Artificial Intelligence in Geosciences 2 (2021) 171–178

Contents lists available at ScienceDirect

Artificial Intelligence in Geosciences

journal homepage: www.keaipublishing.com/en/journals/artificial-intelligence-in-geosciences

Arriving at estimates of a rate and state fault friction model parameter using Bayesian inference and Markov chain Monte Carlo

Saumik Dana [a,*], Karthik Reddy Lyathakula [b]

[a] University of Southern California, Los Angeles, 90007, CA, USA
[b] North Carolina State University, Raleigh, 27607, NC, USA

ARTICLE INFO

Keywords:
Fault friction
Rate and state model
Critical slip distance
Bayesian inference
Markov chain Monte Carlo

ABSTRACT

The critical slip distance in rate and state model for fault friction in the study of potential earthquakes can vary wildly from micrometers to few me-ters depending on the length scale of the critically stressed fault. This makes it incredibly important to construct an inversion framework that provides good estimates of the critical slip distance purely based on the observed ac-celeration at the seismogram. To eventually construct a framework that takes noisy seismogram acceleration data as input and spits out robust estimates of critical slip distance as the output, we first present the performance of the framework for synthetic data. The framework is based on Bayesian inference and Markov chain Monte Carlo methods. The synthetic data is generated by adding noise to the acceleration output of spring-slider-damper idealization of the rate and state model as the forward model.

1. Introduction

Inversion of seismogram data to infer subsurface properties is one of the most critical aspects of geophysics research. In a nutshell, the activity at and around discontinuities called faults in the subsurface spawn off seismic waves which are recorded at the surface on a seismogram. Some of these recordings are too weak to enable any inference of subsurface properties, while some are too rich in terms in seismic signatures. The field of waveform inversion deals with the aspects of inference for the latter. The core set of mathematical operations for waveform inversion are going back and forth between the time domain and the frequency domain. To the naked eye, these operations and algorithms to implement them can get a little complicated to the community of researchers who do not have a firm background in inversion geophysics. Over the past few years, Bayesian methods to infer properties from seismogram data have become popular because of the advent of high performance computing and the ability to construct reduced order models of the high fidelity physics using deep learning. At the heart of the Bayesian inversion framework lies the feasibility to perform a lot of fast forward simulations to eventually arrive at robust estimates of the model parameter. Before stepping into any inversion framework, two questions are critical (1) What is the forward model? and (2) What are we infering? In the geophysics realm, the connection between the physics of what happens around the fault and what happens at the fault is achieved using the rate- and state- model for friction evolution, which is considered the gold standard for modeling earthquake cycles on faults (Ruina, 1983; Scholz, 1989; Marone, 1998; Meng and Shi, 2021; Jia et al., 2021; Zhu et al., 2020)

$$\mu = \mu_0 + A\ln\left(\frac{V}{V_0}\right) + B\ln\left(\frac{V_0\theta}{d_c}\right), \qquad \dot{\theta} = 1 - \frac{\theta V}{d_c}, \tag{1}$$

where $V = |\dot{\boldsymbol{d}}|$ is the slip rate magnitude, $a = \dot{V}$ which we hypothesize is of the same order as recorded by seismograph, μ_0 is the steady-state friction coefficient at the reference slip rate V_0, A and B are empirical dimensionless constants, θ is the macroscopic variable characterizing state of the surface and d_c is a critical slip distance over which a fault loses or regains its frictional strength after a perturbation in the loading conditions (Palmer and Rice, 1591). In principle, it determines the maximum slip acceleration and radiated energy to such an extent that it influences the magnitude and time scale of the associated stress breakdown process (Scholz, 2019). Regardless of the importance, it is paradoxical that the values of d_c reported in the literature range from a few to tens of microns as determined in typical laboratory experiments

* Corresponding author.
E-mail address: saumik@utexas.edu (S. Dana).

https://doi.org/10.1016/j.aiig.2022.02.003
Received 8 December 2021; Received in revised form 25 January 2022; Accepted 23 February 2022
Available online 16 March 2022

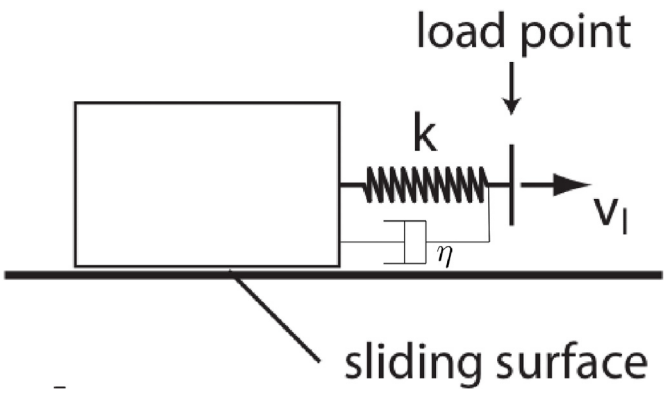

Fig. 1. Spring slider damper idealization of fault behavior. V_l is velocity of load point and V is velocity of the block against the slider surface.

