AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL075925

Key Points:
- We use a coupled flow-geomechanics model of injection-induced seismicity in a Bayesian inversion framework
- We perform probabilistic inference of dynamic friction properties of faults
- We identify stress transfer mechanisms from the interplay between poroelastic expansion and fault slip

Supporting Information:
- Supporting Information S1

Correspondence to:
J. Jagalur-Mohan,
jagalur@mit.edu

Citation:
Jagalur-Mohan, J., Jha, B., Wang, Z., Juanes, R., & Marzouk, Y. (2018). Inferring fault frictional and reservoir hydraulic properties from injection-induced seismicity. *Geophysical Research Letters*, *45*. https://doi.org/10.1002/2017GL075925

Received 4 OCT 2017
Accepted 21 DEC 2017
Accepted article online 10 JAN 2018

©2018. American Geophysical Union. All Rights Reserved.

Inferring Fault Frictional and Reservoir Hydraulic Properties From Injection-Induced Seismicity

Jayanth Jagalur-Mohan[1], Birendra Jha[2], Zheng Wang[1], Ruben Juanes[3,4], and Youssef Marzouk[1]

[1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA, [2]Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, CA, USA, [3]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA, [4]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract Characterizing the rheological properties of faults and the evolution of fault friction during seismic slip are fundamental problems in geology and seismology. Recent increases in the frequency of induced earthquakes have intensified the need for robust methods to estimate fault properties. Here we present a novel approach for estimation of aquifer and fault properties, which combines coupled multiphysics simulation of injection-induced seismicity with adaptive surrogate-based Bayesian inversion. In a synthetic 2-D model, we use aquifer pressure, ground displacements, and fault slip measurements during fluid injection to estimate the dynamic fault friction, the critical slip distance, and the aquifer permeability. Our forward model allows us to observe nonmonotonic evolutions of shear traction and slip on the fault resulting from the interplay of several physical mechanisms, including injection-induced aquifer expansion, stress transfer along the fault, and slip-induced stress relaxation. This interplay provides the basis for a successful joint inversion of induced seismicity, yielding well-informed Bayesian posterior distributions of dynamic friction and critical slip. We uncover an inverse relationship between dynamic friction and critical slip distance, which is in agreement with the small dynamic friction and large critical slip reported during seismicity on mature faults.

1. Introduction

Faults are key to understanding earthquakes, both natural and anthropogenic (Scholz, 1989; Keranen et al., 2014). Yet fault properties and processes are highly uncertain due to the complexity of physical mechanisms controlling fault behavior at depth. Given the importance of fault friction and its evolution with fault slip during an earthquake (Kanamori & Brodsky, 2004; Scholz, 2002), various experimental (Dieterich, 1979; Fulton et al., 2013; Marone, 1998; Ohnaka & Shen, 1999; Scholz, 1989; Toro et al., 2011), theoretical (Campillo & Ionescu, 1997; Rice, 1993; Rubin & Ampuero, 2005; Ruina, 1983; Scholz, 1988), and numerical (Cocco & Bizzarri, 2002) methods have been proposed in seismology and geodynamics to study fault friction. The critical slip distance and the dynamic coefficient of friction have been identified as the two key rheological parameters of the slip-weakening behavior of faults observed during seismic slip (Dieterich, 1979; Ruina, 1983; Scholz, 1988). The values of these two parameters in the field vary widely and are known with little certainty (Fulton et al., 2013; Marone, 1998). Recently, it has been proposed that rigorous modeling of coupled fluid flow and geomechanics, guided by field measurements, can be used to characterize induced seismicity and to forecast the associated seismic hazard (Juanes et al., 2016). In this letter, we demonstrate by means of a synthetic example that a Bayesian inference framework incorporating a coupled flow and geomechanics model of induced seismicity can successfully deduce the critical slip distance and dynamic friction coefficient of a seismogenic fault.

A central challenge of Bayesian inference is the efficient characterization of the *posterior* probability distribution of parameters or properties of interest. This distribution carries more information than the point estimates provided by classical deterministic inversion techniques: from the posterior distribution, one can fully characterize uncertainties and interrelationships among the model parameters and use these uncertainties to make probabilistic predictions (Sivia & Skilling, 2006; Tarantola, 2005). This information is particularly important when observational data are noisy, incomplete, and indirectly related to the quantities of interest,

as in the present problem. Analysis of the posterior naturally quantifies the values of different types of data (e.g., pressure, ground displacement, and fault slip) by describing how they reduce prior uncertainty, individually and in combination. Yet standard methods for Bayesian inference—for instance, Markov chain Monte Carlo (MCMC) sampling—are computationally prohibitive for our problem, as they require repeated evaluations of the coupled flow and geomechanics model; here even a single model evaluation is computationally intensive. Moreover, a particularly large number of samples are required to resolve non-Gaussian features of the posterior, which result from the nonlinear response of the fault during fluid injection. To overcome this computational barrier, we develop principled and convergent approximations of the parameter-to-observable relationships induced by the coupled flow and geomechanics model. These approximations, collectively termed a "surrogate" model, can accelerate posterior sampling by several orders of magnitude (Marzouk et al., 2007; Marzouk & Xiu, 2009). We build our surrogate using state-of-the-art sparse approximation techniques (Conrad & Marzouk, 2013) that *adaptively* explore the response of the forward model, yielding accurate inversion results with a small number of high-fidelity model runs.

