Research Article

An Improved Bayesian Structural Identification Using the First Two Derivatives of Log-Likelihood Measure

Jin Zhou, Akira Mita, and Liu Mei

Department of System Design Engineering, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan

Correspondence should be addressed to Jin Zhou; zhoujin5120@gmail.com

Received 28 November 2014; Accepted 3 March 2015

Academic Editor: Elio Sacco

Copyright © 2015 Jin Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The posterior density of structural parameters conditioned by the measurement is obtained by a differential evolution adaptive Metropolis algorithm (DREAM). The surface of the formal log-likelihood measure is studied considering the uncertainty of measurement error to illustrate the problem of equifinality. To overcome the problem of equifinality, the first two derivatives of the log-likelihood measure are proposed to formulate a new informal likelihood measure for the sake of improving the accuracy of the estimator. Moreover, the proposed measure also reduces the standard deviation (uncertain range) of the posterior samples. The benefit of the proposed approach is demonstrated by simulations on identifying the structural parameters with limit output data and noise polluted measurements.

1. Introduction

Recent years witness the increasing desire of Bayesian estimation for structural parametric system when quantifying the inevitable uncertainties, such as measurement error or structural model error and so forth, as is reviewed by Simoen et al. [1]. In particular, Beck and Au [2] used Laplace's method of asymptotic approximation to obtain a posterior PDF with a small-dimensional parameter space. To solve higher dimensional problems, Muto and Beck [3] developed an adaptive Markov chain Monte Carlo (MCMC) simulation for the Bayesian model updating. Gibbs sampling and transitional Markov chain Monte Carlo (TMCMC) were used by Ching and Chen [4] to obtain the posterior PDF of parameters. Cheung and Beck [5] used a hybrid Monte Carlo method, known as the Hamiltonian Markov chain, to solve higher dimensional model updating problems. Huhtala and Bossuyt [6] explored a Bayesian inference framework to solve the inverse problem of locating structural damage. An et al. [7] proposed a statistical model parameter estimation using Bayesian inference when parameters are correlated and observed that data have noise. Green [8] used a novel MCMC algorithm, data annealing, which is similar to simulated annealing, for the Bayesian identification of a nonlinear dynamic system.

The difficulty of Bayesian estimation lies in the efficiency in the convergence of posterior samples in the Markov chain to the acceptable model set. Moreover, because of the noise corrupted measurement, the surface of the prediction error lies in a hypersurface of a multidimensional parametric space. It will cause the surface of the probability density for the posterior sequences to have multiple regions of attraction and numerous local optima. It thus inevitably yields a biased estimator (no matter what is called maximum likelihood estimator, ML, or maximum a posteriori estimator, MAP). This problem is defined as the "equifinality" [9–11]. The surfaces of the prediction error, using formal likelihood measures, maximum log-likelihood (ML), are studied. From the surfaces of fitness measures, it can be concluded that the formal likelihood measure underestimates or overestimates the uncertain intervals of the posterior samples. The reason is that there are several possible models which can also give high values of likelihood around the neighborhood of the estimator.

In this paper, the bias between the ML/MAP estimator and the actual value is deduced by the Taylor expansion. It is found that the gradient and Hessian matrix of the likelihood measure can bridge the biased estimator and the actual value, which is thus proposed to improve the accuracy of the posterior samples. The parameter estimation

problem is proposed as a two-step strategy. In the first step, the MAP/ML estimator is obtained by the formal Bayesian likelihood measures using the differential evolution adaptive Metropolis-Hastings (DREAM) algorithm. In the second step, a new fitness measure is proposed, which can be seen as the informal likelihood measure under the framework of the generalized likelihood uncertainty estimation (GLUE) [12–14]. Numerical examples of a linear structural system are presented, with which the effectiveness and efficiency of the proposed method are investigated.

2. Problem Statement

2.1. Least Squares (LS) Estimator for the Inverse Problem. Let $y^m(t)$ denote the measured response at each time interval ($t = 1, \dots, N_t$) and $\hat{y}(\mathbf{x}, t)$ denotes the output of candidate models. The difference between the measured response and model outputs is defined as the residual error: $e(\mathbf{x}, t) = y^m(t) - \hat{y}(\mathbf{x}, t)$, where $j = 1, \dots, m$, and m is the number of outputs. The common measure for the inverse problem is to attempt to force the residual vector as close to zero as possible by tuning the model parameter vector, $\mathbf{x}$. Thus, the fitness measure can be defined as follows:

$$\text{LS}(\mathbf{x}) = -\frac{1}{m * N_t} \sum_{j=1}^m \sum_{t=1}^{N_t} e(\mathbf{x}, t)^2, \quad (1)$$

$$\mathbf{x}^* = \arg \max (\text{LS}(\mathbf{x})).$$

This is an n -dimensional optimization issue which maximizes the likelihood measure of SSR (equivalent to minimize the measure of LS formulation). But such measure can only provide an estimate of optimal value of $\mathbf{x}^*$. If we need to quantify the uncertainty of the estimator, it would be a desire to estimate the underlying posterior PDF of parameter, $P(\mathbf{x}(\theta) | y^m)$, which is under the framework of Bayesian probabilistic estimation.

2.2. Bayes Estimate Using Formal Log-Likelihood (LL) Measures. In the Bayesian estimation framework, the model set, M , is a class of probabilistic models, each of which predicts the response of the actual system. The identification problem is to infer the plausibility of each candidate model with a posterior density distribution conditioned by the measured data, $P(\mathbf{x}(\theta) | y^m)$; it is not a quest for the true structural parameters. $\mathbf{x}(\theta)$ is a stochastic parameter vector defining each possible model in the model set. The model set, M , is defined by random parameters, $\mathbf{x}(\theta) = (\hat{x}_1(\theta), \hat{x}_2(\theta), \dots, \hat{x}_n(\theta)) \in \mathbb{R}^n$ ($\theta_{N_s} \in \Omega$, θ is the random variables in probability space Ω), where n is the number of parameters for model $M_m \in M$ and N_s is the number of stochastic samples. The initial plausibility of each model parameterized by $\mathbf{x}(\theta_k)$ ($k = 1, 2, \dots, N_s$) is defined as a prior density function, $P(\mathbf{x}(\theta) | M)$. The updated plausibility of I/O model using Bayes' theorem is as follows:

$$P(\mathbf{x}(\theta) | y^m, M) = P(y^m | \mathbf{x}(\theta), M) \cdot \frac{P_0(\mathbf{x}(\theta) | M)}{P(y^m | M)}. \quad (2)$$

$P(y^m | \mathbf{x}(\theta), M)$ is the likelihood measure, $L(\mathbf{x}(\theta))$. If the measurement error is considered as obeying the Gaussian

distribution with a constant variance, σ_j^2 (j th available observed response), the posterior PDF in (2) is thus as follows:

$$P(\mathbf{x}(\theta) | y^m, M) = \frac{c^{-1} \cdot P_0}{\left(\prod_{j=1}^m \sqrt{2\pi}\sigma_j\right)^{N_t}} e^{[-\sum_{j=1}^m (1/2\sigma_j^2) \sum_{t=1}^{N_t} (y_j^m(t) - \hat{y}_j(\mathbf{x}(\theta), t))^2]}, \quad (3)$$

where c is the evidence of model class ($c = P(y^m | M) = \int P(y^m | \mathbf{x}(\theta), M) \cdot P(\mathbf{x}(\theta) | M) d\theta$), which is a high-dimensional integral. The difficulty in estimating the posterior PDF is none other than approximating the model evidence. Strives to overcome this challenge are the purpose of the Metropolis-Hastings algorithm. For simplicity, (2) is rewritten as $p(\theta) = L(\theta) \cdot p_0(\theta)/c = f(\theta)/c$. The MH algorithm generates the posterior PDF with four steps. (1) Start with initial samples, θ_0 , and compute the likelihood measures $f(\theta_0)$. (2) The updated samples, θ^* , are produced by the jumping distribution, $q(\theta_1, \theta_2)$, which is the probability of returning a value of θ_2 given a previous value of θ_1 . The restriction on the jumping is that the transition is probability symmetric, $q(\theta_1, \theta_2) = q(\theta_2, \theta_1)$. (3) The acceptance ratio at the updated candidate (θ^*) and the source posterior samples (θ_{t-1}) is $\alpha(\theta) = p(\theta^*)/p(\theta_{t-1}) = f(\theta^*)/f(\theta_{t-1}) = L(\theta^*)/L(\theta_{t-1})$. (4) If the acceptance ratio, $\alpha > 1$, accept the candidate sample, θ^* ; if the jump decreases the density, $\alpha < 1$, then it rejects the updating and keeps the current samples, θ_{t-1} . The acceptance ratio is as follows:

$$\alpha(\theta) = \min \left[\frac{L(\theta^*)}{L(\theta_{t-1})}, 1 \right]. \quad (4)$$

