Sparse Bayesian Learning for Structural Damage Detection Using Expectation–Maximization Technique

Rongrong Hou[1], Yong Xia[1*], Xiaoqing Zhou[2], and Yong Huang[3]

[1] *Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China*

[2] *College of Civil Engineering, Shenzhen University, China*

[3] *Key Lab of Structures Dynamic Behavior and Control of the Ministry of Education, School of Civil Engineering, Harbin Institute of Technology, Harbin, China*

*Corresponding author, email: ceyxia@polyu.edu.hk

Abstract

Sparse Bayesian learning (SBL) methods have been developed and applied in the context of regression and classification, in which latent variables and hyper-parameters are iteratively obtained using type-II maximization likelihood. However, this method is ineffective in structural damage detection using modal parameters, which have a nonlinear relation with structural damage. Consequently, the analytical solution of the type-II maximization likelihood is unavailable. In this study, an iterative expectation–maximization (EM) technique is employed to tackle the difficulty. During the iteration, structural damage and hyper-parameters are updated through an expectation and maximization processes alternatively. Two sampling methods are utilized during the expectation procedure. Upon convergence, some hyper-parameters approach infinity and the associated damage variables become zero, resulting in the sparsity of damage. Numerical and experimental examples demonstrate that the proposed SBL method can accurately locate and quantify the sparse damage. The proposed EM technique is easy to implement while contains clear physical meaning.

Keywords

Sparse Bayesian learning, structural damage detection, expectation–maximization, nonlinear inverse problem, sparse recovery, modal parameters

1. Introduction

Structural damage detection has received much attention over the last decades [1–3]. Numerous vibration-based damage detection methods have been proposed using structural modal parameters [4–5] or time history [6].

Structural damage detection always entails uncertainties, which may be categorized as modeling errors, methodology errors, and measurement noise [7−9]. Moreover, operational and environmental variations also cause significant changes in the identified modal parameters [10–11]. Many researchers have proposed probabilistic approaches to tackle the uncertainties in structural damage detection. One category is based on the perturbation technique [8–9, 12–13], which assigns a random variable to each uncertainty and then calculates the statistics of the variable. The other widely used category is the Bayesian approach [14–17], which explicitly quantifies the posterior probability of the uncertainties according to observations and prior information.

Another difficulty in structural damage detection is that the problem is essentially an inverse problem and is typically ill-posed. As the number of available vibration measurements is limited, such detection is usually an underdetermined problem in mathematics [18]. In practice, structural damage commonly appears in a few sections or members only, especially at the early stage. Therefore, it possesses sparsity compared with the numerous elements of the entire structure. Structural damage sparsity is an important prior information that can be exploited. Recently, some researchers (including the authors) [18–24] developed the sparse recovery theory for structural damage detection using the so-called l_1 and l_0 regularization techniques. Sparse recovery theory, however, as shown later in this paper, disregards the relative uncertainties between different variables and causes determination of the regularization parameter difficult.

Bayesian methods provide an efficient way to deal with the ill-posed and underdetermined problem by specifying probability distributions over uncertain parameters, which is equivalent to introducing a regularization term to the optimization problem [25]. Sparse Bayesian learning (SBL) is effective in promoting sparsity in the inferred predictors and has been rapidly developed recently in the context of regression and classification [26–28]. For linear regression problems, the latent variables and associated hyper-parameters are iteratively obtained using type-II maximization likelihood [28]. During the process, most hyper-parameters tend to approach infinity, and thus the corresponding latent variables approach zero, resulting in a sparse regression model. However, SBL has seldom been utilized and explored in damage detection.

In terms of computation of the posterior probability of unknown parameters, the resulting posterior probability density function (PDF) is usually intractable. For example, the high-dimensional integral of the posterior PDF cannot be computed analytically. If the integral cannot be evaluated analytically, then Laplace's approximation method can be used, provided that the model class is globally identifiable according to the available data [29–30]. However, Laplace's approximation has several limitations, one of which is that the accuracy of the approximation depends on the dimension of the parameter vector [31]. If the measurements are considerably fewer than the model parameters, then the corresponding Bayesian updating results may be inaccurate. For unidentifiable and locally identifiable problems, simulation-based techniques, such as the Markov Chain Monte Carlo (MCMC), are proposed to compute the posterior statistics of unknown parameters [32–34]. Lam *et al.* [35] proposed a Bayesian method based on the MCMC algorithm for structural model updating and damage detection. Nichols *et al.* [36] utilized the MCMC approach for the Bayesian parameter estimation of structural systems for damage detection. However, the MCMC algorithm has high computational cost, especially for models with numerous parameters to be inferred. Moreover, evaluating convergence and accuracy for the MCMC methods is difficult, even when conducted empirically.

As the modal data is a nonlinear function of the structural stiffness parameters, the analytical solution of the type-II maximization likelihood is unavailable. Rather than directly tackling this nonlinear problem, Huang *et al.* [37–38] proposed an iterative procedure, which involves a series of coupled linear regression problems. A hierarchical SBL approach combined with Laplace's asymptotic approximation was used to infer the stiffness reductions on the basis of experimentally identified modal parameters. However, all uncertain hyper-parameters of the sophisticated hierarchical Bayesian model should be estimated, which is a nontrivial task. Huang *et al.* [39] converted the nonlinear function of structural stiffness parameters into a series of coupled linear-in-the-parameter problems. Two Gibbs sampling algorithms representing a special case of the MCMC simulation were proposed to provide a full treatment of the posterior PDFs of uncertain parameters for damage assessment. However, the computational efficiency of posterior sampling for the MCMC methods remains a major concern. Mustafa *et al.* [17] utilized linear optimization to identify the posterior statistics of the model parameters for model updating and damage detection, instead of solving the challenging nonlinear optimization problem. A complete review on the recent development of sparse Bayesian learning for structural damage detection and assessment was also provided in Huang *et al*. [40].

In this study, a concise expectation–maximization (EM) technique is employed to tackle the nonlinear problem iteratively, without performing the asymptotic approximation or stochastic simulation. During the expectation step, two sampling methods are used to calculate the expectation of the complete-data likelihood function. The effectiveness of the proposed algorithm is verified using a numerical and experimental example.

2. Structural Model Class

The structural model class, $\mathcal{M}$, is based on a set of linear structural models, where each model has a known mass matrix **M** and an uncertain stiffness matrix **K** parameterized by the stiffness parameters as follows

$$\mathbf{K} = \sum_{i=1}^{n} \alpha_i \, \mathbf{K}_i \tag{1}$$

where $\mathbf{K}_i$ is the *i*th element stiffness matrix which can be obtained from the finite element (FE) model of a structure; α_i is the *i*th element stiffness parameter to be updated according to the observed data; and *n* is the number of structural elements. The change in mass is assumed to be negligible when damage occurs.

Suppose that the element stiffness parameter reduces to $\bar{\alpha}_i$, the stiffness reduction factor (SRF) is then defined as [19, 41]

$$\theta_i = \frac{\bar{\alpha}_i - \alpha_i}{\alpha_i} \tag{2}$$

SRF, as the damage parameter, indicates both damage location and damage severity. As structural damage typically occurs at several sections or members only, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_n]^T$ is a sparse vector with only several non-zero components at the damaged locations.

