LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-225176

# A Markov-Chain Monte-Carlo Based Method for Flaw Detection in Beams

R. E. Glaser, C. L. Lee, J. J. Nitao, T. L. Hickling, W. G. Hanley

October 10, 2006

Journal of Engineering Mechanics

# A Markov-Chain Monte-Carlo Based Method for Flaw Detection in Beams

Ronald E. Glaser[1], Christopher L. Lee[2], John J. Nitao[3],
Tracy L. Hickling[1], William G. Hanley[4]

## Abstract

A Bayesian inference methodology using a Markov Chain Monte Carlo (MCMC) sampling procedure is presented for estimating the parameters of computational structural models. This methodology combines prior information, measured data, and forward models to produce a posterior distribution for the system parameters of structural models that is most consistent with all available data. The MCMC procedure is based upon a Metropolis-Hastings algorithm that is shown to function effectively with noisy data, incomplete data sets, and mismatched computational nodes/measurement points. A series of numerical test cases based upon a cantilever beam is presented. The results demonstrate that the algorithm is able to estimate model parameters utilizing experimental data for the nodal displacements resulting from specified forces.

## CE Database subject headings

Markov chains, Monte Carlo method, Bayesian analysis, structural models, damage assessment, cantilevers

[1] Electronics Engineering Technology Division, Lawrence Livermore National Laboratory, Livermore, CA.

[2] Assistant Professor of Mechanical Engineering, Franklin W. Olin College of Engineering, Needham, MA.

[3] Energy & Environment Directorate, Lawrence Livermore National Laboratory, Livermore, CA.

[4] Deputy Division Leader, Electronics Engineering Technology Division, Lawrence Livermore National Laboratory, Box 808, L-130, Livermore, CA 94551 (corresponding author) E-mail: hanley3@llnl.gov.

## Introduction

The identification of system parameters of computational models of mechanical systems and structures based upon measured data is a fundamental problem related to health monitoring of structures, finite-element model updating, and damage detection (Doebling *et. al.* 1996; Friswell and Mottershead 1995). Parameter identification based upon system output can be classified as an inverse problem, often significantly complicated by issues such as non-linearity, state-space dimensionality, under/over-determined systems, and noisy or dependent data. The severity of these issues often makes classical optimization/inversion algorithms ineffective for estimating system parameters, suggesting that a statistical approach may be more appropriate.

Bayesian inference, an inference methodology in which observed data can be used to update our belief in a particular outcome, has previously been applied to structural systems (Collins *et. al.* 1974; Beck 1989; Sohn and Law 1997; Vanik *et. al*. 2000; Beck and Au 2002; Yuen and Katafygiotis 2002; and Beck and Yuen 2004). In our work, we apply Bayesian inference by employing a Markov Chain Monte Carlo (MCMC) sampling scheme to generate posterior distributions over the space of potential system states. Specifically, through the use of a Metropolis-Hastings (M-H) algorithm (Metropolis *et. al*. 1953; Hastings 1970), plausible system states are proposed and used to generate predictions via a finite element-based forward model that are compared with available observed data. This process yields a sample that allows posterior distributions of system parameters and states to be constructed. The resulting distributions provide quantitative measures of estimation uncertainty, allow for the objective assessment of competing estimates when the available information is insufficient to definitively identify the system state, and provide a reasonable basis for follow-on predictions of structural behavior. This methodology has been successfully applied in the geophysical domain (Aines *et. al*. 2002; Newmark *et. al*. 2003; Ramirez *et. al*. 2005).

## System (Forward) Model: Fixed-Free Cantilever Beam

We demonstrate our approach by considering a fixed-free linearly elastic cantilever beam, discretized as dictated by its *mesh* into $n$ elements of equal size. Each element in the model is homogeneous with respect to its bending stiffness as parameterized by Young's modulus ($E$), whose prior probability distributions for the flawed and unflawed conditions are defined as shown in Figure 1. $E$ values of 'unflawed' elements fall within the interval $[a^{(0)}, b^{(0)}]$ = [1.8e11, 2.05e11], while 'flawed' elements have known $E$ values defined over the interval $[a^{(1)}, b^{(1)})$ = [0.8e11, 1.8e11). Additionally, we will model the frequency of flaw occurrence by placing a prior probability distribution on the number of flawed elements in the beam, i.e., its *level*, assuming no prior knowledge of their locations.

The beam is characterized through the application of $\eta$ (eta) independent static nodal forces at some subset of its $n$ nodes. These forces deflect the beam, and the resulting displacement is measured at each node in $d$ degrees-of-freedom. The equation of static deflection for the $j^{\text{th}}$ applied force is $\mathbf{K}\mathbf{X}^0_j=\mathbf{F}_j$, where $\mathbf{K}$ is the $nd \times nd$ stiffness matrix (each element is a function of $E$), and $\mathbf{X}^0_j$ is the $nd \times 1$ theoretical (i.e., mean) displacement vector induced by the $nd \times 1$ force vector $\mathbf{F}_j$.

The system state is identified by specification of the Young's modulus, $E$, for each of its flawed elements, along with their number, location (i.e., $i = 1, \ldots, n$), and severity. For this purpose we

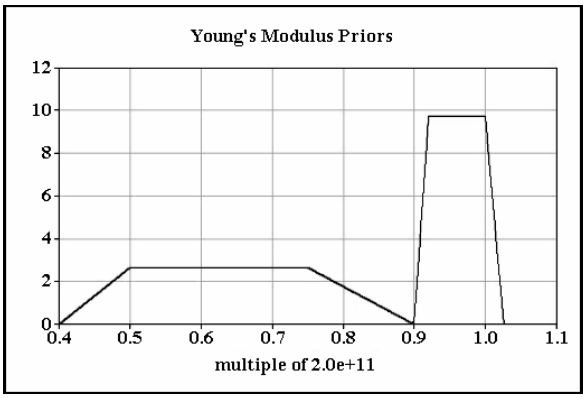


Figure 1: Trapezoidal priors for a beam element's Young's modulus (*E*). Elements with a modulus in the interval [0.8e11, 1.8e11) are "flawed", while those having a modulus in the interval [1.8e11, 2.05e11] are "unflawed".

introduce the notion of a *configuration*, which is a subset of possible states of nature of the system identified by these parameters. In defining a configuration, we do not specify the modulus values of the known unflawed elements. For example, the *k*-flaw configuration, **s**, having flawed elements at locations $i_1, \ldots, i_k$ with respective moduli $\tilde{E}_{i_1},\ldots,\tilde{E}_{i_k}$, consists of the class of possible states $\mathbf{s} = \left\{(E_1,\ldots,E_n): E_{i_1} = \tilde{E}_{i_1},\ldots,E_{i_k} = \tilde{E}_{i_k}, \tilde{E}_j \in [a^{(0)},b^{(0)}], j \neq i_1,\ldots,i_k\right\}$. It is convenient, however, to represent a configuration in terms of *quantiles* of the prior distribution for the flawed condition where the modulus $\tilde{E}_{i_j}$ is the $u_{i_j}$-quantile defined by the equation

$$u_{i_j} = P_1\left(E \le \tilde{E}_{i_j}\right), \tag{1}$$

where the subscript "1" refers to the flawed condition. Thus we can represent **s** using the alternative representation $\mathbf{s} = (\boldsymbol{\omega}, \boldsymbol{\upsilon}, m)$, where $\boldsymbol{\omega}$ (omega) $= (i_1, \ldots, i_k)$ is the naturally ordered set of flawed element locations, $\boldsymbol{\upsilon}$ (upsilon) $= \left(u_{i_1}, \ldots, u_{i_k}\right)$ is the set of quantiles ordered such that the flawed element at $\omega_j$ has the modulus value corresponding to the quantile $\upsilon_j$, and $m$ is the mesh of the finite-element model associated with configuration **s**.