(Scholz, 2019), to 0.1 − 5 m as determined in numerical and seismological estimates based on geophysical observations (Kaneko et al., 2017), and further to several meters as determined in high velocity laboratory experiments (Niemeijer et al., 2011). Understanding the physics that controls the critical slip distance and explains the gap between observations from experimental and natural faults is thus one of the crucial problems in both the seismology and laboratory communities (Ohnaka, 2003). With that in mind, we provide a framework in which synthetic earthquake data is used to quantify uncertainty in critical slip distance. While the resolution and coupled flow and poromechanics (Dana and Wheeler, 2018a; Dana and Wheeler, 2018b; Dana et al., 2018; Dana et al., 2021; DanaJoel Ita and Wheeler, 2020; Dana et al., 2020; Dana, 2018; Dana and Reddy, 2021; Dana and Reddy Lyathakula, 2021) associated with subsurface activity in the realm of energy technologies and concomitant earthquake quantification is a hot topic, in this work, we focus on the effect of a standard trigonometric perturbation with exponentially decreasing amplitude. In section 2, we explain the spring slider damper idealization to infer the influence of critical slip distance on RSF without recourse to complicated elastodynamic equations. In section 3, we explain the Bayesian inference framework to inversely quantify uncertainty in the estimation of critical slip distance. In section 4, we present the results from our investigation. In section 5, we present conclusions and outlook for future work.

2. Forward model for rate and state friction

We first rewrite Eq. (1) as

$$V = V_0 \exp\left(\frac{1}{A}\left(\mu - \mu_0 - B\ln\left(\frac{V_0\theta}{d_c}\right)\right)\right), \qquad \dot{\theta} = 1 - \frac{\theta V}{d_c} \tag{2}$$

As shown in Fig. 1, we model a fault by a slider spring system (Rice and Gu, 1983; Gu et al., 1984; Dieterich, 1992). The friction coefficient of the block is given by

$$\mu = \frac{\tau}{\sigma} = \frac{\tau_l - k\delta - \eta V}{\sigma} \tag{3}$$

where σ is the normal stress, τ the shear stress on the interface, τ_l is the remotely applied stress acting on the fault in the absence of slip, $-k\delta$ is the stress relaxation due to fault slip and η is the radiation damping coefficient (Kanamori and E Brodsky, 2004; McClure and Horne, 2011). We consider the case of a constant stressing rate $\dot{\tau}_l = kV_l$ where V_l is the load point velocity. The stiffness is a function of the fault length l and elastic modulus E as $k \approx \frac{E}{l}$. With $k' = \frac{E}{l\sigma}$, from Eqs. (3) and (2), we get the time derivative of friction coefficient as

$$\begin{aligned}\dot{\mu} &\approx k'(V_l - V) - k''\dot{V} \\ &= k'\left(V_l - V_0\exp\left(\frac{1}{A}\left(\mu - \mu_0 - B\ln\left(\frac{V_0\theta}{d_c}\right)\right)\right)\right) - k''\dot{V}\end{aligned} \tag{4}$$

where $k'' = \frac{\eta}{\sigma}$. We then get acceleration time series as $a \equiv \dot{V}$ as,

$$\begin{aligned}\dot{V} &= \frac{V}{A}\left(\dot{\mu} - \frac{B}{\theta}\dot{\theta}\right) \\ &= \frac{V_0}{A}\exp\left(\frac{1}{A}\left(\mu - \mu_0 - B\ln\left(\frac{V_0\theta}{d_c}\right)\right)\right) \times \left(\dot{\mu} - \frac{B}{\theta}\left(1 - \frac{\theta V}{d_c}\right)\right)\end{aligned} \tag{5}$$

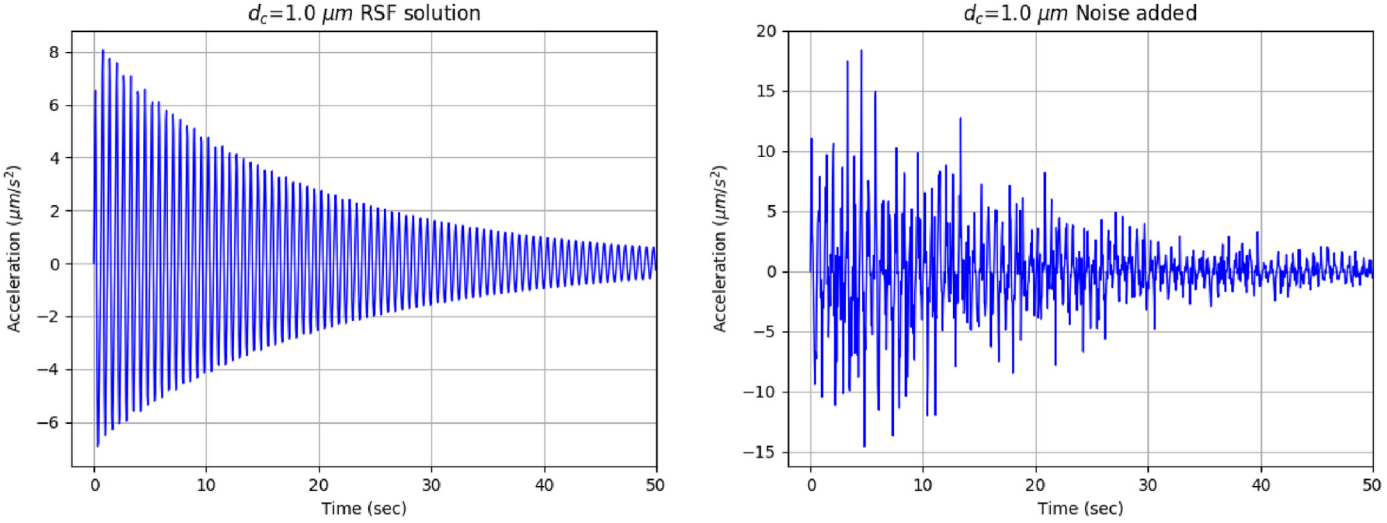

Fig. 2. System response. Before feeding the data to the Bayesian inference framework, we add synthetic noise to the data, because we would always expect some noise to come with the data in a real scenario, and we would want to test the robustness of the inference framework to provide good estimates in such a realistic scenario.

