Computers and Structures xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method

J.M. Nichols [a,*], E.Z. Moore [b], K.D. Murphy [b]

[a] US Naval Research Laboratory, 4555 Overlook Ave., Washington, DC 20375, United States
[b] University of Connecticut, Dept. of Mechanical Engineering, Storrs, CT 06269-3139, United States

ARTICLE INFO

Article history:
Received 21 September 2010
Accepted 25 March 2011
Available online xxxx

Keywords:
Bayesian inference
Population-based Markov Chain Monte Carlo
System identification
Damage identification

ABSTRACT

Estimating damage in structural systems is a challenging problem due to the complexity of the likelihood function describing the observed data. From a Bayesian perspective a complicated likelihood means efficient sampling of the posterior distribution is difficult and standard Markov Chain Monte Carlo samplers may no longer be sufficient. This work describes a population-based Markov Chain Monte Carlo approach for efficient sampling of the damage parameter posterior distributions. The approach is shown to accurately estimate the state of damage in a cracked plate structure using simulated, free-decay response data. The use of this approach in identifying structural damage has not previously been explored.

Published by Elsevier Ltd.

1. Introduction

The problem of identifying damage in structural systems can be effectively cast as an estimation problem. That is to say: given a sequence of observed data collected from one or more sensors, estimate the damage presence, location, and extent. This information is necessary for predicting the future state of the structure in question, i.e. for prognostics.

There are two main classes of estimation approaches available: the method of maximum likelihood and Bayesian estimation. The former produces a single "best" estimate, defined as that which maximizes the probability of having observed a given set of data (e.g. structures' vibrational response). The latter views the parameter to be estimated as a random variable and the goal is to combine the likelihood with prior information to produce the estimated probability density function (PDF) associated with each parameter.

The difficulty with either approach in structural estimation problems is that the likelihood function (the core of both estimation approaches) is often an extremely complicated function of the structure's parameters with many maxima. This severely restricts the use of standard optimization algorithms. For example, a simple gradient ascent algorithm cannot be used to find the maximum likelihood estimate (MLE) in many structural system identification problems. This has been recognized by a number of researchers and solutions proposed. Horibe, for example, used a genetic algorithm (GA) to find the MLE of a structure's parameters in a simple vibration problem [1]. Likewise, Stull et al. considered a GA in generating MLEs of structural parameters in a static shell buckling problem [2]. In the context of damage detection Panigrahi et al. used a GA to perform parameter estimation by minimizing differences between predicted and observed modal properties [3] using a likelihood based on modal properties. Additional work by Hwang et al. also employed a cost function based on modal properties and used a GA to estimate stiffness in a composite structure [4]. The authors point out that for certain problems maximizing the likelihood does not necessarily require a GA. For example, if enough is known about the parameter distributions *a priori* and one has a good initial guess, the perturbation approach of Fonseca et al. [5] or Xu et al. [6] can be used to obtain MLEs of a structure's parameters. In the related field of structural reliability, researchers have also long recognized the challenges in structural optimization problems. The early work of Kiureghian and Dakessian [7], for example, developed a method for avoiding local optima in estimating failure probability in structures. More recently Guo et al. [8] developed an algorithm for efficiently obtaining globally optimal solutions for structural design problems possessing multiple local optima.

In this paper, however, the authors are interested in obtaining the entire parameter probability distribution. This distribution can be used to extract both the parameter estimate *and* credible intervals which provide a measure of confidence in the estimate. The authors view well-defined credible intervals as essential in

* Corresponding author.
E-mail address: jonathan.nichols@nrl.navy.mil (J.M. Nichols).

0045-7949/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.compstruc.2011.03.013

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEP 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Naval Research Laboratory,4555 Overlook Ave,Washington,DC,20375**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Estimating damage in structural systems is a challenging problem due to the complexity of the likelihood function describing the observed data. From a Bayesian perspective a complicated likelihood means efficient sampling of the posterior distribution is difficult and standard Markov Chain Monte Carlo samplers may no longer be sufficient. This work describes a population-based Markov Chain Monte Carlo approach for efficient sampling of the damage parameter posterior distributions. The approach is shown to accurately estimate the state of damage in a cracked plate structure using simulated, free-decay response data. The use of this approach in identifying structural damage has not previously been explored.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

damage detection application. Optimal decisions regarding the maintenance and usage of a structure are predicated on quantifying the uncertainty in the damage estimate. To this end the Bayesian estimation framework is conceptually the most appealing. In Bayesian estimation one combines prior information with the likelihood to produce the desired posterior parameter estimate A number of researchers have had success using a Markov-Chain-Monte-Carlo (MCMC) approach to explore the posterior distribution in structural parameter estimation problems. Numerous works by Beck et al. (see e.g. [9–11]) have used MCMC methods to estimate structural parameters. Additional work by Glaser et al. [12] illustrated the approach in detecting stiffness reduction in beams using static measurements. Zhang and Cho [13] also used the MCMC approach to help design an evolutionary algorithm for performing system identification.

The MCMC approach (to be described) works by repeatedly perturbing the parameter(s) of interest and either accepting or rejecting the proposals based on a well-defined criteria. However, the MCMC algorithm can become "stuck" in a local minima, again due to the complicated likelihood function (see Section 2). This is, of course, why so many researchers use GAs to explore complicated parameter spaces in obtaining MLEs. As a result, recent works have proposed to fuse GAs with MCMC in order to more efficiently explore the posterior distribution. Efforts to this end include the work of Vrugt et al. [14] who combined differential evolution (a genetic algorithm search technique) with MCMC to draw samples from parameters in a complex soil moisture model. Also noted is the work of Zhang and Cho [13] who combined the searching capabilities of an evolutionary algorithm with MCMC to produce faster convergence of parameter estimates using data generated from a laser system. The use of multiple solutions that can exchange information as they explore the posterior have come to be known as "population-based MCMC" methods. A nice overview of population-based MCMC (Pop-MCMC) methods is provided by Jasra et al. [15] where the initial idea was credited to Geyer [16]. Other recent Pop-MCMC works include the imaging application of Kim et al. [17]. Although this approach has seen recent use in these and other complicated estimation problems it has not yet been used in structural dynamics problems despite obvious advantages over conventional MCMC in sampling complicated posterior distributions.

Our goal in this manuscript is therefore to adopt the Pop-MCMC approach to a difficult parameter estimation problem in structural dynamics, namely the estimation of parameters governing crack damage in an aluminum plate. The specific implementation of the pop-MCMC algorithm is new as is the application to structural vibration problems. The population-based approach is shown to provide reliable estimates of the crack parameters under realistic levels of noise. These estimates are not always possible using standard MCMC as will be shown. Additionally, implementation of the Pop-MCMC algorithm requires that the forward model be computationally efficient. Section 4 is therefore devoted to developing a low-dimensional finite element (FE) crack model. The model uses specially tailored elements, resulting in a model that retains the fidelity of the standard FE approach without the computational overhead.