2. Coupled Flow and Geomechanical Model of Injection-Induced Seismicity

We consider a 2-D plane-strain poroelastic domain with a confined aquifer and a fault (Figure 1). The aquifer properties are homogeneous and isotropic, with an aquifer permeability k much higher than that of the overburden and underburden. The fault permeability is 10^{-6} darcy and follows a linear slip-weakening law for the evolution of its friction coefficient μ_f from the static friction coefficient μ_s to the dynamic friction coefficient μ_d as a function of slip with a critical slip distance D_c. The law yields $\mu_f = \mu_d$ when slip exceeds D_c. We fix $\mu_s = 0.45$ and assume that the fault permeability remains unchanged upon fault slip.

A normal faulting stress regime is prescribed through horizontal compression on the right boundary that is 0.7 times the vertical overburden stress, that is, $\sigma_{\mathrm{h}} = 0.7\sigma_{\mathrm{v}}$. The normal faulting boundary condition favors downdip slip on the fault in which the hanging wall containing the injection well slides downward. An injection well located near the left boundary of the aquifer injects CO_2 as a supercritical (dense gas) phase into the aquifer at a constant rate of 0.0061 $m^3/s/m$ in the out-of-plane direction. Our model setup resembles CO_2 injection in the Snøhvit offshore field (Hansen et al., 2013) where CO_2 was injected in the 75–110 m thick Tubaen aquifer situated in a fault block that is sealed by faults on the north and south sides of the aquifer. We solve the coupled problem of flow and fault poromechanics using Pylith-GPRS, a two-way coupled multiphase flow and geomechanics simulator (Jha & Juanes, 2014).

Injection into the confined aquifer leads to a rise in the aquifer pressure, $\Delta p > 0$, and a volumetric expansion of the aquifer due to the coupling between flow and mechanics.

Increase in pressure results in a decrease in the effective compression on the fault, that is, $\Delta(-\sigma_{\mathrm{n}}') < 0$, where the effective normal traction σ_n' (positive in tension) is defined on the fault using the total normal traction and the fault pressure. We define the fault pressure to be the maximum of the two fault block pressures (Jha & Juanes, 2014; Vilarrasa et al., 2016). Volumetric expansion of the aquifer results in an increase in the downdip shear traction, $\Delta\tau_{\mathrm{down}} > 0$, on the fault below the aquifer, and a decrease in the downdip shear traction, $\Delta\tau_{\mathrm{down}} < 0$, above the aquifer. As a result, the bottom point of the aquifer-fault intersection at a depth of 1,550 m reaches the Mohr-Coulomb failure criterion, $\tau = \tau_f = \mu_f\sigma_{\mathrm{n}}'$, before the top point at 1,450 m depth. The time of nucleation of slip at the hypocenter (1,550 m depth) depends on model parameters, including aquifer porosity and permeability, fluid and rock compressibilities, injection rate, and elastic moduli of the layers.

There are three regimes in the evolution of the state of stress on the fault (Figure 2): (1) before slip nucleation, which occurs at the bottom of the aquifer, $t < 15$ days; (2) before slip begins at the top of the aquifer, $15 < t < 29$ days; and (3) initial slip at the top of the aquifer, $t > 29$ days. The effective normal compression decreases monotonically because the pressure increases monotonically and the increase in total normal compression due to aquifer expansion is much smaller than the increase in pressure.

Before nucleation, the evolution is determined by the stress induced by volumetric expansion of the aquifer, which causes the downdip shear traction to increase at the 1,550 m point and decrease at the 1,450 m point. After nucleation, downdip shear at 1,550 m decreases because of slip-induced shear relaxation. Evolution of stress after slip nucleation is determined by a balance between two physical mechanisms: stress due to

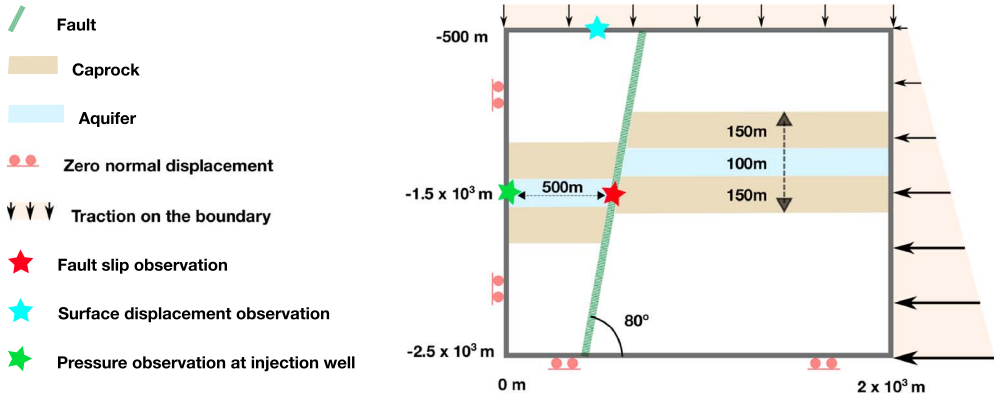

Figure 1. Model of the CO_2 injection-induced seismicity problem (adapted from Cappa & Rutqvist, 2011). The lateral compression is 0.7 times the overburden, and both increase with the lithostatic gradient. CO_2 is injected into the confined aquifer at a depth of 1500 m. The aquifer is bounded at the top and bottom by a low-permeability caprock, and the fault is impermeable to flow. The star-shaped symbols indicate observation points of aquifer pressure p_{obs} at (0, 1,500) m in green color, ground displacements $(u_{x,obs}, u_{y,obs})$ at (541, 500) m in cyan color, and downdip fault slip d_{obs} at (491, 1,550) m in red color.