It is over this space of configurations that we will generate a posterior distribution by applying Bayesian inference via the MCMC sampling scheme. Bayesian inference is rooted in *Bayes' Rule*, which may be expressed as $P(\mathbf{s} \mid D) \propto P(D \mid \mathbf{s})\pi_{\mathbf{s}}$. Conceptually, this means that our posterior belief, $P(\mathbf{s} \mid D)$, that **s** is the "true" beam configuration given the data, *D*, is proportional to the likelihood, $P(D \mid \mathbf{s})$, of observing the data given the configuration **s** weighted by our prior belief, $\pi_{\mathbf{s}}$, that **s** is the true configuration. It is within this paradigm that we employ MCMC sampling to estimate the posterior distribution.

***The Posterior Sampler***

By design, the sampler traverses the space of configurations, rather than all states of nature. To accomplish this, a conditional density, $p(\mathbf{s}' \mid \mathbf{s})$, is introduced for each pair of configurations,

$(\mathbf{s}, \mathbf{s}')$. We typically refer to $p(\mathbf{s}'|\mathbf{s})$ as the *proposal transition density*, $q(\mathbf{s}, \mathbf{s}')$, and we define $q$ such that $\int q(\mathbf{s}, \mathbf{s}')d\mathbf{s}' = 1$. These proposal densities dictate the manner in which the sampler explores the configuration space. A transition from configuration **s** to **s'** is proposed (randomly sampled) according to the proposal density, and the sampler accepts the proposed transition with probability equal to the min (1, $r(\mathbf{s}, \mathbf{s}')$), with $r$ given by

$$r(\mathbf{s}, \mathbf{s}') = \frac{\pi_{\mathbf{s}'} q(\mathbf{s}', \mathbf{s}) L_{\mathbf{s}'}(D)}{\pi_{\mathbf{s}} q(\mathbf{s}, \mathbf{s}') L_{\mathbf{s}}(D)}, \tag{2}$$

where $L_{\mathbf{s}}(D)$ is the likelihood of the measured displacement data, $D$, given **s**, and $\pi_{\mathbf{s}}$ is the prior distribution. If the proposal is rejected, the Markov chain remains at state **s** for that step.

The resulting Markov chain, known as the M-H sampler, can be shown to sample the space of configurations according to the posterior distribution once it has achieved equilibrium (Metropolis *et. al*. 1953; Hastings 1970). In general, to guarantee that equilibrium is achieved, the Markov chain must be irreducible and ergodic (Karlin and Taylor 1975). Moreover, the sampling process must "warm up" so as to forget the initial starting state (usually selected at random) to ensure that the drawn samples are indeed taken from the equilibrium distribution. Specifically, in order to "warm up", MCMC implementations must specify a *burn-in* period, which stipulates the number of samples at the beginning of the process to be discarded. The ensuing samples are then taken to be sampled from the posterior distribution. Note that if $h(\mathbf{s}, \mathbf{s}') = \pi_{\mathbf{s}'} q(\mathbf{s}', \mathbf{s})[\pi_{\mathbf{s}} q(\mathbf{s}, \mathbf{s}')]^{-1} = 1$, then the decision to accept the proposal **s'** is based entirely upon the likelihood ratio. A chain for which $h(\mathbf{s}, \mathbf{s}') = 1 \; \forall \; \mathbf{s}, \mathbf{s}'$ is said to be *reversible*.

***Prior Densities***

In the upcoming discussion, we consider two discretizations of the beam model; a *fine mesh* model of $n_{\mathrm{f}}$ elements versus a *coarse mesh* model of $n_{\mathrm{c}}$ elements (Figure 2). We allocate the prior probabilities of configurations in a quasi-uniform fashion throughout each level. In particular, the prior, $\pi_{\mathbf{s}}$, for a configuration **s** in mesh $m \in \{\mathrm{f}, \mathrm{c}\}$ at level $k \geq 1$ is given by

$$\pi_0 = \lambda_0; \quad \pi_{\mathbf{s}} = \pi^m \lambda_k \binom{n_m}{k}^{-1} \prod_{j=1}^{k} f_1(\tilde{E}_{\omega_j}), \; \mathbf{s} \neq 0, \tag{3}$$

where $\pi^m$ is the prior probability of mesh $m$, $\lambda_k$ is the prior probability of level $k$, and $f_1$ denotes the flawed modulus density. The prior probability for the sole level 0 (zero-flaw) configuration is given by $\pi_0$, where the associated mesh is, by default, that of the data.

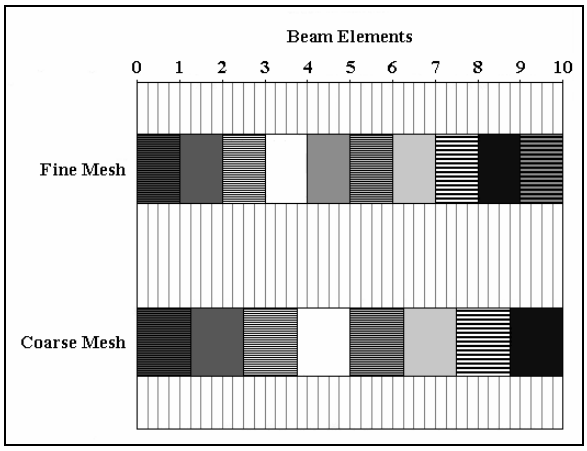


Figure 2: Two finite-element representations for the cantilever beam with fixed end at location 0. The fine mesh has $n_f = 10$ elements with nodes at 1, 2, 3, …, 10, and the coarse mesh has $n_c = 8$ elements with nodes at 1.25, 2.50, 3.75, …, 10.

***Proposal Density***

The proposal function is defined explicitly for all configuration transition pairs. The basic strategy is to consider, in turn, independent configuration changes in level, mesh, modulus, and/or flaw location. The four possible transition types are discussed below, and corresponding proposal densities are presented in the Appendix.

*Transition Type 1.* **s** *and* **s'** *are of the same level and mesh.*

With the level and mesh held constant, we may propose the relocation of exactly one flaw in the current beam configuration to a randomly selected location with a new modulus value that is randomly proposed according to $f_1$. For a configuration with level $k > 1$, we select, at random, one of the flawed elements in **s** to be replaced. When there is exactly one flaw in the system, however, we enable a greater degree of control in the generation of a proposal configuration by introducing the proposal probability, $\beta$, that a *neighbor* will be selected as the new flaw location and modulus. Flaws are considered to be neighbors if they are physically adjacent in the finite-element model and if their moduli are within distance $\delta$ of each other in terms of probability. If this property holds for any two moduli, $\tilde{E}_{i_a}$ and $\tilde{E}_{i_b}$, i.e., $\left| \upsilon_a - \upsilon_b \right| < \delta$ where $\delta \in (0, .5)$, we say that they satisfy the $\delta$*-criterion*. Smaller values for the parameter $\delta$ will, in general, allow a more fine-grained investigation of the state space, and when $\beta$ is relatively large, the sampler will more frequently linger in a local neighborhood around the current single-flaw configuration.

The transition from configuration **s** to **s'** may also, with probability $\theta_k$, exhibit a modulus-only change, where the mesh, level, and flaw locations are held constant. Such a transition is implemented by independently proposing a new modulus for each of the system's $k$ flaws according to the $\delta$-criterion. Like the neighbor-only transition probability, $\beta$, large values of $\theta_k$ encourage the sampler to remain within a local neighborhood of the current configuration.

