
An efficient method for Bayesian system identification based on Markov chain Monte Carlo simulation

Jia-Hua Yang

Department of Disaster Mitigation for Structures,
College of Civil Engineering,
Tongji University,
1239 Siping Road, Shanghai, China
Email: javayang@tongji.edu.cn

Abstract: This paper proposes an efficient method for identifying a dynamic system using measured accelerations. A practical mathematical model of a dynamic system is developed based on modal superposition for response prediction. To explicitly address uncertainties, system identification is treated as a Bayesian inference problem where the objective is to identify the posterior PDF conditional measured data. Unless a very simple system is considered, the posterior PDF is usually complicated in the sense that its significant region is concentrated in the neighbourhood of an extended and extremely complex manifold. An effective Markov chain Monte Carlo algorithm is developed to sample from the posterior PDF. Given the generated samples, a framework is proposed to systematically consider multiple models whose relative plausibility is quantified by the weightings depending on the PDF values of the samples. It is illustrated that the proposed method can handle both globally identifiable and unidentifiable problems.

Keywords: system identification; Bayesian updating; Markov chain Monte Carlo; MCMC; robust prediction.

Reference to this paper should be made as follows: Yang, J-H. (2019) 'An efficient method for Bayesian system identification based on Markov chain Monte Carlo simulation', *Int. J. Lifecycle Performance Engineering*, Vol. 3, No. 1, pp.20–34.

Biographical notes: Jia-Hua Yang is an Assistant Professor at the Department of Disaster Mitigation for Structures, College of Civil Engineering, Tongji University. He performs fundamental and applied research in advanced Markov chain Monte Carlo simulation, Bayesian inference, reliability and risk analysis, and system identification. Specifically, he develops advanced Markov chain Monte Carlo methods with applications to optimisation problems, Bayesian inference, and robust reliability and risk analysis. He also develops efficient algorithms of system identification (operational modal analysis and model updating) with applications to damage detection, robust response prediction, load reconstruction, and wind and earthquake engineering.

1 Introduction

System identification aims at identifying a mathematical model based on measured data for representing a dynamic system so that actual behaviours of this system can be predicted as accurately as possible. The main task comes down to construct a mathematical model with some parameters, then conduct an optimisation process to tune the parameters for minimising the difference between measured and model-predicted response. In deterministic approaches, assuming that the true model of a dynamic system exists, only an optimal model is pinpointed for representing the dynamic system, but uncertainties in the identified model due to measurement noise and modelling error are not accounted for. The effect of uncertainties could bring a large challenge to system identification, which is especially so in practice. This is because in practice, unless the dynamic system under consideration is very simple (but it is seldom the case, e.g., dynamic systems of full-scale structures are not simple), we usually need to identify complicated models with only limited measured data, and the identified models will have relatively large uncertainties, which result non-uniqueness in system identification, i.e., multiple models that can fit the measured data equivalently well exist. Subjective arguments have been used to keep only one model while ruling out all others. However, there seems no theoretical basis ignoring any of these models, since all of them perform well in reproducing the measured data. In addition, ignoring the multiple equivalent models could bias response prediction. An appropriate framework is thus needed to address uncertainties in system identification.

A Bayesian framework has been developed to explicitly address uncertainties in system identification (Beck and Katafygiotis, 1998). A class of system models is first chosen. It is then embedded in a class of probability models that describe the uncertainties due to measurement noise and modelling error. The posterior probability density function (PDF) of a system is derived following Bayes' theorem. System identification is then viewed as a Bayesian inference problem where measured data (given information) is used to update the posterior PDF, which describe the relative plausibility of each system model in a class of models (Ng, 2014; Zhang et al., 2017; Ni and Zhang, 2018; He and Ng, 2017; Yin et al., 2017; Yin et al., 2010). This framework lays down a theoretical basis to handling uncertainties and non-uniqueness, i.e., systematically consider multiple equivalent models while consistent with modelling assumptions. A disadvantage by following the Bayesian framework is that a large amount of computational time is required, especially when identifying the posterior PDF of an unidentifiable problem, where the region of significant probabilities lies on an extended and extremely complex manifold. Some algorithms have been developed to successfully find complex manifolds (Katafygiotis and Beck, 1998; Katafygiotis and Lam, 2002; Katafygiotis et al., 2000, 1998). Efficiently exploring the complex manifold of the posterior PDF can be also done by sampling from the posterior PDF using Monte Carlo simulation. Because the posterior PDF is not a PDF that is convenient to sample from, e.g., Gaussian PDF, directly sampling from the posterior PDF is difficult. A possible solution is to apply importance sampling, where samples are generated from an importance density that is convenient to sample from, with importance weightings quantifying how close each sample lies in the region of significant probabilities. However, the problem is that an importance density that is well approximated to the posterior PDF is very difficult to find, so most of the time samples generated by

importance sampling lie in the region of small probabilities and it will take a long time until enough samples can be used to approximate the posterior PDF. To solve this problem, Markov chain Monte Carlo (MCMC) algorithms based on simulated annealing have been developed and find successful applications (Beck and Au, 2002; Ching and Chen, 2007; Beck and Zuev, 2013; Zuev and Beck, 2013; Lam et al., 2015; Yuen et al., 2004; Lam et al., 2018; Yang et al., 2015; Hu and Yang, 2019). The idea is to decompose sampling into multiple levels (Yang and Lam, 2018) so that the size of parameter space focused in each level is reduced gradually, which makes sure that complicated manifolds can be fully explored. This paper also follows the idea of these MCMC algorithms and proposes an efficient method of system identification for both identifiable and unidentifiable problems.