injection-induced expansion and stress transfer from slip at the hypocenter. Moreover, these two mechanisms are evolving dynamically because aquifer injection and fault slip continue with time. For the values of μ_d, D_c, and k and the injection rate in Figure 2, the injection-induced stress dominates initially, causing an updip slip at 1,450 m at $t = 29$ days that is aligned with the upward expansion of the aquifer. Since the overburden stress is larger than the lateral stress (normal faulting boundary condition), downdip slip at the hypocenter is larger than the updip slip at 1,450 m. This eventually causes the stress transfer from the hypocenter to dominate over the injection-induced stress at the 1,450 m point at late times.

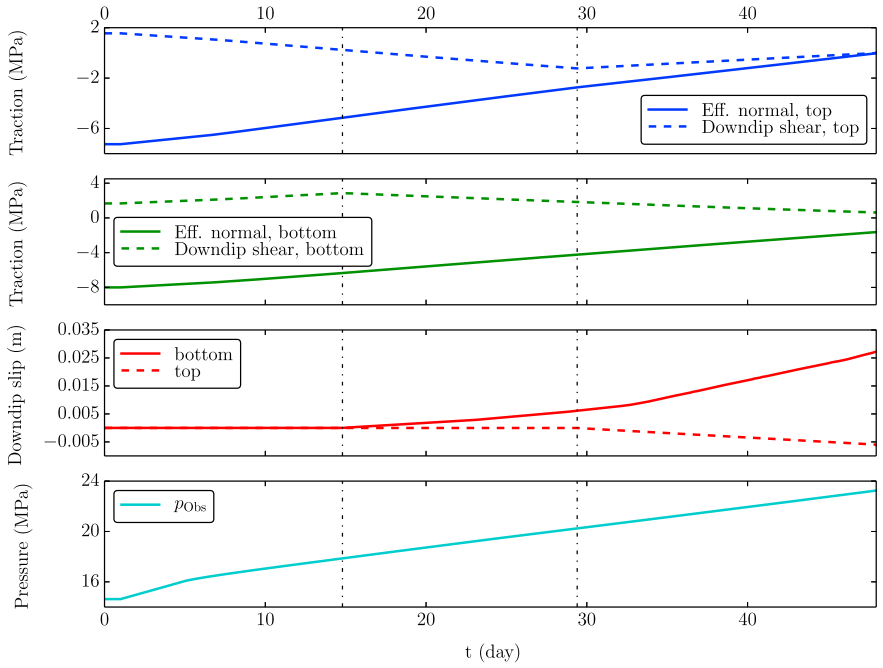

Figure 2. Evolution of fault tractions and slip at the top (1,450 m depth) and bottom (1,550 m depth) points along the aquifer-fault intersection line in a simulation with (μ_d, D_c, k) = (0.2, 0.1 m, 25 millidarcy). The aquifer pressure at the observation point, p_{obs}, increases monotonically with time resulting in a monotonically decreasing effective normal compression. Evolutions of shear traction and slip are nonmonotonic due to complex interplay among different physical mechanisms: injection-induced aquifer expansion, dynamic stress transfer, and slip-induced stress relaxation. The three regimes of stress evolution discussed in the text are delineated by vertical dashed lines at $t = 15$ and 29 days.

3. Bayesian Inversion Approach

The Bayesian statistical approach uses probability distributions to represent knowledge about parameter values. In particular, a prior probability distribution (with density π_{pr}) represents uncertainty about the inversion parameters $\theta \in \mathbb{R}^d$ before incorporating the data, including any relevant physical constraints. In the setting of the coupled flow-geomechanical model of section 2, the parameters to be inferred are $\theta := (\mu_d, D_c, k)$. The results of inversion are captured in the posterior probability distribution (with density π_{pos}), which follows from conditioning on the data $y \in \mathbb{R}^n$ via Bayes' rule:

$$\pi_{\text{pos}}(\theta|y) \propto \pi(y|\theta)\pi_{\text{pr}}(\theta). \tag{1}$$

Here we assume that an additive noise η accounts for the mismatch between data and model predictions, that is, $y = \mathcal{G}(\theta) + \eta$, where $\mathcal{G} : \mathbb{R}^d \to \mathbb{R}^n$ is the forward model. We follow standard practice in inverse problems by letting η be a Gaussian random variable (Kaipio & Somersalo, 2006). Letting π_η denote the probability density of η, the likelihood term in equation (1) reduces to $\pi(y|\theta) = \pi_\eta(y - \mathcal{G}(\theta))$. The data y comprise measurements of p_{obs}, $u_{x,\text{obs}}$, $u_{y,\text{obs}}$, and d_{obs} at the observation points indicated in Figure 1, taken at nine successive times: $t \in \{5, 10, 15, \ldots, 45\}$ days. Fault slip data d_{obs} usually come from the inversion of seismic and/or geodetic data (Kanamori & Brodsky, 2004), which are increasingly available through seismic and geodetic monitoring of injection sites. Each evaluation of $\mathcal{G}$ involves extracting predictions of these observables from a two-way coupled simulation of the injection process.

