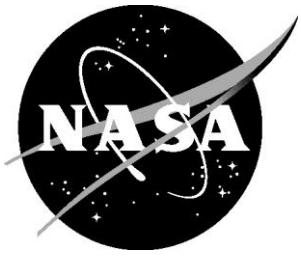
NASA/TP–2016-219169



Probabilistic Damage Characterization using a Computationally-Efficient Bayesian Approach

James E. Warner and Jacob D. Hochhalter
Langley Research Center, Hampton, Virginia

March 2016

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

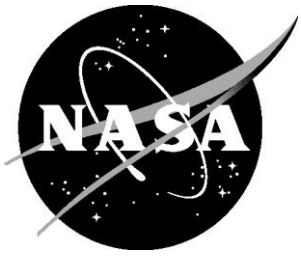
- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TP–2016-219169



Probabilistic Damage Characterization using a Computationally-Efficient Bayesian Approach

James E. Warner and Jacob D. Hochhalter
Langley Research Center, Hampton, Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

March 2016

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Abstract

This work presents a computationally-efficient approach for damage determination that quantifies uncertainty in the provided diagnosis. Given strain sensor data that are polluted with measurement errors, Bayesian inference is used to estimate the location, size, and orientation of damage. This approach uses Bayes' Theorem to combine any prior knowledge an analyst may have about the nature of the damage with information provided implicitly by the strain sensor data to form a posterior probability distribution over possible damage states. The unknown damage parameters are then estimated based on samples drawn numerically from this distribution using a Markov Chain Monte Carlo (MCMC) sampling algorithm. Several modifications are made to the traditional Bayesian inference approach to provide significant computational speedup. First, an efficient surrogate model is constructed using sparse grid interpolation to replace a costly finite element model that must otherwise be evaluated for each sample drawn with MCMC. Next, the standard Bayesian posterior distribution is modified using a weighted likelihood formulation, which is shown to improve the convergence of the sampling process. Finally, a robust MCMC algorithm, Delayed Rejection Adaptive Metropolis (DRAM), is adopted to sample the probability distribution more efficiently. Numerical examples demonstrate that the proposed framework effectively provides damage estimates with uncertainty quantification and can yield orders of magnitude speedup over standard Bayesian approaches.

1 Introduction

The ability to detect and characterize damage is an integral part of structural health management (SHM). For example, an effective SHM system could identify damage (*i.e.*, crack) in an aircraft component well before it neared a critical size in order to recommend inexpensive preventative maintenance rather than more costly repairs down the road. In general, estimating the damage state requires a framework that combines mechanical response information from both on-board sensors and a computational model of the component. Given the critical nature of SHM applications, such methods are held to stringent standards in terms of accuracy as well as computational efficiency.

The task of damage identification is further complicated by uncertainty that manifests itself both in the noisy measurement data acquired and the computational model (boundary conditions, material properties, for example) of the structural component being analyzed. Explicitly characterizing this uncertainty and quantifying its influence on the resulting damage diagnosis is vital for high fidelity prognostics and informed decision-making [1]. However, existing approaches to do so generally come at the expense of substantial computational overhead, making them infeasible for many practical applications. Thus, it remains an open challenge to provide computationally-efficient damage estimates with uncertainty quantification for SHM systems.

Traditionally, damage detection techniques have largely been deterministic in

nature and have identified structural anomalies based on changes in measured mechanical response (*e.g.* vibrations [2, 3], ultrasonic wave characteristics [4, 5], and strains [6,7]). While deterministic approaches have had much success in their ability to accurately locate damage in a computationally-efficient manner, these methods neglect the impact of uncertainty that is ubiquitous in real SHM systems due to instrument noise and the simplifications and errors in computational models. In particular, it has been shown that changes in strain can be used to accurately characterize cracks in relatively few iterations of an optimization algorithm [7], albeit while assuming the synthetic (*i.e.*, simulated) strain data used to be noise free. As an extension of this work, the study proposed herein considers noisy strain data and uses Bayesian inference to take into account the impact of this measurement uncertainty on the resulting crack estimates.

Adopting a Bayesian approach yields damage estimates in the form of probability distributions rather than point estimates, naturally enabling quantification of uncertainty in the predicted damage states. In this way, the framework developed herein uses noisy in-service strain measurements to infer the probability distributions of parameters that describe the location, size, and orientation of a crack. Bayes' Theorem [8] is used to combine any prior knowledge an analyst may have about the nature of the damage with information provided implicitly by the sensor data to formulate a posterior probability distribution over possible damage states. Since the resulting probability distribution is not analytically tractable in general, the crack parameters are then estimated based on samples drawn numerically from the distribution using a Markov chain Monte Carlo (MCMC) [9] sampling algorithm.

There has recently been an increased effort in the problem of uncertainty quantification in damage identification using Bayesian inference. Several studies [10–12] have used noisy vibrational data to detect structural damage, while in one such study [13], the emphasis was on the development of an efficient MCMC algorithm for sampling the resulting solution distribution. Structural crack identification using strain data has also been previously studied within probabilistic frameworks. In one such study [14], the focus was primarily on the problem of optimal sensor placement where only point estimates rather than probability distributions of crack parameters were sought. In a later study [15], a combination of the extended finite element method and MCMC sampling was used to obtain crack parameter probability distributions, first demonstrating the feasibility of quantifying uncertainty in crack estimates from noisy strain data. Compared to deterministic methods, Bayesian approaches have the major advantage of quantifying uncertainty in the estimates provided, but also incur a substantial computational penalty. Here, the computational expense results from the MCMC sampling procedures used, which typically exhibit slow convergence (*i.e.*, require many samples) and involve the evaluation of a potentially intensive computational model (*e.g.* finite element method simulation) for each individual sample drawn.

The Bayesian damage diagnosis strategy presented here is designed to provide significant computational speedup over previous Bayesian approaches, yielding a framework that is better suited for SHM applications. An efficient surrogate model is constructed using sparse grid interpolation [16, 17] to replace a more computa-

tionally expensive finite element simulation that would otherwise be evaluated for each MCMC sample. The overall convergence rate of the sampling process is then improved by modifying the standard Bayesian posterior probability distribution using a weighted likelihood formulation [18, 19]. Additionally, the Delayed Rejection Adaptive Metropolis (DRAM) algorithm [20] is adopted for MCMC sampling, allowing robust and efficient exploration of the posterior distribution with less required user tuning. These improvements yield a Bayesian framework for damage identification that is as computationally-efficient as many existing deterministic approaches while still providing quantitative measures of confidence and uncertainty with its estimates. Numerical examples demonstrate the ability of the method to accurately identify the distributions of damage parameters using noisy strain data and show the potential to provide orders of magnitude computational speedup with respect to traditional Bayesian approaches.

This article is organized as follows. Section 2 provides the background on the general approach for Bayesian damage identification, first detailing the formulation of the probability distribution over the unknown damage parameters and then the MCMC sampling procedure used to explore the distribution is described. Section 3 then details the formulation of the computationally-efficient Bayesian approach proposed here, including the formation of a surrogate model using sparse grid interpolation, the modification of the posterior distribution using a weighted likelihood formulation, and the use of the DRAM MCMC algorithm. Next, results from two numerical examples that demonstrate the effectiveness and efficiency of the proposed framework are presented in Section 4. Finally, the paper is concluded in Section 5 with a summary of the approach and results.

2 Background: The Standard Bayesian Approach

The goal of damage determination from a Bayesian inference standpoint is to obtain a probability distribution for the unknown damage parameters given measured strain data from a limited number of noisy sensors. The information offered implicitly from these strain data is captured in a *likelihood* function that is combined with a *prior* distribution, representing any prior knowledge of the damage state, to produce a *posterior* probability distribution over damage states. In all but the simplest formulations, integration of the resulting probability distribution is not feasible and so MCMC algorithms are used to sample the distribution and estimate the unknown damage parameter statistics. The formulation of the posterior probability distribution and the subsequent MCMC sampling procedure is now given in detail. The proposed computational improvements to the standard Bayesian approach are then detailed in the subsequent section.

