Computer-Aided Civil and Infrastructure Engineering **21** (2006) 242–257

Structural Model Updating and Health Monitoring with Incomplete Modal Data Using Gibbs Sampler

Jianye Ching

Department of Construction Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan

&

Matthew Muto & James L. Beck*

Department of Civil Engineering, Mail Code 104-44, California Institute of Technology, Pasadena, CA 91125, USA

Abstract: *A new Bayesian model updating approach is presented for linear structural models. It is based on the Gibbs sampler, a stochastic simulation method that decomposes the uncertain model parameters into three groups, so that the direct sampling from any one group is possible when conditional on the other groups and the incomplete modal data. This means that even if the number of uncertain parameters is large, the effective dimension for the Gibbs sampler is always three and so high-dimensional parameter spaces that are fatal to most sampling techniques are handled by the method, making it more practical for health monitoring of real structures. The approach also inherits the advantages of Bayesian techniques: it not only updates the optimal estimate of the structural parameters but also updates the associated uncertainties. The approach is illustrated by applying it to two examples of structural health monitoring problems, in which the goal is to detect and quantify any damage using incomplete modal data obtained from small-amplitude vibrations measured before and after a severe loading event, such as an earthquake or explosion.*

*To whom correspondence should be addressed. E-mail: *jimbeck@its.caltech.edu*.

1 INTRODUCTION

Model updating refers to the methodology that determines the most plausible structural model for an instrumented structural system given its measured response and, possibly, its excitation. In recent years, civil engineers have paid much attention to model updating techniques as they have broad applications in structural health monitoring (Natke and Yao, 1988; Hjelmstad and Shin, 1997; Lam et al., 1998; Beck et al., 2001; Chang, 2001; Sohn et al., 2001; Bernal et al., 2002; Casciati, 2002). Among the structural model updating techniques, Bayesian model updating techniques (Beck and Katafygiotis, 1998; Vanik et al., 2000; Yuen et al., 2004; Ching and Beck, 2004) do not just find a single plausible structural model but a set of structural models whose predictions are weighted by the probabilities of those models conditional on the measured data. Due to their ability to consider more than one structural model, Bayesian model updating techniques are robust and are suitable to characterize modeling uncertainties of structural systems.

In terms of structural health monitoring, linear structural models are often used for model updating (Vanik et al., 2000; Sohn and Law, 2001; Caicedo et al., 2004). For instance, much vibration data of structures under investigation are obtained using low-amplitude excitation, e.g., ambient vibration and hammer impact, under which

© 2006 *Computer-Aided Civil and Infrastructure Engineering*. Published by Blackwell Publishing, 350 Main Street, Malden, MA 02148, USA, and 9600 Garsington Road, Oxford OX4 2DQ, UK.

many structures (even damaged structures) behave approximately linearly. In this case, the usual approach is to first identify modal parameters from vibration data, update the stiffnesses of the structural model based on the identified modal data and detect damage locations by comparing the updated stiffness distribution with that from the baseline undamaged structure. Local loss of stiffness in the structural model is taken to indicate damage at the corresponding location in the real structure.

Even with linear models, the problem of model updating with modal data is potentially ill-posed, i.e., there may be more than one optimal model (Katafygiotis and Beck, 1998). The problem becomes even more challenging when only some of the degrees of freedom (DOF) of the model are measured and when modeling errors are explicitly acknowledged. A previous Bayesian structural model updating approach (Beck and Katafygiotis, 1998) has been successful in resolving the aforementioned difficulties when the amount of modal data is sufficiently large so that an asymptotic approximation of the Bayesian predictive integrals is accurate.

However, when the amount of data is smaller, the accuracy of the asymptotic approximation is questionable. Beck and Au (2002) proposed a stochastic simulation approach for these cases that is based on a Markov chain Monte Carlo algorithm (Metropolis-Hastings) that also addresses the aforementioned difficulties. Their approach does not use an asymptotic approximation and the results can be made arbitrarily accurate with increasing number of samples. However, a major limitation is that it is only efficient for lower-dimensional problems.

In this article, a new approach is presented for Bayesian model updating of linear structural models with incomplete modal data that is based on the Gibbs sampler (GS) (Geman and Geman, 1984; Gelfand et al., 1990), a special case of Markov chain Monte Carlo simulation. This approach also does not use any asymptotic approximation and the results can be made arbitrarily accurate with sufficient samples. The most attractive aspect of this new approach is that it is applicable to linear Bayesian model updating problems of arbitrarily high dimensions.

The basic idea of the GS approach presented here is to decompose the uncertain model parameters of the linear structural model into three groups so that exact sampling of one group of parameters is possible when conditional on the other groups and the modal data. The dimension of the uncertain model parameters is then not an issue because the effective dimension is always three. It is shown that the new approach works for globally identifiable and unidentifiable cases with high-dimensional uncertain model parameters. Moreover, the GS approach also estimates the complete mode shapes as well as prediction-error variances. However, it is found that it may not work for locally identifiable cases. The reader is referred to Beck and Katafygiotis (1998) for detailed definitions of globally and locally identifiable, and unidentifiable, problems; one can view these three cases as corresponding, respectively, to a unique maximum likelihood estimate (MLE), multiple but isolated MLEs and a manifold (i.e., continuum) of MLEs in the parameter space.

An obvious application of this new approach is Bayesian model updating for structural health monitoring with low-amplitude vibration data, such as acceleration data generated by ambient vibrations or hammer impacts before and after a severe loading event (e.g., earthquake or explosion) to identify possible damage, including its location and severity. Many structures behave approximately linearly under these weak vibrations, so the linearity assumption of the approach will be justified. Moreover, the GS approach is robust to the dimension of uncertain model parameters, so it may be applied to real structures.

The outline of the article is as follows: First, the Bayesian linear structural model updating problem is defined and the class of linear structural models is introduced. Second, the GS approach is introduced. An index that is useful for damage localization and quantification is then presented. The effectiveness of the new approach is then demonstrated using two structural health monitoring examples: a 2-DOF shear building and the IASC-ASCE Structural Health Monitoring Simulated Benchmark Phase II data. Computational issues for the GS approach are also addressed in the Appendix. Finally, the article concludes with a discussion of the advantages and disadvantages of the GS approach.

2 BAYESIAN LINEAR STRUCTURAL MODEL UPDATING

In this section, a Bayesian procedure is described for updating structural models based on incomplete modal data. It is assumed for the structure under consideration that N_s sets of experimental modal data have been obtained from the structure for N_m dominant modes of vibration using a reliable modal identification technique such as MODE-ID (Beck, 1996; Yuen et al., 2004). The incomplete modal data to be utilized consists of the modal frequencies and mode shape components at the measured DOF for each identified mode (identified modal damping ratios are not used in the damage assessment). The modal data are denoted by $\hat{D} \equiv \{\hat{\omega}_{r,j}, \hat{\psi}_{r,j} : r = 1, \ldots, N_m,\ j = 1, \ldots, N_s\}$, where $\hat{\omega}_{r,j}$ and $\hat{\psi}_{r,j} \in R^{N_O}$ (N_O is the number of measured DOF) are the experimental modal frequency and vector of observed mode-shape components for the rth mode in

the jth modal data set, respectively. The experimental mode shapes are normalized so that their Euclidean norm $\|\hat{\psi}_{r,j}\|_2 = 1$.

2.1 Linear structural identification model

To define the identification model class $\mathcal{M}$, a set of linear structural models with the mass matrix M and stiffness matrix K parameterized in an affine manner is defined as follows:

$$M(\rho) = M_0 + \sum_{i=1}^{N_M} \rho_i M_i \quad K(\theta) = K_0 + \sum_{i=1}^{N_K} \theta_i K_i \qquad (1)$$

where $N_d \geq N_O$ is the number of DOF of the identification model; $M_i \in R^{N_d \times N_d}$ and $K_i \in R^{N_d \times N_d}$ are the prescribed nominal contributions of the ith substructure to the global mass and stiffness matrices, and the uncertain parameters ρ_i and θ_i scale these contributions and are to be updated. Classical normal modes are assumed and thus the damping matrix need not be explicitly modeled because it does not change the eigenvectors specified by the mass and stiffness matrices.

2.1.1 System mode shapes. In the usual situation where the full DOF corresponding to the identification model are not measured, "system mode shapes" (Beck et al., 2001) may be introduced as extra parameters to be updated. These mode-shape components are not constrained to be eigenvectors of the structural model and are introduced to represent the actual underlying mode shapes of the system. The connection between the experimental modal data and the model parameters becomes

$$K(\theta)\phi_r = \hat{\omega}_{r,j}^2 M(\rho)\phi_r + \varepsilon_{r,j} \quad \hat{\psi}_{r,j} = \Gamma\phi_r + e_{r,j} \qquad (2)$$

where $\phi_r \in R^{N_d}$ is the system mode-shape vector of the rth mode; Γ is the matrix that picks the measured DOF from the system mode shape ϕ_r, $\varepsilon_{r,j} \in R^{N_d}$, and $e_{r,j} \in R^{N_O}$ are prediction-error vectors and are modeled as independent Gaussian variables, which can be justified by the maximum differential entropy principle (Jaynes, 1957), that is, the Gaussian probability density function (PDF) gives the largest uncertainty (or least amount of information in the sense of Shannon) for any PDF of unbounded support with specified means and variances,

$$\varepsilon_{r,j} \sim N\big(0, \sigma_r^2 I\big) \quad e_{r,j} \sim N\big(0, \delta_r^2 I\big) \qquad (3)$$

so that the means of the prediction-error vectors are zero and their covariance matrices are scaled versions of the identity matrix I of appropriate order. For notational simplicity, the vector of N_m system mode shapes is denoted by $\phi = [\,\phi_1^T\ \phi_2^T\ \phi_3^T \cdots \phi_{N_m}^T\,]^T$, the structural parameters by $\lambda = [\,\theta^T\ \rho^T\,]^T \in R^{N_\lambda}$, and the prediction-error variances by $\sigma^2 = [\,\sigma_1^2\ \sigma_2^2\ \sigma_3^2 \cdots \sigma_{N_m}^2\,]^T$ and $\delta^2 = [\,\delta_1^2\ \delta_2^2\ \delta_3^2 \cdots \delta_{N_m}^2\,]^T$. Furthermore, the symbol $\mathcal{M}$ for the identification model class will be omitted even though all the PDFs are obviously conditional on the choice of $\mathcal{M}$. Note that (2) and (3) define the PDF $p(\hat{D}\,|\,\lambda, \phi, \sigma^2, \delta^2)$.