The *r*th structural eigenvalue and the corresponding mode shape are governed by the following eigenvalue equation

$$(\mathbf{K} - \lambda_r \mathbf{M})\boldsymbol{\phi}_r = \mathbf{0} \tag{3}$$

Suppose that N_m modes of vibration have been identified from modal testing so that the identified eigenvalues and mode shapes can be expressed as

$$\hat{\boldsymbol{\lambda}} = \{\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_{N_m}\} \tag{4}$$

$$\hat{\boldsymbol{\Psi}} = [\hat{\boldsymbol{\phi}}_1, \hat{\boldsymbol{\phi}}_2, \cdots, \hat{\boldsymbol{\phi}}_{N_m}] \tag{5}$$

where $\hat{\boldsymbol{\phi}}_r \in R^{N_p}$ denotes the identified mode shape of the *r*th mode at N_p measurement points. A set of modal data is expressed as

$$\boldsymbol{\mathcal{D}} = [\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\Psi}}] \tag{6}$$

In the next section, a sparse Bayesian model is developed which automatically promotes sparsity in the inferred damage parameters $\boldsymbol{\theta}$.

3. Sparse Bayesian Modeling

Bayes' theorem is used to develop a PDF for the damage parameters $\boldsymbol{\theta}$**,** conditional on the measured modal data and chosen class of models [42]

$$p(\boldsymbol{\theta}|\boldsymbol{\mathcal{D}}, \mathcal{M}) = c^{-1} p(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta}, \mathcal{M}) p(\boldsymbol{\theta}|\mathcal{M}) \tag{7}$$

where $p(\boldsymbol{\theta}|\boldsymbol{\mathcal{D}}, \mathcal{M})$ is the posterior PDF of the damage parameters given the modal data and model class; $c = p(\boldsymbol{\mathcal{D}}|\, \mathcal{M})$ is a normalizing constant referred to as evidence, which does not affect the shape of the posterior distribution; $p(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta}, \mathcal{M})$ is the PDF of the modal data; and $p(\boldsymbol{\theta}|\mathcal{M})$ is the prior PDF of the damage parameters. The evidence can be used to estimate the hyper-parameters, as described later. To simplify the notation, the dependence of the PDF on $\mathcal{M}$ is dropped hereafter.

3.1 Likelihood functions for the stiffness parameter

According to the axioms of probability, the PDF of the modal data $p(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta})$ in Equation (7) can be expressed as [42]

$$p(\boldsymbol{\mathcal{D}}|\boldsymbol{\theta}) = \prod_{r=1}^{N_m} p(\hat{\lambda}_r|\boldsymbol{\theta}) p(\hat{\boldsymbol{\phi}}_r|\boldsymbol{\theta}) \tag{8}$$

To construct the prior distribution, the measurement error is introduced to the measured eigenvalues and mode shapes as

$$\lambda_r(\boldsymbol{\theta}) = \hat{\lambda}_r(1 + \varepsilon_r) \tag{9}$$

$$\boldsymbol{\phi}_r(\boldsymbol{\theta}) = \hat{\boldsymbol{\phi}}_r + \boldsymbol{e}_r \tag{10}$$

where $\lambda_r(\boldsymbol{\theta})$ and $\boldsymbol{\phi}_r(\boldsymbol{\theta})$ are the eigenvalue and mode shape of the *r*th mode obtained from the analytical model, respectively; and ε_r and $\boldsymbol{e}_r$ are measurement errors of frequencies and mode shapes, respectively.

The measurement errors are treated as independent Gaussian variables as [16, 43]

$$\varepsilon_r \sim N(0, \beta^{-1}) \tag{11}$$

$$\boldsymbol{e}_r \sim N(0,\ \gamma^{-1}\boldsymbol{I}) \tag{12}$$

where hyper-parameters β and γ reflect the precision of the identified eigenvalues $\hat{\boldsymbol{\lambda}}$ and mode shapes $\hat{\boldsymbol{\phi}}$, respectively. In this study, the precision, which is equal to the reciprocal of the variance of the variables, is used instead of the variance for convenience. The resulting likelihood function of $\boldsymbol{\theta}$ based on the measured eigenvalues $\hat{\boldsymbol{\lambda}}$ is

$$p(\hat{\boldsymbol{\lambda}}|\boldsymbol{\theta}, \beta) = \left(\frac{\beta}{2\pi}\right)^{\frac{N_m}{2}} exp\left\{-\frac{\beta}{2}\sum_{r=1}^{N_m}\left[\frac{\lambda_r(\boldsymbol{\theta}) - \hat{\lambda}_r}{\hat{\lambda}_r}\right]^2\right\} \tag{13}$$

and the likelihood function of $\boldsymbol{\theta}$ based on the measured mode shapes $\hat{\boldsymbol{\phi}}$ is

$$p(\hat{\boldsymbol{\phi}}|\boldsymbol{\theta}, \gamma) = \left(\frac{\gamma}{2\pi}\right)^{\frac{N_p \cdot N_m}{2}} exp\left\{-\frac{\gamma}{2}\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|_2^2\right\} \tag{14}$$

3.2 Prior distribution for damage parameters

As introduced previously, the prior PDF for the damage parameters is chosen to provide regularization for the ill-posed inverse damage detection problem. SBL proposes to use the automatic relevance determination (ARD) prior to incorporating a preference for sparser parameters [26, 38, 44, 45]. In addition, the Laplace distribution is also an effective prior to enforce sparsity in the inferred parameters, which is equivalent to l_1 regularization [37, 46].

Accordingly, the ARD prior is adopted in this study to promote sparsity in the damage parameters. The damage parameters $\boldsymbol{\theta}$ are assumed to be Gaussian with zero mean and covariance matrix $\boldsymbol{A} = diag(\alpha_1^{-1}, \cdots, \alpha_n^{-1})$, such that

$$p(\boldsymbol{\theta}|\boldsymbol{\alpha}) = \prod_{i=1}^{n} p(\theta_i|\alpha_i) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \prod_{i=1}^{n}\left[\alpha_i^{\frac{1}{2}}\, exp\left\{-\frac{1}{2}\alpha_i\theta_i^2\right\}\right] \tag{15}$$

where the individual hyper-parameter α_i represents the precision of the associated damage parameter θ_i.