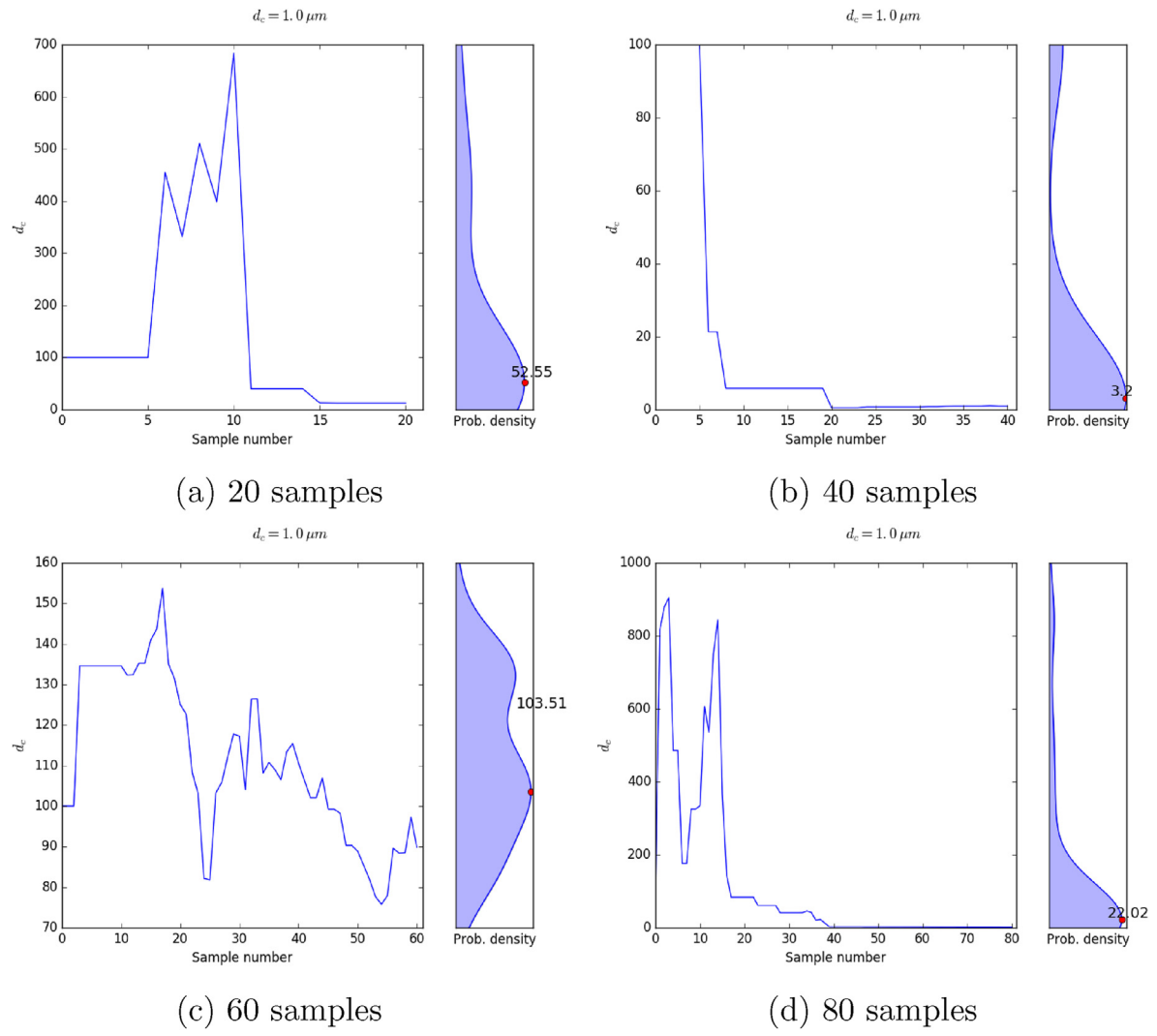

Fig. 3. Output of the Bayesian MCMC framework for model parameter for different sample numbers with true value of $d_c = 1$ μm and an initial guess of 100 μm.

The ballpark values are: $E = 5 \times 10^{10}\ Pa$, $l = 3 \times 10^{-2}\ m$, $\sigma = 200 \times 10^6\ Pa$, $\eta = 20 \times 10^6\ Pa/(m/s)$, $A = 0.011$, $B = 0.014$, $V_0 = 1$ μm/s, $\theta_0 = 0.6$, $\mu_0 = \mu_0 = 0.6$, $t_{start} = 0$, $t_{end} = 50\ s$, $dt = 0.1\ s$, from which the effective stiffness and damping are obtained as $k' \approx 10^{-2}(\mu m)^{-1}$ and $k'' = 10^{-7}\ s/\mu m$. The influence of critical slip distance on system response to a load point perturbation of the form

$$V_l = V_0(1 + \exp(-t/20)\, sin(10t)) \tag{6}$$

is shown in Fig. 2. Before feeding the data to the Bayesian inference framework, we add synthetic noise to the data, because we would always expect some noise to come with the data in a real scenario, and we would want to test the robustness of the inference framework to provide good estimates in such a realistic scenario. The reason for choosing V_l to be varying in time as a sinusoid with an exponentially decreasing magnitude to best replicate the quality of the seismogram velocity reading in V.

