{2A}{\pi},
\]

where \( A \) is the amplitude corresponding to \( \lambda \). Following eq. (1.1), the RMS height of a segment with length \( L = \lambda \) is

\[
h = b \lambda^H.
\]

Equating and substituting the condition \( A(\lambda_{min}) = T/2 \), gives

\[
\lambda_{min} = \left( \frac{T}{\pi b} \right)^{1/H},
\]

where \( T \) is the thickness of the PSZ.
Figure 1-1: (a) A schematic structure of a brittle fault zone in the upper half of the seismogenic layer (modified after Sibson [2003]). (b) The structure of the Gole Larghe fault zone, northern Italy, exhumed from a depth of 9-11 km (modified after Di Toro and Pennacchion [2005]).
Figure 1-2: (a) Map of the main faults in southern California (after Plesch et al. [2007]). (b) Surface rupture map of the 1999, Hector Mine earthquake (after Klinger [2010]).
Figure 1-3: Roughness power spectral density over nine decades of wavelengths. The data include five parallel to slip direction exhumed faults and eight parallel and five normal to slip direction continental earthquake surface rupture traces (modified after Candela et al. [2012]).
Figure 1-4: The geometry of a 20 m rough fault with $b = 0.005$, $\lambda_{\text{min}} = 20$ cm, and $H = 0.8$. The scaled thickness of the black line that represents the fault is 4 cm, while the surrounding medium shown is 22 m x 4 m.
Figure 1-5: The relation between thickness and roughness of the PSZ and the minimum wavelength that should be considered in the mathematical problem for profiles with a thickness of 5 mm and different amplitude of roughness. Each profile is a superposition of sine functions with 10 m and 30 cm wavelengths. The black lines represent profiles with only the larger wavelength. In the case of the smoothest profile, the black line is completely within the PSZ, so the effect of the smaller wavelength is negligible. With increasing roughness, the black line is not restricted only to the PSZ, and the smaller wavelength and its corresponding stress variations also affect the deformation of the PSZ.
Chapter 2

Dynamic mortar finite element method for modeling of shear rupture on frictional rough surfaces

Abstract

This paper presents a mortar-based finite element formulation for modeling the dynamics of shear rupture on rough interfaces governed by slip-weakening (SW) and rate and state (RS) friction laws, focusing on the dynamics of earthquakes. The method utilizes the dual Lagrange multipliers and the primal–dual active set strategy concepts, together with a consistent discretization and linearization of the contact forces and constraints, and the friction laws to obtain a semi-smooth Newton method. The discretization of the RS friction law involves a procedure to condense out the state variables, thus eliminating the addition of another set of unknowns into the system. Several numerical examples of shear rupture on frictional rough interfaces demonstrate the efficiency of the method and examine the effects of the different time discretization schemes on the convergence, energy conservation, and the time evolution of shear traction and slip rate.

2.1 Introduction

Understanding the mechanics of shear rupture on a frictional interface is important for fields and scales ranging from earthquakes to car brakes. In this paper, we introduce a numerical method designed for studying the dynamics of shear rupture during earthquakes, focusing on the effects of the non-planar geometry and the non-linear frictional constitutive laws associated with natural faults. The method, however, can easily be adjusted for other fields and scales.

A common view is that earthquakes occur via a frictional instability, in which the
frictional resistance on the fault decreases with increasing sliding or sliding velocity [e.g. Scholz, 2002]. In the context of earthquakes, the constitutive relations for the evolution of the friction coefficient can be divided into two main groups. In the group of slip-weakening (SW) friction laws, the friction coefficient evolves as a function of slip, while in the group of rate and state (RS) friction laws, it evolves as a function of sliding velocity and state variables. Many studies have examined numerically the effects of these laws on the dynamics of shear rupture during earthquakes (see, e.g., [Bizzarri et al., 2001] and references therein). However, another source of complexity arises from the deviation of faults from planarity, which results in geometric asperities and a locally heterogeneous stress field near the fault. Map traces of the major fault systems, high-resolution maps of large continental strike-slip earthquake surface ruptures [Klinger, 2010; Candela et al., 2012], and measurements of roughness on exhumed faults at scales between 10 μm and 20 m [Brown and Scholz, 1985; Power et al., 1987; Power and Tullis, 1991; Renard et al., 2006; Sagy et al., 2007; Candela et al., 2009, 2012; Bistacchi et al., 2011] show that faults are rough at all scales and can be described as self-affine fractal surfaces.

So far, several numerical studies of dynamic rupture on rough faults governed by frictional constitutive laws have been performed [Dunham et al., 2011; Fang and Dunham, 2013; Shi and Day, 2013; Bruhat et al., 2016]. However, these studies (as well as the numerical studies mentioned earlier) assume that the slip on the fault is small compared to the size of the elements on the fault, thus the grid points are considered as collocated on either side of the fault during all stages of the simulation. While this assumption has a small effect for the amount of slip and the roughness bandwidth considered in these studies, it limits the minimum wavelength of roughness on the fault and may underestimate the variations of the normal stresses on the fault during slip. Therefore, a method in which the interface is governed by friction laws but also allows nonconforming meshes across the fault, with a continuous updating of the contact geometry, is needed.

The most prevalent discretization strategy in the context of large sliding contact problems is the node-to-segment (NTS) approach, in which the nodes of one surface are prohibited from penetrating the segments of the opposing surface. However, this
discretization does not satisfy the contact patch test [Papadopoulos and Taylor, 1992], where a flat contact surface should be able to exactly transmit a spatially constant contact pressure. Moreover, the non-smoothness of the discretized contact surfaces may lead to convergence difficulties and non-physical oscillations of the contact forces [Yang et al., 2005]. Therefore, although various smoothing algorithms for the NTS formulations have been proposed, segment-to-segment discretization techniques have become more attractive, especially the Mortar method. The method was originally introduced in the context domain decomposition method [Bernardi et al., 1993] for coupling of nonconforming discretizations across interfaces. It enforces the continuity of stresses and the contact conditions across the interface in a weak integral sense, rather than as strong, pointwise constraints.

Mortar formulations for finite sliding with penalty or augmented Lagrangian methods to enforce the contact constraints can be found in [Puso and Laursen, 2004; Fischer and Wriggers, 2005, 2006; Yang et al., 2005; Puso et al., 2008]. However, the former method uses unphysical penalty parameters that can affect the accuracy and the latter method involves additional iterative procedures. The traditional direct Lagrange multiplier method avoids these drawbacks and exactly fulfills the contact constraints, but leads to an increased system of equations, with the Lagrange multipliers as additional unknowns. A remedy for this problem was given in [Wohlmuth, 2000], who introduced the dual spaces discretization of the Lagrange multipliers into the Mortar method and enabled an efficient local elimination of the discrete Lagrange multipliers by static condensation. This concept was combined further with the primal–dual active set strategy to give an efficient semi-smooth Newton algorithm for the solution of the nonlinear system of equations [Hueber and Wohlmuth, 2005; Brunssen et al., 2007; Hueber et al., 2008]. The method was extended by [Popp et al., 2009; Gitterle et al., 2010] to quasistatic finite deformation contact problems, including a consistent linearization of the contact virtual work expression and the nonlinear contact constraints. Extension to small deformation dynamic contact problems can be found in [Hager et al., 2008; Hager and Wohlmuth, 2009].

Although some of the Mortar formulations above include sliding on frictional interfaces, only the case of Coulomb friction with a constant coefficient of friction has
been considered. In order to model shear rupture on rough frictional interfaces, this work extends the Mortar formulation in [Popp et al., 2009; Gitterle et al., 2010] to dynamic problems and consistently implements the SW and RS friction laws into the method. While the implementation of the SW friction law is straightforward, the implementation of the RS friction law involves a procedure to condense out the state variables, thus eliminate the addition of another set of unknowns into the system. We believe that the method provides a robust tool to study different scales of the physics of earthquakes, as well as other fields that involve shear rupture on rough frictional interfaces.

The remainder of the paper is organized as follows: In Section 2 we introduce the finite deformation frictional contact problem and describe the SW and RS friction laws. The corresponding weak formulation is presented in section 3. Spatial finite element discretization of the contact virtual work and contact constraints with dual Lagrange multipliers is provided in section 4. Time discretization of the resulting force equilibrium equation and discretized contact constraints is given in section 5. In Section 6, the semi-smooth Newton method for the solution of the resulting discretized system of equations is described, and the discretized form of the friction laws and their associated directional derivatives are provided. In section 7, numerical results are presented to show the accurate implementation of the friction laws, and evaluate convergence energy preservation properties for different time integration schemes. Finally, some conclusions are given in Section 8. It is important to note here that we follow the finite deformation Mortar formulation of [Popp et al., 2009; Gitterle et al., 2010] to make the method more general, but practically to represent the friction laws accurately small time steps have to be adopted and many of the directional derivative calculations in the linearization of the virtual work and the normal and frictional contact constraints can be neglected.

### 2.2 Problem definition

We consider a two-dimensional contact problem with finite deformation and finite frictional sliding on an interface governed by SW or RS friction laws (Figure 2-1). Although only the problem of two contacting bodies is shown here, an extension to multiple bodies or fractures embedded in a continuous domain is straightforward. The initial boundary value problem is given by
\[ \text{Div}(F \cdot S) + b_0 = \rho_0 \ddot{u} \quad \text{in } \Omega, \]
\[ u = \ddot{u}, \quad \text{on } \Gamma_u, \]
\[ \tau = \ddot{\tau}, \quad \text{on } \Gamma_\tau, \]
\[ u^0 = \ddot{u}^0, \quad \text{in } \Omega_0, \]
\[ \ddot{u}^0 = \ddot{u}^0, \quad \text{in } \Omega_0, \]

where \( F \) is the deformation gradient tensor, \( S \) is the second Piola–Kirchhoff stress tensor, \( \rho_0 \) is the density in the reference configuration, \( b_0 \) is the body load, and \( \ddot{\tau} \) is the boundary traction.

In the normal direction, the Karush–Kuhn–Tucker conditions are given by
\[ g_n \geq 0, \quad \tau_n \leq 0, \quad \tau_n g_n = 0. \] (2.2)
Here \( \tau_n \) is the normal component of the current contact tractions \( \tau_c \) and \( g_n \) is a gap function in the normal direction, defined as
\[ g_n = -n \left( x^{(s)}(x^{(s)}) \right) \cdot \left[ x^{(s)}(x^{(s)}) - x^{(m)}(x^{(m)}) \right], \] (2.3)
where \( x^{(m)} \) is the projection of \( x^{(s)} \) onto the current master surface \( y_c^{(m)} \) along the current outward unit normal \( n \) and \( x^{(s)} \) and \( x^{(m)} \) are the corresponding points in the reference configuration (see Figure 2-1).

In the tangential direction, the frictional conditions with a variable friction coefficient \( \mu \) are enforced by
\[ \psi := |\tau_t| - \mu |\tau_n| \leq 0, \quad v_t + \beta \tau_t = 0, \quad \beta \geq 0, \quad \psi \beta = 0. \] (2.4)
Here \( \tau_t \) is the tangential contact traction and \( v_t \) is tangential relative velocity defined as
\[ v_t(X) = t \left( x^{(s)}(x^{(s)}) \right) \cdot \left[ \dot{x}^{(s)}(x^{(s)}) - \dot{x}^{(m)}(x^{(m)}) \right], \] (2.5)
where \( t \) is the current unit tangential vector.

### 2.2.1 Slip-weakening friction

In the simple SW friction law [Jda, 1972; Palmer and Rice, 1973; Andrews, 1976] the coefficient of friction \( \mu \) drops linearly from its static value, \( \mu_s \), to its sliding value, \( \mu_d \), over a specified distance, \( d_c \), (Fig. 2a).
\[ \mu = \begin{cases} \mu_s + \frac{\mu_d - \mu_s}{d_c} u_t, & u_t \leq d_c, \\ \mu_d, & u_t > d_c \end{cases}, \] (2.6)
where $u_t$ is the tangential relative slip along the contact. More complicated slip based friction laws that may also include an initial stage of hardening and then exponential decay of the friction coefficient with slip were also suggested [Ohnaka and Yamashita, 1989; Di Toro et al., 2011]. We consider here only the simple SW friction law, but the other slip-based laws can be implemented in a similar way.

### 2.2.2 Rate and state friction

The response of the friction coefficient to a change in sliding velocity is shown schematically in Figure 2-2b. This behavior was observed experimentally for many materials [Dieterich, 1979] and is the basis of the empirical RS friction laws of [Dieterich, 1979; Ruina, 1983]. With a sudden increase in sliding velocity there is an instantaneous increase of the friction coefficient followed by an evolution stage, in which it decreases to a new steady-state value. These behavior is governed by two material property constants $a$ and $b$, respectively. Frictional instability can occur only if the steady-state velocity dependence in friction coefficient is velocity weakening, i.e. $a - b < 0$.

Several variations of RS friction laws have been proposed, but the aging law is in the best agreement with experimental observations [Beeler et al., 1994; Scholz, 2002]. In this form of the law, the friction coefficient evolves as

$$
\mu = \mu^* + a \ln \left( \frac{v_t}{v^*} \right) + b \ln \left( \frac{v_t \theta}{L} \right),
$$

(2.7)

where $v^*$ is a reference velocity, $\mu^*$ the steady-state friction at $v_t = v^*$, and $\theta$ is a state variable governed by an aging law as

$$
\dot{\theta} = 1 - \frac{\theta v_t}{L},
$$

(2.8)

where $L$ is the critical slip distance. On the micro scale, $\theta$ is interpreted as the average age of contacts and $L$ as the slip necessary to renew surface contacts [Dieterich, 1979].
2.3 Weak form

Using appropriate spaces for the displacements $\mathbf{u}$ and virtual displacements $\delta \mathbf{u}$, the virtual work expression is given by

$$\delta \Pi(\mathbf{u}, \delta \mathbf{u}) = \delta \Pi_{\text{int,ext}}(\mathbf{u}, \delta \mathbf{u}) + \delta \Pi_c(\mathbf{u}, \delta \mathbf{u}).$$

(2.9)

where $\delta \Pi_{\text{int,ext}}(\mathbf{u}, \delta \mathbf{u})$ is the standard virtual work from internal and external forces and $\delta \Pi_c(\mathbf{u}, \delta \mathbf{u})$ the contact virtual work. We use the total Lagrangian formulation of [Bathe, 1996] to compute $\delta \Pi_{\text{int,ext}}$. Exploiting the balance of linear momentum across the contact interface and introducing Lagrange multipliers $\boldsymbol{\lambda} = -\boldsymbol{\tau}_c$ on the “slave” side of the contact, the contact virtual work is expressed as

$$\delta \Pi_c = \int_{\gamma_c} \lambda \cdot (\delta \mathbf{u}^{(1)} - \delta \mathbf{u}^{(2)}) d\gamma.$$  

(2.10)

In the normal direction, the Karush–Kuhn–Tucker conditions are enforced in a weak integral form and point wise as

$$\int_{\gamma_c} \delta \lambda_n g_n d\gamma \geq 0, \quad \lambda_n \geq 0, \quad \lambda_n g_n = 0, \quad \lambda_n g_n = 0,$$

(2.11)

In the tangential direction, the frictional conditions are given by

$$\int_{\gamma_c} \delta \lambda_t (\nu_{\text{rel}} - \beta \lambda_t) d\gamma = 0,$$

$$\psi := |\lambda_t| - \mu |\lambda_n| \leq 0, \quad \beta \geq 0, \quad \psi \beta = 0$$

(2.12)

2.4 Finite element spatial discretization

The geometry, displacements, and displacement time derivatives of the contacting slave and master surfaces are discretized with standard finite element shape functions as

$$\mathbf{u}^{(1)}|_{\gamma_c^{(1)}} = \sum_{j=1}^{n_{sl}} N_j \mathbf{d}_j, \quad \mathbf{u}^{(2)}|_{\gamma_c^{(2)}} = \sum_{j=1}^{n_{mas}} N_j \mathbf{d}_j,$$

(2.13)

where $n_{sl}$ and where $n_{mas}$ are the numbers of nodes on the slave surface on the slave surface $\gamma_c^{(1)}$ and master surface $\gamma_c^{(2)}$.

For the interpolation of the Lagrange multiplier field, dual shape functions $\Phi_j$ are introduced on the slave surface as
\[ \lambda = \sum_{j=1}^{n_{sl}} \phi_j z_j, \]  

(2.14)

where \( z_j \) are the discrete nodal Lagrange multipliers. These shape functions fulfill the so-called biorthogonality condition [Wohlmuth, 2000] as

\[ \int_{y_c^{(1)}} \phi_j N_k^{(1)} \, dy = \delta_{ij} \int_{y_c^{(2)}} N_k^{(1)} \, dy. \]  

(2.15)

A detailed description regarding the construction of the dual shape functions is given in [Hartmann et al., 2007; Popp et al., 2009]. For linear 1D surface elements the dual shape functions are given by:

\[ \phi_1 = \frac{1}{2} (1 - 3 \xi), \quad \phi_2 = \frac{1}{2} (1 + 3 \xi). \]  

(2.16)

Substituting (13–15) into (10) gives [Popp et al., 2009]

\[ \delta \Pi_c = (\delta d^{(1)})^T D_S z - (\delta d^{(2)})^T M_M^T z, \]  

(2.17)

which leads to the discrete vector of contact forces

\[ f_c(d, z) = [0, -M_M, D_S]^T z. \]  

(2.18)

Here \( D_S \in \mathbb{R}^{2n_{sl} \times 2n_{sl}} \) and \( M_M \in \mathbb{R}^{2n_{sl} \times 2n_{mas}} \) are coupling matrices arising from the mortar integrals, evaluated as

\[ D_S[j, j] = D_{jj} = \int_{y_c^{(1)}} N_j^{(1)} \, dy, \quad j = 1, \ldots, n_{sl} \]  

(2.19)

\[ M_M[j, l] = M_{jl} = \int_{y_c^{(2)}} \phi_j N_l^{(2)} \, dy, \quad j = 1, \ldots, n_{sl}, \quad l = 1, \ldots, n_{mas} \]  

(2.20)

The biorthogonality condition results in a diagonal matrix \( D_S \), which allows static condensation of the discrete Lagrange multipliers and simplifies the linearization and solution process. While numerical integration of the mortar matrix \( D_S \) involves simply the integration of the slave side displacement shape functions over the current slave contact, numerical integration of the mortar matrix \( M_M \) is more complex because it involves the product of master side shape functions and slave side dual shape functions over the slave contact surface. To perform this integration, we follow the approach in [Yang et al., 2005; Popp et al., 2009], in which the integration domain is discretized into contact segments, on which both shape functions are defined continuously.
Substituting eq. (2.18) into eq. (2.9), the algebraic form of the force equilibrium equation is given by

\[ \mathbf{M} \ddot{\mathbf{d}} + \mathbf{f}_{\text{int}}(\mathbf{d}) + \mathbf{f}_{\text{c}}(\mathbf{d}, z) - \mathbf{f}_{\text{ext}} = 0, \]  
(2.21)

where \( \mathbf{M} \) represents the mass matrix, \( \mathbf{f}_{\text{int}}(\mathbf{d}) \) is the vector of the deformation dependent internal forces, and \( \mathbf{f}_{\text{ext}} \) is the vector of external forces.