3.3 Posterior distribution for damage parameters

From Equation (7), the posterior PDF of the damage parameters $\boldsymbol{\theta}$ is expressed as

$$p(\boldsymbol{\theta}|\hat{\lambda},\widehat{\boldsymbol{\Psi}},\boldsymbol{\alpha},\beta,\gamma)=c^{-1}p(\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\theta},\beta,\gamma)p(\boldsymbol{\theta}|\boldsymbol{\alpha})=c^{-1}p(\hat{\lambda}|\boldsymbol{\theta},\beta)p(\widehat{\boldsymbol{\Psi}}|\boldsymbol{\theta},\gamma)p(\boldsymbol{\theta}|\boldsymbol{\alpha})$$

$$=c^{-1}\left(\frac{\beta}{2\pi}\right)^{\frac{N_m}{2}}\left(\frac{\gamma}{2\pi}\right)^{\frac{N_p\cdot N_m}{2}}\left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\left(\prod_{i=1}^{n}\alpha_i^{\frac{1}{2}}\right)exp\left\{-\frac{\beta}{2}\sum_{r=1}^{N_m}\left[\frac{\lambda_r(\boldsymbol{\theta})-\hat{\lambda}_r}{\hat{\lambda}_r}\right]^2 -\frac{\gamma}{2}\sum_{r=1}^{N_m}\left\|\widehat{\boldsymbol{\phi}}_r-\boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|_2^2-\frac{1}{2}\sum_{i=1}^{n}(\alpha_i\theta_i^2)\right\} \quad (16)$$

with the distributions on the right-hand side as defined by Equations (13), (14), and (15), respectively. This posterior probability distribution will be used to quantify the plausibility of all possible values of the damage parameters.

4. Bayesian Inference Using EM Algorithm

The posterior PDF of the damage parameters, as defined in Equation (16), depends on the estimates of the hyper-parameters $\boldsymbol{\xi}=\{\boldsymbol{\alpha},\beta,\gamma\}$, which can be obtained by maximizing the evidence $p(\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})$, known in the statistics literature as type-II maximization likelihood [28]. As introduced previously, the evidence is the normalizing term that appears in the denominator in Bayes' theorem. According to the Total Probability Theorem, the evidence is obtained by integrating over the damage parameters as

$$p(\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})=\int p(\boldsymbol{\theta},\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\alpha},\beta,\gamma)d\boldsymbol{\theta}=\int p(\hat{\boldsymbol{\lambda}}|\boldsymbol{\theta},\beta)p(\widehat{\boldsymbol{\Psi}}|\boldsymbol{\theta},\gamma)p(\boldsymbol{\theta}|\boldsymbol{\alpha})d\boldsymbol{\theta} \quad (17)$$

However, the computation of the integral in Equation (18) is intractable, as the frequencies and mode shapes are in nonlinear relations with $\boldsymbol{\theta}$.

In this study, the EM algorithm is proposed to maximize the natural log evidence $\ln p(\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})$ iteratively. The EM algorithm enables parameter estimation in probabilistic models, where the model depends on unobserved latent variables. It alternates between performing an expectation (E) step and a maximization (M) step. $\boldsymbol{\theta}$ is regarded as a latent variable and $\{\boldsymbol{\theta},\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}\}$ is referred to as the complete data set. The complete-data natural log likelihood function is expressed as

$$\ln p(\boldsymbol{\theta},\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})=\ln p(\hat{\boldsymbol{\lambda}}|\boldsymbol{\theta},\beta)+\ln p(\widehat{\boldsymbol{\Psi}}|\boldsymbol{\theta},\gamma)+\ln p(\boldsymbol{\theta}|\boldsymbol{\alpha})$$

$$=\frac{N_m}{2}\ln\left(\frac{\beta}{2\pi}\right)-\frac{\beta}{2}\sum_{r=1}^{N_m}\left[\frac{\hat{\lambda}_r-\lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r}\right]^2+\frac{N_0\cdot N_m}{2}\ln\left(\frac{\gamma}{2\pi}\right) -\frac{\gamma}{2}\sum_{r=1}^{N_m}\left\|\widehat{\boldsymbol{\phi}}_r-\boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|_2^2+\frac{n}{2}\ln\left(\frac{1}{2\pi}\right)+\frac{1}{2}\sum_{i=1}^{n}\ln\alpha_i-\frac{1}{2}\sum_{i=1}^{n}(\alpha_i\theta_i^2) \quad (18)$$

Given the difficulty of direct maximization of $\ln p(\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})$ with respect to $\boldsymbol{\xi}$, the EM algorithm proposes to maximize the expectation of the complete-data $E\{\ln p(\boldsymbol{\theta},\hat{\boldsymbol{\lambda}},\widehat{\boldsymbol{\Psi}}|\boldsymbol{\xi})\}$ instead [28, 47], such that

$$E\{\ln p\,(\boldsymbol{\theta}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}|\boldsymbol{\xi})\} = \frac{N_m}{2}\ln\left(\frac{\beta}{2\pi}\right) - \frac{\beta}{2}E\left\{\sum_{r=1}^{N_m}\left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r}\right]^2\right\}$$

$$+\frac{N_p \cdot N_m}{2}\ln\left(\frac{\gamma}{2\pi}\right) - \frac{\gamma}{2}E\left\{\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|_2^2\right\} + \frac{n}{2}\ln\left(\frac{1}{2\pi}\right) \tag{19}$$

$$+\frac{1}{2}\sum_{i=1}^{n}\ln\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\alpha_i E(\theta_i^2)$$

In practice, the complete data set is not available and the latent variable $\boldsymbol{\theta}$ is given by the posterior distribution $p(\boldsymbol{\theta}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\xi})$. In the E step, given the current value of the hyper-parameter $\boldsymbol{\xi}^{\text{old}}$, the posterior distribution of $\boldsymbol{\theta}$ given by $p(\boldsymbol{\theta}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\xi}^{\text{old}})$ is used to determine the expectation of the complete-data $E\{\ln p\,(\boldsymbol{\theta}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}|\boldsymbol{\xi})\}$. In the subsequent M step, the new estimate $\boldsymbol{\xi}^{\text{new}}$ is obtained by maximizing the expectation with respect to $\boldsymbol{\xi}$. By differentiating Equation (19) with respect to $\boldsymbol{\alpha}$, β, and γ, and then setting these derivatives to zero, we obtain

$$\frac{\partial E\{\ln p\,(\boldsymbol{\theta}, \boldsymbol{\xi}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}})\}}{\partial \alpha_i} = \frac{1}{2\alpha_i} - \frac{1}{2}E(\theta_i^2) = 0 \tag{20}$$

$$\frac{\partial E\{\ln p\,(\boldsymbol{\theta}, \boldsymbol{\xi}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}})\}}{\partial \beta} = \frac{N_m}{2}\frac{1}{\beta} - \frac{1}{2}E\left\{\sum_{r=1}^{N_m}\left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r}\right]^2\right\} = 0 \tag{21}$$

$$\frac{\partial E\{\ln p\,(\boldsymbol{\theta}, \boldsymbol{\xi}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}})\}}{\partial \gamma} = \frac{N_p \cdot N_m}{2}\frac{1}{\gamma} - \frac{1}{2}E\left\{\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|_2^2\right\} = 0 \tag{22}$$

The hyper-parameters are then solved as

$$\alpha_i = \frac{1}{E(\theta_i^2)} \tag{23}$$

$$\beta = \frac{N_m}{E\left\{\sum_{r=1}^{N_m}\left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r}\right]^2\right\}} \tag{24}$$

$$\gamma = \frac{N_p \cdot N_m}{E\left\{\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|^2\right\}} \tag{25}$$

$E\{\cdot\}$ denotes an expectation with respect to the posterior distribution of $\boldsymbol{\theta}$ using the current estimates of the hyper-parameters $\boldsymbol{\xi}^{\text{old}}$.