2. Brief review of Bayes rule and MCMC

Given data from a system of interest, denoted by the vector of N observations $\mathbf{s}$, and a model that describes those data $\mathbf{w}(\boldsymbol{\theta})$, the practitioner's job is to estimate the model parameter vector $\boldsymbol{\theta} \equiv (\theta_1, \theta_2, \ldots, \theta_P)$. Typically one takes "good" estimates to be those that perform well in the face of the inevitable uncertainty (e.g. noise) in the data. In this paper it is assumed that the uncertainty takes the form of additive noise such that our measured data are

$$\mathbf{s} = \mathbf{w}(\boldsymbol{\theta}) + \boldsymbol{\xi} \tag{1}$$

where the noise vector $\boldsymbol{\xi} \equiv \xi_i$, $i = 1, \ldots, T$ is comprised of T, zero-mean, iid entries with joint probability distribution

$$p_H(\boldsymbol{\xi}) = \frac{1}{(2\pi\sigma^2)^{T/2}} e^{-\sum_i^T \xi_i/2\sigma^2} \tag{2}$$

Thus, the noise is assumed to be stationary, Gaussian with variance σ^2. The *likelihood function* is then formed by simply substituting $\xi_i = s_i - w_i(\boldsymbol{\theta})$ giving

$$p_H(\boldsymbol{\xi}) \equiv p_H(\mathbf{s}|\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{T/2}} e^{-\sum_i^T (s_i - w_i(\boldsymbol{\theta}))^2/2\sigma^2} \tag{3}$$

which is the probability of having observed the sequence of data $\mathbf{s}$ given the model, defined by parameters $\boldsymbol{\theta}$. The aforementioned MLEs are obtained by maximizing Eq. (16). However, the goal is to obtain estimates of both the parameter and the amount of uncertainty in the parameter estimate. For this reason the Bayesian estimation philosophy is adopted thus providing estimates of the entire parameter posterior distribution from which credible intervals for our parameters can easily be obtained.

Bayes' rule states that the joint posterior distribution of our model parameters may be found by the relationship

$$p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = \frac{p_H(\mathbf{s}|\boldsymbol{\theta}) p_\pi(\boldsymbol{\theta})}{p_D(\mathbf{s})} \tag{4}$$

where $p_\pi(\boldsymbol{\theta})$ is the joint prior parameter distribution, reflecting any *a priori* information one might have about our parameter, and $p_D(\mathbf{s})$ is the joint distribution of our acquired data. Of course the above expression is for the joint parameter distribution whereas one is typically interested in the marginal (individual) posteriors, i.e. $p_{\Theta_p}(\theta_p)$. Analytically this would require integrating Eq. (4) over the other $P-1$ parameters, e.g.

$$p_{\Theta_p}(\theta_p) = \int_{\mathbb{R}^{P-1}} p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}_{-p} \tag{5}$$

where the notation $\int_{\mathbb{R}^{P-1}} d\boldsymbol{\theta}_{-p}$ denotes the multi-dimensional integral over all parameters other than θ_p. This cannot typically be done in closed form and instead the authors resort to numerical methods.

The Markov Chain Monte Carlo (MCMC) algorithm was proposed by Hastings [18] as a means of drawing samples from Eq. (4) directly (without having to integrate). The MCMC algorithm is now fairly standard and the reader is referred to [19] for a description in the context of structural dynamics. The algorithm builds a Markov chain for each parameter such that samples from the chain are, in fact, samples from the desired posterior distribution. The chain is formed by first proposing state transitions for each parameter via a so-called proposal distribution, $q(\theta_p^*|\theta_p)$. This distribution provides a rule for generating a candidate parameter value θ_p^* given the current value in the chain θ_p. The proposed parameter (keeping all other parameters fixed at their current value) is then accepted or rejected with probability

$$r_{MH} = \min\left\{ \frac{p_H(\mathbf{s}|\theta_p^*) p_{\pi_p}(\theta_p^*)}{p_H(\mathbf{s}|\theta_p) p_{\pi_p}(\theta_p)}, 1 \right\}. \tag{6}$$

This procedure is referred to as the Metropolis–Hastings (MH) algorithm after the originators [20,18]. The same is done for each of the $p = 1, \ldots, P$ parameters in $\boldsymbol{\theta}$ while holding the other $P-1$ parameters fixed. Thus, the approach is really drawing samples from the conditional posterior $p_{\Theta_p}(\theta_p|\boldsymbol{\theta}_{-p})$. In other words, given that the other parameters are fixed to their current values in the chain, sample the posterior for parameter p. This procedure is referred to as Gibbs sampling and eliminates the need to perform the high-dimensional integral required by Eq. (5). The above-described procedure

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

constitutes one iteration of the Markov chain. Starting with initial parameter values chosen from their respective prior distributions $\theta_p(0) \sim p_{\pi_p}(\theta_p)$ this process (accept/reject) repeats K times yielding the samples $\theta_p(k)$, $k = 1 \ldots, K$.

The distribution $q(\theta_p^*|\theta_p)$ can be thought of as a rule for perturbing θ_p. The magnitude of this perturbation will be decided by the parameters associated with $q(\cdot)$. For example, it is common to choose the proposal distribution to be the Uniform distribution centered at θ_p, i.e.

$$q(\theta_p^*|\theta_p) = U(\theta_p - A, \theta_p + A) \tag{7}$$

where A is a real constant that plays the role of the perturbation size. Typically A is tuned dynamically so that on average $r_{MH} = 0.3 - 0.5$ (see [19] for a more thorough discussion and for sample code). Heuristically it is easy to see how the approach works. The algorithm continually perturbs each parameter, checks whether or not a better fit (consistent with the prior) is achieved by computing r_{MH}, and keeps parameter values that do well in this regard. The fact that in the end the chain of values are samples from $p_{\Theta_p}(\theta_p)$ is actually quite remarkable. However, one can immediately see where problems can arise.

Consider a bimodal posterior distribution of a single parameter θ,

$$p_\Theta(\theta) = \frac{1}{2\sqrt{2\pi f\sigma^2}} e^{-(\theta-\mu)^2/(2f\sigma^2)} + \frac{1}{2\sqrt{2\pi(1-f)\sigma^2}} e^{-(\theta+\mu)^2/(2(1-f)\sigma^2)} \tag{8}$$

for known, positive constants f, σ, μ. The constant $f < 1$ specifies the fraction of the distribution variance associated with the distribution peaks at $+\mu$, $-\mu$. For this example the values $\sigma = 1$, $f = 0.3$, were used along with two different values for the peak locations, $\mu = 2,4$. The resulting distribution presents difficulties for the conventional MCMC algorithm for the $\mu = 4$ case as shown in Fig. 1. For $\mu = 2$ the peaks of the distributions are close enough that the algorithm can easily move back and forth between the two regions of high probability. However for $\mu = 4$ the Markov Chain quickly becomes "trapped" in a single portion of the posterior distribution. This is simply due to the fact that the proposal distribution is not capable of moving the chain easily from one peak to the other. The solution, however, is not as simple as changing the proposal distribution to give us larger perturbations to our chain (i.e. increase A). If the proposal values are very far from the existing values they will almost always be rejected, thus the Markov Chains will take an extremely long time to converge. An efficient sampler is one that would allow us to *locally* explore a high probability region of the posterior while simultaneously provide a mechanism for covering large distances in parameter space to reach other high probability regions. The population-based MCMC approach, described next, was designed specifically for this reason.

3. Population-based MCMC

As with standard MCMC, the goal of the Pop-MCMC algorithm is to draw samples from some desired posterior distribution $p_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$. A very nice introduction to the topic of Population-based MCMC (Pop-MCMC) is given by Jasra et al. [15]. The basic idea is to first create a new, composite posterior density

$$p_C(\boldsymbol{\theta}^{(1:N)}) = \prod_{n=1}^{N} p_n(\boldsymbol{\theta}^{(n)}) \tag{9}$$

which is a function of the composite parameter vector $\boldsymbol{\theta}^{(1:N)} = (\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(N)})$. It is required that $p_n(\boldsymbol{\theta}^{(n)}) = p_{\boldsymbol{\Theta}}(\boldsymbol{\theta})$ for at least one n, i.e. one of the posteriors comprising the composite density is the true posterior distribution. In what follows the authors use $n = 1$ to denote the true posterior distribution from which the desired samples will be drawn. If an irreducible, a-periodic Markov chain can be constructed that has $p_C(\boldsymbol{\theta}^{(1:N)})$ has its invariant distribution, then samples from the marginal distribution

$$p_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = \int_{\mathbb{R}^{N-1}} p_C(\boldsymbol{\theta}^{(1:N)}) d\boldsymbol{\theta}^{(2:N)}$$

can be drawn where the notation $\int_{\mathbb{R}^{N-1}} d\boldsymbol{\theta}^{(2:N)}$ denotes the multi-dimensional integral over all parameters vectors other than $\boldsymbol{\theta} \equiv \boldsymbol{\theta}^{(1)}$. This is accomplished numerically by running "N" Markov chains concurrently, each exploring its own posterior distribution $p_n(\boldsymbol{\theta}^{(n)})$. Each chain can be considered in turn, holding the others fixed (Gibbs sampling), and samples from $p_1(\boldsymbol{\theta}^{(1)})$ are retained as samples from the desired joint posterior.

