Model selection for dynamic reduction-based structural health monitoring following the Bayesian evidence approach

T. Yin [1*], H. P. Zhu [2] and S. J. Fu [1,3]

[1] School of Civil Engineering, Wuhan University, Wuhan 430072, P.R. China

[2] School of Civil Engineering & Mechanics, Huazhong University of Science and Technology, Wuhan 430074, P.R. China

[3] School of Civil Engineering, Xijing University, Xi'an 710123, P.R. China

Abstract: There usually exist multitudinous finite element (FE) models with varying level of complexity which can be developed from the engineering judgment for the purpose of structural system identification and health monitoring. By following the theory of Bayesian evidence statistic, this paper proposes a methodology to investigate the issues of FE model-class selection for choosing suitable parameterized structural models utilized in dynamic reduction-based structural health monitoring (SHM). By employing the concept of information divergence, the amount of information needed to be extracted from the measured data is explicitly quantified during the procedure of FE model updating-based structural health monitoring. Then, for achieving a trade-off between the complexity of a parameterized FE model class and that of its corresponding information-theoretic interpretation, such information is utilized for penalizing the complexity of model class to ensure that a relatively simple parameterization scheme can be achieved. The proposed methodology consists of calibration and subsequent monitoring stages, and the information obtained in the former stage is utilized as pseudo-data which is learned by the latter stage to improve the model parameter estimation by implementating the delayed rejection adaptive Metropolis algorithm. Through numerical case studies conducted for a four-storey two-bay steel frame structure considering semi-rigid connections as well as laboratory experiment performed for a two-storey bolt-connected steel frame model, the feasibility and validity of proposed methodology is demonstrated.

Keywords: Model-class selection; structural health monitoring; finite element model reduction; Bayesian evidence inference; information theory.

[*] Corresponding author: Dr. Tao YIN, Associate Professor, School of Civil Engineering, Wuhan University, Wuhan, 430072, P.R. China. E-mail: tyin@whu.edu.cn

1. INTRODUCTION

The development of a methodology for the accurate and reliable assessment of structural damage is very essential to ensure the integrity and stability of structures, reduce the cost of maintenance and prevent catastrophic failure. There has been great effort in developing structural health monitoring (SHM) methodologies utilizing vibration measurements over the last few decades [1,2]. The majority of the methods in the literature have been validated by various types of structural components or systems, such as truss-type structures [3-6], beam-type structures [7-11], plate-type structures [12-15], frame structures [16-21], shear building structures [22-25], bridge structures [26-28], and periodic structural systems [29-32]. Most of these mentioned research works are based on the framework of finite element (FE) model updating by utilizing dynamic characteristics [33,34], where an objective function is usually defined in terms of the discrepancies between the experimental measurements and those calculated from a FE model, and it is then minimized for the examination of the stiffness parameter changes to indicate the structural damage.

For the FE model updating-based SHM procedure mentioned above, the accuracy of FE model is essential for its successful implementation. However, due to assumption and uncertainty arisen from the theoretical hypothesis, boundary condition, and geometric and material properties, there is an unavoidable mismatch between the model-predicted and measured dynamic properties. In order to obtain higher accuracy for structural analysis, the FE model of the target structure tends to be fine enough to approximate the structural details, leading to the increase of the model complexity. This, however, is not beneficial for the efficient implementation of FE model updating-based SHM, which is typically an inverse problem of structural dynamics. Since the repeated solution of the large eigenvalue problem is usually required in the process of FE model updating, the computational cost of this procedure eventually resulted unaffordable when dealing with complex models with large amount of degrees of freedom (DOFs). In addition, the measured modal data is incomplete due to the problems of limited number of sensors, measurement noise and truncation error of higher order modes, etc., which cause the inability to capture the full dynamics of the structure, rendering the inverse problem of FE model updating procedure uncertain and ill-posed [35,36]. Moreover, for forming the above-mentioned measure-of-fit objective function by utilizing modal parameters, it is necessary to ensure that the measured modes are well matched with the calculated ones by employing the modal assurance criterion (MAC) technique, which is especially difficult in the situation of a complex structural model together with a limited number of measured DOFs. Such mode matching becomes even more difficult by noting the

fact that the present of damage might sometimes cause changes in the order of the modes.

In order to resolve the difficulties arisen from the limited number of measurement channels, the methods usually start from computing the missing components of the mode shapes through mode shape expansion procedure [37]. However, it has been revealed that this would aggregate the modeling error, experimental noise and other sources of uncertainties into the resultant mode shapes [38,39], affecting substantially the damage detection results. In such circumstances, the FE model reduction method originally developed for the purpose of reducing the computation effort for large-scale structural models [40-43], particularly for the dynamic-reduction method [43], becomes a more practical alternative since it does not introduce any error in the transformation process within a certain frequency range [20,44,45], and also avoids the mode matching problem for using modal data. By employing the FE model reduction technique, the original full system matrices are condensed to the reduced system matrices corresponding to only the measured DOFs, so the matrix dimension of the inverse problem for damage identification could be significantly reduced, especially efficient for large-scale structural models with a huge number of DOFs. Nevertheless, there usually exist multitudinous FE models with varying level of complexity that can be developed from the engineering judgment. The inverse problem may be non-uniquely solvable for FE models with higher parameterization complexity due to the large number of uncertain parameters to be identified as compared to the limited measurement information available. Thus, it is particularly important to choose the FE model with the suitable complexity for the purpose of the FE model updating-based SHM.

In this paper, a methodology is proposed to investigate the problem of choosing suitable class of parameterized FE models for the dynamic reduction-based SHM following the Bayesian evidence inference method. Within the concept of information theory, the amount of information needed to be extracted from the measured data for the prescribed set of parameterized model classes is explicitly quantified during the dynamic model reduction-based SHM procedure. Then, this information measure is utilized for penalizing the parameterization complexity of a given FE model to ensure a FE parameterization scheme with suitable complexity. There are two stages involved in the proposed methodology. In the calibration stage, Bayesian approach is employed for updating the initial FE model with modal parameters from the intact structure to obtain the most probable values and associated uncertainties of model parameters by implementing the delayed rejection adaptive Metropolis (DRAM) algorithm. In the subsequent monitoring stage, the information achieved from the previous stage is utilized as pseudo-data in order to further improve the model parameter estimation.

The proposed methodology is validated through both the numerical case studies and laboratory experiment conducted for steel frames with semi-rigid connections.

2. THE PROPOSED METHODOLOGY

2.1 Dynamic reduction-based damage identification

The eigen-system equations of a N_d-DOF FE model can be written as

$$\mathbf{K}(\boldsymbol{\theta})\boldsymbol{\phi}_n = \lambda_n \mathbf{M}\boldsymbol{\phi}_n, \quad \mathbf{K}(\boldsymbol{\theta}) = \mathbf{K} - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}^{(i)} \tag{1}$$

where the global stiffness matrix $\mathbf{K}(\boldsymbol{\theta})$ is parameterized by the non-dimensional stiffness scaling parameter vector $\boldsymbol{\theta} = \{\theta_1, \theta_2, \cdots, \theta_{N_\theta}\}^{\mathrm{T}} \in \mathbb{R}^{N_\theta \times 1}$ describing reduction of stiffness for the potentially damaged structural elements or substructures. It allows the nominal stiffness matrix given by $\mathbf{K}(\boldsymbol{\theta}_0)$ to be updated based on the identified modal parameters through the model updating procedure. N_θ is the number of uncertain scaling parameters to be updated. $\mathbf{K}^{(i)}$, for $i = 1,2,\cdots,N_\theta$, denotes the contribution of the ith element or substructure to the global stiffness matrix of the FE model, and θ_i is the ith element in the uncertain parameter vector $\boldsymbol{\theta}$. $\mathbf{K}$, $\mathbf{M} \in \mathbb{R}^{N_d \times N_d}$ are constant matrices independent of $\boldsymbol{\theta}$. $\lambda_n = (2\pi f_n)^2$, and f_n is the nth natural frequency, for $n = 1,2,\cdots,N_t$. N_t is the number of measured modes. λ_n and $\boldsymbol{\phi}_n \in \mathbb{R}^{N_d \times 1}$ are the nth eigenvalue and eigenvector, respectively. In this study, damage is due to the reduction of rotational stiffness of semi-rigid connections, and it is assumed that there is no change in the mass distribution of the structural system before and after the occurrence of damage.

According to the master (or measured) set of DOF (index m) and the slave (or unmeasured) set of DOF (index s) of the FE model, Eq. (1) can be partitioned into the master and slave sets of DOFs, respectively as

$$\begin{bmatrix} \mathbf{K}_{mm} - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}_{mm}^{(i)} & \mathbf{K}_{ms} - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}_{ms}^{(i)} \\ \mathbf{K}_{sm} - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}_{sm}^{(i)} & \mathbf{K}_{ss} - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}_{ss}^{(i)} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\phi}_{n,m} \\ \boldsymbol{\phi}_{n,s} \end{Bmatrix} = \lambda_n \begin{bmatrix} \mathbf{M}_{mm} & \mathbf{M}_{ms} \\ \mathbf{M}_{sm} & \mathbf{M}_{ss} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\phi}_{n,m} \\ \boldsymbol{\phi}_{n,s} \end{Bmatrix} \tag{2}$$

where $\boldsymbol{\phi}_{n,m} \in \mathbb{R}^{N_m \times 1}$ and $\boldsymbol{\phi}_{n,s} \in \mathbb{R}^{N_s \times 1}$ are the measured and unmeasured parts of full mode shape vector $\boldsymbol{\phi}_n$, respectively, $N_m + N_s = N_d$. From the second set of equations in Eq. (2), one obtains that

$$\boldsymbol{\phi}_{n,s} = \mathbf{D}_n \boldsymbol{\phi}_{n,m} \tag{3}$$

where $\mathbf{D}_n \in \mathbb{R}^{N_s \times N_m}$ depending on both eigenvalue λ_n and parameter vector $\boldsymbol{\theta}$, is given by

$$\mathbf{D}_n = \left(\mathbf{K}_{ss} - \lambda_n \mathbf{M}_{ss} - \sum_{i=1}^{N_\theta} \theta_i \, \mathbf{K}_{ss}^{(i)} \right)^{-1} \left(\mathbf{K}_{sm} - \lambda_n \mathbf{M}_{sm} - \sum_{i=1}^{N_\theta} \theta_i \, \mathbf{K}_{sm}^{(i)} \right) \tag{4}$$

Based on Eq. (3), the full mode shape vector $\boldsymbol{\phi}_n$ can be represented by

$$\boldsymbol{\phi}_n = \begin{Bmatrix} \boldsymbol{\phi}_{n,m} \\ \boldsymbol{\phi}_{n,s} \end{Bmatrix} = \mathbf{T}_n \boldsymbol{\phi}_{n,m} = \begin{bmatrix} \mathbf{I}_{N_m} \\ \mathbf{D}_n \end{bmatrix} \boldsymbol{\phi}_{n,m} \tag{5}$$

where $\mathbf{T}_n \in \mathbb{R}^{N_d \times N_m}$ expanding the measured partial mode shape vector to the full model shape vector of the FE model is the system transformation matrix for the nth mode. $\mathbf{I}_{N_m}$ is the identity matrix with dimension $N_m \times N_m$.

Substituting Eq. (5) into Eq. (2), and pre-multiplying the system transformation matrix $\mathbf{T}_n$ to both sides of the resultant equations leads to the following eigen-system equations for the reduced system model corresponding to the N_m master DOFs, i.e.,

$$\left(\mathbf{K}_n^R - \sum_{i=1}^{N_\theta} \theta_i \mathbf{K}_n^{(i)R} \right) \boldsymbol{\phi}_{n,m} = \lambda_n \mathbf{M}_n^R \boldsymbol{\phi}_{n,m} \tag{6}$$

where $\mathbf{K}_n^R,\ \mathbf{M}_n^R \in \mathbb{R}^{N_m \times N_m}$ are the reduced system stiffness and mass matrices with respect to the nth mode, respectively. $\mathbf{K}_n^{(i)R} \in \mathbb{R}^{N_m \times N_m}$, related to the nth mode, is reduced matrix of $\mathbf{K}^{(i)}$. These three matrices are given by

$$\mathbf{K}_n^R = \mathbf{T}_n^{\mathrm{T}} \overline{\mathbf{K}} \mathbf{T}_n, \quad \mathbf{M}_n^R = \mathbf{T}_n^{\mathrm{T}} \overline{\mathbf{M}} \mathbf{T}_n, \quad \mathbf{K}_n^{(i)R} = \mathbf{T}_n^{\mathrm{T}} \overline{\mathbf{K}}^{(i)} \mathbf{T}_n \tag{7}$$

and

$$\overline{\mathbf{K}} = \begin{bmatrix} \mathbf{K}_{mm} & \mathbf{K}_{ms} \\ \mathbf{K}_{sm} & \mathbf{K}_{ss} \end{bmatrix}, \quad \overline{\mathbf{M}} = \begin{bmatrix} \mathbf{M}_{mm} & \mathbf{M}_{ms} \\ \mathbf{M}_{sm} & \mathbf{M}_{ss} \end{bmatrix}, \quad \overline{\mathbf{K}}^{(i)} = \begin{bmatrix} \mathbf{K}_{mm}^{(i)} & \mathbf{K}_{ms}^{(i)} \\ \mathbf{K}_{sm}^{(i)} & \mathbf{K}_{ss}^{(i)} \end{bmatrix} \tag{8}$$

Thus, by gathering all the measured modes, Eq. (6) can be expressed in a matrix form as

$$\mathbf{H}(\boldsymbol{\theta})\boldsymbol{\theta} = \mathbf{h}(\boldsymbol{\theta}) \tag{9}$$

where both the matrix $\mathbf{H} \in \mathbb{R}^{(N_m N_t) \times N_\theta}$ and vector $\mathbf{h} \in \mathbb{R}^{(N_m N_t) \times 1}$ are the implicit functions

of the parameter vector $\boldsymbol{\theta}$ and they are given below as

$$\mathbf{H}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_{N_t} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_1^{(1)R}\boldsymbol{\phi}_{1,m} & \mathbf{K}_1^{(2)R}\boldsymbol{\phi}_{1,m} & \cdots & \mathbf{K}_1^{(N_\theta)R}\boldsymbol{\phi}_{1,m} \\ \mathbf{K}_2^{(1)R}\boldsymbol{\phi}_{2,m} & \mathbf{K}_2^{(2)R}\boldsymbol{\phi}_{2,m} & \cdots & \mathbf{K}_2^{(N_\theta)R}\boldsymbol{\phi}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}_{N_t}^{(1)R}\boldsymbol{\phi}_{N_t,m} & \mathbf{K}_{N_t}^{(2)R}\boldsymbol{\phi}_{N_t,m} & \cdots & \mathbf{K}_{N_t}^{(N_\theta)R}\boldsymbol{\phi}_{N_t,m} \end{bmatrix} \tag{10}$$

and

$$\mathbf{h}(\boldsymbol{\theta}) = \begin{Bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_{N_t} \end{Bmatrix} = \begin{Bmatrix} (\mathbf{K}_1^R - \lambda_1\mathbf{M}_1^R)\boldsymbol{\phi}_{1,m} \\ (\mathbf{K}_2^R - \lambda_2\mathbf{M}_2^R)\boldsymbol{\phi}_{2,m} \\ \vdots \\ (\mathbf{K}_{N_t}^R - \lambda_{N_t}\mathbf{M}_{N_t}^R)\boldsymbol{\phi}_{N_t,m} \end{Bmatrix} \tag{11}$$

where $\mathbf{H}_n \in \mathbb{R}^{N_m \times N_\theta}$ and $\mathbf{h}_n \in \mathbb{R}^{N_m \times 1}$ for $n = 1,2,\cdots,N_t$ are the nth submatrix and subvector of the matrix $\mathbf{H}$ and vector $\mathbf{h}$, respectively. By implementing Eq. (9) with the measured modal parameters before and after the present of damage through the model updating procedure, the health status of the structure being monitored can be quantified by the change of unknown parameter vector $\boldsymbol{\theta}$. It is noted that there is still a question yet to be solved, i.e., the class of parameterized models of the reduced structural system should be specified before carrying out the monitoring procedure, which will be addressed in a statistic manner in the following part of the paper.