As shown in [Popp et al., 2009], the discretized form of the normal conditions in (2.11) is equivalent to the following set of pointwise conditions

\[ \bar{g}_{nj} \geq 0, \quad z_{nj} \geq 0, \quad z_{nj} \bar{g}_{nj} = 0, \]  
(2.22)

where the discrete weight gap function in the normal direction is given by

\[ \bar{g}_{nj} = -n_j^T \mathbf{D}_S[j,j] \mathbf{x}_j^{(s)} + n_j^T \sum_{l=1}^{n_{mas}} \mathbf{M}_M[l,l] \mathbf{x}_l^{(m)}. \]  
(2.23)

Following [Gitterle et al., 2010], the nodal tangential contact conditions are given by

\[ \psi_j = |z_{ij} - \mu_j z_{nj}| \leq 0, \quad \bar{v}_t - \bar{\beta}_j z_{tj} = 0, \quad \bar{\beta}_j \geq 0, \quad \psi_j \bar{\beta}_j = 0, \]  
(2.24)

with the weighted tangential relative velocity defined as

\[ \bar{v}_t = t_j^T \mathbf{D}_S[j,j] \mathbf{x}_j^{(s)} - t_j^T \sum_{l=1}^{n_{mas}} \mathbf{M}_M[l,l] \mathbf{x}_l^{(m)}, \]  
(2.25)

and

\[ \bar{\beta}_j = \int_{\gamma_{ij}} \phi_j d\gamma \beta_j. \]  
(2.26)

It is important to note that the definition of the weighted tangential relative velocity here is slightly different from that of [Gitterle et al., 2010], who used the time derivatives of \( \mathbf{D}_S \) and \( \mathbf{M}_M \) to guaranty frame indifference also for large rotations during a given time step. In this study this effect is negligible because the displacements during the time steps must be maintained small in order to accurately model the evolution of the frictional stress with slip or slip rate. Moreover, we do not aim in this study to address problems with very large rotation.
2.5 Time discretization

The force equilibrium equation is discretized in time with the Hilber-Hughes-Taylor (HHT) scheme [Hilber et al., 1977] as following

\[ r = M\ddot{d}^{t+\Delta t} + f_{int}(d^{t+\alpha}) + f_c(d^{t+\alpha}, z^{t+\alpha}) - f_{ext}^{t+\alpha}, \]  
(2.27.a)

\[ d^{t+\alpha} = (1 - \alpha)d^t + \alpha d^{t+1}, \]  
(2.27.b)

\[ \ddot{d}^{t+\alpha} = (1 - \alpha)d^t + \alpha \ddot{d}^{t+1} = (1 - \alpha)d^t + \alpha \left( \frac{d^{t+1} - d^{t}}{\Delta t} - \frac{1 - \beta/\gamma}{\beta/\gamma} d^t \right) \]  
(2.27.c)

\[ \ddot{d}^{t+1} = \frac{d^{t+1} - d^{t}}{\gamma \Delta t} - \frac{1 - \gamma}{\gamma} \ddot{d}^t = \frac{d^{t+1} - d^{t}}{\beta \Delta t^2} - \frac{1}{\beta \Delta t} d^t - \left( \frac{1}{\gamma} - 1 \right) \ddot{d}^t. \]  
(2.27.d)

In general, a term with the superscript \( t + \alpha \) is discretized as in (2.27.b), while a term with the superscript \( t + \alpha \) is the actual value calculated at time \( t + \alpha \). Substituting in eq. (2.27.d) in eq. (2.27.a) to eliminate the accelerations gives

\[ r = \frac{1}{\beta \Delta t^2} M\ddot{d}^{t+\Delta t} + f_{int}(d^{t+\alpha}) + f_c(d^{t+\alpha}, z^{t+\alpha}) - f_{ext}^{t+\alpha} - R, \]  
(2.28)

where

\[ R = M \left[ \frac{1}{\beta \Delta t^2} d^t + \frac{1}{\beta \Delta t} d^t + \left( \frac{1}{\gamma} - 1 \right) \ddot{d}^t \right]. \]  
(2.29)

Note that the scheme reduces to the family of Newmark integration schemes [Newmark, 1959] for \( \alpha = 1 \) and to the midpoint rule for \( \alpha = 1/2, \beta = 1/2, \) and \( \gamma = 1 \). Focusing on the contact, we consider here only linear elastic materials, thus the computation of \( f_{int}(d^{t+\alpha}) \) is straightforward; an extension to other elastic materials is provided in [Gonzalez, 2000; Laursen and Meng, 2001]. The contact force time discretization is approximated as \( f_c(d^{t+\alpha}, z^{t+\alpha}) = [0, -M_c^{t+\alpha}, D_s^{t+\alpha}]z^{t+\alpha}. \)

Following [Laursen and Chawla, 1997], to conserve energy in dynamic simulations, the persistency condition is added to the set of conditions in the normal direction to give

\[ \dot{g}_{nj}^{t+\alpha} \quad \Rightarrow \quad z_{nj}^{t+\alpha} = 0, \quad \gamma_n^{t+\alpha} \quad \Rightarrow \quad \dot{\gamma}_{nj}^{t+\alpha} = 0, \]  
(2.30)

\[ g_{nj}^{t} \leq 0 \quad \Rightarrow \quad \dot{\gamma}_{nj}^{t+\alpha} \geq 0, \quad z_{nj}^{t+\alpha} \geq 0, \quad z_{nj}^{t+\alpha} \dot{\gamma}_{nj}^{t+\alpha} = 0, \]  
with the gap rate in the normal direction defined as
\[
\begin{align*}
\dot{g}_{n_j}^{t+\alpha} &= -(n_j^{t+\alpha})^T D_s^{t+\alpha}[j,j] \dot{d}_j^{t+\alpha} + (n_j^{t+\alpha})^T \sum_{l=1}^{n_{mas}} M_M^{t+\alpha}[j,l] \dot{d}_l^{t+\alpha}.
\end{align*}
\] (2.31)

This set of conditions ensures that the expression \( z_{n_j}^{t+\alpha} \dot{g}_{n_j}^{t+\alpha} = 0 \) holds also at time steps when nodes come into contact or are released, thus the contact energy in the normal direction is always zero. However, it is important to note that this formulation of the constraints may result in small penetrations, especially for relatively large time steps.

In the tangential direction, the frictional conditions are enforced at time \( t + \alpha \) as

\[
\begin{align*}
\psi_j^{t+\alpha} := |z_j^{t+\alpha}| - \mu_j^{t+\alpha} |z_j^{t+\alpha}| \leq 0, \\
\beta_j^{t+\alpha} - \beta_j^{t+\alpha} z_j^{t+\alpha} = 0, \quad \beta_j^{t+\alpha} \geq 0, \quad \psi_j^{t+\alpha} \beta_j^{t+\alpha} = 0,
\end{align*}
\] (2.32)

where

\[
\beta_j^{t+\alpha} = (t_j^{t+\alpha})^T D_s^{t+\alpha}[j,j] \dot{d}_j^{t+\alpha} - (t_j^{t+\alpha})^T \sum_{l=1}^{n_{mas}} M_M^{t+\alpha}[j,l] \dot{d}_l^{t+\alpha}.
\] (2.33)

We note that \( \beta_j^{t+\alpha} \) and \( \mu_j^{t+\alpha} \) in the case of RS friction involve the calculation of nodal velocities also in a quasi-static formulation. In this case one would omit the acceleration term in (2.27.a) and take \( \alpha = 1 \) and \( \beta / \gamma = 1 \), to obtain a backward Euler scheme.

### 2.6 Solution with a semi-smooth Newton method

Aiming to obtain a Newton-type algorithm for the solution of the discretized system of nonlinear algebraic equations in eqn. (2.28), (2.30), and (2.32), the concept of dual Lagrange multipliers is combined with the primal-dual active set strategy and the contact conditions in eqn. (2.30) and (2.32) are replaced by equivalent nonlinear semi-smooth complementarity (NCP) functions equations. These functions reformulate these conditions as equality conditions that enable the treatment of all sources of nonlinearities in a single iterative scheme, including the categorization of all potential contact nodes into not in contact, sticking, and slipping nodes.

#### 2.6.1 Non-smooth complementarity functions

Similarly to [Hueber and Wohlmuth, 2005; Hueber et al., 2008; Popp et al., 2009; Gitterle et al., 2010], the complementarity functions for the normal and tangential
conditions are defined for each slave node \( j \in S \) as

\[
C_{nj}(z_{nj}^t + \tilde{a}, d^t + \tilde{a}) = z_{nj}^t + \tilde{a} - \max\{0, z_{nj}^t + \tilde{a} - c_n \tilde{g}_{nj}^{\text{test}}\} = 0
\]  

(2.34)

and

\[
C_{ij}(z_{ij}^t + \tilde{a}, d^t + \tilde{a}) = \max(\mu_j z_{ij}^t + \tilde{a} - c_n \tilde{g}_{ij}^{\text{test}}, |z_{ij}^t + \tilde{a}|)z_{ij}^t + \tilde{a} - \mu_j z_{ij}^t + \tilde{a} - c_n \tilde{g}_{ij}^{\text{test}} \leq 0, \quad \mu_j > 0,
\]  

(2.35)

respectively, where \( \tilde{g}_{ij}^{\text{test}} \) in eq. (2.34) is defined as

\[
\tilde{g}_{ij}^{\text{test}} := \begin{cases} 
\frac{z_{nj}^t}{c} + \tilde{g}_{nj}^t & \text{if } \tilde{g}_{nj}^t > 0, \\
\tilde{g}_{nj}^t \Delta t & \text{if } \tilde{g}_{nj}^t \leq 0.
\end{cases}
\]

The algorithmic parameters \( c_n \) and \( c_t \) do not affect the accuracy, but to achieve a good convergence behavior, they should be on the order of \( E\Delta t/l_{\text{slave}} \), where \( E \) is the Young's modulus of the medium near the fault and \( l_{\text{slave}} \) is the average length of the slave elements. Note that in a quasi-static formulation (\( \alpha = 1 \)), the complementarity functions should involve the normal gap function rather than its rate, thus \( \tilde{g}_{ij}^{\text{test}} = \tilde{g}_{ij}^{t+1} \).

### 2.6.2 Consistent linearization within the semi-smooth Newton method

To solve \( d^{t+\Delta t}, z^{t+\tilde{a}} \), an iterative semi-smooth Newton method is applied to the nonlinear system of equations of eqn. (2.28), (2.34), and (2.35) as

\[
\Delta r^t(k d^{t+\tilde{a}}, k z^{t+\tilde{a}}) = -k r^{t+\tilde{a}},
\]  

(2.36.a)

\[
\Delta C_{nj}(k d^{t+\tilde{a}}, k z^{t+\tilde{a}}) = -k C_{nj}^{t+\tilde{a}} \quad \forall j \in S,
\]  

(2.36.b)

\[
\Delta C_{ij}(k d^{t+\tilde{a}}, k z^{t+\tilde{a}}) = -k C_{ij}^{t+\tilde{a}} \quad \forall j \in S,
\]  

(2.36.c)

with the update

\[
k^{k+1} d^{t+1} = k d^{t+1} + \Delta k d^{t+1}, \quad k^{k+1} z^{t+\tilde{a}} = k z^{t+\tilde{a}} + \Delta k z^{t+\tilde{a}},
\]  

(2.37)

where the superscript \( ^{k+1} \) stands for the current iteration and the directional derivative \( \Delta \)
The linearization of the force equilibrium in eq. (2.36.a) is given by

\[
\Delta k^t+\vec{\alpha} = \Delta \left( \frac{1}{\beta \Delta t^2} M k^t+\Delta t + f_{int}\left(k^t+\alpha, k^z_t+\vec{\alpha}\right) \right)
\]

\[
= \left( \frac{1}{\beta \Delta t^2} M + kK^{int}_t+\vec{\alpha} \right) \Delta d + \Delta k^t+\vec{\alpha} = kK^{t+\vec{\alpha}} \Delta d + \Delta k^t+\vec{\alpha} \quad (2.39)
\]

where \( kK^{t+\vec{\alpha}} \) is the tangent stiffness matrix and \( kK^{t+\vec{\alpha}} \) is an effective stiffness matrix defined as \( kK^{t+\vec{\alpha}} = \left( \frac{1}{\beta \Delta t^2} M + kK^{int}_t+\vec{\alpha} \right) \).

The linearization of the contact forces can be expressed as

\[
\Delta k^t+\vec{\alpha} = \left[ 0, \Delta \left( -kM^{t+\alpha^T} k^z_t+\vec{\alpha} \right), \Delta \left( kD_S^{t+\alpha^T} k^z_t+\vec{\alpha} \right) \right]
\]

\[
= \left[ 0, \Delta \left( -\alpha kM^{t+1^T} k^z_t+\vec{\alpha} \right), \Delta \left( \alpha kD_S^{t+1^T} k^z_t+\vec{\alpha} \right) \right]
\]

\[
+ \left[ 0, -kM^{t+\alpha^T} k^z_t+\vec{\alpha}, kD_S^{t+\alpha^T} k^z_t+\vec{\alpha} \right] \Delta k^t+\vec{\alpha}
\]

\[
= \alpha k\bar{C} \Delta k^t+\vec{\alpha} + \left[ 0, -kM^{t+\alpha^T} k^z_t+\vec{\alpha}, kD_S^{t+\alpha^T} k^z_t+\vec{\alpha} \right] \Delta k^t+\vec{\alpha}
\]

\[
= -1 \cdot \left[ 0, -kM^{t+\alpha^T} k^z_t+\vec{\alpha}, kD_S^{t+\alpha^T} k^z_t+\vec{\alpha} \right] k^z_t+\vec{\alpha} = -k^t+\vec{\alpha}
\]

where the matrix \( \bar{C} \in \mathbb{R}^{(2n_{sl}+2n_{mas}) \times (2n_{sl}+2n_{mas})} \) includes the directional derivatives of mortar matrices \( M_M \) and \( D_S \) multiplied by the current Lagrange multiplier values \( k^z_t+\vec{\alpha} \) and \( \Delta k^t+\vec{\alpha} \) are the corresponding incremental displacements of slave (S) and master (M) nodes. The directional derivatives of mortar matrices \( M_M \) and \( D_S \) are given in [Popp et al., 2009].

We linearize eqn. (2.36.b) and (2.36.c) similarly to [Popp et al., 2009; Gitterle et al., 2010], but account for a variable friction coefficient and dynamic time discretization. In the normal direction, eqn. (2.34) and (2.36.b) yield the separation of the slave nodes into an inactive node set \( kI \) and an active node set \( kA \) as

\[
kI := \{ j \in | \Delta k^z_{n,j} - c_n k^t+\vec{\alpha}_j \leq 0 \}
\]

\[
(2.41)
\]
\[ kA := \{ j \in kI | kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\text{test}} > 0 \}, \]

which leads to [Hueber et al., 2008; Popp et al., 2009]

\[
\begin{align*}
\{ kZ_j^{t+\bar{\alpha}} &= 0 \quad \forall j \in kI \\
\Delta k\bar{g}_{nj}^{t+\bar{\alpha}} &= - k\bar{g}_{nj}^{t+\bar{\alpha}} \quad \forall j \in kA.
\end{align*}
\]  

(2.42.a)  

(2.42.b)

Here, the directional derivative of the gap function rate is given by

\[
\Delta k\bar{g}^{t+\bar{\alpha}}_{nj} = - k\eta j^{t+\alpha}T \left( kD_s^{t+\alpha}[j, j] k\bar{d}_{j}^{t+\alpha} - \sum_{l=1}^{n_{\text{mas}}} kM_M^{t+\alpha}[j, l] k\bar{d}_{l}^{t+\alpha} \right) 
- \alpha \Delta k\eta j^{t+\alpha}T \left( kD_s^{t+\alpha}[j, j] k\bar{d}_{j}^{t+\alpha} - \sum_{l=1}^{n_{\text{mas}}} kM_M^{t+\alpha}[j, l] k\bar{d}_{l}^{t+\alpha} \right)
- k\eta j^{t+\alpha}T \left( \alpha \Delta kD_s^{t+\alpha}[j, j] k\bar{d}_{j}^{t+\alpha} - \sum_{l=1}^{n_{\text{mas}}} \alpha \Delta kM_M^{t+\alpha}[j, l] k\bar{d}_{l}^{t+\alpha} \right),
\]  

(2.43)

where \( \Delta k\bar{d}^{t+\alpha} = \frac{\Delta y}{\Delta t} \Delta k\bar{d}^{t+1} \) and the directional derivative of the unit normal vector \( k\eta j^{t+1} \) is given in [Popp et al., 2009].

In the tangential direction, the directional derivative of eq. (2.35) also splits the slave nodes into inactive and active node sets defined in eq. (2.41). For the inactive node set, eq. (2.36.c) also becomes

\[ kZ_j^{t+\bar{\alpha}} = 0 \quad \forall j \in kI. \]  

(2.44)

Similar to [Hueber et al., 2008; Gitterle et al., 2010], the active node set branches into a stick node set \( kSt \) and a slip node set \( kSl \) as

\[
kSt := \{ j \in kA | ( kZ_{nj}^{t+\bar{\alpha}} + c_n k\bar{g}_{nj}^{t+\bar{\alpha}} ) - k\eta j^{t+\bar{\alpha}} ( kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\text{test}} ) < 0 \}
kSl := \{ j \in kA | ( kZ_{nj}^{t+\bar{\alpha}} + c_n k\bar{g}_{nj}^{t+\bar{\alpha}} ) - k\eta j^{t+\bar{\alpha}} ( kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\text{test}} ) \geq 0 \}.
\]  

(2.45)

The directional derivatives of eq. (2.36.c) with a variable friction coefficient for the sticking and slipping nodes becomes

\[
\begin{align*}
\Delta k\eta j^{t+\bar{\alpha}}_{St} &= - k\eta j^{t+\bar{\alpha}} ( kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\bar{\alpha}} ) c_t \Delta k\bar{v}_{nj}^{t+\bar{\alpha}} \\
- k\eta j^{t+\bar{\alpha}} ( \Delta kZ_{nj}^{t+\bar{\alpha}} - c_n \Delta k\bar{g}_{nj}^{t+\bar{\alpha}} ) c_t k\bar{v}_{nj}^{t+\bar{\alpha}} \\
- \Delta k\eta j^{t+\bar{\alpha}} ( kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\bar{\alpha}} ) c_t k\bar{v}_{nj}^{t+\bar{\alpha}} \\
= k\eta j^{t+\bar{\alpha}} ( kZ_{nj}^{t+\bar{\alpha}} - c_n k\bar{g}_{nj}^{t+\bar{\alpha}} ) c_t k\bar{v}_{nj}^{t+\bar{\alpha}} = - k\eta j^{t+\bar{\alpha}}_{St}, \quad j \in kSt
\end{align*}
\]  

(2.46)
and

\[
\Delta k C^t_{ij,Sl} = \left| k z^t_{ij} + c_t k v^t_{ij} \right| \Delta k z^t_{ij} + c_t \Delta k v^t_{ij},
\]

\[
= -k_{ij}^t (k z^t_{ij} - c_n k g^t_{nj}) \left( \Delta k z^t_{ij} + c_t \Delta k v^t_{ij} \right) -k_{ij}^t (k z^t_{ij} - c_n k g^t_{nj}) \left( k z^t_{ij} + c_t k v^t_{ij} \right)
\]

(2.47)

respectively, where the directional derivative of the weighted tangential relative velocity 
\( k v^t_{ij} \) is similar to eq. (2.43) but with \( k t^t_{ij} \) replacing \( k n^t_{ij} \) and \( \Delta k t^t_{ij} = \mathbf{e}_3 \times \Delta k n^t_{ij} \). The directional derivatives of the normal and tangential components of the 
Lagrange multiplier are given by

\[
\Delta k z^t_{nj} = \Delta k n^t_{j+} \cdot k z^t_{j+} + k n^t_{j+} \cdot \Delta k z^t_{j+},
\]

(2.48.a)

\[
\Delta k z^t_{ij} = \Delta k t^t_{j+} \cdot k z^t_{j+} + k t^t_{j+} \cdot \Delta k z^t_{j+},
\]

(2.48.b)

where \( \Delta k z^t_{j+} = k z^t_{j+} - k z^t_{j+} \).

Both (2.46) and (2.47) involve the coefficient of friction and its directional derivative. In 
the following sections we derive the numerical approximation of these quantities for SW 
and RS friction laws.