There exists a good deal of freedom in choosing the $p_n(\cdot)$ and in selecting different types of proposals for exploring the parameter space. With regard to the former one would like to wisely choose the $p_n(\cdot)$ so as to facilitate easy exploration of the parameter space. While other approaches can be used (see [15]), in this paper the so-called *tempered* sequence of distributions is used

$$p_n(\boldsymbol{\theta}^{(n)}) \equiv p_{\boldsymbol{\Theta}}^{\zeta_n}(\boldsymbol{\theta}^{(n)}) \tag{10}$$

for $\zeta_n \in (0,1]$. For $\zeta_1 = 1$, of course, one has the true posterior. For smaller values of ζ_n one obtain successively smoother versions of the original posterior. The idea is that the smoothed distributions are easier for their corresponding Markov chains to explore, yet they are still related to the true posterior and therefore still carry information about high probability regions of the parameter space. As an example consider $N = 4$ separate instances of the bi-modal distribution (8) raised to the $\zeta_n = 1$, 0.5, 0.1, 0.0125 powers, respectively. These distributions are shown in Fig. 2. The standard MCMC algorithm will have an easier time exploring these smoother parameter spaces. Of course only samples from the distribution with $\zeta_1 = 1$ are of interest, however the additional distributions can clearly facilitate efficient sampling if one has a means of passing information between chains. The chain associated with the true posterior needs to be informed by the chains exploring the smoother distributions.

Perhaps the simplest type of move to accomplish this is the so-called *swap* move. This is similar to a Metropolis–Hastings move where the proposal is to consider swapping the parameter values in chains u, v. Assuming an equal probability of selecting chains u, v from the N possibilities, this swap is accepted with probability

$$r_{swap} = \min\left\{\frac{p_u(\boldsymbol{\theta}^{(v)}) p_v(\boldsymbol{\theta}^{(u)})}{p_u(\boldsymbol{\theta}^{(u)}) p_v(\boldsymbol{\theta}^{(v)})}, 1\right\}. \tag{11}$$

Typically swap moves are not proposed after every iteration in the Markov Chains but are performed with some probability. Returning to the bi-modal example, consider $N = 4$ chains exploring the composite target

$$p_C(\boldsymbol{\theta}^{(1:4)}) = \prod_{n=1}^{4} p_{\boldsymbol{\Theta}}^{\zeta_n}(\boldsymbol{\theta}^{(n)}), \tag{12}$$

thus one has four Markov chains running concurrently with $\zeta_n = 1$, 0.75, 0.5, 0.25, respectively. After each iteration in the Markov chain a swap move is performed with 50% probability. This move consists of uniformly selecting two of the chains and evaluating Eq. (11). If the proposal is accepted the values in the chains are exchanged and the algorithm continues to the next iteration in the Markov chains. Fig. 3 shows the results of this sampler All four Markov chains are informing each other as to the presence of multiple peaks in the distribution. The end result is that the chain associated with $\zeta_1 = 1$ contains samples from the desired posterior distribution; this

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

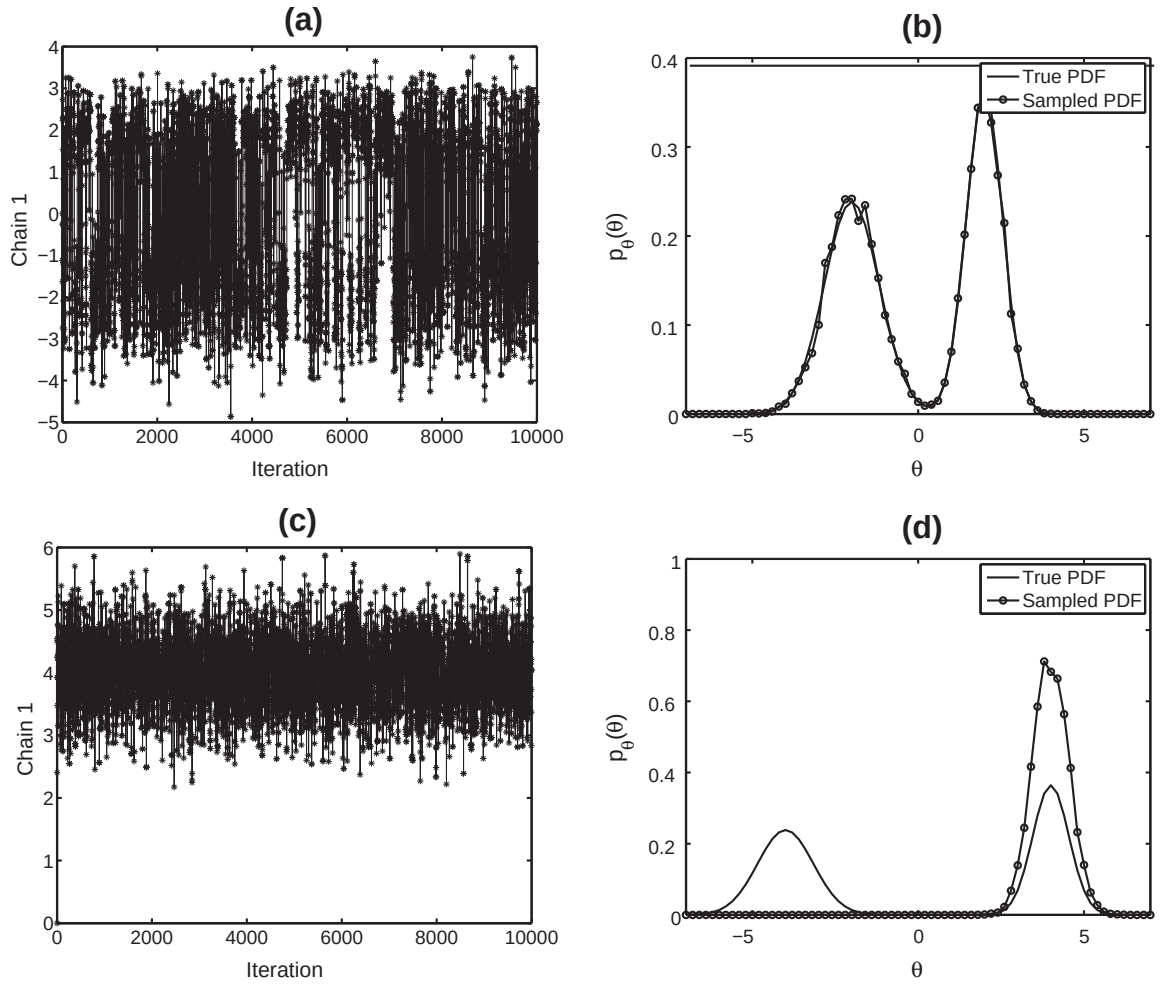

Fig. 1. Markov Chains and estimated bi-modal posterior distribution given by Eq. (8) with $\mu = 2$ (a,b) and $\mu = 4$ (c,d), respectively.