3. The formalism of Bayesian inference with Markov chain Monte Carlo sampling

The Bayesian inference framework works on the basic tent of uncovering a distribution centered around the true value and starts off with an initial guess for the distribution also called "prior" $\mathcal{D}$ to eventually get to the most accurate distribution possible also called "posterior" $\mathcal{P}$ through a likelihood $\mathcal{L}$. The Bayes theorem in a nutshell is:

$$\mathcal{P} = \frac{\mathcal{D} \times \mathcal{L}}{\int \mathcal{D} \times \mathcal{L}} \tag{7}$$

The prior is typically taken to be a Gaussian distribution and the likelihood carries information about the forward model. The quantity that makes evaluation of the posterior difficult is the integral term in the denominator. Since direct evaluation of the integral using quadrature rules is expensive, sampling methods like Markov chain Monte Carlo (MCMC) (Shapiro, 2003; Keith Hastings, 1970; Haario et al., 2001; Mueller, 2010) are used. To put it mathematically, if we were evaluating an integral, then the sampling would apply to points at which we know the value of integrand, and then proceed to evaluate the integral. But if the integrand at each of those points is a distribution rather than a value, it makes the sampling and subsequent averaging significantly more complicated. By constructing a Markov chain that has the desired distribution as its equilibrium distribution,

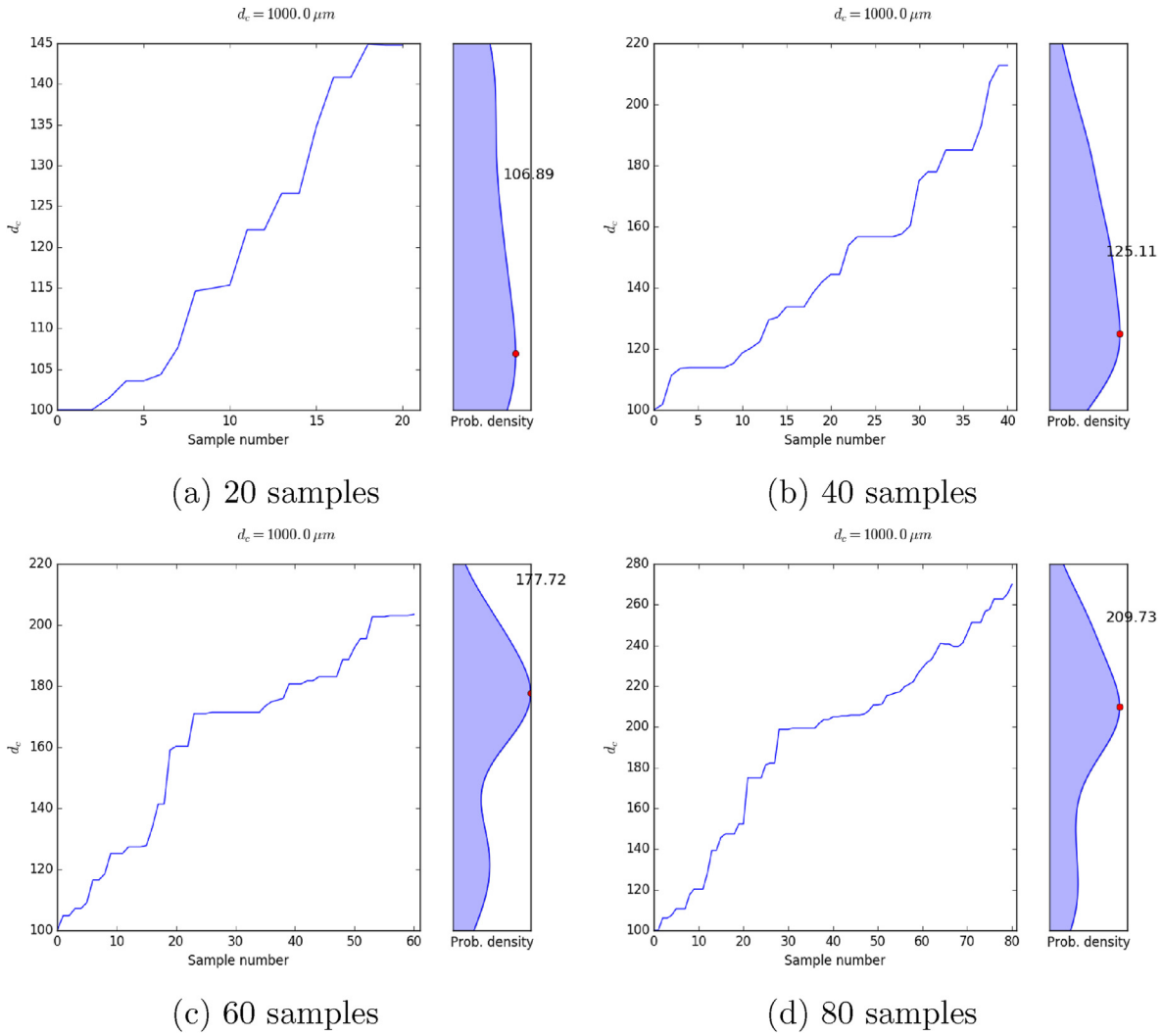

Fig. 4. Output of the Bayesian MCMC framework for model parameter for different sample numbers with true value of $d_c = 1000$ μm and an initial guess of 100 μm.

one can obtain a sample of the desired distribution by recording states from the chain. The more steps are included, the more closely the distribution of the sample matches the actual desired distribution.