To make the solution of the inverse problem tractable, we replace the high-fidelity forward model $\mathcal{G}$ in the likelihood with a surrogate model $\widetilde{\mathcal{G}}$. Our surrogate consists of generalized polynomial chaos (gPC) expansions (LeMaître & Knio, 2010; Xiu, 2010), which approximate each component $\mathcal{G}^{(j)}$ of $\mathcal{G}$ (for $j = 1 \ldots n$, corresponding to each observable and time) as a weighted sum of polynomials orthogonal with respect to π_{pr}. The surrogate model in turn induces a surrogate posterior $\widetilde{\pi}_{\text{pos}}(\theta|y) \propto \pi_\eta(y - \widetilde{\mathcal{G}}(\theta))\,\pi_{\text{pr}}(\theta)$ (compare to equation (1). The surrogate posterior is a close approximation to the true posterior insofar as $\widetilde{\mathcal{G}}$ is a good approximation to $\mathcal{G}$. Convergence of the surrogate model to the true forward model guarantees convergence of the surrogate posterior to the true posterior, measured using various divergences or distances between distributions (Cotter et al., 2010; Marzouk & Xiu, 2009).

gPC surrogate models offer several advantages over other choices such as Gaussian process regression models. They converge exponentially fast (in polynomial degree) when the outputs of the forward model depend smoothly on the input parameters. Moreover, through the use of sparse grids based on the Smolyak construction (Smolyak, 1963), we can create surrogates that scale to high-dimensional parameter spaces. These methods are advantageous even in the present problem, where $d = 3$. The Smolyak construction employs a carefully chosen combination of full-tensor polynomial approximations, taking advantage of certain mixed regularity in the function to be approximated. While Smolyak approaches in general seek to omit higher-order interactions among model parameters, it is computationally advantageous to identify the most important interactions via anisotropic adaptation (Gerstner & Griebel, 2003). In this work, we employ the adaptive pseudospectral approximation formulated by Conrad and Marzouk (2013). This formulation guarantees that certain aliasing errors (due to the approximation of polynomial coefficients from a finite number of model runs) remain small (Constantine et al., 2012), while providing useful error indicators for greedy adaptation of the underlying polynomial basis. Since our parameters θ will be endowed with uniform priors, the pseudospectral approximation employs Gauss–Patterson quadrature rules, which are nested by construction—further reducing the number of model runs used to build the surrogate. More details on this construction are in Conrad and Marzouk (2013).

After constructing the surrogate $\widetilde{\mathcal{G}}$ using model runs at parameter values dictated by our anisotropic adaptation procedure, we characterize the corresponding surrogate posterior distribution using MCMC sampling with long chains, since $\widetilde{\pi}_{\text{pos}}$ is inexpensive to evaluate (Marzouk & Xiu, 2009).

4. Numerical Results

In this section, we first validate our surrogate of the two-way coupled flow and geomechanical model. Then we perform several inversion studies to characterize the parameters governing poromechanical behavior of the fault. Computations are performed with the help of the open-source uncertainty quantification software MUQ (Parno et al., 2017).

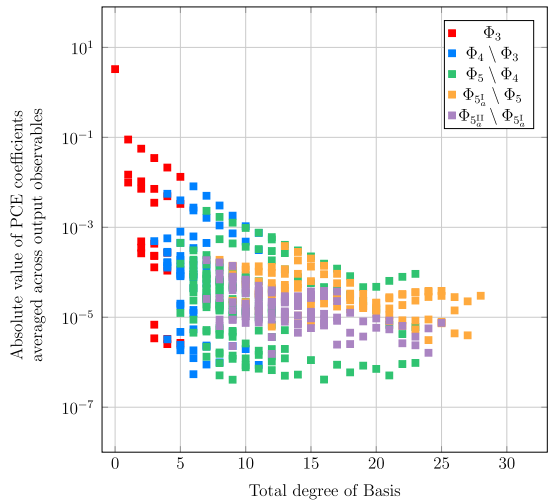

Figure 3. Magnitudes of coefficients in the surrogate model, versus total degree of the corresponding polynomial. Colors identify successive "generations" of basis functions, matching those in Figure S1. Coefficients added in later stages of the adaptive procedure are relatively small and correspond to refinements of the surrogate that are well below the observational noise magnitude.

4.1. Validation of the Adaptive Surrogate Model

To successfully employ the surrogate model in inversion, we must ensure its accuracy over the prior distribution π_{pr}; here the prior is chosen uniform on $\mu_d \in [0.1, 0.45]$, $D_c \in [0.05, 1.0]$ m, and $k \in [10, 200]$ millidarcy.