The system mode shapes ϕ can be regarded as a bridge connecting the identification problem with full mode shape information to the one with partial mode shape information. There are, however, several other advantages to expanding the identification model class $\mathcal{M}$ by introducing the system mode shapes:

1. Because of the constraints of the assumed mathematical structure built into $\mathcal{M}$, it might not be possible for any structural model in this class to produce theoretical mode shapes that will give a good match to the experimental mode shapes. The system mode shapes provide extra flexibility in this aspect.
2. Their introduction also turns out to remove any need to match experimental and model modes during the identification, thereby avoiding a common difficulty in applications.

In the Bayesian framework, all model parameters in the linear structural model, including mass, stiffness, and system mode shapes, are updated. Prior PDFs are specified for the model parameters to reflect the relative plausibilities of their values in the absence of any measurement data. The goal of the Bayesian model updating technique is to construct an updated (posterior) PDF for the parameters using the prior PDFs and the experimental modal data.

2.1.2 Selection of prior PDF. The prior PDF for λ and ϕ is taken to be the product of independent Gaussian PDFs where

$$\lambda \sim N(\lambda^0, P^0) \quad \phi \sim N(\phi^0, Q^0) \qquad (4)$$

and where λ^0 and ϕ^0 are the chosen most probable values of λ and ϕ, respectively, and P^0 and Q^0 are the chosen variances to express the initial uncertainty in these parameters. The variance parameters $\{\delta_r^2 : r = 1, \ldots, N_m\}$ are directly estimated from the sample variance of the experimental mode shape data,

$$\hat{\delta}_r^2 = \left(\sum_{j=1}^{N_s} \left\| \hat{\psi}_{r,j} - \left(\sum_{m=1}^{N_s} \hat{\psi}_{r,m} \Big/ N_s \right) \right\|_2^2 \right) \Big/ (N_s N_o) \qquad (5)$$

where $\|\cdot\|_2$ stands for the Euclidean norm. The prior PDF for $\sigma^2 \equiv \{\sigma_r^2 : r = 1, \ldots, N_m\}$ is taken to be the product of independent inverse gamma PDFs, $IG(\alpha_r, \beta_r)$,

$$p(\sigma_r^2) \sim IG(\alpha_r, \beta_r) \propto \left(1/\sigma_r^2\right)^{\alpha_r - 1} e^{-(\beta_r/\sigma_r^2)} \qquad (6)$$

When $\alpha_r = \beta_r = 0$, the inverse gamma prior becomes the usual Jeffreys' non-informative prior, i.e., $p(\sigma_r^2) \sim 1/\sigma_r^2$.

These choices of the prior PDFs (Gaussian and inverse gamma) facilitate the use of the GS. In fact, these prior PDFs are the Bayesian conjugate priors (Gelman et al., 1995) for the mean and variance of a Gaussian PDF and their use enables us to do exact sampling with the GS algorithm.

3 GIBBS SAMPLER

From the structural health monitoring point of view, our goal is to compute $p(\theta \mid \hat{D})$, the PDF of the stiffness parameters after it is updated by including the information from the modal data $\hat{D}$. Note that $p(\theta \mid \hat{D})$ is the marginal PDF obtained from the PDF $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ by integrating out ρ, ϕ, and σ^2. The latter PDF is given by Bayes' Theorem,

$$p(\lambda, \phi, \sigma^2 \mid \hat{D}) = \frac{p(\lambda)p(\phi)p(\sigma^2)p(\hat{D} \mid \lambda, \phi, \sigma^2)}{\int p(\lambda)p(\phi)p(\sigma^2)p(\hat{D} \mid \lambda, \phi, \sigma^2)\, d\lambda\, d\phi\, d\sigma^2} \qquad (7)$$

where all PDFs on the right-hand side of this equation have already been specified.

Unfortunately, it is not feasible to compute $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ analytically, nor numerically, because the integral in the denominator in (7) is difficult to evaluate due to the high dimensionality of $\{\lambda, \phi, \sigma^2\}$. Therefore, our strategy is to sample from $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ so that the θ components of these samples are distributed as $p(\theta \mid \hat{D})$. Any quantity of interest regarding θ, such as the sample mean and variance of θ conditional on $\hat{D}$, can be estimated using the samples. Difficulties arise because direct sampling from $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ is not feasible, so the GS algorithm (Geman and Geman, 1984; Gelfand et al., 1990) is used.

3.1 Gibbs sampler algorithm

1. Initialize the samples with $\{\hat{\lambda}^{(0)}, \hat{\phi}^{(0)}, \hat{\sigma}^{2(0)}\}$ drawn from the prior PDFs and let $k = 1$.
2. a. Sample the system mode shapes $\hat{\phi}^{(k)} \sim p(\phi \mid \hat{\lambda}^{(k-1)}, \hat{\sigma}^{2(k-1)}, \hat{D})$.
   b. Sample the prediction-error variances $\hat{\sigma}^{2(k)} \sim p(\sigma^2 \mid \hat{\lambda}^{(k-1)}, \hat{\phi}^{(k)}, \hat{D})$.
   c. Sample the structural parameters $\hat{\lambda}^{(k)} \sim p(\lambda \mid \hat{\sigma}^{2(k)}, \hat{\phi}^{(k)}, \hat{D})$.
3. Let $k = k + 1$, go back to Step 2 and cycle until N samples $\{\hat{\lambda}^{(k)}, \hat{\phi}^{(k)}, \hat{\sigma}^{2(k)} : k = 1, \ldots, N\}$ are obtained.

Therefore, the basic idea of the GS is to decompose the uncertain model parameters $\{\lambda, \phi, \sigma^2\}$ into three groups, i.e., λ, ϕ, and σ^2. With the choices of the conjugate priors, exact sampling from one parameter group conditional on the other groups and the modal data can be achieved. Thus, although the dimension of $\{\lambda, \phi, \sigma^2\}$ is huge, the effective dimension is always three. The GS is a special case of Markov Chain Monte Carlo methods; however, unlike the common Markov Chain Monte Carlo methods, the GS never rejects samples.

When k gets large, the Markov chain samples $\{\hat{\lambda}^{(k)}, \hat{\phi}^{(k)}, \hat{\sigma}^{2(k)}\}$ are distributed as $p(\lambda, \phi, \sigma^2 \mid \hat{D})$, assuming that the Markov chain created by the GS is ergodic (Gelman et al., 1995). According to our experience, when the regions of high values of $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ are effectively connected, the Markov chain samples are distributed as $p(\lambda, \phi, \sigma^2 \mid \hat{D})$ when k is large no matter how the GS algorithm is initialized, i.e., an ergodic Markov chain is obtained. Globally identifiable and unidentifiable cases (Beck and Katafygiotis, 1998) are of this kind. Therefore, it is postulated that the proposed GS approach is useful for globally identifiable and unidentifiable problems. For locally identifiable problems (Beck and Katafygiotis, 1998), the samples obtained using a single GS run may get trapped in the neighborhood of only one of the optimal models.

3.2 Burn-in period and convergence diagnosis

The Markov chain samples obtained from the GS are distributed as the target PDF when the Markov chain reaches its stationary state. The time period required for the Markov chain to reach its stationary state is called the burn-in period. Usually, a burn-in period is determined by visual inspection by plotting the Markov chain samples through time and identifying the transient period. The Markov chain samples within the burn-in period are discarded because they are not distributed as the target PDF.

The Markov chain samples after the burn-in period are collected, and the quantities of interest can be estimated using these samples. However, there is a question of when to stop the GS so that the obtained Markov chain samples well represent the target PDF. This question arises because the Markov chain samples are dependent. There are several available convergence diagnosis criteria that can be used to determine if the Markov chain samples are enough (e.g., the Gelman-Rubin convergence statistic; Gelman and Rubin, 1992). In this article, the following simple method of determining the convergence is chosen: the convergence of the estimates of the quantities of interest is visually inspected and the GS stops if the convergence appears to be reached.

3.3 Full conditional PDF

To implement the GS, the full conditional PDFs, i.e., $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$, $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$, and $p(\sigma^2 \mid \hat{\lambda}, \hat{\phi}, \hat{D})$, need to be derived. These full conditional PDFs are also Gaussian and inverse gamma because conjugate priors are adopted so that they are readily sampled. First, $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$ and $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ are Gaussian. More specifically, with λ and σ^2 fixed, (2) can be written in the following form:

$$\hat{Y} = AX + E \tag{8}$$

where A and $\hat{Y}$ are a fixed matrix and vector depending on the fixed λ and the modal data; E is the prediction error, distributed as $N(0, \Sigma)$, where Σ is a fixed covariance matrix depending on the fixed σ^2; the prior PDF of X is $N(X^0, \Sigma^0)$. For the current discussion, X is identical to ϕ so $X^0 = \phi^0$ and $\Sigma^0 = Q^0$. According to Gelfand et al. (1990)

$$\begin{aligned} \mathrm{E}(X \mid \hat{Y}) &= X^0 + \Sigma^0 A^T (A\Sigma^0 A^T + \Sigma)^{-1} (\hat{Y} - AX^0) \\ \mathrm{Cov}(X \mid \hat{Y}) &= \Sigma^0 - \Sigma^0 A^T (A\Sigma^0 A^T + \Sigma)^{-1} A\Sigma^0 \end{aligned} \tag{9}$$

where $\mathrm{E}(X \mid \hat{Y})$ and $\mathrm{Cov}(X \mid \hat{Y})$ are, for the current discussion, exactly the first two moments of the Gaussian PDF $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$. The first two moments of $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ can be obtained in a similar manner.