2.2 Bayesian evidence statistic

In principle, the proposed methodology is intended to efficiently monitor the service status of large-scale structures with the FE parameterization scheme of suitable complexity from a probabilistic point of view. Provided the sets of measured dynamic data $\mathbb{D}_N$, the goal of model class selection is to use $\mathbb{D}_N$ to select the most probable class of models representing the system out of N_M given classes of parameterized models $\mathbb{M} = \{\mathbb{M}_1, \mathbb{M}_2, \cdots, \mathbb{M}_{N_M}\}$, where $\mathbb{M}_j$ specifies not only a class of deterministic dynamic models derived from Eq. (9), but also the probability descriptions for the prediction error variance. It is noted that the parameter vector $\boldsymbol{\theta}_j \in \mathbb{R}^{N_j \times 1}$ depends on the model class $\mathbb{M}_j$ even though it is not explicitly reflected in the symbol for simplicity, and N_j denotes the number of uncertain stiffness scaling parameters to be identified with respect to the model class $\mathbb{M}_j$. Let $\mathbb{D}_N$ denote the modal parameters measured from the target structural system through modal testing, i.e.,

$$\mathbb{D}_N = \bigcup_{s=1}^{N} \left\{ \left(\hat{f}_n^{(s)}, \widehat{\boldsymbol{\phi}}_{n,m}^{(s)} \right), \ n = 1, \cdots, N_t \right\} \tag{12}$$

where N is the number of repeatedly measured modal parameter pairs, which can be identified from recorded time-domain responses with sufficient duration by utilizing suitable modal parameter identification techniques [46,47]. The superscript ^ denotes the measured quantities.

By utilizing the set of measured data $\mathbb{D}_N$ and Eq. (9), the eigen-equation errors $\mathbf{e}_{n,j} \in \mathbb{R}^{N_m \times 1}$ of the reduced structural system for the nth mode ($n = 1, \cdots, N_t$) can be expressed as

$$\mathbf{H}_n(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j)\boldsymbol{\theta}_j - \mathbf{h}_n(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j) = \mathbf{e}_{n,j} \tag{13}$$

where the covariance matrix $\boldsymbol{\Sigma}_{n,j} \in \mathbb{R}^{N_m \times N_m}$ of $\mathbf{e}_{n,j}$ controls the size of the eigen-equation errors of the nth mode with an assumption of Gaussian probability model. In addition, it is also assumed that the uncertainty in the equation errors of the reduced structural system for the nth mode are modeled as independent and identically distributed, so the covariance matrix of $\mathbf{e}_{n,j}$ can be given by

$$\boldsymbol{\Sigma}_{n,j} = \sigma_j^2 \mathbf{I}_{N_m} \tag{14}$$

where $\mathbf{I}_{N_m}$ is the N_m-dimensional identity matrix, and σ_j^2 denotes the variance of eigen-equation errors with respect to the jth class of models.

By providing the dynamic data set $\mathbb{D}_N$ and specifying the model class $\mathbb{M}_j$, the likelihood function $p(\mathbb{D}_N|\boldsymbol{\theta}_j, \mathbb{M}_j)$ is defined by

$$p(\mathbb{D}_N|\boldsymbol{\theta}_j, \mathbb{M}_j) = \left(2\pi\sigma_j^2\right)^{-\frac{N_m N_t}{2}} \exp\left[-\frac{N_m N_t}{2\sigma_j^2} \mathcal{J}(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j) \right] \tag{15}$$

where $\mathcal{J}(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j)$ is the measure-of-fit function quantifying the eigen-equation errors of the reduced structural system, i.e.,

$$\mathcal{J}(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j) = \frac{1}{N_m N_t} \sum_{n=1}^{N_t} \left\| \mathbf{H}_n(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j)\boldsymbol{\theta}_j - \mathbf{h}_n(\boldsymbol{\theta}_j; \mathbb{D}_N, \mathbb{M}_j) \right\|^2 \tag{16}$$

It is noted that, the proposed health monitoring methodology consists of two stages, i.e., the

calibration stage and the subsequent monitoring stage, denoted by Stage-I and Stage-II, respectively. In the first stage, based on the identified modal parameters of the intact structure, Bayesian approach is employed for updating the initial FE model to find the MAP (maximum a posteriori) estimate of stiffness scaling parameter vector $\widehat{\boldsymbol{\theta}}_j^c$ and the associated diagonal covariance matrix $\widehat{\boldsymbol{\Sigma}}_j^c = \mathrm{diag}[(\widehat{\boldsymbol{\sigma}}_j^c)^2] \in \mathbb{R}^{N_j \times N_j}$. Where $\widehat{\boldsymbol{\sigma}}_j^c = \{\hat{\sigma}_{j,1}^c, \hat{\sigma}_{j,2}^c, \cdots, \hat{\sigma}_{j,N_j}^c\}^{\mathrm{T}} \in \mathbb{R}^{N_j \times 1}$ is the identified standard deviation of all parameters, and $\hat{\sigma}_{j,i}^c$ is the ith component of $\widehat{\boldsymbol{\sigma}}_j^c$. It is further assumed that the MAP estimate $\widehat{\boldsymbol{\theta}}_j^c$ obtained from the first stage is unique with a large amount of dynamic measurements available for simplicity. In the subsequent stage, the values of stiffness scaling parameters $\boldsymbol{\theta}_j$ are continuously monitored by utilizing the measured dynamic data $\mathbb{D}_N$ to update the previously obtained refined FE model. In the proposed methodology, motivated by the concept of automatic relevance determination (ARD) [48-50] and also utilizing the MAP estimate $\widehat{\boldsymbol{\theta}}_j^c$ achieved in the foregoing calibration stage, one can get that

$$p\left(\Delta\boldsymbol{\theta}_j|\boldsymbol{\alpha}_j, \mathbb{M}_j\right) = \mathbb{N}\left(\Delta\boldsymbol{\theta}_j|\mathbf{0}, \mathbf{A}_j\right) \tag{17}$$

where $\Delta\boldsymbol{\theta}_j = \boldsymbol{\theta}_j - \widehat{\boldsymbol{\theta}}_j^c$ represents the change of stiffness scaling parameters, and $\mathbf{A}_j = \mathrm{diag}(\boldsymbol{\alpha}_j) \in \mathbb{R}^{N_j \times N_j}$ is a diagonal covariance matrix defined by $\boldsymbol{\alpha}_j = \{\alpha_{j,1}, \alpha_{j,2}, \cdots, \alpha_{j,N_j}\}^{\mathrm{T}} \in \mathbb{R}^{N_j \times 1}$, which is the hyper-parameter vector for the jth class of models $\mathbb{M}_j$ and represents the prediction-error variance for the stiffness change vector $\Delta\boldsymbol{\theta}_j$.

A likelihood function for $\boldsymbol{\theta}_j$ can be further defined to exploit the information that $\Delta\boldsymbol{\theta}_j$ should be reasonably a sparse vector with most of its components being zero during the early stage of damage development. By defining the pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c = \widehat{\boldsymbol{\theta}}_j^c$, this likelihood function can be defined as

$$p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j, \boldsymbol{\alpha}_j, \mathbb{M}_j\right) = (2\pi)^{-N_j/2}\det\left(\mathbf{A}_j\right)^{-1/2}\exp\left[-\frac{1}{2}\mathcal{G}\left(\boldsymbol{\theta}_j; \widehat{\boldsymbol{\vartheta}}_j^c, \boldsymbol{\alpha}_j, \mathbb{M}_j\right)\right] \tag{18}$$

with

$$\mathcal{G}\left(\boldsymbol{\theta}_j; \widehat{\boldsymbol{\vartheta}}_j^c, \boldsymbol{\alpha}_j, \mathbb{M}_j\right) = \left(\boldsymbol{\theta}_j - \widehat{\boldsymbol{\theta}}_j^c\right)^{\mathrm{T}}\mathbf{A}_j^{-1}\left(\boldsymbol{\theta}_j - \widehat{\boldsymbol{\theta}}_j^c\right) \tag{19}$$

It is noted that the above choice of likelihood function based on the pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c$ would lead to a sparse representation of the parameter change vector $\Delta\boldsymbol{\theta}_j$ during the optimization of the hyper-parameter vector $\boldsymbol{\alpha}_j$. This choice is motivated by the sparse Bayesian learning framework [48] which is known to provide an effective tool for pruning large numbers of

irrelevant or redundant features in a linear regression model that are not supported by the data.

By using the Bayes' theorem [51], the pseudo-posterior PDF of the model parameter vector $\boldsymbol{\theta}_j$ conditional on the pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c$ can be given as

$$p\left(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)=c_0 p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)p\left(\boldsymbol{\theta}_j|\mathbb{M}_j\right) \tag{20}$$

where c_0 is a normalizing constant for ensuring that the integration of the PDF over predefined domain is equal to unity. $p\left(\boldsymbol{\theta}_j|\mathbb{M}_j\right)$ is the prior PDF specified by the user, and it is chosen as a non-information distribution within the predefined parameter domain here. It should be pointed out that the pseudo-posterior PDF of scaling parameter $\boldsymbol{\theta}_j$ given in Eq. (20) is treated as a prior PDF for these uncertain parameters within the Bayesian statistic framework in the later part of the paper.

By employing the Bayes' theorem, the posterior PDF of unknown model parameters conditional on the observed quantities can be given by

$$p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)=c_1 p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right) \tag{21}$$

where $c_1=p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)^{-1}$ is also a normalizing constant to ensure the integration of the PDF over predefined domain to be unity, and $p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)$, denoting the evidence for model class $\mathbb{M}_j$ provided by both the measured data $\mathbb{D}_N$ and pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c$, can be obtained by following the theorem of total probability.

By substituting Eq. (20) into Eq. (21), the posterior PDF of model parameter $\boldsymbol{\theta}_j$ for the jth model class $\mathbb{M}_j$ can be further written by

$$p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\propto p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)p\left(\boldsymbol{\theta}_j|\mathbb{M}_j\right) \tag{22}$$

It is noted that the most probable values of uncertain model parameters $\boldsymbol{\theta}_j$ can be estimated by maximizing the posterior PDF $p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)$ given in Eq. (22), which is equivalent to maximizing the product of two individual likelihood functions, i.e., $p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)$ and $p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)$. In practice, the most probable values of unknown parameters can be conveniently obtained by maximizing the logarithm of the product of these two likelihoods, i.e.,

$$\begin{aligned}\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\right] &= \ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)\right]+\ln\left[p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\right]\\ &= -\frac{N_m N_t}{2}\ln(2\pi)-N_m N_t\ln\left(\sigma_j\right)-\frac{N_m N_t}{2\sigma_j^2}\mathcal{J}\left(\boldsymbol{\theta}_j;\mathbb{D}_N,\mathbb{M}_j\right)-\frac{N_j}{2}\ln(2\pi)\\ &\quad -\frac{1}{2}\ln\left[\det\left(\mathbf{A}_j\right)\right]-\frac{1}{2}\mathcal{G}\left(\boldsymbol{\theta}_j;\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\end{aligned} \tag{23}$$

For a given parameter vector $\boldsymbol{\theta}_j$, maximization $\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\right]$ in Eq. (23) with respect to the ith hyper-parameter $\alpha_{j,i}$, for $i=1,\cdots,N_j$, requires that

$$\frac{\partial\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\right]}{\partial\alpha_{j,i}}=\frac{\partial\ln\left[p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)\right]}{\partial\alpha_{j,i}}=0 \tag{24}$$

It follows:

$$\frac{\partial\ln\left[\det\left(\mathbf{A}_j\right)\right]}{\partial\alpha_{j,i}}+\frac{\partial\mathcal{G}\left(\boldsymbol{\theta}_j;\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)}{\partial\alpha_{j,i}}=0 \tag{25}$$

where

$$\frac{\partial\ln\left[\det\left(\mathbf{A}_j\right)\right]}{\partial\alpha_{j,i}}=\frac{1}{\det\left(\mathbf{A}_j\right)}\frac{\partial\det\left(\mathbf{A}_j\right)}{\partial\alpha_{j,i}}=\mathrm{tr}\left(\mathbf{A}_j^{-1}\frac{\partial\mathbf{A}_j}{\partial\alpha_{j,i}}\right)=\alpha_{j,i}^{-1} \tag{26}$$

and

$$\frac{\partial\mathcal{G}\left(\boldsymbol{\theta}_j;\widehat{\boldsymbol{\vartheta}}_j^c,\boldsymbol{\alpha}_j,\mathbb{M}_j\right)}{\partial\alpha_{j,i}}=\frac{\partial\left(\boldsymbol{\theta}_j-\widehat{\boldsymbol{\theta}}_j^c\right)^{\mathrm{T}}\mathbf{A}_j^{-1}\left(\boldsymbol{\theta}_j-\widehat{\boldsymbol{\theta}}_j^c\right)}{\partial\alpha_{j,i}}=-\alpha_{j,i}^{-2}\left(\theta_{j,i}-\hat{\theta}_{j,i}^c\right)^2 \tag{27}$$

By substituting Eqs. (26) and (27) into Eq. (25), one can obtain for $i=1,\cdots,N_j$ that

$$\hat{\alpha}_{j,i}\left(\theta_{j,i}\right)=\left(\theta_{j,i}-\hat{\theta}_{j,i}^c\right)^2 \tag{28}$$