2.6.2.1 Slip-weakening friction discretization

The discrete form of the SW friction law given in eq. (2.6) for a node \( j \) on the slave 
surface is given by

\[
k_{ij}^t \Delta z^t_{ij}, = \mu_s \Delta u_{rel,ij}^t, \quad k_{ij}^t \Delta v^t_{ij} \leq d_c
\]

(2.49)

with the total relative slip of node \( j \) on the slave surface \( k u^t_{rel,ij} \) defined as
\[
k_{t+r}^{t+\alpha} = u_{rel,j}^{t+\alpha} + k_{t}^{t+\alpha} \cdot \left[ (k_{d_i}^{t+\alpha} - d_i^j) - (k_{d_i}^{t+\alpha} - d_i^j) \right]
\]
\[
= u_{rel,j}^{t+\alpha} + k_{t}^{t+\alpha} \cdot \left[ (k_{d_i}^{t+\alpha} - d_i^j) - \sum_{l=1}^{n_{\text{mas}}} k_{N_l}^{t+\alpha} (k_{d_i}^{t+\alpha} - d_i^j) \right]
\]
\[
= u_{rel,j}^{t+\alpha} + k_{t}^{t+\alpha} \left[ (k_{d_i}^{t+\alpha} - d_i^j) - \sum_{l=1}^{n_{\text{mas}}} k_{N_l}^{t+\alpha} (k_{d_i}^{t+\alpha} - d_i^j) \right].
\] (2.50)

where \( d_i^j \) is the displacement of node \( \mathbf{x}_i \) corresponding to the projection of the slave node on the master surface (see Figure 2-1), \( d_i \) is the displacement of the nodes associated with a surface element that includes \( \mathbf{x}_i \), \( N_1 = 0.5 \left( 1 - \xi (\mathbf{x}_i) \right) \) and \( N_2 = 0.5 \left( 1 + \xi (\mathbf{x}_i) \right) \) are the corresponding shape functions, and \( N \) is a matrix defined as
\[
N[j,l] = \left( N_{1,j} + N_{2,j} \right).I.
\]

The directional derivative of (49) is given by
\[
\Delta \left( k_{d_i}^{t+\alpha} \right) = \frac{\mu_d - \mu_s}{d_c} \Delta \left( k_{u_{rel,j}}^{t+\alpha} \right), \quad u_{rel,j}^{t+\alpha} \leq d_c,
\]
\[
0, \quad u_{rel,j}^{t+\alpha} > d_c.
\] (2.51)

where
\[
\Delta \left( k_{u_{rel,j}}^{t+\alpha} \right) = k_{t}^{t+\alpha} \cdot \left[ \alpha \Delta k_{d_i}^{t+1} - \sum_{l=1}^{n_{\text{mas}}} k_{N_l}^{t+\alpha} [j,l] \alpha \Delta k_{d_i}^{t+1} \right]
\]
\[
+ \alpha \Delta k_{t}^{t+1} \cdot \left[ (k_{d_i}^{t+\alpha} - d_i^j) - \sum_{l=1}^{n_{\text{mas}}} k_{N_l}^{t+\alpha} [j,l] (k_{d_i}^{t+\alpha} - d_i^j) \right]
\]
\[
+ k_{t}^{t+\alpha} \left[ - \sum_{l=1}^{n_{\text{mas}}} \alpha \Delta k_{N_l}^{t+1} [j,l] (k_{d_i}^{t+\alpha} - d_i^j) \right].
\] (2.52)

The directional derivative \( \Delta N_{l} \) involves the directional derivative of \( \xi (\mathbf{x}_i) \), which is given in [Popp et al., 2009].

2.6.2.2 Rate and state friction discretization

The discrete form of the RS friction law (eqn. (2.7) and (2.8)) for node \( j \) on the slave surface is given by
\[ k_{\mu j}^{t+\alpha} = \mu^* + a \ln \left( \frac{k_{\mu_{rel,j}^{t+\alpha}} + v_{th}}{v^*} \right) + b \ln \left( \frac{k_{\theta_j^{t+\alpha}}}{\theta^*} \right). \]  

(2.53)

with the state variable evolving as

\[ k_{\theta_j^{t+\alpha}} = 1 - \frac{k_{\theta_j^{t+\alpha}}(k_{\mu_{rel,j}^{t+\alpha}} + v_{th})}{L}. \]  

(2.54)

and the slip rate defined as

\[ k_{\mu_{rel,j}^{t+\alpha}} = k_{\mu j}^{t+\alpha} \cdot \left[ k_{\theta_j^{t+\alpha}} - \sum_{i=1}^{n_{\text{mas}}} k_{N}^{t+\alpha}[j, l] k_{d_i^{t+\alpha}} \right]. \]  

(2.55)

The threshold velocity term, \( v_{th} \), is added to avoid singularity at slip rate of \( \dot{u}_{rel,j}^{t+\alpha} = 0 \).

The directional derivative of eq. (2.53) is simply

\[ \Delta \mu^{t+\alpha} = \frac{a}{k_{\mu_{rel,j}^{t+\alpha}} + v_{th}} \Delta k_{\mu_{rel,j}^{t+\alpha}} + \frac{b}{k_{\theta_j^{t+\alpha}}} \Delta k_{\theta_j^{t+\alpha}}, \]  

(2.56)

with the directional derivative of the slip rate given by

\[ \Delta k_{\mu_{rel,j}^{t+\alpha}} = k_{\mu j}^{t+\alpha} \cdot \left[ \alpha \Delta k_{d_i^{t+1}} - \sum_{i=1}^{n_{\text{mas}}} k_{N}^{t+\alpha}[j, l] k_{d_i^{t+\alpha}} \right] \]

\[ + \alpha \Delta k_{t_j^{t+1}} \cdot \left[ k_{\theta_j^{t+\alpha}} - \sum_{i=1}^{n_{\text{mas}}} k_{N}^{t+\alpha}[j, l] k_{d_i^{t+\alpha}} \right] \]

\[ + k_{\mu j}^{t+\alpha} \cdot \left[ - \sum_{i=1}^{n_{\text{mas}}} \alpha \Delta k_{N}^{t+1}[j, l] k_{d_i^{t+\alpha}} \right]. \]  

(2.57)

Similarly to [Liu and Borja, 2009], to avoid an additional set of variables, we aim to express \( k_{\theta_j^{t+\alpha}} \) and \( \Delta k_{\theta_j^{t+\alpha}} \) as a function of the slip rate. We discretize the state variable in time similar to the nodal displacement and velocity time discretization in eq. (2.27):

\[ k_{\theta_j^{t+\alpha}} = (1 - \alpha) \theta_j^t + \alpha \theta_j^{t+1}, \]  

(2.58.a)

\[ k_{\theta_j^{t+\alpha}} = \frac{\alpha}{\beta/\gamma} \theta_j^{t+1} - \frac{\beta/\gamma - \alpha}{\beta/\gamma} \theta_j^t. \]  

(2.58.b)

Equating eq. (2.58.b) with eq. (2.54) together with some algebra gives

\[ k_{\theta_j^{t+\alpha}} = \left[ 1 + \frac{\theta_j^t}{(\beta/\gamma)\Delta t} - \frac{\beta/\gamma - \alpha}{\beta/\gamma} \right] \frac{(\beta/\gamma)L\Delta t}{L + (\beta/\gamma)\Delta t \left( k_{\mu_{rel,j}^{t+\alpha}} + v_{th} \right)} \]  

(2.59)

and
\[ \Delta k \theta_{j+a} = - \left[ 1 + \frac{1}{\beta / \gamma \Delta t} \right] \left( \frac{\beta / \gamma - \alpha}{\beta / \gamma} \varphi_j \right) \frac{L(\Delta t)^2}{(L + (\beta / \gamma) \Delta t (\frac{k_i}{u_{rel,j}} + u_{th}))} \Delta k \varphi_{rel,j}. \] (2.60)

### 2.6.3 Algebraic representation

Finally, the global algebraic representation of eq. (2.36) to be solved in each iteration is derived. The matrix and vector blocks of this linear system are defined by the five sets \( N \), \( M \), \( I \), \( St \) and \( Sl \). We drop the iteration and time indices here for ease of notation.

\[
\begin{bmatrix}
K_{NN} & K_{NM} & K_{NI} & K_{NSI} & 0 & 0 & 0 \\
K_{MN} & K_{MM} & K_{MI} & K_{MIS} & -M_{MI} & -M_{MST} & -M_{MST} \\
K_{IN} & K_{IM} & K_{II} & K_{ISI} & D_{SI} & 0 & 0 \\
K_{STN} & K_{STM} & K_{ST} & K_{STSI} & 0 & D_{SSI} & 0 \\
K_{SIN} & K_{SIM} & K_{SI} & K_{SISI} & 0 & 0 & D_{SSI} \\
0 & 0 & 0 & 0 & I_l & 0 & 0 \\
0 & S_{AM} & S_{Al} & S_{AS} & S_{Al} & 0 & 0 \\
0 & F_{STM} & F_{St} & F_{STSI} & F_{St} & 0 & P_{St} \\
0 & G_{SIM} & G_{SII} & G_{SISI} & G_{SISI} & 0 & 0 & L_{SI} \\
\end{bmatrix}
\begin{bmatrix}
\Delta d_N \\
\Delta d_M \\
\Delta d_I \\
\Delta d_{St} \\
\Delta d_{St} \\
\Delta d_{St} \\
\Delta d_{St} \\
0 \\
0 \\
\end{bmatrix}
= - \begin{bmatrix}
\Gamma_N \\
\Gamma_M \\
\Gamma_I \\
\Gamma_{St} \\
\Gamma_{Sl} \\
\Phi_{nA} \\
\Phi_{t,St} \\
\end{bmatrix}
\] (2.61)

The first five rows can be identified as the linearized algebraic form of the force equilibrium equation in eq. (2.36a), where \( \bar{R} = K + \bar{C} \). The sixth row represents the contact constraint condition for nodes of the inactive set \( I \). In the seventh row, matrix \( S_{A} \in \mathbb{R}^{n_a \times (2n_{mas}+2n_{sl})} \) is the assembly of all linearizations of \( g_{nA} \) (eqn. (2.36b) and (2.42.b)), where \( n_a \) is the number of active slave nodes. In the eighth row, \( F \in \mathbb{R}^{n_{stick} \times (2n_{mas}+2n_{sl})} \) is the assembly of all linearizations of \( C_{t,St} \) with respect to displacements and \( P_{St} \in \mathbb{R}^{n_{stick} \times 2n_{stick}} \) is the assembly of all linearizations with respect to the Lagrange multipliers (eqn. (2.36c) and (2.46)). In the ninth row, \( G \in \mathbb{R}^{n_{slip} \times (2n_{mas}+2n_{sl})} \) is the assembly of all linearizations of \( C_{t,Sl} \) with respect to displacements and \( L_{Sl} \in \mathbb{R}^{n_{slip} \times 2n_{slip}} \) is the assembly of all linearizations with respect to the Lagrange multipliers (eqn. (2.36c) and (2.47)).

This system contains both displacement and Lagrange multiplier degrees of freedom. For efficient solution of the system, the Lagrange multipliers are condensed out in the
following two stages. First, because the Lagrange multipliers of the inactive nodes are zero, the sixth row and column are eliminated. Second, the diagonality of $D_5$ enables expressing the Lagrange multipliers of the stick and slip nodes as

$$z_{St} = D_{St}^{-1}(r_{St} - K_{StN} \Delta d_N - K_{SIM} \Delta d_M - K_{StI} \Delta d_I - K_{StS} \Delta d_{St} - K_{StS} \Delta d_{Sl})$$

$$z_{Sl} = D_{Sl}^{-1}(r_{Sl} - K_{SlN} \Delta d_N - K_{SlM} \Delta d_M - K_{SlI} \Delta d_I - K_{SlS} \Delta d_{Sl} - K_{SlS} \Delta d_{St})$$

(2.62)

Substituting into eq. (2.61), a reduced system with only displacement degrees of freedom is obtained as

$$\begin{align*}
\begin{bmatrix}
K_{NN} & K_{NM} & K_{NI} & K_{NST} & K_{NSI} \\
K_{MN} & \tilde{R}_{MM} & R_{MI} & \tilde{R}_{MST} & \tilde{R}_{MTS} \\
+M_{MST}^T \tilde{K}_{STN} & +M_{MST}^T \tilde{K}_{STI} & R_{II} & \tilde{R}_{IST} & \tilde{R}_{IST} \\
+M_{MTS}^T \tilde{K}_{STN} & +M_{MTS}^T \tilde{K}_{STI} & \tilde{R}_{IST} & R_{II} & \tilde{R}_{IST} \\
K_{IN} & \tilde{R}_{IM} & \tilde{R}_{ISt} & \tilde{R}_{ISt} & \tilde{R}_{ISt} \\
0 & 0 & 0 & 0 & 0 \\
P_{St} D_{St}^{-1} K_{StN} & P_{St} D_{St}^{-1} K_{StI} & -F_{St} & 0 & 0 \\
L_{St} D_{St}^{-1} K_{StI} & L_{St} D_{St}^{-1} K_{StI} & 0 & -G_{St} & 0
\end{bmatrix}
\begin{bmatrix}
\Delta d_N \\
\Delta d_M \\
\Delta d_I \\
\Delta d_{St} \\
\Delta d_{Sl}
\end{bmatrix}
= \begin{bmatrix}
r_N \\
r_I \\
r_{St} \\
r_{Sl}
\end{bmatrix}
\begin{bmatrix}
\tilde{M}_{MST}^T r_{St} + \tilde{M}_{MTS}^T r_{Sl} \\
P_{St} D_{St}^{-1} r_{St} - C_{t,St} \\
L_{St} D_{St}^{-1} r_{St} - C_{t,Sl}
\end{bmatrix}
\end{align*}
$$

(2.63)

where $\tilde{M} = D^{-1}M$.

### 2.6.4 Primal–dual active set algorithm

1. **Initialize** $d^0, \bar{d}^0 \Rightarrow \bar{d}^0$ and $z^0$.
2. **Loop over all time steps**
   - For a given time step at $t + \alpha$
     1. Set $k = 0$ and initialize $0d = d^t, \ 0z^{t+\alpha} = z^{(t-1)+\alpha}, \ 0l = l^t, \ 0s_t = St^t, \ 0s_l = Sl^t, \ 0a = St \cup Sl$
     2. Find $\Delta^k d$ and $\Delta^k z^{t+\alpha}$ by solving
\[
\Delta^{k+1}r^{t+\alpha} = -k^{t+\alpha} \tag{2.64.a}
\]
\[
k^{+1}z_{j}^{t+\alpha} = 0, \quad j \in kI \tag{2.64.b}
\]
\[
\Delta^{k}g_{j}^{t+\alpha} = -k^{t+\alpha}, \quad j \in kA \tag{2.64.c}
\]
\[
\Delta^{k}c_{t,tk}^{t+\alpha} = -k^{t+\alpha}, \quad j \in kSt \tag{2.64.d}
\]
\[
\Delta^{k}c_{t,sk}^{t+\alpha} = -k^{t+\alpha}, \quad j \in kSl \tag{2.64.e}
\]

2. Update \(k^{+1}d^{t+1} = k^{t}d^{t+1} + \Delta^{k}d\)

3. Update \(k^{+1}I, k^{+1}A, k^{+1}St,\) and \(k^{+1}Sl\) as
   \[
   k^{+1}I := \{j \in [^{k+1}z_{n}^{t+\alpha} - c_{n}^{k+1}g_{j}^{test} \leq 0]\}
   \]
   \[
   k^{+1}A := \{j \in [^{k+1}z_{n}^{t+\alpha} - c_{n}^{k+1}g_{j}^{test} > 0]\}
   \]
   \[
   k^{+1}St := \{j \in k^{+1}A\big|$\left[(^{k+1}z_{n}^{t+\alpha} + c_{t}^{k+1}v_{tj}^{t+\alpha}) -^{k+1}\mu_{j}^{t+\alpha}(^{k+1}z_{n}^{t+\alpha} - c_{n}^{k+1}g_{j}^{test}) \leq 0\right]$\big]\}
   \[
   k^{+1}Sl := \{j \in k^{+1}A\big|$\left[(^{k+1}z_{n}^{t+\alpha} + c_{t}^{k+1}v_{tj}^{t+\alpha}) -^{k+1}\mu_{j}^{t+\alpha}(^{k+1}z_{n}^{t+\alpha} - c_{n}^{k+1}g_{j}^{test}) \geq 0\right]$\big]\}
\[
(2.65)
\]

4. If \(k^{+1}I = kI, k^{+1}St = kSt, k^{+1}Sl = kSl\) and the convergence criterion is satisfied continue to stage 5, else set \(k = k + 1\) and go to stage 1

5. Update \(d^{t+\Delta t}\) and \(\ddot{d}^{t+\Delta t}\) with the time discretization scheme

2.7 Examples

In this section, several numerical examples are provided in order to demonstrate the capabilities of the method, examine the accuracy of implementation of the highly nonlinear RS friction law, and to evaluate the convergence and energy preservation properties of the method for different time discretization schemes.

2.7.1 Quasi-static benchmark for rate and state friction

Although this paper mostly focuses on the dynamic response of a contact with variable friction, to verify the implementation of RS friction into the Mortar method, we begin with the following quasi-static numerical test. A 5 x 10 cm rectangular body with a fault at an orientation of 45° and non-matching grid is subjected to the boundary conditions shown in Figure 2-3a. The fault is governed by RS friction with \(\mu^{*} = 0.6, \nu^{*} = 1 \times 10^{-6}\) m/s, \(\nu_{th} = 10^{-9}\), \(a = 0.01\), \(b = 0.012\), and \(L = 50 \mu m\). At the beginning the fault is
locked, then with increasing slip of the upper edge, and consequently shear stress on the fault, the upper block begins to slide relative to the lower block and shortly approaches a steady state constant sliding velocity of $v = 1 \times 10^{-6}$ m/s. At this stage we increase the velocity of the upper edge by a factor of 10 and examine the response of the friction coefficient on the central node of the upper side of the fault (Figure 2-3b). Because the relative slip rate on the fault does not increase immediately by a factor of 10 as on the boundary, the peak value of the numerical friction coefficient is slightly lower than that of the analytic, but after two time steps the numerical solution converges to the analytic solution quite well.

2.7.2 Dynamic rupture with slip-weakening friction

In this section, we demonstrate the effectiveness of the method in studying dynamic shear rupture problems on rough faults governed by SW friction. We consider the problem of a 20 m self-affine rough fault embedded in a 60 x 30 m elastic domain subjected to a simple shear loading conditions. The mechanical properties of the domain and the loading conditions are shown in Figure 2-4a. This setup results in a gradual increase of the shear load on the fault, thus the nucleation of the rupture is completely spontaneous. The geometry of the fault is shown in Figure 2-4b and the fault is governed by the following SW parameters: $\mu_s = 0.6$, $\mu_d = 0.55$, and $D_c = 0.3$ mm. In order to represent properly the chosen minimum wavelength of roughness (20 cm) with the mesh, we use hanging nodes to gradually refine the quadrilateral element from a size of about 1 x 1 m near the boundaries of the model to about 1.56 x 1.56 cm around the fault. This leads to 1281 nodes on each side of the fault. We perform four simulations with different values of the time scheme parameters in (27) as follows: (1) Midpoint scheme ($\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 1$); (2) average acceleration Newmark scheme ($\alpha = 1$, $\beta = 0.25$, and $\gamma = 0.5$); (3) Newmark scheme with a small damping ($\alpha = 1$, $\beta = 0.3$, and $\gamma = 0.6$); and (4) Newmark scheme with a larger damping ($\alpha = 1$, $\beta = 0.5$, and $\gamma = 1$). In order to load the fault, all simulations begin with several large quasi-static time steps, then the time step size is reduced to $\Delta t = 20$ $\mu$s and the simulations continue dynamically.