2.1 Formulation of the Posterior Probability Distribution

To begin, the damage considered in this study is represented by a simple crack in a two-dimensional plate that is parameterized with four independent parameters

$$\mathcal{C} = [x, y, a, \theta] \tag{1}$$

where (x, y) denote the coordinates of the crack center, a is the crack length, and θ is the crack orientation (*i.e.*, its angle with respect to the x-axis). A schematic of this setup is shown in Figure 1. It is assumed that strain data are obtained at M locations throughout the plate, yielding a data set

$$\mathcal{D} = \{\hat{\mathbf{S}}_i\}_{i=1}^{M} \tag{2}$$

from which the crack parameters in Equation 1 are to be estimated. Here, $\hat{\mathbf{S}}_i = [\epsilon_i^{xx}, \epsilon_i^{yy}, \gamma_i^{xy}]^T$ is the vector of strain components from sensor i. Furthermore, a computational model, $\mathcal{M}$, of the plate being analyzed is required that can return the simulated strains $\mathbf{S}_i$ at sensor i for a given estimate of the crack configuration $\mathcal{C}$:

$$\mathbf{S}_i = \mathcal{M}(\mathcal{C}, i) \tag{3}$$

The formulation to follow is done largely in terms of the general damage parameter array $\mathcal{C}$, response dataset $\mathcal{D}$, and computational model $\mathcal{M}$. Thus, it is important to note that the proposed framework allows the possibility of including more complex damage parameterizations as well as different types of mechanical response data, provided there is a corresponding computational model to estimate the measured quantity.

With these preliminaries in place, the formulation of a conditional probability distribution of the crack parameters (Equation 1) given the measured strain data (Equation 2), $p(\mathcal{C}|\mathcal{D})$, can now be described. The foundation of the approach is Bayes' Theorem [8], which provides a simple but powerful relation between conditional probability distributions as follows

$$p(\mathcal{C}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathcal{C})p(\mathcal{C})}{p(\mathcal{D})} \propto p(\mathcal{D}|\mathcal{C})p(\mathcal{C}) \tag{4}$$

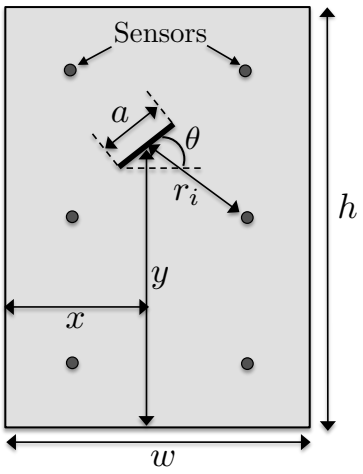


Figure 1. Schematic of the plate geometry and parameterization of the structural crack.

where the normalization constant $p(\mathcal{D})$ is typically neglected for convenience as it is not required by the MCMC sampling procedure to follow. Thus, the posterior distribution $p(\mathcal{C}|\mathcal{D})$ over the unknown crack parameters is obtained by combining information from two sources: 1) a likelihood function, $p(\mathcal{D}|\mathcal{C})$, describing the probability of measuring the strain data for a given crack configuration, and 2) a prior distribution, $p(\mathcal{C})$, reflecting any *a priori* information one might have about the crack parameters before obtaining data.

With the posterior distribution expression provided by Bayes' Theorem in Equation 4, the specification of the likelihood $p(\mathcal{D}|\mathcal{C})$ and prior distribution $p(\mathcal{C})$ will complete the Bayesian damage identification formulation. Beginning with the latter, this study assumes a non-informative prior distribution representing a situation where the analyst has no credible knowledge *a priori* about the nature of possible damage. In this way, a uniform distribution over the crack parameters is specified that simply enforces any practical bounds on the unknowns (*e.g.* the crack lies within the domain regardless of orientation, the crack length is positive):

$$p(\mathcal{C}) \propto \begin{cases} 1, & \text{if } \left(\frac{a}{2} \leq x \leq (w - \frac{a}{2})\right) \cap \left(\frac{a}{2} \leq y \leq (h - \frac{a}{2})\right) \cap (\ 0 \leq a) \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

While the choice of a uniform prior distribution is motivated by the desire to avoid any subjectivity in the demonstration of the proposed framework, it is noted that the prior distribution can be a powerful mechanism for incorporating an analyst's insight in practical SHM applications. For example, one could assume that a crack is more probable near a stress concentration or that cracks are generally more likely to be oriented perpendicular to an applied load.

The likelihood function is now formulated as the joint probability of measuring the strain values in $\mathcal{D}$ given a particular crack configuration, $\mathcal{C}$:

$$p(\mathcal{D}|\mathcal{C}) = p(\hat{\mathbf{S}}_1, ..., \hat{\mathbf{S}}_M|\mathcal{C}) \tag{6}$$

To obtain an explicit expression for Equation 6, the common and simplifying assumption is first made that the strain measurements are polluted with random noise, $\boldsymbol{\eta}$, that is independent and identically distributed (i.i.d) according to a zero-mean Gaussian distribution [12–14]

$$\hat{\mathbf{S}}_i = \mathbf{S}_i + \boldsymbol{\eta}_i, \quad \boldsymbol{\eta}_i \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \tag{7}$$

where $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is a multivariate Gaussian (Normal) distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\mathbf{I} \in \mathbb{R}^3$ is the identity matrix, and σ is the standard deviation of the measurement error distribution, interpreted here as the noise level. Using Equations 3 and 7, the probability of a single strain measurement at sensor i for a given crack configuration, $\mathcal{C}$, can therefore be written as

$$p(\hat{\mathbf{S}}_i|\mathcal{C}) = N(\mathcal{M}(\mathcal{C}, i), \sigma^2\mathbf{I}) \tag{8}$$

Now, according to the assumed independence of $\boldsymbol{\eta}_i$, the likelihood function (Equation

6) can be obtained by factoring Equation 8 over the M measurements

$$p(\mathcal{D}|\mathcal{C}) = \prod_{i=1}^{M} p(\hat{\mathbf{S}}_i|\mathcal{C})$$
$$= \frac{1}{(2\pi\sigma^2)^{M/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{M} \|\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)\|^2\right) \tag{9}$$

where

$$\sum_{i=1}^{M} \|\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)\|^2 = \sum_{i=1}^{M} [\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)]^T \cdot [\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)] \tag{10}$$

is the sum of squares error between measured and computed strains for a given guess of crack parameters. It is clear from Equation 9 that a particular crack configuration $\mathcal{C}$ that produces large errors between the computational model and measurement strains will receive lower probability and vice versa, as desired.

With the prior distribution and likelihood function specified, the posterior distribution over the crack configurations is obtained by simply combining Equations 4, 5, and 9 as

$$p(\mathcal{C}|\mathcal{D}) \propto \begin{cases} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{M} \|\hat{\mathbf{S}}_i - \mathcal{M}_i(\mathcal{C}))\|^2\right), & \text{if } (L \leq \mathcal{C} \leq U) \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

where the normalization constant in the likelihood has been omitted since it is independent of $\mathcal{C}$ and the notation for the upper and lower bounds in the prior distribution have been simplified for brevity. It is noted that an additional strength of the Bayesian approach used here is the ability to treat the noise level, σ^2, as an unknown and to infer it from the data in the common situation where it is difficult to estimate *a priori*. This is done by rewriting the posterior distribution (Equation 11) as the joint probability of the unknown crack configuration and noise level given the strain data using the product probability rule

$$p(\mathcal{C}, \sigma^2|\mathcal{D}) \propto p(\mathcal{D}|\mathcal{C}, \sigma^2)p(\sigma^2) \tag{12}$$

Here, a prior probability on σ^2 must be prescribed to complete the formulation. A common choice, which is adopted here, is the inverse-gamma distribution

$$p(\sigma^2) \propto (\sigma^2)^{(-\alpha-1)} \exp(-\frac{\beta}{\sigma^2}) \tag{13}$$

where α and β are the shape and scale parameters, respectively. The inverse-gamma distribution represents an uninformative prior distribution with a suitable selection of α and β, limiting any subjectivity by the analyst, and also benefits from algorithmic convenience when employing MCMC sampling methods [11]. To simplify the remaining formulation, however, the form of the posterior distribution given in Equation 11 will be assumed without loss of generality.