2 Methodology

2.1 Mathematical model of a dynamic system

A mathematical model of a dynamic system is needed to predict the response of the system so that the predicted response can be fit to measured response for identifying the mathematical model, and thus the dynamic system. Response of a structural dynamic system under dynamic loads is governed by the equation of motion:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t) \quad (1)$$

where $\mathbf{M} \in R^{N_d \times N_d}$, $\mathbf{C} \in R^{N_d \times N_d}$ and $\mathbf{K} \in R^{N_d \times N_d}$ denote the mass, damping and stiffness matrices of the structural system, respectively; N_d is the number of degrees of freedom; $\ddot{\mathbf{x}} \in R^{N_d}$, $\dot{\mathbf{x}} \in R^{N_d}$ and $\mathbf{x} \in R^{N_d}$ denote the acceleration, velocity and displacement of the system, respectively, i.e., the dynamic response; $\mathbf{f} \in R^{N_d}$ is the excitation to the system. Given material properties and topological information of the system, \mathbf{M} , \mathbf{C} and \mathbf{K} can be constructed conveniently using finite element method. \mathbf{f} is assumed to be known from measurement. Equation (1) can then be solved directly using time integration method such as the Newmark-beta method (Au, 2017). However, this will cost large computational power, because the system matrices in practical problems are in large size and response needs to be calculated for many number of time steps. For the efficient application in practical problems, the modal superposition method is used to reduce the size of equation (1). The idea is that in practice usually only the first several modes contribute to system response, so instead of solving for a full system, the eigenvalue problem related to equation (1) is first solved to get natural frequencies, damping ratios and mode shapes. Considering only the first N_n dominant modes, equation (1) is then transformed to modal coordinates, rendering solving N_n uncoupled single-degree-of-freedom (SDOF) equations of motion, the n^{th} of which is:

$$\ddot{q}_n(t) + 2\zeta_n\omega_n\dot{q}_n(t) + \omega_n^2q_n(t) = a_n(t) \quad (2)$$

where q_n is the n^{th} modal coordinate; ω_n and ζ_n are the natural frequency and damping ratio of the n^{th} mode, respectively; a_n is the modal acceleration excitation in the n^{th} modal coordinate given by

$$a_n(t) = \frac{p_n(t)}{m_n} \quad (3)$$

p_n is the modal excitation of the n^{th} mode

$$p_n(t) = \boldsymbol{\varphi}_n^T \mathbf{f}(t) \quad (4)$$

where $\boldsymbol{\varphi}_n$ is the mode shape of the n^{th} mode. m_n is the modal mass of the n^{th} mode

$$m_n = \boldsymbol{\varphi}_n^T \mathbf{M} \boldsymbol{\varphi}_n \quad (5)$$

After solving equation (2) for the considered N_m modes, the predicted system acceleration response is obtained by superposing the modal acceleration response:

$$\ddot{\mathbf{x}}(t) = \sum_{n=1}^{N_m} \boldsymbol{\varphi}_n \ddot{q}_n(t) \quad (6)$$

Solving these SDOF equations of motion is much more efficient than solving the original full system. The increase of computational efficiency makes it possible for practical applications.

2.2 Bayesian system identification

2.2.1 Posterior PDF

System identification is viewed as a Bayesian inference problem where the posterior PDF of the uncertain parameters of a dynamic system is identified. This section derives the posterior PDF by embedding the system model in a probability model based on measured data. Let the measured data be $D = \{\hat{\mathbf{x}}_t \in R^{N_o}, \hat{\mathbf{f}}_t \in R^{N_f} : t = 1, 2, \dots, N_t\}$, where $\hat{\mathbf{x}}_t$ is the vector containing accelerations at the t^{th} time step measured at N_o DOFs; $\hat{\mathbf{f}}_t$ is the vector containing excitation at the t^{th} time step measured at N_f DOFs; and N_t is the total number of time steps. If the model of a system described in the previous section is parameterised by a model parameter vector $\boldsymbol{\theta}$, let $\ddot{\mathbf{x}}_t(\boldsymbol{\theta}) \in R^{N_d}$ denote the model-predicted acceleration at the time step t , and assume the model-predicted acceleration and the measured one differ with an error:

$$\hat{\mathbf{x}}_t = \mathbf{L} \ddot{\mathbf{x}}_t(\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_t \quad (7)$$

where $\mathbf{L} \in R^{N_o \times N_d}$ is a selection matrix used to consider the situation that only part of the DOFs of a model can be measured in practice and it has only one non-zero element, equal to unity in each row; and $\boldsymbol{\varepsilon}_t \in R^{N_o}$ is the prediction error at the time step t . According to Bayes' theorem, the posterior PDF of $\boldsymbol{\theta}$ is written as

$$p(\boldsymbol{\theta} | \mathbf{D}) = \frac{p(\boldsymbol{\theta}) p(\mathbf{D} | \boldsymbol{\theta})}{p(\mathbf{D})} \quad (8)$$

where $p(\boldsymbol{\theta} | \mathbf{D})$ is the posterior PDF of $\boldsymbol{\theta}$ that quantifies the relative plausibility of different models in the parameter space conditional on the measured data \mathbf{D} ; $p(\boldsymbol{\theta})$ is the prior PDF of $\boldsymbol{\theta}$ that gives a measure of the initial relative plausibility of the models before

any data is given; $p(\mathbf{D} | \boldsymbol{\theta})$ is the likelihood function that expresses the probability of getting data \mathbf{D} based on the mathematical model; and is called the evidence of the mathematical model class given the data:

$$p(\mathbf{D}) = \int_{\Theta} p(\boldsymbol{\theta}) p(\mathbf{D} | \boldsymbol{\theta}) d\boldsymbol{\theta} \quad (9)$$

and it works as a normalising constant so that the integration of $p(\boldsymbol{\theta} | \mathbf{D})$ over the parameter space Θ equals to unity. If we want the inference to fully rely on the measured data, $p(\boldsymbol{\theta})$ is taken to be a uniform PDF, so equation (8) can be written as

$$p(\boldsymbol{\theta} | \mathbf{D}) = c p(\mathbf{D} | \boldsymbol{\theta}) \quad (10)$$

where c is a constant. To formulate the likelihood function, the mathematical model is embedded into a probabilistic model by assuming that the prediction errors at different DOFs and different time steps follow i.i.d (independent and identically distributed) zero-mean normal distributions, $\mathcal{N}(0, \sigma^2)$, with a mean 0 and a variance σ^2 , so

$$\rho(\mathbf{D} | \boldsymbol{\theta}) = \prod_{t=1}^{N_t} \prod_{i=1}^{N_o} p(\hat{x}_{i,t} | \boldsymbol{\theta}) \quad (11)$$

where $\hat{x}_{i,t}$ is the measured acceleration of the i^{th} DOF at the t^{th} time step and

$$p(\hat{x}_{i,t} | \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (\hat{x}_{i,t} - \ddot{x}_{i,t}(\boldsymbol{\theta}))^2\right) \quad (12)$$

The posterior PDF is thus obtained:

$$p(\mathbf{D} | \boldsymbol{\theta}) = c_0 \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^{N_t} \sum_{i=1}^{N_o} (\hat{x}_{i,t} - \ddot{x}_{i,t}(\boldsymbol{\theta}))^2\right) \quad (13)$$

where c_0 absorbs all the constants.

2.2.2 Markov chain Monte Carlo

Bayesian system identification is to identify the posterior PDF in equation (13). It is almost impossible to obtain the close-form formulation for the posterior PDF, so Monte Carlo simulation is resorted to for approximating the posterior PDF. There are mainly two problems. First, it is difficult to describe the complicated topology of the posterior PDF in the vast parameter space, i.e., how to generate samples in regions with high probability. Second, it is difficult to generate samples from the posterior PDF, because directly sampling from equation (13) is simply impossible. To solve the first problem, a strategy based on simulated annealing is adopted (Beck and Au, 2002; Lam et al., 2015, 2018). The idea is to set-up a sequence of bridge PDFs, $p_1, p_2, \dots, p_g, \dots$, with each of their regions of high probability reduced gradually. Sampling is then conducted in multiple levels, and in each level g , samples are generated according to p_g . According to equation (13), the bridge PDF p_g is constructed with a form of equation (13):

$$p_g = c_g \exp\left(-\frac{1}{2\sigma_g^2} \sum_{t=1}^{N_t} \sum_{i=1}^{N_o} (\hat{x}_{i,t} - \ddot{x}_{i,t}(\boldsymbol{\theta}))^2\right) \quad (14)$$

with the variance σ_g^2 controlling the parameter space covered by p_g . The change of σ_g^2 between two levels is as follows so that the parameter space covered by p_g is reduced gradually:

$$\sigma_g^2 = \frac{\sigma_{g-1}^2}{b} \quad (15)$$

where b is an algorithmic constant larger than 1. Then, samples are drawn from equation (14) in each level.

Considering that p_g is known up to a constant, the Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970) is suitable for sampling. To apply the MH algorithm in the g^{th} level, first, a proposal PDF $q(\boldsymbol{\theta}|\boldsymbol{\theta}_g^{h-1})$ that is convenient to sample from is constructed to draw a candidate $\boldsymbol{\theta}^*$ of the current sample $\boldsymbol{\theta}_g^{(h)}$ conditional on the previous sample $\boldsymbol{\theta}_g^{(h-1)}$, where h is the sample index. The candidate is accepted subjected to the acceptance ratio r in the following:

$$r = \frac{p_g(\boldsymbol{\theta}^*)/q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_g^{(h-1)})}{p_g(\boldsymbol{\theta}_g^{(h-1)})/q(\boldsymbol{\theta}_g^{(h-1)}|\boldsymbol{\theta}^*)} \quad (16)$$

where $q(\boldsymbol{\theta}_g^{(h-1)}|\boldsymbol{\theta}^*)$ is the PDF of the previous sample $\boldsymbol{\theta}_g^{(h-1)}$ conditional on the candidate. If $r > 1$, the candidate is accepted as the current sample, i.e., $\boldsymbol{\theta}_g^{(h)} = \boldsymbol{\theta}^*$. If $r < 1$, the candidate is accepted as the current sample with probability r ; and with probability $1 - r$, the candidate is rejected and the previous sample is taken as the current sample, i.e., $\boldsymbol{\theta}_g^{(h)} = \boldsymbol{\theta}_g^{(h-1)}$. The MH algorithm will be used in the g^{th} level to draw the desired N_g samples.