This shows how the most probable hyper-parameter $\widehat{\boldsymbol{\alpha}}_j$ depends on the uncertain scaling parameters $\boldsymbol{\theta}_j$. Further substituting Eq. (28) into Eq. (23) yields:

$$\begin{aligned}&\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\widehat{\boldsymbol{\alpha}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\right]\\&\quad=-\frac{N_mN_t}{2}\ln(2\pi)-N_mN_t\ln\left(\sigma_j\right)-\frac{N_j}{2}\ln(2\pi)-\frac{N_j}{2}\\&\quad-\frac{N_mN_t}{2\sigma_j^2}\mathcal{J}\left(\boldsymbol{\theta}_j;\mathbb{D}_N,\mathbb{M}_j\right)-\frac{1}{2}\ln\left[\det\left(\widehat{\mathbf{A}}_j(\boldsymbol{\theta}_j)\right)\right]\end{aligned}\tag{29}$$

It should be pointed out that the above procedure does not consider the uncertainty of stiffness scaling parameter estimate $\widehat{\boldsymbol{\theta}}_j^c$, i.e., treating the parameters $\widehat{\boldsymbol{\theta}}_j^c$ to be a deterministic vector. To be more practical, by also involving the covariance matrix $\widehat{\boldsymbol{\Sigma}}_j^c$ achieved from the calibration stage as the pseudo-data, the diagonal covariance matrix $\widehat{\mathbf{A}}_j(\boldsymbol{\theta}_j)$ defined by the most probable hyper-parameter vector $\widehat{\boldsymbol{\alpha}}_j(\boldsymbol{\theta}_j)$ in Eq. (29) should be replaced by the diagonal covariance matrix $\widehat{\mathbf{B}}_j\left(\boldsymbol{\theta}_j\right)=\mathrm{diag}\left[\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j)\right]=\widehat{\mathbf{A}}_j\left(\boldsymbol{\theta}_j\right)+\widehat{\boldsymbol{\Sigma}}_j^c$, and meanwhile the pseudo-data is updated to be $\widehat{\boldsymbol{\vartheta}}_j^c=\{\left(\widehat{\boldsymbol{\theta}}_j^c\right)^{\mathrm{T}},(\widehat{\boldsymbol{\sigma}}_j^c)^{\mathrm{T}}\}^{\mathrm{T}}$ accordingly.

In globally identifiable cases, the posterior PDF $p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)$ provided in Eq. (22) given a large amount of data $\mathbb{D}_N$ may be approximated accurately by a Gaussian distribution, so the evidence $p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)$ can be approximated by using Laplace's method for asymptotic expansion [52]. However, this asymptotic expansion would be not valid for the general case where the posterior PDF may not be approximated by the Gaussian distribution. Motivated by the procedure proposed in [53] and also noticing Eq. (21), the following identity of logarithm of evidence can be considered as

$$\begin{aligned}&\ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]=\\&=\int\ln\left[\frac{p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\boldsymbol{\theta}_j|\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)}{p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)}\right]p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\mathrm{d}\boldsymbol{\theta}_j\end{aligned}\tag{30}$$

By utilizing Eq. (20), the log-evidence can be further expressed as

$$
\begin{aligned}
&\ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]\\
&\propto\int\ln\left[\frac{p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)p\left(\boldsymbol{\theta}_j|\mathbb{M}_j\right)}{p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)}\right]p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\mathrm{d}\boldsymbol{\theta}_j\\
&=\int\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\right]p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\mathrm{d}\boldsymbol{\theta}_j\\
&-\int\ln\left[\frac{p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)}{p\left(\boldsymbol{\theta}_j|\mathbb{M}_j\right)}\right]p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\mathrm{d}\boldsymbol{\theta}_j
\end{aligned}
\tag{31}
$$

The first term in the right hand side of Eq. (31) is a measure of the average data-fit of the model class $\mathbb{M}_j$, accounting for the log-goodness of fit for different combinations of the parameters weighted by the posterior PDF. The second term is the Kullback–Leibler information, which is a non-negative measure of the information gain about $\mathbb{M}_j$ from the data $\mathbb{D}_N$. If the selection of a model class is solely determined by the data-fit term, then more complex models will usually be preferred over simpler ones. This often leads to over-fitting of the data and the updated model depending too much on the details of the specific data will be unreliable. The combination of these two factors in the log evidence for $\mathbb{M}_j$ provides a mathematically rigorous and robust way to build in a trade-off between the data-fit of the model class and its information-theoretic complexity.

The probability of model class $\mathbb{M}_j$ conditional on both the dynamic data $\mathbb{D}_N$ and pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c$ can be obtained by following the Bayes' theorem as

$$
p\left(\mathbb{M}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}\right)=\frac{p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)p\left(\mathbb{M}_j|\mathbb{M}\right)}{\sum_{j=1}^{N_M}p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)p\left(\mathbb{M}_j|\mathbb{M}\right)}
\tag{32}
$$

where the denominator is given by the theorem of total probability. $p(\mathbb{M}_j|\mathbb{M})$ represents the judgment on the initial plausibility of the model classes expressed as the prior plausibility $p(\mathbb{M}_j|\mathbb{M})$ on the model class $\mathbb{M}_j$, $j=1,2,\cdots,N_M$. The prior plausibilities are normalized as $\sum_{j=1}^{N_M}p(\mathbb{M}_j|\mathbb{M})=1$, and it is simply assumed here that each class of models has the same initial plausibility. The evidence $p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)$ expressing how likely the data $\mathbb{D}_N$ is obtained with the specified model class $\mathbb{M}_j$ is obtained previously by Eq. (31).

It is shown in Eq. (32) that the most probable class of models is the one that maximizes $p\left(\mathbb{M}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}\right)$ which is equivalent to maximizing the evidence $p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)$ with respect to j. However, it should be noted that for evaluating the plausibility of different model

classes, directly taking the exponential of the log-evidence and then normalizing the plausibility may cause computational problems due to the existence of zero or infinity values. As a solution, the maximum log-evidence is calculated and subtracted from the log-evidence of each individual model class and then taking the exponential of the resulting quantities [54]. This operation does not affect the relative plausibility between different model classes and the problem of numerical difficulties can be well resolved. Thus, the plausibility of the jth model class $\mathbb{M}_j$ can be obtained through the following normalizing form as

$$p\left(\mathbb{M}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}\right)=\frac{\exp\left(\ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]-\max\limits_j \ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]\right)}{\sum_{j=1}^{N_M}\exp\left(\ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]-\max\limits_j \ln\left[p\left(\mathbb{D}_N|\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)\right]\right)} \tag{33}$$

2.3 Delayed rejection adaptive Metropolis algorithm

The high-dimensional integrals in log-evidence provided in Eq. (31) can be evaluated numerically by the MCMC algorithm. For this purpose, the samples are required to be drawn from the posterior PDF $p\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)$, and they are obtained by implementing the DRAM algorithm in the present paper. This algorithm combines two powerful ideas, i.e., adaptive Metropolis algorithm and delayed rejection strategy [55], which has been proved to be more effective than traditional MCMC algorithms. To ensure that this paper is self-contained, the principle of implementing DRAM methods to solve the specific problem presented in the previous section is briefly summarized, and interested readers are directed to Ref. [56] for detailed explanation. Hereafter, the logarithm of product of two likelihood functions is defined for convenience by

$$\mathcal{L}\left(\boldsymbol{\theta}_j|\mathbb{D}_N,\widehat{\boldsymbol{\vartheta}}_j^c,\mathbb{M}_j\right)=\ln\left[p\left(\mathbb{D}_N|\boldsymbol{\theta}_j,\mathbb{M}_j\right)p\left(\widehat{\boldsymbol{\vartheta}}_j^c|\boldsymbol{\theta}_j,\widehat{\boldsymbol{\beta}}_j(\boldsymbol{\theta}_j),\mathbb{M}_j\right)\right] \tag{34}$$

with the present of the measured data set $\mathbb{D}_N$, the pseudo-data $\widehat{\boldsymbol{\vartheta}}_j^c$ and the specified class of modes $\mathbb{M}_j$.

Initially the DRAM algorithm requires the selection of a covariance design parameter s_d, the adaptation interval length k_0 and the maximum number of allowed chain iterates K. The parameter k_0 determines when the covariance matrix $\boldsymbol{\Sigma}^{(k)}$ of the chain should be updated. The choice of k_0 is critical in adaptive Metropolis for achieving a good balance mixing in the initial stages and for generating non-singular covariance matrices. The fundamental steps are briefly summarized as following:

1. To begin with, set design parameters s_d and k_0 and the number of simulations K.

2. Estimate the initial state $\boldsymbol{\theta}_j^{(0)}$ by

$$\boldsymbol{\theta}_j^{(0)} := \arg\min_{\boldsymbol{\theta}_j}\{-\mathcal{L}(\boldsymbol{\theta}_j|\mathbb{D}_N, \widehat{\boldsymbol{\vartheta}}_j^c, \mathbb{M}_j)\} \tag{35}$$

and also specify a positive define covariance matrix $\boldsymbol{\Sigma}^{(0)}$.

3. For the kth simulation step, generate the candidate state $\boldsymbol{\theta}_j^{*1}$ by utilizing the current state $\boldsymbol{\theta}_j^{(k)}$ and the specified proposal distribution as $\boldsymbol{\theta}_j^{*1} \sim q_1\left(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{(k)}\right)$.

4. Calculate the acceptance ratio for the current candidate $\boldsymbol{\theta}_j^{*1}$ as

$$\tau_1\left(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{(k)}\right) = \min\left\{1, \frac{\mathcal{L}(\boldsymbol{\theta}_j^{*1}|\mathbb{D}_N, \widehat{\boldsymbol{\vartheta}}_j^c, \mathbb{M}_j) q_1\left(\boldsymbol{\theta}_j^{(k)}|\boldsymbol{\theta}_j^{*1}\right)}{\mathcal{L}\left(\boldsymbol{\theta}_j^{(k)}|\mathbb{D}_N, \widehat{\boldsymbol{\vartheta}}_j^c, \mathbb{M}_j\right) q_1\left(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{(k)}\right)}\right\} \tag{36}$$

5. Sample from a uniform distribution, i.e., $u_1 \sim \mathbb{U}(0,1)$. If $u_1 < \tau_1$, accept $\boldsymbol{\theta}_j^{*1}$ by setting $\boldsymbol{\theta}_j^{(k+1)} = \boldsymbol{\theta}_j^{*1}$.

6. Otherwise, enter the delayed rejection algorithm, and generate another candidate sate $\boldsymbol{\theta}_j^{*2}$ through the proposal distribution as $\boldsymbol{\theta}_j^{*2} \sim q_2\left(\boldsymbol{\theta}_j^{*2}|\boldsymbol{\theta}_j^{(k)}, \boldsymbol{\theta}_j^{*1}\right)$.

7. Calculate the acceptance ratio τ_2 by

$$\begin{aligned}
&\tau_2\left(\boldsymbol{\theta}_j^{*2}|\boldsymbol{\theta}_j^{(k)}, \boldsymbol{\theta}_j^{*1}\right) \\
&= \min\left\{1, \frac{\mathcal{L}(\boldsymbol{\theta}_j^{*2}|\mathbb{D}_N, \widehat{\boldsymbol{\vartheta}}_j^c, \mathbb{M}_j) q_1(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{*2}) q_2\left(\boldsymbol{\theta}_j^{(k)}|\boldsymbol{\theta}_j^{*1}, \boldsymbol{\theta}_j^{*2}\right)\left[1 - \tau_1(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{*2})\right]}{\mathcal{L}\left(\boldsymbol{\theta}_j^{(k)}|\mathbb{D}_N, \widehat{\boldsymbol{\vartheta}}_j^c, \mathbb{M}_j\right) q_1\left(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{(k)}\right) q_2\left(\boldsymbol{\theta}_j^{*2}|\boldsymbol{\theta}_j^{(k)}, \boldsymbol{\theta}_j^{*1}\right)\left[1 - \tau_1\left(\boldsymbol{\theta}_j^{*1}|\boldsymbol{\theta}_j^{(k)}\right)\right]}\right\}
\end{aligned} \tag{37}$$

8. Sample from a uniform distribution again, i.e., $u_2 \sim \mathbb{U}(0,1)$. If $u_2 < \tau_2$, accept $\boldsymbol{\theta}_j^{*2}$ by setting $\boldsymbol{\theta}_j^{(k+1)} = \boldsymbol{\theta}_j^{*2}$; Otherwise, reject $\boldsymbol{\theta}_j^{*2}$, and set $\boldsymbol{\theta}_j^{(k+1)} = \boldsymbol{\theta}_j^{(k)}$.

9. If $k > k_0$, follow the adaptive Metropolis rule, update the covariance matrix $\boldsymbol{\Sigma}^{(k+1)} = s_d \mathrm{Cov}\left(\boldsymbol{\theta}_j^{(0)}, \boldsymbol{\theta}_j^{(1)}, \cdots, \boldsymbol{\theta}_j^{(k+1)}\right)$; Otherwise, set $\boldsymbol{\Sigma}^{(k+1)} = \boldsymbol{\Sigma}^{(k)}$.

It should be pointed out that, although only two proposal stages are outlined here, the above algorithm allows for as many proposal stages to be performed as desired in principle.

3. NUMERICAL SIMULATIONS

3.1 Introduction of the four-storey two-bay frame example

For verifying the proposed methodology, a four-storey two-bay plane frame with standard 10# I-steel columns and beams as shown in Fig. 1 is employed in this section. The sectional and material properties of this frame are provided in Table 1. In this present study, the column-base and beam-column connections of the frame are all treated as semi-rigid with rotational stiffness. For convenience, the semi-rigid connections for both column-base and beam-column connections are simulated by beam elements of very short length with relatively smaller flexural rigidity as compared to the regular beam and column components. The connection statuses are quantified by the flexural stiffness of these short beam elements, and they are further considered as uncertain modeling parameters to be identified. Referring to Fig. 1, there are 19 such semi-rigid connections, denoted by S1, S2, …, and S19, respectively, where S1 to S3 are utilized to simulate the semi-rigid column-base connections by three short beams, and S4 to S19 are designed to quantify all the beam-column semi-rigid connections.

The four-storey steel frame is discretized by plane beam elements (each node with two translational and one rotational DOFs) into two FE models with different discretizing schemes in order to purposely introduce modeling error effect. The first FE model (FEM1, with 94 nodes and 273 DOFs) with a level of 5% Gaussian white noise of zero-mean added into the sectional properties is utilized for generating simulated modal data, while the second one (FEM2, with 54 nodes and 153 DOFs) is employed for connection damage detection using simulated data got from FEM1. It is noted that the various numbers next to the markers shown in Fig. 1 denote the FE node numbers, and two numbers separated by a slash represent the corresponding nodes belonging to FEM1 and FEM2, respectively.