2.2 Sampling with Markov Chain Monte Carlo

While an explicit expression has been formulated for the posterior probability distribution of the crack configuration in Equation 11, the calculation of an analytical solution for the unknown parameters remains intractable due to the dependence on the computational model $\mathcal{M}$. The standard approach is thus to resort to sampling algorithms to numerically explore the complex probability distribution [8]. After a collection of samples $\{\mathcal{C}^{(j)}\}$ has been drawn from $p(\mathcal{C}|\mathcal{D})$, both point and interval estimates of the unknowns can be calculated. For example, the expected value of the crack configuration can be approximated by the sample mean

$$\mathbb{E}[\mathcal{C}] = \int_L^U p(\mathcal{C}|\mathcal{D})\mathcal{C}d\mathcal{C} \approx \frac{1}{N}\sum_{j=1}^{N}\mathcal{C}^{(j)} \tag{14}$$

where N is the total number of samples drawn. Furthermore, credible intervals can be established based on the empirical distributions of the samples, providing a measure of confidence in the crack parameter estimates.

MCMC is a powerful and general framework for sampling high dimensional and complex probability distributions [9] such as Equation 11. Rather than directly drawing independent samples from the target distribution, a dependent sequence of samples is generated based on a fixed number of previous samples (*i.e.*, a Markov chain) and an easily-sampled proposal distribution. Under mild assumptions about the choice of this proposal distribution and the adoption of a simple acceptance criteria for the proposal samples drawn, it can be shown that MCMC produces samples that follow the target probability distribution. In this section, however, the discussion is limited to the practical aspects of MCMC sampling, while the references cited can be consulted for coverage of the theoretical justifications of the framework.

A popular and straightforward implementation of MCMC known as the Metropolis algorithm is now described. Algorithm (1) displays the pseudocode for drawing N samples of the crack configuration $\{\mathcal{C}^{(j)}\}_{j=1}^{N}$ from the posterior distribution $p(\mathcal{C}|\mathcal{D})$ using the Metropolis algorithm. For each iteration, the approach simply draws a trial sample $\mathcal{C}^*$ from the proposal distribution $q(\mathcal{C}^*|\mathcal{C}^{(j-1)})$, and then decides whether to accept or reject this sample based on the *acceptance probability*, $A(\mathcal{C}^*, \mathcal{C}^{(j-1)})$.

The Metropolis algorithm assumes that the proposal distribution is symmetric and thus one appropriate and common choice is a Gaussian distribution centered at the previous sample

$$q(\mathcal{C}^*|\mathcal{C}^{(j-1)}) = N(\mathcal{C}^{(j-1)}, \Sigma_q^2) \tag{15}$$

where Σ_q^2 is the user-specified covariance matrix. Here, drawing a candidate sample from q can be equivalently viewed as adding a random perturbation to the current crack configuration parameters, with perturbation size and shape being prescribed by Σ_q. With regard to the acceptance probability A, it can be seen that a trial sample $\mathcal{C}^*$ with higher probability than the previous sample $\mathcal{C}^{(j-1)}$ according to $p(\mathcal{C}|\mathcal{D})$ will always be accepted (with probability $A = 1$). On the other hand, if the trial sample is less probable according to $p(\mathcal{C}|\mathcal{D})$, it can still be accepted, but with

Algorithm 1 The Metropolis MCMC Algorithm

Initialize $\mathcal{C}^{(0)}$
for $j = 1 : N$ **do**
  Sample $u \sim \text{Uniform}(0, 1)$
  Sample $\mathcal{C}^* \sim q(\mathcal{C}^*|\mathcal{C}^{(j-1)})$
  if $u < A(\mathcal{C}^*, \mathcal{C}^{(j-1)}) = \min\{1, \frac{p(\mathcal{C}^*|\mathcal{D})}{p(\mathcal{C}^{(j-1)}|\mathcal{D})}\}$ **then**
    $\mathcal{C}^{(j)} = \mathcal{C}^*$
  else
    $\mathcal{C}^{(j)} = \mathcal{C}^{(j-1)}$
  end if
end for

decreasing probability according to the ratio $\frac{p(\mathcal{C}^*|\mathcal{D})}{p(\mathcal{C}^{(j-1)}|\mathcal{D})}$. Since the target probability distribution only appears in the Metropolis algorithm through this ratio, it need be only specified up to a normalization constant as in Equation 4. It is also noted that the computational model $\mathcal{M}$ must be evaluated for each trial sample $\mathcal{C}^*$ in order to compute its probability according to $p(\mathcal{C}|\mathcal{D})$.

While the Metropolis algorithm and other MCMC approaches can be expressed in just a few lines of pseudocode and have straightforward implementations, tuning the algorithms for good performance is notoriously difficult [20]. This is primarily due to the selection of an appropriate proposal distribution through the specification of Σ_q in Equation 15. Here, Σ_q effectively controls the expected range of the trial samples in the parameter space or, as mentioned, the size and shape of the perturbation added to the previous sample to generate the trial sample. If the range selected is too small, a thorough exploration of the posterior probability distribution can take a huge number of samples and the algorithm is more susceptible to being trapped in local modes of the distribution. Increasing the range too far, however, also increases the rejection rate of the trial samples and can significantly decrease the rate of convergence.

3 Formulation: Improvements for Computational Efficiency

The Bayesian approach to damage identification with MCMC sampling described in the previous section provides an effective way to quantify how the uncertainty in the noisy strain measurement data propagates to estimates of the damage state (*i.e.*, crack parameters). However, the well known drawback of such approaches is the substantial computational overhead incurred by the sampling process. Three primary factors that increase solution time for MCMC-based Bayesian approaches are: 1) the evaluation of the potentially intensive computational model $\mathcal{M}$ for each sample drawn from the posterior distribution $p(\mathcal{C}|\mathcal{D})$, 2) the complex, multimodal nature of the resulting posterior distribution itself, and 3) the inefficiency and limited robustness of standard MCMC sampling algorithms.

This work seeks to substantially improve the efficiency of the standard Bayesian approach by addressing each of these factors with the following: 1) an efficient surrogate model using sparse grid interpolation to replace the costly computational model, 2) a reformulation of the posterior distribution using a weighted likelihood approach that improves sampling convergence, and 3) the adoption of the Delayed Rejection Adaptive Metropolis (DRAM) algorithm to improve the robustness of the sampling process itself. The remainder of this section details each of these three computational enhancements.

3.1 Surrogate Modeling with Sparse Grid Interpolation

It is clear from Algorithm (1) and the expression for the posterior distribution in Equation 11 that each iteration of the Metropolis algorithm requires the evaluation of $p(\mathcal{C}|\mathcal{D})$ and thus the solution of the computational model $\mathcal{M}$. For expensive models such as three-dimensional (3D) finite element analysis codes, obtaining the computed strains throughout the sampling process can have a prohibitively long computation time, especially for SHM applications. Hence, to improve the efficiency of the Bayesian damage detection approach, a surrogate model is substituted for the computational model that can be evaluated in a fraction of the time.

In this work, sparse grid interpolation [17] is employed for surrogate modeling. This method generates an approximation to a given multivariate function by interpolating between stored values of the function on a predefined grid of points. Sparse grid interpolation relies on Smolyak's algorithm [16] for selecting an optimal, reduced set of points and appropriate basis functions for representing the underlying function. This results in a method that scales better with increasing dimension, *i.e.*, the number of required grid points can be reduced by orders of magnitude with respect to full grid interpolation while retaining a similar degree of accuracy.

In the context of the Bayesian damage detection framework presented here, the goal is to construct sparse grid interpolants that can efficiently approximate the strains at each sensor location for a given crack configuration. For example, the strain vector at sensor i is obtained through the evaluation of an interpolant $\mathcal{I}_i$ rather than the original computational model:

$$\mathbf{S}_i = \mathcal{M}(\mathcal{C}, i) \approx \mathcal{I}_i(\mathcal{C}) \tag{16}$$

Here, each component of strain is viewed as a function of crack configuration $\mathcal{C}$ for which an interpolant of the following form can be constructed

$$\epsilon_i^{xx}(\mathcal{C}) \approx \mathcal{I}_i^{xx}(\mathcal{C}) = \sum_{k=1}^{n} a_k(\mathcal{C})\epsilon_i^{xx}(\bar{\mathcal{C}}^{(k)}) \tag{17}$$

where $\{\bar{\mathcal{C}}^{(1)}, ..., \bar{\mathcal{C}}^{(n)}\}$ represent a predefined sparse grid of n points in the four dimensional crack configuration space and $a_k(\mathcal{C})$ are appropriate basis functions satisfying $a_k(\bar{\mathcal{C}}^{(k)}) = 1$ and $a_k(\bar{\mathcal{C}}^{(j)}) = 0 \ \forall k, \ k \neq j$. The same logic applies to the strain components ϵ^{yy} and γ^{xy} at each sensor.