Figure 2-5 shows snapshots of the distribution of shear stress around the fault at four different stages of the rupture for a simulation with time discretization scheme #3.
The first stage corresponds to the end of the quasi-static loading stage. At this stage, a few small regions with preferable local orientation of the fault begin to slip with a decrease in the friction coefficient and shear stress. At the second stage, a 2m long rupture nucleates from one of these regions with further decreases in shear stress and development of stress concentrations at its tips. At the third stage, the rupture expands at velocities of about 2200 m/s and 2400 m/s to the left and right sides, respectively. Finally, the entire fault ruptures and slips, and large shear stresses are observed at the tips, as well as significant variations in shear stress with the local geometry of the fault.

To study the effects of the time discretization parameters, we examine the evolution of sliding velocity and shear traction with time in all simulations at the two nodes on the fault shown in Figure 2-5. The locations are chosen to represent both the nucleation (node A) and propagation (node B) phases of the rupture. As expected, the latter shows a narrower velocity curve with a larger peak (Figures 2-6a and 2-6b), which slightly decreases with increasing damping in the time discretization schemes. In both locations, the general behavior of the velocity curve is similar for all of the time discretization schemes we tested, where the differences between schemes #1 and #2 are negligible and the other schemes damp mostly the high frequency content of the curve. The shear traction at node A decreases from its initial value to its residual value over about 3 ms (Figure 2-6c), while, at node B, the shear traction initially increases over about 0.5 ms and then decreases sharply to a residual value (Figure 2-6d). In both locations, there is a further moderate decrease in shear traction, followed by larger variations resulting from the arrest of the rupture at the tips of the fault. Differences among the time schemes are observed mostly at this last stage.

Next, we examine the effects of the time discretization schemes on the energy components in the system for the problem described above. A scheme conserves energy if at a given time step

\[ E_{\text{pot}}^{t+1} + E_{\text{kin}}^{t+1} - (E_{\text{ext}}^{t+1} - E_{\text{con}}^{t+1}) = 0, \]  

(66)

where the potential energy is given by

\[ E_{\text{pot}}^{t+1} = E_{\text{pot}}^t + 0.5 \Delta d^T (f_{\text{int}}^t + f_{\text{int}}^{t+1}), \]  

(67)

the kinetic energy is given by
the contact work is given by

\[ E_{ct}^{t+1} = \begin{cases} E_c^t + 0.5d^T(d_c^t + d_{c+1}^t), & \alpha = 1 \\ E_c^t + \Delta d^T(f_c^t + f_{c+1}^t), & \alpha = 1/2 \end{cases} \]  

(69)

and the external work is given by

\[ E_{ext}^{t+1} = E_{ext}^t + 0.5\Delta d^T(f_{ext}^t + f_{ext}^{t+1}). \]  

(70)

As expected, schemes #3 and #4, which involve algorithmic damping, slightly dissipate energy with a smaller decrease of the potential energy and correspondingly a smaller increase of the contact work and kinetic energy with time compared to schemes #1 and #2 (Figure 2-7). However, it is important to note that this effect decreases with decreasing time step size and vice versa. While scheme #2 exactly conserves energy, a slight growth in the total energy is observed for scheme #1.

To examine the convergence behavior of the method, Figure 2-8 shows the number of iterations and number of slipping nodes during the simulation for schemes #1 and #4. Both schemes show excellent convergence despite the large number of nodes on the fault. Scheme #1 shows more changes between slipping and sticking nodes and consequently slightly more iterations. Between time step 250 and time step 800, when most of the rupture process occurs, the average number of iterations per time step of scheme #1 is 3.9, while that of scheme #4 is 3.2. Schemes #2 and #3 show similar behavior with average numbers of iterations of 3.7 and 3.25, respectively.

### 2.7.3 Dynamic rupture with rate and state friction

Using the same problem setup as in Figure 2-4a, we demonstrate the capability of the method to study physical problems that involve frictional instability on a rougher fault governed by RS friction. We use the same RS friction parameters as in section 7.1 and increase the roughness amplitude by a factor of two. In order to model the evolution of the friction coefficient accurately during the nucleation and propagation phases of the rupture, the time step varies such that the maximum relative slip in a given time step is smaller than half of the critical slip distance \( L \). This leads to a significant reduction in time step size from a value of \( \Delta t = 1000 \) s to \( \Delta t = 10 \) \( \mu \)s during the nucleation phase and additional decrease to a value of \( \Delta t = 3 \) \( \mu \)s during the propagation phase. To examine the
energy conservation property also at the stage of the rupture arrest, we do not allow the
time step size to increase back, and fix it at a value of $\Delta t = 10 \mu s$. The transition from
quasi-static to dynamic time integration is performed when the time step size decreases
below 0.001 s.

Figure 2-9a shows the time evolution of slip along the fault for a simulation with
time discretization scheme #2. The initial quasi-static stage is represented by red contours
at decreasing time intervals, while the dynamic stage is represented by black contours at
the time interval set to a value of $5 \times 10^{-4}$ s. At the end of the quasi-static stage most of the
slip occurs along a portion of the fault between 10.5 and 14 m. With the transition to the
dynamic stages we observe a complex behavior of the rupture, with asymmetric
expansion of rupture and large spatial variations of the final slip that correspond to the
local geometry of the fault. Moreover, the rupture velocity $V_r$ varies significantly with the
local geometry of the fault (Figure 2-9b).

Similar to section 7.2, we examine the effect of the time discretization schemes on
the evolution of sliding velocity and shear traction at nodes $A$ and $B$. Although the sliding
velocity is generally larger than that obtained with SW, the difference between the time
discretization schemes is smaller (Figures 2-10a and 2-10b) and is observed mostly at the
stage where the rupture decelerates. In general, the combination of larger roughness
together with RS friction results in a complex behavior of the shear traction with large
temporal variations, including an initial strengthening stage also at node $A$ and very high
traction concentrations at node $B$ (Figures 2-10c and 2-10d). In both locations, the final
stage of the simulations with no damping involves large oscillations in the shear traction.
It is important to note that these oscillations are not the result of numerical errors, but are
the result of the propagation of waves in the domain and the arrest of the rupture at the
tips of the fault.

Figure 2-11 shows the energy partitioning during the rupture process. The initial
stage of strengthening in the RS friction law leads to a larger potential energy compared
to the case of the SW friction law and consequently the kinetic energy and contact work
components are larger. During the propagation phase of the rupture the energy dissipation
in damping schemes (#3 and #4) is quite small because of the small time step size. An
increase in energy dissipation is observed during the arrest of the rupture, where the time
step size is larger. Similar to the case of the SW friction law, scheme #2 exactly conserves energy and a slight growth in the total energy is observed for scheme #1.

To study the convergence behavior of the method, Figure 2-12 shows the number of iterations and number of slipping and inactive nodes during the simulation for schemes #1 and #4. In general good convergence is observed, although the high nonlinearity of the RS friction law and the larger roughness amplitude result in slower convergence rate compared to the previous example. Moreover, small portions of the fault open near the end of the simulations and 30 nodes become inactive. The algorithmic damping improves the convergence rate, with an average number of iterations per time step of 6.4 for scheme #4 and 7.4 for scheme #1. Scheme #4 also shows a more steady convergence rate. Schemes #2 and #3 show similar behavior with an average number of iterations of 8.3 and 7, respectively. All schemes show smaller convergence rates near the end of the simulation. At this stage the slip rate along the fault is small (see Figure 2-10), but with the RS friction law small changes in the slip rate result in large variations in the friction coefficient. These variations, together with variations in the normal and shear tractions on the fault because of propagating elastic waves in the medium, result in a slower convergence rate.

2.8 Conclusions

We extend the 2D finite deformation mortar formulation to dynamic problems and implement SW and RS friction laws into the method. We utilize the dual Lagrange multipliers and the primal–dual active set strategy concepts and accordingly discretize and linearize the friction laws to obtain a semi-smooth Newton method. Moreover, the discretization of the RS friction law involves a procedure to condense out the state variables, thus eliminating the addition of another set of unknowns into the system.

Several numerical examples are provided in order to demonstrate the capabilities of the method for modeling shear rupture on rough surfaces governed by SW and RS friction laws. The effect of the different time discretization schemes on the convergence, energy conservation, and the time evolution of shear traction and slip rate is examined. The method shows excellent convergence for the SW friction law with efficient detection between the slipping and sticking states of the nodes despite the large number of nodes.
on the fault. A good convergence is also obtained for the RS friction law, but because of its high nonlinearity and because it involves significant variations of the friction coefficient with small change in slip rate, more iterations are needed before convergence. For both friction laws, the total energy is exactly conserved with the non-damping Newmark scheme and experiences very small growth with the mid-point scheme. The amount of energy dissipation in the damping schemes is quite small. It decreases with decreasing time step size and affects mostly the very high frequency variations in the shear traction and slip rate.
Figure 2-1: Notation for finite deformation contact problem.
Figure 2-2: (a) A linear SW friction law. (b) The change in friction in response to sudden increase in sliding velocity as predicted by the RS friction law.
Figure 2-3: Benchmark problem for RS friction. (a) Geometry and boundary conditions. (b) Friction coefficient versus relative slip at the central node of the upper side of the fault in response to a change in the velocity of the upper edge and consequently the slip rate on the fault. The red circles represent the numerical solution, while the black line is the analytic solution. Note that we begin with a non-matching grid on the fault.
Figure 2-4: (a) The problem set up: a 60 x 30 m elastic medium with a 20 m long fault is subjected to a prescribed slow horizontal velocity $\pm V_b$ at the top and bottom and zero vertical displacements on all boundaries, resulting in a simple shear loading conditions. (b) The geometry of the fault: self-affine fractal roughness geometry with the minimum wavelength of roughness set to 20 cm.
Figure 2-5: The distribution of shear stress around the fault at four different stages of a simulation with time discretization parameters of $\alpha = 1$, $\beta = 0.3$, and $\gamma = 0.6$. The reference for the time shown is the end the quasi-static loading stage (stage 1). The black circles show the locations where the sliding velocity and shear traction are measured in Figure 2-6.
Figure 2-6: The effect of the time discretization parameters on the time evolution of sliding velocity and shear traction at nodes A and B for a fault governed by SW friction. The nodes are located inside and outside the nucleation region, respectively (see Figure 2-5). Note that in all schemes we choose $\gamma = 0.5\beta$ and the time is calculated from the end of the quasi-static loading stage.
Figure 2-7: Potential (a), kinetic (b), contact (c), and the energy balance (d) vs. time for the four time discretization schemes. Note that in all schemes we choose $\gamma = 0.5\beta$ and the time is calculated from the end of the quasi-static loading stage.
Figure 2-8: Number of slipping nodes (black) and iterations (red) during the simulation for scheme #1 (solid) and scheme #4 (dashed).
Figure 2-9: (a) The evolution of slip along the fault with time for a simulation with time discretization scheme #2. The initial quasi-static stage is represented by red contours, with decreasing time intervals between the contours, while the dynamic stage is represented by black contours with the time interval between the contours set to a value of $5 \times 10^{-4}$ s. (b) The rupture velocity $V_r$ along the fault.
Figure 2-10: The effect of the time discretization parameters on the time evolution of sliding velocity and shear traction at nodes A and B for fault governed by RS friction. Note that in all schemes we choose $\gamma = 0.5\beta$ and the time is calculated from the end of the quasi-static stage.
Figure 2-11: Potential (a), kinetic (b), contact (c), and the energy balance (d) vs. time for the four time discretization schemes. Note that in all schemes we choose $\gamma = 0.5 \beta$ and the time is calculated from the end of the quasi-static stage.
Figure 2-12: Number of slipping (black) and inactive (blue) nodes and number of iterations (red) during the simulation for scheme #1 (solid) and scheme #4 (dashed).
Chapter 3

The slip behavior and source parameters of spontaneous slip events on rough faults subjected to slow tectonic loading

Abstract

We study the response to slow tectonic loading of rough faults governed by velocity weakening rate and state friction, using a 2-D plane strain model. Our numerical approach accounts for all stages in the seismic cycle, and in each simulation we model a sequence of two earthquakes or more. We focus on the global behavior of the faults and find that as the roughness amplitude increases and the minimum wavelength of roughness decreases, there is a transition from seismic slip to aseismic slip, in which the load on the fault is released by more slip events but with lower slip rate, lower seismic moment per unit length $M_{0,1d}$, and lower average static stress drop on the fault $\Delta \tau_s$. Even larger decreases with roughness are observed when these source parameters are estimated only for the dynamic stage of the rupture. The source parameters $M_{0,1d}$ and $\Delta \tau_s$ decrease mutually such that the linear relationship between $\Delta \tau_s$ and the average fault strain $\bar{D}/L_{rup}$ is preserved, except for fast slip events on faults with large amplitudes of roughness, where the decrease in slip is larger than that of $\Delta \tau_s$. A consistent effect of the length of the fault on $\Delta \tau_s$ is observed only for smooth faults, thus we speculate that roughness may be one of the reasons for the stress drop being independent of magnitude.

3.1 Introduction

The deviation of natural faults from planarity [Brown and Scholz, 1985; Power et al., 1987; Power and Tullis, 1991; Renard et al., 2006; Sagy et al., 2007; Candela et al., 2009, 2012; Bistacchi et al., 2011; Brodsky et al., 2011] results in geometric asperities
and a locally heterogeneous stress field, which affect the nucleation and propagation of shear rupture and consequently earthquake source parameters, such as the maximum slip rate, seismic moment, and static stress drop. However, the exact effect of roughness on these parameters is not yet clear.

Numerical static analyses of the response of rough faults to an external increase in shear load [Dieterich and Smith, 2009] or a prescribed uniform shear stress reduction on the fault [Zielke et al., 2017] show that the slip on the fault (as well as the area of the rupture in the latter study) decreases with increasing roughness amplitude. Using a quasi-static cascade rupture model for an earthquake, Candela et al., [2011] suggest that, the stress drop decreases with increasing fault dimension because of the self-affine fractal geometry of faults. The quasi-static analysis of Bailey and Ben-Zion [2009] shows that frictional heterogeneities on the fault, which can represent geometrical variations to some extent, result in decreases in the average stress drop on the fault. However, although these studies provide important first order information regarding the effects of roughness on the source parameters, they do not take into account the dynamics of the rupture and the effects of friction laws and may be oversimplified. Moreover, they do not provide any information regarding the maximum slip rate on the fault during the events, thus whether they happen seismically or aseismically.

Several numerical studies of dynamic rupture on rough faults governed by frictional laws have been performed, all of them with strongly rate-weakening friction laws and visco-plasticity for the bulk material [Dunham et al., 2011a; Fang and Dunham, 2013; Shi and Day, 2013; Bruhat et al., 2016]. These studies focus on the rupture process rather than the macroscopic source parameters and show that roughness promotes the development of self-healing rupture pulses, substantial fluctuations in rupture velocity, heterogeneous slip distribution, inelastic deformation, and diverse rupture styles, such as rupture arrests, secondary slip pulses that rerupture previously slipped fault sections, and supershear transitions. In the context of the size of earthquakes, Fang and Dunham, [2013] point out that most ruptures in their simulations stop naturally before reaching the ends of the computational domain, usually when they encounter unfavorable stress conditions on compressional bends.

The dynamic simulations above are limited to a single rupture and adopt a
constant time step size and an artificial nucleation procedure. The question then arises: would the modeled ruptures nucleate spontaneously into fast seismic events when the faults are only loaded by the tectonic stress. Moreover, these models assume a spatially uniform initial stress field and friction parameters (state variable), but the slip and stress heterogeneities at the end of slip events affect the rupture process at subsequent events. In addition, the contact formulation assumes that the grid points are collocated on either side of the fault during all stages of the simulation, thus the range of roughness wavelengths is limited and the variations of the normal stresses on the fault during slip may be underestimated.

The method developed in chapter 2 enables the use of a variable time step size and transitions between quasi-static and fully dynamic implicit time integration schemes to model the whole seismic cycle, including the slow aseismic nucleation, the dynamic propagation, and the arrest of the rupture. It is a mortar-based finite element formulation, thus non-matching meshes are allowed across the fault, the contacts are continuously updated, and the variation in the contact stresses is accurately modeled. It is important to note, however, that because our model currently accounts only for an elastic medium, the amount of slip relative to the minimum wavelength of roughness is also limited in this study and consequently also the number of slip events in the cycle. Otherwise, unphysical stresses will be developed in the medium surrounding the fault. Moreover, the deformation of the fault geometry and the surrounding medium during the seismic cycle are controlled by other physical process, which are currently not accounted for in the our model or in any of the existing dynamic simulators, such as wear process and brittle damage. Another limitation of the method adopted here is that the computational cost is much larger, a problem that we partially address by refining the mesh toward the fault with hanging nodes.

We study the response of rough faults governed by rate and state friction to slow tectonic loading over time equivalent to about two seismic cycles, if a smooth fault were considered, and examine the effects of roughness on the stress drop, seismic moment, and whether the fault slips seismically or aseismically. We focus on the scale of small earthquakes, and choose the minimum roughness wavelength to be at the size close to the lab samples (20 cm) and thus use the observed lab-scale rate and state friction laws
without up-scaling the constitutive parameters, assuming that the experimental data already include the effects of smaller wavelengths of roughness. We provide some of the quantities in dimensionless form, and believe that trends obtained here are also valid for larger earthquakes, but it is important to note that the complexity and heterogeneity of the medium around large faults are expected to be larger, and that the friction laws should be modified to include the large reduction in the friction coefficient observed in large slip experiments [e.g. Di Toro et al., 2011].

3.2 Model description

We study the nucleation and propagation of shear rupture on a finite fault with length, $L_f$, embedded in a 2-D elastic medium with dimensions $3L_f \times 1.5L_f$ (Figure 3-1a). We assume plane strain boundary conditions and load the medium with a prescribed slow velocity of $V_b = \pm \frac{L_f}{20} \cdot 5 \times 10^{-10}$ m/s. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. This setup enables completely spontaneous nucleation of the rupture. We do not explicitly consider pore fluids in this study and assume that these are effective stresses if the medium is saturated with fluids. The fault geometry is a self-affine fractal with Hurst exponent of $H = 0.8$ and roughness pre-factor, $b_r$, ranges between 0.001 and 0.01. For reference, we also run simulations with smooth faults. The fault is governed by rate and state friction with an aging law for the evolution of the state variable. The mechanical properties of the medium, friction law parameters, and initial stresses are given in Table 3-1.

Currently, the numerical method accounts only for elastic rheology of the medium surrounding the fault. This approximation limits the amount of deformation that the medium can experience before unrealistic stresses larger than the Coulomb failure criteria are accumulated around the fault. Therefore, we set the total time of simulation to the time that two large events occur, when a smooth fault is considered. Because we do not model a large sequence of earthquakes, the initial conditions have a significant effect on the results, despite the spontaneous nucleation. We conceptually begin the simulation at
the end of an earthquake that ruptured the whole fault and choose the initial friction parameters accordingly. We assume that the earthquake approached slip rate on the order of 1 m/s during the rupture, that the slip rate is currently zero, and that the state variable and the friction coefficient had no time to evolve and are equal to their steady state values at a slip rate of 1 m/s, which are given by \( \theta_0 = L/(v_t = 1) \) and \( \mu_0 = (a - b) \ln(v_t = 1) \), respectively. The initial shear stress is chosen such that \( \sigma_{xy0} = \mu_0 \sigma_{yy0} \). As the roughness of the fault increases, some segments along the fault may begin to slip under smaller shear stress. In these cases, the initial shear stress is smaller than \( \mu_0 \sigma_{yy0} \), but in order to maintain similar initial conditions in all simulations, we do not allow the state variable to evolve until the remote shear stress exceeds the initial shear stress calculated for a smooth fault.