There are twelve sensors utilized for this frame as shown in Fig. 1. The first four sensors are used to measure the horizontal vibration of the left column, while the latter eight are employed to monitor the vertical motion of the two beams. Although the complete set of modal parameters can be easily obtained from eigen-system solution for this numerical model, only the first two modes with partial DOFs are employed to investigate the performance of this approach with relatively less information about the dynamics of the structure. The measurement includes the eigenvalues and partial mode shapes at the measured DOFs of the first two modes under both healthy and damaged situations. They are calculated from FEM1 with measurement error assumed to be an i.i.d. Gaussian white noise of zero-mean with a level of 1% without losing the generality. In the following calibration as well as damage detection stages, the identification model, i.e., FEM2 is employed.

Table 2 shows the six classes of models considered for this four-storey frame. The complexity of model parameterization increases gradually from $\mathbb{M}_1$ to $\mathbb{M}_5$. Specifically, $\mathbb{M}_1$ with one parameter θ_1 to scale all rotational stiffness possesses the lowest degree of complexity among all considered model classes. $\mathbb{M}_2$ is a little bit more complex than $\mathbb{M}_1$, and the former has two parameters θ_1 and θ_2 to scale the rotational stiffness of all the column-base and beam-column connections, respectively. With one more parameters involved, $\mathbb{M}_3$ is proposed to distinguish the beam-column connections of the first two stories from the last two. $\mathbb{M}_4$ is developed based on $\mathbb{M}_3$ to further discriminate the semi-rigid stiffness of middle column-base connection (S1) from the remaining two column-base connections (S1 and S3). As the most complex parameterization scheme among all six model classes, $\mathbb{M}_5$ uses five scaling parameters to update all column-base connections as well as all beam-column connections with respect to each storey. $\mathbb{M}_6$ is designed to be a comparison model class originated from $\mathbb{M}_2$ with one more parameters involved to discriminate the middle column-base connection from the rest two. This model class is intended to be comparable with $\mathbb{M}_3$ for investigating the effect of different parameterization scheme by keeping the same number of unknown parameters.

There are three cases considered for this four-storey frame as showed in Table 3. It is assumed that the rotational stiffness of column-based and beam-column connections are consistent for convenience. k_c denotes the magnitude of nominal or baseline values of rotational stiffness, and it is set to be 0.02 times of the flexural rigidity of the beam and column elements. Case 1 considers the damage on single column-based connection, and the rotational stiffness of left beam-column connection S1 of the first storey is set to be 0.7 times of nominal values, while all the remaining unknown modeling parameters are kept to be their baseline values. Case 2 investigates the situation that all the column-base connections are damaged with different extents. Case 3 is also a multi-damage case, where six damaged beam-column connections are located in two adjacent columns, which is employed to verify the ability of proposed methodology in a more difficult situation for detecting larger number of damaged connections.

3.2 Numerical simulation results

The prior probability distributions for the stiffness scaling parameters are chosen as a zero-mean Gaussian distribution with standard deviation of 0.3 to ensure that most of the prior samples fall within a prescribed range between -1 and 1, and the samples going beyond this scope are discarded. It is recalled that the proposed methodology consists of two stages, i.e., the calibration stage (denoted by Stage-I) and subsequent monitoring stage (denoted by Stage-II). In the first stage, by applying the proposed methodology to each class of models,

FEM2 corresponding to various FE parameterization schemes are updated by utilizing the simulated modal parameters obtained from FEM1 under intact status, and the pseudo-posterior samples of uncertain scaling parameters in this stage can be achieved for each model class. It is noted that the so-called MAP estimates and associated uncertainties of the uncertain stiffness parameters in Stage-I are obtained from the statistical properties of MCMC samples achieved in this stage. Then, in the second stage, FEM2 for each model class is updated again by using simulated modal data from FEM1 in damaged situation as well as the information provided by the previously obtained samples of uncertain parameters.

By applying the proposed methodology, a total of 3×10^4 samples are obtained, and those within the 'burn-in' period are further excluded. Table 4 shows the results of Bayesian model class selection performed for different cases with the prescribed set of model classes, where the results corresponding to the most probable model class are highlighted in bold. As for Case 1, it is found that the most probable class of models is $\mathbb{M}_1$, possessing the simplest parameterization scheme with one single parameter. Referring to the damage configuration of this case as shown in Table 3, this seems to be some counterintuitive at the first glance by inspecting the log-evidence value. As it is our common sense that this simplest modeling strategy is unable to distinguish the rotational stiffness reduction of column-base connections from that of beam-column connections, and we would expect more complex class of models be selected in this case. This result may be understandable if we examine the measure of data-fit as well as the Kullback information of the uncertain parameters with respect to the chosen prior probability distributions. In this case, as shown in Table 4, data-match of model class $\mathbb{M}_2$ is larger than $\mathbb{M}_1$ since one more parameter is involved. Accordingly, larger amount of information is extracted for $\mathbb{M}_2$ to update its extra parameter as indicated by the information-gain value. However, from model classes $\mathbb{M}_1$ to $\mathbb{M}_2$, the amount of the information extraction (4.22-2.24=1.98) exceed the benefits of the data-match improvement (-194.00+195.53=1.53), resulting in a lower log-evidence for model class $\mathbb{M}_2$. With continuous increase of the model complexity, more information is generally expected to be extracted to update the excessive parameters, which overshadow the improvement of data-fit measure, making $\mathbb{M}_1$ to be the most probable class of modes for this case.

Case 2 investigates more serious damage situation where all the column-base connections are damaged with various extent as compared to the previous case. As shown in Table 4, $\mathbb{M}_2$ is identified as the most probable class of models in this case by utilizing the proposed Bayesian evidence measure, while the plausibility of the simplest model class $\mathbb{M}_1$ is extremely low for this case. This is expected and can be explained by examining the values of both data-match

and information gain. Specifically, it is clearly found that data-match of $\mathbb{M}_1$ is far less than others in the prescribed set of model classes. This is not surprised since the present of damage in all column-base connections for this case induces more significant change of structural dynamic properties as compared to the previous case. In such circumstances, the matching between the single-parameter model class $\mathbb{M}_1$ and the damaged structural model is very poor, making the log-evidence of $\mathbb{M}_1$ to be at a very low level regardless of its relatively less information extraction. For other model classes in this case, although more complex models generally match experimental data better, the proposed Bayesian evidence measure penalizes the complexity of model parameterization by assessing the amount of information extracted. Thus, $\mathbb{M}_2$ outperforms the other pre-defined classes of models for this case.

Case 3 considers a more complex damage configuration, where damaged beam-column connections are located within the first three storeys. For this case, it is seen from Table 4 that $\mathbb{M}_3$ is identified to be the most probable class of models class. This is not surprised since it's well anticipated that this model class possessing a relative simple parameterization scheme discriminates the beam-column connections belonging to two separated groups (i.e., the first two and last two stories), and matches well with the prescribed damage configuration. In addition, it is also observed that $\mathbb{M}_2$ is chosen as the second-best model class and outperforms the most complex one $\mathbb{M}_5$. This is also expected by examination of the data-fit measure as well as Kullback information for this case. It is found that although the data-match measure of $\mathbb{M}_2$ is not as good as that of $\mathbb{M}_5$ due to its deficiency for identifying the beam-column connection damage located in each storey, the amount of information extracted to update its two uncertain parameters is obviously less than that required by $\mathbb{M}_5$ with the five parameters. As for the model class $\mathbb{M}_6$, although its parameterization complexity is comparable with $\mathbb{M}_3$, i.e., both with three parameters, it is clearly found that log-evidence of the former is less than the latter. The main reason is due to the fact that there is only single parameter utilized in $\mathbb{M}_6$ for characterizing the status of all the beam-column connections, the data-fit measure of $\mathbb{M}_6$ would be much worse than $\mathbb{M}_3$, even though the amount of information extracted of the former model class is relatively less, rendering $\mathbb{M}_6$ to be less plausible as compared to $\mathbb{M}_3$.

By taking the most probable model class $\mathbb{M}_3$ in Case 3 as an example, Fig. 2 shows the samples of prior and posterior distribution of all stiffness scaling parameters after excluding the samples in 'burn-in' period within both calibration and monitoring stages. It's obvious that the degree of dispersion of the samples obtained from Stage-II is relatively smaller than those from the previous stage. This is anticipated since in the proposed methodology, the Markov chain samples of unknown parameters achieved in the calibration stage are taken as the pseudo-data

for further estimating the distribution of these uncertain parameters in the subsequent monitoring stage. This implies that a certain amount of information has been learned from the samples of the calibration stage to update the estimation of unknown parameters during the monitoring stage, and the uncertainty of parameter estimation in the latter stage is thus reduced. In addition, it can also be found that there exist obvious shifts along positive direction between samples of two stages for the scaling parameters θ_2 and θ_3, indicating that the statuses of these parameters are highly suspected to be changed due to the present of damage, which coincides with the actual damage configuration for this case. This can be further verified through the marginal posterior PDF of the most probable model class $\mathbb{M}_3$ for this case as shown in Fig. 3. It is clear that the uncertainties of all scaling parameters in the second stage are less than the first stage due to the present of pseudo-data provided in the first stage. Also, the changes of stiffness parameters θ_2 and θ_3 are found to be more obvious in this figure, and thus damage occurred in the corresponding element group is validated.

With the generated samples, Fig. 4 shows a normal probability plot of all scaling parameters in $\mathbb{M}_3$ for Case 3 from both two stages. It plots the empirical cumulative distribution of the sample data versus the theoretical cumulative distribution function of a normal distribution, which is employed to assess graphically whether these samples of each parameter come from a population with a normal distribution. It's clearly found that the sample data for each scaling parameter appears to be a normal distribution. In addition, since the line slope indicates the relative uncertainty, it's obviously shown that the uncertainties of θ_1 and θ_3 are very similar due to their similar slopes, which can also be verified through the results of Fig. 3. Furthermore, in order to better interpret the damage identification results obtained by the proposed methodology, based on the marginal posterior PDFs achieved from both the calibration and monitoring stages as presented in Fig. 3, the probability that the ith stiffness scaling parameter θ_i has been reduced by a certain fraction d compared to the intact status of the structure [22] is suggested to be calculated. In this note, an asymptotic Gaussian approximation is utilized to give

$$\begin{aligned} p_i^{dam}\left(d|\mathbb{D}_N,\mathbb{M}_j\right) &= p\left(\hat{\theta}_i > d|\mathbb{D}_N,\mathbb{M}_j\right) \\ &= \frac{1}{\sqrt{2\pi}\hat{\sigma}_i}\int_0^d \exp\left[-\frac{1}{2\hat{\sigma}_i^2}\left(\theta_i-\hat{\theta}_i\right)^2\right]\mathrm{d}\theta_i \end{aligned} \tag{38}$$

where $i = 1,2,\cdots,N_\theta$, $\hat{\theta}_i$ and $\hat{\sigma}_i$ denote the most probable value and standard deviation of the ith stiffness scaling parameter θ_i obtained from either calibration or monitoring stage.

The corresponding results are shown in Fig. 5, and the offset of probability curves between two stages indicates the damage extent of the corresponding element groups. It's clearly revealed again that the actually damaged element groups for the four-storey frame are well identified, and the corresponding change of probability curves of the scaling parameters related to the potentially damaged element groups is more significant than others.

4. EXPERIMENTAL VERIFICATIONS

In order to further verify the proposed methodology, a laboratory two-storey bolt-connected steel frame is utilized in this section, as shown in Fig. 6, for demonstration. The span of the frame is 1.95 m and the height is 1.25 m for each storey. All beams and columns are built with I-steel possessing the same size, and corresponding sectional and material properties are shown in Table 5. Fig. 7 shows the details of beam-column and column-base connections for this two storey frame. The beams and columns are connected through two pairs of angle steel with different sizes, which are used to fix the top and bottom flanges of beam to the inner flange of column, and also connect the beam web with the column inner flange. Each column is firstly welded to a square steel plate and then fixed to the ground by four bolts, and all bolts in beam-column and column-base connections are tightened by utilizing the wrench. The column-base and beam-column connections of the frame are considered as semi-rigid and they are treated in the same way in the FE model as in the previous numerical simulations. Damage on column-base connection is considered for the laboratory two-storey frame, and two bolts (marked in red circle shown in Fig. 7(e)) of the left column-based connection are loosen by using the wrench.

The experimental equipment utilized for the laboratory two-storey frame is shown in Fig. 8. A DYTRAN 5805A impact hammer with a load cell, as shown in Fig. 8(a), is used to excite the frame, and the excitation points are chosen to avoid the stationary nodes of the first few modes. The dynamic responses are acquired by KISTLER accelerometers (Type 8776A50M3) shown in Fig. 8(b) with sensitivity around 100mV/g attached to both columns and beams, and then collected by the 16-channeled DEWETRON signal conditioning box together with data acquisition software DEWESoft (referring to Fig. 8(c) to (f)). It is noted that only the free vibration responses are measured for modal parameter identification by utilizing the ERA technique.

Fig. 9 shows the FE model of the experimental two-storey bolt-connected frame utilized for damage detection. By using the same plane beam element as in the previous numerical example, the frame is discretized into a FE model of 30 elements and 30 nodes with 84 DOFs.

Six potential damaged connections, i.e., two column-base connections and four beam-column connections denoted by S1, S2, …, and S6, respectively. In addition, as shown in the same figure, fourteen measurement channels are utilized. The first eight sensors are used to measure the horizontal vibration of the frame, while the left ones are employed to monitor the vertical motion of the two beams. Table 6 shows the natural frequencies of the first four modes calculated from the FE model and identified from experimental modal testing. It is clear that the results of FE model matches well with those of the experimental measurement, implying that the proposed FE model is sufficient to capture the dynamic characteristics of the laboratory frame in its healthy situation. In addition, it is also obvious that influence of the prescribed damage configuration, i.e., loosening of column-base connection, on the fundamental mode is much more significant than others. In addition, Fig. 10 shows the measured mode shapes of the first four modes under both healthy and damaged situations together with the calculated results from the FE model of the two-storey frame plotted for reference. It's seen that the measured mode shapes are well matched with those predicted by the FE model in the intact status. It's also noted that although the damage-induced changes due to the connection bolt loosening in mode shapes are found to be very small as expected, it can still be noticed that the effect of bolt loosening of left column-based connection with respect to the first two modes seems to be a little more obvious than the latter two ones.