On the other hand, $p(\sigma^2 \mid \hat{\lambda}, \hat{\phi}, \hat{D})$ is inverse gamma (Gelfand et al., 1990):

$$p\left(\sigma_r^2 \mid \hat{\lambda}, \hat{\phi}, \hat{D}\right) = IG\left(\alpha_r + \frac{N_d N_s}{2}, \beta_r + \frac{1}{2} \sum_{j=1}^{Ns} \left\| \hat{K}\hat{\phi}_r - \hat{\omega}_{r,j}^2 \hat{M}\hat{\phi}_r \right\|_2^2\right) \tag{10}$$

Some computational issues that may occur when implementing the GS approach are discussed in the Appendix.

4 EVALUATION OF DAMAGE PROBABILITY

For the purpose of structural health monitoring, it is essential to locate possible damage and provide an index indicating the severity of the damage. This is achieved by estimating the probability that any stiffness parameter has decreased more than a fraction d. Note that the GS samples $\{\hat{\lambda}^{(k)}, \hat{\phi}^{(k)}, \hat{\sigma}^{2(k)} : k = 1, \ldots, N\}$ will ultimately be distributed as $p(\lambda, \phi, \sigma^2 \mid \hat{D})$. This implies that the extracted stiffness samples $\{\hat{\theta}^{(k)} : k = 1, \ldots, N\}$ are ultimately distributed as $p(\theta \mid \hat{D})$. With the θ samples conditional on the modal data obtained from an undamaged state ($\hat{D}^{\mathrm{ud}}$) and a possible damaged state of a structure ($\hat{D}^{\mathrm{pd}}$), the probability that the damage fraction in the ith stiffness parameter exceeds threshold d can be estimated using the following approximation:

$$\begin{aligned} P\left(\theta_i^{\mathrm{pd}} < (1-d)\theta_i^{\mathrm{ud}} \mid \hat{D}^{\mathrm{ud}}, \hat{D}^{\mathrm{pd}}\right) \\ \approx \frac{1}{M} \sum_{n=1}^{M} I\left[\hat{\theta}_{i,n}^{\mathrm{pd}} < (1-d)\hat{\theta}_{i,n}^{\mathrm{ud}}\right] \end{aligned} \tag{11}$$

where $I[.]$ is the indicator function, which is unity if the condition is satisfied, otherwise it is zero; θ_i^{pd} and θ_i^{ud} are the ith stiffness parameters in a possibly damaged state and undamaged state, respectively; $\hat{\theta}_{i,n}^{\mathrm{pd}}$ and $\hat{\theta}_{i,n}^{\mathrm{ud}}$ are the nth samples independently resampled from the Markov chain samples of $p(\theta \mid \hat{D}^{\mathrm{pd}})$ and $p(\theta \mid \hat{D}^{\mathrm{ud}})$, respectively. Such resampling is conducted M times (i.e., $n = 1, \ldots, M$), where M is a large integer.

5 EXAMPLES

5.1 Simulated 2-DOF linear shear building

The first example is a 2-DOF linear shear building in which the lumped masses of the two floors are equal and the ratio of each inter-story stiffness to the mass is 2,000 in appropriate units. The building is subject to simulated white-noise ambient excitation at each degree of freedom, and the measured response of the structure, which is equal to the actual acceleration plus 10% root-mean-square (RMS) noise, is simulated before and after damage. The goal of this example is to use the GS approach to locate and quantify the damage given the measured response and the structural model.

Damage is simulated by reducing the column stiffness. There are two damage cases: the damage for the first and second cases is a 30% reduction of the inter-story stiffness of the first and second stories, respectively. For the undamaged case and the two damaged cases, 110 seconds of ambient-vibration acceleration data at the two floors are generated using an implicit time-integration method.

5.1.1 Modal identification. Ten sets of modal data are extracted from the acceleration data using a modal identification procedure called MODE-ID (Beck, 1996), so $N_S = 10$ (the 110 seconds of data are divided into 10 segments with the first 10 seconds neglected). The method assumes a linear model with classical normal modes of vibration. In the case where the excitation is unknown (e.g., ambient vibration), it is modeled as a weakly stationary stochastic process where the current excitation is assumed to be uncorrelated from the past response. It can be shown that the cross-correlation functions of the model responses then satisfy the original equation of motion for the structure in free vibration where the time

Table 1
Mean modal data for the 2-DOF shear building example

	Frequency (Hz)	
	Mode 1	*Mode 2*
Undamaged	4.37	11.59
COV (%)	6.41	3.88
30% 1st-story damage	3.95	11.05
COV (%)	5.32	2.90
30% 2nd-story damage	4.24	10.53
COV (%)	4.95	3.70

lag serves as a pseudo-time (Beck et al., 1994). The identified modal parameters from these pseudo-free vibration data are the modal frequencies, damping ratios, and mode shape components at the measured DOFs. Table 1 shows the average values and the coefficients of variation (COV) of the identified modal frequencies. Mode shapes are also identified using MODE-ID although they are not shown.

5.1.2 Identification model. The structural model used for the GS is a 2-DOF linear shear building model, where the uncertain parameters are $\{\rho_1, \rho_2\}$, $\{\theta_1, \theta_2\}$, $\{\phi_1, \phi_2\}$, and $\{\sigma_1, \sigma_2\}$ (see (1) and (2)). The variances $\{\delta_1^2, \delta_2^2\}$ are directly estimated from the experimental mode shapes identified by MODE-ID. The number of uncertain parameters is therefore 10 because the system mode shapes each have two components. The prior PDFs of $\{\rho_1, \rho_2\}$ are chosen to be independent Gaussian PDFs with means equal to 1 and COV equal to 10%. Those of $\{\theta_1, \theta_2\}$ are independent Gaussian PDFs with means equal to 1 and COV equal to 30%. Flat independent prior PDFs are taken for $\{\phi_1, \phi_2\}$, and independent inverse gamma prior PDFs with large COV (100%) are taken for $\{\sigma_1^2, \sigma_2^2\}$. It is found that the GS results are insensitive to the means of the inverse gamma prior PDFs as their COV are large.

5.1.3 Results from the Gibbs sampler. Following the GS procedure, Markov chain samples of the masses, stiffnesses, system mode shapes, etc. are obtained. The initial point for the Markov chain is generated from the prior PDFs of the uncertain parameters. Figure 1a shows the Markov chain samples of $\{\theta_1, \theta_2\}$ for the undamaged case, which shows that the Markov chain seems to reach its stationary state very quickly, i.e., the burn-in period is less than 50 time steps. In Figure 2a, all the samples for the undamaged case are plotted, excluding those in the burn-in period, in the $\{\theta_1, \theta_2\}$ space. The results show that the model is globally identifiable (GID).

Also of interest are the GS results when fewer modal data are available. Figure 1b shows the Markov chain samples of θ_1 and θ_2 for the undamaged case when only the roof acceleration is measured and also only the modal

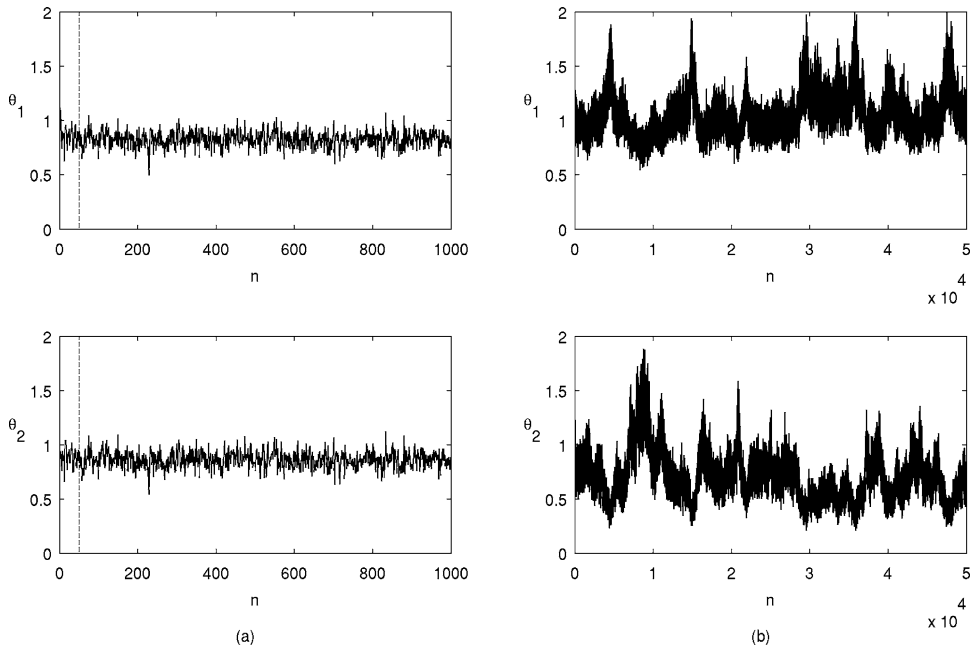

Fig. 1. Markov chain samples for the stiffness parameters of the undamaged case in the (a) globally identifiable and (b) unidentifiable cases. The index n denotes the time step of the Markov chain. The dashed line indicates the end of the burn-in period at $n = 50$ (not visible for the unidentifiable cases because of scale).