4.1 Posterior sampling

Posterior sampling is conducted to approximate the expectations in Equations (23), (24), and (25). We first approximate the conditional posterior PDF $p(\boldsymbol{\theta}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma)$ of stiffness parameter $\boldsymbol{\theta}$ in Equation (16) by a multivariate Gaussian distribution using Laplace's method of asymptotic approximation [29]. This is based on the assumption that the modal data $\{\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}\}$ available is sufficient to constrain the updated parameter $\boldsymbol{\theta}$ to give a globally identifiable model class. The mean of the Gaussian distribution is the maximum posterior (MAP) estimate $\tilde{\boldsymbol{\theta}}$, which is calculated by maximizing $\ln p\,(\boldsymbol{\theta}|\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\xi})$, or equivalently minimizing the following objective function

$$J(\boldsymbol{\theta}) = \beta \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r} \right]^2 + \gamma \sum_{r=1}^{N_m} \sum_{j=1}^{N_p} \left[\hat{\phi}_{j,r} - \phi_{j,r}(\boldsymbol{\theta}) \right]^2 + \sum_{i=1}^{n} (\alpha_i \theta_i^2) \tag{26}$$

where $\hat{\phi}_{j,r}$ and $\phi_{j,r}(\boldsymbol{\theta})$ denote the rth measured and analytical mode shapes at the jth point, respectively. The covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ of the approximated Gaussian distribution is equal to the inverse of the Hessian matrix $\mathcal{H}(\tilde{\boldsymbol{\theta}})$ calculated at $\tilde{\boldsymbol{\theta}}$, where the (i, j) component of the Hessian matrix $\mathcal{H}(\tilde{\boldsymbol{\theta}})$ is given by

$$\mathcal{H}_{i,j}(\tilde{\boldsymbol{\theta}}) = \left. \frac{\partial^2 J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \right|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \tag{27}$$

Since the dimension of $\boldsymbol{\theta}$ is large, the obtained covariance matrix is not positive semi-definiteness. In this regard, we calculate the variance for each damage parameter θ_i independently [42]. Then we generate samples from the posterior PDF $p(\boldsymbol{\theta} | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma)$ and the probabilistic information encapsulated in $p(\boldsymbol{\theta} | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma)$ is characterized by the posterior samples $\boldsymbol{\theta}^{(k)}, k = 1, \dots, K$ (we will choose the number of samples $K = 5{,}000$ in later examples). The expectations in (23), (24) and (25) are finally approximated by (28), (29) and (30), respectively

$$E(\theta_i^2) = \int \theta_i^2 p(\theta_i | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma) d\theta_i \approx \frac{1}{K} \sum\nolimits_{k=1}^{K} \left((\theta_i)^{(k)} \right)^2 \tag{28}$$

$$\begin{aligned} E\left\{ \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r} \right]^2 \right\} &= \int \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r} \right]^2 p(\theta_i | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma) d\boldsymbol{\theta} \\ &\approx \frac{1}{K} \sum\nolimits_{k=1}^{K} \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta}^{(k)})}{\hat{\lambda}_r} \right]^2 \end{aligned} \tag{29}$$

$$\begin{aligned} E\left\{ \sum_{r=1}^{N_m} \left\| \hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta}) \right\|^2 \right\} &= \int \sum_{r=1}^{N_m} \left\| \hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta}) \right\|^2 p(\theta_i | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma) d\boldsymbol{\theta} \\ &\approx \frac{1}{K} \sum_{r=1}^{N_m} \left\| \hat{\boldsymbol{\phi}}_r - \boldsymbol{\phi}_r(\boldsymbol{\theta}^{(k)}) \right\|^2 \end{aligned} \tag{30}$$

4.2 Likelihood sampling

Considering the complexity of the posterior PDF $p(\boldsymbol{\theta} | \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\psi}}, \boldsymbol{\alpha}, \beta, \gamma)$, which might not be Gaussian, another sampling method is proposed based on the likelihood function of $\boldsymbol{\theta}$. N_s sets of modal data $\boldsymbol{\mathcal{D}}_j = [\hat{\boldsymbol{\lambda}}^{(j)}, \hat{\boldsymbol{\psi}}^{(j)}]$ $(j = 1, 2, \dots, N_s)$, are generated according to the measured modal data following Gaussian distribution, that is, Equations (11) and (12). The mean of the Gaussian distribution is equal to the measured modal data with assigned variance. For each data set, given the current estimates of the hyper-parameters $\boldsymbol{\xi}$, the MAP estimate $\tilde{\boldsymbol{\theta}}$ is calculated by minimizing the objective function in Equation (26). The expectation is then performed with respect to the MAP values of $\boldsymbol{\theta}$ as

$$E(\theta_i^2) = E(\tilde{\theta}_i^2) \tag{31}$$

$$E\left\{ \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r} \right]^2 \right\} = E\left\{ \sum_{r=1}^{N_m} \left[\frac{\hat{\lambda}_r - \lambda_r(\tilde{\boldsymbol{\theta}})}{\hat{\lambda}_r} \right]^2 \right\} \tag{32}$$

$$E\left\{\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r-\boldsymbol{\phi}_r(\boldsymbol{\theta})\right\|^2\right\}=E\left\{\sum_{r=1}^{N_m}\left\|\hat{\boldsymbol{\phi}}_r-\boldsymbol{\phi}_r(\widetilde{\boldsymbol{\theta}})\right\|^2\right\} \tag{33}$$

4.3 Summary

Each iteration uses estimates of $\boldsymbol{\xi}$ to determine the posterior distribution of the latent variable $\boldsymbol{\theta}$. The current distribution of $\boldsymbol{\theta}$ is in turn utilized to improve the estimates of $\boldsymbol{\xi}$. The proposed EM algorithm is implemented as follows:

1. Initialize the hyper-parameters $\boldsymbol{\xi}^{(0)}$ and latent variable $\boldsymbol{\theta}^{(0)}$.
2. At the ith iteration,

   E step: Compute the MAP estimates of $\widetilde{\boldsymbol{\theta}}^{(i)}$ through minimizing $J(\boldsymbol{\theta})$ in Equation (26) given hyper-parameters $\boldsymbol{\xi}^{(i-1)}$; Calculate the expectations in Equations (23)~(25) using (28)~(30) for the posterior sampling and (31)~(33) for the likelihood sampling;

   M step: Through maximization of $E\{\ln p\left(\boldsymbol{\theta},\hat{\boldsymbol{\lambda}},\hat{\boldsymbol{\psi}}|\boldsymbol{\xi}\right)\}$ with respect to $\boldsymbol{\alpha}$, β, and γ, update the hyper-parameters to $\boldsymbol{\xi}^{(i)}$ according to Equations (23), (24), and (25), given $\widetilde{\boldsymbol{\theta}}^{(i)}$.
3. Repeat Step 2 for the (i+1)th iteration until the following convergence criterion is met:

$$\left\|\widetilde{\boldsymbol{\theta}}^{(i)}-\widetilde{\boldsymbol{\theta}}^{(i-1)}\right\|_2/\left\|\widetilde{\boldsymbol{\theta}}^{(i)}\right\|_2 \le Tol$$

It is notes that the posterior sampling is conducted after E step once $\widetilde{\boldsymbol{\theta}}^{(i)}$ is obtained at each iteration step, whereas the likelihood sampling is conducted at the initialization stage only.