There are six model classes considered for this two-storey portal frame as shown in Table 7. The complexity of model parameterization increases gradually from $\mathbb{M}_1$ to $\mathbb{M}_6$. For instance, $\mathbb{M}_1$ with one parameter θ_1 to scale all six rotational stiffness has the lowest degree of complexity among all considered model classes. $\mathbb{M}_2$ is a little bit more complex than $\mathbb{M}_1$, and it has two separate parameters θ_1 and θ_2 to scale the rotational stiffness of all the column-base and beam-column connections, respectively. $\mathbb{M}_6$ is the class of models with the most complex parameterization among all six model classes, and there are six scaling parameters in this class of models to update all column-base and beam-column connections separately.

The prior probability distributions for the stiffness scaling parameters are taken to be the same as the previous numerical case studies. Table 8 shows the results of model class selection performed for this laboratory frame by applying the proposed methodology with a total of 3×10^4 samples. The results corresponding to the identified most probable model class are highlighted in bold, and it's clear that the most probable model class is $\mathbb{M}_2$. In the parameterization scheme of this model class, the semi-rigid connections are divided into two groups, which are parameterized by θ_1 and θ_2, respectively, to discriminate the column-base

connections from beam-column connections. Referring to the actual damage configuration, it is not surprised that the model class $\mathbb{M}_2$ is chosen to be the most probable one by the proposed evidence measure as this class of models is able to discriminate the damaged element group of column-base connections with a relatively simple parameterization. When further inspecting the data-match and information gain, it is found that the data-fit measure of model class $\mathbb{M}_2$ with respect to the real damage case is comparable with other model classes, while the amount of information extracted to update its two parameters of $\mathbb{M}_2$ are much less than its counterparts possessing more complex parameterization scheme. Thus, the combination of data-match and information gain measure in the proposed Bayesian evidence measure pinpoints the most probable model class in a mathematically rigorous way.

Fig. 11 shows the samples of prior and posterior distribution of some selected stiffness scaling parameters for the model class $\mathbb{M}_6$ within both two stages. Similar to the previous numerical case studies, it is also found that the degree of dispersion of the samples obtained from the monitoring stage is relatively smaller than the calibration stage. The uncertainty reduction is also due to the fact that some information is learned from the Markov chain samples achieved in the calibration stage. In addition, it is also clear from Fig. 11(a) that for the scaling parameter θ_1, there exists a certain extent of movement along positive direction for samples of the monitoring stage with respect to the calibration stage, implying that the damage is highly suspected to occur in the left column-base connection, which accords with the actual damage configuration. Moreover, by comparing Fig. 2 with Fig. 11, it's also very clear that the scaling parameters obtained in experimental situation are more uncertain than those in the numerical simulation cases. This is not surprised as there are more unpredictable uncertain sources which would affect the parameter estimation in real application circumstance.

Figs. 12 and 13 show the marginal posterior PDF and the probability of damage for model class $\mathbb{M}_6$, respectively. It is obvious that the uncertainties of all scaling parameters in the second stage are lower than those of the first stage due to the present of pseudo-data provided in the first stage. Also, as clearly shown in both figures, the changes of stiffness parameter θ_1 are more obvious than others, indicating the present of damage in the corresponding element or element group.

Table 9 shows the mean values and standard deviations of identified stiffness scaling parameters during the calibration stage and the subsequent monitoring stage for all considered model classes. By comparing the most probable model class $\mathbb{M}_2$ with its most complex counterpart $\mathbb{M}_6$, it is clear that the uncertainty of the former is less than the latter for both two stages, implying that with less information extracted, the parameters of the most probable

model class can be identified with a relatively higher accuracy. Thus, the identified most probable model class by the proposed Bayesian evidence measure presents a reasonable compromise related to the degree of agreement with the experimental values and the amount of extracted information. In addition, in order to verify the updated FE model corresponding to the most probable model class identified by the proposed methodology, by utilizing the difference between mean values of identified stiffness scaling parameters from both stages in Table 9, the predicted natural frequencies from the updated most probable class of models is calculated to be 14.04 Hz, 64.40 Hz, 105.48 Hz, and 110.12 Hz, respectively, which well match the natural frequencies measured in the damaged status as listed in Table 6.

5. CONCLUSIONS

This paper addresses the issues of choosing suitable parameterized models for dynamic reduction-based structural health monitoring through the model-class selection procedure by utilizing the Bayesian evidence inference. The concept of information divergence is employed to quantify the amount of information needed to be extracted from the measured data for achieving a trade-off between the parameterization complexity of a prescribed model class and that of its corresponding information-theoretic interpretation. The proposed methodology is verified through both numerical simulations of a four-storey two-bay frame with semi-rigid connections and experimental studies conducted for a laboratory two-storey bolt-connected steel frame. In the proposed methodology, only noisy incomplete modal parameters with limited number of sensors are utilized, and one of the significant features is that the method does not require the mode matching, which is vital for real applications as the mode order may occasionally change due to the present of damage.

The obtained results show that the combination of data-fit measure and Kullback information measure provides an efficient and mathematically rigorous way to select the most probable class of reduced FE models possessing a relatively simple parameterization scheme for the purpose of structural health monitoring. In addition, it is also found that the degree of dispersion of the samples obtained from the monitoring stage is relatively smaller than that from the calibration stage. This phenomenon of uncertainty reduction is mainly due to the fact that there is some information learned from the Markov chain samples achieved in the calibration stage for refining the subsequent estimation during the monitoring stage within the proposed framework of model selection. Furthermore, the estimation of unknown parameters in the most plausible class of models is found to be comparatively less uncertain and also more efficient than those models with more parameterization complexity. The reason of efficiency mainly comes from the fact that the required computation effort of the inverse problem to

identify the unknown parameters of a relatively simpler model is generally less than that of more complex models with a larger amount of uncertain parameters. This is especially crucial for the fast diagnose of structural damage for the in-service structural health monitoring. It should be pointed out that the results of model class selection depend on the form of prior probability distributions assumed for the stiffness scaling parameters, which will be further investigated in the coming publications. In addition, it is also noted that this paper mainly focuses on the identification of damage in a group of elements by inspecting the information-theoretic complexity of all prescribed model classes to ensure a relatively simple parameterization scheme, and the issue of fine localization of some specific damage within the suspected damaged groups of elements is worth further research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the National Natural Science Foundation of China (Grants No. 51778506, 51838006, and 51578260), and a scholarship from the China Scholarship Council (File No. 201806275091) while the first author visiting the Center for Engineering Dynamics and Institute for Risk and Uncertainty in the University of Liverpool. The authors would also like to thank the Editor and the anonymous reviewers for their constructive comments and valuable suggestions to improve the quality of the article.

REFERENCES

[1] E.P. Carden, F. Fanning, Vibration based condition monitoring: A review, Struct. Health Monit. 3 (2004) 355–377.

[2] O.S. Salawu, Detection of structural damage through changes in frequency: A review, Eng. Struct. 19 (1997) 718–723.

[3] H.F. Lam, K.V. Yuen, J.L. Beck, Structural health monitoring via measured Ritz vectors utilizing artificial neural networks, Comput.-Aided Civ. Infrastruct. Eng. 21 (2006) 232–241.

[4] S.S. Law, K. Zhang, Z.D. Duan, Structural damage detection from coupling forces between substructures under support excitation, Eng. Struct. 32 (2010) 2221–2228.

[5] G. Yan, S.J. Dyke, A. Irfanoglu, Experimental validation of a damage detection approach on a full-scale highway sign support truss, Mech. Syst. Signal Pr. 28 (2012) 195–211.

[6] Y. Xu, Y. Qian, J. Chen, G. Song, Probability-based damage detection using model updating with efficient uncertainty propagation, Mech. Syst. Signal Pr. 60-61 (2015) 958–970.

[7] J.T. Kim, Y.S. Ryu, H.M. Cho, N. Stubbs, Damage identification in beam-type structures:

Frequency-based method vs mode-shape-based method, Eng. Struct. 25 (2003) 57–67.

[8] A. Teughels, G.D. Roeck, Damage detection and parameter identification by finite element model updating, Arch. Comput. Method E. 12 (2005) 123–164.

[9] N.M.M. Maia, J.M.M. Silva, E.A.M. Almas, R.P.C. Sampaio, Damage detection in structures: from mode shape to frequency response function methods, Mech. Syst. Signal Pr. 17 (2003) 489–498.

[10] M. Radzieński, M. Krawczuk, M. Palacz, Improvement of damage detection methods based on experimental modal parameters, Mech. Syst. Signal Pr. 25 (2011) 2169–2190.

[11] J.B. Hansen, R. Brincker, M. Lópezaenlle, C.F. Overgaard, K. Kloborg, A new scenario-based approach to damage detection using operational modal parameter estimates, Mech. Syst. Signal Pr. 94 (2017) 359–373.

[12] L.H. Yam, Y.Y. Li, W.O. Wong, Sensitivity studies of parameters for damage detection of plate-like structures using static and dynamic approaches, Eng. Struct. 24 (2002) 1465–1475.

[13] H.F. Lam, T. Yin, Statistical detection of multiple cracks on thin plates utilizing dynamic response, Eng. Struct. 32 (2010) 3145–3152.

[14] T. Yin, H.F. Lam, H.M. Chow, A Bayesian probabilistic approach for crack characterization in plate structures, Comput.-Aided Civ. Infrastruct. Eng. 25 (2010) 375–386.

[15] M. Chandrashekhar, R. Ganguli, Damage assessment of composite plate structures with material and measurement uncertainty, Mech. Syst. Signal Pr. 75 (2016) 75–93.

[16] H.P. Chen, Application of regularization methods to damage detection in large scale plane frame structures using incomplete noisy modal data, Eng. Struct. 30 (2008) 3219–3227.

[17] E. Grande, M. Imbimbo, A multi-stage approach for damage detection in structural systems based on flexibility, Mech. Syst. Signal Pr. 76-77 (2016) 455–475.

[18] L. Yu, T. Yin, Damage identification in frame structures based on FE model updating, J. Vib. Acoust. 132 (2010) 1741–1757.

[19] J.R. Wu, Q.S. Li, Structural parameter identification and damage detection for a steel structure using a two-stage finite element model updating method, J. Constr. Steel Res. 62 (2006) 231–239.

[20] T. Yin, Q.H. Jiang, K.V. Yuen, Vibration-based damage detection for structural connections using incomplete modal data by Bayesian approach and model reduction technique, Eng. Struct. 132 (2017) 260–277.

[21] H.H. Khodaparast, J.E. Mottershead, K.J. Badcock, Interval model updating with irreducible uncertainty using the Kriging predictor, Mech. Syst. Signal Pr. 25 (2011) 1204–1226.

[22] M.W. Vanik, J.L. Beck, S.K. Au, Bayesian probabilistic approach to structural health monitoring, J. Eng. Mech.-ASCE 126 (2000) 738–745.

[23] H.F. Lam, C.T. Ng, The selection of pattern features for structural damage detection using an extended Bayesian ANN algorithm, Eng. Struct. 30 (2008) 2762–2770.

[24] W.C. Su, C.S. Huang, S.L. Hung, L.J. Chen, W.J. Lin, Locating damaged storeys in a shear building based on its sub-structural natural frequencies, Eng. Struct. 39 (2012) 126–138.

[25] K.V. Yuen, J.L. Beck, L.S. Katafygiotis, Unified probabilistic approach for model updating and damage detection, J. Applied Mech. 73 (2006) 555–564.

[26] T. Yin, H.P. Zhu, Probabilistic damage detection of a steel truss bridge model by optimally designed Bayesian neural network, Sensors 18 (2018) 3371, doi:10.3390/s18103371.

[27] H.P. Wan, W.X. Ren, A residual-based Gaussian process model framework for finite element model updating, Comput. Struct. 156 (2015) 149–159.

[28] H.P. Wan, W.X. Ren, Stochastic model updating utilizing Bayesian approach and Gaussian process model, Mech. Syst. Signal Pr. 70–71 (2016) 245–268.

[29] T. Yin, K.V. Yuen, H.F. Lam, H.P. Zhu, Entropy-based optimal sensor placement for model identification of periodic structures endowed with bolted joints, Comput.-Aided Civ. Infrastruct. Eng. 32 (2017) 1007–1024.

[30] H. Zhu, M. Wu, The characteristic receptance method for damage detection in large mono-coupled periodic structures, J. Sound Vib. 251 (2002) 241–259.

[31] H.P. Zhu, Y.L. Xu, Damage detection of mono-coupled periodic structures based on sensitivity analysis of modal parameters, J. Sound Vib. 285 (2005) 365–390.

[32] T. Yin, H.P. Zhu, S.J. Fu, Damage identification of periodically-supported structures following the Bayesian probabilistic approach, Int. J. Struct. Stab. Dyn. 19 (2019) 1940011.

[33] M.I. Friswell, J.E. Mottershead, Finite element model updating in structural dynamics, Kluwer Academic Publishers, 1995.

[34] J.E. Mottershead, M. Link, M.I. Friswell, The sensitivity method infinite element model updating: a tutorial, Mech. Syst. Signal Pr. 25 (2011) 2275–2296.

[35] B.N. Datta, Finite element model updating, eigenstructure assignment and eigenvalue embedding techniques for vibrating systems, Mech. Syst. Signal Pr. 16 (2002) 83–96.

[36] L. Mthembu, T. Marwala, M.I. Friswell, S. Adhikari, Model selection in finite element model updating using the Bayesian evidence statistic, Mech. Syst. Signal Pr. 25 (2011), 2399–2412.

[37] K.V. Yuen, J.L. Beck, L.S. Katafygiotis, Efficient model updating and health monitoring methodology using incomplete modal data without mode matching, Struct. Control Hlth. 13 (2006) 91–107.

[38] A.W. Zhu, G.J. Qu, Y.N. Gao, Matrix reduction method for updating dynamic model, Chin. Space Sci. Technol. (2003) 6–10 [in Chinese].

[39] K.V. Yuen, Updating large models for mechanical systems using incomplete modal measurement, Mech. Syst. Signal Pr. 28 (2012) 297–308.

[40] R.J. Guyan, Reduction of stiffness and mass matrices, J. AIAA 3 (1965) 380.

[41] J.C. O'Callahan, A procedure for an improved reduced system (IRS) model, In: Proceedings of 7th international modal analysis conference, (1989) 17–21.

[42] M.I. Friswell, S.D. Garvey, J.E.T. Penny, Model reduction using dynamic and iterated IRS techniques, J. Sound Vib. 186 (1995) 311–323.

[43] R.L. Kidder, Reduction of structural frequency equations, J. AIAA 11 (1973) 892.

[44] T. Yin, H.F. Lam, H.M. Chow, H.P. Zhu, Dynamic reduction-based structural damage detection of transmission tower utilizing ambient vibration data, Eng. Struct. 31 (2009) 2009–2019.

[45] H.F. Lam, T. Yin, Dynamic reduction-based structural damage detection of transmission towers: Practical issues and experimental verification, Eng. Struct. 33 (2011) 1459–1478.