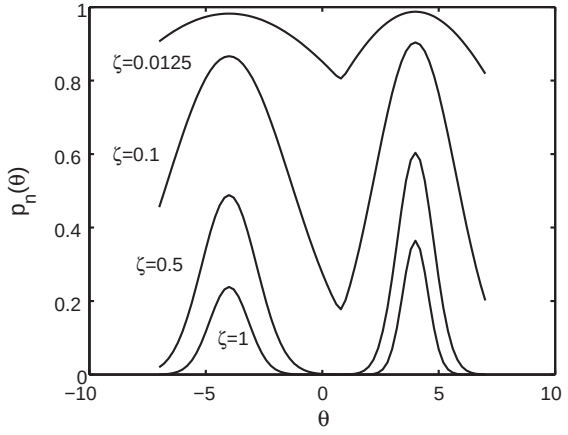

Fig. 2. Successive $p_n(\theta) = p_\theta^{\zeta_n}(\theta)$ corresponding to the bi-modal posterior distribution Eq. (8). The true posterior distribution from which samples are sought is given by $\zeta_1 = 1$.

distribution is shown in Fig. 3b and compares favorably to the true posterior density.

While the swap move is an effective means of communicating among chains, a more sophisticated type of move is required for the structural estimation problem considered next. Particularly for multivariate parameter estimation, it is useful to design a move that can hold multiple parameters fixed while allowing other groups of parameters to move simultaneously. Such a move is not allowed with the standard Metropolis-within-Gibbs sampling (one parameter moved at a time) or by the swap move which moves all parameters simultaneously. A particularly effective move is to use *differential evolution* to generate a trial vector. Differential evolution (DE) is the engine of a popular GA used in searching complex parameter space [21]. The approach draws at random three members of the population, $\boldsymbol{\theta}^{(u)}$, $\boldsymbol{\theta}^{(v)}$, $\boldsymbol{\theta}^{(w)}$ and generates the trial vector

$$\boldsymbol{\theta}^{(u')} = \boldsymbol{\theta}^{(u)} + \gamma(\boldsymbol{\theta}^{(v)} - \boldsymbol{\theta}^{(w)}) \tag{13}$$

where γ is a user-defined constant. Each of the P elements in this trial vector replace the elements of the original vector $\boldsymbol{\theta}^{(u)}$ with 50% probability. This final step (keep new element or retain old) emulates the "cross-over" step common to most GAs. Once the trial vector has been generated it is accepted/rejected using r_{MH} (Eq. 6). However, in order to differentiate among the types of moves the acceptance ratio for the DE move will be denoted r_{DE}.

It should mentioned that others have proposed a straight *cross-over* move whereby the parameter vector is split at a point $m < P$ for two "parent" vectors and generating the trial vectors:

$$\begin{aligned}\boldsymbol{\theta}^{(u')} &= [\theta_1^{(u)}, \ldots, \theta_m^{(u)}, \theta_1^{(v)}, \ldots, \theta_P^{(v)}] \\ \boldsymbol{\theta}^{(v')} &= [\theta_1^{(v)}, \ldots, \theta_m^{(v)}, \theta_1^{(u)}, \ldots, \theta_P^{(u)}]\end{aligned} \tag{14}$$

and accepting with probability r_{swap} (Eq. 11)[15]. This is indeed a useful way of exchanging information between chains, however

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

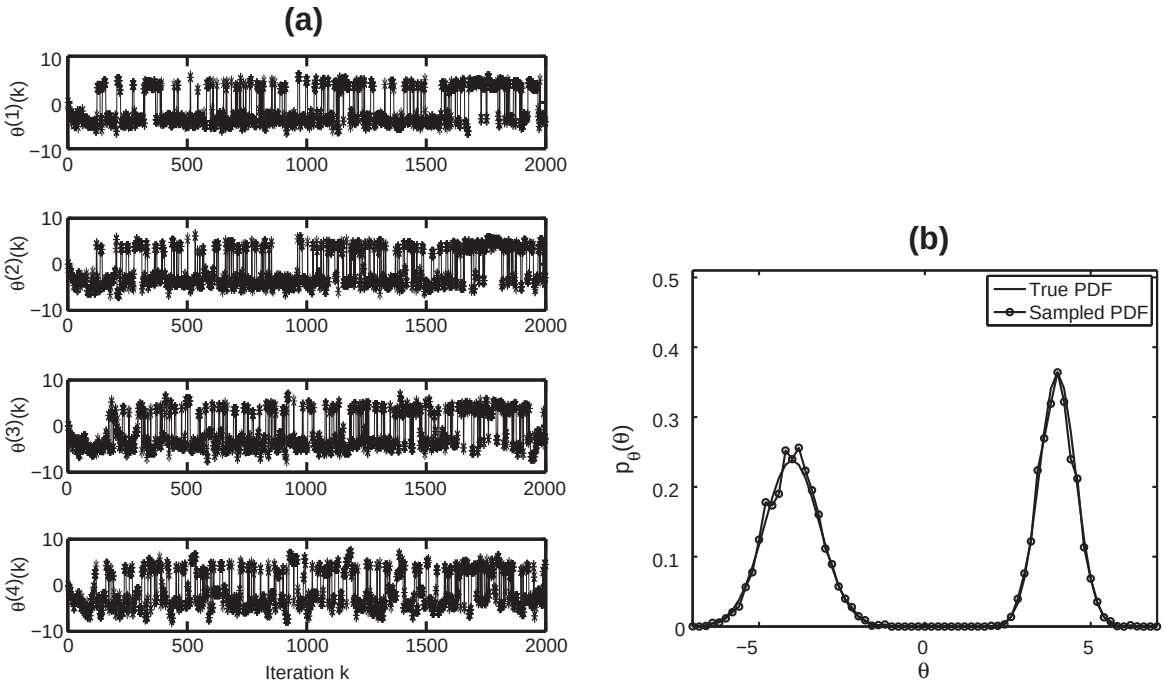

Fig. 3. (a) Markov chains and (b) estimated posterior distribution associated with the bi-modal distribution given by Eq. (8).

the authors have found that this mechanism is already provided for in the differential evolution move, thus this type of move is not used in this paper.

The above-described moves clearly borrows from the genetic algorithm (GA) approach to optimization problems. In fact, Pop-MCMC is effectively combining the efficient search capabilities of GAs with the power of the Bayesian MCMC approach to sampling. Both types of moves, swap and differential evolution, will be used in the structural dynamics example presented in subsequent sections.

4. An efficient cracked plate model

One critical component of this process is a computationally fast model that is also flexible enough to incorporate a variety of damage modes. The model predicts the lateral deflection of the plate at a given point in space and time, $w(x,y,t)$. Efficiency of the forward model is essential as the Pop-MCMC algorithm requires a very large number of model-to-date comparisons, i.e. evaluate r_{MH}, r_{swap}, r_{DE}. Flexibility is also desired because while the current study considers damage consisting of only a single straight crack, it is the model alone that limits the type of damage that can be considered by this technique. A model that could handle multiple cracks, branched cracks, edge cracks, or corrosion would be valuable. A purely analytical model entailing virtually no computation time would be ideal on both counts. Unfortunately, analytical solutions for the dynamic response of a rectangular plate with an arbitrary crack (position, orientation, and size) do not presently exist to the knowledge of the authors. Solecki [22] produced a solution for the natural frequencies of a plate with a single crack, but did not extend it to the myriad of other situations of future interest. A second possible family of models are finite element solutions. These clearly have the desired flexibility, but involve time marching and so are computationally intense. In this present work, a compromise approach is taken. A finite element eigen-solution is used to build in the features of the crack singularity. Building off of these numerical frequencies and mode shapes, modal analysis is used to describe the actual motion in space and time, $w(x,y,t)$. The eigen-solution to this finite element model gives the modes in terms of the nodal displacements. The values at the instrumentation locations are interpolated via the natural neighbor method. Standard modal analysis proceeds from there. This hybrid solution is many times faster than time marching, while retaining the flexibility inherent in the finite element approach.