*Transition Type 2.* **s** *and* **s'** *are of different levels but of the same mesh.*

We require our algorithm to vary the level of a configuration by only one flaw, if a level change is proposed. In this scenario, the proposal probability for transitioning from level $i$ to level $j$ is

specified by the probability $a_{ij} > 0$, for $|i - j| = 1$, $i, j \in \{0, 1, \ldots, K\}$. When the configuration level is decreased, a single flaw in the current state is randomly selected for removal, while all remaining flaw locations and moduli remain constant. If a flaw is added to the system, an unflawed element in the current state is randomly selected and assigned a random modulus according to the flawed modulus density, $f_1$.

*Transition Type 3.* **s** *and* **s'** *are of the same level but of different meshes.*
The proposal probability of transitioning from a mesh $m$ to a mesh $m'$ is given by the parameter $p_{mm'}$. If $m$ and $m'$ are different meshes, we say that a position sequence, $\boldsymbol{\omega}'$, of **s'** is viable with respect to the current state **s** if elements $\omega_j$ and $\omega'_j$ overlap for $j = 1, \ldots, k$ as shown in Figure 3. The configuration **s'** is itself a viable proposal with respect to **s** if and only if its position sequence is viable and, moreover, the corresponding moduli satisfy the $\delta$-criterion. We enforce the overlap constraint by weighting the proposal probability for each viable position sequence by the extent of its overlap with the position sequence of the current system state. If we denote the overlap of elements $\omega_j$ and $\omega'_j$ by $O_j$ (uppercase oh), a measure of this overall overlap is the product, $\prod_{j=1}^{k} O_j$. Let $\Omega_{\mathbf{s}}$ be the set of viable position sequences in the opposite mesh generated by **s**. Then we select the sequence $\hat{\boldsymbol{\omega}}' \in \Omega_{\mathbf{s}}$ with probability

$$\gamma^{\mathbf{s}}_{\hat{\boldsymbol{\omega}}'} = \prod_{j=1}^{k} O_j^{\hat{\omega}'} \left( \sum_{l \in \Omega_{\mathbf{s}}} \prod_{j=1}^{k} O_j^{l} \right)^{-1}, \tag{4}$$

so that sequence pairs with a higher degree of overlap are afforded higher proposal probabilities.

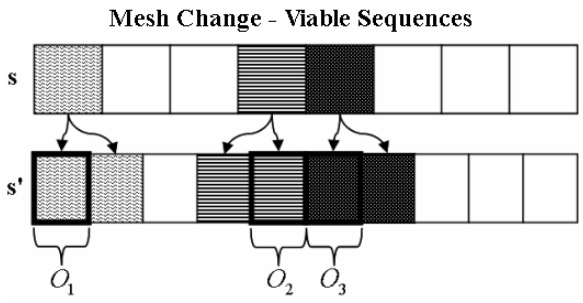


Figure 3: An example of a viable proposal configuration in the mesh-change scenario. The heavy black boundaries identify the set of proposed flaws with the greatest overlap with the current state.

In Figure 3, the elements in **s'** with a heavy black border exhibit the greatest degree of overlap with **s** and comprise the location sequence with the highest proposal probability. The associated moduli, $\hat{\boldsymbol{\upsilon}}'$, are proposed independently for each location $\hat{\omega}'_j$ according to the $\delta$-criterion.

*Transition Type 4.* **s** *and* **s'** *are of different levels and different meshes.*
The proposal algorithm for varying both the level and mesh is the combination of operations

developed in transition scenarios 2 and 3. The mesh change is always performed at the lower level (e.g., a mesh change is followed by a level increase). Since two operations take place

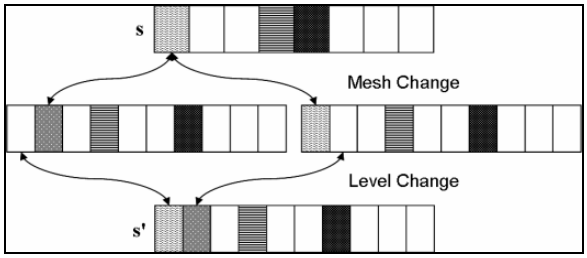


Figure 4: An example of a viable proposal configuration with a mesh change and level increase. Two possible intermediate configurations are shown.

serially in this scenario, there may frequently be multiple intermediate configurations that can produce a specific proposal configuration. Figure 4 illustrates this, where the mesh change is implemented prior to increasing the configuration level. Because the leftmost flaw in the current state, **s**, overlaps two flaws in the proposal configuration, **s'**, there are actually two intermediate configurations that can ultimately lead to the same proposal. This potential for multiple intermediate states demands that we sum the transition probability over all of these states.

In the above transitions, the Young's modulus value frequently changes stochastically and independently at a flawed element according to the parameter $\delta$. Now we will present this concept more formally. Suppose the modulus at a particular flawed element of **s** is $E$, which is the $u$-quantile of the prior distribution for the flawed condition. We wish to transform $u \in (0,1)$ to a value $u'$ distributed uniformly within the range $(u-\delta,\, u+\delta)$. Under circumstances where these intervals only partially overlap, any portion of $(u-\delta,\, u+\delta)$ lying outside (0, 1) is reflected back into (0, 1), doubling the density for values sufficiently close to the reflection points. The resulting density for movement to $u'$ from $u$ is as follows:

$$g_u(u') = \begin{cases} \frac{1}{2\delta} & \text{if } |u-\delta| < u' < 1-|1-u-\delta| \\ \frac{1}{\delta} & \text{if } u+u' < \delta \text{ or } 2-u-u' < \delta. \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

It can be easily shown that $g_u$ is symmetric with respect to $u$ and $u'$, i.e., we have the general relationship $g_u(u') = g_{u'}(u)$. The proposal modulus $E'$ is obtained from $u'$ by transformation (1) and hence the density for this proposal is $g_E^*(E') = g_u(u') f_1(E')$. Note that the key ratio $r(\mathbf{s}, \mathbf{s}')$ of (2) depends upon $g_E^*$ only through $g_u$. User selection of the modulus step size $\delta$ plays a critical role in *mixing efficiency*, as appropriately small or large modulus changes can help the sampler better explore and characterize few or many modes of the likelihood surface.

***The Likelihood Function***

The likelihood for a configuration is a measure of the discrepancy between the set of observed beam displacements and the corresponding set of expected (fitted or forward model) displacements, taking into account the finite-element model fit variation with respect to the variability in unflawed element moduli and measurement error. The total number of these displacements, $T$, is the product of the number of nodes ($n$), the number of displacement measurements per node ($d$), and the number of applied forces ($\eta$).

The finite-element model variation for a particular configuration that is induced by varying the unflawed moduli can be estimated by a discretized representation that approximates the configuration according to a binning of the flawed moduli. We partition the range of flawed modulus values into $M$ equally probable ranges according to the prior density $f_1$; e.g., if $M = 5$, bin 2 represents the 20th to the 40th percentiles of the distribution. Then for every possible configuration defined by some combination of mesh, level, flawed element locations, and the $M^k$ modulus bin assignments for the given location set, we characterize the model variation by executing a finite element code (NIKE3D, a 3D nonlinear, implicit code) forward model random sample: fix the flawed moduli at nominal values (medians within the respective bins) and vary the moduli for the unflawed elements according to the unflawed modulus prior density, $f_0$. The sample runs generate expected displacement components. Summary statistics of the runs generate an estimated bias and variance for each *configuration bin*, whose correspondence with a configuration is determined by the set of modulus bins containing the flawed element moduli.