An important issue of the MH algorithm is to construct the proposal PDF $q(\boldsymbol{\theta}|\boldsymbol{\theta}_g^{h-1})$. A bad proposal PDF will generate most of the samples in regions of low probability, which in turn will give a bad approximation of the posterior PDF. It is proposed to construct the proposal PDF using the previous samples, so that the characteristics of the posterior PDF are tracked and updated continuously. To be specific, given the samples of the previous level $g - 1$, $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$, the proposal PDF used in the current level g is constructed as a kernel density (Beck and Au, 2002; Lam et al., 2015, 2018; Yang et al., 2015):

$$q(\boldsymbol{\theta}|\boldsymbol{\theta}_g^{(h-1)}) = \sum_{h=1}^{N_g} W_{g-1}^{(h)} \mathcal{N}(\boldsymbol{\theta}_{g-1}^{(h)}, \mathbf{C}_{g-1}^{(h)}) \quad (17)$$

Equation (17) is a weighted sum of Gaussian PDFs with each Gaussian PDF centred at the h^{th} sample of the previous level $g - 1$ (i.e., $\boldsymbol{\theta}_{g-1}^{(h)}$), the sample covariance matrix of

each Gaussian PDF (i.e., $\mathbf{C}_{g-1}^{(h)}$) calculated using $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$, and the weighting of each Gaussian PDF calculated using p_{g-1} that describes the relative plausibility of each sample in $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$. Given enough samples, the kernel density in equation (17) can well approximate the bridge PDF in each level. Note that the weightings $\{W_{g-1}^{(h)} : h = 1, 2, \dots, N_g\}$ are normalised so that $\sum_{h=1}^{N_g} W_{g-1}^{(h)} = 1$, so to draw a candidate from $q(\boldsymbol{\theta} | \boldsymbol{\theta}_g^{(h-1)})$ using equation (17), first, draw an index $j \in \{1, 2, \dots, N_g\}$ with the corresponding probability $\{W_{g-1}^{(h)} : h = 1, 2, \dots, N_g\}$. Second, draw the candidate $\boldsymbol{\theta}^*$ from the Gaussian PDF $N(\boldsymbol{\theta}_{g-1}^{(j)}, \mathbf{C}_{g-1}^{(j)})$.

The above procedures of constructing the proposal PDF are applicable for a general level g except for the 1st level, because there are no samples yet for the 1st level. The proposal PDF in the 1st level is, thus, chosen as a uniform PDF. The overall procedures of the proposed MCMC method are summarised as follows.

- 1 In the 1st level, σ_1^2 can be set to the variance of the measured accelerations. N_g samples are drawn according to the PDF in equation (14) using the MH algorithm with the proposal PDF being a uniform PDF. The MH algorithm work as follows.

Given the previous sample $\boldsymbol{\theta}_1^{(h-1)}$,

- a generate a candidate $\boldsymbol{\theta}^* \sim q(\boldsymbol{\theta} | \boldsymbol{\theta}_1^{(h-1)})$, where $q(\boldsymbol{\theta} | \boldsymbol{\theta}_1^{(h-1)})$ is a uniform PDF
- b take $\boldsymbol{\theta}_1^{(h)} = \boldsymbol{\theta}^*$ with probability $\rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_1^{(h-1)})$, and $\boldsymbol{\theta}_1^{(h)} = \boldsymbol{\theta}_1^{(h-1)}$ with probability $1 - \rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_1^{(h-1)})$, where $\rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_1^{(h-1)}) = \min\left(\frac{p_1(\boldsymbol{\theta}^*)/q(\boldsymbol{\theta}^* | \boldsymbol{\theta}_1^{(h-1)})}{p_1(\boldsymbol{\theta}_1^{(h-1)})/q(\boldsymbol{\theta}_1^{(h-1)} | \boldsymbol{\theta}^*)}, 1\right)$
- c repeat steps a and b until N_g samples are generated.

- 2 In a general level g , N_g samples are drawn according to the PDF in equation (14) using the MH algorithm with the proposal PDF constructed by equation (17), based on the samples in the previous level $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$. The MH algorithm works as follows.

Given the previous sample $\boldsymbol{\theta}_g^{(h)}$ and the samples in the previous level $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$,

- a Construct the proposal PDF $q(\boldsymbol{\theta} | \boldsymbol{\theta}_g^{(h-1)})$ according to equation (17) using $\{\boldsymbol{\theta}_{g-1}^{(h)}\}$.
- b Generate a candidate $\boldsymbol{\theta} \sim q(\boldsymbol{\theta} | \boldsymbol{\theta}_g^{(h-1)})$. To do this,
 - draw an index $j \in \{1, 2, \dots, N_g\}$ with the corresponding probability $\{W_{g-1}^h : h = 1, 2, \dots, N_g\}$
 - draw the candidate $\boldsymbol{\theta}^*$ from the Gaussian PDF $N(\boldsymbol{\theta}_{g-1}^{(j)}, \mathbf{C}_{g-1}^{(j)})$

- c Take $\boldsymbol{\theta}_g^{(h)} = \boldsymbol{\theta}^*$ with probability $\rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_g^{(h-1)})$, and $\boldsymbol{\theta}_g^{(h)} = \boldsymbol{\theta}_g^{(h-1)}$ with probability $1 - \rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_g^{(h-1)})$, where $\rho(\boldsymbol{\theta}^*, \boldsymbol{\theta}_g^{(h-1)}) = \min\left(\frac{p_1(\boldsymbol{\theta}^*)/q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_g^{(h-1)})}{p_1(\boldsymbol{\theta}_g^{(h-1)})/q(\boldsymbol{\theta}_g^{(h-1)}|\boldsymbol{\theta}^*)}, 1\right)$.
- d repeat steps b and c until N_g samples are generated.
- 3 Update the level index $g = g + 1$ and then the variance $\sigma_g^2 = \sigma_{g-1}^2/b$.
- 4 Repeat steps (2) and (3) until convergence. This is reflected by which the samples are stable within certain regions and the predicted accelerations can match the measured ones.
- 5 The posterior PDF of the uncertain parameters is approximated by the kernel density constructed by equation (17) using the samples in the final level L . The predicted accelerations can be obtained by the weighted sum of the accelerations calculated by each sample of $\boldsymbol{\theta}$, i.e.,

$$\ddot{\mathbf{x}}_t = \sum_{h=1}^{N_g} W_L^{(h)} \ddot{\mathbf{x}}_t(\boldsymbol{\theta}_L^{(h)}) \quad (18)$$

where the subscript L indicates that the samples in the final level are used and the weightings $\{W_L^{(h)}\}$ are obtained when the kernel density is constructed.