3.1. Applied to our problem

In this particular inverse problem, the acceleration response of the model is known and the goal is to find the critical slip distance parameter d_c. To formalize the problem, consider the relationship between acceleration $a(t)$ and the model response by the following statistical model

$$a(t) = f(t, \theta, \mu, A, B, d_c) + \varepsilon \tag{8}$$

where ε is the noise. Assuming the $\varepsilon \sim N(0, \sigma^2)$ as unbiased, independent and identical normal distribution with standard deviation σ allows us to conveniently generate the synthetic data as shown in Fig. 2. The goal of the inverse problem is to determine the model parameter distribution as follows

$$\pi(d_c|a(t_1), \ldots, a(t_n)) = \frac{\pi(a(t_1), \ldots, a(t_n)|d_c)\pi_0(d_c)}{\int_{d_c} \pi(a(t_1), \ldots, a(t_n)|d_c)\pi_0(d_c)dd_c} \tag{9}$$

where $\pi_0(d_c)$ is the prior distribution and $\pi(a(t_1), \ldots, a(t_n)|d_c)$ is the likelihood given by

$$\pi(a(t_1), \ldots, a(t_n)|d_c) = \prod_{i=1}^{n} \pi(a(t_i)|d_c) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{a(t_i) - f(V, t_i, \theta, \mu, A, B, d_c)}{\sigma}\right)^2} \tag{10}$$

3.2. Adaptive metropolis algorithm

The adaptive Metropolis algorithm (Haario et al., 2001) explores the parameter space with specified limits ranging from a high of d_c^u to a low of d_c^l and starts from a random initial guess of the model parameter d_c^m, where m is the iteration number. The initial covariance matrix in the adaptive Metropolis algorithm is constructed using the initial parameter

$d_c^{m=0}$. At each iteration, the steps are

- A random parameter sample d_c^* is generated from the proposal distribution
- If d_c^* is not within the specified limits, $d_c^* \notin (d_c^l, d_c^u)$, the iteration is passed without moving to the next steps, and the previous sample is considered as the new sample, $d_c^{m+1} = d_c^m$
- If d_c^* is within the specified limits, $d_c^* \in (d_c^l, d_c^u)$, a new value of standard deviation associated with d_c^* is generated using the inverse-gamma distribution
- d_c^* is accepted as the new sample $d_c^{m+1} = d_c^*$ if a criterion which involves the standard devation is met
- The covariance matrix is updated if the iteration number is an exact multiple of m_0 using the previous m_0 model parameters (Smith, 2013)

The simulation is repeated for n iterations and the parameter samples resulting from all these iterations represent the parameter posterior distribution. The construction of the d_c^* is only based on the current parameter, d_c^m, which is the Markov process. The computational time of the MCMC sampling method is proportional to the number of generated samples n. To put it more succinctly, the algorithm is elucidated in Algorithm 1.

4. Results

Figs. 3 and 4 show a gradual improvement in the estimated model parameter as the number of samples is increased. It is evident that the framework has a lot more difficulty converging to the true value if the initial guess is far away from the true value. In reality, it is not really known what the true value is, and arbitrary initial guesses will lend a solution that is way off target regardless of the number of sampling points. We observe from Fig. 3 that the response is not necessarily monotonic, in the sense that the estimated value with the maximum probability does not necessarily monotonically converge to the true value as the number of sampling points is increased. In all of the simulations, we observe in all the simulations that the Bayesian framework does an excellent job of dropping to the ballpark of the expected true value within the first few samples.

4.1. With burn-in

As an MCMC algorithm is rarely initialized from its invariant distribution, there might be some concern that its initial values might bias results even if it does approach this equilibrium distribution later on. To compensate for this, a burn-in period is often implemented: the first N samples being discarded, with N being chosen to be large enough that the chain has reached its stationary regime by this time. We take that burn-in

Algorithm 1: Metropolis-Hastings algorithm

```
m ← 0;
d_c^m ← d_c^0;                                   // d_c^0 is an initial guess
V ~ d_c^m;                                       // Construct covariance matrix
for j=1,2,...,n do
    d_c^* ~ d_c^m, V;                            // Randomly sample from a distribution with
                                                    covariance V
    if d_c^* ∉ (d_c^l, d_c^u) then
        continue;                                // Move on if the guess is off the specified limits
    else
        γ ~ d_c^*;                               // Get standard deviation
        if f(γ, d_c^*) then
            d_c^{m+1} ← d_c^*;                   // Accept sample based on a condition f(γ, d_c^*)
                                                    being satisfied
        else
            d_c^{m+1} ← d_c^m;                   // Reject sample
    m ← m + 1;
    if m%m_0 then
        V ~ {d_c^m, d_c^{m-1}, .., d_c^{m-m_0}};  // Update covariance matrix every
                                                    m_0 iterations
```