[46] S.K. Au, F.L. Zhang, Y.C. Ni, Bayesian operational modal analysis: Theory, computation, practice, Comput. Struct. 126 (2013) 3–14.

[47] F. Zhang, S.K. Au, H.F. Lam, Assessing uncertainty in operational modal analysis incorporating multiple setups using a Bayesian approach, Struct. Control Hlth. 22 (2015) 395–416.

[48] M.E. Tipping, Sparse Bayesian learning and the relevance vector machine, J. Mach. Learn. Res. 1 (2001) 211–244.

[49] C.M. Bishop, Pattern Recognition and Machine Learning, Berlin: Springer, 2006.

[50] Y. Huang, J.L. Beck, Hierarchical sparse Bayesian learning for structural health monitoring with incomplete modal data, Int. J. Uncertain Quan. 5 (2015) 139–169.

[51] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, J. Eng. Mech.-ASCE, 124 (1998) 455–461.

[52] C. Papadimitriou, J.L. Beck, L.S. Katafygiotis, Asymptotic expansions for reliability and moments of uncertain systems, J. Eng. Mech.-ASCE, 123 (1997) 1219–1229.

[53] J. Ching, M. Muto, J.L. Beck, Structural model updating and health monitoring with incomplete modal data using Gibbs sampler, Comput.-Aided Civ. Inf. 21 (2006) 242–257.

[54] K.V. Yuen, Bayesian Methods for Structural Dynamics and Civil Engineering, John Wiley & Sons, Ltd, 2010.

[55] L. Tierney, A. Mira. Some adaptive Monte Carlo methods for Bayesian inference, Stat. Med. 18 (1999) 2507–2515.

[56] H. Haario, M. Laine, A. Mira, E. Saksman, DRAM: Efficient adaptive MCMC, Stat. Comput. 16 (2006) 339–354.

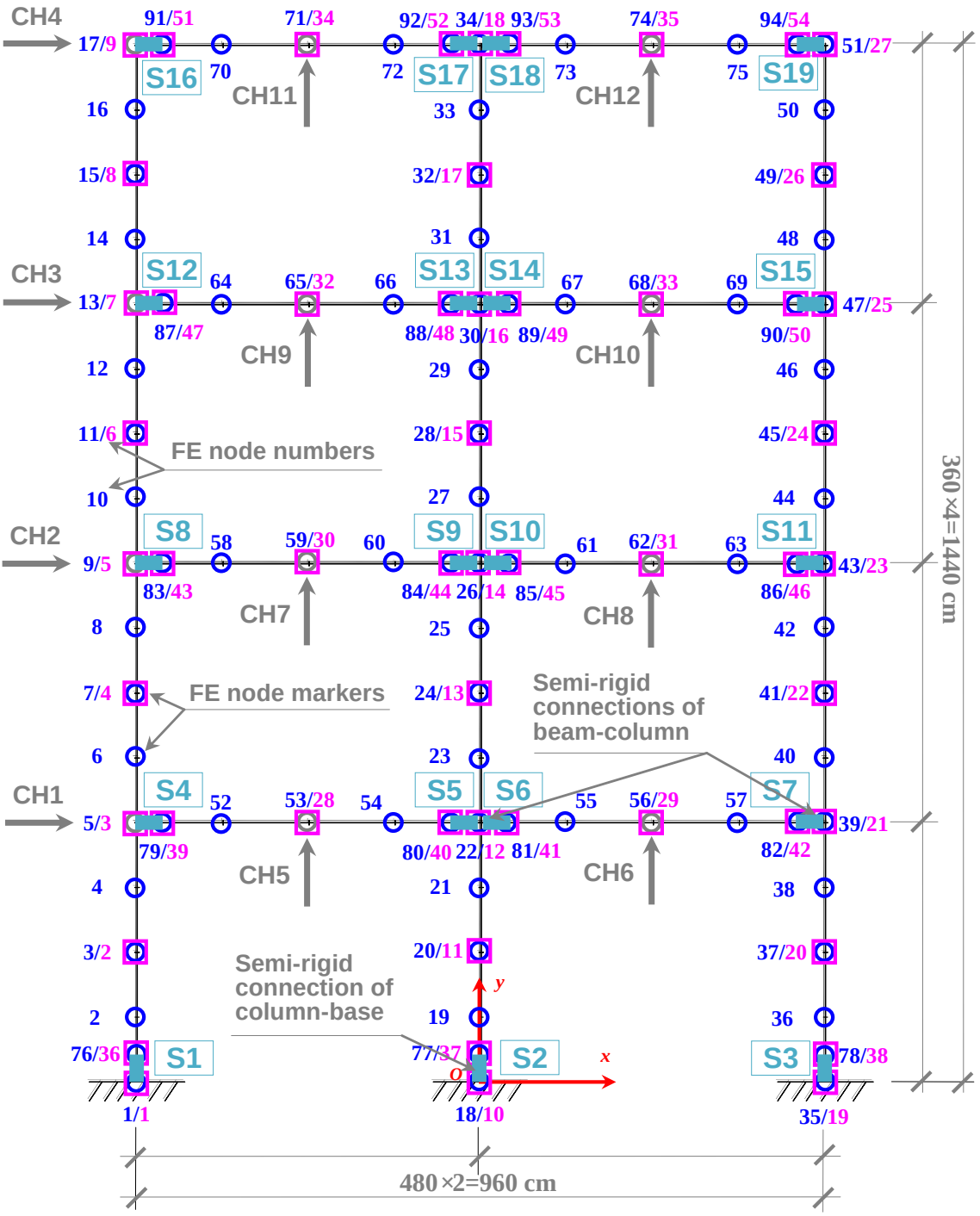


Fig. 1. Two FE models for the four-storey portal frame (FEM1 (nodes marked with blue circles) for generating simulated modal data; FEM2 (nodes marked with pink squares) for damage identification)

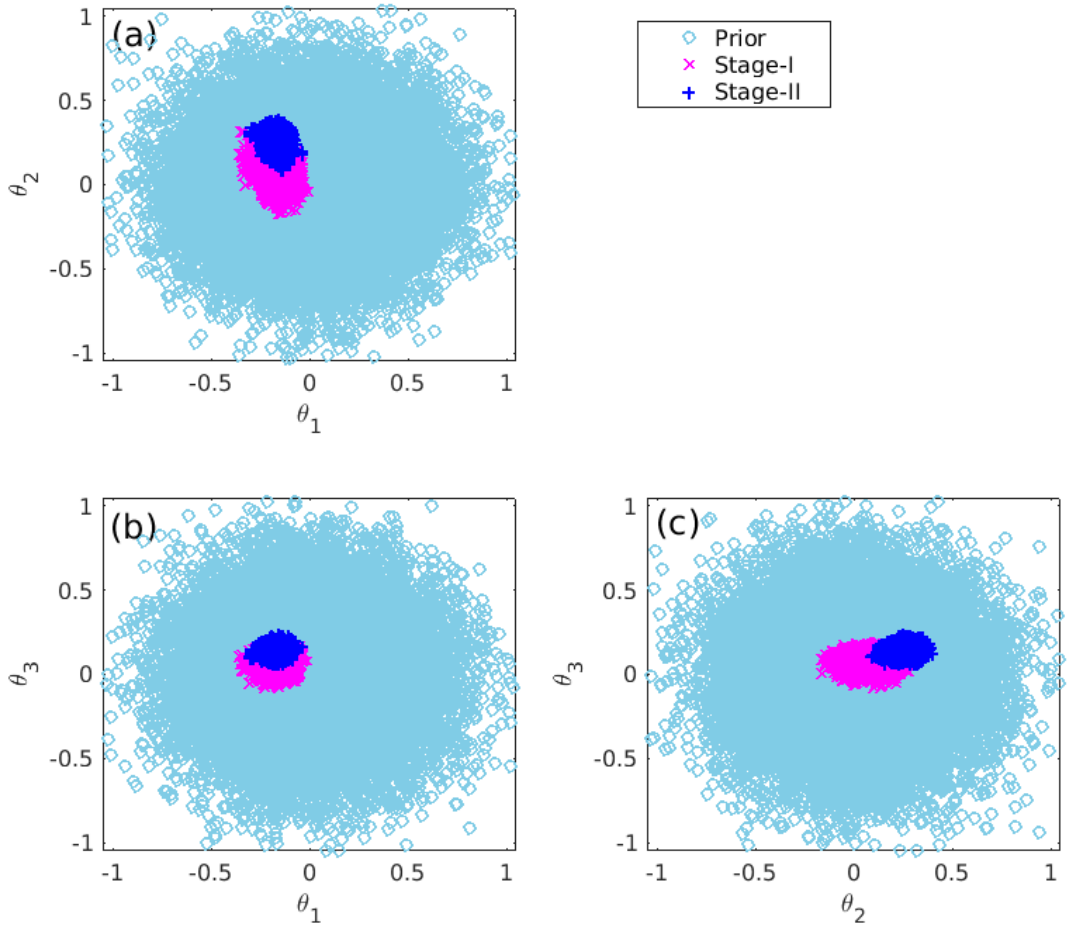


Fig. 2. Samples of prior and posterior PDFs for the most probable class of models $\mathbb{M}_3$ in Case 3: (a) $\{\theta_1, \theta_2\}$; (b) $\{\theta_1, \theta_3\}$; (c) $\{\theta_2, \theta_3\}$

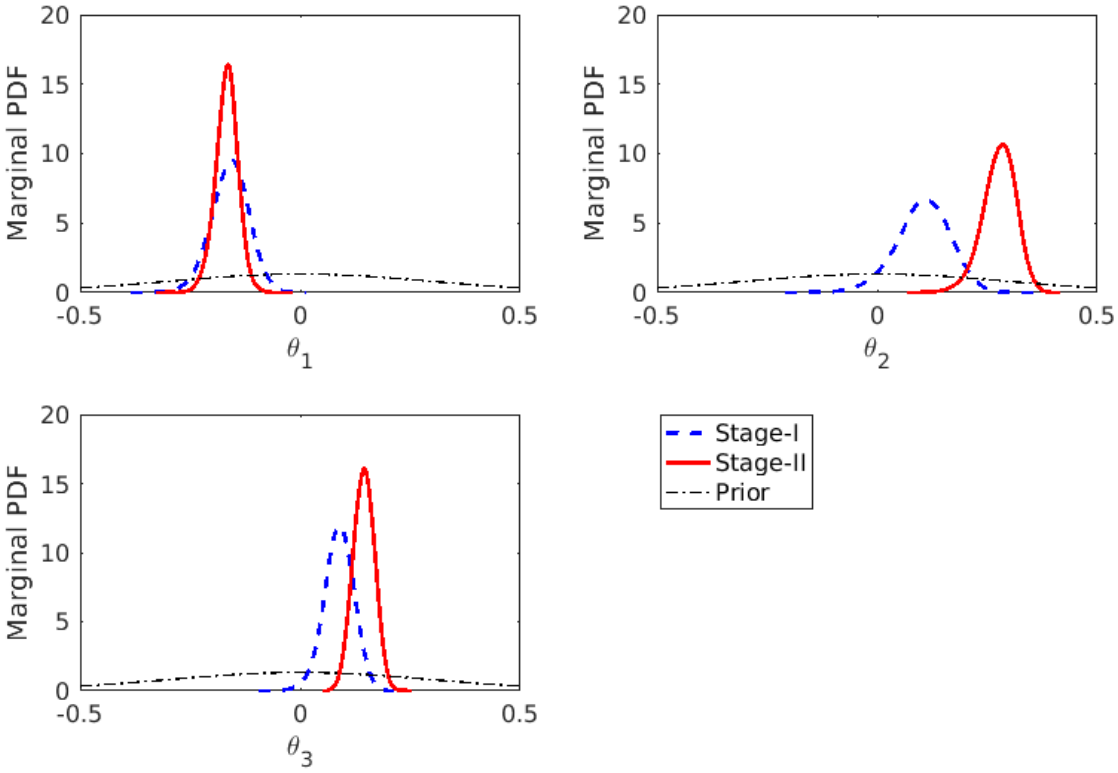


Fig. 3. Prior and marginal posterior PDFs of the two stages for the model class $\mathbb{M}_3$ in Case 3.

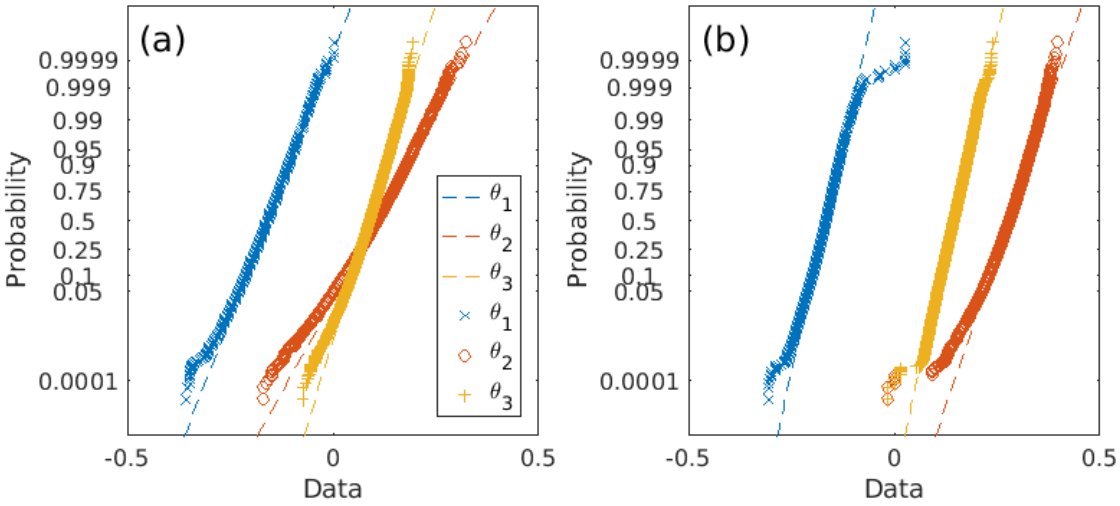


Fig. 4. Normal probability plots of scaling parameters from posterior distribution of two stages for the model class $\mathbb{M}_3$ in Case 3: (a) Stage-I; (b) Stage-II.