To construct the interpolant Equation 17, a sparse grid of crack configurations $\{\bar{\mathcal{C}}^{(1)}, ..., \bar{\mathcal{C}}^{(n)}\}$ is first defined based on the level of refinement specified and the upper

and lower bounds on possible crack parameters. Next, for each of the n cracks in the sparse grid, the original computational model is evaluated and the strains at each of the sensor locations are stored. Then, during the MCMC sampling process, an approximate posterior distribution depending on the sparse grid interpolation surrogate model is evaluated at each iteration

$$p(\mathcal{C}|\mathcal{D}) \approx \tilde{p}(\mathcal{C}|\mathcal{D}) \propto \begin{cases} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{M}\|\hat{\mathbf{S}}_i - \mathcal{I}_i(\mathcal{C})\|^2\right), & \text{if } (L \leq \mathcal{C} \leq U) \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

Although the surrogate model $\mathcal{I}$ for computed strains can require many simulations using $\mathcal{M}$ to generate, it is important to note that this is considered an *offline* cost. Specifically, for a given structural component with specified sensor locations, a series of n expensive computational model evaluations is performed just once initially so that an arbitrary number of subsequent analyses can be carried out in a fraction of the time for SHM applications. Furthermore, the n simulations required are independent of one another and can thus be executed simultaneously across multiple processors.

3.2 Weighted Likelihood Formulation

Another significant challenge of Bayesian damage detection approaches is that the posterior distribution is often an extremely complicated function with many modes (*i.e.*, maxima). This complexity is a reflection of the likelihood function that describes the observed data and the fact that several damage states can produce comparable structural responses [13]. Each plausible damage state then corresponds to a mode of the likelihood and posterior distributions. As a result, MCMC sampling algorithms are susceptible to becoming "trapped" in these local modes while exploring the posterior distribution, substantially increasing the number of samples needed for accurate damage estimates. While Nichols and coworkers [13] focused on developing a MCMC method that could better explore a complicated posterior distribution, the framework presented here instead reformulates the posterior distribution to make it more amenable to sampling.

The reformulation of the posterior distribution done here can be seen as a transformation of the deterministic damage detection approach [7] to a Bayesian context using a mechanism known as weighted likelihood [18, 19]. To illustrate this, it is first noted that typical deterministic approaches seek a point estimate of damage by maximizing the likelihood function (Equation 9), or equivalently, finding the crack configuration that minimizes the errors between computed and measured strains (Equation 10). However, it was reported [7] that minimizing a weighted objective function of the form

$$\sum_{i=1}^{M} r_i^2 \|\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)\|^2, \tag{19}$$

where r_i is the distance from the current crack sample $\mathcal{C}$ and sensor i (Figure 1), yielded faster convergence rates.

In the Bayesian context, the weighted likelihood method has been used in regression problems to provide a formal mechanism to trade bias for precision by allowing observed data to have varying weight [18]. A weighted likelihood formulation is adopted in this work in an effort to accelerate the convergence of MCMC sampling algorithms for Bayesian damage detection in a manner analogous to the use of Equation 19 [7]. To this end, it can be seen that minimizing Equation 19 is equivalent to maximizing a *weighted likelihood* of the form

$$p_w(\mathcal{C}|\mathcal{D}) = \prod_{i=1}^{M} p(\hat{\mathbf{S}}_i|\mathcal{C})^{r_i^2} \tag{20}$$

where the standard likelihood in Equation 9 has been modified to weight the strain data at each sensor i by r_i. Therefore, the posterior distribution using a weighted likelihood formulation is given by

$$p_w(\mathcal{C}|\mathcal{D}) \propto \begin{cases} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{M} r_i^2 \|\hat{\mathbf{S}}_i - \mathcal{M}(\mathcal{C}, i)\|^2\right), & \text{if } (L \leq \mathcal{C} \leq U) \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

and the final form of the posterior distribution used in this work with both the surrogate model $\mathcal{I}$ and the weighted likelihood function is

$$\tilde{p}_w(\mathcal{C}|\mathcal{D}) \propto \begin{cases} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{M} r_i^2 \|\hat{\mathbf{S}}_i - \mathcal{I}_i(\mathcal{C})\|^2\right), & \text{if } (L \leq \mathcal{C} \leq U) \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

It is demonstrated in the numerical examples that replacing the posterior distribution in Equation 11 with Equation 21 or Equation 22 results in probability distributions for $\mathcal{C}$ that are far less complex and multimodal, providing substantial improvement in the convergence of MCMC algorithms.

3.3 Efficient Sampling with the DRAM Algorithm

As mentioned in Section 2.2, it can be very challenging to appropriately tune MCMC methods such as the Metropolis algorithm to get good sampling performance with Bayesian approaches. A user must carefully specify a suitable proposal distribution (Equation 15) that strikes a balance between an efficient exploration of the posterior probability distribution and an acceptable rejection rate of proposed samples. Failure to do so can drastically increase the number of samples required for accurate crack parameter estimates and therefore the overall solution time in a Bayesian damage detection framework. In this work, a self-tuning, efficient version of the original Metropolis algorithm known as the Delayed Rejection Adaptive Metropolis (DRAM) algorithm [20] is adopted to sample the posterior distribution more effectively.

The DRAM algorithm combines two powerful and complementary MCMC ideas: 1) adaptive Metropolis sampling and 2) delayed rejection. The adaptive Metropolis concept allows for automatic, on-line tuning of the proposal distribution by explicitly updating Σ_q^2 in Equation 15 using the empirical covariance of the previously

drawn samples. This adaptation results in trial samples that are more likely to follow the target posterior distribution and diminishes the impact of a poorly chosen initial proposal distribution. Delayed rejection modifies the standard Metropolis algorithm to allow more than one stage of proposal distributions so that there are multiple opportunities to accept a trial sample (*i.e.*, delay its rejection). For example, the variance or range of the proposal distributions can be decreased at higher stages to ensure that some samples will be accepted, ultimately decreasing the rejection rate. Since the adaptation component of the algorithm depends on accepted samples to update the proposal covariance, the delayed rejection idea complements this process by increasing the probability of accepting a sufficient number of trial samples. Furthermore, it can be proven that with these modifications, the DRAM algorithm still converges to the desired probability distribution. More details can be found in [20].

4 Numerical Examples

The effectiveness of the proposed Bayesian damage detection framework is now demonstrated in two numerical examples of strain-based crack identification. The approach was implemented in MATLAB [21], using the `spinterp` [17] software package to implement the sparse grid interpolation surrogate model and the DRAM [20] MATLAB implementation for MCMC sampling. Abaqus [22] was used to generate the computed strains using the finite element method to construct the surrogate model with `spinterp`.

In both examples, strain data were generated synthetically by storing strain values generated with Abaqus at select sensor locations for the reference crack configuration and adding random Gaussian noise to mimic measurement errors. In Example 1, only the crack location is considered unknown. Here, an in-depth comparison of the efficiency of the proposed method is made against a standard Bayesian approach, demonstrating the significant computational speedup provided by both surrogate modeling and the weighted likelihood formulation. In Example 2, a crack of unknown location, size, and orientation is identified using strain data with varying noise levels. Here, the ability of the framework to accurately and efficiently estimate the four crack parameters is illustrated, while also quantifying the increasing uncertainty introduced by measurement errors.

4.1 Example 1: Crack Location Identification

In this example, noisy simulated strain data are used to estimate the location of a crack with known size and orientation using the proposed Bayesian framework. A schematic of the reference configuration for this problem can be seen in Figure 2(a). Dimensionless quantities are assumed in this work for simplicity, where a two-dimensional rectangular domain with width $(w) = 6$ and height $(h) = 12$ and crack location $(x, y) = (4, 9)$ is considered. Fifteen virtual sensor locations, represented by the white dots in Figure 2(a), are placed uniformly throughout the domain. The strains due to a uniform applied displacement of 0.5 are computed at these locations

using Abaqus and 5% random Gaussian noise is then added to produce synthetic measurement data for the problem. The finite element mesh used throughout this work had approximately 10^4 degrees of freedom. The crack size and orientation are fixed at $a = 1.0$ and $\theta = 0^\circ$, respectively, and thus the goal is to estimate $p(\mathcal{C}|\mathcal{D})$ where $\mathcal{C} = [x, y]$ and $\mathcal{D} = \{\hat{\mathbf{S}}_i\}_{i=1}^{15}$. This simplified setting with just two unknowns allows for clearer visualization of the framework and comparison with a standard Bayesian approach, while the treatment of general crack identification with unknown size and orientation is considered subsequently in Example 2.

