Supporting Information for "Critical evaluation of state evolution laws in rate and state friction: fitting large velocity steps in simulated fault gouge with time-, slip- and stress-dependent constitutive laws"

Contents of this file

1. Text

2. Figures S1 and S2

Introduction

This supplementary material file describes the details of the Markov Chain Monte Carlo (MCMC) algorithm used in the accompanying article. We describe below the non-standard modifications we needed to introduce to the classical Metropolis-Hastings approach to ensure fast converging and well-mixed chains. In addition, we also outline our philosophy and methodology for treating data-errors in our experiments, and how this information was used to construct a properly weighted likelihood function.

Details of our MCMC algorithm

The MCMC algorithm used here was based on the Metropolis-Hastings (MH) algorithm [*Metropolis et al.*, 1953; *Hastings*, 1970] and makes use of the strategy of sampling the

Corresponding author: Pathikrit Bhattacharya, Department of Geosciences, Princeton University, Guyot Hall, Princeton, NJ 08544, USA. (pathikri@princeton.edu)

posterior using Bayes' theorem. It is worthwhile to spell out here our philosophy behind attempting these detailed parameter estimations. We believe that the laboratory derived rate and state friction laws are likely inaccurate. We want to study how closely the best fitting parameters for each of the state evolution laws explain the various experimental datasets. Our repeated experiments with the downhill simplex algorithm lead us to believe that the posteriors are multi-modal for some of these inverse problems. Therefore our central focus is on the existence and positions of these modes/mode. The most general approach to characterizing these modes, is to attempt to construct the full marginal posteriors for each parameter, a problem well-suited to Bayesian inference.

Before introducing the core structure of the MH variant used in this study, we need to list some background information pertaining and specific to the inversion of typical laboratory friction data. At the outset, it is important to recognize that the models used here do not allow for negative values of any of the parameters of interest for a realistic experimental dataset, with the possible but rare exception of b, which can be negative in some cases [*Carpenter et al.*, 2011]. This permits us to transform the parameter space to logarithmic scales, allowing us to restrict sampling only to the positive real axis. For setting up the sampling procedure, we make a few more simplifications. Firstly, we choose our priors to be uniform distributions in the logarithmic domain over many orders of magnitudes. The upper and lower bounds of this uniform distribution are not needed in practice due to their appearance in both the numerator and denominator terms in the acceptance ratio formulated from Bayes' theorem. We further make the assumption that the priors of the different parameters are independent of each other. This is definitely not true in the case of the Nagata law. This choice is dictated by the fact that the

covariances characterizing the joint prior are unknown to us. Finally, the time sampling of the stress measurements during our experiments is such that the number of data points recorded during sliding at a particular slip rate is a function of slip rate. Typically, slow sliding contributes many more data points than fast sliding. This a-priori bias is relaxed by using weights that are scaled according to the number of points within the particular phase of sliding such that all phases (each velocity increase and decrease) have equal contribution to the misfit. Further, these weights are scaled by numerical factors linearly proportional to $d\mu/d\delta\,(= V^{-1}d\mu/dt)$. Within a particular phase of sliding, when $V^{-1}d\mu/dt$ falls below a critical absolute value at a particular time, an exponential taper is applied for all subsequent times for as long as steady state sliding lasts (or until the end of the phase). This style of weight distribution is designed to promote better fits to non-steady state sliding and is typical of all the fits described in the text. Note that the steady state part of the data is independently used to constrain the steady state velocity dependence $a - b$ to a fixed value in most of our inversions.

With this background set, we are now in a position to list the elements that are used to construct the MH acceptance ratio α for rejection sampling in an n-dimensional parameter space given a N-tuple data vector:

1. Prior: In the logarithmic domain the prior is represented by a uniform distribution spanning many orders of magnitude about the initial guess. The initial guess comes from the Nelder-Mead inversions:

$$P(\mathbf{x_a}) = \prod_i \frac{1}{x^i_{\max} - x^i_{\min}}, \tag{S1}$$

where $x^i_{\max}$ and $x^i_{\min}$ are the upper and lower bounds on the i^{th} parameter in the logarithmic domain.