An in-house finite element code was developed for this particular application, so as to more easily wrap the MCMC process (code) around the time series solution. The model parameters are the location of the center of the crack $(x_{\text{crack}}, y_{\text{crack}})$, the crack length (a), and the orientation of the crack measured from the positive x-axis (α). See Fig. 4c. The elements away from the crack tip are standard eight-noded quadrilateral Mindlin serendipity elements, with nodes at each corner and at the middle of each side. Adjacent to the crack tip the eight-noded quads are modified as described in Refs. [23,24]. In short, two corner nodes and the node in between them are moved to a single location, leaving a triangular element. The adjacent side nodes are then moved from the midpoints to one quarter of the side length. The result is an element with a stress field that varies as $\frac{1}{\sqrt{r}}$, where r is the distance from the collapsed node (to be placed on the crack tip). See Fig. 4a and b for an illustration. This stress distribution exactly matches the Mitchell solution[25], which is the analytic stress field near a static crack tip. Because these augmented triangular elements capture the crack tip behavior, they are placed around the crack tip in a pinwheel fashion. See Fig. 4d. This permits a more sparse mesh near the crack tip than would otherwise be necessary. This reduces the number of degrees of freedom required and, hence, the computation time.

The cracked plate model involves several standard assumptions. First, the model is linear and it is assumed that material properties are known exactly (though these could be left as unknown parameters and found via the Bayesian/MCMC approach). The deflections are presumed small and crack growth is not considered; the latter is reasonable under small deflections. The crack is also assumed to remain open, such that impacts at the crack interface are ignored. Any mass lost loss due to the crack is

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

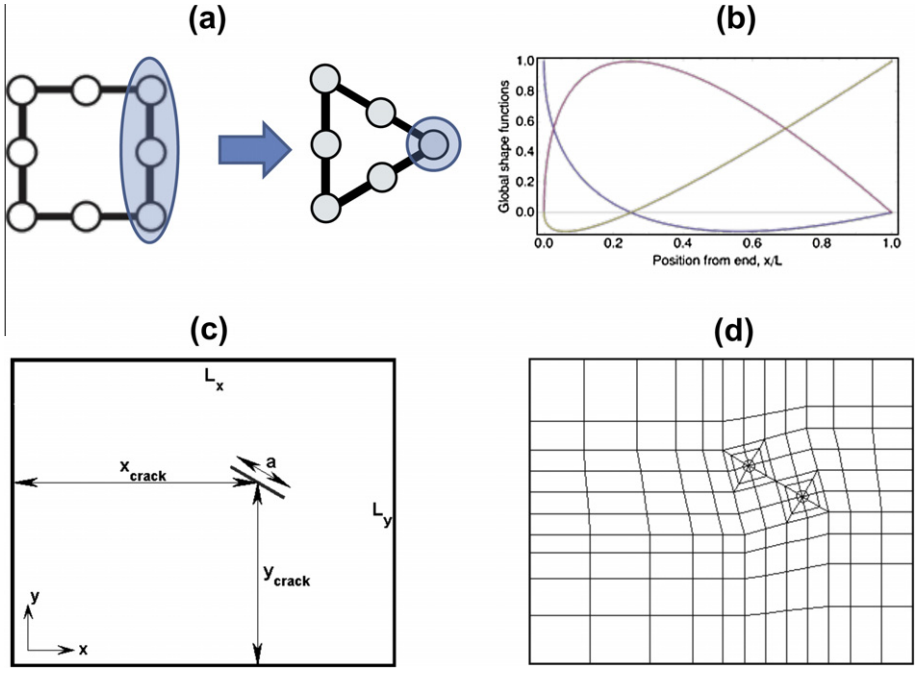

Fig. 4. (a) Nodal configuration of the Mindlin serendipity element, (b) resulting interpolation functions, (c) schematic of plate with crack parameters and (d) resulting course FEM mesh.

presumed negligible. Finally, there is only one crack and the tips are at least one half of the crack length from the edges of the plate. Since finding smaller cracks is the more interesting and useful problem, the allowable zone is still a very large fraction of the plates area. Even taken together, the authors do not believe that these requirements are too onerous.

As with any finite element analysis, convergence of the mesh must be verified. Due to the large number of iterations, it is not practical to verify convergence for every perturbed set of parameters. Instead, because all the meshes used are qualitatively similar, convergence was checked on several representative parameter vectors.

5. Implementation

In this paper the structure of interest assume a clamped rectangular plate measuring 1.25 m × 1 m with thickness $h = 0.01$ m with material properties $E = 209$ GPa (Young's modulus), $\nu = 0.3$ (Poisson's ratio), and $\rho = 7850$ kg/m^3 (density). It will also be assumed that the plate has been instrumented with M displacement sensors capable of sampling the plate's response to an input at $i = 1,\ldots,T$ equally spaced points in time. The observed signal model is therefore written (as before)

$$s_{ij} = w_{ij}(\boldsymbol{\theta}) + \xi_{ij}, \quad i = 1,\ldots T,\ j = 1,\ldots,M \tag{15}$$

where the ξ_{ij} are taken as realizations of an iid, zero-mean, Gaussian random process, i.e. each $\xi_{ij} \sim N(0,\sigma^2)$. The model (described in the previous section) is therefore evaluated at the same times $i = 1,\ldots,T$ and locations $j = 1,\ldots,M$ as the observed data. Under this noise model the likelihood function for the data is given by [19]

$$p_H(\mathbf{s}|\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{TM/2}} e^{-\frac{1}{2\sigma^2}\sum_i^T\sum_j^M (s_{ij}-w_{ij}(\boldsymbol{\theta}))^2} \tag{16}$$

Denoting the sum-squared error over all sensors by

$$Q(\mathbf{s},\boldsymbol{\theta}) = \sum_i^T \sum_j^M (s_{ij} - w_{ij}(\boldsymbol{\theta}))^2,$$

the sequence of tempered posterior distributions is formed as

$$p_n(\boldsymbol{\theta}^{(n)}) = p_H(\mathbf{s}|\boldsymbol{\theta}^{(n)})^{\zeta_n} p_\pi(\boldsymbol{\theta}^{(n)}) = \frac{p_\pi(\boldsymbol{\theta}^{(n)})}{(2\pi\sigma_n^2)^{\zeta_n TM/2}} e^{-\frac{\zeta_n}{2\sigma_n^2}Q(\mathbf{s},\boldsymbol{\theta}^{(n)})}$$
$$n = 1,\ldots,N \tag{17}$$

using the sequence $\zeta_1 = 1.0$,

$$\zeta_{n+1} = \zeta_n - \frac{1}{N}, \quad n = 1,\ldots,N-1$$

as suggested by Jasra et al. [15]. The idea here, as in the toy example, is to explore a composite posterior where the "smoothed" marginal posteriors $n = 2,\ldots,N$ are related to the true posterior of interest, $p_1(\boldsymbol{\theta}^{(1)})$.