To execute the proposed Bayesian framework, the surrogate model for computed strains using sparse grid interpolation was first constructed using the `spinterp` MATLAB package. Interpolants were constructed for different levels of refinement with increasing numbers of nodes (n) in the sparse grid. The accuracy of the surrogate model $\mathcal{I}$ with respect to the finite element model $\mathcal{M}$ for each refinement level was evaluated using the average percent error in computed strains

$$e(\mathcal{C}) = \frac{1}{M} \sum_{i=1}^{M} \frac{|\mathcal{I}_i(\mathcal{C}) - \mathcal{M}(\mathcal{C}, i)|}{|\mathcal{M}(\mathcal{C}, i)|} \times 100 \tag{23}$$

for particular crack parameters $\mathcal{C} = [x, y]$. This metric was evaluated for ten randomly drawn crack configurations, where the average result is plotted as a function of grid refinement in Figure 3. Here, a refinement level of 8 with 1,537 grid points produced a surrogate model that was deemed sufficiently accurate for the subsequent analysis, as it had an average error (4.38%) that was within the noise level (5%) for the problem.

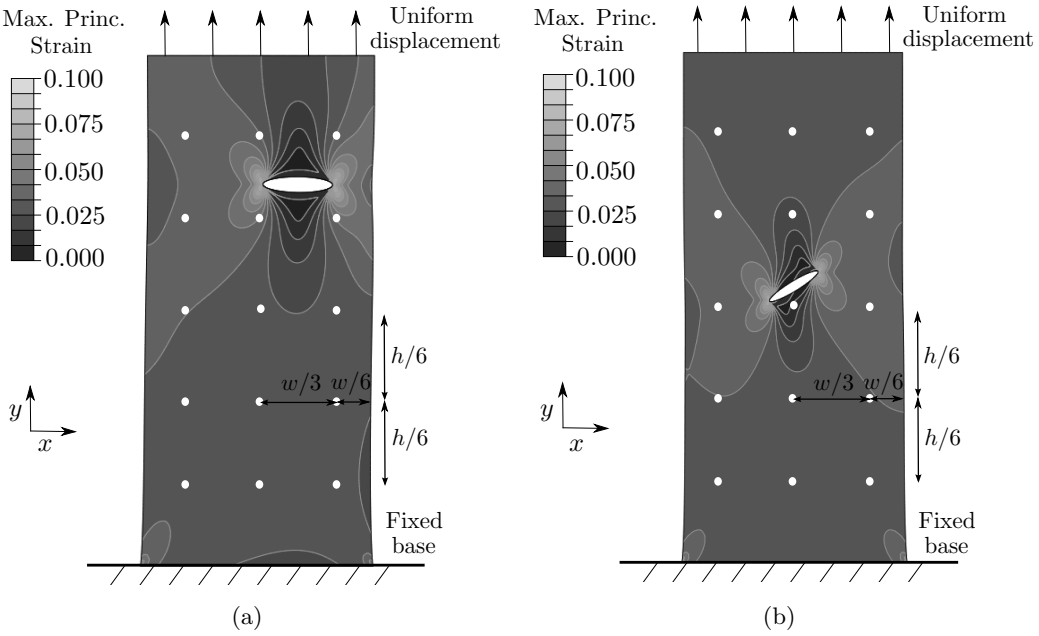


Figure 2. Reference configurations for (a) Example 1 and (b) Example 2, where $w = 6$ and $h = 12$. The white dots indicate assumed strain sensor locations.

For illustration purposes, the posterior probability distribution of the unknown crack location $p(x, y|\mathcal{D})$ is now visualized. The crack location probability was evaluated on a uniform 50×100 grid across the problem domain and is displayed in Figure 4, comparing distributions for (a) the standard Bayesian formulation (Equation 11), (b) the weighted likelihood formulation (Equation 21), and (c) the weighted likelihood formulation with the sparse grid interpolation surrogate model (Equation 22). The complexity of the posterior distribution for a standard Bayesian approach is evident, where the multimodal nature of the contours seems to indicate the plausibility of a crack residing in a number of areas in addition to the true location. Figure 4(b) represents the substantial simplification of the posterior distribution when using the weighted likelihood formulation, displaying just one primary mode around the true crack location with negligible probabilities far from that area. Figure 4(c) verifies the effectiveness of the constructed surrogate model, where there is good agreement between this approximate posterior distribution using sparse grid interpolation and the "true" posterior in 4(b) using the finite element model. It is noted that while a "brute force" evaluation of the posterior distribution on a full grid in two dimensions is plausible here for demonstration, this is not a viable solution method in general for problems with higher dimensional unknowns as the number of computations grows exponentially.

Next, the DRAM algorithm was employed to sample the proposed posterior distribution described by Equation 22 and displayed in Figure 4(c), starting from an initial guess of $(x_0, y_0) = (2, 2)$. The resulting sample-based estimates of the unknown crack coordinates using 2,000 samples are shown in Figure 5. Figures 5(a) and 5(b) show the estimated probability density functions of the x and y coordinate

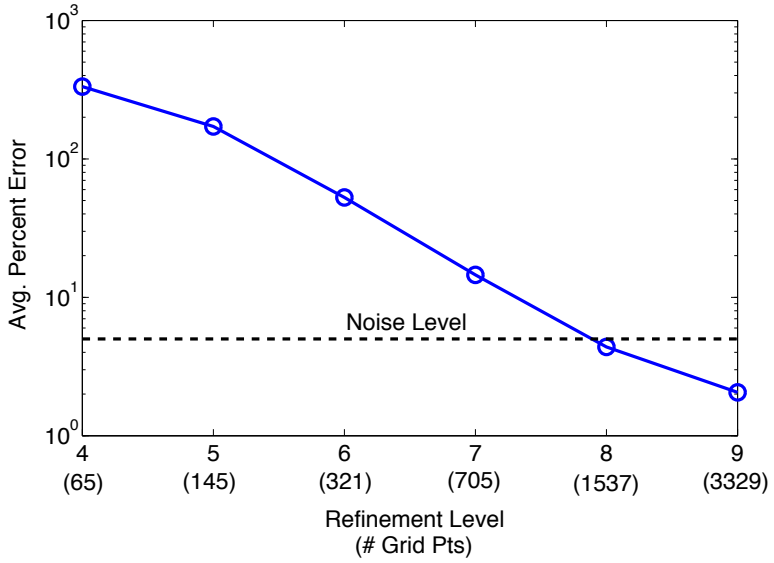


Figure 3. Average surrogate model error for ten randomly drawn crack configurations (Equation 23) versus refinement level of the sparse grid used in Example 1.