We use the mortar finite element method described in chapter 2 and refine the mesh around the fault with hanging nodes in order to represent the geometry of the fault properly. The resulting elements on the fault have dimensions of about 1.56 x 1.56 cm (Figure 3-1b). We use a variable time step and based on the current values of slip rates we estimate the time step size at the next time step such that the average incremental slip of the 40 fastest nodes along the fault will not be larger than half of the size of the critical distance. This procedure results in a time step size that represents the evolution of the friction coefficient well without reducing the time step size to values that lead to simulations with an exaggerated number of time steps. The method switches between quasi-static and dynamic time integration schemes when the average slip rate at the 40 fastest nodes on the fault is larger than \( 5 \times 10^{-5} \) m/s. We use the Newmark time integration scheme with small algorithmic damping (\( \beta = 0.35 \) and \( \gamma = 0.7 \)) for the dynamic stage and also add an absorbing layer with gradual Raleigh damping near the boundaries of the model.

In order to study the effects of the geometry of the fault on the rupture process, we perform 75 simulations, in which we vary \( L_f \), \( b_r \), and the minimum wavelength of roughness, \( \lambda_{min} \) (Figure 3-2). We consider fault lengths of \( L_f = 20 \text{m}, 30 \text{m}, \) and \( 40 \text{m} \) and three different geometries for each length. For each geometry we generate eight profiles, in which \( b_r \) ranges between \( b_r = 0.001, 0.002, 0.005 \) and \( 0.01 \) and \( \lambda_{min} \) ranges between 0.2 m and 1 m. For reference, we also run a simulation with a smooth fault for each length.
3.3 Results

3.3.1 The effect of roughness on the earthquake sequence

In order to study the effect of fault geometries on whether they experience seismic or aseismic fault slip, Figure 3-3 shows the time evolution of the maximum slip rate and average shear traction on the fault for different fault geometries. For brevity we show only the results of Geo-1 (20m), Geo-4 (30m), and Geo-7 (40m) here. Similar trends are observed for the other six geometries. We determine the initial and final stages of the slip events from the curves of the evolution of the average shear traction on the fault. The initial stages correspond to peaks in the curve, while the final stages correspond to the stages just before the beginning of the loading stages, in which the average stress increases linearly with time. Note that in slow slip events, the final stages coincide with the minima of the curves, while in fast slip events, those stages are slightly after the minima because of dynamic effects.

In general, as $b_r$ increases, faults experience more slip events but with lower slip rates and lower stress drops. A more rigorous estimation of the stress drops is provided is section 3.3, but the major trends are clearly observed in Figure 3-3. There is no clear relationship between $\lambda_{min}$ and $L_f$ and the total number of events (Figure 3-4), but the slip rates and stress drops generally increase with increasing $\lambda_{min}$, mostly for large values of $b_r$, and with increasing $L_f$ between $L_f = 20$ m and $L_f = 30$ m. Faults with $b_r \leq 0.001$ slip only seismically. They experience two fast slip events with slip rates of few meters per second. Faults with $b_r = 0.002$ experience 2 – 3 slip events, most of them fast. Faults with $b_r = 0.005$ experience both slow and fast slip events, with at least two fast events in the case of $\lambda_{min} = 1$ m or $L_f \geq 30$ m. The roughest faults ($b_r = 0.01$) experience fast slip events only for $\lambda_{min} = 1$ m. For faults with $b_r = 0.01$ and $\lambda_{min} = 0.2$ m, the maximum slip rates in the slip events are generally between $10^{-8}$ and $10^{-4}$ m/s. In two slip events on fault with Geo-8 (40), the maximum slip rates approach larger values, but these values are observed only on a small portion of about 4 m along the fault. On the rest of the faults the slip rates are always below $10^{-5}$ m/s during these two slip events. Correspondingly, the average stress drop is quite small.
Under the same initial and loading conditions, rougher faults include segments with favorite orientation that start to slip earlier and partially release the shear load. In the case of $b_r \geq 0.005$, some of the load is initially released by a few localized slow slip events. As the roughness decreases, the fault accumulates more shear load before the first slip events and consequently the peak stress, stress drop, and the maximum slip rate of the first slip event increase with decreasing roughness. At later stages, the slip events depend on the final loading and frictional conditions of prior events, and these parameters do not always decrease with increasing roughness. For example, in the case of a fault with $L_f = 20 \text{ m}$ and $\lambda_{\text{min}} = 0.2 \text{ m}$ (Figures 3-3b), a fault with $b_r = 0.002$ experiences larger peak stress and stress drop during the third slip event than those observed for a fault with $b_r = 0.001$ during the second slip event. Moreover, for faults with $b_r = 0.002$ and 0.005, these parameters are usually larger as the slip event is later in the sequence, while for smoother faults there is no clear trend.

### 3.3.2 Seismic moment

In this section, we analyze each slip event in the 75 earthquake sequences mentioned above and examine the effect of roughness on the seismic moment. Because we use a 2-D plane strain model, we examine the moment per unit length, $M_{0,1d}(t)$, which is given by

$$M_{0,1d}(t) = G \int_{L_f} u(x, t) dx = GL_f \bar{D}$$

(3.1)

where $G$ is the shear modulus, $L_f$ is the fault length, $\bar{D}$ is the average slip on the fault, and $u(x, t)$ is the slip accumulated from the beginning of the slip event. While the seismic moment defined above may involve also the slip accumulated over time scales of hours and days, for fast slip events, it is important to estimate how much of the total seismic moment is released during the dynamic stage of the rupture. To estimate the moment release during the dynamic stage of the rupture, $M_{0\text{dyn},1d}(t)$, we examine the evolution of the moment rate, which is given by

$$\dot{M}_{0,1d}(t) = G \int_{L_f} v(x, t) dx$$

(3.2)
where $v(x,t)$ is the slip rate. Figure 3-5a shows the evolution of $M_{0,1d}$, $M_{0dyn,1d}$, and $\dot{M}_{0,1d}$ during the first slip event of the sequence for fault with Geo-4 (30 m), $b_r = 0.001$, and $\lambda_{min} = 0.2$ m. We define the beginning of the dynamic stage of the rupture at the time when the moment rate exceeds a threshold value of $\dot{M}_{th,ini} = 5 \times 10^9$ N/s and calculate $M_{0dyn,1d}(t)$ with the slip accumulated from that stage. Because we are interested in comparing $M_{0dyn,1d}$ with the static stress drop during the dynamic stage of the rupture, we set the end of the dynamic stage to the time when the moment rate decreases below a lower threshold value of $\dot{M}_{th,fin} = 5 \times 10^7$ N/s, thus allowing for the dynamic effects on the shear traction to decay (Figure 3-5b).

To provide physical meaning for the threshold values, as well as more insight into the rupture process, Figures 3-5c and 3-5d show the evolution of slip rate and slip along the fault, respectively. The nucleation stage, where $\dot{M}_{0,1d} < 5 \times 10^9$, is represented by red contours at decreasing time intervals. The dynamic stage is represented by black contours for the propagation phase and gray contours for the arrest phase, both with time intervals of $1 \times 10^{-3}$ s. The displacements are calculated from the end of the previous event (in this case, the beginning of the simulation), with the first contour corresponding to the beginning of the slip event, which is the reference for $M_{0,1d}$. The rupture begins with localization of slip and slip rate between 17 and 21 m and expands asymmetrically. Three times a new slip pulse with a larger rupture velocity initiates within the existing rupture. The threshold moment $\dot{M}_{th,ini}$ corresponds to a stage in which the whole rupture is about 9 m and the slip rates along a 4 – 5 m portion of the rupture are on the order of a few cm/s.

With increasing roughness pre-factor, the complexity of the rupture increases and it may include multiple transitions between the “slow” and dynamic stages. For example, Figure 3-6a shows the evolution of $M_{0,1d}$, $M_{0dyn,1d}$, and $\dot{M}_{0,1d}$ during the third slip event of the sequence obtained for fault with Geo-4 (30 m), $b_r = 0.01$, and $\lambda_{min} = 1$ m. Most of the event is aseismic, but the figure represents a time range of 85 seconds, which includes three short sub-events where the rupture propagates dynamically ($\dot{M}_{0,1d} > 5 \times 10^9$). The time between the last two sub-events is slightly longer than a second. Between the sub-events, $\dot{M}_{0,1d}$ decreases significantly by a few orders of magnitude, and there is an
additional slow sub-event. The slip contours along the fault (Figure 3-5b) show a very complex behavior. The dynamic sub-events are localized to small portions of the fault, and the slip accumulated during these sub-events is much smaller than the total slip of the event. The propagations of the ruptures during the dynamic sub-events occur mostly on portions of the fault preferably oriented for slip (Figure 3-6c) and usually while rerupturing a portion of the fault that slipped previously during the event. For example, during the first dynamic sub-event, the rupture is locked on a barrier on the right and propagates dynamically only to the left, rerupturing a portion of the fault with a negative slope. It stops propagating dynamically at a local peak of the fault topography.

Figure 3-7.a shows the total seismic moment of all the slip events in the 75 earthquake sequences, as well as the average value of the moment for each combination of $b_r$, $\lambda_{min}$, and $L_f$. It is important to note that in the case of $b_r \geq 0.005$, there might be slip events that are not included here because they are too small to affect the behavior of the average shear traction on the fault. In general, the average and the maximum values of the seismic moment decrease with increasing roughness pre-factor and decreasing fault length and minimum wavelength, where the effect of the latter is mostly observed for $b_r \geq 0.005$. For $\lambda_{min} = 1$ m, the change in the moment of the largest event with $b_r$ is quite small and observed mostly at $b_r = 0.01$, where the maximum moment is about 60 – 70 % of that of a smooth fault. For $\lambda_{min} = 0.2$ m, a small decrease is observed also for $b_r = 0.005$ and a decrease of about one order of magnitude is observed between a smooth fault and a fault with $b_r = 0.01$. A significant decrease in the average moment is already observed at smaller values of $b_r$, where for some values of $\lambda_{min}$ and $\eta$, the largest decrease is observed at $b_r$ values between 0.002 and 0.005. In the case of smooth faults, $M_{0,1d}$ and $M_{odyn,1d}$ of the second event in the sequence are smaller than those of the first one. The finiteness of the fault leads to stress concentrations near the tips of the fault at the end of the first event; thus the second event nucleates near the ends of the fault and the propagation of the rupture is more complex. This effect decreases with increasing $b_r$.

Figure 3-7.b shows the effect of roughness on the moment ratio $M_{odyn,1d}/M_{0,1d}$ for the fast slip events. In the few cases where the fast slip events include several
dynamic sub-events (see, for example, Figure 3-6), $M_{0\text{dyn,1d}}$ is calculated with the sum of the dynamic moment release in all the sub-events. Similarly to the total seismic moment, the average and the maximum values of the moment ratio generally decrease with increasing roughness pre-factor and decreasing minimum wavelength. The effect of the length of the fault is less consistent. The moment ratio generally increases between $L_f = 20 \text{ m}$ and $L_f = 30 \text{ m}$ (Note that in the case of $L_f = 20 \text{ m}$ and $\lambda_{\text{min}} = 1 \text{ m}$ there is only one fast event for $b_r = 0.01$), but moment ratios of faults with $L_f = 40 \text{ m}$ are not always higher than those of faults with $L_f = 30 \text{ m}$. However, it is important to note that the dynamic moment itself, which can be obtained by multiplying the data in Figure 3-7.b with the corresponding data in Figure 3-7.a, clearly increases with increasing length of the fault.

### 3.3.3 Static stress drop

Another important earthquake source quantity is the static stress drop, $\Delta \tau_t$, which is defined as the difference between the average shear stress on the fault before and after the slip event. Because the local stress change varies over the fault, the averaging methodology may affect the estimated $\Delta \tau_t$. Noda et al. [2013] compared three different measures for averaging heterogeneous stress drop distributions: the moment-based, the rupture area-based, and the energy-based stress drop measures. For small levels of heterogeneity, the three measures of the average stress drop were similar, but with increasing heterogeneity significant differences were observed. Similarly to Bailey and Ben-Zion [2009] and Cocco et al. [2016], we use a rupture area-based measure, which averages the local stress drops over the portions of the fault that slipped more than a threshold value, $u_{th}$. For natural earthquakes the actual constitutive law that governs the fault is unknown and $u_{th}$ is defined as a certain percentage of the average or maximum slip on the fault [e.g. Cocco et al., 2016]. In this numerical study the fault is governed by rate and state friction with a known critical slip distance, thus we simply define the threshold slip as $u_{th} = 5L$.

Figure 3-8a shows the evolution of shear traction, $\tau_t$, with the normalized slip $u/L$ at six locations on the fault, as well as the average behavior during the first slip event in the sequence obtained for fault with $L_f = 30 \text{ m}$, $b_r = 0.001$, and $\lambda_{\text{min}} = 1 \text{ m}$ (Geo-4).
Because in this case the whole fault slips, the static stress drop and the static stress of the dynamic stage, $\Delta \tau_{t(dyn)}$, are directly obtained from the plot showing the average behavior. Note not to confuse the latter with the dynamic stress drop, which is the difference between the initial average shear stress and the average residual sliding resistance and, in this case, is smaller than $\Delta \tau_t$ by about 18% due to dynamic overshoot. The slip and the shear traction are plotted from the beginning of the simulation, and on each curve three circles denote different stages of the event. The black circle denotes the beginning of the slip event, which corresponds the maximum of the average shear traction, $\tau_{p,av}$. The blue circle denotes the beginning of the dynamic stage of the rupture. The red circle denotes the end of the slip event, which coincides with the end of the dynamic stage for this case. At the beginning of the event, the node located at 18.5 m along the fault already experienced slip of about $u = 15L$ and some reduction in $\tau_t$. By the beginning of the dynamic stage, this node experiences another increase and reduction in $\tau_t$ and approaches steady state friction. The node located at 22.5 m also experiences some slip and reduction in $\tau_t$ at this stage.

To have a better understanding of the rupture process at the stages described above, the spatial distributions of shear traction along the fault at these stages are shown in Figure 3-8b. At the stage corresponding to the beginning of the event, the rupture nucleates at a region on the fault between 17.5 m and 20.5 m with a reduction in $\tau_t$ and the development of stress concentrations at the tips. At the stage corresponding to the transition to the dynamic stage of the rupture, the rupture is 9.5 m long with a single pulse propagating to the left and two pulses propagating to the right. The region that already slips shows small spikes in $\tau_t$ at intervals of $\lambda_{min}/2$. These spikes are due to changes in the normal traction on the fault as it slips. At the end of the slip event, the rupture extends over the whole fault and the magnitudes of spikes increase.

Figure 3-9 shows $\Delta \tau_t$ and $\Delta \tau_{t(dyn)}$ in all of the slip events, as well as their average value for each combination of $b_r$, $\lambda_{min}$, and $L_f$. In general, $\Delta \tau_t$ and $\Delta \tau_{t(dyn)}$ decrease with increasing $b_r$, with a large decrease already between a smooth fault and a fault with $b_r = 0.001$. While the average values show a completely consistent decrease, the maximum values show an increase between $b_r = 0.001$ and 0.002 in some cases. The
effect of $\lambda_{\text{min}}$ is significant for $b_r \geq 0.005$, where $\Delta \tau_t$ and $\Delta \tau_{t(dyn)}$ increase with increasing $\lambda_{\text{min}}$. On the contrary to the seismic moment, the effect of $L_f$ is not consistent, except for the smooth faults which show an increase in $\Delta \tau_t$ and $\Delta \tau_{t(dyn)}$ with increasing $L_f$.

3.4 Discussion

3.4.1 Scaling relations

While both the moment and the stress drop clearly decrease with increasing roughness, it is important to examine the role of roughness in the scaling of the two. In static theories [e.g. Kanamori and Anderson, 1975], the constant stress change $\Delta \sigma$ on a fault with characteristic dimension $L = A^{1/2}$ is given by

$$\Delta \sigma = CG \frac{\bar{D}}{L}, \quad (3.3)$$

and is related to the seismic moment $M_0 = GA\bar{D}$ via

$$\Delta \sigma = C \frac{M_0}{L^3}, \quad (3.4)$$

where $C$ is a shape factor and $A$ is the area of the fault. These relations, together with the observational data, which shows that the stress drop is independent of the seismic moment [Aki, 1967; Kanamori and Anderson, 1975; Abercrombie and Leary, 1993; Abercrombie, 1995; Allmann and Shearer, 2009; Cocco et al., 2016 and references therein], suggest that the rupture process is self-similar with a constant ratio $\bar{D}/L$ for small and large earthquakes.

Considering the effect of roughness on $M_{0,1d}$ (see Figure 3-7), the ratio $\bar{D}/L = M_{0,1d}/G \bar{L}^2$ clearly decreases with increasing roughness for a given $L_f$, where $\bar{L}$ here is equal to the actual length of the rupture, $L_{rup}$, which may be smaller than the length of the fault. This was also shown by Dieterich and Smith [2009], who performed static analysis with the boundary element method. However, the break of self-similarity with roughness does not suggest that the ratio $\bar{D}/L_{rup}$ changes with the size of the fault. Moreover, it is
important to understand the relationship between this ratio and $\Delta \tau_t$, which is a measured quantity in the spontaneous dynamic rupture simulations performed is this study.

Figure 3-10a summarizes the values of $\Delta \tau_t$ vs. $\bar{D}/L_{rup}$ for all the slip events in the 75 earthquake sequences simulated in this study. As a reference, the figure also shows the static relationship in eq. (3.3) for a smooth 1-D finite fault subjected to different values of uniform shear stress reduction. Because of the finite dimensions of the model, this solution deviates from that of Starr [1928] by 14%. In general, $\Delta \tau_t$ and $\bar{D}/L_{rup}$ decrease mutually with $b_r$ such that the linear relationship is preserved, especially for fast slip events with large values of $\bar{D}/L_{rup}$ and for the slow slip events. At $\bar{D}/L_{rup}$ values between $1.7 \times 10^6$ and $3 \times 10^6$, which correspond to fast slip events on faults with $b_r \geq 0.005$, the simulated data slightly diverge from the static solution with larger values of $\Delta \tau_t$ compared to those expected by the static solution. Similar trends are also observed for the parameters $\Delta \tau_t(dyn)$ and $\bar{D}_{dyn}/L_{rup}$ of the dynamic stages of the fast slip events (Figure 3-10b), where $\bar{D}_{dyn}$ is the average of the slip accumulated during the dynamic stage of the rupture. Because the dynamic parameters are smaller, the deviation from the static solution occurs at smaller values of $\Delta \tau_t$.

The effect of the length of the faults on $\Delta \tau_t$, $\bar{D}/L_{rup}$, $\Delta \tau_t(dyn)$, and $\bar{D}_{dyn}/L_{rup}$ is consistent only for smooth faults (See also Figure 3-9). In the case of dynamic rupture on smooth faults, as $L_f$ increases, the rupture propagates to a larger distance with larger slip rates behind the fronts. Thus the average slip rate on the fault is larger and consequently also the values $\bar{D}/L_{rup}$ and $\Delta \tau_t$. Roughness introduces more complexity into the rupture process, which decreases this effect and may partially explain the seismological observations of self-similar behavior of the ratio $\bar{D}/L$ and the independence of stress drop on the moment. It is important to note that plasticity or damage on smooth faults also decreases the growth of slip rate with distance [e.g. Dunham et al., 2011b]. Moreover, in this study, we consider only faults with $L_f \leq 40$. With increasing $L_f$, the complexity and heterogeneity of the medium around the fault are expected to be larger and the friction laws should be modified to include the large reduction in the friction coefficient observed at large slip experiments [e.g. Di Toro et al., 2011]. We do not observe the decrease in
stress drop with increasing fault dimension suggested by Candela et al. [2011], but, as mentioned above, the range of $L_f$ in our study is quite small.