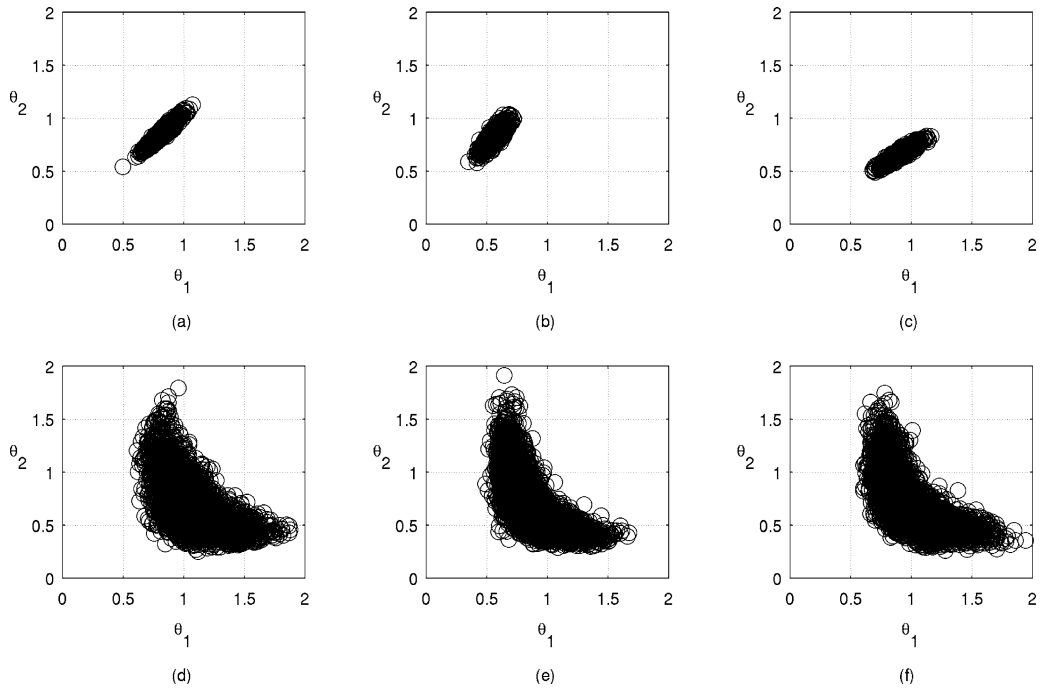

Fig. 2. Stiffness parameter samples plotted in the $\{\theta_1, \theta_2\}$ space for the (a) GID undamaged, (b) GID 30% 1st story damage, (c) GID 30% 2nd story damage, (d) UID undamaged, (e) UID 30% 1st story damage, and (f) UID 30% 2nd story damage cases.

data from the first mode are available. The burn-in period is roughly 50 time steps. Similarly, Figure 2d plots all the samples for the undamaged case in the $\{\theta_1, \theta_2\}$ space. It is clear that the samples spread out in the $\{\theta_1, \theta_2\}$ space. In fact, the 2-DOF model is unidentifiable (UID), showing there is insufficient information provided when only the first mode data from acceleration measurements at the roof is available.

Figures 2b and c show the Markov chain samples in the $\{\theta_1, \theta_2\}$ space for the first and second damaged cases when the full set of the modal data are available (GID case), Figures 2e and f show the samples for the two damaged cases when only first mode data from the roof accelerations is available (UID case), which again show that the samples spread out in the $\{\theta_1, \theta_2\}$ space.

Figures 3 and 4 show the probability that the damage fractions of the two stiffness parameters exceeds the threshold d for the two damaged cases, as estimated using (11). It is clear that the results of damage detection are reliable when the full set of modal data are available (GID case) but not so when only the roof acceleration and the first mode data are available (UID case). The GID mean damage fraction for the first and second story damage states are 30% and 22%, respectively, compared to the actual damage fraction of 30% for both damage states. However, in the unidentifiable case, damage to the first story is detected with an estimated mean damage fraction of 20%, but damage to the second story is detected as damage to the first story with a mean damage fraction of less than 10%. This is because it is not possible to accurately localize the damage without mode shape information. Performance is better for the first story damage case because the first natural frequency is much more sensitive to the first story stiffness, so the difference between the damaged and undamaged states is greater than for the second story damage case. It is found that the convergence of the estimates of the probability of damage is fast for the GID cases (1,000 Markov chain samples are sufficient, as more samples do not improve the results), although for the UID cases, it requires more than 10,000 samples for convergence, as indicated by the results in Figures 3 and 4.

5.2 IASC-ASCE Phase II Simulated Benchmark studies

For the second example, the GS approach is applied to the connection-damage cases in the IASC-ASCE Phase II Simulated Structural Health Monitoring Benchmark problems (Bernal et al., 2002). In the benchmark problems, data are generated by a finite-element structural model with 120 DOF, which is referred to as the benchmark model and is shown in Figure 5. It is a 3D four-story 2-bay by 2-bay unbraced model structure. The y-direction is the strong direction of the columns, the base is fixed, and the floors are rigid in plane. The nominal mass from the first to the fourth story is 3,242, 2,652,

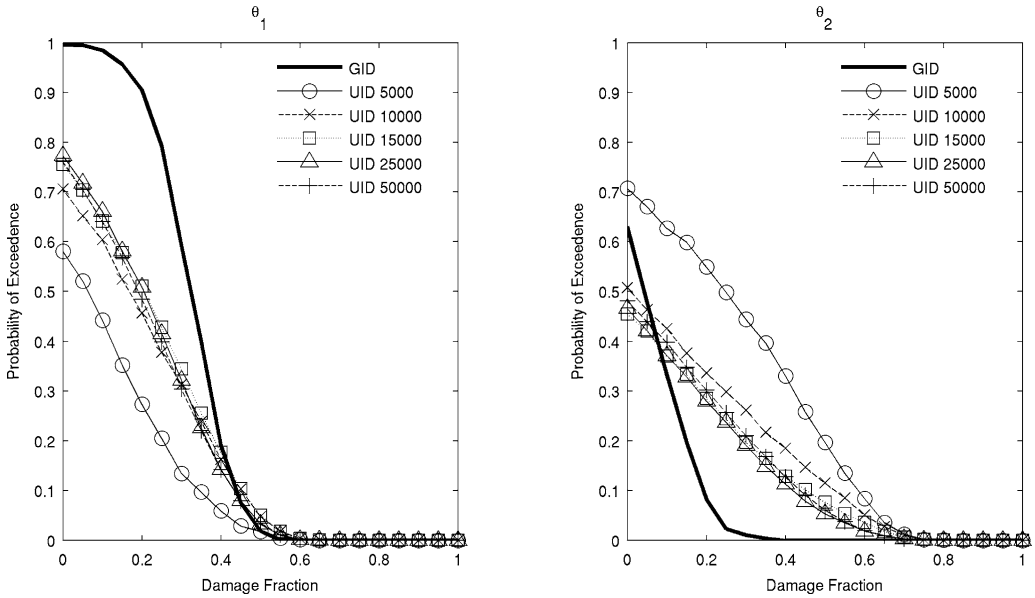

Fig. 3. Estimated damage probability curves for 30% 1st story damage for the GID and UID cases. The curves for the UID case are calculated using 1,000, 10,000, 15,000, 25,000, and 50,000 samples (samples before burn-in are discarded). The curve for the GID case is calculated using 1,000 samples.

2,652, and 1,809 kg. The nominal properties of the beams and columns are listed in Table 2. Actual masses are modeled by randomly selecting all floor mass from a uniform distribution over [0.9, 1.1] of the nominal value; the center of the floor mass deviates from the geometrical floor center by randomly selecting a factor from a uniform distribution over [−0.05, 0.05] of the floor width. These random factors in the benchmark model that are used to generate the simulated test data are unknown to investigators.

The connection-damage cases in the Phase II Benchmark problems involve detection and assessment of simulated beam-column connection damage with different severities and at different locations in the benchmark

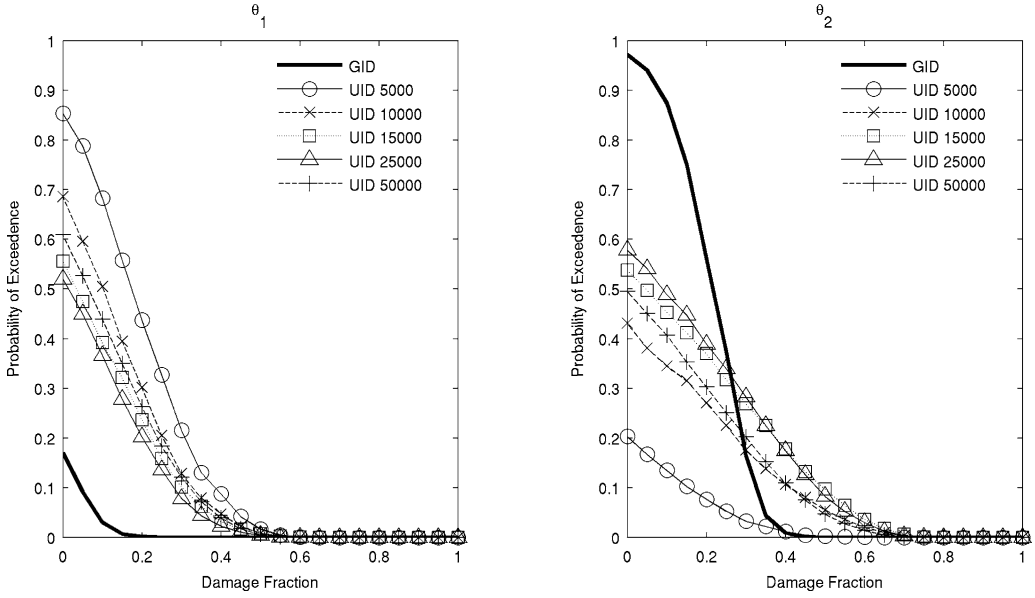

Fig. 4. Estimated damage probability curves for 30% 2nd story damage for the GID and UID cases. The curve for the GID case is calculated using 1,000 samples.