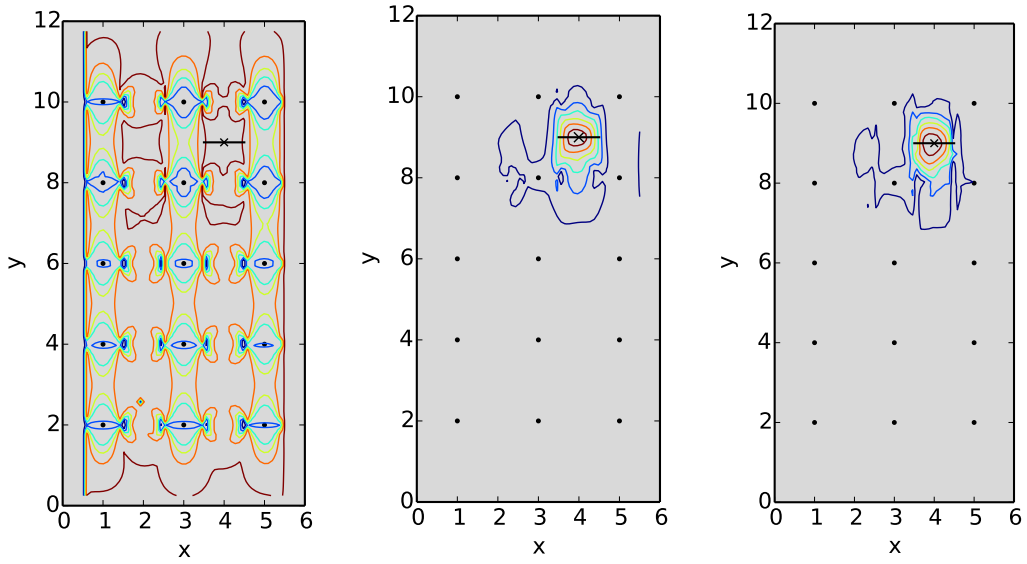


Figure 4. Posterior probability distributions for the crack location in Example 1 using three different expressions: a) Standard Bayesian formulation (Equation 11), b) weighted likelihood formulation (Equation 21), and c) surrogate model-approximated weighted likelihood formulation (Equation 22). Virtual strain sensor locations are illustrated as block dots.

of the crack, respectively, while Figure 5(c) displays the mean estimate (circle) and true location ("x") of the crack along with 50% and 95% credible intervals. It is observed that the true crack location is accurately predicted by the mean estimate $(3.95, 8.86)$ while the Bayesian approach also quantifies the effect of the 5% measurement noise uncertainty through the estimated distributions. It is noted that fast convergence was observed in the sampling process and that the change in statistics beyond 2,000 samples was negligible.

4.1.1 Computational Efficiency

A comparison of the proposed framework to a standard Bayesian damage detection approach was also made to demonstrate the improvement in computational efficiency provided. As the solution time in a MCMC-based Bayesian approach is governed by the sampling process, the focus was on how replacing a finite element model with a surrogate model decreases the computation time to draw a single sample from the posterior distribution and how the weighted likelihood formulation reduces the total number of samples required for an accurate solution. Given the simple nature of the two-dimensional geometry assumed here, the computation time for the finite element method simulation in Abaqus was approximately 30 seconds. However, evaluating the surrogate model was substantially faster, taking around 0.1 seconds in this example to compute Equation 17 for all strain sensor locations.

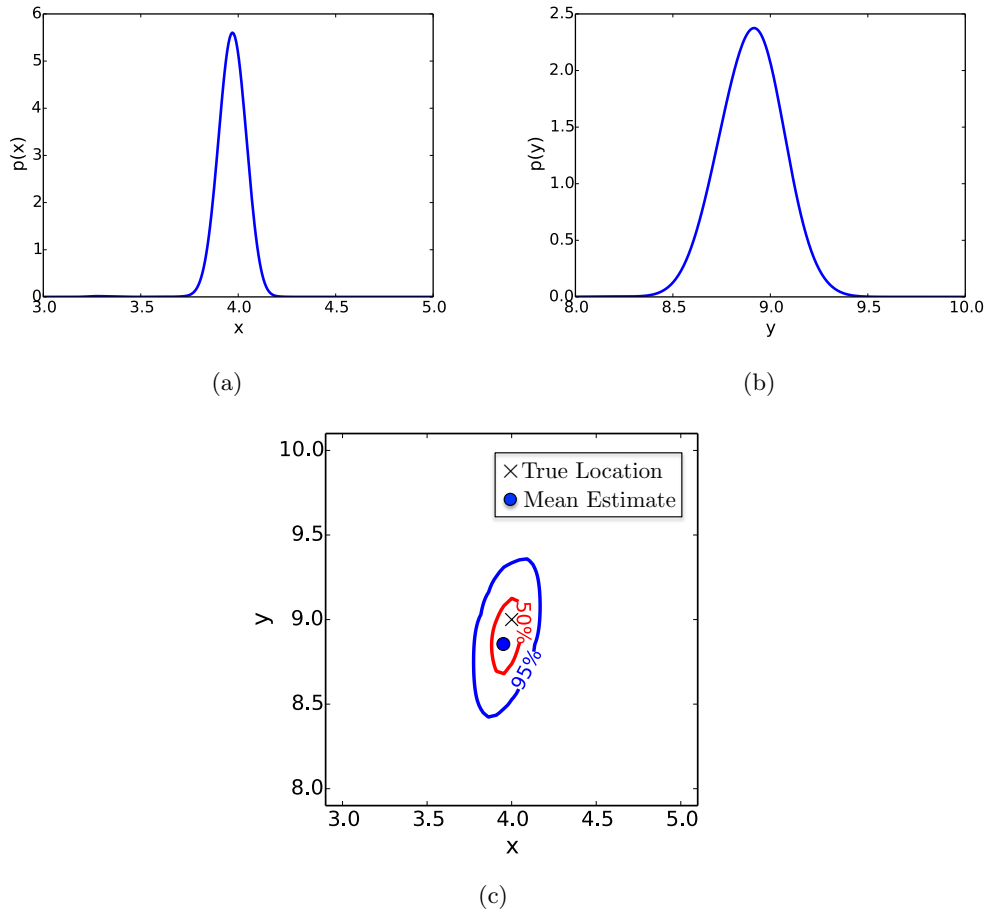


Figure 5. Sample-based estimates of the unknown crack location in Example 1 using 2,000 MCMC samples. a) The empirical probability distribution of the x coordinate. b) The empirical probability distribution of the y coordinate. c) The true location and mean estimate of the crack location, along with the 50% and 95% credible intervals.

With regard to the weighted likelihood formulation, a brief study was performed to quantify how much the approach improves the convergence of the sampling process. Typically, a large number of samples must be discarded from the beginning of the Markov chain during a so-called "burn-in" period to remove the effect of the initial guess, ultimately having a significant impact on solution time. Hence, the aim of the convergence study was to measure the size of the burn-in period necessary for both the standard Bayesian and proposed weighted likelihood posterior distributions in Example 1. The DRAM algorithm collected 5,000 samples and postprocessing was done to record how many samples it took for the chain to travel from an initial starting point to within a 0.5 radius of the true crack location. Three different starting points were considered, and since MCMC sampling is inherently random, five trials of the sampling process were performed from each point.

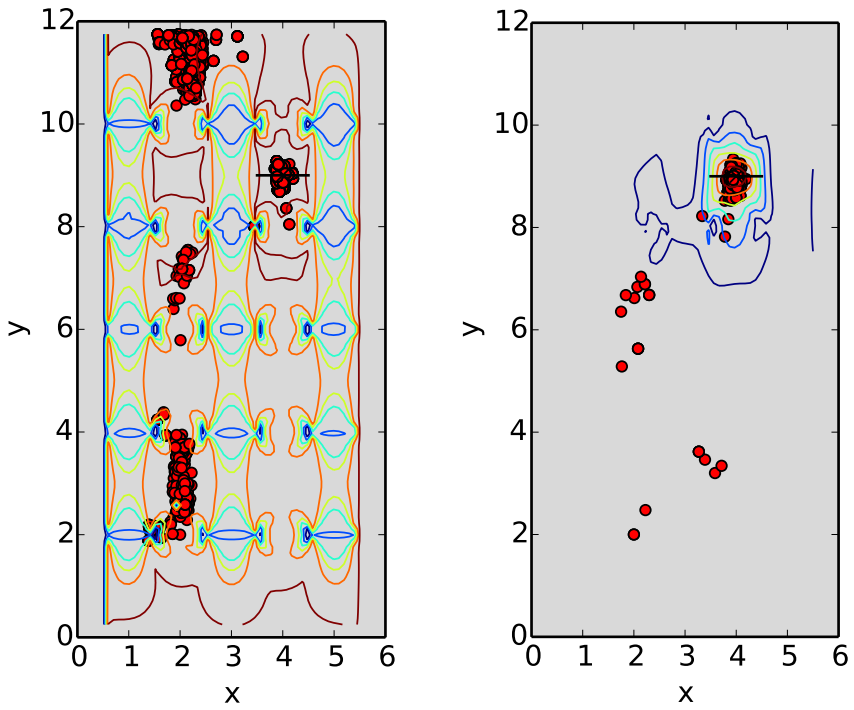
Table 1 displays the results of this convergence study, comparing the minimum, maximum and average number of samples it required to reach the crack location vicinity with both approaches for the three initial guesses. A substantial convergence improvement is observed for the weighted likelihood approach, showing roughly a 100-fold decrease in average number of samples for each initial guess. It is noted that for each initial guess, at least one of the five trials failed to reach the solution in the first 5,000 samples for the standard Bayesian approach, while the maximum reported value for the weighted likelihood approach was just 46 samples to reach the solution for an initial guess of $(2, 2)$. With such fast convergence using the weighted likelihood approach, the use of a burn-in period was neglected in this study with little impact on the sample-based estimates of the crack parameters.