2. Proposal: The proposal function is a center shifting, symmetric Gaussian about the present point in the Markov chain. Again, all parameters were deemed to be independent with regards to the proposal. The mathematical form of the distribution is:

$$P(\mathbf{x_p}|\mathbf{x_q}) = \frac{1}{2\pi^{n/2}}|\Sigma|^{-1}\exp\left[-\frac{1}{2}(\mathbf{x_p}-\mathbf{x_q})^T\Sigma^{-1}(\mathbf{x_p}-\mathbf{x_q})\right], \tag{S2}$$

where $\mathbf{x_p}$ is the proposed vector state and $\mathbf{x_q}$ is the last accepted one. For typical MH algorithms, the parameter covariance matrix Σ is assumed to be diagonal in the absence of a-priori knowledge of its full structure, i.e. $\Sigma \sim \boldsymbol{\sigma}\mathbf{I_n}$ where $\boldsymbol{\sigma}$ is the vector of (non-degenerate) standard deviations fixed by experimentation. For our typical Markov chains, especially with the strongly correlated parameters of the Nagata law, such diagonal parameter covariances led to many wasted samples and, therefore, to very slow mixing and convergence. To get around this problem, we ran long trial chains to 'learn' the proposal covariance matrix specific to a particular problem. In such a run, we initiated the chain with a diagonal parameter covariance but, beyond a large prescribed number of initial samples, calculated the proposal covariance from the parameter values accepted in the chain. For the remainder of the trial chain, we continued adaptively updating the proposal covariance on the run to end up with a relatively stationary estimate. This proposal was then scaled for optimal convergence following the prescriptions in *Bai* [2009a, b]; *Rosenthal* [2011] and used to construct the posterior across various different chain initiations. In some cases, this approach led to more than an order of magnitude faster convergence.

3. Likelihood: The standard Gaussian likelihood was used under the assumption of point wise independence of the data set. Additionally, a scalar standard deviation, S_e, is used in the order of magnitude sense to represent data error. It is worthwhile to remind ourselves that the main source of the misfit in these models is not data uncertainty but model inadequacy. Given this, and the inherent lack of repeatability of laboratory friction experiments, we decided not be more rigorous in characterizing data errors. However, even the value of S_e is not known a-priori and, due to lack of many repeated friction experiments, is not trivial to estimate. One can, however, use the fact that steady state friction is assumed to be a constant frictional strength and treat consecutive measurements of the shear stress during steady state sliding as repeated measurements of the same physical state. Therefore any fluctuations in shear stress about this constant value can be used to construct a distribution of data errors (Figure S1). This likely leads to a lower bound on data errors because typical repeated experiments on the same material can lead to much larger fluctuations in the data values than the standard deviation of the resultant error distribution. Additionally, especially for gouge, there are often linear trends superposed on this constant steady state friction value which probably make our assertion regarding steady state and repeated measurements less tenable. We have checked that the distribution of errors about a linear trend is not very different from the distribution of errors about a constant steady state value. But we treat this estimate of the standard error strictly as a lower bound on the data error and typically use about 10-40 times this estimate as the value of S_e in our inversions. We explain the rationale behind this scaling in greater detail in Appendix C. We typically run chains with different values of S_e from

this range, and we choose the minimum value above which the posterior structure does not change conspicuously.

Once we have fixed S_e, the likelihood can be calculated as:

$$P(\mathbf{d}|\mathbf{x_p}) = \frac{1}{\sqrt{2\pi}^N S_e} \exp\left[-\frac{1}{2}(N-1)\chi^2\right], \tag{S3}$$

where χ is the weighted RMSE scaled by S_e. With this formulation, the acceptance ratio is defined as

$$\alpha = \min\left[1, \frac{P(\mathbf{d}|\mathbf{x_p})}{P(\mathbf{d}|\mathbf{x_q})}\right] \tag{S4}$$

When trying to construct the full posterior, one cannot rule out the possibility of multi-modality of the posterior, especially with the Nagata law which admits more than one local minimum.

Finally, it is worth mentioning that for well separated modes in the posterior, even an adaptive proposal MH algorithm might lead to a chain that is not well-mixed, e.g. the 'learned' parameter covariance might be deficient if the trial chains are not well mixed. One possible way to deal with this problem was suggested by *Guan et al.* [2006]. They proposed a hybrid local-global proposal distribution which could populate well separated modes by using the form:

$$Q(\mathbf{x_p}|\mathbf{x_q}) = (1-p)P(\mathbf{x_p}|\mathbf{x_q}) + \frac{p}{|S|}, \tag{S5}$$

where the local proposal from $P(\mathbf{x_p}|\mathbf{x_q})$ is occasionally superceded by a wild, uniform distribution proposal over a given space of measure $|S|$ with a small probability p. *Guan et al.* [2006] prescribe an empirical estimate of the parameter p as

$$p \approx \frac{1}{\sqrt{Mr}}, \tag{S6}$$

where M is the length of the Markov chain and r is the ratio of the area under the posterior contributed by the modes to the total area. This strategy has been observed to lead to rapid stabilization of the acceptance ratio of the chain for posteriors with well separated modes [*Guan et al.*, 2006]. Generally, the burn-in period (the initial portion of the chain not used for statistical inference) was decided by the amount of time it takes the code to saturate its acceptance rate to a stationary value. A typical converged Nagata law Markov Chain and its diagnostics are shown in Figure S2.