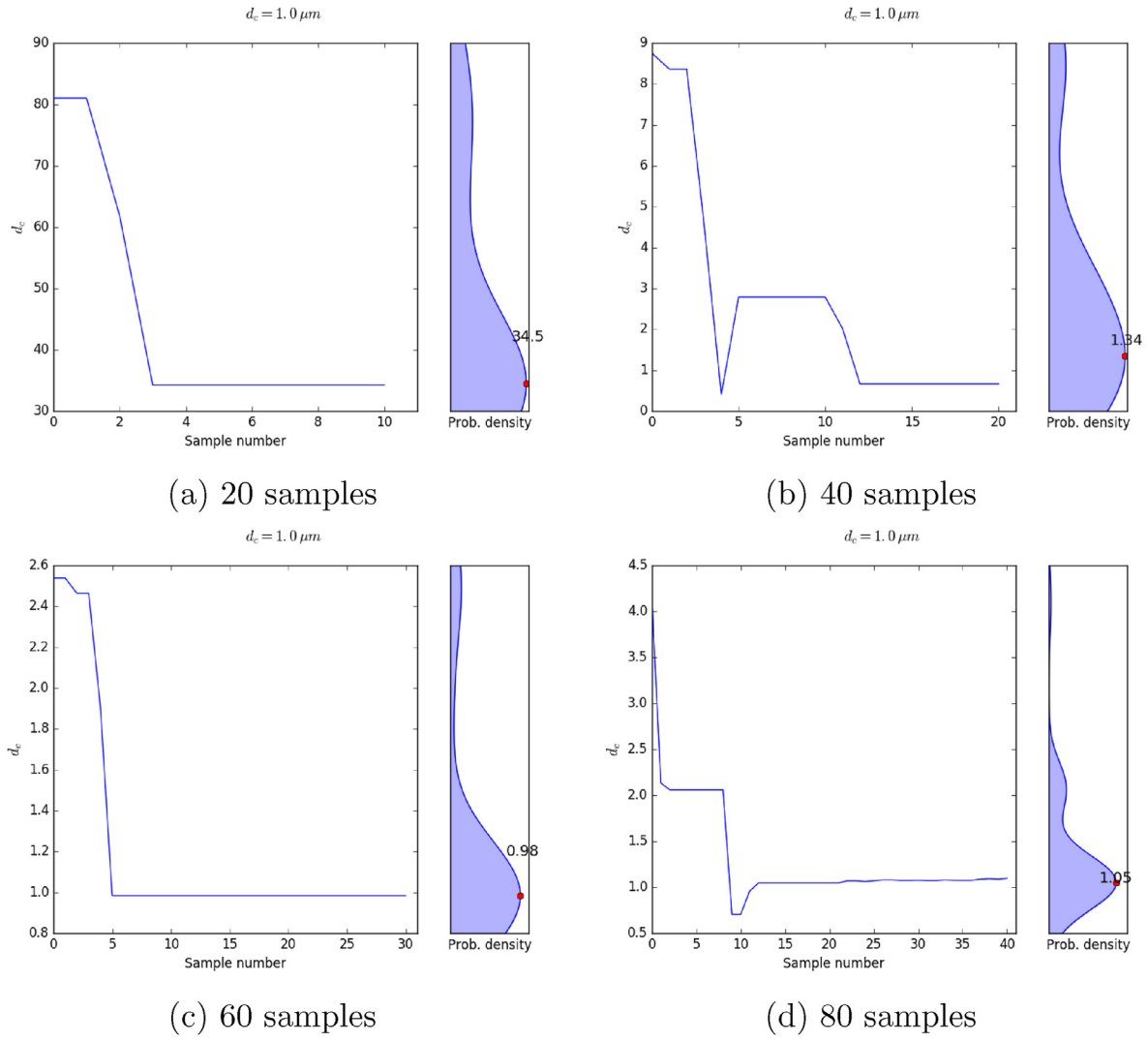

Fig. 5. Output of the Bayesian MCMC framework for model parameter for different sample numbers with true value of $d_c = 1$ μm and an initial guess of 100 μm with half the number of samples burnt-in.

number to be half the number of samples. A comparison between Figs. 3 and 5 as well as Figs. 4 and 6 clearly demonstrates the improved accuracy of the estimate when a certain number of samples are burnt-in.

5. Conclusions and outlook

The rate and state model for fault friction evolution is a piece of the puzzle in which forward simulations are used to arrive at the seismic impact of fault slip on earthquake activity. Typically, the earthquakes are measured with seismograms and geophones on the surface as P-waves and S-waves, and then these readings are used to calibrate the seismic activity for constant monitoring. The acceleration field around the fault slip activity translates to these waves recorded on the surface, and forward simulations with wave propagation bridge that gap. That being said, a seismic recording on the surface cannot easily be backtraced to acceleration around the fault, and finally to the source of the fault slip. That is precisely inverse modeling, and the Bayesian framework coupled with MCMC allows us to put such a framework in place. The accelerations are field quantities, and the inverse estimation of the field around the fault from the time series at the seismogram and/or geophone is not a trivial task. With that in mind, we test the robustness of the Bayesian/MCMC framework to inversely estimate a value instead of a field. The thing is that any inversion framework works on data as the input, and since subsurface data is not available other the sensors at the wells, the forward simulations are used to generate this data with all the computational physics put in place. This generated data is then used to inversely estimate the acceleration field from the data at the surface. Although such forward simulations to generate data are infeasible in real-time scenarios where we need estimates of what is happening in the subsurface from the reading on the surface almost immediately, the framework robustness would eventually lend itself to that scenario. The scenario is that the recording at the seismogram and/or geophone would be fed into the Bayesian/MCMC fraemwork as an input, and the framework would provide an estimate of the acceleration field around the fault as the output. In this work, we work on arriving at estimates of critical slip distance in the rate and state model by running a spring-slider-damper as the forward model instead of using a full-fledged forward simulator. In the future, we will be deploying the coupled flow and geomechanics simulator as the forward model, and gradually arrive at the aforementioned scenario.