For illustration purposes, Figures 6(a) and 6(b) show the results from one DRAM trial with the initial guess of $(2, 2)$ for the standard Bayesian and weighted likelihood approaches, respectively. Here, the first 2,000 samples are superimposed over each respective posterior distribution contour. The difficulty of sampling the standard posterior distribution is obvious, as the algorithm is temporarily trapped in three distinct modes before it ultimately reaches the mode corresponding to the true solution. On the other hand, the DRAM samples for the weighted likelihood formulation travel relatively directly to the true crack location where they remain for the duration of the sampling process.

A summary of the computational efficiency for identifying crack location in Example 1 is displayed in Table 2, comparing the standard Bayesian approach using the finite element method and the proposed weighted likelihood formulation using

Table 1: Convergence comparison between sampling a standard Bayesian and weighted likelihood posterior distribution in Example 1.

	Number of Samples to Solution					
Initial Guess	Standard Bayesian			Weighted Likelihood		
(x_0, y_0)	min.	max.	avg.	min.	max.	avg.
$(2, 2)$	1,447	5,000	3,319	18	46	35.8
$(3, 6)$	261	5,000	2,502	13	32	20.2
$(4, 2)$	119	5,000	2,177	14	43	25



(a) (b)

Figure 6. Results of the first 2,000 samples for one trial of the DRAM sampling algorithm superimposed over the posterior probability distribution contours for a) the standard Bayesian formulation and b) the weighted likelihood formulation. Both sampling trials started from an initial guess of (2,2).

a sparse grid interpolation surrogate model. It is shown that the proposed Bayesian framework can provide orders of magnitude speedup over a typical Bayesian approach, reducing the solution time for this example from 1.74 days to just 3.33 minutes. Note that the computation time needed to generate the training data for the surrogate model using finite element simulation is not included in this figure, as it is considered an offline cost. The number of samples for each approach in this comparison is approximate, though it was shown previously (Figure 5) that $2,000$ samples was sufficient for accurate solutions with the weighted likelihood approach. The number of samples required for the standard Bayesian solution was selected conservatively low (5,000) under the assumption that it takes roughly 2,000 additional samples to approximate the underlying distribution after the average burn-in period (3,319) specified in Table 1. It is further noted that the realized computational benefit of the proposed framework would increase significantly for more complex 3D models. In this case, the computational cost of the finite element model increases drastically while the surrogate model is relatively unaffected by the added complexity.

4.2 Example 2: General Crack Identification

The goal of Example 2 is to identify a crack of unknown location, size, and orientation using the proposed Bayesian framework. A schematic of the reference configuration for the problem is shown in Figure 2(b). The plate geometry, sensor placement, and boundary conditions remained unchanged from Example 1, but the crack parameters to be estimated were given by $[x, y, a, \theta] = [3, 6.5, 0.8, 25^\circ]$ in this case. The probability distributions of these four crack parameters were estimated for strain data with increasing noise levels $(1\%, 3\%, 5\%)$ to study the impact of measurement noise on the solutions. The surrogate model for the problem was again constructed beforehand using Abaqus and the `spinterp` package with a four dimensional sparse grid of 18,945 points. In this case, substantially more grid points were required to maintain sufficient accuracy due to the increase in dimension from two to four.

Starting from an initial guess of $\mathcal{C}_0 = [2, 2, 1.0, 0^\circ]$, the DRAM algorithm drew $2,000$ samples from the posterior distribution (Equation 22) using simulated strain data for each of the three noise levels. These samples were then used to calculate empirical probability distributions for each of the unknown crack parameters. These distributions are shown in Figure 7, where the true value of each parameter is indicated by the vertical dashed line. The mean estimates for each parameter were also calculated and are displayed in Table 3 along with the relative errors. As was

Table 2: Comparison of the computational efficiency for the proposed method versus a standard Bayesian approach in Example 1.

Approach	# Samples	Time per Sample	**Solution Time**
Standard Bayesian + FEM	5000	30 sec	**1.74 Days**
Weighted Likelihood + Surrogate	2000	0.1 sec	**3.33 Mins**

Table 3: Crack parameter mean estimates and errors in Example 2 for different noise levels.

Noise Level	x		y		a		θ	
	Mean	% Error	Mean	% Error	Mean	% Error	Mean	% Error
1%	3.003	0.13	6.505	0.08	0.403	0.50	27.48	9.93
3%	2.985	0.50	6.457	0.66	0.406	1.50	21.25	14.99
5%	2.989	0.37	6.289	3.25	0.438	9.49	23.96	4.13

the case in Example 1, fast convergence was observed in the sampling process with DRAM, where the change in the estimated crack parameter statistics was negligible after $2,000$ samples.

It can be seen that the proposed Bayesian framework effectively captures the growing uncertainty in the measurement data used, which is reflected in the widening of the estimated probability distributions in Figure 7 as the noise level increased. A general trend of decreasing accuracy in the mean estimates with increasing noise level can also be seen in Table 3, as expected. However, the mean estimates are seen to be generally accurate, especially for the location coordinate estimates. The crack orientation θ was the least accurately estimated, with an error as high as 14.99% for the 3% noise level. This is largely due to the fact that the strain data are least sensitive to θ in comparison with the other crack parameters, as also reflected in the support size of the distributions $p(\theta)$ in Figure 7(d). It is noted that the results presented in this problem were calculated in just under 30 minutes. Here, the increase in computation time was due to the added complexity of the surrogate model in four dimensions, which took roughly 0.9 seconds to evaluate in this case. However, a standard Bayesian analysis using the finite element method would still take well over one day of computation time.

5 Conclusion

In this study, a computationally-efficient Bayesian approach to damage determination was proposed and demonstrated for the problem of identifying structural cracks using noisy strain sensor data. Bayesian inference was used to postulate a probability distribution over the parameters describing the damage and a MCMC sampling procedure was then used to approximate this distribution numerically. Since the resulting damage diagnosis is in the form of a probability distribution rather than a point estimate, the method naturally quantifies the uncertainty induced by measurement errors and thus enables notions of credibility along with its predictions.

As this work was motivated by structural health management applications, the primary focus was on improving the computational efficiency of the method with respect to traditional Bayesian approaches by speeding up the MCMC sampling-based solution process. To this end, a surrogate modeling technique using sparse grid interpolation was employed to replace a more costly finite element simulation that would otherwise be evaluated for each sample drawn by the MCMC algorithm.