The main advantages of the EM algorithm are its simplicity and ease of implementation. Moreover, the EM algorithm is proven to be stable and robust [48]. In this study, the integral of the nonlinear inverse problem is avoided through an iterative procedure. The MAP values of the damage parameters are determined adaptively, and the sparsity of the damage parameters is secured automatically. As examples in following sections show, during the iterative process, many hyper-parameters α_i approach infinity, and the corresponding θ_i approaches zero.

The SBL has some similarities with the sparse recovery theory that uses the regularization technique. In Equation (26), the first two terms are equivalent to the data-fitting terms with different weights and the third term to the regularization term with α as the regularization parameter in sparse recovery. In this study, each damage parameter θ_i is assigned with an individual hyper-parameter α_i, indicating that each damage parameter has a unique regularization parameter. A large α_i carries a significant weight in the corresponding θ_i and thus enforces it to be close to zero during the optimization, thereby achieving the sparse solution to θ. From Equation (23), the regularization parameter is at the similar level of $1/\theta^2$. Therefore, Equation (26) is equivalent to the iteratively reweighted least squares (IRLS) [49–51], which closely resembles the l_0 regularization. If all damage parameters $\boldsymbol{\theta}$ are assumed to have a uniform Gaussian prior distribution, the hyper-parameter $\boldsymbol{\alpha}$ becomes a single value. The resulting objective function (Equation [26]) is equivalent to the l_2 regularization, which may cause the identified damage distributed to many structural elements and fail to obtain the sparse solution [18–19]. If a Laplace prior is applied to the damage parameter, then the corresponding objective function is equivalent to the l_1 regularization [46]. Therefore, the EM-based SBL technique is more general and more flexible relative to sparse recovery theory. Moreover, the hyper-parameters possess a clear physical meaning that represents the precision of the uncertainties. The hyper-parameters are updated automatically, thereby avoiding the tricky selection of the regularization parameter in sparse recovery theory.

5. Comparison of new SBL method with the methods in [37], [38] and [39]

For the purpose of Bayesian damage identification and assessment, several SBL methods were proposed in [37], [38] and [39], where the SBL method in [38] is an improvement of that in [37] by removing some unreliable approximations. The present study differs from the above methods in following aspects:

(1) In the presented SBL method, the EM algorithm is employed for learning the hyper-parameters $\{\boldsymbol{\alpha}, \beta, \gamma\}$ to maximize the evidence function in (17), while direct differentiation was used in [37, 38, 39]. In fact, it is intractable to maximize the current form of the evidence function in (17) with respect to the hyper-parameters $\{\boldsymbol{\alpha}, \beta, \gamma\}$ by direct differentiation. Therefore, the EM algorithm employed here has broader applications for real problems. The present study provides an alternative approach to SBL for model updating and damage assessment.

(2) The Bayesian modeling and inference formulation presented in the paper are simpler and easier to be implemented, compared with those in [37, 38, 39]. For example, for the characterization of the posterior PDF of the stiffness parameters $\boldsymbol{\theta}$, only three hyper-parameters $\{\boldsymbol{\alpha}, \beta, \gamma\}$ need to be learned here. In [38], however, ten parameters including the high-dimensional system mode shape parameters, are required to be optimized simultaneously to find their MAP values. Therefore, the iterative method for the optimization in the present paper tends to converge faster and likely to the global maximum of the solution. As shown in the numerical example later, the convergence occurs in three or four iterations only. In [39], a Gibbs sampling based stochastic simulation method was presented by generating samples to characterize the posterior PDF of $\boldsymbol{\theta}$, in which high computational efforts are required.

(3) The new proposed method still works when only the measured eigenvalues $\hat{\boldsymbol{\lambda}}$ are available, as shown in the numerical and experimental examples in subsequent sections. This is quite useful for real applications when accurate structural mode shapes $\hat{\boldsymbol{\psi}}$ are not available. In [37, 38, 39], the measured mode shapes $\hat{\boldsymbol{\psi}}$ must be available and utilized for a reliable inference of the stiffness parameters $\boldsymbol{\theta}$.

6. Numerical Example

6.1 Model description

As shown in Figure 1, a cantilever beam, is first utilized as a numerical preliminary study. The mass density and Young's modulus are 7.67×10^3 kg/m^3 and 7.0×10^{10} N/m^2, respectively. The beam is modeled with 45 equal Euler–Bernoulli beam elements (i.e., $n = 45$), each 20 mm long. Elements 1 and 45 are located close to the clamped end and free end, respectively. The damage is simulated by the reduction of the bending stiffness while mass remains unchanged. In this study, Elements 1 at the clamped end and 23 at the mid-span are damaged by 50%, i.e., $\theta_1 = \theta_{23} = -0.5$ and all other $\theta = 0$.

Figure 1 Geometric configuration of the beam structure (unit: mm)

6.2 Damage detection

In this numerical example, the natural frequencies only are used for damage detection. The first six natural frequencies of the beam before and after damage are listed in Table 1. Given that only the natural frequencies are used for damage detection, the objective function in Equation (26) is simplified as

$$J_1(\boldsymbol{\theta}) = \frac{\beta}{2}\sum_{r=1}^{N_m}\left[\frac{\hat{\lambda}_r - \lambda_r(\boldsymbol{\theta})}{\hat{\lambda}_r}\right]^2 + \frac{1}{2}\sum_{i=1}^{n}(\alpha_i\theta_i^2) \tag{34}$$

Table 1 Frequencies of the beam in the undamaged and damaged states

Mode no.	Undamaged Freq. (Hz)	Damaged Freq. (Hz)	Change ratio (%)
1	6.02	5.75	−4.56
2	37.75	35.67	−5.50
3	105.73	102.44	−3.11
4	207.25	197.69	−4.61
5	342.70	333.96	−2.55
6	512.07	492.45	−3.83
Average of frequency change (%)			−4.03

To implement the EM process, the hyper-parameters $\boldsymbol{\alpha}$ and β should be initialized by setting them to the precision of the uncertainties. The uncertainty level of the damage parameters is assumed as 10% of the exact damage parameter. Therefore, the initial value $\alpha_i^{(0)} = 1/(10\%)^2 = 100$ $(i = 1, 2, \ldots, 45)$. In practical vibration tests, natural frequencies may typically contain 1% noise [9, 52–54]. Consequently $\beta^{(0)} = 1/(1\%)^2 = 1 \times 10^4$. The ratio of α_i to β is analogous to the regularization parameter used in the IRLS algorithm, which is $\alpha_i^{(0)} / \beta^{(0)} = 0.01$ in the current example. The initial damage parameters $\boldsymbol{\theta}^{(0)}$ are set at their nominal values $\boldsymbol{\theta}^{(0)} = \{0, \ldots, 0\}^T$, indicating that no damage is present.