### 3.4.2 Dimensionality

Although the simulations in this study are performed in 2-D plane strain, we believe that they provide important information regarding the role of roughness in the rupture process, especially given that the amplitude of roughness in natural faults is significantly smaller in the direction of slip than in the perpendicular direction [Sagy et al., 2007; Candela et al., 2012]. We focus here on the global behavior of the earthquakes and because we obtain 1-D slip distributions, we study the effects of the roughness on the moment per unit length, rather than extrapolating the 1-D slip distribution in the $z$ direction and calculating the 2-D moment [e.g. Lapusta and Rice 2003]. However, we believe that the trends obtained here with plane strain modeling, which is equivalent to a fault with a very large dimension in the $z$ direction, are also adequate for faults with dimension in the $z$ direction comparable to $L_f$ and that the obtained 1-D slip distribution could be extrapolated in the $z$ direction to give a 2-D seismic moment.

To examine this assumption, we consider the static problem of a square fault with dimensions $L_f \times L_f$ under constant reduction of shear stress. Using 3-D numerical models, Parsons et al. [1988] and Noda et al. [2013] found that the shape factor for this geometry is $C = 2.53$. We calculate an equivalent shape factor $C_{eq}$ by extrapolating the 1-D slip distribution $u(x)$ obtained by the plane strain solution [Starr, 1928] to a 2-D slip distribution $u(x, z)$ on a square fault. Assuming an elliptic slip distribution in the $z$ direction, with $u(x, 0) = u(x)$ and $u(x, \pm L_f/2) = 0$, the average slip on the fault is given by

$$
\bar{D}(x, z) = \frac{L_f \Delta \sigma}{4 \sqrt{\frac{3}{2}}} \int_{-L_f/2}^{L_f/2} \sqrt{1 - \left(\frac{x - L_f/2}{L_f/2}\right)^2} \sqrt{1 - \frac{z^2}{(L_f/2)^2}} \, dx \, dz / L_f^2, \quad 0 \leq x \leq L_f
$$

and $C_{eq} = 2.16$, which is only 15% smaller than that obtained with the 3-D models. This simple analysis may suggest that the trends obtained by the 2-D model (1-D fault) can be
extrapolated to a 2-D fault with similar dimension in the $z$ direction, but with small overestimation of the slip obtained for a given stress drop and vice versa. It is important to note that this analysis doesn’t take into account the dynamics of the rupture and the effect of roughness on the extrapolation in the perpendicular dimension.

### 3.5 Conclusions

We study the response of rough faults governed by velocity weakening rate and state friction to slow tectonic loading, using the mortar-based method developed in chapter 2. The method enables the modeling the whole seismic cycle, including the slow aseismic nucleation, the dynamic propagation, and the arrest of the rupture.

We focus on the global behavior of the faults and find that as the roughness pre-factor $b_r$ increases, there is a transition from seismic slip behavior to aseismic slip behavior, in which the load on the fault is released by more slip events but with lower slip rate, seismic moment $M_{0,1d}$, and the average static stress drop $\Delta \tau_t$. These parameters also decrease with decreasing minimum wavelength of roughness $\lambda_{\text{min}}$, especially for large values of $b_r$.

The source parameters $M_{0,1d}$ and $\Delta \tau_t$ decrease mutually such that the linear relationship between $\Delta \tau_t$ and the average fault strain $\bar{D}/L_{\text{rup}}$ is preserved, except for fast slip events on faults with large values of $b_r$, where the decrease in slip is larger than that of $\Delta \tau_t$ and vice versa.

Because the total deformation in fast slip events involves also aseismic slip and may also occur over time scales of hours and days, we also estimate the seismic moment $M_{\text{dyn,1d}}$ and stress drop $\Delta \tau_t(\text{dyn})$ of the dynamic stage of the ruptures for these events, based on the evolution of moment rate on the fault. We find even larger decreases of these parameters with increasing roughness compared to their corresponding total values.

A consistent effect of the length of the fault on $\Delta \tau_t$ and $M_{0,1d}/L_f^2$ is observed only for smooth faults, thus we speculate that roughness may be one of the reasons for the stress drop being independent of magnitude.
Figure 3-1: (a) The problem set up: a finite fault with length $L_f$ is embedded in a 2D elastic medium with dimensions $3L_f \times 1.5L_f$, which is subjected to a prescribed slow horizontal velocity $V_b = \pm \frac{L_f}{20} \times 5 \times 10^{-10}$ m/s at the top and bottom, and initial stresses $\sigma_{xx0}$, $\sigma_{yy0}$, and $\sigma_{xy0}$. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. The model includes an absorbing layer with gradual Rayleigh damping. (b) The mesh structure around the left side portion of a rough fault with length of 30 m. The mesh is refined with hanging nodes to give elements on the fault with dimensions of about 1.56 x 1.56 cm.
Figure 3-2: The fault profiles examined in this study: We consider a total of nine different general geometries for fault lengths of $L_f = 20\text{m}$ (a), $30\text{m}$ (b), and $40\text{m}$ (c). For each geometry eight profiles are generated with roughness pre-factor values of $b_r = 0.001$, $0.002$, $0.005$ and $0.01$ and minimum wavelength of $\lambda_{min} = 0.2\text{ m}$ and $1\text{ m}$. For reference, we also run a simulation with a smooth fault for each length.
Figure 3-3: The evolution of maximum slip rate (a, c, and e) and average shear traction (b, d, and f) obtained for the fault geometries Geo-1 (20m), Geo-4 (30m), and Geo-7 (40m) shown in Figure 3-2, as well as for smooth faults.

Figure 3-4: The number of events vs. $b_r$ for different values of $L_f$ and $L_{min}$. For rough faults, there are three different sequences for each of $b_r$, $L_f$, and $L_{min}$, and the number of event is the sum of all the events in these sequences divided by 3.
Figure 3-5: (a) The evolution of $M_{0,1d}$ (black), $M_{0,1d}$ (blue), and $M_{0,1d}$ (red) during 0.1 s of the first slip event of the sequence obtained for fault with Geo-4 (30 m), $b_r = 0.001$, and $\lambda_{\text{min}} = 0.2$ m. We define the beginning of the dynamic stage of the rupture at the stage when the moment rate exceeds a threshold value of $\dot{M}_{\text{th,ini}} = 5 \times 10^9$ N/s and calculate $M_{0,1d}(t)$ with the slip accumulated from that stage. (b) The evolution of shear traction at six locations along the fault. (c) Profiles of slip rate along the fault. (d) Profiles of slip. The stage where $\dot{M}_{0,1d} < 5 \times 10^9$ is represented by red contours at decreasing time intervals. The dynamic stage is represented by black contours for the propagation phase and gray contours for the arrest phase, both with time intervals of $1 \times 10^{-3}$ s.
Figure 3-6: (a) The evolution of $M_{0,1d}$, $M_{0,1d}$, and $\dot{M}_{0,1d}$ during the third slip event of the sequence obtained for fault with Geo-4 (30 m), $b_r = 0.01$, and $\lambda_{\text{min}} = 1$ m. Most of the event is aseismic, but the figure represents a time range of 85 seconds, which includes three short sub-events where the rupture propagates dynamically ($\dot{M}_{0,1d} > 5 \times 10^9$). (b) Contours of the evolution of slip along the fault. The stage where $\dot{M}_{0,1d} < 5 \times 10^9$ is represented by red contours at decreasing time intervals. The dynamic stage is represented by black contours with time intervals of $1 \times 10^{-3}$ s. (c) The geometry of the fault.
Figure 3-7: (a) $M_{0,1d}$ vs. $b_r$ for $L_f = 20$ m (blue), $L_f = 30$ m (green), and $L_f = 40$ m (red) in all the slow (‘x’ symbols) and fast (‘o’ symbols) slip events in the 75 simulated earthquake sequences, as well as the average values (solid curves). The events on faults with $\lambda_{min} = 0.2$ are shown on the left and on those with $\lambda_{min} = 1$ are shown on the right. (b) The moment ratio $M_{0,dyn,1d}/M_{0,1d}$ vs. $b_r$ for $\lambda_{min} = 0.2$ and $\lambda_{min} = 1$ for all the fast slip events, as well as the average values. Note that in the case of $b_r = 0.01$, $\lambda_{min} = 1$ m, and $L_f = 20$ m, there is only one fast slip event, so the average value is not representative.
Figure 3-8: The evolution of $\tau_t$ vs. normalized slip ($u/L$) at six locations on the fault, as well as the average behavior, during the first slip event in the sequence obtained for fault with $L_f = 30$ m, $b_r = 0.001$, and $\lambda_{min} = 1$ m (Geo-4). The slip and the shear traction are plotted from the beginning of the simulation, and circles denote the beginning of the slip event (black) and the dynamic stage of the rupture (blue) and end of the slip event (red). (b) The spatial distributions of shear traction along the fault at these stages.
Figure 3-9: (a) $\Delta \tau_t$ vs. $b_r$ for $L_f = 20$ m (blue), $L_f = 30$ m (green), and $L_f = 40$ m (red) in all the slow ('x' symbols) and fast ('o' symbols) slip events in the 75 simulated earthquake sequences, as well as the average values (solid curves). The events on faults with $\lambda_{\min} = 0.2$ are shown on the left and on those with $\lambda_{\min} = 1$ are shown on the right. (b) $\Delta \tau_{t(dyn)}$ vs. $b_r$ for $\lambda_{\min} = 0.2$ and $\lambda_{\min} = 1$ for all the fast slip events, as well as the average values. Note that in the case of $b_r = 0.01$, $\lambda_{\min} = 1$ m, and $L_f = 30$ m, there is only one fast slip event, so the average value is not representative.
Figure 3.10: $\Delta\tau_s$ vs. $\bar{D}/L_{rup}$ (a) and $\Delta\tau_{t(syn)}$ vs. $\bar{D}_{dyn}/L_{rup}$ (b) for all the slow (filled symbols) and fast (open symbols) slip events in the 75 simulated earthquake. The color of the symbols denotes different values of $b_r$, while the type denotes different fault lengths. Slip events on fault with $\lambda_{min} = 0.2$ and $\lambda_{min} = 1$ are shown with thin and thick symbols, respectively. The static solution for the case a uniform shear stress change on a smooth 1-D finite fault embedded in a 2-D domain is also shown (solid line).
Table 3-1. Model Parameter Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frictional properties</strong></td>
<td></td>
</tr>
<tr>
<td>Direct-effect parameter</td>
<td>$a$</td>
</tr>
<tr>
<td>Evolution-effect parameter</td>
<td>$b$</td>
</tr>
<tr>
<td>Reference velocity</td>
<td>$v^*$</td>
</tr>
<tr>
<td>Reference friction</td>
<td>$\mu^*$</td>
</tr>
<tr>
<td>Critical slip distance</td>
<td>$L$</td>
</tr>
<tr>
<td>Initial friction</td>
<td>$\mu_0$</td>
</tr>
<tr>
<td>Initial state variable</td>
<td>$\theta_0$</td>
</tr>
<tr>
<td><strong>Bulk properties</strong></td>
<td></td>
</tr>
<tr>
<td>Young’s modulus</td>
<td>$E$</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td>$\nu$</td>
</tr>
<tr>
<td>Density</td>
<td>$\rho$</td>
</tr>
<tr>
<td><strong>Initial remote stresses</strong></td>
<td></td>
</tr>
<tr>
<td>Horizontal stress</td>
<td>$\sigma_{xx0}$</td>
</tr>
<tr>
<td>Vertical stress</td>
<td>$\sigma_{yy0}$</td>
</tr>
<tr>
<td>Shear stress</td>
<td>$\sigma_{xy0}$</td>
</tr>
</tbody>
</table>
Chapter 4

The effects of roughness on the nucleation process

Abstract

We study numerically the effect of roughness on the nucleation process during earthquake sequences of at least two events. The faults are governed by a rate and state friction law with an aging evolution law. The roughness introduces local barriers that complicate the nucleation process and result in asymmetric expansion of the rupture, non-monotonic increase in the slip rates on the fault, and the generation of multiple slip pulses. These complexities are reflected as irregular fluctuations in the moment rate. There is a large difference between first slip events in the sequences, where the initial conditions are homogenous, and later events, where the initial stress field and friction conditions are determined by the rupture growth and arrest in previous slip events. In the first events there is a large increase in the nucleation length with increasing $b_r$ for $b_r \leq 0.002$, and a transition to aseismic or mostly aseismic deformation for larger values, where $b_r$ is the roughness pre-factor. For slip events later in the sequence there is a trade off between the effects of the finiteness of the fault and the roughness. For $b_r \leq 0.002$, the effect of the finiteness of the fault on the initial stresses is a more dominant factor in the nucleation process, thus the nucleation length barely changes with $b_r$, and most of the events initiate close to the ends of the fault. For larger values of $b_r$, the roughness seems to play a larger role, thus the nucleation length increases with $b_r$ and the location where the events initiate is more variable.

4.1 Introduction

A common view is that earthquakes occur via a shear rupture instability, in which the frictional resistance on a pre-existing fault decreases with increasing sliding or sliding
velocity [e.g. Scholz, 2002]. As the instability occurs, the rupture propagates dynamically at a high speed close to the wave speed and with slip velocities much larger than the loading rate. Laboratory experiments on preexisting faults show that the development of the unstable rupture occurs via a nucleation process, which is characterized by roughly two phases: A quasi-static phase in which the rupture grows at a steady slow velocity with accelerating slip and a phase where the rupture grows with accelerating speed [e.g. Dieterich, 1978; Okubo and Dieterich, 1984; Ohnaka and Shen, 1999; Nielsen et al., 2010; McLaskey and Kilgore, 2013].

Rate and state friction constitutive laws [Dieterich, 1979; Ruina, 1983] have emerged as powerful tools for investigating various earthquake phenomena, including earthquake nucleation [Marone, 1998]. Numerical and theoretical models with rate and state fault friction are generally consistent with laboratory observations, but provide additional insight into the nucleation process, as well as up-scaling of the lab observations to natural faults. The models show that the exact behavior of the rupture during nucleation, and specifically how and to what extent the rupture expands between the initial localization of slip and slip rate and the dynamic stage of the rupture, is highly affected by the loading and initial conditions and the rate and state parameters and evolution laws [Dieterich, 1992; Lapusta and Rice, 2003; Rubin and Amuero, 2005; Amuero and Rubin, 2008; Kaneko and Lapusta, 2008; Fang et al., 2010; Kaneko and Amuero, 2011].

A source of complexity in the nucleation process is expected to arise from the deviation of faults from planarity. High-resolution map traces of large continental strike-slip earthquake surface ruptures and measurements of roughness on exhumed faults [Brown and Scholz, 1985; Power et al., 1987; Power and Tullis, 1991; Renard et al., 2006; Sagi et al., 2007; Candela et al., 2009, 2012; Klinger, 2010; Bistacchi et al., 2011; Brodsky et al., 2011] show that faults are rough at all scales and can generally be described as self-affine fractal surfaces. Experimental observations show that, on the lab-scale, roughness affects frictional constitutive parameters such as stress drop and critical slip weakening distance and results in larger nucleation zones [Okubo and Dieterich, 1984; Ohnaka and Shen, 1999; Ohnaka, 2003]. On the scale of natural faults, the effects of roughness are not yet clear. Numerical studies have focused so far only on the effects
of roughness on the dynamic stage of the rupture [e.g. Dunham et al., 2011; Fang and Dunham, 2013; Shi and Day, 2013; Bruhat et al., 2016].

In this study, we use the numerical approach developed in chapter 2 to explore the effects of roughness on the nucleation process of faults governed by rate and state friction laws. The simulations include sequences of at least two seismic cycles, thus enable us to examine the effect of the stress state and frictional conditions resulting from the rupture growth and arrest for a given slip event on the nucleation process at subsequent events. We focus on the scale of small earthquakes and consider faults with a length of 40 m. We choose the minimum roughness wavelength to be at a size close to lab samples (20 cm) and thus use the observed lab-scale rate and state friction laws without up scaling the constitutive parameters, assuming that the experimental data already include the effects of smaller wavelengths of roughness.

4.2 Model description

We examine effects of roughness on the nucleation of shear rupture on a 40 m long finite fault embedded in a 2-D elastic medium with dimensions 120 x 80 m (Figure 4-1a). We assume plane strain boundary conditions and load the top and bottom boundaries of the medium with a prescribed slow velocity of $V_b = \pm 10^{-9}$ m/s. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. This setup enables completely spontaneous nucleation of the rupture. We do not explicitly consider pore fluids in this study and assume that the stresses are effective stresses if the medium is saturated with fluids. We perform 25 simulations. We consider three different fault geometries of self-affine fractals with Hurst exponent of $H = 0.8$ (Figure 4-1b). For each geometry we generate eight profiles, in which $b_r$ ranges between 0.001, 0.002, 0.005 and 0.01 and $\lambda_{min}$ ranges between 0.2 m and 1 m. For reference, we also run a simulation with a smooth fault. The fault is governed by rate and state friction, which is given in the standard aging formulation by [Dieterich, 1979; Ruina, 1983]

$$\mu = \mu^* + a \ln \left( \frac{v + v_{th}}{v^*} \right) + b \ln \left( \frac{L \theta}{v^*} \right)$$ (4.1)
\begin{equation}
\dot{\theta} = 1 - \frac{\theta(v + v_t)}{L},
\end{equation}

where $a$ and $b$ are rate and state constitutive parameters, $v$ is the slip rate, $v^*$ is a reference slip rate, $\mu^*$ the steady-state friction at $v = v^*$, $\theta$ is a state variable, and $L$ is the characteristic sliding distance. We add a threshold velocity term, $v_{th} = 10^{13}$ m/s to avoid singularity at $v = 0$. We do not consider the effect of normal stress variations on $\theta$. The mechanical properties of the medium, friction law parameters, and initial stresses are given in Table 3-1.

Currently, the numerical method accounts only for elastic rheology of the medium surrounding the fault. This approximation limits the amount of deformation that the medium can experience before unrealistic stresses larger than Coulomb failure criteria are accumulated in the material around the fault. Therefore, we set the total time of the simulation to the time that two large events would occur, if a smooth fault were considered. Because we do not model a large sequence of earthquakes, the initial conditions have a significant effect on the results, despite the spontaneous nucleation. We conceptually begin the simulation at the end of an earthquake that ruptured the whole fault and choose the initial friction parameters accordingly. We assume that the earthquake approached a slip rate in the order of 1 m/s during the rupture, that the slip rate is currently zero, and that the state variable and the friction coefficient had no time to evolve and are equal to their steady state values at slip rate of 1 m/s, $\theta_0 = L/(v_t = 1)$ and $\mu_0 = (a - b) \ln(v_t = 1)$, respectively. The initial shear stress is chosen such that $\sigma_{xy0} = \mu_0 \sigma_{yy0}$. As the roughness of the fault increases, some segments along the fault may begin to slip under smaller shear stress. In these cases, the initial shear stress is smaller than $\mu_0 \sigma_{yy0}$, but in order to maintain similar initial conditions in all simulations, we do not allow the state variable evolve until the remote shear stress exceeds $\sigma_{xy}$. We use the Mortar Finite elements method described in chapter 2 and refine the mesh around the fault with hanging nodes in order to represent the geometry of the fault properly. This results in 2561 nodes along the fault with a mesh spacing of about 1.56 cm. We use a variable time step and based on the current values of slip rates we estimate the time step size at the next time step such that the average incremental slip of the 40 fastest nodes along the fault will not be larger than half of the size of the critical distance.
This procedure results in a time step size that represents the evolution of the friction coefficient well without reducing the time step size to values that lead to simulations with an exaggerated number of time steps. The method switches between quasi-static and dynamic time integration schemes when the average slip rate at the 40 fastest nodes on the fault is larger than $5 \times 10^{-5}$ m/s. We use the Newmark time integration scheme with small algorithmic damping ($\beta = 0.35$ and $\gamma = 0.7$) for the dynamic stage and also add an absorbing layer with gradual Raleigh damping near the boundaries of the model.

4.3 Results

4.3.1 The nucleation process

To study the nucleation process, we need to define when the nucleation process begins and when nucleation ends and the dynamic stage begins. For the friction parameters adopted in this study, the nucleation process involves stages both of localization and expansion of the rupture with accelerating slip rates [Rubin and Ampuero, 2005]. The determination of the exact stage in which those processes start is somewhat subjective, and becomes more complex with the addition of roughness. Therefore, we determine the beginning of the nucleation stage, from peaks in curves of the time evolution of the average shear traction on the fault, $\tau_{av}$ (Figure 4-2a). With this choice, the nucleation process may begin after the beginning of the localization, but this has no effect on the findings described in this paper.