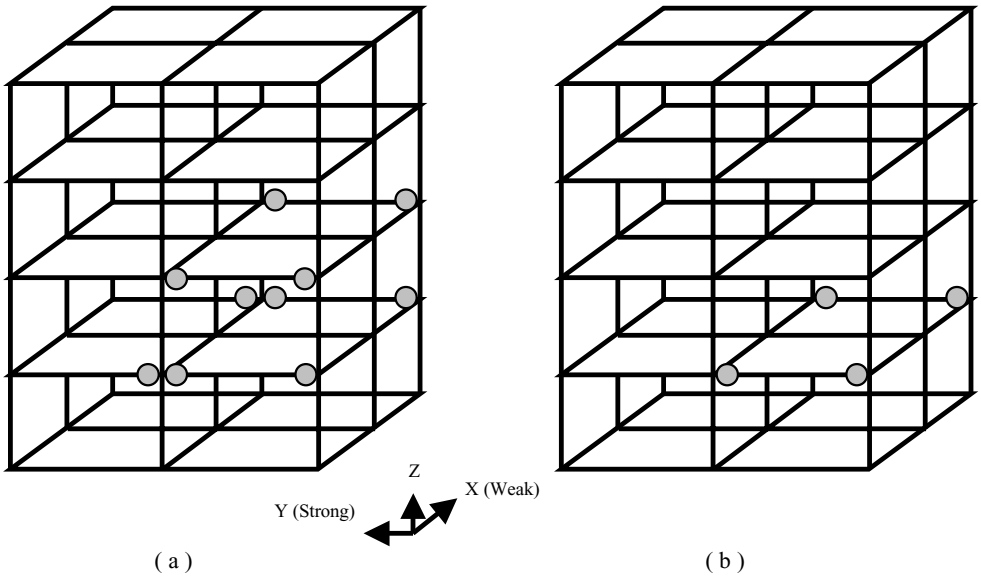

Fig. 5. The diagram of the benchmark structure, showing damage locations for (a) DP 1 and (b) DP 2. The circles indicate connections with reduced rotational stiffness.

model. The damage is simulated by reducing the stiffnesses of the undamaged (Undmg) finite-element model at the ends of some beams, that is, the stiffnesses of some rotational springs whose axes in the x-direction are reduced so that the model becomes more flexible in the strong y-direction. There are two damage patterns: (1) DP1: complete loss of rotational stiffness at six first-story beam-column connections and four second-story beam-column connections, as shown in Figure 5a; (2) DP2: complete loss of rotational stiffness at four first-story beam-column connections, as shown in Figure 5b. The connection damage cases are the most challenging in the simulated benchmark study, as the rotational DOF are not directly observed.

The simulated ambient-vibration data for all cases are generated by simulated wind forces consisting of broadband stationary excitation of the model at each floor; this excitation is unknown to the investigator. The data can be downloaded from the ASCE benchmark website (Dyke, 2001). Acceleration measurements are available at the center of each side at each floor with the directions parallel to the side in either the positive x- or y-direction (see Figure 5).

Time histories with a sampling interval of 0.002 seconds and total duration of 210 seconds are generated for all damaged and undamaged cases. They are partitioned temporally into 10 sets ($N_s = 10$) of equal duration of 20 seconds (ignoring the first 10 seconds) to yield 10 sets of independent estimates of the modal data for each case. The first 10 seconds are ignored as it contains non-stationary transient responses. When generating the time histories, damping ratios are taken to be equal to 1% for all modes and simulated measurement noise is added equal to 10% RMS of the actual acceleration at the measured DOF.

5.2.1 Modal identification. Eight modes ($N_m = 8$), four in the strong (y) direction and four in the weak (x) direction, of the structure are identified using MODE-ID for all cases. Table 3 shows the average value of the modal

Table 2
The properties of the structural elements in the analytical model (Johnson et al., 2004)

Property		*Columns*	*Floor Beams*
S*ection type*		B100 × 9	S75 × 11
Cross-sectional area	A (m^2)	1.133×10^{-3}	1.43×10^{-3}
Moment of inertia (strong direction)	I_1 (m^4)	1.97×10^{-6}	1.22×10^{-6}
Moment of inertia (weak direction)	I_2 (m^4)	0.664×10^{-6}	0.249×10^{-6}
St. Venant torsion constant	J (m^4)	8.01×10^{-9}	38.2×10^{-9}
Young's Modulus	E (Pa)	2×10^{11}	2×10^{11}
Mass per unit length	ρ (kg/m)	8.89	11.0

Table 3
Modal data for the ASCE Benchmark example

	Frequency (Hz)							
	(W1)	*(S1)*	*(W2)*	*(S2)*	*(W3)*	*(S3)*	*(W4)*	*(S4)*
Undmg	3.19	3.98	9.79	13.41	16.66	25.15	23.72	39.28
COV (%)	1.72	0.33	0.46	0.29	1.08	0.34	0.37	0.20
DP1	3.20	3.42	9.78	12.91	16.69	24.68	23.72	39.11
COV (%)	2.45	0.91	0.60	0.57	0.41	0.39	0.38	0.30
DP2	3.19	3.79	9.79	13.13	16.72	25.15	23.72	39.17
COV (%)	2.02	0.57	0.53	0.58	0.45	0.30	0.38	0.28

**"W" and "S" denote "weak" and "strong" directions.

frequencies identified from the 10 sets of time histories. The corresponding COV are shown in percent in these tables. The modal frequencies in the weak direction are expected to be unchanged because the damage patterns weaken only the strong direction. The 16 acceleration measurements, four on each face of the structure, are used to identify the Euclidean-normalized mode shape components at the "+x" "−x" "+y" "−y" faces of the structure, which are plotted in Figure 6 for the undamaged case. Here, "+x" "−x" "+y" "−y" represent the directions of the outward normal of the faces. In these plots, the identified mode shape components are joined by straight-line segments.

5.2.2 Identification model. A 3-D 36-DOF model that assumes rigid floors in the x–y plane and allows rotation along the x- and y-axes is proposed. Each floor is assumed to be rigid in plane, that is, with respect to rotation along the z-axis and for x- and y-direction translations, giving three of the nine DOF for each story. Nodes in each floor are allowed to rotate along the x- and y-axes in a constrained way: nodes with the same x-coordinates or same

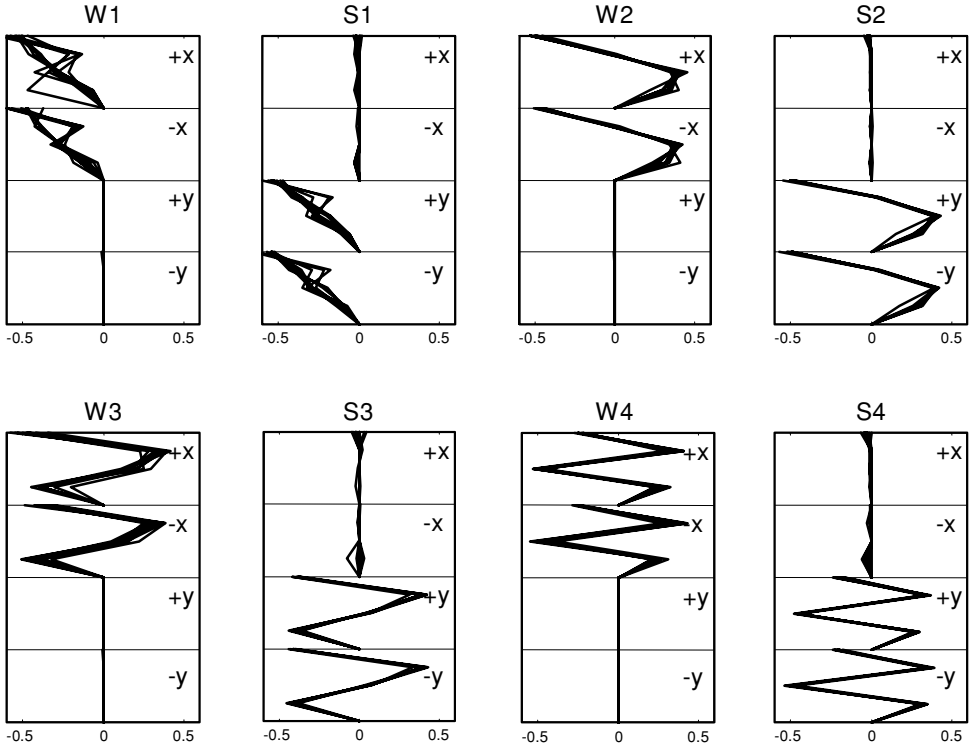

Fig. 6. The experimental mode shape components of the structure for the undamaged case. "W" and "S" denote "weak" and "strong" directions.

y-coordinates are assumed to have the same amount of rotation along the y-axis and x-axis, respectively, to give the remaining six DOF for each floor. Translation along the z-direction is not allowed in this model.