This implementation of the pop-based algorithm proceeds as follows. The parameter vectors for each of the N chains are initialized by drawing samples from the priors. Then, for each iteration in the Markov chain, one of the chains $n \in N$ is selected with uniform probability, and a standard MH update is performed for each of the P parameters in $\boldsymbol{\theta}^{(n)}$ using the Gibbs sampling strategy. Thus, for each parameter $p = 1,\ldots,P$ one generates a candidate value using Eq. (7), evaluates the ratio

$$r_{MH} = \mathrm{Exp}\left[-\frac{\zeta_n}{2\sigma_n^2}\left(Q(\mathbf{s},\theta_p^{*(n)}|\boldsymbol{\theta}_{-p}^{(n)}) - Q(\mathbf{s},\theta^{(n)}|\boldsymbol{\theta}_{-p}^{(n)})\right)\right]\frac{p_\pi(\theta_p^{*(n)})}{p_\pi(\theta_p^{(n)})}$$

and accepts with probability $\min(r_{MH},1)$. The (unknown) noise variance associated with the nth chain, σ_n, also needs to be sampled. It has been demonstrated [19] that by choosing a vague prior for this parameter, one may directly sample from the posterior via

$$\sigma_n^2 \sim \frac{1}{\Gamma(MT/2, 2/(\zeta_n Q(\mathbf{s},\boldsymbol{\theta}^{(n)}))}$$

where Γ(a,b) denotes the Gamma distribution with parameters a,b. Once each of the parameters (including the noise variance) have been

updated for chain n, two different chains u, $v \in [1,N]$ are selected at random with uniform probability. With 50% probability either a "swap" move or a "DE" move between these chains is performed. For the swap move, the prior ratios will cancel, thus one evaluates

$$r_{swap} = \mathrm{Exp}\left[-\zeta_u\left(\frac{1}{2\sigma_v^2}Q(\mathbf{s},\boldsymbol{\theta}^{(v)}) - \frac{1}{2\sigma_u^2}Q(\mathbf{s},\boldsymbol{\theta}^{(u)})\right) - \zeta_v\left(\frac{1}{2\sigma_u^2}Q(\mathbf{s},\boldsymbol{\theta}^{(u)}) - \frac{1}{2\sigma_v^2}Q(\mathbf{s},\boldsymbol{\theta}^{(v)})\right)\right], \tag{18}$$

accepting the move with probability $\min(r_{swap},1)$. For the DE move one requires three randomly drawn chains (see Eq. 13) to generate the trial vector $\boldsymbol{\theta}^{(u')}$. The ratio of priors in the acceptance criteria is also required as they do not cancel out. One therefore evaluates

$$r_{DE} = \mathrm{Exp}\left[-\frac{\zeta_u}{2\sigma_u^2}\left(Q(\mathbf{s},\boldsymbol{\theta}^{(u')}) - Q(\mathbf{s},\boldsymbol{\theta}^{(u)})\right)\right] \times \frac{\prod_p^P p_\pi(\theta_p^{(u')})}{\prod_p^P p_\pi(\theta_p^{(u)})}, \tag{19}$$

and accept the move with probability $\min(r_{DE},1)$. To summarize, the algorithm picks a chain at random and performs a standard MH update on each of the parameters, including the noise variance. The algorithm then selects chains at random and performs either a swap move (Eq. 18) or crossover (Eq. 19) with 50% probability. This

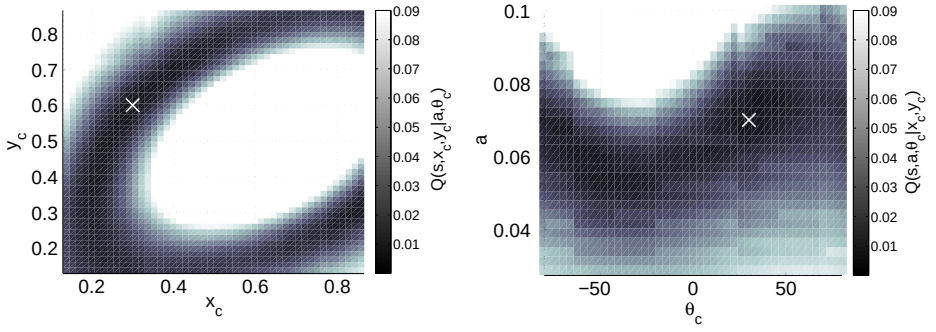

Fig. 5. Main argument of the likelihood function, $Q(\mathbf{s},\boldsymbol{\theta})$, plotted as a function of (a) crack location on the plate and (b) the parameters a (crack length) and θ_c, crack orientation. For each plot the remaining parameters were held fixed at their true values. True parameter values are denoted with an "X".

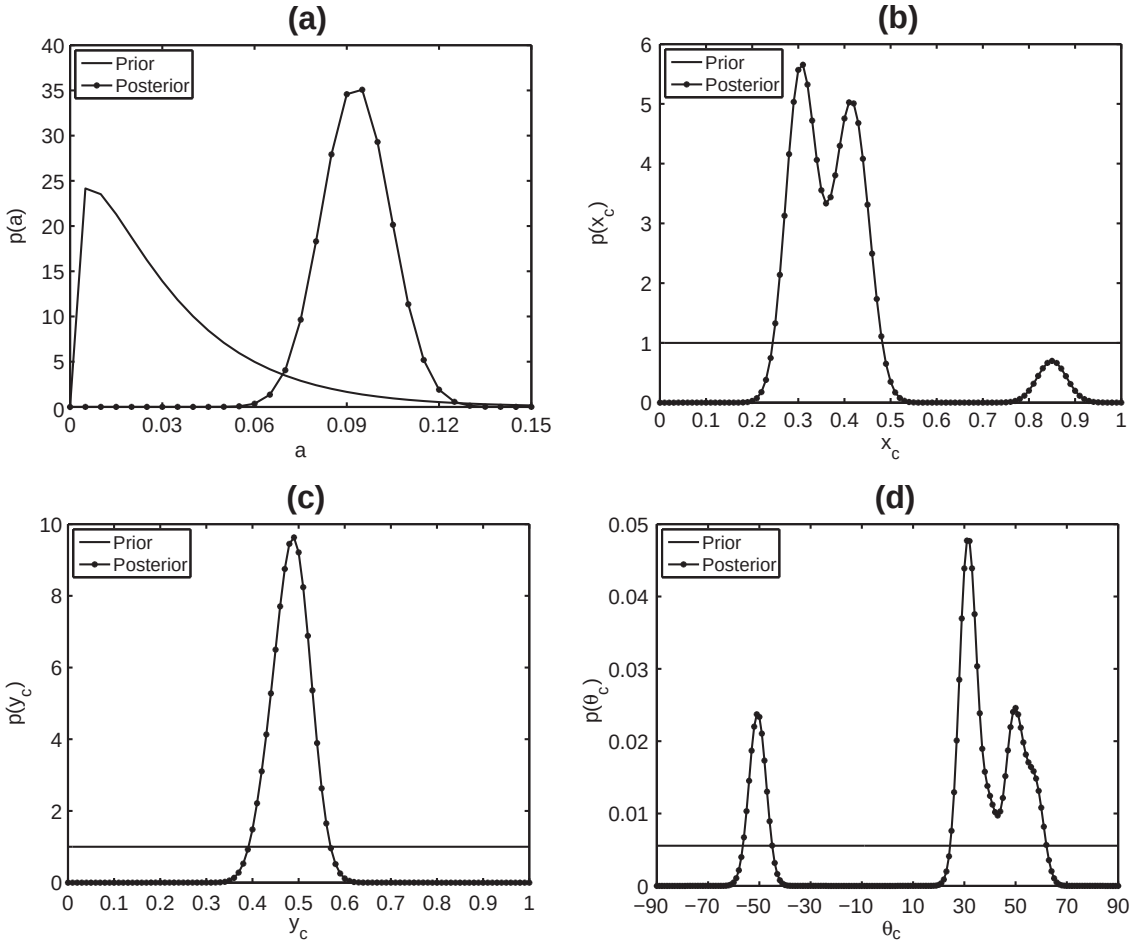

Fig. 6. Estimated posterior densities associated with (a) crack length, (b,c) crack location, and (d) crack orientation. The true parameter values are $a = 0.1$, $(x_c,y_c) = (0.3,0.5)$, and $\theta_c = 30$.

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

procedure repeats for some number of iterations until enough samples have been drawn from the posterior distribution.