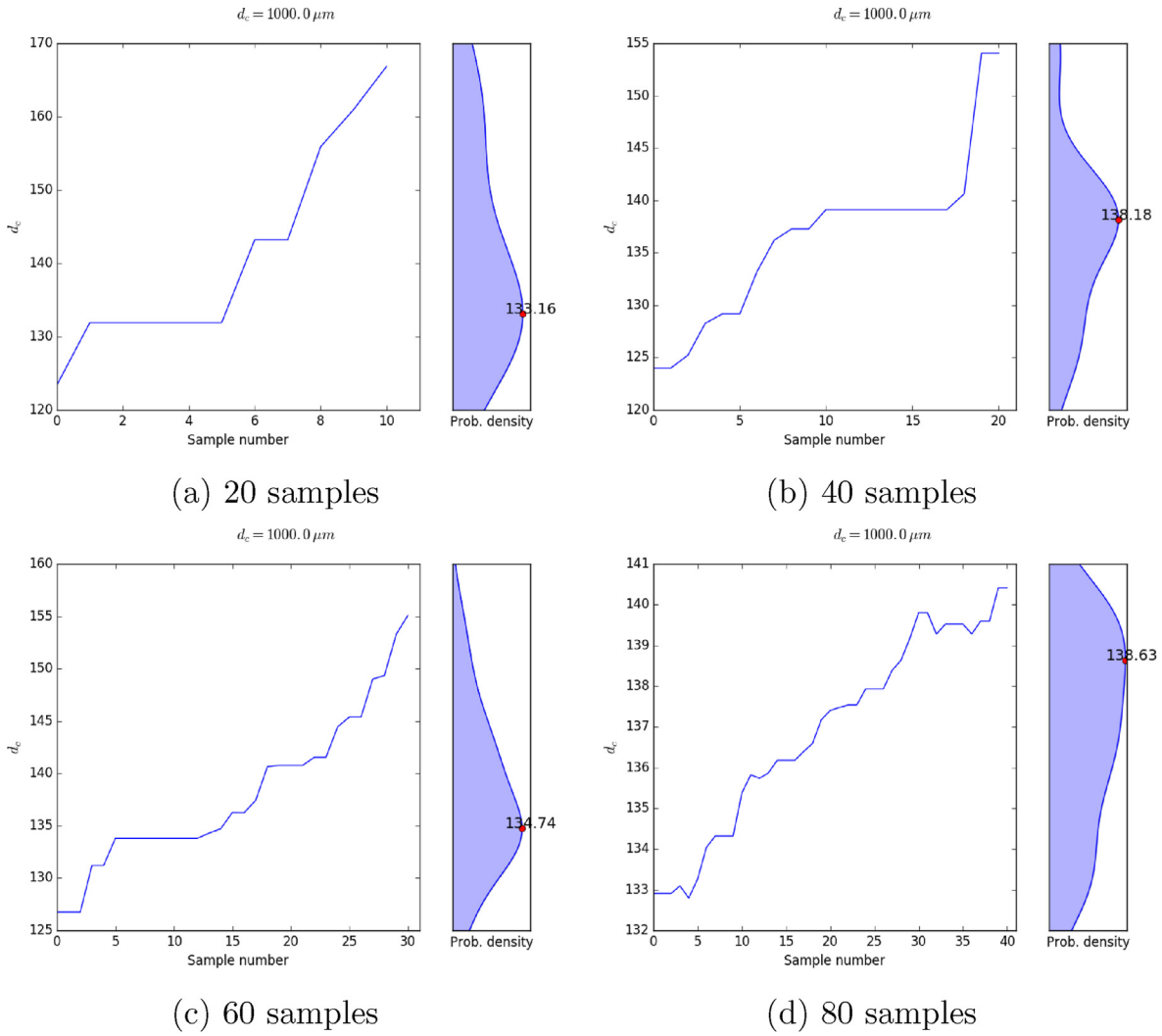

Fig. 6. Output of the Bayesian MCMC framework for model parameter for different sample numbers with true value of $d_c = 1000$ μm and an initial guess of 100 μm with half the number of samples burnt-in.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Dana, S., 2018. Addressing Challenges in Modeling of Coupled Flow and Poromechanics in Deep Subsurface Reservoirs. PhD thesis. The University of Texas at Austin.

Dana, Saumik, Reddy, Karthik, 2021. Bayesian Inference and Markov Chain Monte Carlo Based Estimation of a Geoscience Model Parameter. *arXiv preprint arXiv:2201.01868*.

Dana, Saumik, Reddy Lyathakula, Karthik, 2021. Uncertainty quantification in friction model for earthquakes using bayesian inference. *arXiv preprint arXiv:2104.11156*.