Recall that the surrogate consists of a sparse sum of tensorized orthonormal polynomial basis functions, adaptively chosen. Results of this adaptation procedure, and their relationship to the physics encoded in the forward model, are detailed in the supporting information. Here we focus on assessing the accuracy of the surrogate, which requires examining not only the choice of basis functions but the magnitudes of their coefficients. Figure 3 shows the absolute value of each polynomial coefficient in our final approximation as a function of the total degree of the corresponding polynomial. Since the surrogate comprises a distinct polynomial expansion for each output (though using a common set of basis functions), we plot the mean—across all the observables—of the absolute value of each coefficient. Additionally, since the pressure, displacement, and fault slip values vary by several orders of magnitude in physical units, our approximations are constructed after scaling the corresponding outputs of the forward model to be $\mathcal{O}(1)$. As in Figure S1 of the supporting information, we have color coded the markers using five distinct sets. The key takeaway from Figure 3 is that the coefficients corresponding to higher-order basis functions (added later in the adaptive procedure) are significantly smaller in magnitude than the lower-order coefficients. Based on the magnitudes of these coefficients, and the $\mathcal{O}(10^{-2})$ standard deviation of our normalized observational noise (see below), we conclude that the accuracy of the surrogate model is more than sufficient. Building this surrogate required 703 high-fidelity model evaluations. Any further refinements to the surrogate, which involve adding coefficients of even smaller magnitude, will not impact the inversion study at this noise level.

4.2. Results of Bayesian Inversion

We solve the inverse problem using synthetic data, generated by recording the response of the high-fidelity forward model $\mathcal{G}$ at $[\overline{\mu}_d, \overline{D}_c, \overline{k}] = [0.2, 0.1, 25]$, which we henceforth refer to as the *truth* parameter values. We assume an uncorrelated noise with a signal-to-noise ratio of 1%, reflective of practical observation scenarios. The synthetic data are perturbed with a realization of this noise. We explore the posterior distribution induced by the perturbed data and our surrogate model, using 500×10^3 steps of delayed rejection adaptive Metropolis MCMC (Haario et al., 2006) and discarding the first 5,000 samples as burn-in.

In Figure 4 we plot two-dimensional marginals of the posterior distribution for each pair of inference parameters. The contour lines represent different levels of the two-dimensional probability density function of each pair, with red corresponding to higher values than blue. We consider the case when the posterior is determined using all the observables described in section 3. The resulting distribution is clearly non-Gaussian, justifying the need for a fully Bayesian approach. The data are quite informative, in the sense that the posterior distribution is much narrower than the prior distribution in all parameter directions. (Compare the prior ranges, specified in section 4.1, with the much more concentrated regions of high posterior probability in Figure 4.) A striking feature of the posterior distribution is the negative correlation between μ_d and D_c. The posterior also suggests some positive correlation between k and D_c, particularly for the lower values of D_c that are favored by the data. These relationships—between D_c and μ_d and between D_c and k—uncovered by the Bayesian inversion can be interpreted from the physics of fault slip. A larger critical slip distance D_c decreases the rate of shear relaxation, and thereby the rate of slip, and thus increases frictional stability of the fault. Honoring the timing of fault slip d_{obs} then increases the likelihood of a lower dynamic friction coefficient μ_d (which decreases the frictional strength of the fault) and a higher reservoir permeability k (which accelerates pore pressure increase at the fault from fluid injection). More generally, the fault slip is also affected by mechanical parameters such as the rock compressibility and Poisson's ratio (Jha et al., 2015), which in the present context we assume to be known. We plan to include uncertainties in mechanical parameters in our future studies.

It is interesting that the posterior mode in Figure 4 does not coincide precisely with the *truth* parameter value (illustrated with a star symbol) used to generate the synthetic data, though it is close—especially relative

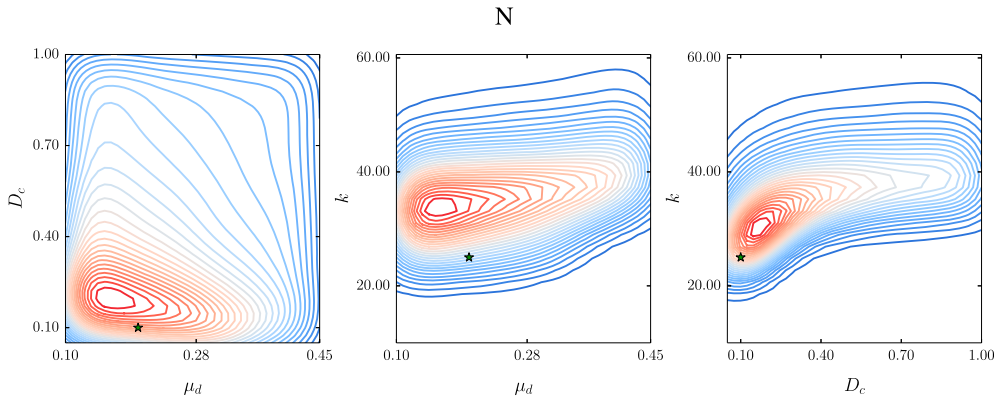

Figure 4. Pairwise marginals of the posterior distribution, with the *truth* parameters indicated by the star.

to the scale of the prior. Yet this mismatch is not surprising, as in general these two values should not coincide in Bayesian inference with finite data (Stark & Tenorio, 2010). In particular, the data here are perturbed by observational noise and the forward model has a relatively flat response to input parameter variations in certain regions of the prior.