To minimize the time and storage restrictions caused by combinatoric expansion, we limit $k$ to 2. A feasible alternative for $k > 2$ is to average the biases and variances for all $k(k-1)/2$ flaw pairs. Specifically, consider a configuration **s** at level $k > 2$, and let $\mu_{\mathbf{s}}$ denote the forward model fit for the nominal state of nature defined by using the unflawed modulus median, $E_0$, for each unflawed element, along with the flawed moduli $E_{i_j}$, $j = 1,...,k$, of **s**. Furthermore, for each displacement $t$, $t = 1,...,T$, let $\left(B_{\mathbf{s}t}, S^2_{\mathbf{s}t}\right)$ denote the averaged pairwise (bias, variance) described above for appropriate modulus bins and flawed element location pairs. Then the measured $t^{th}$ displacement for an arbitrary state of nature (based upon a random set of unflawed moduli) within configuration **s** has estimated mean and variance

$$\xi_{\mathbf{s}t} = \mu_{\mathbf{s}t} + B_{\mathbf{s}t} \qquad\qquad \tau^2_{\mathbf{s}t} = S^2_{\mathbf{s}t} + \sigma^2_t, \tag{6}$$

where $\sigma^2_t$ denotes the measurement error variance (assumed known in applications). Based on normality assumptions and displacement independence, the likelihood function for **s** is given by

$$L_{\mathbf{s}}(D) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\tau_{\mathbf{s}t}} \exp\left(-\frac{1}{2\tau^2_{\mathbf{s}t}}\left(D_t - \xi_{\mathbf{s}t}\right)^2\right). \tag{7}$$

***Example Simulations***

We carried out a number of simulations in order to assess the performance of our algorithms, and in every case the posterior distribution of the location and stiffness of flaws concentrated at the truth. As an example, data were simulated from the fine mesh model (10 elements), with Young's modulus values (1.895453e+11, 1.955167e+11, 1.379704e+11, 1.077513e+11, 1.934691e+11, 1.336336e+11, 1.839496e+11, 1.996551e+11, 1.979272e+11, 1.820129e+11), respectively. We use the trapezoidal priors of Figure 1, which assume a stiffness threshold of 1.80e+11; thus our example is a configuration with 3 flaws, in positions 3, 4, and 6. Expressed as percentiles of the prior distribution, these flawed element moduli are 63.96, 23.67, and 58.18, respectively. The data consisted of vertical and angular displacements measured at the nodes, based upon three static forces independently applied at the beam center and/or free end. (Note that we assume that both vertical and angular measurements are available. However, in practice, angular displacements may be difficult to obtain. Under these circumstances, forces can be incrementally applied with small changes to the linearly elastic model in the tangential direction, resulting in primarily normal displacements that can then be easily measured.) Random Gaussian measurement error (noise) with noise-to-signal ratio $\sigma = 0.01$ was added to the finite-element model displacements to create the simulated data. The M-H sampler was initialized at a random configuration and proceeded to sample the posterior distribution of flaw locations and moduli. Both the fine and coarse meshes were assigned prior probability 0.5.

The results for a representative run of 5000 iterations are summarized in Figures 5 and 6. We have conservatively used 1000 as the burn-in length; therefore the figures pertain to the last 4000 iterations. Figure 5 shows that accurate configurations, with respect to the mesh, level, and flaw locations, account for 95.4% of the posterior probability. The remaining 4.6% is distributed among fine mesh configurations that include the truth and coarse mesh configurations that heavily overlap the truth.

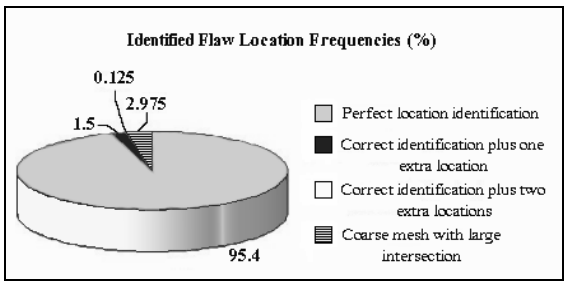


Figure 5: Performance of the posterior sampler in locating the flaws in a three-flaw configuration simulated from the fine mesh, assuming equal priors for the fine and coarse meshes.

Figure 6 summarizes the distribution of sampled flawed modulus values for the 3816 samples that correctly identified the flaw locations. As a precision metric, we use the average modulus error for the 3 flaws, as measured in percentile points of the prior distribution of flaw modulus. The figure shows that the sampler concentrated around the true stiffness values, such that the average stiffness error never exceeded 15 percentile points, and approximately 30% of the samples had absolute stiffness errors of only one percentile point. Clearly the sampler has succeeded in discovering not only the flaw locations, but also the degree of severity of the flaws.

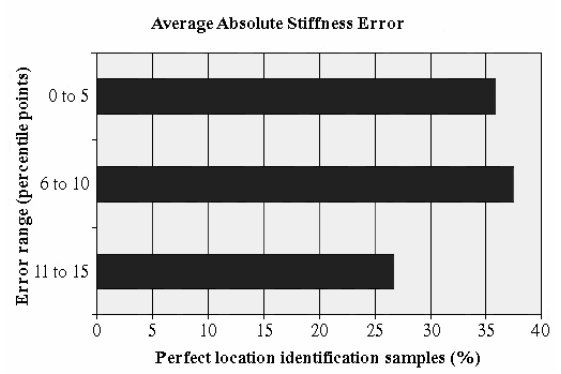


Figure 6: Distribution of the average modulus error for the samples that correctly identified the flaw locations in the scenario presented in Figure 5.

***Issues Regarding Sampler Performance***

To analyze sampler performance more efficiently, we limit the set of possible configurations to $2n^2 + 1$ by allowing, at most, two flaws in a configuration and by restricting the Young's modulus of an element to one of the three values, $e_0$, $e_1$, $e_2$, where $e_0$ is an unflawed modulus value, and $e_1$ and $e_2$ are flawed modulus values. The proposal function used in these simulations is not quite the discrete analog of those outlined above for the continuous parameterization. We use a function (Glaser *et. al.* 2003) which effects a reversible Markov chain sampler of the prior, in the spirit of Mosegaard (Mosegaard 1998), so that the posterior sampler depends entirely upon the likelihood ratio for its acceptance of proposals.

***Convergence of the Sampler***

For the discretized setting induced by three possible moduli, the posterior probability for a configuration **s** may be expressed according to Bayes' rule as

$$\rho_{\mathbf{s}} = P(\mathbf{s} \mid D) = \frac{\pi_{\mathbf{s}} L_{\mathbf{s}}(D)}{\sum_{j=1}^{N} \pi_j L_j(D)} , \tag{8}$$

where $N$ is the number of configurations. Thus the posterior distribution can be evaluated analytically given the prior distribution and likelihood function parameters. As a result, we may compare sampled versions of $\rho_{\mathbf{s}}$ (rho) with their analytic counterparts of (8). Our simulations have clearly demonstrated that we are sampling the posterior. We have typically used simulations of ten million Markov chain steps, following a burn-in period of ten thousand steps. Without exception, we have observed agreement to four significant figures between empirical simulated relative frequencies and corresponding theoretical posterior probabilities as defined by (8).

***Noise in the Data and Selection of the Prior***

The ability of the sampler to identify the underlying configuration depends primarily upon the degree of variation in the observed measurements relative to the theoretical values predicted by the equations of motion. To effect an easily interpreted noise relationship, we modify the

additive measurement error model implicit in (7) to a multiplicative version $X_{jt} = X_{jt}^{0}\left(1+\sigma Z_{jt}\right), t = 1,...,nd$, where $X^{0}$ represents the mean displacement in the equation of static deflection, $X$ is the measured displacement, and $j$ is the applied force index. The $Z$'s are independent standard normal variates, where the scalar parameter, $\sigma$, is effectively the noise-to-signal ratio. A value of $\sigma = 0.10$ implies that the standard deviation of the measured displacement for a given component is 10% of the mean displacement as given by the equations of motion. The parameter $\sigma$ is propagated in the likelihood function and becomes crucial in defining the posterior distribution, $\rho$, as well as its exploration by the sampler. In effect, $\sigma$ competes with the prior distribution $\pi$ in characterizing the posterior distribution. For sufficiently small $\sigma$, the observed data outweigh prior considerations, and the posterior probability is concentrated at the actual configuration, whereas, by the Bayesian paradigm, prior beliefs must carry increased weight if the observed data is sufficiently noisy to induce ambiguity. In such noisy scenarios the resulting posterior is a blend of the prior and the data that may suggest a number of plausible configurations.