2.3 Case studies

Two numerical case studies were conducted to illustrate that the proposed method can efficiently handle system identification, for both identifiable and unidentifiable problems. In the first case, Bayesian system identification was conducted for a simple two-story shear building. It is a globally identifiable problem. In the second case, a complicated three-dimension steel tower was considered to illustrate that the proposed method can also handle unidentifiable problems.

2.3.1 Bayesian identification of a two-story shear building

A two-story shear building with story masses $m_1 = m_2 = 1$ kg, damping ratios $\zeta_1 = \zeta_2 = 2\%$ and inter-story stiffness $k_1 = k_2 = 1,000$ N/m is considered in this case. The accelerations were simulated by inputting the El Centro earthquake acceleration (see Figure 1) to the foundation. The accelerations were calculated with two modes considered and the time step being 0.02 s. Gaussian noise with the noise level 3% was then added to the accelerations to simulate the real situation. The level of the added noise is expressed as the percentage ratio of its standard deviation over the standard deviation of the model acceleration. It was assumed that the accelerations at the two stories were measured for system identification. The uncertain parameters to be identified were $\boldsymbol{\theta}(1)$ and $\boldsymbol{\theta}(2)$ that scale the nominal stiffness k_1 and k_2 , respectively, while the masses and damping ratios were assumed to be known. Note that $\boldsymbol{\theta}(1)$ and $\boldsymbol{\theta}(2)$ are the two elements of the vector $\boldsymbol{\theta}$. With the measured accelerations at the two stories and the earthquake input, the MCMC

method was applied for Bayesian system identification. Figure 2 shows the samples of θ in the final level, based on which the posterior marginal PDFs could be constructed (see Figure 3). It is observed that with the amount of data large enough, θ can be globally identified (Beck and Katafygiotis, 1998). The predicted accelerations at the two stories were then calculated using the samples based on equation (18) and compared to the measured accelerations. The comparison is shown in Figure 4. It can be seen that the predicted accelerations (the red dash lines) almost overlap with the measured ones (the blue solid lines), indicating the good performance of the proposed method.

Figure 1 El Centro earthquake input to the shear building (see online version for colours)

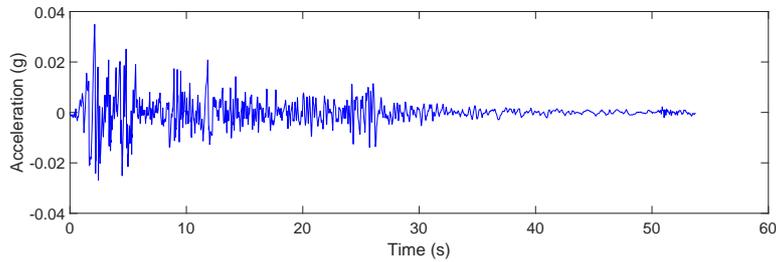


Figure 2 The samples of θ in the final level for the shear building (see online version for colours)

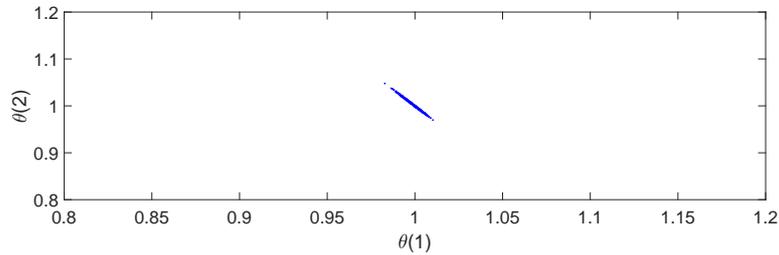


Figure 3 The posterior marginal PDFs for the shear building (see online version for colours)

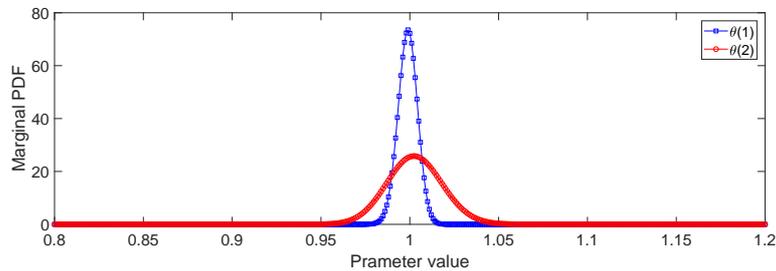
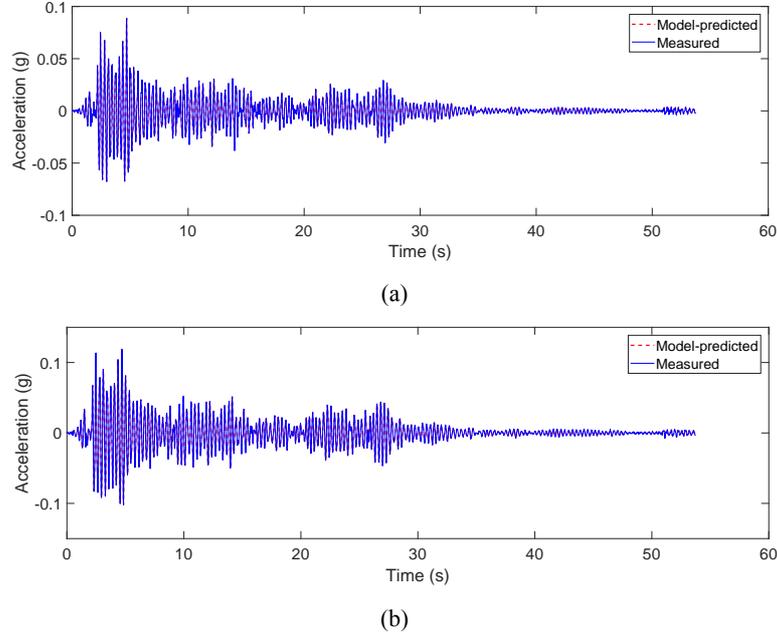