Two parameters are used for the rotational stiffness in each floor: it is assumed that the rotational stiffness of all beam-column connections in the same floor along the x-axis (or along the y-axis) is identical. This assumption is imposed as the mode shape data (e.g., Figure 6) indicate that the mode shape components in the x and y directions are mostly decoupled; therefore, in practice it is unlikely that the face sustaining the damage can be distinguished. The rotational stiffnesses with rotation axes along the x-axis and y-axis are called the strong and weak direction rotational stiffnesses, respectively. Besides the rotational stiffness, four parameters are used for the column stiffnesses, one for each story, to give 12 stiffness parameters in total:

$$K(\theta) = \sum_i \theta_{c,i} \bar{K}_{c,i} + \sum_i \sum_j \theta_{ij} \bar{K}_{ij} \tag{12}$$

where $i = 1, \ldots, 4$, $j = 'x', 'y'$, and the indices i and j denote the story number and the axis along which the rotational stiffness is active, respectively; $\bar{K}_{ij}$ are the nominal rotational stiffness matrices computed based on the model assumptions for the original undamaged structure; $\bar{K}_{c,i}$ is the nominal stiffness matrix contributed by the columns in the ith story. It follows that under this model, damage pattern DP1 corresponds to 2/3 loss of θ_{1y} and 1/3 loss of θ_{2y}, and DP2 corresponds to 1/3 loss of θ_{1y}.

Due to the fact that the stiffness matrix of the benchmark structure is dominated by the columns and the fact that the columns also provide rotational stiffness, slight errors in the identified column stiffness parameter will significantly influence the values of the identified rotational stiffness parameters. This indicates that to reliably detect and assess rotational stiffness damage, sufficient prior information about the column stiffness is needed. Therefore, the prior PDF on the column stiffness parameter is taken to be independent Gaussian with mean and COV equal to 1% and 2%. The prior PDF of the eight rotational stiffness parameters is taken to be independent Gaussian with mean and COV equal to 1% and 20%. These choices would be reasonable for a real structure where there is likely to be a strong belief that structural damage will occur in the connections and not in the columns themselves.

In calculating the nominal story masses, the mass of the columns is lumped at the floors that they are connected to. One mass parameter is used for each story to give four mass parameters:

$$M(\rho) = \sum_i \rho_i \bar{M}_i \tag{13}$$

where $i = 1, \ldots, 4$ represents the story number and the $\bar{M}_i$ are the nominal mass matrices computed based on the original undamaged structure. The prior PDF for the uncertain mass parameters is assumed to be independent Gaussian with mean equal to unity and COV equal to 1%. This means that the mass will not vary much from its most probable (nominal) value.

5.2.3 Results from the Gibbs sampler. Following the GS procedure, Markov chain samples of the masses, stiffnesses, system mode shapes, etc. are obtained. The total number of uncertain parameters is 312. To examine whether the Markov chain created by the GS is ergodic or not, five parallel GSs are conducted to obtain five independent Markov chains. The initial point for each Markov chain is generated from the prior PDFs of the uncertain parameters. It is found that all the five Markov chains converge to the same region in the parameter space, suggesting that the Markov chain is ergodic.

Figure 7 shows the samples of the eight rotational stiffnesses for the undamaged case from one of the parallel Markov chains. The Markov chain seems to reach its stationary state after a burn-in period of 1,000 time steps. For the two damaged cases, the results are similar except that the sample mean of the rotational stiffnesses may be different. In Figure 8, the kernel PDFs of the rotational stiffnesses built by the Markov chain samples (excluding the 1,000 in the burn-in period) are plotted for the Undmg, DP1, and DP2 cases. It is clear that the damage of the strong direction rotational stiffnesses in the first and second stories is correctly detected for the DP1 case, and the damage in the first story is also correctly detected for the DP2 case. Moreover, the amount of the identified stiffness loss is roughly correct, i.e., for the most probable values, there is about 55% loss in θ_{1y} and 40% loss in θ_{2y} for DP1 and 30% loss in θ_{1y} for DP2. Figures 9 and 10 show the sample mean of the system mode shapes for the three cases, which shows that the identified mode shapes do not change after the connection damage except for slight changes in the rotational components of the mode shapes in the strong direction, the damaged direction.

Figures 11 and 12 show the probability that the damage fractions of the eight rotational stiffness parameters exceed threshold d for the two damaged cases, as estimated using (11) with samples from one Markov chain for each case (Undmg., DP1, DP2). It is found that 4,000 samples are sufficient for convergence. As seen from the figures, all the actual damage locations are correctly

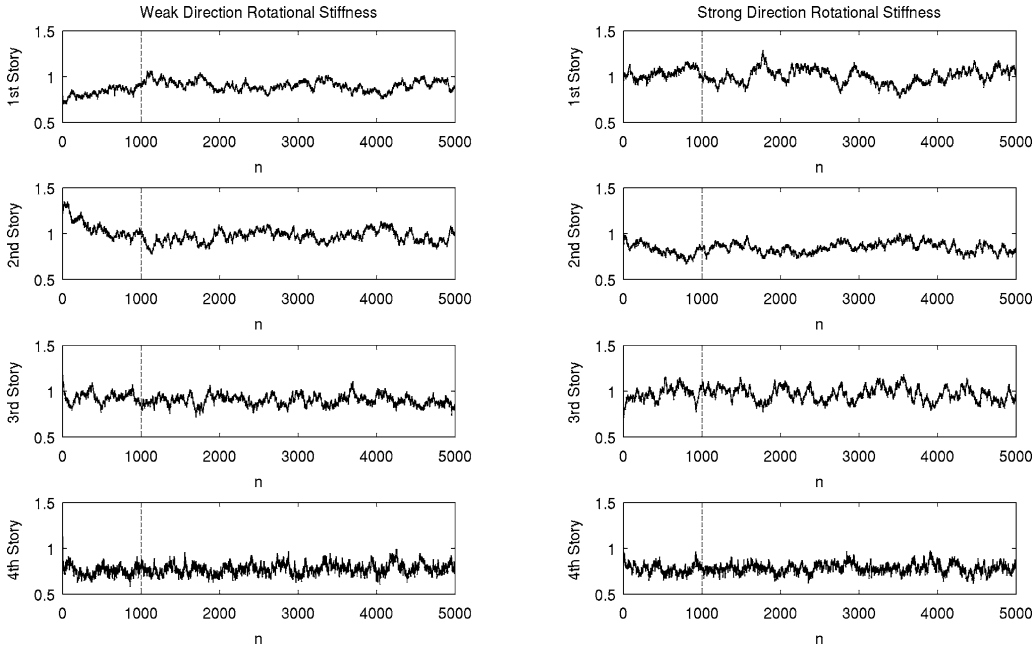

Fig. 7. Markov chain samples for the rotational stiffness parameters in the undamaged case. The index n denotes the time step of the Markov chain. The dashed line indicates the end of the burn-in period at $n = 1{,}000$.

detected (i.e., in the first and second stories in the strong direction for the DP1 case and the first story only for the DP2 case). For each of the undamaged locations, the median loss of rotational stiffness never exceeds about 3%.

6 CONCLUSION AND DISCUSSION

A GS approach for linear Bayesian structural model updating has been presented. The results from the two examples show that the GS approach is effective in

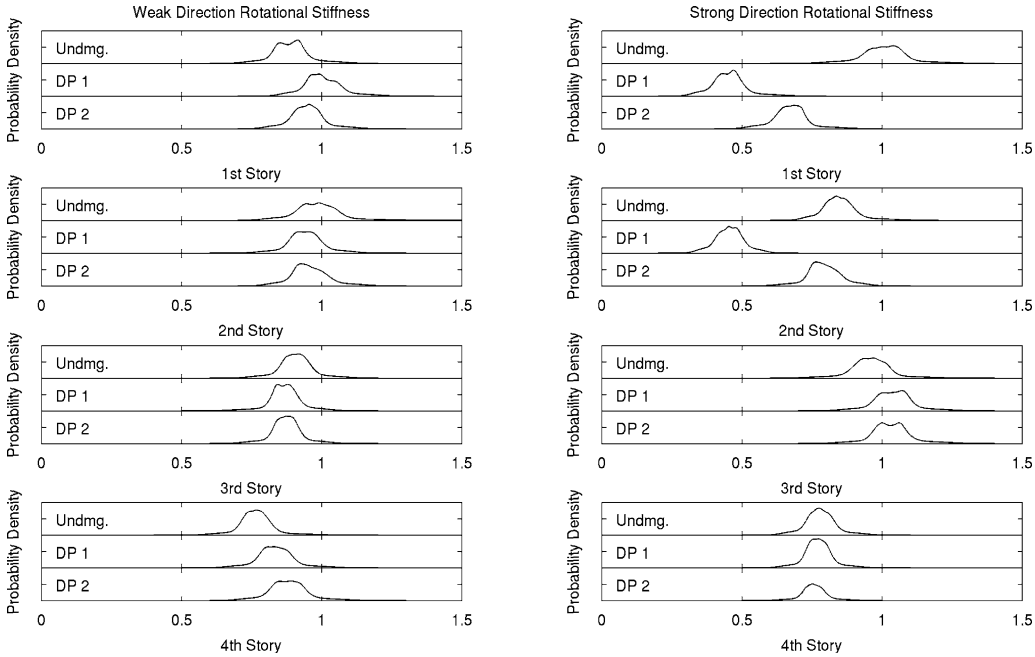

Fig. 8. Kernel probability densities built from the rotational stiffness parameter samples from five parallel Markov chains with 5,000 samples each.