In Figure 5 we plot marginal posterior distributions for one inference parameter at a time. We distinguish between the posterior obtained using all the data simultaneously, $\widetilde{\pi}_{\text{pos}}$, and the posteriors obtained using only observables of a single kind: $\widetilde{\pi}^{\text{p}}_{\text{pos}}$ for pressure measurements only, $\widetilde{\pi}^{\text{d}}_{\text{pos}}$ for ground displacements only, and $\widetilde{\pi}^{\text{fs}}_{\text{pos}}$ for fault slip measurements only. In Figure 5 (first and second panels), the marginals for parameters μ_d and D_c corresponding to the displacement (blue) and pressure (green) observations are not shown, as the corresponding data sets are relatively uninformative and we obtain the prior up to a sampling error. These two data sets do, however, inform the permeability k to some extent. In contrast, the fault slip observations are extremely informative about all the parameters. We also note that k is particularly well identified: its posterior distribution is very concentrated relative to its prior range. Pressure observations do inform the value of k, but fault slip observations are even more informative.

These "experiments" with different combinations of data illustrate the utility of performing Bayesian inference with different kinds of observations and—perhaps more importantly—the value of a two-way coupled model in *interpreting* these observations. Without a coupled model in the likelihood function, it would not be possible for fault slip observations to inform the aquifer permeability; nor would it be feasible to learn the parameters μ_d and D_c of the slip-weakening law, as honoring the observations hinges on capturing the interplay of injection-induced expansion and slip-induced stress transfer.

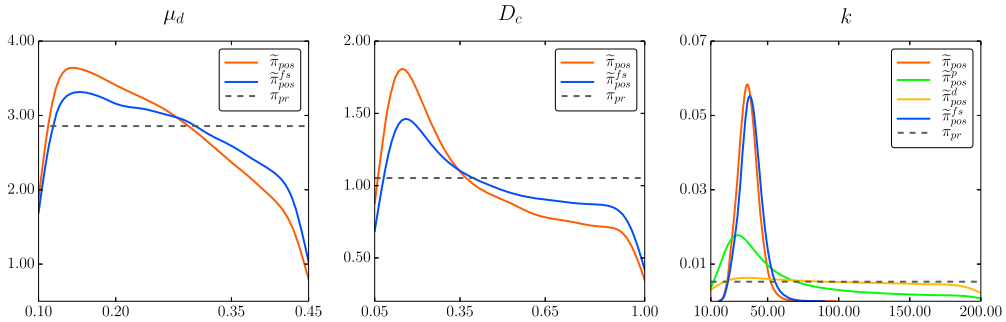

Figure 5. Comparison of single-parameter posterior marginal distributions and the prior. $\widetilde{\pi}_{\text{pos}}$ is obtained using all the observables; $\widetilde{\pi}^{\text{p}}_{\text{pos}}$ is obtained using pressure measurements only, $\widetilde{\pi}^{\text{d}}_{\text{pos}}$ is obtained using surface displacement measurements only, and $\widetilde{\pi}^{\text{fs}}_{\text{pos}}$ is obtained using fault slip measurements only. $\widetilde{\pi}^{\text{p}}_{\text{pos}}$ and $\widetilde{\pi}^{\text{d}}_{\text{pos}}$ in the first and second panels are essentially the prior density π_{pr} up to a sampling error and have not been plotted.

5. Discussion and Conclusions

We have presented a coupled flow-geomechanics Bayesian inversion framework to learn the dynamic friction properties of faults subject to destabilization from injection. The framework is unique in combining the rigorous multiphysics simulation of induced seismicity with a sophisticated surrogate-based Bayesian inference methodology. We infer the dynamic friction coefficient and the critical slip distance, two of the most uncertain and seismically influential fault parameters, along with the aquifer permeability between the injection well and the fault. Our inference framework reveals an inverse relationship between the two fault parameters, suggesting that the dynamic friction coefficient is smaller for faults with larger critical slip distances—a phenomenon that has previously been reported for mature faults such as the San Andreas. More generally, our results point to the benefit of accounting for the coupling between flow and geomechanics for improved identification of model parameters in the subsurface and elucidation of the mechanisms potentially responsible for injection-induced seismicity. In the future, we intend to apply our framework to a real site of injection-induced seismicity with site-specific values of the model parameters

Acknowledgments
The authors are grateful for support from the Eni-MIT Alliance research program. Additionally, J. Jagalur Mohan and Y. Marzouk acknowledge support for software infrastructure on Bayesian inversion from the National Science Foundation SI2-SSI program, under award 1550487. The data used in this work are synthetic, produced by simulations that are fully described in the paper. Computations pertinent to the inversion process were performed using the open-source uncertainty quantification software MUQ (http://muq.mit.edu). Any other details can be obtained from the corresponding author.

References

Campillo, M., & Ionescu, I. R. (1997). Initiation of antiplane shear instability under slip dependent friction. *Journal of Geophysical Research*, *102*, 363–371.