Figure 4 Comparison of the predicted and measured accelerations, (a) story 1 (b) story 2 (see online version for colours)



2.3.2 Bayesian identification of a three-dimension steel tower

In this case, a complicated and practical structure, a three-dimension steel tower (see Figure 5), was considered to illustrate that the proposed method can also handle the challenging unidentifiable problems. This steel tower was also studied in Yang et al. (2015) and its dimension and material properties can be found there. Here, the measured data was numerically simulated by inputting the earthquake record in Figure 1 to the finite element model of this tower. The earthquake excitation was applied at the foundation and along the x direction. To simulate the accelerations, ten modes were considered with 2% modal damping ratio used for each mode. The time step was set to 0.02 second. Gaussian noise with the noise level 3% was then added to the simulated accelerations to give the measured data. For Bayesian system identification, the vector $\boldsymbol{\theta}$ includes six uncertain parameters, with $\boldsymbol{\theta}(1)$ to $\boldsymbol{\theta}(5)$ scaling the Young's modulus of the columns of floors 1 to 5, respectively, and $\boldsymbol{\theta}(6)$ scaling the Young's modulus of all the columns of floors 6 to 8. The arrangement of the uncertain parameters can be also seen in Figure 5. It was assumed that the x-direction accelerations at points 2 to 9 and the earthquake excitations were measured. It was also assumed that the modal damping ratios were known for system identification. The MCMC method was applied to update the finite element model, and the posterior marginal PDFs of the six uncertain parameters were constructed based on the samples in the final sampling level (see Figure 6). It is observed that the posterior marginal PDFs of $\boldsymbol{\theta}(1)$, $\boldsymbol{\theta}(3)$, $\boldsymbol{\theta}(4)$, and $\boldsymbol{\theta}(5)$ have clear peaks and these four parameters are globally identifiable. The posterior marginal PDFs of $\boldsymbol{\theta}(2)$ and $\boldsymbol{\theta}(6)$ are flat across a wide region of the parameter space, indicating that an infinite number of models that have almost equal plausibility is available for response prediction,

and these two parameters are unidentifiable. This tower has many DOFs but only several DOFs were measured. The measured data may not contain enough information for globally identifying $\theta(2)$ and $\theta(6)$. It would be difficult to pinpoint a single model [equivalently pinpoint a single value for $\theta(2)$ or $\theta(6)$] for predicting the accelerations, because all the models in the parameter space have almost equal plausibility. Moreover, artificially pinpointing a single model will bias the predicted accelerations. The proposed method provides a rigorous theoretical basis to consider all equally important models in the parameter space by including all generated samples for response prediction, so the difficult choice of pinpointing a single model is not required.

Figure 5 The finite element model of the steel tower [19] (see online version for colours)

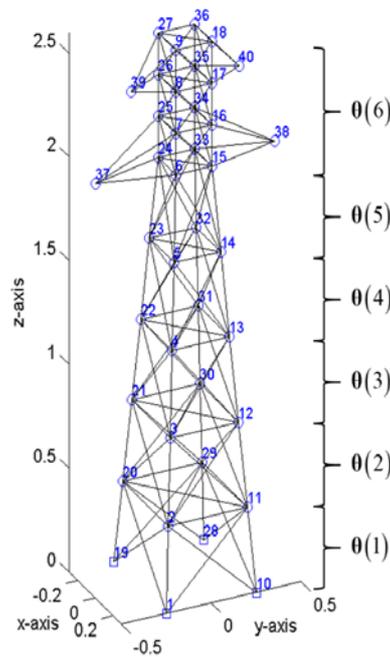
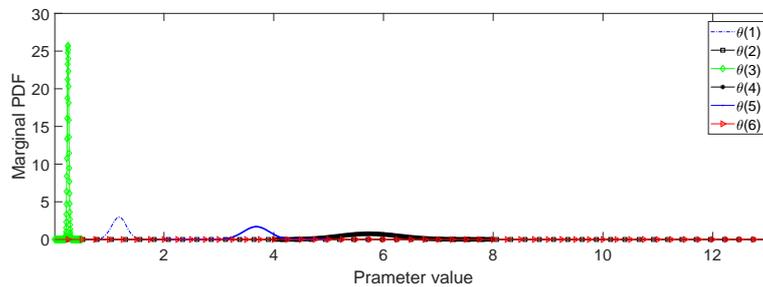


Figure 6 The posterior marginal PDFs for the steel tower (see online version for colours)



By using equation (18), the accelerations were calculated and compared to the measured ones. It can be seen in Figure 7 that the predicted accelerations can match the measured ones quite well. This shows that the proposed method can handle system identification

for unidentifiable problems. To demonstrate the superiority of the proposed method over the methods that pinpoint one single model, a model was arbitrarily selected (i.e., a sample of θ) in the parameter space and only this model was used for response prediction. The predicted acceleration at point 7 by this model was compared to the corresponding measured one in Figure 8. When comparing Figure 8 to Figure 7(c), observable bias of the predicted acceleration is noticed. This also illustrates that if the multiple equally-important models in an unidentifiable problem are not properly considered, response prediction will be not reliable.