With the minimization of Equation (26), the MAP values of the damage parameters are then obtained. Using the posterior sampling, 5000 samples of $\boldsymbol{\theta}^{(k)}$ are generated and then the hyper-parameters are calculated through Equations (23)~(25) and (28)~(30). With the proposed iterative EM, the MAP values of $\boldsymbol{\theta}$ in each iteration are updated and shown in Figure 2. In the first iteration, a number of elements have non-zero SRFs. After one more iteration, the identified damage parameters tend to approach the actual values and the process converges after three iterations only.

(a) Iteration no.1 (b) Iteration no.2

(c) Iteration no.3

Figure 2 Damage identification results during the iterative process using posterior sampling

The likelihood sampling is then also applied, in which 50 sets of natural frequencies are generated (i.e., N_s=50) with Gaussian distribution having a zero mean and 1% standard deviations of the true values.

Within EM, each set of sampled natural frequencies results in one set of MAP values of the damage parameters, from which the expectations are calculated according to Equations (31)~(33). The mean of the MAP values of $\boldsymbol{\theta}$ in each iteration are shown in Figure 3. Similar to the above results, the damage parameters converge after four iterations only and the actual damaged elements are correctly located and quantified. Although the likelihood sampling approach takes one more iteration, the identified damage parameters are more accurate than those of the posterior sampling.

(a) Iteration no.1 (b) Iteration no.2
(c) Iteration no.3 (d) Iteration no.4

Figure 3 Damage identification results during the iterative process using likelihood sampling

During the iteration, the hyper-parameters also change continuously. β and a few representative α_i are shown in Figure 4 for the two sampling methods. As the iteration proceeds, the hyper-parameters associated with the damaged elements, i.e., α_1 and α_{23}, decrease quickly and converge after only a few iterations. For the undamaged Element 45, the corresponding hyper-parameter α_{45} increases rapidly to a sizeable number (a logarithmic coordinate is used). The variation of β is not significant during the process.

(a) Posterior sampling (b) Likelihood sampling

Figure 4 Variation of hyper-parameters during the iterative process

To investigate the effect of noise on damage detection accuracy, a different level of Gaussian noise of the frequencies is introduced, namely, Noise Level 2 has a zero mean and 2% standard deviation of the exact natural frequencies. In this regard, the initial estimate of the hyper-parameter β becomes 2.5×10^3. Therefore, the equivalent regularization parameters is 0.04, with the initial $\alpha_i^{(0)} = 100$ $(i = 1, 2, \dots, 45)$ remaining unchanged.

Using the same EM procedure, the identified MAP values of the damage parameters are shown in Figures 5 and 6 using two sampling methods. The convergence occurs in three and four iterations, respectively. In the case of the posterior sampling, the true damage at Element 1 is detected correctly, whereas Element 23 at the mid-span cannot. Using the likelihood sampling, both damaged elements are located accurately and the identified SRFs are close to their actual values.

The above numerical results show that the proposed EM algorithm is effective in locating and quantifying structural damage. For the posterior sampling, the convergence rate is faster; while the damage identification results obtained are not accurate as compared with the likelihood sampling, especially at the higher noise level. Previous studies indicated that the main disadvantage of the EM algorithm is its slow convergence in some cases [16, 47]. However, the proposed EM algorithm for damage detection has fast convergence. All converge within four iterations for the two noise levels.

(a) Iteration no.1 (b) Iteration no.2
(c) Iteration no.3

Figure 5 Damage identification results for Noise level 2 using posterior sampling

(a) Iteration no.1 (b) Iteration no.2
(c) Iteration no.3 (d) Iteration no.4

Figure 6 Damage identification results for Noise Level 2 using likelihood sampling

7. Experimental Example

7.1 Model description

The experimental cantilever beam used here was presented in [54]. The geometric dimensions of the beam are shown in Figure 7. The mass density and Young's modulus are estimated as 7.67×10^3 kg/m^3 and 2.0×10^{11} N/m^2, respectively. Three saw cuts were sequentially introduced into the beam model, corresponding to four damage scenarios (DSs) (Figure 7). The saw cuts have the same length (b = 10 mm) but have varied depths to simulate different damage severities.

10 accelerometers were equidistantly mounted on the beam to measure the out-of-plane horizontal vibration. Modal testing was first conducted on the intact beam, and then repeated for each damage state. The first six natural frequencies and mode shapes were extracted accordingly, as listed in Table 2.

Table 2 Modal data of the undamaged and damaged beam structures

Mode no.	Undamaged	DS1		DS2		DS3		DS4	
	Freq. (Hz)	Freq. (Hz)	MAC	Freq. (Hz)	MAC	Freq. (Hz)	MAC	Freq. (Hz)	MAC
1	3.53	3.49 (−1.24)	99.99	3.38 (−4.41)	99.97	3.33 (−5.91)	99.98	3.36 (−4.91)	99.99
2	21.77	21.39 (−1.72)	99.95	20.85 (−4.26)	99.84	20.29 (−6.81)	99.86	19.76 (−9.22)	99.95
3	60.78	59.46 (−2.16)	99.88	58.93 (−3.04)	99.83	58.38 (−3.95)	99.57	54.37 (−10.55)	99.60
4	119.46	118.31 (−0.96)	99.88	116.01 (−2.88)	99.51	113.35 (−5.12)	99.23	106.31 (−11.01)	99.06
5	194.78	191.98 (−1.44)	99.78	188.74 (−3.10)	99.17	188.46 (−3.25)	98.87	187.17 (−3.91)	99.14
6	292.82	281.56 (−3.84)	98.07	286.76 (−2.07)	94.95	275.08 (−6.06)	98.26	267.45 (−8.66)	97.26
Average		(−1.90)	99.59	(−3.29)	98.88	(−5.18)	99.30	(−8.04)	99.17

Note: Values in parentheses are the frequency change ratios (%) between the damaged and undamaged states. MAC = modal assurance criterion.

In the FE model, the beam is divided into 100 Euler–Bernoulli beam elements, each 10 mm long. Given that the length of each cut is identical to the length of one element, the damage severity of each cut is equal to the reduction in the moment of inertia of the cross section, which is equal to the reduction in the depth of the cut section. In this manner, the damage severity is quantified. Table 3 lists the damage locations and severities quantified by SRF for the four DSs. At most, three damaged elements have non-zero SRF values. Therefore, the SRF vector is very sparse compared to the 100 elements of the beam model.

Figure 7 Overview of the beam structure (unit: mm)

Table 3 Damage locations and severities for four DSs

Damage Scenario	Cut no.	Cut depth (mm)	SRF (θ)
DS1	1	10	θ_1 =−40%
DS2	1	15	θ_1 =−60%
DS3	1	15	θ_1 =−60%
	2	15	θ_{50} =−60%
DS4	1	15	θ_1 =−60%
	2	15	θ_{50} =−60%
	3	20	θ_{75} =−80%

7.2 Damage detection

For the current experimental beam, only one set of measured modal data are available. In the likelihood sampling, 50 sets of modal data $\mathcal{D}_j = [\hat{\boldsymbol{\lambda}}^{(j)}, \hat{\boldsymbol{\Psi}}^{(j)}]$ $(j = 1, 2, \dots, 50)$ are generated as

$$\hat{\lambda}_r^{(j)} \sim N\left(\hat{\lambda}_r, \left(0.01\hat{\lambda}_r\right)^2\right) \tag{35}$$

$$\hat{\boldsymbol{\phi}}_r^{(j)} \sim N(\hat{\boldsymbol{\phi}}_r, (0.05)^2\boldsymbol{I}) \tag{36}$$

where the uncertainty levels of 1% and 5% are adopted for natural frequencies and mode shapes, respectively as natural frequencies are generally measured more accurately than mode shapes,. The hyper-parameters are thus initialized as $\beta^{(0)} = 1 \times 10^4$, $\gamma^{(0)} = 400$ and $\alpha_i^{(0)} = 100$ $(i = 1,2, \dots, 100)$.