Three noise levels for measured displacements are considered for the following examples, which use only the fine mesh beam model: low noise ($\sigma = 0.01$); medium noise ($\sigma = 0.05$); and high noise ($\sigma = 0.10$). The possible modulus values are ($e_0$, $e_1$, $e_2$) = (2.0e+11, 7.0e+10, 1.0e+10). For brevity, we refer to flawed modulus values $e_1$ and $e_2$ as type 1 and type 2, respectively, and specific flaws will be expressed by the ordered pair (flaw location, flaw type).

The subsequent figures combine contributions from both flaw types in summary three-dimensional portrayals of the posterior distribution. Recall that there are $N = 2n_f^2 + 1 = 201$ possible configurations for a fine mesh beam with two flaw types, assuming there are at most two flawed elements in the beam. There are 56 coordinates in each graph, where the leftmost corner position represents the zero-flaw case, the 10 points along the diagonal represent the various flaw locations for the one-flaw case, and the remaining 45 points represent the different flaw location pairs for the two-flaw case. At each coordinate is a cone whose height equals the total posterior probability, expressed as a percentage, for the configuration associated with the designated flaw location(s).

We illustrate the effects of noise in Figures 7 and 8. In Figure 7, the actual simulated configuration has a single flawed element, namely (3, 2). For low noise, we see a cone with height 100 at the coordinate representing a single flaw at element number 3. This tells us that the posterior distribution has apportioned probability 1 between the configurations (3,1) and (3, 2). In fact, though not shown in the graph, the posterior has correctly placed probability 1 at (3, 2) and probability zero at (3, 1). We have observed that, in low-noise conditions, any significant posterior probability is concentrated at only one of the flaw type possibilities for a given flaw location description. However, in high-noise scenarios, we have observed cases in which significant posterior probability has been spread between two different stiffness combinations.

The simulation results displayed in the top row of Figure 7 show the overwhelming effectiveness of the sampler, regardless of the degree of noise, when the prior favors the truth, in this case, the fact that the beam contains exactly one flawed element. This prior is characterized by $(\lambda_0, \lambda_1, \lambda_2) = (0.1, 0.8, 0.1)$. In the bottom row of Figure 7, however, we see the effect of a prior

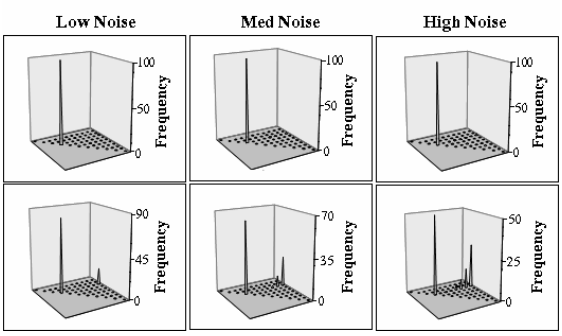


Figure 7: The effects of noise on the posterior distribution for a one-flaw scenario in which the prior favors one flaw (top) and two flaws (bottom). The flawed element is at position 3. The noise levels are $\sigma$ = 0.01, 0.05, and 0.10.

which favors configurations contrary to truth. Once again the one-flaw configuration (3, 2) represents truth, but now the assumed prior distribution, $(\lambda_0, \lambda_1, \lambda_2) = (0.02, 0.23, 0.75)$, favors configurations of two flaws. The observed data in this scenario are not sufficiently compelling to dismiss some two-flaw configurations. The Bayesian paradigm conducts a competition between observed data and prior beliefs; i.e. noisier data is less influential in amending the prior. At each noise level in this example, the greatest probability is placed at the actual configuration, (3, 2). In addition, there is significant probability placed at a number of two-flaw configurations which include (3, 2). If this were a real-life application, we would conclude that there is clearly a flaw of type 2 at element 3. Moreover, there could be an additional flaw at a location near the end of the beam that might warrant further investigation.

Figure 8 presents examples based upon the configuration {(3, 2), (4, 1)}. The results are similar to those of the single-flaw scenarios shown in Figure 7. In the low-noise setting, the data overwhelms the prior and the actual configuration is discovered, while for increased noise, there exists greater posterior uncertainty. The effect of a prior contrary to truth, in this case a prior favoring a single flaw, is to put significant posterior probability at a single-flaw subset of the truth, (3, 2) in the high-noise case.

The preceding examples demonstrate that the M-H sampler can readily determine the actual flaw configuration in the low-noise environment. In noisy situations, it will formulate a posterior distribution that heavily supports the actual configuration while allowing for some alternative plausible configurations, motivated to some extent by prior beliefs.

***Missing Data***

In practice, whether by accident or design, data may not be collected or available at each node. For the M-H sampler, missing data does not, in general, present an impasse. The posterior distribution can still be sampled or calculated, but it tends to display increased uncertainty as the amount of missing data increases. The calculation of the posterior by MCMC sampling, or by analytic means, merely requires a modification of the likelihood function to include only those terms *jt* for which measured data $D_{jt}$ is available.

We have examined the relationship between the quantity of observed data and the quality of prediction, as measured by the spread of the posterior distribution, for the examples depicted with complete data in Figure 8 of the previous section. We now consider incomplete data in

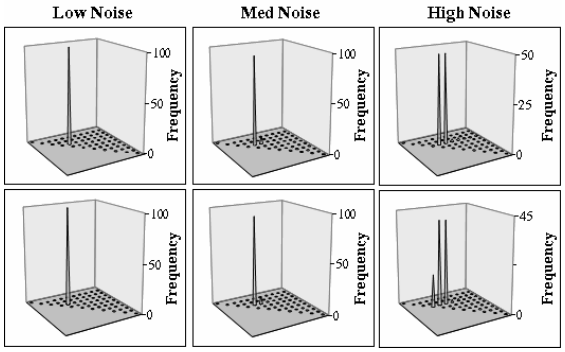


Figure 8: The effects of noise on the posterior distribution for a two-flaw scenario in which the prior favors two flaws (top) and one flaw (bottom). The flawed elements are at positions 3 and 4. The noise levels are $\sigma = 0.01$, 0.05, and 0.10.

which all $d\eta = 6$ measurements are absent for each node within a selected subset of nodes. Our results are shown in Figures 9, 10, and 11 for low, medium, and high noise, respectively. In the low-noise setting there is negligible loss of precision when up to 4 of the 10 available nodes

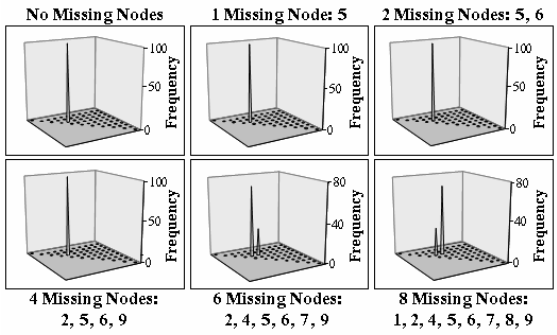


Figure 9: The effect of missing data on the posterior distribution for a low-noise ($\sigma = 0.01$) two-flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

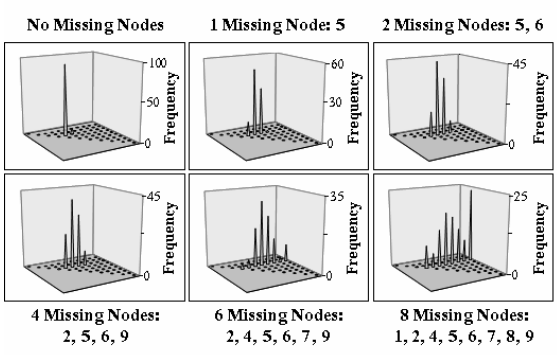


Figure 10: The effect of missing data on the posterior distribution for a medium-noise ($\sigma = 0.05$) two-flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