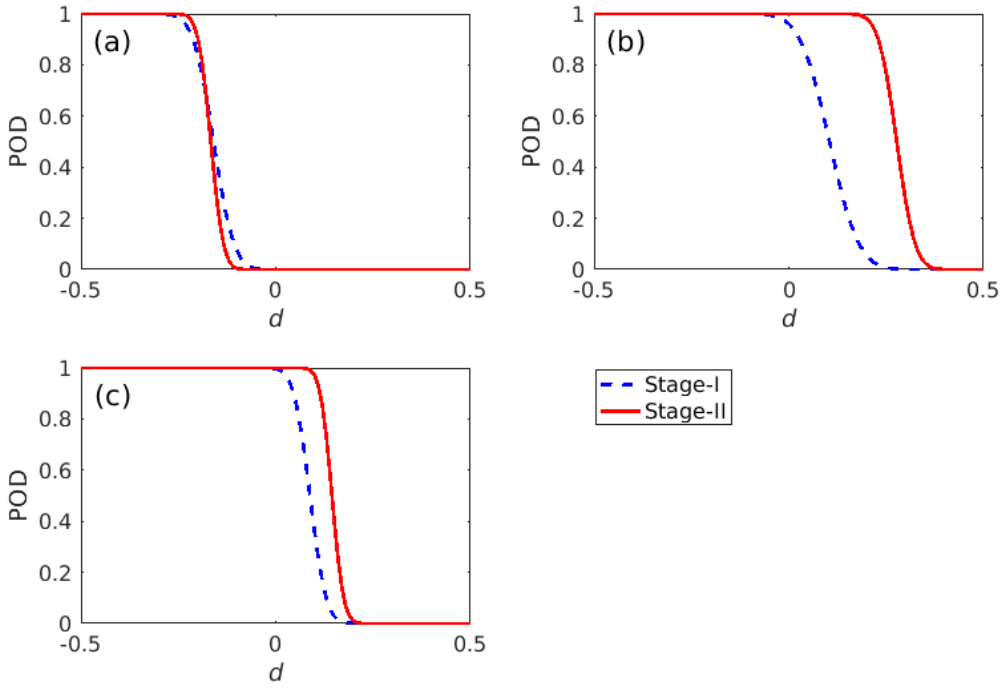


Fig. 5. Probability of damage (POD) of the two stages for the model class $\mathbb{M}_3$ in Case 3: (a) θ_1; (b) θ_2; (c) θ_3

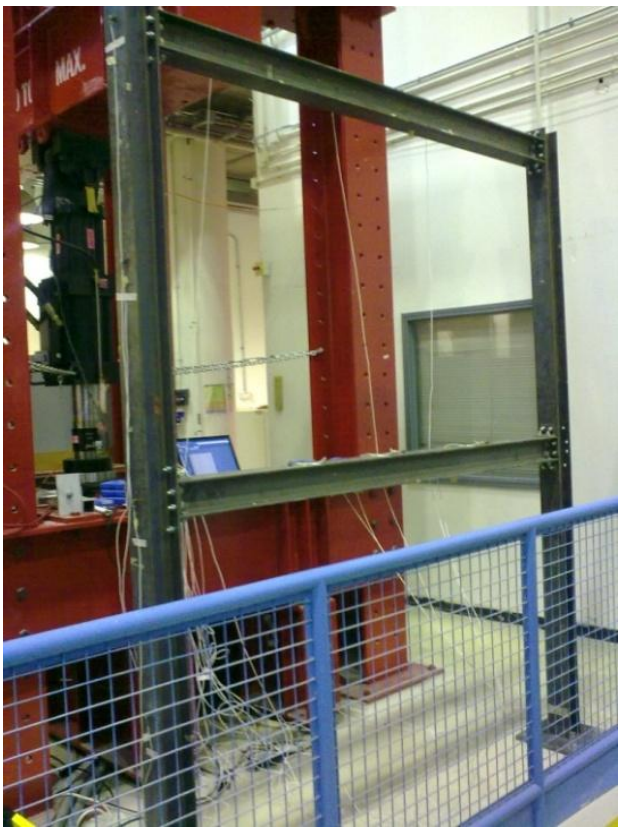
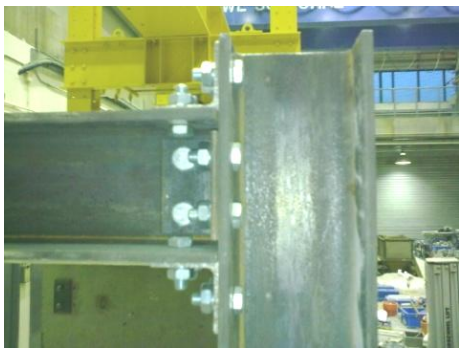


Fig. 6. The two-storey bolt-connected steel frame in laboratory.

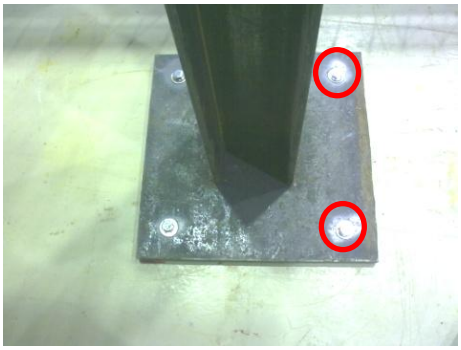
(a) Beam-column connection on the left hand side of the second storey



(b) Beam-column connection on the right hand side of the second storey

(c) Beam-column connection on the left hand side of the first storey

(d) Beam-column connection on the right hand side of the first storey

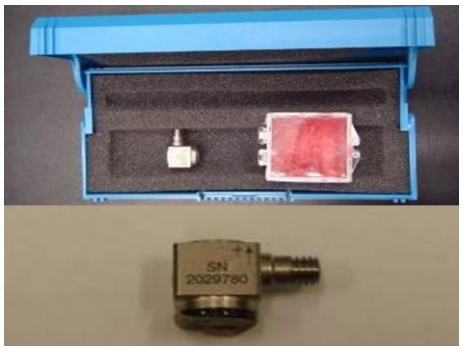


(e) Column-base connection on the left

(f) Column-base connection on the right

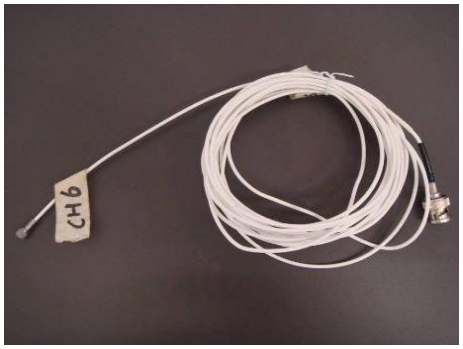
Fig. 7. Beam-column and column-base connections of the two-storey frame.

(a) Impact hammer



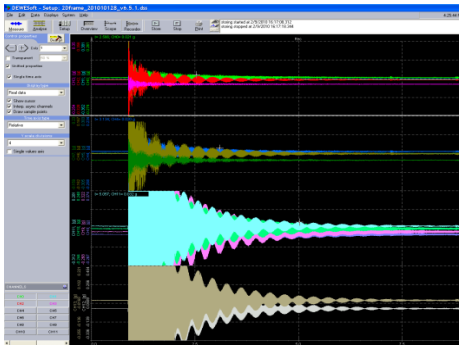
(b) KISTELER accelerometers

(c) DEWETRON signal conditioning box



(d) Cables

(e) Computer system with DEWESoft



(f) Response in DEWESoft

Fig. 8. Main experimental instruments of the laboratory two-storey frame.

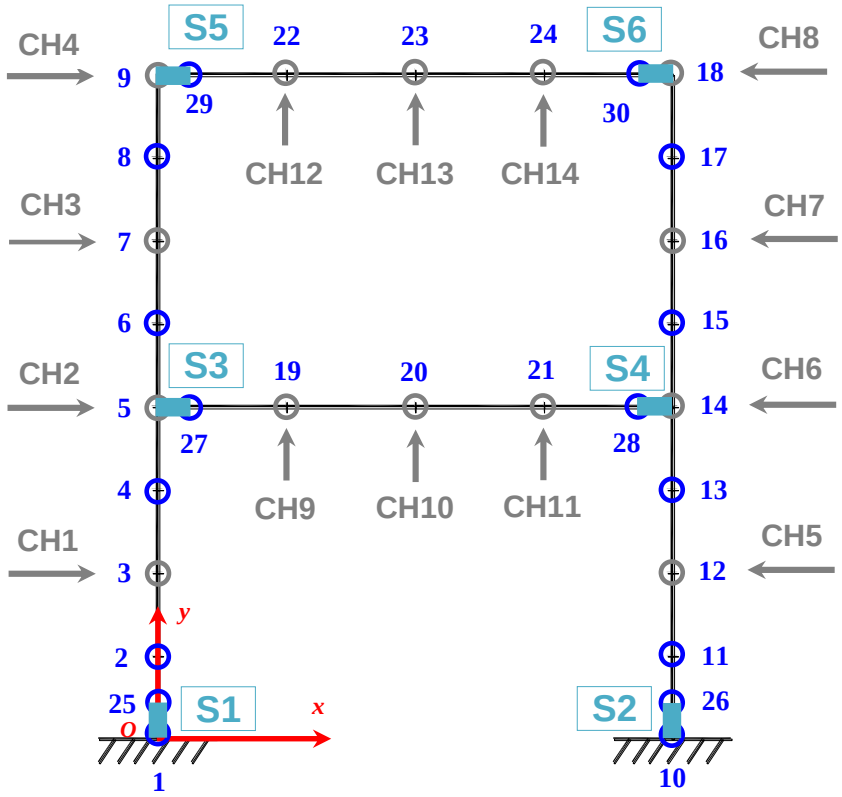


Fig. 9. FE model of the experimental two-storey bolt-connected steel frame.

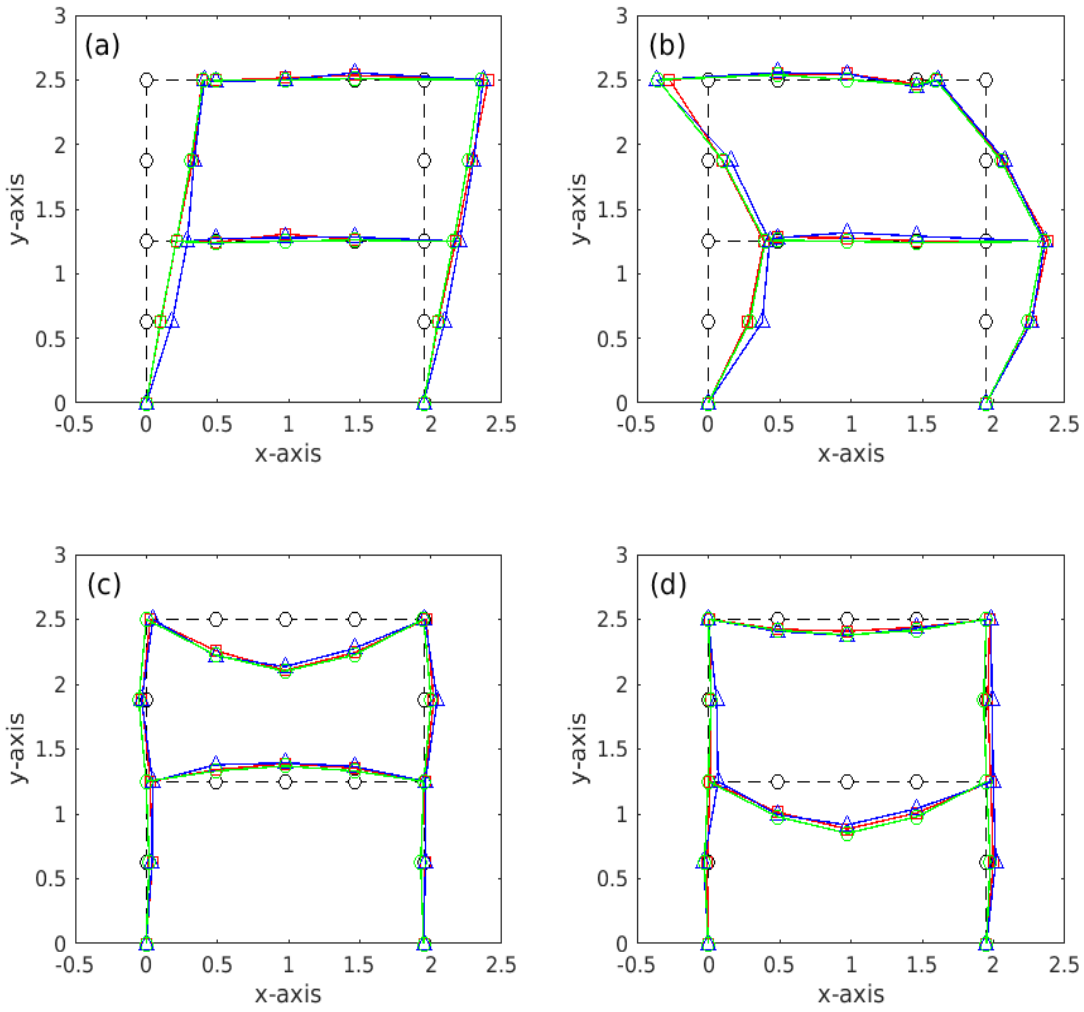


Fig. 10. Experimental and FE mode shapes of the two-storey bolt-connected frame (unit: m; red line with square marker: experimental results in healthy status; blue line with triangle marker: experimental results in damaged status; green line with circle marker: FE results in healthy status; black dotted line with circle marker: original undeformed shape): (a) to (d) modes 1 to 4.