First, only the natural frequencies are utilized for damage detection. Following the iterative procedures summarized in Section 4, the most plausible values of the damage parameters can be obtained using the two proposed sampling methods. To quantify the damage identification accuracy, the identification error δ is defined as

$$\delta = \frac{\|\tilde{\boldsymbol{\theta}} - \bar{\boldsymbol{\theta}}\|_2}{\|\bar{\boldsymbol{\theta}}\|_2} \tag{37}$$

where $\bar{\boldsymbol{\theta}}$ denotes the actual damage parameters in the experiment. The damage identification results of the four DSs using the likelihood sampling are shown in Figure 8. For DS1 and DS2, the actual damage is located and quantified with good accuracy. For DS3, the two damaged elements are located successfully, while the severity of the damage at the mid-span is smaller than the true value, i.e., with 60% reduction. For DS4, the actual damage locations are detected successfully with some errors in the identified severities. However, when using the posterior sampling, only the damaged Element 1 can be detected in all four DSs and damaged Elements 50 and 75 cannot be detected at all in DS3 and DS4. The results are not shown here for brevity.

(a) DS1 ($\bar{\theta}_1 = -0.4,\ \delta = 1.00\%$) (b) DS2 ($\bar{\theta}_1 = -0.6,\ \delta = 0.83\%$)
(c) DS3 ($\bar{\theta}_1 = \bar{\theta}_{50} = -0.6,\ \delta = 2.36\%$) (d) DS4 ($\bar{\theta}_1 = \bar{\theta}_{50} = -0.6,\ \bar{\theta}_{75} = -0.8,\ \delta = 2.36\%$)

Figure 8 Damage identification result for four DSs using frequencies only

When both the natural frequencies and mode shapes are used, the optimal damage parameters are calculated by minimizing the objective function in Equation (26) iteratively. The damage identification results using the likelihood sampling are shown in Figure 9. For the single damage scenarios, i.e., DS1 and DS2, the damage identification results are almost the same as those using frequencies only. For DS3, the identified damage severity of the damage at the mid-span becomes closer to the actual value and the error reduces to 4.12%. For DS4, the identification error reduces to 2.27%. Therefore the damage identification accuracy is improved significantly by incorporating the mode shape data. For all DSs using different kinds of modal parameters, the results converge within five iterations. Again the results using the posterior sampling are not accurate, for example, element 50 is not detected in DS4.

(a) DS1 (Actual damage $\bar{\theta}_1 = -0.4,\ \delta = 0.83\%$) (b) DS2 ($\bar{\theta}_1 = -0.6,\ \delta = 0.80\%$)
(c) DS3 ($\bar{\theta}_1 = \bar{\theta}_{50} = -0.6,\ \delta = 1.03\%$) (d) DS4 ($\bar{\theta}_1 = \bar{\theta}_{50} = -0.6,\ \bar{\theta}_{75} = -0.8,\ \delta = 1.06\%$)

Figure 9 Damage identification results for four DSs using frequencies and mode shapes

8. Conclusions and Discussions

An SBL method has been proposed in this study for probabilistic structural damage detection using modal parameters. The sparsity of structural damage has been exploited as important prior information from the Bayesian perspective. The integral of the nonlinear eigenvalue equation is avoided through an iterative procedure based on the EM algorithm. Moreover, two sampling methods have been proposed and compared to approximate the evidence. Without employing a complicated Bayesian model with asymptotic approximation or utilizing stochastic simulation, this proposed method is concise and easy to implement.

Numerical and experimental examples demonstrate that the proposed method is effective in identifying single and multiple damage elements, even when the measurement data are much fewer than the damage parameters. The damage identification results also show that the likelihood sampling is more robust to noise and accurate than the posterior sampling. The convergence using the EM for damage detection is very rapid. Moreover, the damage identification accuracy has been improved by incorporating mode shape data, especially for multiple damage scenarios.

Compared with the regularization-based sparse recovery theory, the EM-based SBL is more general as more hyper-parameters are used and own clear physical meaning. Moreover, the hyper-parameters can be updated automatically, thereby avoiding the selection of the regularization parameter in the sparse recovery.

Acknowledgements

This research was supported by the PolyU Research Grant (Project No. 1-ZVFH) and the National Natural Science Foundation of China (Project No. 51678364).

References

1. Doebling SW, Farrar CR, Prime MB, Shevitz DW. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review, *Los Alamos National Laboratory Report*, 1996.
2. Sohn H, Farrar CR, Hemez FM, Shunk DD, Stinemates DW, Nadler BR, Czarnecki JJ. A review of structural health monitoring literature: 1996–2001, *Los Alamos National Laboratory Report*, 2003.
3. Fan W, Qiao P. Vibration-based damage identification methods: a review and comparative study. *Structural Health Monitoring* 2011; **10**(1):83–111.
4. Salawu OS. Detection of structural damage through changes in frequency: a review. *Engineering Structures* 1997; **19**(9):718–723.
5. Shi ZY, Law SS, Zhang LM. Damage localization by directly using incomplete mode shapes. *Journal of Engineering Mechanics* 2000; **126**(6):656–660.
6. Xu YL, Zhang J, Li JC, Xia Y. Experimental Investigation on Statistical Moment-based Structural Damage Detection Method. *Structural Health Monitoring* 2009; **8**(6):555–571.
7. Hemez FM. *Uncertainty quantification and the verification and validation of computational models.* In Damage Prognosis: For Aerospace, Civil and Mechanical Systems, Inman DJ (ed.). John, Wiley and Sons: West Sussex, England, 2005.
8. Xia Y, Hao H, Brownjohn JMW, Xia PQ. Damage identification of structures with uncertain frequency and mode shape data. *Earthquake Engineering and Structural Dynamics* 2002; **31**(5):1053–1066.
9. Xia Y, Hao H. Statistical damage identification of structures with frequency changes. *Journal of Sound and Vibration* 2003; **263**(4):853–870.
10. Sohn H, Dzwonczyk M, Straser E, Kiremidjian A, Law K, Meng T. An experimental study of temperature effect on modal parameters of the alamosa canyon bridge. *Earthquake Engineering and Structural Dynamics* 1999; **28**:878–897.
11. Xia Y, Hao H, Zanardo G, Deeks AJ. Long Term Vibration Monitoring of an RC Slab: Temperature and Humidity Effect. *Engineering Structures* 2006; **28**(3):441–452.
12. Liu PL. Identification and damage detection of trusses using modal data. *Journal of Structural Engineering* 1995; **121**(4):599–608.
13. Hua XG, Ni YQ, Chen ZQ, Ko JM. An improved perturbation method for stochastic finite element model updating. *International Journal for Numerical Method in Engineering* 2008; **73**(3):1845–1864.
14. Sohn H, Law KH. A Bayesian probabilistic approach for structure damage detection. *Earthquake Enineering and Structural Dynamics* 1997; **26**(12):1259–81.
15. Beck JL, Au SK, Vanik MW. Monitoring structural health using a probabilistic measure. *Computer-Aided Civil and Infrastructure Engineering* 2001; **16**(1):1–11.
16. Ching J, Beck JL. Bayesian analysis of the phase II IASC–ASCE structural health monitoring experimental benchmark data. J*ournal of Engineering Mechanics* 2004; **130**(10):1233–44.
17. Mustafa S, Debnath N, Dutta A. Bayesian probabilistic approach for model updating and damage detection for a large truss bridge. *International Journal of Steel Structures* 2015; **15**(2):473–485.
18. Hou RR, Xia Y, Zhou XQ. Structural damage detection based on l_1 regularization using natural frequencies and mode shapes. *Structural Control and Health Monitoring* 2018; **25**(3),e2107.
19. Zhou XQ, Xia Y, Weng S. L_1 regularization approach to structural damage detection using frequency data. *Structural Health Monitoring* 2015; **14**(6):571–582.