References

Bai, Y. (2009a), An adaptive directional Metropolis-within-Gibbs algorithm, Preprint.

Bai, Y. (2009b), Convergence of adaptive Markov Chain Monte Carlo algorithms, Phd thesis, University of Toronto.

Carpenter, B. M., C. Marone, and D. M. Saffer (2011), Weakness of the San Andreas Fault revealed by samples from the active fault zone, *Nat. Geosci.*, *4*, http://dx.doi.org/10.1038/ngeo1089, doi:10.1038/ngeo1089.

Guan, Y., R. Fleissner, P. Joyce, and S. M. Krone (2006), Markov chain monte carlo in small worlds, *Stat. Comp.*, *16*(2), 193–202, doi:10.1007/s11222-006-6966-6.

Hastings, W. K. (1970), Monte carlo sampling methods using markov chains and their applications, *Biometrika*, *57*(1), 97–109.

Metropolis, N., A. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*, 1087–1092.

Rathbun, A. P., and C. Marone (2013), Symmetry and the critical slip distance in rate and state friction laws, *J. Geophys. Res.: Solid Earth*, *118*(7), 3728–3741, doi:10.1002/jgrb.50224.

Rosenthal, J. S. (2011), Optimal proposal distribution and adaptive MCMC, in *Handbook of Markov Chain Monte Carlo*, edited by S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, pp. 93–112, Chapman and Hall/CRC Press.