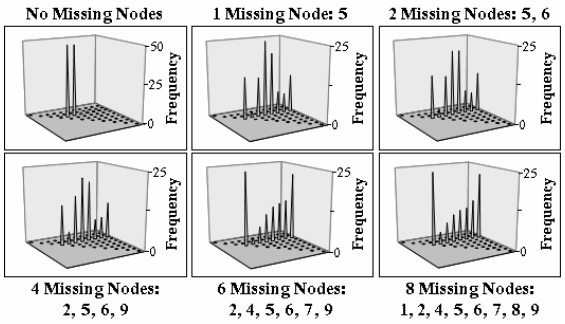


Figure 11: The effect of missing data on the posterior distribution for a high-noise ($\sigma = 0.10$) two-flaw scenario in which the prior favors two flaws. The flawed elements are at positions 3 and 4. The nodes without data are as noted.

remain unmeasured, and respectable performance when 6 nodes are not measured. In the medium- and high-noise settings, however, there is an immediate degradation in the ability to correctly identify both flaws: although flaw (3, 2) is correctly identified, significant posterior probability is placed at configurations away from the other actual flaw, (4, 1). We have conducted alternative simulations with the same amounts of missing data, but with different selections of missing nodes. The results are nearly identical to those shown here, indicating that the number of missing nodes is more important than their specific locations.

***Mesh Coarseness***

The accuracy of a finite-element model depends heavily upon mesh selection, since homogeneity is assumed within each element. In the following examples, we examine simulation results from samplers pitting the fine and coarse meshes, each with prior probability 0.5. The coarse mesh model has $N = 2n_c^2 + 1 = 129$ possible configurations if we assume, as in the fine mesh case, that there can be no more than two flawed elements. We assume the measured displacements are at nodes of the model generating the data. If, for example, the data are generated from the fine mesh model, in order to evaluate the coarse likelihood we must obtain coarse mean displacements at nodes of the fine mesh. Our approach is to linearly interpolate the coarse model mean displacements, while selecting applied forces for the simulated data that are restricted to the common nodes, namely locations 5 and 10 (see Figure 2).

We have engaged the competing mesh sampler in a variety of scenarios with the previous noise levels and discovered that the posterior probability distribution concentrates itself almost entirely on the mesh that generates the simulated data. This is not terribly surprising, because the data is fitted best by the mesh geometry upon which it was created, and any configuration of the opposite mesh suffers from the drawback that it is necessarily an incorrect model. Moreover, the fact that the opposite mesh stiffness matrices produce mean displacements at nodes that coincide with the data nodes in only two places require us to interpolate these displacements to estimate mean displacements at a majority of the data mesh nodes. This (linear) interpolation may be sufficiently poor that no opposing mesh configuration can realistically compete with the data mesh configurations. A fair competition would be to simulate data from a model that is not correctly represented by any configuration in either mesh.

To see how the coarse mesh estimates the flaw properties of the beam, we ran the sampler restricted to the 129 coarse mesh configurations, invoking the bounding case $\pi^c = 1$. The results

are displayed in Figure 12, using the underlying one- and two-flaw models of Figures 7 and 8. It is noteworthy that the coarse posterior places essentially all of its probability on a single configuration for each noise level. In the left-hand figure, we have in truth a fine mesh with a

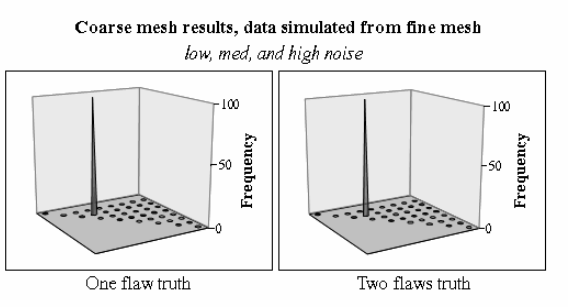


Figure 12: Coarse mesh posteriors for the one-flaw scenario of Figure 7 (left) and the two-flaw scenario of Figure 8 (right) in which the data are simulated from the fine mesh. The results are the same for all the noise levels, $\sigma = 0.01$, 0.05, and 0.10.

single flaw at (fine 3, 2). The coarse posterior places all its probability at the two-flaw configuration, {(coarse 2, 1), (coarse 3, 2)}, because fine element 3 is bisected by coarse elements 2 and 3. In the right-hand figure, we have a fine mesh with two flawed elements, namely {(fine 3, 2), (fine 4, 1)}. The coarse posterior puts all its probability in this case at the two-flaw configuration, {(coarse 2, 1), (coarse 3, 2)}. This pairing is closest to the truth in terms of mesh overlap, and the flaw types seem to have been adjusted in some compensatory fashion.

***Violation of Prior Assumptions***

The M-H sampler is a Bayesian algorithm that characterizes the posterior distribution, a synthesis of the prior distribution and the data. It is mathematically impossible for the posterior to place positive probability on a configuration that has been given zero probability in the prior. Hence if we postulate a prior that puts all its probability on configurations with no more than two flawed elements, the posterior will necessarily put all its probability on configurations with no more than two flawed elements, despite any evidence in the data that supports the case for more than two flaws.

| True Flaw Configuration | Mesh | Posterior Mode |
|---|---|---|
| (3,1), (5,1), (10,2) | fine | (3,1), (5,1) |
| (3,2), (5,2), (9,1), (10,2) | fine | (3,2), (5,2) |
| (1,2), (5,1), (8,1), (9,1), (10,2) | fine | (1,2), (5,1) |
| (5,1), (6,2), (7,1) | coarse | (5,1), (6,2) |
| (4,1), (5,1), (7,1), (8,1) | coarse | (4,1), (5,1) |
| (4,2), (5,1), (6,1), (7,2), (8,1) | coarse | (4,2), (7,2) |

Table A. Posterior distribution for six examples of a prior that does not include the true configuration.

In Table A, we investigate six examples of the effect of a prior that does not include the truth. In each case we simulate displacement data for a beam with more than two flawed elements and see what posterior distribution the sampler develops, given the prior assumption that there can be no more than two flaws. Interestingly, the posterior distribution in each case puts probability essentially 1 at a single two-flaw configuration.

In reviewing the results in the Table A, we see that for all but the last case, the posterior selects the two flaws nearest the fixed end from among the actual flawed elements. This makes intuitive sense because a flawed element affects displacements of elements to its right (toward the free end) more than elements to its left (toward the fixed end). The exception is the last case, which opts for the leftmost flawed element but skips two type 1 flaws and selects a type 2 flaw. This may be explained by the fact that the type 2 modulus suggests a more severe flaw.

It is perhaps most noteworthy that the posteriors are virtually degenerate, placing probability nearly 1 at a single configuration. This may be explained by the nature of the likelihood function, whose quadratic term penalizes large deviations from the truth. As a consequence, in our scenario in which each candidate configuration is incorrect, we have a collection of widely varying but tiny likelihoods, of which one dominates. In general, a degenerate posterior is cause for elation, because it indicates the true state of nature has been discovered. However, as we see in these examples, such an assertion is meaningless if we undermine the Bayesian paradigm by positing a prior that excludes the truth.