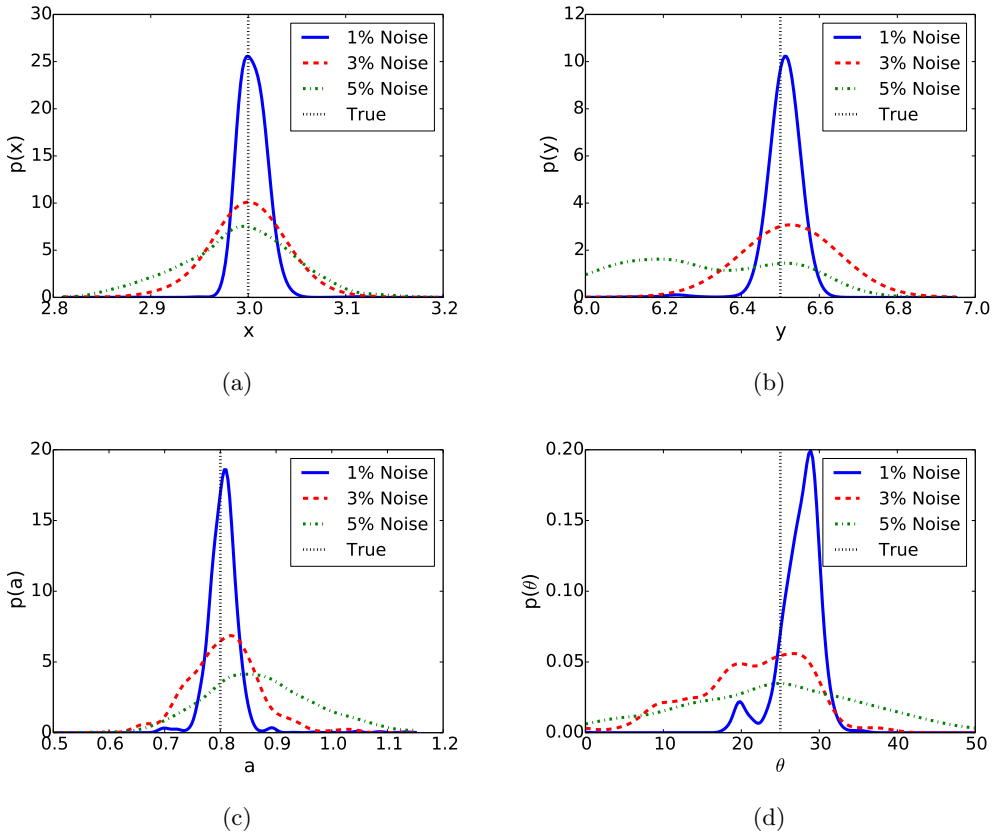


Figure 7. Empirical probability distributions of the crack parameters $\mathcal{C} = [x, y, a, \theta]$ from 2,000 MCMC samples for different noise levels in Example 2. Crack location in a) x and b) y, with length c) a, and orientation d) θ.

While surrogate modeling significantly reduced the time to draw each sample, a new weighted likelihood formulation of the posterior solution distribution was shown to increase the convergence of the sampling process, greatly reducing the total number of samples required for accurate estimates of the damage parameters. Additionally, the Delayed Rejection Adaptive Metropolis (DRAM) algorithm was adopted in this work for robust and effective MCMC sampling of the posterior distribution. It was shown through numerical examples that the proposed framework can provide orders of magnitude speedup with respect to a standard Bayesian approach and effectively captures the effects of measurement uncertainty in damage identification problems.

References

1. Sankararaman, S.; and Goebel, K.: Uncertainty in prognostics and health management: An overview. *2nd European Conference of the Prognostics and Health Management Society*, Nantes, France, July 2014.

2. Kim, J. T.; and Stubbs, N.: Improved damage identification method based on modal information. *Journal of Sound and Vibration*, vol. 252, 2002, pp. 223–238.

3. Mal, A. K.; Ricci, F.; Banerjee, S.; and Shih, F.: A conceptual structural health monitoring system based on vibration and wave propagation. *Structural Health Monitoring*, vol. 4, 2005, pp. 283–293.

4. Wang, L.; and Yuan, F. G.: Active damage localization technique based on energy propagation of Lamb waves. *Smart Structures and Systems*, vol. 3, 2007, pp. 201–217.

5. Kehlenbach, M.; and Hanselka, H.: Automated structural integrity monitoring based on broadband Lamb wave excitation and matched filtering. *Proceedings of the 44th AIAA, ASME, ASCE, AHS, ASC Structures, Structural Dynamics, and Materials Conference*, no. AIAA 2003-1563, Norfolk, VA, April 2003.

6. Krishnamurthy, T.; and Gallegos, A. M.: Damage characterization using the extended finite element method for structural health management. *52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 13th AIAA Non-Deterministic Approaches Conference*, no. AIAA-2011-1701, Denver, CO, April 2011.

7. Hochhalter, J. D.; Krishnamurthy, T.; Aguilo, M. A.; and Gallegos, A. M.: Strain-based damage determination using finite element analysis for structural health management. *Structural Health Monitoring*, In review.

8. Gelman, A.; Carlin, J. B.; Stern, H. S.; Dunson, D. B.; Vehtari, A.; and Rubin, D. B.: *Bayesian Data Analysis*. Chapman and Hall/CRC, Boca Raton, Florida, Third ed., 2013.

9. Gamerman, D.; and Lopes, H. F.: *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall/CRC, Boca Raton, Florida, Second ed., 2006.

10. Moore, E. Z.; Murphy, K. D.; and Nichols, J. M.: Crack identification in a freely vibrating plate using Bayesian parameter estimation. *Mechanical Systems and Signal Processing*, vol. 25, 2011, pp. 2125–2134.

11. Nichols, J. M.; Link, W. A.; Murphy, K. D.; and Olson, C. C.: A Bayesian approach to identifying structural nonlinearity using free-decay response: Application to damage detection in composites. *Journal of Sound and Vibration*, vol. 329, 2010, pp. 2995–3007.

12. Huhtala, A.; and Bossuyt, S.: A Bayesian approach to vibration based structural health monitoring with experimental verification. *Journal of Structural Mechanics*, vol. 44, no. 4, 2011, pp. 330–344.

13. Nichols, J. M.; Moore, E. Z.; and Murphy, K. D.: Bayesian identification of a cracked plate using a population-based Markov Chain Monte Carlo method. *Computers and Structures*, vol. 89, 2011, pp. 1323–1332.

14. Gaitanaros, S.; Karaiskos, G.; and P: A Bayesian methodology for crack identification in structures using strain measurements. *International Journal of Reliability and Safety*, vol. 4, no. 2, 2010, pp. 206–237.

15. Yan, G.: A Bayesian approach for identification of structural crack using strain measurements. *Sixth European Workshop on Structural Health Monitoring*, Dresden, Germany, July 2012.

16. Smolyak, S. A.: Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.*, vol. 4, 1963, pp. 240–243.

17. Klimke, A.; and Wohlmuth, B.: Algorithm 847: `spinterp`: Piecewise multilinear hierarchical sparse grid interpolation for MATLAB. *ACM Transactions on Mathematical Software*, vol. 31, no. 4, 2005, pp. 561–579.

18. Wang, X.; Eeden, C. V.; and Zidek, J. V.: Asymptotic properties of maximum weighted likelihood estimators. *Journal of Statistical Planning and Inference*, vol. 119, 2004, pp. 37–54.

19. Wang, X.: Approximating Bayesian inference by weighted likelihood. *The Canadian Journal of Statistics*, vol. 34, no. 2, 2006, pp. 279–298.

20. Haario, H.; Laine, M.; and Mira, A.: DRAM: Efficient adaptive MCMC. *Statistics and Computing*, vol. 16, no. 4, 2006, pp. 339–354.

21. The MathWorks Inc., Natick, Massachusetts: *MATLAB Version 7.10.0 (R2010a)*. 2010.

22. SIMULIA: *ABAQUS Documentation Version 6.12*. 2012.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-03-2016	Technical Publication	

4. TITLE AND SUBTITLE

Probabilistic Damage Characterization using a Computationally-Cfficient Bayesian Approach

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Warner, James E.; Hochhalter, Jacob D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

694478.02.93.02.14.58.23

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER

L-20667

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)

NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

NASA/TP-2016-219169

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified
Subject Category 65
Availability: STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This work presents a computationally-efficient approach for damage determination that quantifies uncertainty in the provided diagnosis. Given strain sensor data that are polluted with measurement errors, Bayesian inference is used to estimate the location, size, and orientation of damage. This approach uses Bayes' Theorem to combine any prior knowledge an analyst may have about the nature of the damage with information provided implicitly by the strain sensor data to form a posterior probability distribution over possible damage states. The unknown damage parameters are then estimated based on samples drawn numerically from this distribution using a Markov Chain Monte Carlo (MCMC) sampling algorithm. Since the resulting damage diagnosis is in the form of a probability distribution rather than a point estimate, it naturally enables notions of credibility and uncertainty in the predictions. Motivated by structural health management applications, the proposed damage identification strategy is designed to provide significant computational speedup over traditional Bayesian approaches. This is achieved by both decreasing the computational time to generate a single sample from the solution distribution and by reducing the total number of samples required to obtain accurate damage estimates. For the former, an efficient surrogate model is constructed using sparse grid interpolation to replace a costly finite element model that must otherwise be evaluated for each sample drawn.

15. SUBJECT TERMS

Bayesian inference; Damage detection; Surrogate modeling; Uncertainty quantification

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Help Desk (email: help@sti.nasa.gov)
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	28	(757) 864-9658