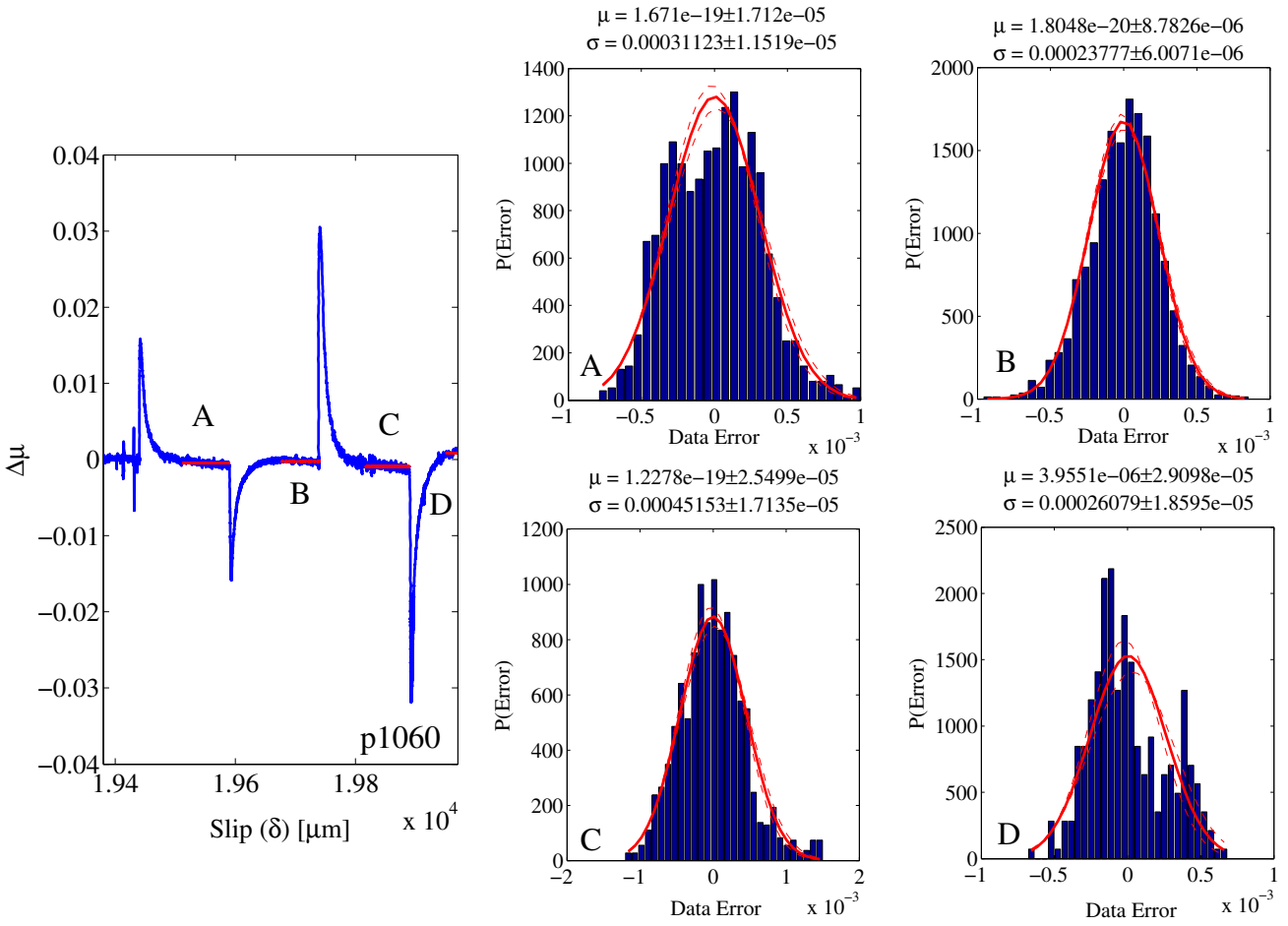

Figure S1. Estimate of standard error of the data from the steady state portions of the experimental data (dataset p1060). We assume that shear stress during steady state sliding following a velocity step should be constant. Therefore the consecutive measurements of shear stress during steady state sliding could be considered as repeated measurements of the same stress state. This means that the distribution of shear stress fluctuations about such a constant value could give us some estimate, most likely the lower bound, of the data error. The histogram of errors for the sections A, B, C and D of the data (labeled in the left hand panel showing the stress data) are plotted in parts (A), (B), (C) and (D). Maximum-likelihood estimates of Gaussian parameters for these error distributions are overlain on these histograms with the corresponding fits shown as red solid lines and their confidence intervals as red dashed lines.

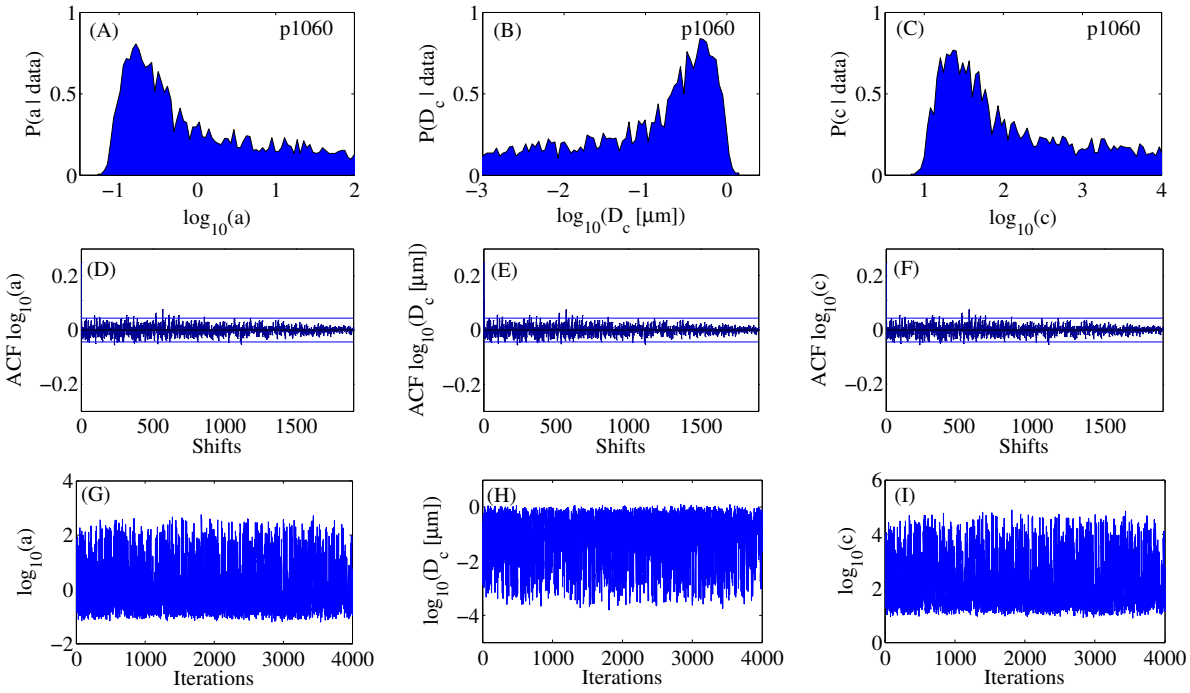

Figure S2. A typical converged Markov Chain for the Nagata law. This is the same chain shown in Figures 3 and 7 in the main text for dataset p1060. (A)–(C) show the posteriors for a, D_c and c, (D)–(F) show the auto correlation function (ACF) along the chain for each parameter along with 5% confidence level shown as horizontal blue lines, (G)–(I) show the trace plots along the chain for each parameter. The ACF estimates show that the samples are mostly uncorrelated at the 5% level. The trace plots show that the sequence is well mixed above a lower bound in c.