The algorithm is fairly simple to implement in software, however is clearly computationally intensive. Each evaluation of a ratio (r_{MH}, r_{swap} or r_{DE}) requires solving the forward model described in Section 4. The efficiency of the forward model is therefore of great importance in using MCMC in structural system identification problems.

6. Cracked plate identification

There are four parameters that determine the state of damage in the plate model: crack length a, crack location (x_c, y_c), and crack orientation θ_c. These damage parameters are fixed to the values $a = 0.1$ m, $x_c = 0.3$ m, $y_c = 0.5$ m, and $\theta_c = 30°$. Further assume that four displacement sensors have been placed on the surface of the plate at the x–y locations (0.375, 0.375), (0.375, 0.862), (0.862, 0.375), (0.862, 0.862). The acquired data will consist of each sensor's response to four separate impacts (hammer strikes) at locations (0.29, 0.275), (0.29, 0.725), (0.96, 0.275), and (0.96, 0.275). These sensor locations were chosen so as to maximize the sum of the first four modes in the response. No claims of optimality are made regarding this choice, in fact finding sensor locations that produce a more well-defined likelihood is an active area of research.

The data used in the identification procedure consist of four simulated impulse response signals, sampled at 4 kHz for a duration of 2 s ($T = 8000$ observations). The first nine modes were used in generating the solution, all well below the Nyquist frequency of 2 kHz. The signal to noise ratio was set at 17 dB (50:1) corresponding to the level of noise observed in previous experiments.

Given this response the goal is to estimate the posterior distributions associated with each of the crack parameters. Using the standard MCMC algorithm the authors were unable to consistently identify the parameters due to the aforementioned problem with exploring complicated likelihood functions. Fig. 5 shows the argument of the likelihood, $Q(\mathbf{s}, \boldsymbol{\theta})$ as a function of different parameter combinations (holding the others fixed at their true values) As with the simple 1-D example (Fig. 1), one can immediately see the difficulty presented by this estimation problem. Consider Fig. 5a. The location parameter x_c has a minimum at the true location, $x_c = 0.3$, however, there is also a "trough" of minima with a particularly low value at ($x_c = 0.85$, $y_c = 0.28$). A similarly complicated likelihood is shown for the parameters a, θ_c (Fig. 5b). These plots illustrate the fundamental difficulty in parameter identification for structural systems: *multiple damage states can yield very similar structural response data.*

For this reason the authors have gravitated toward the population-based approach to sampling the posterior. In the example that follows, $N = 10$ chains were used for each of the parameters a, θ_c, x_c, y_c. The prior for the crack parameter a is take to be a Gamma distribution, $p_{\pi_a}(a) = a^{\alpha-1}\frac{\mathrm{Exp}[-a/\beta]}{\Gamma(\alpha)\beta^{\alpha}}$ with parameters $\alpha = 1.25$, $\beta = 0.025$. Thus, essentially no damage is assumed at the outset. Certainly this

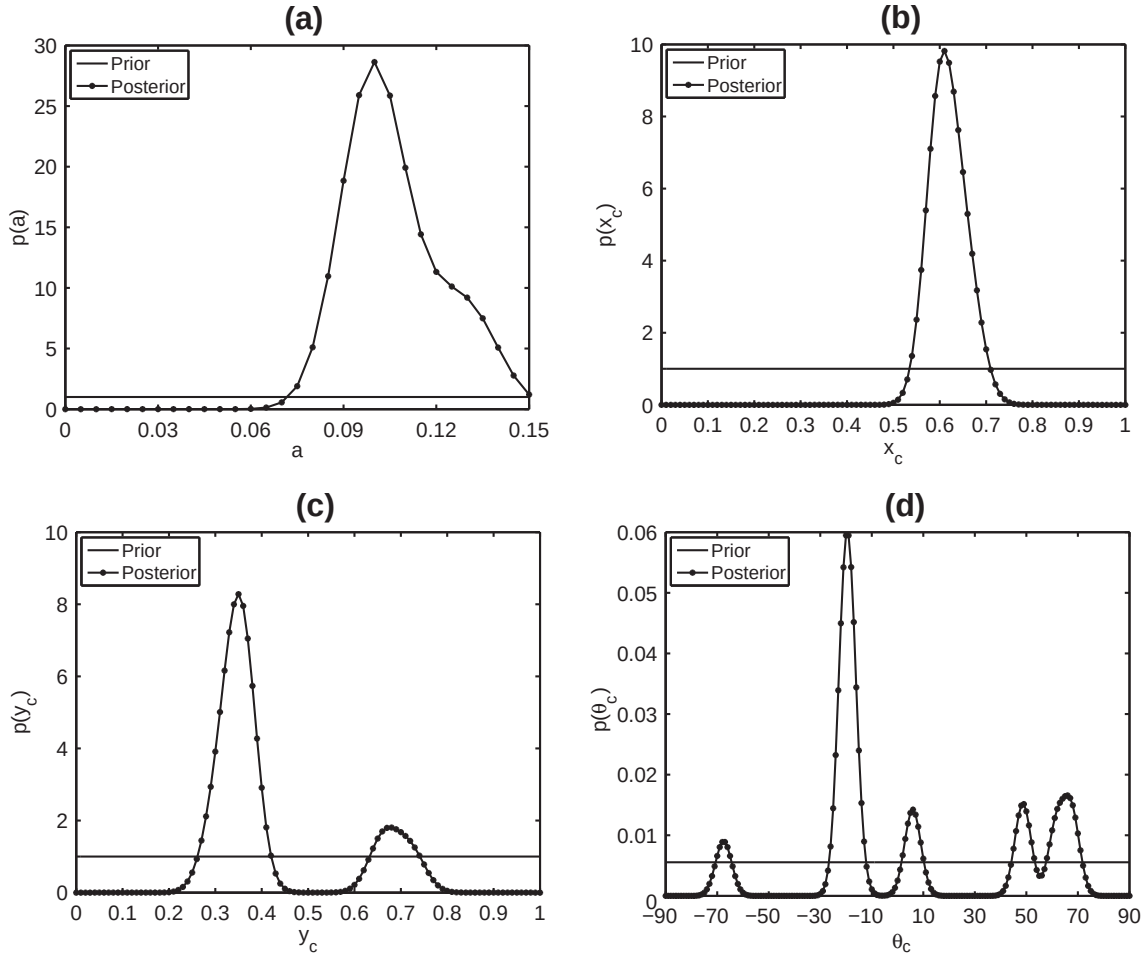

Fig. 7. Estimated posterior densities associated with (a) crack length, (b,c) crack location, and (d) crack orientation. The true parameter values are $a = 0.1$, $(x_c, y_c) = (0.6, 0.35)$, and $\theta_c = -20$.

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

distribution can be altered to reflect a known level of damage, however the more typical case is to assume the plate is healthy and allow the data (likelihood) to drive the posterior. The form of this prior (Gamma) was chosen based on the fact that the crack cannot be negative. One might also have chosen a Beta prior which has finite support on the plate, however for the prior parameters chosen there is essentially no probability of the crack extending off the end of the plate. For the crack location parameters x_c, y_c uniform priors were chosen over the span of the crack, i.e. x_c $U(0,1.25)$, y_c $U(0,1)$ reflecting the fact that no *a prior* knowledge about the crack location exists. Similarly the prior on crack angle was taken as θ_c $U(-90,90)$ as in general one will not know the crack orientation *a priori* either. The prior distributions are displayed in Fig. 6.