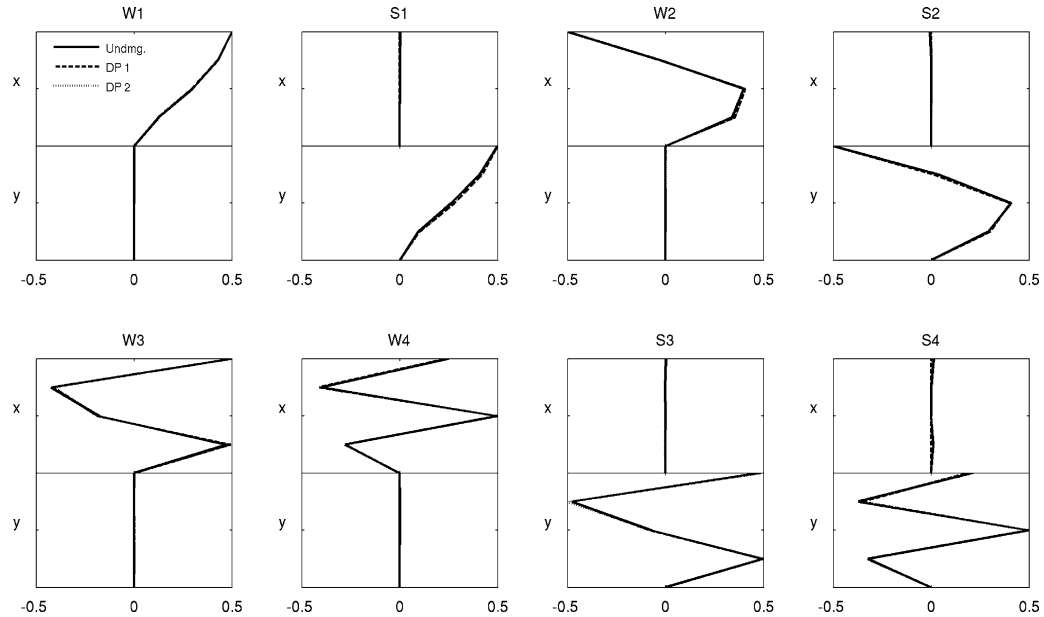

Fig. 9. Translational degrees of freedom for the mean system mode shapes (first four modes in each direction and for the three damage cases).

detecting and locating damage in an instrumented structure by using small-amplitude vibration data from before and after a severe loading event, as long as there are sufficient sensors so that the model is globally identifiable. The major advantage of the GS approach is that its efficiency does not degrade with an increasing number of uncertain parameters in the structural model. This property is rare among the most common stochastic simulation approaches, e.g., importance sampling, Metropolis-Hasting, rejection sampling, etc. Therefore, in principle,

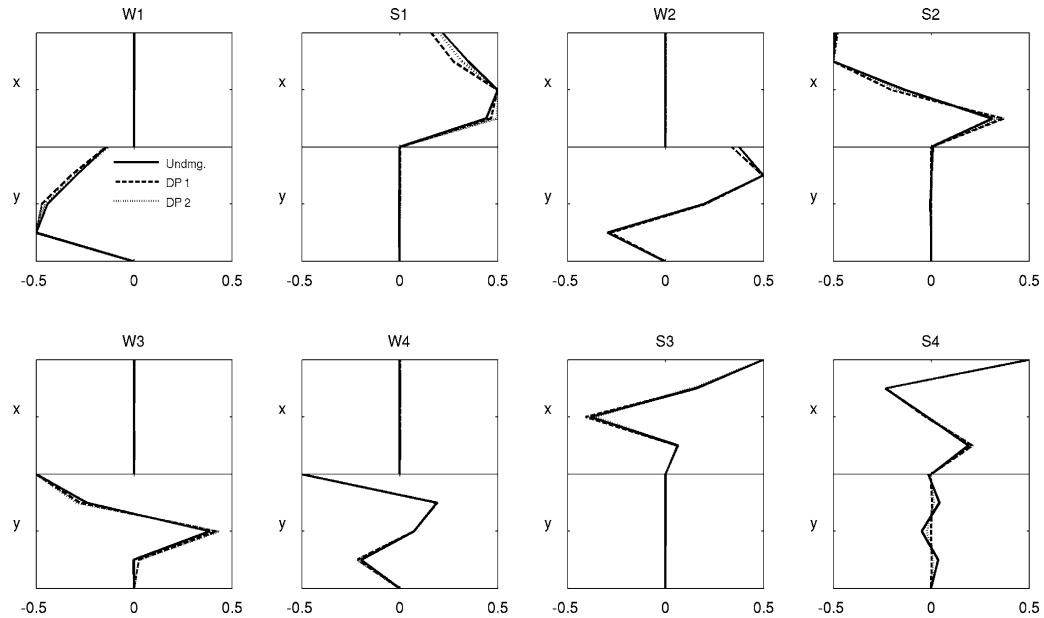

Fig. 10. Rotational degrees of freedom of the mean system mode shapes (first four modes in each direction and for the three damage cases). The average values of the three rotational mode shape components at each floor and in each direction are shown in this figure.

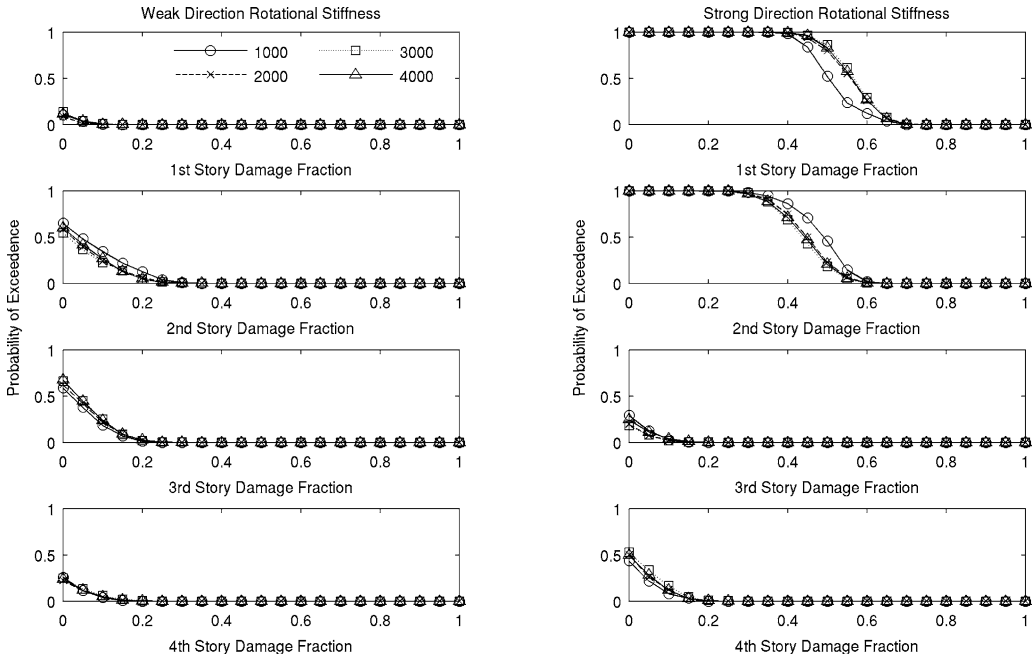

Fig. 11. Damage probability curves for DP1, using 1,000, 2,000, 3,000, and 4,000 post burn-in samples.

the GS approach may be applied to the health monitoring of real structures with high-dimensional uncertain parameters.

There are no false damage detections in the usual strict sense, only in a soft probabilistic sense, i.e., there is a significant probability of substantial damage (say, more than 10% stiffness reduction) for some locations where there is no damage (but these probabilities are less than 0.5, so the inference of substantial damage is more likely to be not true). These situations mostly occur in the unidentifiable cases, which are the cases where there are not enough sensor measurements, so

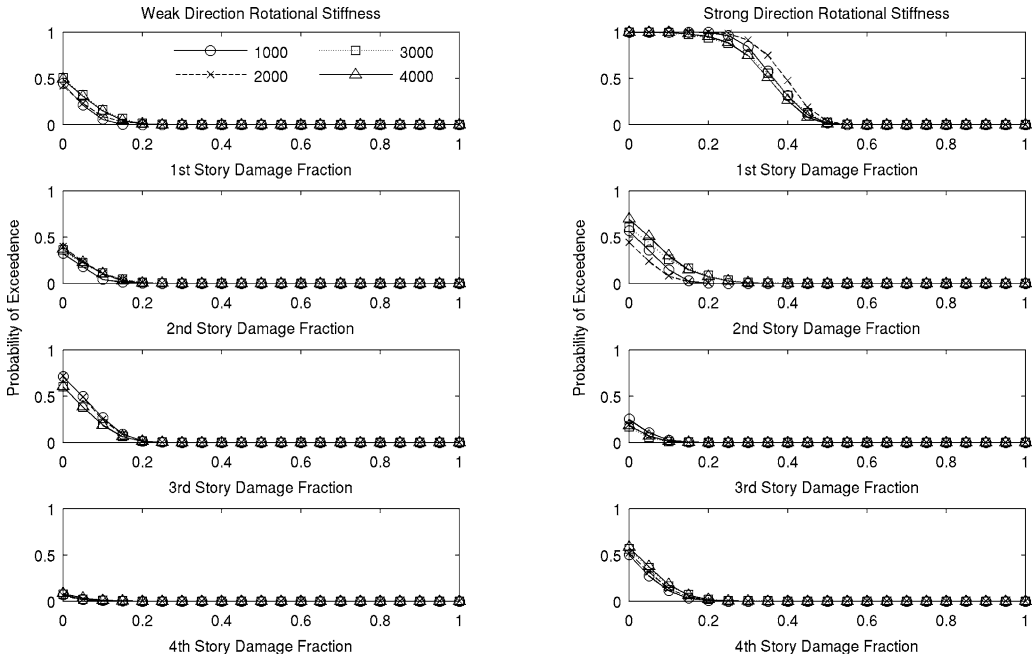

Fig. 12. Damage probability curves for DP2, using 1,000, 2,000, 3,000, and 4,000 post burn-in samples.

it is due to an inherent lack of information rather than a deficiency of the method. Similar comments can be made about missed damage detections, which occur only in a soft probabilistic sense and only for the unidentifiable cases (see Figures 3 and 4).