It is clear that the advantage of MH algorithm lies in the fact that when computing acceptance ratio there is no need to obtain the model evidence since the constant c cancels out. The transition of samples generates a Markov chain ($\theta_0, \theta_1, \dots, \theta_k, \dots$). Following a burn-in period, the Markov chain approaches its stationary distribution, and the samples after the burn-in period converge into the posterior PDF, $p(\theta)$, as that in (2). From (4), it can be found that the Bayesian estimate relies on the likelihood measure, $P(y^m | \mathbf{x}(\theta), M)$. It is more convenient to use the logarithm of the likelihood measures (LL($\mathbf{x}(\theta)$)) rather than the likelihood function itself:

$$\text{LL}(\mathbf{x}(\theta)) = -\frac{N_t}{2} \ln(2\pi) - \frac{N_t}{2} \sum_{j=1}^m \ln(\sigma_j^2) - \frac{1}{2} \sum_{j=1}^m \sigma_j^{-2} \cdot \sum_{t=1}^{N_t} (y_j^m(t) - \hat{y}_j(\mathbf{x}(\theta), t))^2. \quad (5)$$

Either the log-likelihood measure, as is in (5), or the least square measure, as is in (1), obeys the rule of "goodness-of-fit." This is because only the model with high probabilistic value of likelihood in the MH method will be accepted.

2.3. The Surface of the Likelihood Measures. To illustrate the problem of "equifinality," the surfaces of the common-used likelihood measure, the LL($\mathbf{x}(\theta)$) as is in (5), are simulated in identifying the stiffness parameters of a 2-DOF linear

dynamic system. The state space of the system is written as follows:

$$\begin{bmatrix} \dot{\mathbf{v}}_1(t) \\ \dot{\mathbf{v}}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ -\mathbf{I} \end{bmatrix} \Gamma^T \mathbf{u}(t), \quad (6)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the mass, damping, and the stiffness matrices, $\mathbf{I}$ is $n \times n$ identity matrix, and $\Gamma = [1, 1, \dots, 1]^T$ is $1 \times n$ position vector. $\mathbf{v}_1(t)$ and $\mathbf{v}_2(t)$ are state space vectors, respectively, representing the displacement and velocity, and $\mathbf{u}(t)$ is the input of the system. The system output is an acceleration which is assumed to be contaminated by Gaussian white noise $w_j(t) \sim N(0, \sigma_j(\theta))$, ($j = 1, \dots, m$). The measured output vector is thus

$$\mathbf{y}(t) = [-\mathbf{M}^{-1}\mathbf{K} \quad -\mathbf{M}^{-1}\mathbf{C}] \begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} - \Gamma^T \mathbf{u}(t) + \mathbf{w}(\sigma(\theta), t). \quad (7)$$

The mass and stiffness of each DOF are defined as 100 kg and 1000 N/m. Equation (7) includes a Rayleigh damping C Mita [15], where the first two-modal damping ratio (ζ_r) is set as 5%:

$$\begin{aligned} \mathbf{C} &= \alpha \mathbf{M} + \beta \mathbf{K}, \\ \zeta_r &= \frac{\alpha}{2\omega_r} + \frac{\beta\omega_r}{2}, \quad (r = 1, 2). \end{aligned} \quad (8)$$

The parametric domain is meshed by 5% deviation of the true value. The output acceleration (acc.) with different noise levels (7) was used in the simulation, in which the noise level (nl.) was defined as $\sigma_i = \text{nl.} \times \sigma_{\text{acc.}}$. The contour plots of the likelihood measure, as is in (5), in the scenarios of noise-free and different noise level scenarios are exhibited in Figure 1.

From Figure 1, we can find that only when the measurement error is ignored, the center of the posterior samplings model set (the ML/MAP estimator) will be unbiased; however, taking the noise into account, the optimal solution with maximum PDF is biased. Moreover, around the search neighborhoods of the optimal solution, there are many local optimums. It can conclude that when considering the measurement error the common likelihood measures as in (1) and in (5) are weak to solve the problem of equifinality. The bias of the estimator will increase with adding the number of the parametric dimensions and the noise level. It is thus necessary to improve the identified ML/MAP estimator.