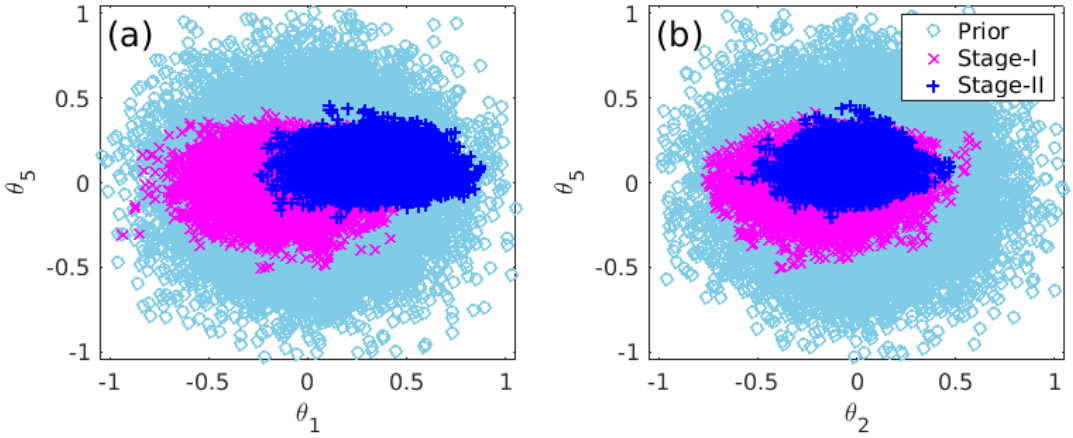


Fig. 11. Samples of prior and posterior PDFs of the two stages for the model class $\mathbb{M}_6$: (a) $\{\theta_1, \theta_5\}$; (b) $\{\theta_2, \theta_5\}$

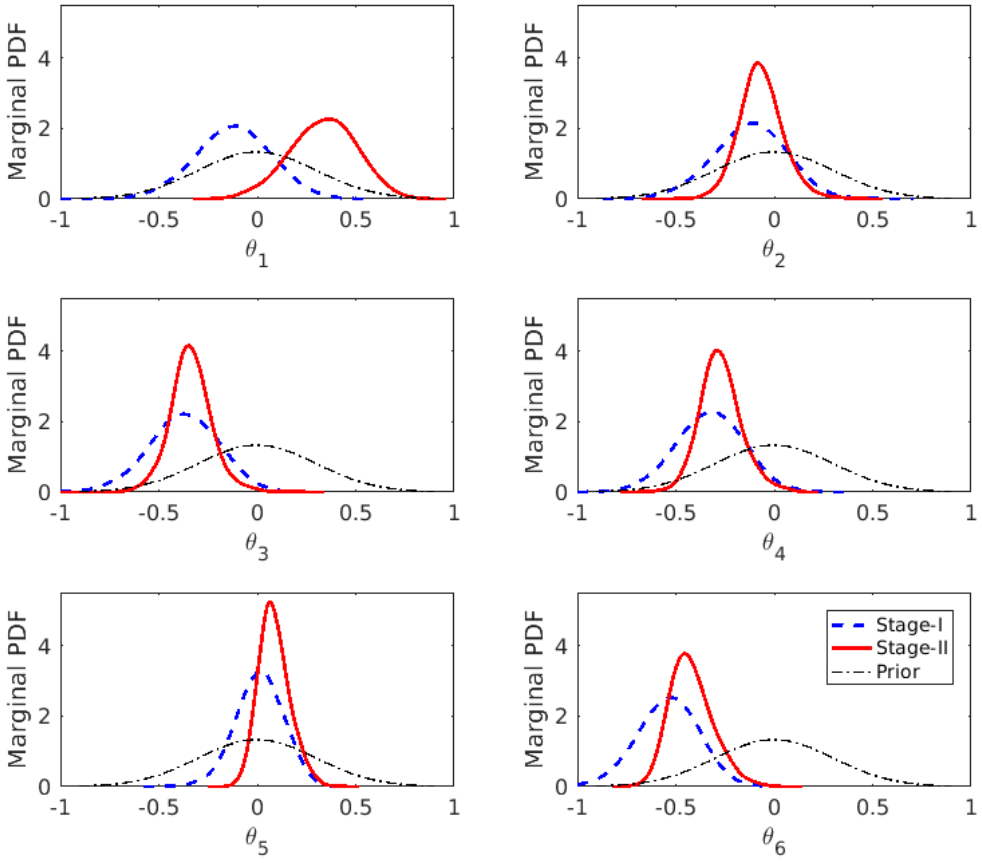


Fig. 12. Prior and marginal posterior PDFs of the two stages for the model class $\mathbb{M}_6$.

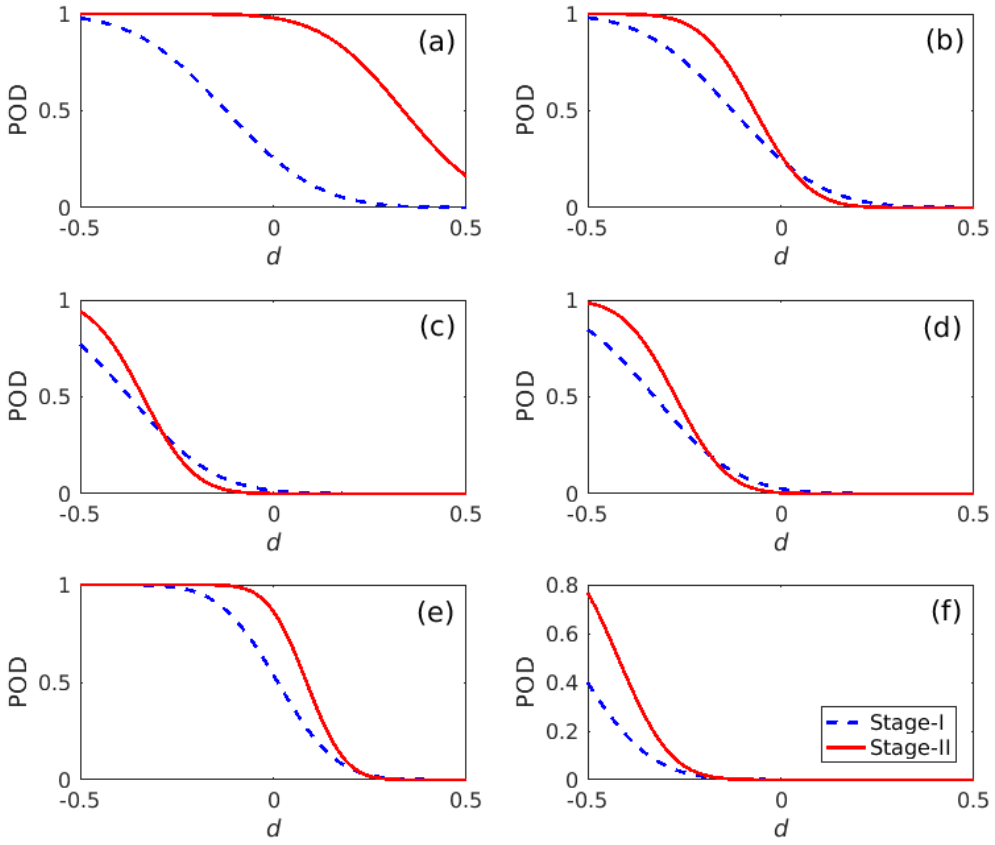


Fig. 13. Probability of damage (POD) of the two stages for the model class $\mathbb{M}_6$: (a) θ_1; (b) θ_2; (c) θ_3; (d) θ_4; (e) θ_5; (f) θ_6

Table 1

Sectional and material properties of the four-storey bolt-connected frame.

Parameter names	Values
Overall height of cross section	100 mm
Flange width of cross section	68 mm
Web thickness of cross section	4.5 mm
Young's modulus	2.00×10^{11} N/m^2
Mass density	7.85×10^{3} kg/m^3
Cross-sectional area	1.43×10^{-3} m^2
Moment of inertia	2.45×10^{-6} m^4

Table 2

Definition of different model classes for the four-storey frame.

Model classes	No. parameters	Parameter names	Element grouping
$\mathbb{M}_1$	1	θ_1	$\{S1-S12\}$
$\mathbb{M}_2$	2	θ_1	$\{S1-S3\}$
		θ_2	$\{S4-S19\}$
$\mathbb{M}_3$	3	θ_1	$\{S1-S3\}$
		θ_2	$\{S4-S11\}$
		θ_3	$\{S12-S19\}$
$\mathbb{M}_4$	4	θ_1	$\{S1,S3\}$
		θ_2	$\{S2\}$
		θ_3	$\{S4-S11\}$
		θ_4	$\{S12-S19\}$
$\mathbb{M}_5$	5	θ_1	$\{S1-S3\}$
		θ_2	$\{S4-S7\}$
		θ_3	$\{S8-S11\}$
		θ_4	$\{S12-S15\}$
		θ_5	$\{S16-S19\}$
$\mathbb{M}_6$	3	θ_1	$\{S1,S3\}$
		θ_2	$\{S2\}$
		θ_3	$\{S4-S19\}$

Table 3

Cases considered for the numerical simulations.

	Case descriptions
Case 1	$k_1 = 0.7k_c,\ k_2 = k_3 = \cdots = k_{19} = k_c$
Case 2	$k_1 = 0.7k_c,\ k_2 = 0.8k_c,\ k_3 = 0.6k_c,\ k_4 = k_5 = \cdots = k_{19} = k_c$
Case 3	$k_4 = 0.7k_c,\ k_5 = 0.8k_c,\ k_8 = 0.6k_c,\ k_9 = 0.8k_c,\ k_{12} = 0.7k_c,\ k_{13} = 0.6k_c$ $k_1 = k_2 = k_3 = k_6 = k_7 = k_{10} = k_{11} = k_{14} = k_{15} = \cdots = k_{19} = k_c$

Table 4

Results of model class selection for each numerical case.

	Model classes	$\boldsymbol{p(\mathbb{D}_N \mid \widehat{\vartheta}_j^c, \mathbb{M}_j)}$	**Log-evidence**	**Data-match**	**Information gain**
Case 1	$\mathbb{M}_1$	**26.85%**	**-197.77**	**-195.53**	**2.24**
	$\mathbb{M}_2$	17.12	-198.22	-194.00	4.22
	$\mathbb{M}_3$	11.59	-198.61	-192.75	5.86
	$\mathbb{M}_4$	15.96	-198.29	-192.28	6.01
	$\mathbb{M}_5$	12.68	-198.52	-191.36	7.16
	$\mathbb{M}_6$	15.80	-198.30	-193.63	4.67
Case 2	$\mathbb{M}_1$	0.01%	-208.28	-205.99	2.29
	$\mathbb{M}_2$	**24.72%**	**-200.47**	**-195.48**	**4.99**
	$\mathbb{M}_3$	23.05%	-200.54	-194.25	6.29
	$\mathbb{M}_4$	20.44%	-200.66	-193.81	6.85
	$\mathbb{M}_5$	12.91%	-201.12	-192.80	8.32
	$\mathbb{M}_6$	18.87%	-200.74	-195.17	5.57
Case 3	$\mathbb{M}_1$	0.00%	-212.81	-211.07	1.74
	$\mathbb{M}_2$	20.95%	-203.22	-199.26	3.96
	$\mathbb{M}_3$	**39.73%**	**-202.58**	**-196.72**	**5.86**
	$\mathbb{M}_4$	14.47%	-203.59	-197.57	6.02
	$\mathbb{M}_5$	17.15%	-203.42	-195.92	7.50
	$\mathbb{M}_6$	7.71%	-204.22	-200.11	4.11

Table 5

Sectional and material properties of the two-storey bolt-connected frame in laboratory.

Parameter names	Values
Overall height of cross section	100 mm
Flange width of cross section	100 mm
Web thickness of cross section	5×10^{-3} m
Flange thickness of cross section	7.5×10^{-3} m
Young's modulus	2.10×10^{11} N/m^2
Mass density	7.85×10^{3} kg/m^3
Cross-sectional area	1.93×10^{-3} m^2
Moment of inertia	3.47×10^{-6} m^4

Table 6

FE and experimental natural frequencies (Hz) for the laboratory two-storey frame.

		Mode 1	Mode 2	Mode 3	Mode 4
FE model		15.31	66.75	106.88	112.34
Experiment	Healthy	15.53	66.27	105.10	111.79
	Damaged	13.98	65.28	104.73	111.29
	Change (%)	-9.98	-1.49	-0.35	-0.45

Table 7

Definition of different model classes for the laboratory two-storey frame.

Model classes	No. parameters	Parameter names	Element grouping
$\mathbb{M}_1$	1	θ_1	$\{S1-S6\}$
$\mathbb{M}_2$	2	$\theta_1,\ \theta_2$	$\{S1,S2\},\ \{S3-S6\}$
$\mathbb{M}_3$	3	$\theta_1,\ \theta_2,\ \theta_3$	$\{S1,S2\},\ \{S3,S4\},\ \{S5,S6\}$
$\mathbb{M}_4$	4	$\theta_1,\ \theta_2,\ \theta_3,\ \theta_4$	$\{S1\},\ \{S2\},\ \{S3,S4\},\ \{S5,S6\}$
$\mathbb{M}_5$	5	$\theta_1,\ \theta_2,\ \theta_3,\ \theta_4,\ \theta_5$	$\{S1,S2\},\ \{S3\},\ \{S4\},\ \{S5\},\ \{S6\}$
$\mathbb{M}_6$	6	$\theta_1,\ \theta_2,\ \theta_3,\ \theta_4,\ \theta_5,\ \theta_6$	$\{S1\},\ \{S2\},\ \{S3\},\ \{S4\},\ \{S5\},\ \{S6\}$

Table 8

Results of model class selection for the laboratory two-storey frame.

Model classes	$p(\mathbb{D}_N \mid \hat{\vartheta}_j^c, \mathbb{M}_j)$	Log-evidence	Data-match	Information gain
$\mathbb{M}_1$	12.70%	-128.17	-127.44	0.73
$\mathbb{M}_2$	**21.42%**	**-127.64**	**-126.34**	**1.30**
$\mathbb{M}_3$	17.34%	-127.86	-125.60	2.25
$\mathbb{M}_4$	15.47%	-127.97	-125.16	2.81
$\mathbb{M}_5$	17.57%	-127.84	-123.88	3.96
$\mathbb{M}_6$	15.52%	-127.97	-123.57	4.39

Table 9

Identified model parameters of various model classes for the laboratory two-storey frame.

Model classes			θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
$\mathbb{M}_1$	Stage-I	MEAN	-0.3128					
		STD	0.0792					
	Stage-II	MEAN	-0.0680					
		STD	0.0763					
$\mathbb{M}_2$	Stage-I	MEAN	-0.1356	-0.3522				
		STD	0.1522	0.0922				
	Stage-II	MEAN	0.1983	-0.2186				
		STD	0.1417	0.0811				
$\mathbb{M}_3$	Stage-I	MEAN	-0.1446	-0.4129	-0.2703			
		STD	0.1528	0.1382	0.1041			
	Stage-II	MEAN	0.1919	-0.3573	-0.1447			
		STD	0.1365	0.0865	0.0844			
$\mathbb{M}_4$	Stage-I	MEAN	-0.1038	-0.1441	-0.4155	-0.2690		
		STD	0.1899	0.1868	0.1373	0.1068		
	Stage-II	MEAN	0.3500	-0.0835	-0.3585	-0.1392		
		STD	0.1666	0.1149	0.0850	0.0846		
$\mathbb{M}_5$	Stage-I	MEAN	-0.1509	-0.3831	-0.3237	0.0137	-0.5345	
		STD	0.1481	0.1789	0.1689	0.1153	0.1510	
	Stage-II	MEAN	0.2054	-0.3444	-0.2711	0.0895	-0.4247	
		STD	0.1372	0.1068	0.0998	0.0788	0.1057	
$\mathbb{M}_6$	Stage-I	MEAN	-0.1223	-0.1238	-0.3717	-0.3269	0.0115	-0.5390
		STD	0.1874	0.1821	0.1735	0.1691	0.1195	0.1535
	Stage-II	MEAN	0.3273	-0.0625	-0.3385	-0.2695	0.0901	-0.4175
		STD	0.1682	0.1131	0.1056	0.1074	0.0797	0.1062