Cappa, Frédéric, & Rutqvist, Jonny (2011). Impact of CO_2 geological sequestration on the nucleation of earthquakes. *Geophysical Research Letters*, *38*, L17313. https://doi.org/10.1029/2011GL048487

Cocco, M., & Bizzarri, A. (2002). On the slip-weakening behavior of rate- and state dependent constitutive laws. *Geophysical Research Letters*, *29*, 111–114. https://doi.org/10.1029/2001GL013999

Conrad, P., & Marzouk, Y. M. (2013). Adaptive Smolyak pseudospectral approximations. *SIAM Journal on Scientific Computing*, *35*(6), A2643–A2670.

Constantine, P. G., Eldred, M. S., & Phipps, E. T. (2012). Sparse pseudospectral approximation method. *Computer Methods in Applied Mechanics and Engineering*, *229–232*, 1–12.

Cotter, S. L., Dashti, M., & Stuart, A. M. (2010). Approximation of Bayesian inverse problems for PDEs. *SIAM Journal on Numerical Analysis*, *48*(1), 322–345.

Dieterich, J. H. (1979). Modeling of rock friction, 1. Experimental results and constitutive equations. *Journal of Geophysical Research*, *84*, 2161–2168.

Fulton, P. M., Brodsky, E. E., Kano, Y., Mori, J., Chester, F., & Ishikawa, T. (2013). Low coseismic friction on the Tohoku-Oki fault determined from temperature measurements. *Science*, *342*, 1214–1217.

Gerstner, T., & Griebel, M. (2003). Dimension-adaptive tensor-product quadrature. *Computing*, *71*(1), 65–87. https://doi.org/10.1007/s00607-003-0015-5

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing*, *16*(4), 339–354. https://doi.org/10.1007/s11222-006-9438-0

Hansen, O, Gilding, D., Nazarian, B., Osdal, B., Ringrose, P., Kristoffersen, J.-B., … Hansen, H. (2013). SnøHVIT: The history of injecting and storing 1 Mt CO_2 in the Fluvial Tubåen Fm. *Energy Procedia*, *37*(Supplement C), 3565–3573.

Jha, B., Bottazzi, F., Wojcik, R., Coccia, M., Bechor, N., McLaughlin, D., … Juanes, R. (2015). Reservoir characterization in an underground gas storage field using joint inversion of flow and geodetic data. *International Journal for Numerical and Analytical Methods in Geomechanics*, *39*(14), 1619–1638. https://doi.org/10.1002/nag.2427

Jha, B., & Juanes, R. (2014). Coupled multiphase flow and poromechanics: A computational model of pore pressure effects on fault slip and earthquake triggering. *Water Resources Research*, *50*(5), 3776–3808. https://doi.org/10.1002/2013WR015175

Juanes, R., Jha, B., Hager, B. H., Shaw, J. H., Plesch, A., Astiz, L., … Frohlich, C. (2016). Were the May 2012 Emilia–Romagna earthquakes induced? A coupled flow-geomechanics modeling assessment. *Geophysical Research Letters*, *43*, 6891–6897. https://doi.org/10.1002/2016GL069284

Kaipio, J., & Somersalo, E. (2006). *Statistical and computational inverse problems* (Vol. 160). New York: Springer.

Kanamori, H., & Brodsky, E. E. (2004). The physics of earthquakes. *Reports on Progress in Physics*, *67*(8), 1429–1496.

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., & Ge, S. (2014). Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. *Science*, *345*(6195), 448–451.

LeMaître, O., & Knio, O. M. (2010). *Spectral methods for uncertainty quantification*. Netherlands: Springer.

Marone, C. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*, 643–696.

Marzouk, Y., & Xiu, D. (2009). A stochastic collocation approach to Bayesian inference in inverse problems. *Communications in Computational Physics*, *6*(4), 826–847.

Marzouk, Y. M., Najm, H. N., & Rahn, L. A. (2007). Stochastic spectral methods for efficient Bayesian solution of inverse problems. *Journal of Computational Physics*, *224*(2), 560–586.

Ohnaka, M., & Shen, L. (1999). Scaling of the shear rupture process from nucleation to dynamic propagation: Implications of geometry irregularity of the rupturing surfaces. *Journal of Geophysical Research*, *104*, 817–844.

Parno, M., Davis, A., Conrad, P., & Marzouk, Y. M. (2017). MIT Uncertainty Quantification Library (MUQ). Retrieved from http://muq.mit.edu

Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research*, *98*, 9885–9907.

Rubin, A. M., & Ampuero, J.-P. (2005). Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research*, *110*, B11312. https://doi.org/10.1029/2005JB003686

Ruina, A. L. (1983). Slip instability and state variable friction laws. *Geophysical Research Letters*, *88*, 359–370.

Scholz, C. H. (1988). The critical slip distance for seismic faulting. *Nature*, *336*, 761–763.

Scholz, C. H. (1989). Mechanics of faulting. *Annual Review of Earth and Planetary Sciences*, *17*(1), 309–334. https://doi.org/10.1146/annurev.ea.17.050189.001521

Scholz, C. H. (2002). *The mechanics of earthquakes and faulting*. New York: Cambridge University Press.