20. Bao Y, Li H, Chen Z, Zhang F, Guo A. Sparse l_1 optimization-based identification approach for the distribution of moving heavy vehicle loads on cable-stayed bridges. *Structural Control and Health Monitoring* 2016; **23**(1):144–155.
21. Zhang CD, Xu YL. Comparative studies on damage identification with Tikhonov regularization and sparse regularization. *Structural Control and Health Monitoring* 2016; **23**(3):560–579.
22. Yang Y, Nagarajaiah S. Structural damage identification via a combination of blind feature extraction and sparse representation classification. *Mechanical Systems and Signal Processing* 2014; **45**(1):1–23.
23. Yang Y, Nagarajaiah S. Output-only modal identification with limited sensors using sparse component analysis. *Journal of Sound and Vibration* 2013; **332**(19):4741–4765.
24. Chen ZP, Yu L. Application of sparsity in structural damage detection. *Proceedings of the 8th International Conference on Structural Health Monitoring of Intelligent Infrastructure*, Brisbane, Australia, 2017; 1297–1307.
25. Williams PM. Bayesian regularization and pruning using a Laplace prior. *Neural Computation* 1995; **7**(1):117–143
26. Tipping ME. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 2001; 1:211–244.
27. Wipf D, Palmer J, Rao B. Perspectives on sparse Bayesian learning. *Advances in Neural Information Processing Systems (NIPS 16)*, Whistler, Canada, December 2003.
28. Bishop CM. *Pattern Recognition and Machine Learning*. Berlin: Springer, 2006.
29. Beck JL, Katafygiotis LS, Updating models and their uncertainties. I: Bayesian statistical framework, *Journal of Engineering Mechanics* 1998; **124**(4):455–461.
30. Beck JL. Bayesian system identification based on probability logic. *Structural Control and Health Monitoring* 2010; 17:825–847.
31. Onar A. Laplace Approximations in Bayesian Lifetime Analysis. Wiley StatsRef: *Statistics Reference Online*, 2014.
32. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 1953; 21:1087–1092.
33. Ching JY, Chen YC. Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection and model averaging. *Journal of Engineering Mechanics* 2007; 133:816–832.
34. Cheung SH, Beck JL. Calculation of the posterior probability for Bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Computer-Aided Civil and Infrastructure Engineering* 2010; 25:304–321.
35. Lam HF, Yang JH, Au SK. Markov chain Monte Carlo-based Bayesian method for structural model updating and damage detection. *Structural Control Health Monitoring* 2018; **25**:e2140.
36. 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. *Journal of Sound and Vibration* 2010; **329**(15):2995–3007.
37. Huang Y, Beck JL. Hierarchical sparse Bayesian learning for structural health monitoring with incomplete modal data. *International Journal for Uncertainty Quantification* 2015; **5**(2):139–69.
38. Huang Y, Beck JL, Li H. Hierarchical sparse Bayesian learning for structural damage detection: Theory, computation and application. *Structural Safety* 2017; **64**:37–53.
39. Huang Y, Beck JL, Li H. Bayesian system identification based on hierarchical sparse Bayesian learning and Gibbs sampling with application to structural damage assessment. *Computer*

Methods in Applied Mechanics and Engineering 2017; **318**:382–411.
40. Huang Y, Shao C, Wu B, Beck JL, Li H. State-of-the-art review on Bayesian inference in structural system identification and damage assessment. *Advances in Structural Engineering,* 1369433218811540.
41. Xia Y, Hao H, Deeks AJ, Zhu X. Condition assessment of shear connectors in slab-girder bridges via vibration measurements. *Journal of Bridge Engineering* 2008; **13**(1):43–54.
42. Vanik MW, Beck JL, Au SK. Bayesian probabilistic approach to structural health monitoring. *Journal of Engineering Mechanics* 2000; **126**(7):738–745.
43. Jaynes ET. *Probability Theory: The Logic of Science*. Cambridge, UK: Cambridge University Press: 2003.
44. Mackay DJC. *Bayesian methods for adaptive models*. Ph.D. Thesis in Computation and Neural Systems. Pasadena, California: California Institute of Technology; 1992.
45. Huang Y, Li H, Wu S, Yang YC. Fractal dimension based damage identification incorporating multi-task sparse Bayesian learning. *Smart Materials and Structures* 2018; 27:075020.
46. Babacan S, Molina R, Katsaggelos A. Bayesian compressive sensing using Laplace priors. *IEEE Transactions on Image Processing* 2010; **19**(1):53–63.
47. Dempster AP, Laird N, Rubin D. Maximum likelihood for incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 1977; **39**:1–38.
48. Couvreur C. The EM Algorithm: A Guided Tour. In: *Proceedings of the second IEEE European workshop on computer-intensive methods in control and signal processing*, Pragues, Czech Republik, August 1996; 115–120.
49. Daubechies I, DeVore R, Fornasier M, GÜNTÜRK CS. Iteratively reweighted least squares minimization for sparse recovery. *Communications on Pure and Applied Mathematics* 2010; **63**(1):1–38.
50. Chartrand R, Yin W. Iteratively reweighted algorithms for compressive sensing. In: *Proceedings of the 33nd IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, 2008; 3869–3872.
51. Chartrand R, Staneva V. Restricted isometry properties and nonconvex compressive sensing. *Inverse Problems* 2008; **24**:035020.
52. Adewuyi AP, Wu Z, Serker NHMK. Assessment of vibration-based damage identification methods using displacement and distributed strain measurements. *Structural Health Monitoring* 2009; **8**(6):443–461.
53. Mottershead J, Friswell M. Model updating in structural dynamics: a survey, *Journal of Sound and Vibration* 1993; **167**(2):347–375.
54. Hou RR, Xia Y, Bao YQ, Zhou XQ. Selection of regularization parameter for l_1-regularized damage detection. *Journal of Sound and Vibration* 2018; 423:141–160.