Dana, S., Wheeler, M.F., 2018a. Convergence analysis of fixed stress split iterative scheme for anisotropic poroelasticity with tensor biot parameter. Comput. Geosci. 22 (5), 1219–1230.

Dana, S., Wheeler, M.F., 2018b. Convergence analysis of two-grid fixed stress split iterative scheme for coupled flow and deformation in heterogeneous poroelastic media. Comput. Methods Appl. Mech. Eng. 341, 788–806.

Dana, Saumik, Ganis, Benjamin, Wheeler, Mary F., 2018. A multiscale fixed stress split iterative scheme for coupled flow and poromechanics in deep subsurface reservoirs. J. Comput. Phys. 352 (1–22).

Dana, Saumik, Srinivasan, Shriram, Karra, Satish, Makedonska, Nataliia, Hyman, Jeffrey D., O'Malley, Daniel, Viswanathan, Hari, Srinivasan, Gowri, 2020. Towards real-time forecasting of natural gas production by harnessing graph theory for stochastic discrete fracture networks. J. Petrol. Sci. Eng. 195, 107791.

Dana, Saumik, Jammoul, Mohamad, Wheeler, Mary F., 2021. Performance studies of the fixed stress split algorithm for immiscible two-phase flow coupled with linear poromechanics. Comput. Geosci. 1–15.

Dana, Saumik, Joel Ita, Wheeler, Mary F., 2020. The correspondence between voigt and reuss bounds and the decoupling constraint in a two-grid staggered algorithm for consolidation in heterogeneous porous media. Multiscale Model. Simul. 18 (1), 221–239.

Dieterich, James H., 1992. Earthquake nucleation on faults with rate-and state-dependent strength. Tectonophysics 211 (1–4), 115–134.

Gu, Ji-Cheng, Rice, James R., Ruina, Andy L., Tse Simon, T., 1984. Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. J. Mech. Phys. Solid. 32 (3), 167–196.

Haario, Heikki, Saksman, Eero, Tamminen, Johanna, 2001. An Adaptive Metropolis Algorithm. Bernoulli, pp. 223–242.

Jia, Yunzhong, Tang, Jiren, Lu, Yiyu, Lu, Zhaohui, 2021. The effect of fluid pressure on frictional stability transition from velocity strengthening to velocity weakening and critical slip distance evolution in shale reservoirs. Geomechanics and Geophysics for Geo-Energy and Geo-Resources 7 (1), 1–13.

Kanamori, Hiroo, E Brodsky, Emily, 2004. The physics of earthquakes. Rep. Prog. Phys. 67 (8), 1429.

Kaneko, Yoshihiro, Fukuyama, Eiichi, James Hamling, Ian, 2017. Slip-weakening distance and energy budget inferred from near-fault ground deformation during the 2016 mw7. 8 kaikōura earthquake. Geophys. Res. Lett. 44 (10), 4765–4773.

Keith Hastings, W., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications.

Marone, C., 1998. Laboratory-derived friction laws and their application to seismic faulting. Annu. Rev. Earth Planet Sci. 26, 643–696.

McClure, Mark W., Horne, Roland N., 2011. Investigation of injection-induced seismicity using a coupled fluid flow and rate/state friction model. Geophysics 76 (6), WC181–WC198.

Meng, Wei, Shi, Pengcheng, 2021. Synchronization of earthquake cycles of adjacent segments on oceanic transform faults revealed by numerical simulation in the framework of rate-and-state friction. J. Geophys. Res. Solid Earth 126 (1), e2020JB020231.
Mueller, Christian Lorenz, 2010. Exploring the common concepts of adaptive mcmc and covariance matrix adaptation schemes. In: *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum fÃ1/4r Informatik.
Niemeijer, André, Di Toro, Giulio, Nielsen, Stefan, Di Felice, Fabio, 2011. Frictional melting of gabbro under extreme experimental conditions of normal stress, acceleration, and sliding velocity. J. Geophys. Res. Solid Earth 116 (B7).
Ohnaka, Mitiyasu, 2003. A constitutive scaling law and a unified comprehension for frictional slip failure, shear fracture of intact rock, and earthquake rupture. J. Geophys. Res. Solid Earth 108 (B2).
Palmer, Andrew Clennel, Rice, James Robert, 1591. The growth of slip surfaces in the progressive failure of over-consolidated clay. Proc. Roy. Soc. Lond. A. Math. Phys. Sci. 332, 527–548, 1973.
Rice, James R., Gu, Ji-cheng, 1983. Earthquake aftereffects and triggered seismic phenomena. Pure Appl. Geophys. 121 (2), 187–219.
Ruina, A.L., 1983. Slip instability and state variable friction laws. Geophys. Res. Lett. 88, 359–370.
Scholz, C.H., 1989. Mechanics of faulting. Annu. Rev. Earth Planet Sci. 17, 309–334.
Scholz, Christopher H., 2019. The Mechanics of Earthquakes and Faulting. Cambridge university press.
Shapiro, Alexander, 2003. Monte Carlo sampling methods. Handb. Oper. Res. Manag. Sci. 10, 353–425.
Smith, Ralph C., 2013. Uncertainty Quantification: Theory, Implementation, and Applications, vol. 12. Siam.
Zhu, Weiqiang, Allison, Kali, Dunham, Eric, Yang, Yuyun, 09 2020. Fault valving and pore pressure evolution in simulations of earthquake sequences and aseismic slip. Nat. Commun. 11, 4833.