3. The Proposed Accuracy-Improving Method

3.1. The First Two Deviations of the Likelihood Measure. With Taylor's expansion, the likelihood measure can be deduced as

$$\begin{aligned} \mathbf{L}(\hat{\mathbf{x}}(\theta)) &= \mathbf{L}(\hat{\mathbf{x}}(\theta_o)) + \mathbf{L}'(\hat{\mathbf{x}}(\theta_o)) \Delta \hat{\mathbf{x}} \\ &+ \frac{1}{2} \Delta \hat{\mathbf{x}}^T \mathbf{L}''(\hat{\mathbf{x}}(\theta_o)) \Delta \hat{\mathbf{x}} + \mathbf{o}(\Delta \hat{\mathbf{x}}^2), \end{aligned} \quad (9)$$

where $\Delta \hat{\mathbf{x}}$ denotes $\hat{\mathbf{x}}(\theta) - \hat{\mathbf{x}}(\theta_o)$. With the derivative of (9) to $\hat{\mathbf{x}}(\theta)$ and ignoring the high order derivative series, one obtains:

$$\mathbf{G}(\hat{\mathbf{x}}(\theta)) = \mathbf{G}(\hat{\mathbf{x}}(\theta_o)) + \Delta \hat{\mathbf{x}} \mathbf{H}(\hat{\mathbf{x}}(\theta_o)), \quad (10)$$

where $\mathbf{G}(\hat{\mathbf{x}}(\theta))$ is the first order derivative of likelihood measure, which is the gradient matrix of $\mathbf{L}(\hat{\mathbf{x}}(\theta))$; $\mathbf{H}(\hat{\mathbf{x}}(\theta))$ is the second order derivative of likelihood measure, which is the Hessian matrix of $\mathbf{L}(\hat{\mathbf{x}}(\theta))$. $\mathbf{G}(\hat{\mathbf{x}}(\theta_o))$ and $\mathbf{H}(\hat{\mathbf{x}}(\theta_o))$ are the gradient and Hessian matrix at the point of true value, $\hat{\mathbf{x}}(\theta_o)$. Since $\mathbf{G}(\hat{\mathbf{x}}(\theta_{\text{ML}})) = \mathbf{0}$, then it yields that

$$\hat{\mathbf{x}}(\theta_{\text{ML}}) = \hat{\mathbf{x}}(\theta_o) - \mathbf{G}(\hat{\mathbf{x}}(\theta_o)) [\mathbf{H}(\hat{\mathbf{x}}(\theta_o))]^{-1}. \quad (11)$$

As seen in (11), the bias, $\hat{\mathbf{x}}(\theta_{\text{ML}}) - \hat{\mathbf{x}}(\theta_o)$, is equal to the negative product of the gradient and the inverse Hessian matrix at the actual value. Equation (11) is thus proposed in this study to formulate the informal likelihood measure. Let $d(\hat{\mathbf{x}}(\theta)) = \hat{\mathbf{x}}(\theta) - \hat{\mathbf{x}}(\theta_{\text{ML}})$ denote the bias between the MAP/ML estimator and each of the posterior samples. The proposed Bayesian updating of the posterior samples using DREAM algorithm can be divided into two steps. The first step is to find the ML/MAP estimator, $\hat{\mathbf{x}}(\theta_{\text{ML}})$. The second step is to search the optimal point, $\hat{\mathbf{x}}(\theta^*)$, of the proposed criteria around the neighborhood of the ML/MAP estimator. The informal likelihood measure can be written as follows:

$$\begin{aligned} L^*(\hat{\mathbf{x}}(\theta)) &= \frac{1}{(\|d(\hat{\mathbf{x}}(\theta)) - \mathbf{G}(\hat{\mathbf{x}}(\theta)) [\mathbf{H}(\hat{\mathbf{x}}(\theta))]^{-1}\| + \varepsilon)}, \\ \hat{\mathbf{x}}(\theta^*) &= \arg \max(L^*), \end{aligned} \quad (12)$$

where ε denote the credible range, which can be decided by the user as the stop criteria.

3.2. Illustration of the Proposed Likelihood Measure. From (12), it can be found that the extreme points of the likelihood measure, $L^*(\hat{\mathbf{x}}(\theta))$, are the actual value, $\hat{\mathbf{x}}(\theta_o)$, and the MAP/ML estimator, $\hat{\mathbf{x}}(\theta_{\text{ML}})$, because the former extreme value is the null point of (11) and the latter item meets the condition of $\mathbf{G}(\hat{\mathbf{x}}(\theta_{\text{ML}})) = \mathbf{0}$. Since the MAP/ML estimator, $\hat{\mathbf{x}}(\theta_{\text{ML}})$, has already been obtained in the first step after the distribution of the posterior samples has become stationary, the posterior samples on the Markov chains in the second step will converge into the neighborhood around the actual value due to the proposed likelihood measure. The accuracy of the estimator is thus improved and the uncertain range of the posterior samples will be narrowed around the actual value. To verify the proposed likelihood measure, as is in (12), the surface and contour for the 2-DOF linear system, of which the simulation in 100% noise scenario is the same as that in Section 2.3, are shown in Figure 2. The accuracy of the posterior samples, seen in Figure 2, is improved comparing with the MAP estimator, comparing with Figure 1(b).