**Conclusions**

A stochastic simulation methodology for the identification of mechanical and structural systems has been presented. The methodology is a M-H implementation of a Bayesian MCMC algorithm. Using a finite-element model of a uniform fixed-free linearly elastic cantilever beam as the forward model, the algorithm is shown to identify probable configurations of the beam by estimating the stiffness of each beam element through an identification of its elastic modulus. System inputs are static forces applied to the nodes of the beam; associated measured deflections constitute the data. Each beam element may have a modulus value within a range, including flawed and unflawed sub-ranges. Results show that the methodology successfully calculates posterior probability distributions across the possible configurations of elastic moduli of the beam elements. The distributional information in turn effectively identifies flaws and their severity by providing a probability measure for plausible configurations of flawed elements and their modulus values. Furthermore, the methodology is shown, in the simpler, discretized (three possible modulus values) example cases, to be robust to the presence of noise in the data, missing data, and competing mesh model descriptions.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

**Notation**

$a_{ij}$ = prob. of changing level $i$ to $j$, $a_{ii} \in [0,1)$; $a_{ij} \in (0,1)$ if $|i-j|=1$, $0 \le i, j \le K$;

$B_{\mathbf{s}t}$ = average bias for all flaw pairs in configuration $\mathbf{s}$ for the $t^{th}$ displacement;

$b, b'$ = probability of state transition in neighbor-only setting (see Appendix);

$D_t$ = $t^{th}$ displacement measurement of $T$ total displacements;

$d$ = number of displacement measurements per node;

$\tilde{E}_j$, $\tilde{E}'_j$ = Young's modulus value at element $j$ in a configuration $\mathbf{s}, \mathbf{s'}$;

$\mathbf{F}_j$ = $j^{\text{th}}$ force vector applied to the beam;

$f_1$ = modulus density for the flawed condition;

$\mathbf{K}$ = $nd \times nd$ stiffness matrix for the equation of static deflection;

$k$ = configuration level, $k = 1, \ldots, K$, $K \le \min\{n_c, n_f\}$;

$L_{\mathbf{s}}(D)$ = likelihood of the data, $D$, given configuration $\mathbf{s}$ as truth;

$m, m'$ = mesh associated with a configuration $\mathbf{s}, \mathbf{s'}$, $m, m' \in \{n_c, n_f\}$;

$n_c, n_f$ = number of elements in the coarse / fine mesh beam model;

$O_j$ = overlap of elements $\omega_j$ and $\omega'_j$ when mesh change is proposed;

$p_{mm'}$ = proposal probability of moving from mesh $m$ to mesh $m'$, $p_{mm'} \in [0,1]$;

$S^2_{\mathbf{s}t}$ = average variance for all flaw pairs in configuration $\mathbf{s}$ for the $t^{th}$ displacement;

$\mathbf{s}, \mathbf{s'}$ = configurations, defined by a 3-tuple $\mathbf{s} = (\boldsymbol{\omega}, \boldsymbol{\upsilon}, m)$; $\mathbf{s'}$ is proposed by $q(\mathbf{s}, \mathbf{s'})$;

$\mathbf{s'}(l)$ = configuration $\mathbf{s'}$ following removal of the flaw at location $l$;

$u_j$, $u'_j$ = quantile at modulus value $\tilde{E}_j$, $\tilde{E}'_j$;

$\mathbf{X}^0_j$ = $nd \times 1$ theoretical displacement vector induced by applied force $\mathbf{F}_j$;

$x(\hat{\boldsymbol{\omega}}', j)$ = location of $\mathbf{s'}$ that matches up with the $j^{\text{th}}$ element of $\hat{\boldsymbol{\omega}}'$;

$\alpha, \alpha'$ = unshared flawed element location in $\mathbf{s}, \mathbf{s'}$;

$\beta$ = proposal probability of neighbor-only location change at level 1, $\beta \in [0,1]$;

$\gamma^{\mathbf{s}}_{\hat{\boldsymbol{\omega}}'}$ = probability of selecting sequence $\hat{\boldsymbol{\omega}}'$ from the set $\Omega_{\mathbf{s}}$;

$\delta$ = maximum modulus step size, measured in probability, $\delta \in (0, \frac{1}{2})$;

$\eta$ = number of independent static nodal forces applied to the beam;

$\theta_k$ = proposal probability of a modulus-only change at level $k$, $\theta_k \in [0,1)$;

$\lambda_k$ = prior probability of level $k$, $\lambda_k \in (0,1)$;

$\mu_{\mathbf{s}}$ = forward model fit for $\mathbf{s}$, using modulus median for each unflawed element;

$\xi_{\mathbf{s}t}$ = estimated mean of the $t^{th}$ displacement for a random state of nature within $\mathbf{s}$;

$\pi^m$ = prior probability of mesh $m$, $\pi^m \in [0,1]$;

$\pi_{\mathbf{s}}$ = prior density for configuration $\mathbf{s}$;

$\rho_{\mathbf{s}}$ = posterior probability for configuration $\mathbf{s}$;

$\sigma^2_t$ = measurement error variance for $t^{th}$ displacement;

$\tau^2_{\mathbf{s}t}$ = estimated variance of the $t^{th}$ displacement for a random state of nature within $\mathbf{s}$;

$\boldsymbol{\upsilon}, \boldsymbol{\upsilon}'$ = set of quantiles $\upsilon_j, \upsilon'_j$ corresponding to moduli at flaw locations $\omega_j, \omega'_j$;

$\hat{\boldsymbol{\upsilon}}'$ = vector of moduli corresponding the flaw locations in the proposed sequence $\hat{\boldsymbol{\omega}}'$;

$\Omega_{\mathbf{s}}$ = set of viable position sequences in the opposite mesh generated by $\mathbf{s}$;

$\boldsymbol{\omega}, \boldsymbol{\omega}'$ = naturally ordered set of flawed element locations in configuration $\mathbf{s}, \mathbf{s}'$;

$\hat{\boldsymbol{\omega}}', \hat{\boldsymbol{\omega}}$ = proposed flaw location sequence from the set $\Omega_{\mathbf{s}}, \Omega_{\mathbf{s}'}$.

**References**

Aines, R., Nitao, J., Newmark, R., Carle, S., Ramirez, A., Harris, D., Johnson, J., Johnson, V., Ermak, D., Sugiyama, G., Hanley, W., Sengupta, S., Daily, W., Glaser, R., Dyer, K., Fogg, G., Zhang, Y., Yu, Z., and Levine, R. (2002). "The Stochastic Engine Initiative: Improving Prediction of Behavior in Geologic Environments We Cannot Directly Observe." Lawrence Livermore National Laboratory, Livermore, CA, Report UCRL-ID-148221.

Beck, J. L. (1989). "Statistical System Identification of Structures." *Proc., 5th Int. conf. Structural Safety and Reliability*, ASCE, New York, 1395-1402.

Beck, J. L., and Au, S-K. (2002). "Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation." *J. Eng. Mech.,* 128(4), 380-391.

Beck, J. L., and Yuen, K-V. (2004). "Model Selection using Response Measurements: Bayesian Probabilistic Approach." *J. Eng. Mech.,* 130(2), 192-203.

Collins, J. D., Hart, G. C., Hasselman, T. K., and Kennedy, B. (1974). "Statistical identification of structures." *AIAA J*., 12, 185-190.

Doebling, S. W., Farrar, C. R., Prime, M. B., and Shevitz, D. W. (1996). "Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in their Vibration Characteristics: A Literature Review." *Tech. Rep. LA-13070-MS,* Los Alamos National Laboratory, Los Alamos, NM.

Friswell, M. I., and Mottershead, J. E. (1995). *Finite-element model updating in structural dynamics,* Kluwer Academic Publishers, Boston, MA.

Glaser, R. E., Hanley, W. G., Lee, C. L., and Nitao, J. J. (2003). "A Markov Chain Monte Carlo Based Method for System Identification." Lawrence Livermore National Laboratory, Livermore, CA, Report UCRL-JC-150494.