Figure 7 Comparison of the predicted and measured accelerations, (a) x direction at point 3 (b) x direction at point 5 (c) x direction at point 7 (d) x direction at point 9 (see online version for colours)

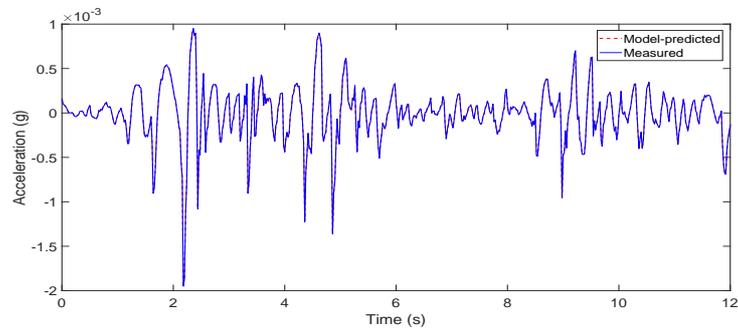
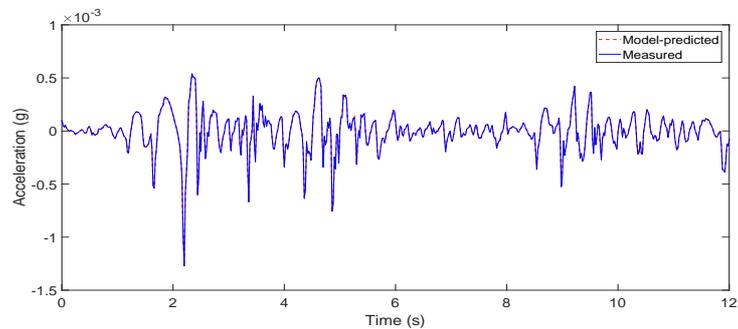
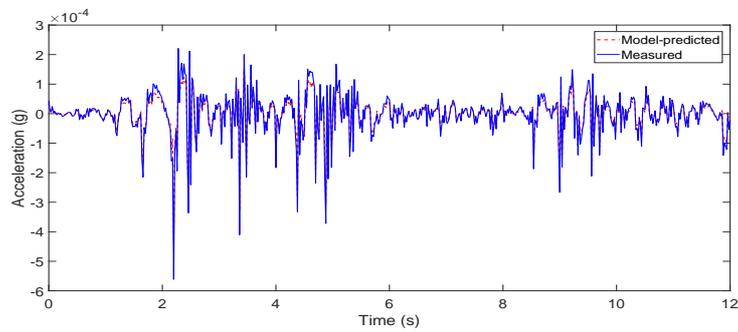
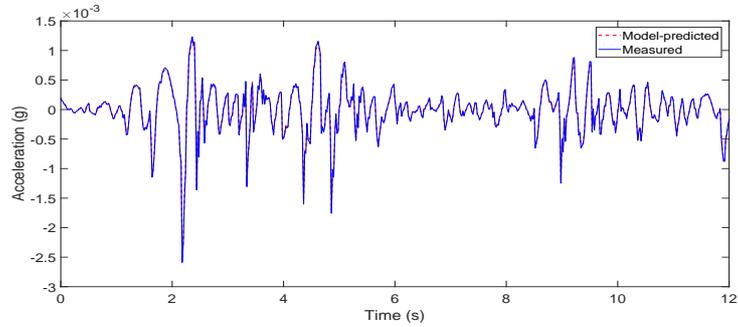
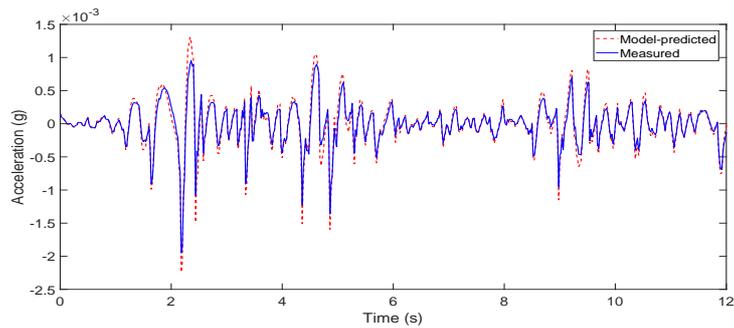


Figure 7 Comparison of the predicted and measured accelerations, (a) x direction at point 3 (b) x direction at point 5 (c) x direction at point 7 (d) x direction at point 9 (see online version for colours) (continued)



(d)

Figure 8 Comparison of the x-direction acceleration at point 7 using an arbitrary model (see online version for colours)



3 Concluding remarks

An efficient MCMC method Bayesian system identification using time-domain data is developed in this paper. The mathematical model of a dynamic system based on modal superposition is first derived so that dynamic response can be efficiently calculated for practical problems. After deriving the posterior PDF of the uncertain parameters, a multi-level sampling method based on the idea of simulated annealing is developed. This sampling method ensures that the high-dimension parameter space can be efficiently explored. Two numerical cases are used to verify the proposed method. The first case is Bayesian system identification of a two-story shear building, which illustrates that the proposed method can handle globally identifiable problems. The second case is Bayesian system identification of a three-dimension steel tower, which emphasises the importance of using the proposed method for unidentifiable problems.