Initializing the Markov chains using these priors, the population-based MCMC algorithm was run for 60,000 iterations with a burn-in of 50,000 iterations. The remaining 10,000 iterations were stored and used to form the posterior densities shown in Fig. 6. In this simulation, and in others the authors have looked at, crack length is perhaps the most easily identified parameter. The Markov chains in all populations tend to the true value ($a = 0.1$) after only a few thousand iterations. Likewise, the location parameter y_c exhibits a uni-modal posterior distribution. The other parameters, x_c, θ_c are significantly more challenging to estimate as both show the clear presence of multiple maxima in the likelihood. The location parameter x_c exhibits multi-modal behavior. In particular one can see the multiple minima observed in Fig. 5. The two peaks located at $x_c = 0.3$, $x_c = 0.4$ presented a significant challenge for the regular MCMC algorithm. Because crack lengths were initially assumed small, the first minimum encountered by the Markov chain was the one located at $x_c = 0.4$ and the solution would often remain here. Running multiple chains easily overcomes this problem. Perhaps the greatest utility of the approach, however, can be seen in the identified orientation θ_c. Regardless of crack geometry, there are nearly always multiple solutions for θ_c that come close to maximizing the likelihood. The result are multiple well-defined local maxima that easily trap the standard MCMC algorithm. By contrast, Fig. 6 captures the relative heights of these maxima indicating that the highest probability for crack orientation is at the true value, $\theta_c = 30°$.

As a second example, Fig. 7 shows the identified posterior distributions for the case where the true crack parameters were set to the values $a = 0.1$, $x_c = 0.6$, $y_c = 0.35$, $\theta_c = -20$. Additionally, the Gamma prior (biased toward no crack) was changed to a Uniform prior in order to demonstrate the insensitivity of the approach to this choice. It could be that the practitioner has no *a prior* information regarding the presence and length of a crack, thus a uniform prior would be appropriate. All parameters are again correctly identified, provided that the final estimate is taken as the maximum *a posterior* value. Again, one sees multiple "good" solutions which can often trap the standard MCMC algorithm. For example, the crack location parameter y_c has a fairly high probability of being $y_c = 0.7$ despite the fact that the true value is a factor of two different. Similarly multiple peaks for the orientation parameter θ_c can be seen. The authors note that the standard MCMC algorithm was not able to sample from these complex posterior distributions.

The complexity of the posterior distributions is simply a consequence of trying to identify parameters that have little affect on the global vibrations. It is obvious from the likelihood plots (Fig. 5) that varying crack configurations can lead to nearly the same vibrational response. This is part of the physics of structural identification problems and the Bayesian approach, using the pop-MCMC algorithm, is correctly reflecting this uncertainty. Rather than producing point estimates which may or may not coincide with the true parameter value, the approach described here allows for a full accounting of the probability associated with different crack configurations. The authors view this as a strength in difficult system identification problems, particularly those in damage detection where quantifying uncertainty is critical to making informed decisions regarding the health of a structure. Choosing to use more modes in the model helps the situation, however, in practice it becomes increasingly difficult to match model to data for higher modes. In this paper the focus was on the first five modes of the response.

7. Conclusions

This paper proposes a population-based Markov Chain Monte Carlo method for solving complicated system identification problems in structural dynamics. Both the application and specific implementation of the method are new. The Pop-MCMC method is well-suited to the types of posterior parameter distributions often found in structural dynamics problems. These distributions tend to be multi-modal and can easily confuse standard MCMC approaches. The population-based approach, by contrast, uses concepts from genetic algorithm search routines to more efficiently search this parameter space. The sampling is done in such a way as to avoid becoming stuck in locally optimal solutions. The efficacy of the approach has been demonstrated in identifying crack location, length, and orientation using only simulated impulse-response data from a cracked-plate model. Because the Pop-based MCMC method requires repeated iterations of the forward model, great care was taken in developing an efficient model. This was accomplished by using tailored "serendipity" elements to describe the stress field near the crack tip. This allows for good model convergence with many fewer elements than are used in standard finite element codes. The combination of efficient modeling and effective parameter identification routines can provide a wealth of information about the state of a structure. Both the parameter estimates and the credible intervals associated with those estimates are obtained. This allows for confidence-based decisions regarding the maintenance of a structure and also provides information needed in prognostics models for damage evolution in structures.

Acknowledgments

The authors would like to acknowledge the Office of Naval Research under Contract No. N00014-09-WX-2-1002 for providing funding for this work.

References

[1] Horibe T, Watanabe K. Crack identification of plates using genetic algorithm. JSME Int J 2006;49:403–10.
[2] Stull CJ, Earls CJ, Aquino W. A posteriori initial imperfection identification in shell buckling problems. Comput Methods Appl Mech Eng 2008;198:260–8.
[3] Panigrahi SK, Chakraverty S, Mishra BK. Vibration based damage detection in a uniform strength beam using genetic algorithm. Meccanica 2009;44:697–710.
[4] Hwang S-F, Wu J-C, He R-S. Identification of effective elastic constants of composite plates based on a hybrid genetic algorithm. Compos Struct 2009;90:217–24.
[5] Fonseca JR, Friswell MI, Mottershead JE, Lees AW. Uncertainty identification by the maximum likelihood method. J Sound Vib 2005;288:587–99.
[6] Xu GY, Zhu WD, Emory BH. Experimental and numerical investigation of structural damage detection using changes in natural frequencies. J Vib Acoust 2007;129:686–700.
[7] Kiureghian AD, Dakessian T. Multiple design points in first and second-order reliability. Struct Safety 1998;20:37–49.
[8] Guo X, Bai W, Zhang W, Gao X. Confidence structural robust design and optimization under stiffness and load uncertainties. Comput Methods Appl Mech Eng 2009;198:3378–99.
[9] Beck JL, Au S-K. Bayesian updating of structural models and reliability using Markov Chain Monte Carlo simulation. J Eng Mech – ASCE 2002;128(4):380–91.

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013

[10] Cheung SH, Beck JL. Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters. J Eng Mech – ASCE 2009;135(4):243–55.
[11] Au SK, Ching J, Beck JL. Application of subset simulation methods to reliability benchmark problems. Struct Safety 2007;29:183–93.
[12] Glaser RE, Lee CL, Nitao JJ, Hickling TL, Hanley WG. Markov chain Monte Carlo-based method for flaw detection in beams. ASCE J Eng Mech 2007;133(12):1258–67.
[13] Zhang B-T, Cho D-Y. System identification using evolutionary Markov Chain Monte Carlo. J Syst Architect 2001;47:587–99.
[14] Vrugt JA, ter Braak CJF, Diks CGH. Accelerating Markov Chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int J Nonlinear Sci Numer Simul 2009;10(3):273–90.
[15] Jasra A, Stephens DA, Holmes CC. On population-based simulation for static inference. Stat Comput 2007;17:263–79.
[16] Geyer CJ. Markov chain maximum likelihood. In: Keramigas E, editor. Computing science and statistics: the 23rd symposium on the interface. Fairfax: Interface Foundation; 1991. p. 156–63.
[17] Kim W, Park J, Lee KM. Stereo matching using population-based mcmc. Int J Comput Vision 2009;83:195–209.
[18] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57(1):97–109.
[19] Nichols JM, Link WA, Murphy KD, Olson CC. A Bayesian approach to identifying structural nonlinearity using free-decay response: application to damage detection in composites. J Sound Vib 2010;329(15):2995–3007.
[20] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. J Chem Phys 1953;21(6):1087–92.
[21] Storn R, Price R. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 1997;11: 341–59.
[22] Solecki R. Bending vibration of a rectangular plate with arbitrarily located rectilinear crack. Eng Fract Mech 1985;22(4):687–95.
[23] Henshell RD, Shaw KG. Crack tip finite elements are unnecessary. Int J Numer Methods Eng 1975;9:495–509.
[24] Barsoum RS. On the use of isoparametric finite elements in linear fracture mechanics. Int J Numer Methods Eng 1976;10:25–37.
[25] Mitchell J. Elementary distributions of plane stress. Proc Lond Math Soc 1901;32:35–61.

Please cite this article in press as: Nichols JM et al. Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. Comput Struct (2011), doi:10.1016/j.compstruc.2011.03.013