The GS approach is not effective for locally identifiable cases, where the regions of high values of the posterior PDF are well separated. To the authors' knowledge, all stochastic simulation approaches suitable for locally identifiable cases, such as importance sampling and the Markov chain Monte Carlo approach developed by Beck and Au (2002), are not efficient for problems with high-dimensional uncertain parameters. It remains an open research subject to sample a high-dimensional PDF with well-separated support regions. It is noted, though, that locally identifiable cases will be relatively rare in practice compared with globally identifiable, or unidentifiable, cases where the sensor data either provide more, or less, independent information than needed to constrain the updated parameter values (Katafygiotis and Beck, 1998).

ACKNOWLEDGMENTS

The first and second authors would like to acknowledge the financial support from Caltech's George W. Housner Postdoctoral Fellowship and Caltech's Harold Hellwig Graduate Fellowship, respectively.

REFERENCES

Beck, J. L. (1996), System identification methods applied to measured seismic response, in *Proceedings of the 11th World Conference on Earthquake Engineering*, Acapulco, Mexico.

Beck, J. L. & Au, S. K. (2002), Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation, *Journal of Engineering Mechanics*, **128**, 380–91.

Beck, J. L., Au, S. K. & Vanik, M. W. (2001), Monitoring structural health using a probabilistic measure, *Computer-Aided Civil and Infrastructure Engineering*, **16**, 1–11.

Beck, J. L. & Katafygiotis, L. S. (1998), Updating models and their uncertainties—Bayesian statistical framework, *Journal of Engineering Mechanics*, **124**, 455–61.

Beck, J. L., May, B. S. & Polidori, D. C. (1994), Determination of modal parameters from ambient vibration data for structural health monitoring, in *Proceedings of the First World Conference on Structural Contro*l, Pasadena, California, TA3:3–TA3:12.

Bernal, D., Dyke, S. J., Beck, J. L. & Lam, H. F. (2002), Phase II of the ASCE benchmark study on structural health monitoring, in *Proceedings of the 15th Engineering Mechanics Conference of the American Society of Civil Engineer*s, New York.

Caicedo, J. M., Dyke, S. J. & Johnson, E. A. (2004), Natural excitation technique and eigensystem realization algorithm for Phase I of the IASC-ASCE Benchmark Problem: Simulated data, *Journal of Engineering Mechanics*, **130**, 49–60.

Casciati, F. (ed.) (2002), in *Proceedings of the 3rd World Conference on Structural Control*, Como, Italy.

Chang, F. K. (ed.) (2001), in *Proceedings of the 3rd International Workshop on Structural Health Monitoring*, Stanford University, California.

Ching, J. & Beck, J. L. (2004), Bayesian analysis of the Phase II IASC-ASCE structural health monitoring experimental benchmark data, *Journal of Engineering Mechanics*, **130**, 1233–44.

Dyke, S. J. (2001), Website for ASCE Structural Health Monitoring Committee: Benchmarks. Available at http://wusceel.cive.wustl.edu/asce.shm/benchmarks.htm.

Gelfand, A. E., Hills, S. E., Racine-Poon, A. & Smith, A. F. M. (1990), Illustration of Bayesian inference in normal data models using Gibbs sampling, *Journal of American Statistical Association*, **85**, 972–85.

Gelman, A., Carlin, J. S., Stern, H. S. & Rubin, D. B. (1995), *Bayesian Data Analysis*, New York: Chapman & Hall.

Gelman, A. B. & Rubin, D. B. (1992), Inference from iterative simulation using multiple sequences, *Statistical Science*, **7**, 457–72.

Geman, S. & Geman, D. (1984), Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721–41.

Hjelmstad, K. D. & Shin, S. (1997), Damage detection and assessment of structures from static response, *Journal of Engineering Mechanics*, **123**, 568–76.

Jaynes, E. T. (1957), Information theory and statistical mechanics, *Physical Review*, **106**, 620–30. Also in Rosenkrantz, R. D. (ed.) (1982). *E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics*, Dordrecht, Holland: D. Reidel Publishing.

Johnson, E. A., Lam, H. F., Katafygiotis, L. S. & Beck, J. L. (2004), Phase I IASC-ASCE structural health monitoring benchmark problem using simulated data, *Journal of Engineering Mechanics*, **130**, 3–15.

Katafygiotis, L. S. & Beck, J. L. (1998), Updating models and their uncertainties: Model identifiability, *Journal of Engineering Mechanics*, **124**, 463–67.

Lam, H. F., Ko, J. M. & Wong, C. W. (1998), Localization of damaged structural connections based on experimental modal and sensitivity analysis, *Journal of Sound and Vibration*, **210**, 91–115.

Natke, H. G. & Yao, J. T. P. (eds.) (1988), in *Proceedings of the Workshop on Structural Safety Evaluation Based on System Identification Approaches*, Wiesbaden: Vieweg and Sons.

Sohn, H. & Law, K. H. (2001), Damage diagnosis using experimental Ritz vector, *Journal of Engineering Mechanics*, **127**, 1184–93.

Sohn, H., Farrar, C. R., Hunter, N. F. & Worden, K. (2001), Structural health monitoring using statistical pattern recognition techniques, *ASME Journal of Dynamic Systems, Measurement and Control*, **123**, 706–11.

Vanik, M. W., Beck, J. L., & Au, S. K. (2000), A Bayesian probabilistic approach to structural health monitoring, *Journal of Engineering Mechanics*, **126**, 738–45.

Yuen, K. V., Au, S. K. & Beck, J. L. (2004), Two-stage structural health monitoring approach for Phase I benchmark studies, *Journal of Engineering Mechanics*, **130**, 16–33.

APPENDIX—COMPUTATIONAL ISSUES

In the GS algorithm, to sample $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$ and $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$, their first two moments are needed, which, in turn, requires matrix inversion (see (9) for the term $(A\Sigma^0 A^T + \Sigma)^{-1}$). The size of this matrix is $N_s N_m(N_o + N_d) \times N_s N_m(N_o + N_d)$ for $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$ and $N_s N_m N_d \times N_s N_m N_d$ for $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ so the inversion of these matrices is computationally expensive. A sequential way of computing the first two moments of $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$ and $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ is proposed without the need to invert these huge matrices, which will be explained using $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ as an example. The first equation in (2) can be written as

$$\hat{Y}_{r,j} = A_{r,j}\lambda + \varepsilon_{r,j} \tag{A.1}$$

where

$$\begin{aligned} \hat{Y}_{r,j} &= K_0\hat{\phi}_r - \hat{\omega}_{r,j}^2 M_0\hat{\phi}_r \in R^{N_d} \\ A_{r,j} &= \left[-K_1\hat{\phi}_r \ \cdots \ -K_{n_K}\hat{\phi}_r \ \ \hat{\omega}_{r,j}^2 M_1\hat{\phi}_r \ \cdots \ \hat{\omega}_{r,j}^2 M_{n_M}\hat{\phi}_r\right] \\ &\quad \in R^{N_d \times n_\lambda} \end{aligned} \tag{A.2}$$

It follows that $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D}) = p(\lambda \mid \hat{\sigma}^2, \{\hat{D}_{r,j} : r = 1, \ldots, N_m, j = 1, \ldots, N_s\})$, where $\hat{D}_{r,j} \equiv \{\hat{Y}_{r,j}, A_{r,j}\}$. Instead of conditioning on $\{\hat{D}_{r,j} : r = 1, \ldots, N_m, j = 1, \ldots, N_s\}$ all together, first condition on $\hat{D}_{1,1}$ to obtain $p(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1})$ by computing its first two moments $\mathrm{E}(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1})$ and $\mathrm{Cov}(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1})$, which are readily computed using (9). During the computation, an $N_d \times N_d$ matrix inversion is required. Second, obtain $p(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}, \hat{D}_{1,2})$ by calculating its first two moments,

$$\begin{aligned} &\mathrm{E}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}, \hat{D}_{1,2}\right) \\ &\quad = \mathrm{E}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) + \mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) A_{1,2}^T \\ &\qquad \times \left[A_{1,2}\mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) A_{1,2}^T + \hat{\sigma}_1^2 I\right]^{-1} \\ &\qquad \times \left[\hat{Y}_{1,2} - A_{1,2}\mathrm{E}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right)\right] \\ &\qquad \times \mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}, \hat{D}_{1,2}\right) \\ &\quad = \mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) - \mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) A_{1,2}^T \\ &\qquad \times \left[A_{1,2}\mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) A_{1,2}^T + \hat{\sigma}_1^2 I\right]^{-1} \\ &\qquad \times A_{1,2}\mathrm{Cov}\left(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}\right) \end{aligned} \tag{A.3}$$

Note that in (A.3), only an $N_d \times N_d$ matrix inversion is needed. The sequential procedure continues for $p(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}, \hat{D}_{1,2}, \hat{D}_{1,3})$, $p(\lambda \mid \hat{\sigma}^2, \hat{D}_{1,1}, \hat{D}_{1,2}, \hat{D}_{1,3}, \hat{D}_{1,4}) \ldots$ until $p(\lambda \mid \hat{\sigma}^2, \hat{\phi}, \hat{D})$ is reached. For $p(\phi \mid \hat{\lambda}, \hat{\sigma}^2, \hat{D})$, the same sequential method is used but the size of the matrix inversion is now $(N_o + N_d) \times (N_o + N_d)$.