Sivia, D., & Skilling, J. (2006). *Data analysis: A Bayesian tutorial*. Oxford: Oxford University Press.
Smolyak, S. A. (1963). Quadrature and interpolation formulas for tensor products of certain classes of functions. *Doklady Akademii Nauk SSSR*, *4*(123), 240–243.
Stark, P. B., & Tenorio, L. (2010). A primer of frequentist and Bayesian inference in inverse problems. In L. T. Biegler et al. (Eds.), *Large-scale inverse problems and quantification of uncertainty* (pp. 9–32). New York: John Wiley. https://doi.org/10.1002/9780470685853.ch2
Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. Philadelphia: Society for Industrial and Applied Mathematics.
Toro, G. D., Han, R., Hirose, T., Paola, N. De, Nielsen, S., Mizoguchi, K., … Shimamoto, T. (2011). Fault lubrication during earthquakes. *Nature*, *471*, 494–499.
Vilarrasa, V., Makhnenko, R., & Gheibi, S. (2016). Geomechanical analysis of the influence of CO_2 injection location on fault stability. *Journal of Rock Mechanics and Geotechnical Engineering*, *8*(6), 805–818.
Xiu, D. (2010). *Numerical methods for stochastic computations: A spectral method approach*. Princeton, NJ: Princeton University Press.

Supporting Information for "Inferring fault frictional and reservoir hydraulic properties from injection-induced seismicity"

Jayanth Jagalur-Mohan[1], Birendra Jha[2], Zheng Wang[1], Ruben Juanes[3,4], and Youssef Marzouk[1]

[1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA.

[2]Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, CA, USA.

[3]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.

[4]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA.

Contents

This document provides additional numerical results supporting our Bayesian inversion study of the poromechanical behavior of the aquifer-fault system.

Text S1: Adaptive construction of the surrogate model

Recall that the objective of adaptation is to iteratively refine the surrogate by adding relevant basis functions until a desired accuracy is achieved. Adaptation should thus reflect the physics encoded in the forward model—identifying parameters that yield a more nonlinear response in the observables, parameters that interact closely and hence require coupled basis functions, and parameters that act alone. Beginning with a basis set defined by an isotropic sparse grid of prescribed total order, we incrementally incorporate basis functions of higher order into the approximation. In the present case we consider a level–5 sparse grid before starting the anisotropic adaptation. Fig. S1 visualizes the basis functions in our final surrogate model. Each cube corresponds to a unique basis function, whose polynomial degree in each parameter dimension is indicated by the three axes. The cubes are color-coded by the *stage* of adaptation at which they are added. We highlight five distinct sets of basis functions and annotate the plot with the cardinality of each set. The set Φ_3, colored red, corresponds to a coarse surrogate model defined by a level–3 isotropic sparse grid. The basis set $\Phi_4 \setminus \Phi_3$, colored blue, is introduced in the level–4 isotropic sparse grid. The remaining sets can be understood in a similar manner. In particular, the sets $\Phi_{5_a^I} \setminus \Phi_5$ and $\Phi_{5_a^{II}} \setminus \Phi_{5_a^I}$ contain the basis functions introduced after two successive anisotropic adaptation steps, beyond the level–5 isotropic sparse grid. A key observation emerging from the adaptation process is that the friction coefficient μ_d and critical slip distance D_c *interact* nonlinearly; in other words, mixed high-order polynomials in both parameters are needed to resolve the forward model response. The permeability k is comparatively less coupled.

Corresponding author: J. Jagalur-Mohan, jagalur@mit.edu

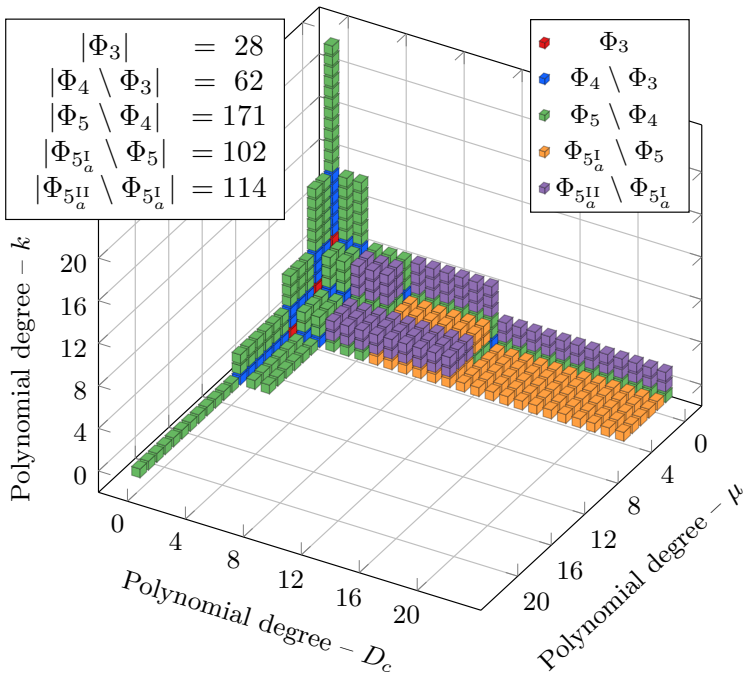

Figure S1. Polynomial basis functions used to construct a surrogate of the coupled flow–geomechanical model. Each cube corresponds to a polynomial whose degree in each parameter direction is indicated by the three axes. Colors identify different "generations" of basis functions added by the adaptive procedure. The pattern reflects the anisotropic dependence of the model outputs (pressure, ground displacements, fault slip) on the input parameters.