The transition to the dynamic stage of the rupture is defined at the time at which both the moment rate per unit length, $\dot{M}_{0,1d} = G \int_{L_f} v(x, t) dx$, is larger than a threshold value of $5 \times 10^9$ N/s and an active slip pulse moves at speed larger than a threshold of 20% of the shear wave speed of the surrounding medium. $G$ is the shear modulus of the medium. Because with increasing roughness there are many fluctuations in the rupture velocity $v_r$, the latter threshold is larger than that used by Kaneko and Lapusta [2008]. Figure 4-2b shows the evolution of $\tau_{av}$ during the first slip events in the sequences obtained for faults with $\lambda_{min} = 0.2$ m and $b_r = 0.001$ and 0.002, as well as the transition to the dynamic stage. The nucleation begins with a very small decrease in $\tau_{av}$, over a time of
about 5,500 s and 27,000 s for $b_r = 0.001$ and 0.002, respectively. For $b_r = 0.002$, the nucleation includes another stage of moderate decrease in $\tau_{\alpha}$ over about 1000 s, in which there are some fluctuations in the rate of decrease. After the transition to the dynamic stage of the rupture, there is a rapid decrease in $\tau_{\alpha}$. For $b_r = 0.001$, there is a sharper transition into the dynamic stage.

To examine the effect of roughness on the nucleation process, Figure 4-3 shows the evolution of slip $u$, slip rate $v$, and the friction coefficient $\mu$ along the fault during the sequences obtained for a smooth fault and for a rough fault with Geo-1, $b_r = 0.001$, and $\lambda_{\text{min}} = 0.2$ m, where for the evolution of $\mu$ is shown only for the first slip event in each sequence. We show here the evolution of $\mu$, rather than the shear stress, because the roughness the large spatial variations in the normal and shear stress mask the behavior of the rupture. The contours of $u$ are plotted for the loading, nucleation, and dynamic (propagation and arrest) stages of the seismic cycle. In the case of the first event of the sequence obtained for a smooth fault, the contours of $v$ and $\mu$ are plotted for the nucleation and the whole dynamic stage. However, for better visualization of the nucleation stage, the contours of $v$ and $\mu$ are shown only for the nucleation stage and the beginning of the dynamic stage in other slip events.

We begin with examining the behavior of the smooth fault (Figures 4-3a, 4-3c, and 4-3e). Because the fault is finite, at the beginning of the first slip event on the smooth fault, the profiles of $u$ and $v$ have a maximum value at the center of the fault. Correspondingly, $\mu$ shows a very small reduction of about 0.02 % at the center. Consequently, there is a localization of $u$ and $v$ and further decrease in $\mu$ at the center, which is followed by expansion of the rupture in a crack-like fashion, with peaks in $v$ and $\mu$ near the tips and relatively uniform values in the interior. This behavior was also observed in the simulations of Rubin and Ampuero [2005] and Fang et al. [2010] when similar $a/b$ values where used. Both $v$ and $v$, and increase as the rupture expands, and the nucleation length is about 5.6 m, where the nucleation length, $l_{\text{nuc}}$, is defined as the total length of the rupture at the transition to the dynamic stage. Note that we use the total length of the rupture, while some studies refer to the half rupture length. The complex arrest of the rupture and the dynamic effects result in non-uniform distributions of $\mu$ and
v at the end of the slip event, with maximum values about 5 m from the ends of the fault. These become the initial condition for the second slip event, which begins with a localization of $u$ and $v$ at two locations 5.5 m from the left and right ends of the fault. However, because of small numerical errors during the dynamic stage of the rupture, the profiles of $u$ and $v$ are not perfectly symmetric with slightly larger values at the right hand side. This leads to the nucleation of the rupture on the right side of the fault, again in a crack-like fashion, but with some effects of the location being close to the end of the fault. The nucleation length is larger, with $2l_c \approx 8.5$ m. It is important to note that, practically, the effect of numerical errors, which leads to the break in the symmetry of the second slip event, is negligible because it is much smaller than the effect of any heterogeneity of the fault.

In the case of a rough fault with Geo-1, $b_r = 0.001$, and $\lambda_{\text{min}} = 0.2$ m (Figures 4-3b, 4-3d, and 4-3f), the nucleation process is more complex and highly affected by local geometric barriers, which slow down the expansion of the rupture and the increase in slip rate. At the beginning of the first slip event, there are already localizations of $u$ and $v$, with the maximum values on a 5 m long segment at the center of the fault. A portion of this segment also experiences a reduction of about 2% in $\mu$. It is important to note that the rupture initiates in this location because of the local slope of the fault (see Figure 4-1), rather than the effect of the finiteness of the fault. The slip rates at the beginning of the expansion process are a few orders of magnitude smaller than those in the initial event on the smooth fault. In general, smaller slip rates at this stage promote more expansion of the rupture before the transition to the dynamic stage [Fang et al., 2010]. The rupture expands to both sides and develops peaks in $v$ and $\mu$ near the tips, but then it approaches a barrier on the right, and expands much faster to the left as a unilateral pulse. When the slip rate on the left front is about $10^3$ m/s, a new slip pulse initiates on the left and propagates to right on the existing rupture. Because of a barrier on the left front of the rupture, the new pulse becomes more dominant and has a larger $v$. At the transition to the dynamic stage of the rupture, the total length of the rupture is about 9 m and there is a new pulse that propagates to the left. As new pulses initiate later in the nucleation process, the rerupturing process is associated with a smaller peak in $\mu$, as well as a smaller reduction. As they join the stationary pulses at the fronts of the rupture, the peaks
in $\mu$ instantaneously increase. Similar to the smooth fault, the second slip event also starts next to one of the ends of the fault. The nucleation process during the second event is somewhat less complex than that of the first. Note that for the second slip event, the slip rates at the beginning of the expansion process are similar to those in the second slip event on the smooth fault and consequently they both have similar nucleation lengths.

In general, the complexity of the nucleation process increases as $b_r$ increases. In the first slip event on the fault with Geo-1, $b_r = 0.002$, and $\lambda_{min} = 0.2$ m (Figures 4-4a, 4-4c, and 4-4e), the nucleation stage involves many new slip pules, as well as stages where $v$ decreases as the rupture grows. Moreover, the complexity in the nucleation process and the lower slip rates in the beginning of the expansion process lead to a large nucleation length of $2l_c \approx 22$ m. Similar to the fault with $b_r = 0.001$, the second slip event shows simpler behavior and larger slip rates at the beginning of the expansion process, thus the nucleation length is small. In the case of the rough fault with Geo-1, $b_r = 0.005$, and $\lambda_{min} = 0.2$ m (4-4b, 4-4d, and 4-4f), there are three slip events. The first slip event is almost completely aseismic, with only 3% of the total slip occurring during the dynamic stage of the rupture. Moreover, the loading stage involves much slip, which is partly accumulated during very slow slip events with a maximum slip rate of $v \sim 10^{-8}$ m/s (see Figure 4-2c). Because these events barely affect the evolution of the average shear stress on the fault, they are not considered here as separate slip events. The nucleation lengths of the second and third slip events are $2l_c \approx 15$ and 33 m, respectively, and about 40% of the slip accumulated during the third slip event is aseismic.

As $b_r$ increases and the slip events are later in the sequences, the number of nodes that are not in contact increases. When the active rupture passes through these nodes, they slip at slip rates similar to their neighboring nodes. However, non-contact nodes that are not on the active rupture experience substantial variation in $v$ because of the dynamic waves radiated by the rupture, including transitions to negative values, where the logarithm is undefined. Moreover, the value of $\mu$ for the non-contact nodes is meaningless. Therefore, we do not plot the values of $\mu$ and $v$ at the stages when they are not in contact.

At stages when a portion of the fault already slips at a large slip rate, there are spikes with large values of $v$ at regions on the fault that are not part of the active rupture.
These spikes correspond to nodes where the normal stress continuously decreases. As the normal stress decreases, there is positive feedback: the decrease in the resistive forces leads to an increase in \( v \) and \( u \), which leads to a further decrease in the normal stress. The slip rate at the spikes is significantly larger than in the surrounding nodes, but is generally smaller than \( 10^{-4} \) m/s. At these slip rates, the strain is so small that the deformation can be very local. Note also that the fault includes 2561 nodes, thus spikes that include several nodes also seem extremely localized in the plots.

### 4.3.2 Nucleation length

The nucleation lengths, \( l_{nuc} \), of the fast slip events in the 25 simulated earthquake sequences are summarized in Figure 4-5. Only slip events where the slip accumulated during the dynamic stage of the rupture is more than half of the total slip are considered. In general, \( l_{nuc} \) increases with increasing slope as \( b_r \) increases, but there is also substantial variability in the values as \( b_r \) increases. A significant effect is observed for whether the slip event is first or later in the sequence, where except for the smooth fault, the values of \( l_{nuc} \) in the first slip events are larger than those of later events. In the case of the first slip events, \( l_{nuc} \) increases rapidly with \( b_r \) for \( b_r \leq 0.002 \) and there are no fast slip events for larger \( b_r \) values. For slip events later in the sequence, there is a very small increase in \( l_{nuc} \) with \( b_r \) for \( b_r \leq 0.002 \) and a large increase for higher \( b_r \) values. The effect of \( \lambda_{min} \) on \( l_{nuc} \) is negligible for \( b_r \leq 0.002 \). As \( b_r \) increases, \( l_{nuc} \) significantly increases with decreasing \( \lambda_{min} \), where for \( b_r = 0.01 \), there are fast slip events only for \( \lambda_{min} = 1 \).

In the case of the smooth fault, the nucleation length of the first slip event is about half of the estimate of Rubin and Ampuero [2005], who suggest that for a smooth fault with homogenous initial conditions:

\[
l_{nuc} = \frac{2GLb}{(1 - v)\pi \sigma_n (b - a)^2}.
\]  

(4.3)

However, the value obtained here is consistent with the study of Fang et al. [2010], which shows that the loading rate and the initial conditions of \( v \) and \( \theta \) significantly affect the nucleation-zone expansion process and that the nucleation length is smaller than that predicted in Rubin and Ampuero [2005] for most initial conditions, including those in natural earthquake cycles.
4.3.3 Initiation of the rupture

4.3.3.1 Location on the fault

In all slip events, the nucleation process begins with localization of slip and slip rate and local reduction in the frictional resistance (see Figures 4-3 and 4-4). In this section, we examine the effects of roughness on the location where the nucleation process of fast slip events initiates, as well as the effects of the final conditions of the preceding slip events when the slip events are later in the sequence. As mentioned earlier, we define the beginning of the slip events at peaks in curves of the time evolution of the average shear stress on the fault. However, as the roughness increases, more slip is accumulated during the loading stage and the localization process begins in multiple regions before the beginning of the slip event (as we defined it), although with very small slip rates. We are interested in the region in which the actual rupture initiates, and study the location of this region in the rest of this section. To define an exact location, we search for a peak in the slip accumulated during the loading stage in this region (Figure 4-6). It is important to note that the complex nucleation process on rough faults may lead to a completely different location of the final slip pulse at the transition to the dynamic stage.

In the first fast slip events in the earthquake sequences, the ruptures initiate at the center of the fault for Geo-1, 6 – 8 m from the ends of the fault for Geo-2, and 2.5 m from the center of the fault for Geo-3 (Figure 4-7). These locations correspond to regions on the faults where the orientation is preferable for slip, i.e. regions with relatively large negative slopes of the fault topography and with no adjacent geometric barriers with a large positive slope. Note that the figure shows the absolute distance of the initial localization from the center of the fault $|x_{ini}|$. In the case of Geo-2 and $\lambda_{min} = 0.2$ m, the initiation of the actual rupture occurs at the left side of the fault for $b_r = 0.001$ (Figure 4-6). For $b_r = 0.002$, the slope of this section is a more negative, thus it starts to slide under smaller loading and the available energy is smaller. Moreover, the rupture has to propagate through barriers with larger amplitude. Therefore, the slip rate on this section does not accelerate and the actual rupture initiates on a section on the right that accumulated less slip during the loading stage.
For slip events that are later in the sequences, \( |x_{\text{in}}| \) is determined by both the state of stress at the end of earlier events and the geometry of the fault. For \( b_r \leq 0.001 \), the later events initiate 4 – 6 m from the ends of the fault, where the shear stress at the end of preceding slip events is maximal. As \( b_r \) increases, the effect the geometry becomes more important. In the case of faults with \( b_r \geq 0.005 \) and Geo-1, the location where the later events initiate is always next to the center, but for Geo-2 and Geo-3 it is more scattered between few regions where the geometry of the fault is preferable for slip.

### 4.3.3.2 Loading stage

To examine further the initiation of the rupture in later events, Figure 4-8 shows the evolution of shear tractions \( \tau, v, \theta, \) and \( \mu \) at six loading stages between the end of the first slip event and the beginning of the second slip event for a smooth fault and a fault with Geo-1, \( b_r = 0.002 \) and \( \lambda_{\text{min}} = 1 \text{ m} \). The figure also shows the six loading stages on curves of the average shear stress on the fault vs. time. For both geometries, the distributions of \( v, \theta, \) and \( \mu \) at stage #1 are not correlated with the local geometry of the fault. They result from the finiteness of the fault and the complex arrest of the rupture during the first slip event. The maximum values of \( v \) and \( \mu \) at this stage are at both \( X = 5 \) and 35 m for the smooth fault and \( X = 5 \) m for the rough fault. For the rough fault, there are large spatial variations in \( \tau \), which correlate with \( \lambda_{\text{min}}/2 \). These are due to changes in the normal tractions on the fault as it slipped during the first slip event. Note that stage #1 is defined at the stage where the average stress drop begins to increase consistently with time, which is 500 – 800 s after the dynamic stage of the rupture (see Figure 4-8).

Consider the smooth fault. At stage #2, all the nodes on the smooth fault are in a stuck state, thus the slip rate considered for the friction calculation is the threshold velocity (eqn. 1 and 2) and \( \mu \) is spatially constant. However, this does not seem to be an important issue because the shear tractions are not affected by the threshold velocity. At stages 3 – 5, \( \tau, v, \) and \( \mu \) increase with time, maintaining their distribution from stage 1. At stages 1 – 4, the term \( v\theta/L \) in eq. (4.2) is much smaller than one, thus \( \theta \) is spatially constant and equal to the time from the dynamic stage of the first slip event. At stage #5, \( \theta \) varies spatially with a distribution that is a mirror image of the distributions of \( \tau, v, \) and
At stage #6, the localization in $v$ begins, with a maximum value at $X \approx 34.5$ m. Correspondingly, there are reductions in $\tau$, $\theta$, and $\mu$ in this region.

For the rough fault, some of nodes are at a stuck state with constant value of $\mu$ at stage #2, while other nodes slip with different values of $v$ and $\mu$. The large-scale distributions of $v$ and $\mu$ at the slipping nodes correspond to their distributions at the end of the first slip event, but there are also small-scale variations that correspond to the roughness and the spatial variations of the normal tractions. At stages 3 – 5, $\tau$, $v$, and $\mu$ increase with time, but the amplitude of the small-scale variations in $v$ and $\mu$ decreases, especially on the left side of the fault, where both are larger. At stages 1 – 3, $\theta$ is spatially constant and equal to the time from the dynamic stage of the first slip event, while at stages 4 – 5 it varies spatially with a distribution that is a mirror image of the distributions of $v$ and $\mu$. At stage #6, $v$ does not localize exactly at the region where the values of $v$ and $\mu$ are maximum at stage #1, but there is a shift of 2.5 m to the right because the localization cannot take place on a region where the slope of the fault is positive, especially as $b$, increases. The small-scale variations in $v$ that correspond to $\lambda_{min}/2$ are not observed in the region where the localization occurs. The localization is accompanied by reductions in $\tau$, $\theta$, and $\mu$ at the same region and the initial development of peaks in $\mu$ at the ends of this region.

4.4 Discussion

4.4.1 The effect of the initial conditions and fault geometry on the nucleation process

The location of the slip event in the earthquake sequence has large a effect on the nucleation process, and especially whether the slip event is first or later in the sequence. Moreover, the effect of roughness is generally larger for the first slip events. The first slip events share similar initial conditions with initially homogeneous stress along the fault. As $b_s$ increases, segments with preferable orientation slip earlier and partially release the shear load, thus the available energy for the rupture process is smaller and the transition between the localization and expansion processes of the rupture is earlier. The earlier
transition is accompanied by smaller slip rates (see Figures 4-3 and 4-4), which enable more expansion of the rupture before the transition to the dynamic stage. Moreover, the rupture has to propagate through barriers with larger amplitude, which complicates the nucleation process. For $b_r \leq 0.002$, these effects lead to significant increases in $l_{mc}$ with $b_r$ for the first events. For larger values of $b_r$, the first events are aseismic, or involve mostly aseismic deformation.

For slip events that are later in the sequences, the initial stresses are not constant along the fault and are determined by the rupture process in the previous slip event, which is itself determined by the roughness and finiteness of the fault. The latter has a large effect on the arrest stage of the rupture. The non-constant initial stresses together with the roughness govern the loading and nucleation stages of the rupture in the later slip events. The small change in $l_c$ with $b_r$ for $b_r \leq 0.002$ for the later events, suggest that the effect of the finiteness of the fault on the initial stresses is the most dominant factor in the nucleation process for these values of $b_r$. This is also reflected in the location where most of these events initiate for $b_r \leq 0.001$ and for some of the slip events on fault with $b_r = 0.002$. For larger values of $b_r$, the roughness seems to play a larger role.

We consider here only the elastic response of the medium surrounding the fault. For the first slip events in the sequences, the inclusion of more complex rheologies, such as damage, plasticity, or viscoelasticity, is not expected to have a large effect on the nucleation process because the deformations in the medium around the fault are not large yet. At later events there are two competing effects. On the one hand, they would decrease the stress concentrations at the tips of the fault and, in general, promote more homogeneous initial stresses on the fault, which would lead to a larger effect of roughness during the nucleation stage, as discussed above. On the other hand, the geometric barriers, which introduce substantial complexity into the nucleation process, are expected to be weaker with the inclusion of these rheologies.

We consider only faults with a length of 40 m in this study. For smooth faults with homogenous initial conditions, $l_{mc}$ generally decreases with increasing fault length. However, the rate of decrease is quite small for most initial and loading conditions, including those of this study, and declines with increasing fault length [Fang et al.,
2010]. Moreover, we perform simulations for smooth faults with lengths of 20 and 30 m and calculate nucleation lengths of $l_{nc} = 7.1$ and 6.1 m, respectively. Considering that the for 40 m long smooth fault, $l_{nc} = 5.6$, we do not expect a significant decrease in $l_{nc}$ for larger faults. For rough faults, there is no correlation between the nucleation lengths and whether the rupture initiates next to the center or near the end of the fault; thus we do not expect different values of $l_{nc}$ for larger fault lengths. A question arises: whether the slip events on roughest faults do not evolve into fast events just because the nucleation length is larger? If this is the case, fast slip events would occur also for faults with $b_r = 0.01$ and $\lambda_{\min} = 0.2$ m for larger faults. In the case of the first slip events in the sequences, the rupture does not reach the ends of the faults, thus fast slip events would not occur also on larger faults with these rupture parameters. The similar number of fast slip events for fault lengths of 30 m and 40 m in the case of faults with $b_r = 0.005$ and with $b_r = 0.01$ and $\lambda_{\min} = 1$ m (see chapter 3) may suggest that the later events would also not evolve into fast slip events for faults $b_r = 0.01$ and $\lambda_{\min} = 0.2$ m and larger fault lengths.