Acknowledgements

The author is funded by National Natural Science Foundation of China (Grant No. 51808400), Shanghai Sailing Program (Grant No. 18YF1424500), and the Fundamental Research Funds for the Central Universities (Grant No. 22120180007). The generous support is gratefully acknowledged.

References

- Au, S.K. (2017) *Operational Modal Analysis: Modeling, Bayesian Inference, Uncertainty Laws*, Springer, Singapore.
- Beck, J.L. and Au, S.K. (2002) 'Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation', *Journal of Engineering Mechanics*, Vol. 128, No. 4, pp.380–391.
- Beck, J.L. and Katafygiotis, L.S. (1998) 'Updating models and their uncertainties: I: Bayesian statistical framework', *Journal of Engineering Mechanics*, Vol. 124, No. 4, pp.455–461.
- Beck, J.L. and Zuev, K.M. (2013) 'Asymptotically independent Markov sampling: a new Markov chain Monte Carlo scheme for Bayesian inference', *International Journal for Uncertainty Quantification*, Vol. 3, No. 5, pp.445–474.
- Ching, J. and Chen, Y.C. (2007) 'Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging', *Journal of Engineering Mechanics*, Vol. 133, No. 7, pp.816–832.
- Hastings, W.K. (1970) 'Monte Carlo sampling methods using Markov chains and their applications', *Biometrika*, Vol. 57, No. 1, pp.97–109.
- He, S. and Ng, C.T. (2017) 'Guided wave-based identification of multiple cracks in beams using a Bayesian approach', *Mechanical Systems and Signal Processing*, Vol. 84, Part A, pp.324–345.
- Hu, J. and Yang, J.H. (2019) 'Operational modal analysis and Bayesian model updating of a coupled building', *International Journal of Structural Stability and Dynamics*, Vol. 19, No. 1, p.1940012.
- Katafygiotis, L.S. and Beck, J.L. (1998) 'Updating models and their uncertainties: II: model identifiability', *Journal of Engineering Mechanics*, Vol. 124, No. 4, pp.463–467.
- Katafygiotis, L.S. and Lam, H.F. (2002) 'Tangential-projection algorithm for manifold representation in unidentifiable model updating problems', *Earthquake Engineering & Structural Dynamics*, Vol. 31, No. 4, pp.791–812.
- Katafygiotis, L.S., Lam, H.F. and Papadimitriou, C. (2000) 'Treatment of unidentifiability in structural model updating', *Advances in Structural Engineering*, Vol. 3, No. 1, pp.19–39.
- Katafygiotis, L.S., Papadimitriou, C. and Lam, H.F. (1998) 'A probabilistic approach to structural model updating', *Soil Dynamics and Earthquake Engineering*, Vol. 17, Nos. 7–8, pp.495–507.
- Lam, H.F., Yang, J.H. and Au, S.K. (2015) 'Bayesian model updating of a coupled-slab system using field test data utilizing an enhanced Markov chain Monte Carlo simulation algorithm', *Engineering Structures*, Vol. 102, pp.144–155.
- Lam, H.F., Yang, J.H. and Au, S.K. (2018) 'Markov chain Monte Carlo-based Bayesian method for structural model updating and damage detection', *Structural Control and Health Monitoring*, Vol. 25, No. 4, p.e2140.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) 'Equation of state calculations by fast computing machines', *The Journal of Chemical Physics*, Vol. 21, No. 6, pp.1087–1092.
- Ng, C.T. (2014) 'On the selection of advanced signal processing techniques for guided wave damage identification using a statistical approach', *Engineering Structures*, Vol. 67, pp.50–60.

- Ni, Y.C. and Zhang, F.L. (2018) 'Fast Bayesian approach for modal identification using forced vibration data considering the ambient effect', *Mechanical Systems and Signal Processing*, Vol. 105, pp.11–128.
- Yang, J.H. and Lam, H.F. (2018) 'An efficient adaptive sequential Monte Carlo method for Bayesian model updating and damage detection', *Structural Control and Health Monitoring*, Vol. 25, No. 12, p.e2260.
- Yang, J.H., Lam, H.F. and Hu, J. (2015) 'Ambient vibration test, modal identification and structural model updating following Bayesian framework', *International Journal of Structural Stability and Dynamics*, Vol. 15, No. 7, p.1540024.
- Yin, T., Jiang, Q.H. and Yuen, K.V. (2017) 'Vibration-based damage detection for structural connections using incomplete modal data by Bayesian approach and model reduction technique', *Engineering Structures*, Vol. 132, pp.260–277.
- Yin, T., Lam, H.F. and Chow, H.M. (2010) 'A Bayesian probabilistic approach for crack characterization in plate structures', *Computer-Aided Civil and Infrastructure Engineering*, Vol. 25, No. 5, pp.375–386.
- Yuen, K.V., Beck, J.L. and Au, S.K. (2004) 'Structural damage detection and assessment by adaptive Markov chain Monte Carlo simulation', *Structural Control and Health Monitoring*, Vol. 11, No. 4, pp.327–347.
- Zhang, F.L., Ni, Y.C. and Lam, H.F. (2017) 'Bayesian structural model updating using ambient vibration data collected by multiple setups', *Structural Control and Health Monitoring*, Vol. 24, No. 12, p.e2023.
- Zuev, K.M. and Beck, J.L. (2013) 'Global optimization using the asymptotically independent Markov sampling method', *Computers & Structures*, Vol. 126, pp.107–119.