Hastings, W. K. (1970). "Monte Carlo Sampling Methods using Markov Chains and their Applications." *Biometrika*, 57, 97-109.

Karlin, S. and Taylor, H. M. (1975). *A First Course in Stochastic Processes,* Academic Press, San Diego, CA.

Mosegaard, K. (1998). "Resolution Analysis of General Inverse Problems through Inverse Monte Carlo Sampling." *Inverse Problems*, 14, 405-426.

Metropolis, N., Rosenbluth, M. N., Teller, A. H., Teller, E. (1953). "Equations of State Calculations by Fast Computing Machines." *J of Chemical Physics*, 21, 1087-1092.

Newmark, R., Aines, R., Nitao, J., Hanley, W., Carle, S., Ramirez, A., Sengupta, S., Harris, D. (2002). "Stochastic Engine: Direct Incorporation of Measurements Into Predictive Simulations." Lawrence Livermore National Laboratory, Livermore, CA, Report UCRL-JC-145116.

Ramirez, A. J., Nitao, J. J., Hanley, W. G., Aines, R., Glaser, R. E., Senguta, S. K., Dyer, K. M., Hickling, T. L., and Daily, W. D. (2005). "Stochastic Inversion of Electrical Resistivity Changes using a Markov Chain Monte Carlo Approach." *J. Geophys. Res.*, 110, B02101.

Sohn, H., and Law, K. H. (1997). "A Bayesian Probabilistic Approach for Structure Damage Detection." *Earthquake Eng. Struct. Dyn.,* 26, 1259-1281.

Vanik, M. W., Beck, J. L., and Au, S. K. (2000). "Bayesian Probabilistic Approach to Structural Health Monitoring." *J. Eng. Mech.,* 126(7), 738-745.

Yuen., K-V. and Katafygiotis, L. S. (2002). "Bayesian Model Updating using Complete Input and Incomplete Response Noisy Measurements." *J. Eng. Mech.,* 128(3), 340-350.

**Appendix**

*Transition Type 1.* **s** *and* **s'** *are of the same level and mesh.*
The proposal density for the modulus-only transition is given by

$$q(0,0) = a_{00}; \quad q(\mathbf{s},\mathbf{s'}) = a_{kk}\, p_{mm'} \theta_k \prod_{j=1}^{k} g^{*}_{\tilde{E}_{\omega_j}}\left(\tilde{E}'_{\omega_j}\right). \tag{9}$$

The proposal density for a single flaw location change when $k > 1$ is given in Eq. (10). It can be shown that the proposals described by equations (9) and (10) are both reversible, i.e., $h(s,s') = 1$.

$$q(\mathbf{s},\mathbf{s'}) = a_{kk}\; p_{mm'} (1-\theta_k) \tfrac{1}{k} \tfrac{1}{n_m - k} f_1\left(\tilde{E}'_{\omega'_{\alpha'}}\right). \tag{10}$$

When $k = 1$, the proposal density for relocating the single flaw is given in Eq. (11), where $b$, given in Eq. (12), is the neighbor selection probability, and $\alpha'$ (alpha) denotes the flaw location in **s'** which is not shared by **s**. Note that $\beta$, the probability of a neighbor-only transition, is used as a mixing parameter to control the magnitude of the change in proposal state, so we have

$$q(\mathbf{s},\mathbf{s'}) = a_{11}\, p_{mm'} (1-\theta_1)\left(\beta b g^{*}_{\tilde{E}_{\omega_\alpha}}\left(\tilde{E}'_{\omega'_{\alpha'}}\right) + (1-\beta)\tfrac{1}{n_m - 1} f_1\left(\tilde{E}'_{\omega'_{\alpha'}}\right)\right), \tag{11}$$

And

$$b = \begin{cases} \tfrac{1}{2} & \text{if } 2 \le \omega_\alpha \le n_m - 1 \\ 1 & \text{otherwise} \end{cases}. \tag{12}$$

It can be shown that $h(s,s') = 1$ if and only if the flaw is an interior element of **s** and **s'**.

*Transition Type 2.* **s** *and* **s'** *are of different levels but of the same mesh.*
The proposal densities for transitions with a level change are given by

$$\begin{aligned} q(0,\mathbf{s'}) &= a_{01} \pi^m \tfrac{1}{n_m} f_1\left(\tilde{E}'_{\omega'_{\alpha'}}\right) \\ q(\mathbf{s},0) &= a_{10} \\ q(\mathbf{s},\mathbf{s'}) &= a_{k,k+1}\; p_{mm'} \tfrac{1}{n_m - k} f_1\left(\tilde{E}'_{\omega'_{\alpha'}}\right) \\ q(\mathbf{s'},\mathbf{s}) &= a_{k+1,k}\; p_{m'm} \tfrac{1}{k+1} \end{aligned}. \tag{13}$$

*Transition Type 3.* $\mathbf{s}$ *and* $\mathbf{s'}$ *are of the same level but of different meshes.*
The proposal density for a transition with a mesh change is given by

$$q(\mathbf{s},\mathbf{s'}) = a_{kk}\, p_{mm'} \gamma^{\mathbf{s}}_{\hat{\boldsymbol{\omega}}'} \prod_{j=1}^{k} g^{*}_{\tilde{E}_{\omega_j}}\left(\tilde{E}'_{\omega'_j}\right), \tag{14}$$

where $\gamma^{\mathbf{s}}_{\hat{\boldsymbol{\omega}}'}$ is the probability that $\hat{\boldsymbol{\omega}}'$ will be selected from the set of all viable sequences.

*Transition Type 4.* $\mathbf{s}$ *and* $\mathbf{s'}$ *are of different levels and different meshes.*
The proposal density for a transition including a mesh change and an increase in level is given by

$$q(\mathbf{s},\mathbf{s'}) = a_{k,k+1}\, p_{mm'} \sum_{\substack{\hat{\boldsymbol{\omega}}' \in \Omega_{\mathbf{s}} \\ *}} \gamma^{\mathbf{s}}_{\hat{\boldsymbol{\omega}}'} \left( \left[ \prod_{j=1}^{k} g^{*}_{\tilde{E}_{\omega_j}}\left(\tilde{E}'_{\omega'_{x(\hat{\boldsymbol{\omega}}',j)}}\right) \right] \frac{1}{n_{m'}-k} f_1\left(\tilde{E}'_{\boldsymbol{\omega}'\setminus\hat{\boldsymbol{\omega}}'}\right) \right), \tag{15}$$

where $x(\hat{\boldsymbol{\omega}}', j)$ denotes the location of $\mathbf{s'}$ which corresponds to the $j^{\text{th}}$ element of the proposed location sequence $\hat{\boldsymbol{\omega}}'$, and $\boldsymbol{\omega}'\setminus\hat{\boldsymbol{\omega}}'$ is a set difference operation referring to the flawed element location added during the level increase. Since there may be multiple intermediate states (after the mesh change) which can yield $\mathbf{s'}$ following the level increase, we sum over the viable location sequences which can ultimately lead to this successful transition. The proposal density for a transition including both a mesh change and a decrease in level is given in Eq. (16), where $\mathbf{s'}(l)$ denotes the configuration $\mathbf{s'}$ with flaw $l$ (ell) removed. Since there may be multiple elements which, when removed, can yield $\mathbf{s'}$ upon completion of the mesh change, we sum the proposal probability over these elements, which gives

$$q(\mathbf{s'},\mathbf{s}) = a_{k+1,k}\, p_{m'm} \sum_{\substack{l=1 \\ *}}^{k+1} \frac{1}{k+1} \gamma^{\mathbf{s'}(l)}_{\hat{\boldsymbol{\omega}}(l)} \prod_{j=1}^{k} g^{*}_{\tilde{E}'_{\omega'_j}|l}\left(\tilde{E}_{\omega_j}\right) \quad . \tag{16}$$