### 4.4.2 Detection of the nucleation stage at its relationship to the final size of the event

Near-source observations suggest that the seismic nucleation phase can be detected and that it is characterized by relatively small moment rate with irregular fluctuations, which are followed by quadratic growth in the moment rate as rupture begins to propagate [Ellsworth and Beroza, 1995, 1998]. Lapusta and Rice [2003] performed simulations of earthquake sequences in a 2-D antiplane framework and showed that large earthquakes may have irregular moment rate in the beginning of the dynamic stage because of the heterogeneous stress field caused by arrest of previous events.

The complexities in the nucleation process in the case of rough faults, such as irregular evolutions of the slip rates and the rate of expansion of the rupture, also result in fluctuations in the moment rate. Figure 4-9a shows curves of the evolution of the moment rate per unit length, $\dot{M}_{0.1d}$, in the slip events shown in Figures 4-3 and 4-4, except for the first slip event for $b_r = 0.005$. To show all slip events in a single plot, the curves are aligned such that they all begin when $\dot{M}_{0.1d}$ exceeds the value of $10^8$ N/s for the last time before it approaches the maximum value. At this scale, $\dot{M}_{0.1d}$ shows irregular behavior.
mostly for $b_r = 0.005$, with large variations that may be considered as sub-events. To examine the behavior near to the transition to the dynamic stage, in Figure 4-9b, the curves are aligned such that they all begin when $\dot{M}_{0,1d}$ exceeds value of $2 \times 10^6$ N/s for the last time before it approaches the maximum value. In the first slip events on a smooth fault and a fault with $b_r = 0.001$, $\dot{M}_{0,1d}$ increases monotonically up to the peak value, while in other events it exhibits irregular fluctuations, where the number of the fluctuations and their magnitudes generally increase with increasing $b_r$. In the two events on the fault with $b_r = 0.002$ and the second event on fault with $b_r = 0.005$, there is a relatively long stage of small and irregular $\dot{M}_{0,1d}$, which is followed by a stage of rapid increase toward the peak value. Note that the transition between these two stages occurs after the beginning of the dynamic stage of the rupture defined in this study.

The behavior of $\dot{M}_{0,1d}$ at the two stages somewhat resemble the seismic nucleation phase observed by Umeda [1990] and Ellsworth and Beroza [1995, 1998]. While they observe that the duration, source dimension, and average slip associated with the nucleation phase scale with the moment of the eventual earthquake, we generally observe that the final size of the slip event decreases as $\dot{M}_{0,1d}$ shows a longer stage of irregular behavior. However, we study the behavior of small faults with the largest slip event equivalent to a magnitude 1.5 earthquake, while they examine moderate to large earthquakes, in which the nucleation phases themselves are much larger than the slip events in our study. Nakatani et al. [2000] analyzed the velocity waveforms of 17 microearthquakes ($0.3 \leq M \leq 2.1$) in Japan and showed that microearthquakes that start with a stronger initial rupture tend to grow larger, which is consistent with the trends obtained in our study. It is important to note, however, that other studies do not show a consistent relationship between the beginning of the nucleation phase and the final size of the earthquakes [e.g. Mori and Kanamori, 1996].

The irregular behavior of $\dot{M}_{0,1d}$ for $b_r \geq 0.002$, as well as the increase in $l_c$ with $b_r$ at these $b_r$ values, increase the potential for detection of the nucleation phase, at least near the beginning of the dynamic stage, when the rupture is larger and the fluctuations are associated with larger deformations on the fault than at earlier stages. However, the lab-scale value of $L = 20 \mu m$ for the characteristic slip distance used in our study
leads to small nucleation zones, which are difficult to detect, even for the roughest faults. If the values of $L$ for natural faults are much larger than the values obtained in laboratory experiments [e.g. Scholz, 1988; Marone and Kilgore, 1993], the deformation during the nucleation should be large enough to be detected. However, it is not clear if the trends observed here for rough faults remain if similar roughness parameters are used, but with larger values of $L$.

4.5 Conclusions

We study numerically the effects of roughness on the nucleation of earthquakes on faults governed by rate and state friction, and subjected to slow loading. Our numerical approach accounts for all stages in the seismic cycle, and in each simulation we model a sequence of two earthquakes or more. This enables to study the effects of heterogeneities left by the arrest of a slip event on the nucleation process in a subsequent event.

The roughness introduces local barriers that complicate the nucleation process and result in asymmetric expansion of the rupture, stages where the rupture expands but the slip rates on the fault decrease, and the generation of new slip pulses, which rerupture regions that already slipped.

A significant effect is observed for whether the slip events are first or later in the earthquake sequence, with larger effects of the roughness for the first events, where the initial conditions are homogenous. For the first events, there is a large increase in the nucleation length with $b_r$ for $b_r \leq 0.002$, and a transition to aseismic or mostly aseismic deformation for larger $b_r$ values. Moreover, in the first events the ruptures always initiate where the local geometry of the fault is most preferable for slip. For slip events later in the sequence, the initial stress field and frictional conditions are determined by the rupture growth and arrest in previous slip events, which are themselves determined by the finiteness of the fault and the roughness. This leads to a trade off between the effects of the finiteness of the fault and the roughness. For $b_r \leq 0.002$, the effects of finiteness of the fault on the initial stresses is a more dominant factor in the nucleation process, the nucleation length barely changes with $b_r$, and most of the events initiate close to the ends of the fault. For larger values of $b_r$, the roughness seems to play a larger role, the
nucleation length increases with $b_r$, and location where the events initiate is more variable.

The complexities in the nucleation process are reflected as irregular fluctuations in the moment rate, especially for $b_r \geq 0.002$. The irregular behavior of $\dot{M}_{0,1d}$ and the increase in the nucleation length at these $b_r$ values increase the potential for detection of the later stages in the nucleation process, especially if the values of $L$ for natural faults are indeed larger than the values obtained in laboratory experiments.
Figure 4-1: (a) The problem set up: a 40 m long finite fault is embedded in a 2D elastic medium with dimensions 120 x 80, which is subjected a prescribed slow horizontal velocity $V_b = \pm 10^{-9}$ at the top and bottom, and initial stresses $\sigma_{xx0}$, $\sigma_{yy0}$, and $\sigma_{xy0}$. The other boundary and loading conditions are prescribed such that the remote horizontal and vertical stresses are spatially and permanently constant and the remote shear stress is spatially constant but increases gradually with time. The model includes an absorbing layer with gradual Rayleigh damping. (b) The fault profiles examined in this study. We consider a total of three different general geometries, and for each geometry, eight profiles are generated with roughness pre-factor values of $b_r = 0.001$, 0.002, 0.005 and 0.01 and minimum wavelength of $\lambda_{min} = 0.2$ m and 1 m. For reference, we also run a simulation with a smooth fault.
Figure 4-2: (a) The evolution of the $r_{sv}$ vs. time for the eight profiles of Geo-1, as well as for a smooth fault. (b) The evolution of the $r_{sv}$ vs. time during the first slip events in the sequences obtained for faults with $\lambda_{min} = 0.2$ m, and $b_r = 0.001$ (right) and 0.002 (left). (c) The evolution of the maximum slip rate on the fault vs. time for the eight profiles of Geo-1, as well as for smooth faults.
Figure 4-3: Profiles of $u$, $v$, and $\mu$ along the fault for a smooth fault (a, c, and e) and a rough fault with Geo-1, $b_r = 0.001$, and $\lambda_{min} = 0.2$ m (b, d, and f). The contours of $\mu$ are shown only for the first slip event in each sequence. The contours of $u$ are plotted for all the stages in the seismic cycle, i.e. the loading (dashed purple), nucleation (red), and dynamic (black) stages, where the dynamic stage includes both the propagation and arrest of the rupture. The final stage of each slip event is shown in blue. In the case of the first event of the sequence obtained for a smooth fault, the contours of $v$ and $\mu$ are plotted for the whole dynamic stage, with gray contours for the arrest phase. However, for better visualization of the nucleation stage, the contours of $v$ and $\mu$ are shown only for the nucleation stage and the beginning of the dynamic stage in other slip events.
Moreover, the contours of $\mu$ are shown only for the first event in each sequence and we plot only six contours (with different colors) for the nucleation stage.

Figure 4-4: Profiles of $u$, $v$, and $\mu$ along rough faults with Geo-1, $\lambda_{\text{min}} = 0.2$ m, and $b_r = 0.002$ (a, c, and e) and 0.005 (b, d, and f). The contours of $u$ are plotted for all the stages in the seismic cycle, i.e. the loading (dashed purple), nucleation (red), and dynamic (black) stages, where the dynamic stage includes both the propagation and arrest of the rupture. The final stage of each of the slip events is shown in blue. For $b_r = 0.005$, the contours of $v$ are shown only for the first and second slip events of the sequence. For better visualization of the nucleation stage, the contours of $v$ and $\mu$ are shown only for the nucleation stage and the beginning of the dynamic stage.
Moreover, the contours of $\mu$ are shown only for the first event in each sequence and we plot only six contours (with different colors) for the nucleation stage.

![Graph](image)

**Figure 4-5:** The nucleation lengths of the fast slip events in the 25 simulated earthquake sequences. Only slip events where the slip accumulated during the dynamic stage of the rupture is more than half of the total slip are considered. Slip events that are the first in the sequence are shown in blue, while the later events are shown in red, both with open circles for faults with $\lambda_{\text{min}} = 0.2$ m and '+' symbols for faults with $\lambda_{\text{min}} = 1$ m. The solid curves represent the average values for the first (blue) and later (red) slip events on faults with $\lambda_{\text{min}} = 0.2$ m (solid) and $\lambda_{\text{min}} = 0.2$ m (dashed).
Figure 4-6: Profiles of $u$ and $v$ for the loading (purple) and the beginning of the nucleation stage (red) during the first slip events in the sequences obtained for faults with Geo-2, $\lambda_{min} = 0.2$ m, and $b_r = 0.001$ (a and c) and 0.002 (b and d).
Figure 4-7: The locations where the fast slip events in the 25 simulated earthquake sequences initiate at the beginning of the nucleation stage. The plot shows the absolute distance of the locations from the center of the fault. Slip events that are first in the sequence are shown in blue, while the later events are shown in red, both with open circles for faults with $\lambda_{\text{min}} = 0.2$ m and '+' symbols for faults with $\lambda_{\text{min}} = 1$ m.
Figure 4-8: The evolution of the shear tractions $\tau$, $v$, $\theta$, and $\mu$ at six loading stages between the end of the first slip event and the beginning of the second slip event for a smooth fault and a fault with Geo-1, $b_r = 0.002$ and $\lambda_{\text{min}} = 1$ m. The six loading stages are shown on curves of the average shear stress on the fault vs. time.
Figure 4-9: (a) Curves of the evolution of $\dot{M}_{0,1d}$ during 18 s near their maximum values for the first (solid) and later (dashed) slip events in the earthquake sequences shown in Figures 4-3 and 4-4, except for the first slip event for $b_r = 0.005$. To show all slip events in a single plot, the curves begin at the time when the moment rate exceeds a value of $\dot{M}_{0,1d} = 10^6$ N/s for the last time before it approaches its maximum value. The black circles denote the transition between the nucleation and the dynamic stages, as is defined in this study. (b) The evolution of $\dot{M}_{0,1d}$ during 0.05 s near its maximum value for the same slip events. The curves are aligned such that they all begin when $\dot{M}_{0,1d}$ exceeds a value of $2 \times 10^6$ N/s for the last time before it approaches the maximum value. The black circles denote the transition between the nucleation and the dynamic stages.
Chapter 5

Conclusion and future directions

5.1 Conclusion

Field and laboratory observations show that faults are rough at all scales and can be described as self-affine fractals. This thesis examines the role of roughness in the seismic cycle numerically, focusing on the scale of small earthquakes, in which realistic geometry and friction law parameters can be incorporated more easily. The minimum roughness wavelength in the modeled faults is at the scale of lab samples, assuming that the friction law provided by laboratory observations already include the effects of smaller wavelengths.

To model all stages in the seismic cycle and to accurately model the effect of the roughness, I extend the 2-D large sliding mortar finite element formulation to dynamic problems in chapter 2 and implement slip-weakening (SW) and rate and state (RS) friction laws into the method. The method utilizes the dual Lagrange multipliers and the primal–dual active set strategy concepts and accordingly discretizes and linearizes the friction laws to obtain a semi-smooth Newton method. Moreover, the discretization of the RS friction law involves a procedure to condense out the state variables, thus eliminating the addition of another set of unknowns into the system. The method involves both quasi-static and dynamic time discretization schemes that are implicit; thus the implementation of a variable time stepping is straightforward, and the whole seismic cycle can be modeled, including a completely spontaneous nucleation process.

Several numerical examples are provided in chapter 2 in order to demonstrate the capabilities of the method. The method shows excellent convergence for the SW friction law with efficient detection between the slipping and sticking states of the nodes despite the large number of nodes on the fault. A good convergence is also obtained for the RS friction law, but because of its high nonlinearity and because it involves significant variations of the friction coefficient with small change in slip rate, more iterations are
needed before convergence. For both friction laws, the total energy in the system is exactly conserved with the non-damping Newmark scheme. The amount of energy dissipation in the damping schemes is quite small. It decreases with decreasing time step size and affects mostly the very high frequency variations in the shear traction and slip rate with time.

In chapter 3 and 4, the developed methodology is used to study the response of rough faults governed by rate and state friction to slow tectonic loading, where in each simulation, the earthquake sequence include at least two seismic cycles. Chapter 3 focuses on the global behavior of the faults. In general, as the roughness amplitude increases and the minimum wavelength of roughness decreases, there is a transition from seismic slip behavior to aseismic slip behavior, in which the load on the fault is released by more slip events but with lower slip rate, seismic moment, and average static stress drop. Moreover, these source parameters decrease mutually such that the linear relationship between the stress drop and the average fault strain is generally preserved. The results also suggest that the roughness may have a role in constant stress drop for small and large earthquakes shown in seismological observations.

In chapter 4, I study the effect of roughness on the nucleation process. The roughness introduces local barriers that complicate the nucleation process and result in asymmetric expansion of the rupture, stages where the rupture expands but the slip rates on the fault decrease, and the generation of new slip pulses, which rerupture regions that already slipped. These complexities reflected as irregular fluctuations in the moment rate and are discussed in the context of a possible detection of the later stages of the nucleation process in seismological observations.

There is large difference in the nucleation process between initial slip events in the sequences, where the initial conditions are homogenous, and the later events, where the initial stress field and friction conditions are determined by the rupture growth and arrest in previous slip events. In the initial events there is large increase in the nucleation length with increasing $b_r$ for $b_r \leq 0.002$, and transition to aseismic or mostly aseismic deformation for larger values, where $b_r$ is the roughness pre-factor. Moreover, the ruptures always initiate where the local geometry of the fault is most preferable for slip. For slip events later in the sequence there is a trade off between the effects of the
finiteness of the fault and the roughness. For $b_r \leq 0.002$, the effect of finiteness of the fault on the initial stresses is a more dominant factor in the nucleation process, thus the nucleation length barely changes with $b_r$ and most of the events initiate close to the ends of the fault. For larger values of $b_r$, the roughness seems to play a larger role, thus the nucleation length increases with $b_r$ and location where the events initiate is more sporadic.

5.2 Future directions

This thesis is an initial step in developing a realistic physics-based model of the seismic cycle that accounts for the complexities associated with the fault zone. I believe that further development of such a model requires the following components:

**Improving the numerical technique**

Currently the method can account for fault roughness, rate and state and slip-weakening friction laws, and variations of the lithologies and the frictional properties along the fault. The method should be extended to account for non-linear material laws for the medium surrounding the fault, such as the damage models of Lyakhovsky et al. [1997] and Bhat et al. [2012]. The method should include additional friction laws, which will be based on data from large slip experiments [e.g. Di Toro et al., 2011]. The computational efficiency of the method should be improved, especially if extension of the method to 3-D problems is considered. The method is currently implemented in Matlab. Although the code involves only matrix operations (except for the time step and iteration loops), its parallel performance is not good.

**Qualitative study on the influence of different fault zone properties on the earthquake cycle**

In this thesis I study the effect of roughness in the scale of small earthquakes. It is important to upscale the findings of this study to larger earthquakes and to study the effect of other fault zone complexities. To accomplish that, a systematic set of simulations should be performed in order to study the effects of the fault geometry, the rheology of fault zone rocks, and fault friction laws on the earthquake cycle and specifically on the transition from aseismic to seismic slip, the source parameters of
earthquakes, and shear rupture properties such as rupture velocity, nucleation length, high- and low-frequency radiation, and the transition between crack-like and the pulse-like rupture styles.

Constraints from observational data

Seismic and geodetic networks in proximity to fault zones provide detailed information regarding the temporal and spatial distribution of seismicity, the inter-, co- and post-seismic deformation, and the style of the rupture. The next step is to preform simulations with the specific fault zone structure and loading conditions of tectonically active regions as an input and examine if the main features in the seismic and geodetic observations can be reproduced and what is the role of different fault zone complexities in these features.

The methodology developed in this study can also be used to address the following fundamental questions in earthquake physics:

- What is a self-consistent physical model for the experimental observations of rock friction? For example, although the concept of rate and state friction was introduced to the earthquake community almost 40 years ago, there is no clear physical description of the state variable and how it evolves and relates to the small-scale asperities. With some adjustments, this problem can be addressed with the numerical approach developed in this thesis, which allows slip larger than the scale of roughness.

- How do fault geometry and off-fault deformation evolve during the seismic cycle at both experimental and natural scales? This can be addressed by enhancing the numerical approach with constitutive laws of wear processes as well as more complex rheologies.
References


Dieterich, J. H. (1992), Earthquake nucleation on faults with rate-and state-


Hager, C., S. Hüeber, and B. I. Wohlmuth (2008), A stable energy-conserving


Laursen, T., and V. Chawla (1997), Design of Energy Conserving Algorithms for
1996), 863–886, doi:10.1002/(SICI)1097-0207(19970315)40:5<863::AID-
NME92>3.0.CO;2-V.

Laursen, T. A., and X. N. Meng (2001), A new solution procedures for application of
energy-conserving algorithms to general constitutive models in nonlinear

Liu, F., and R. I. Borja (2009), An extended finite element framework for slow-rate

Lyakhovsky, V., and Y. Ben-Zion (2014), Damage-breakage rheology model and

Lyakhovsky, V., Y. Ben-Zion, and A. Agnon (1997), Distributed damage, faulting, and

Mair, K., and C. Marone (1999), Friction of simulated fault gouge for a wide range of
velocities and normal stresses, *J. Geophys. Res. Solid Earth*, 104(B12), 28899–
28914.

Marone, C. (1998), Laboratory-derived friction laws and their application to seismic

Marone, C., and B. Kilgore (1993), Scaling of the critical slip distance for seismic
faulting with shear strain in fault zones, *Nature*, 362(6421), 618–621,
doi:10.1038/362618a0.

McLaskey, G. C., and B. D. Kilgore (2013), Foreshocks during the nucleation of stick-
slip instability, *J. Geophys. Res. Solid Earth*, 118(6), 2982–2997,

Mori, J., and H. Kanamori (1996), Initial rupture of earthquakes in the 1995
Ridgecrest, California sequence, *Geophys. Res. Lett.*, 23(18), 2437–2440,

Nakatani, M., S. Kaneshima, and Y. Fukao (2000), Size-dependent microearthquake
initiation inferred from high-gain and low-noise observations at Nikko district,


Nielsen, S., J. Taddeucci, and S. Vinciguerra (2010), Experimental observation of
stick-slip instability fronts, *Geophys. J. Int.*, 180(2), 697–702,

Noda, H., N. Lapusta, and H. Kanamori (2013), heterogeneous stress change and


Di Toro, G., and G. Pennacchioni (2005), Fault plane processes and mesoscopic...


Umeda, Y. (1990), High-amplitude seismic waves radiated from the bright spot of an earthquake., 175, 81–92.