3.3. Two Steps of the DREAM Based Bayesian Estimation

3.3.1. Step 1: MAP Estimator Using DREAM Algorithm. The DREAM algorithm combines the DE mutation strategy into

FIGURE 1: Contour plot of the likelihood measures. (“ $\circ$ ” denotes the actual value; “ $*$ ” denotes the MAP/ML estimator.)

FIGURE 2: The surface and contour plot of the proposed likelihood measures.

the updating of Markov chains. In the transition of Markov chains, the posterior samples are updated by the revised DE mutation strategy. Let $\Delta_{\hat{x}}^i(\theta) = \hat{x}_{k+1}^i(\theta) - \hat{x}_k^i(\theta)$ denote the jump scale between the updating state ($k+1$) and current state (k) of the i th of the Markov chain. The samples are updated as follows:

$$\Delta_{\hat{x}}^i = (I_d + e_d) \gamma(\delta, d) \left[\sum_{j=1}^{\delta} \hat{x}_k^{r_1(j)} - \sum_{n=1}^{\delta} \hat{x}_k^{r_2(n)} \right] + \varepsilon_r, \quad (13)$$

where ε_r is a small random vector that is drawn from $N_d(0, \hat{\Sigma})$; δ is the number of chosen pairs; and $r_1(j)$ and $r_2(n)$ are, respectively, different and random integers that are chosen from the integer set $\{1, 2, \dots, i-1, i+1, \dots, N_s\}$. The term I_d signifies the d -dimensional identity matrix, and e_d signifies a small random vector drawn from a uniform distribution to assure the ergodicity of the samples on the Markov chain. The scaling factor γ is decided by the values of δ and d , where d is the parametric dimension. The DREAM algorithm also explores the DE crossover strategy in d dimensions of the current samples, $\hat{x}_k^i(\theta)$, the updated posterior samples, $\hat{x}_{k+1}^i(\theta)$, and a trial sample $\hat{x}_{k+1}^{ij}(\theta)$ is generated with

$$\hat{x}_{k+1}^{ij}(\theta) = \begin{cases} \hat{x}_k^{ij}(\theta), & \text{if } \text{und}(j) \leq 1 - \text{CR}, \\ \hat{x}_{k+1}^i(\theta), & \text{otherwise,} \end{cases} \quad (14)$$

where $j = 1, 2, \dots, d$; $\text{und}(j)$ is the j th independent random number uniformly distributed in the range of $[0, 1]$. CR is a crossover probability defined by the user. The DREAM accepts the candidate state $\hat{x}_{k+1}(\theta)$ with probability $\min(1, \alpha(\hat{x}_{k+1}(\theta), \hat{x}_k(\theta)))$ and keeps the current state $\hat{x}_k(\theta)$ with probability $1 - \min(1, \alpha(\hat{x}_{k+1}(\theta), \hat{x}_k(\theta)))$. The MH acceptance probability, $\alpha(\hat{x}_{k+1}(\theta), \hat{x}_k(\theta))$, is computed by (4). The posterior samples are updated by (13) and (14) and approaches to a stationary after a burn-in period. The convergence of posterior samples is checked by the Gelman-Rubin criteria [16]. The $\hat{R}_j$ -statistic, $\hat{R}$, is computed by using the last 50% of the samples in each chain. The convergence criteria, $\hat{R}$, are

$$\hat{R} = \sqrt{\frac{k+1}{k} \times \frac{\hat{\sigma}^2}{W} - \frac{N_s - 1}{k \times N_s}}, \quad (15)$$

where k is the number of the used posterior samples; W equals $\sum_{i=1}^k s_i^2/k$, where s_i^2 is the variance of the sequence; the posterior variance is estimated as $\hat{\sigma}^2 = ((n-1)/n)W + (1/n)B$, where B is the variance between the sequences; B can be computed as $B = N_s \times \sum_{i=1}^k (\hat{x}_i - \bar{\hat{x}})^2 / (k-1)$.

3.3.2. Step 2: Density Updating of the Samples That Satisfy the Proposed Criteria. When $\hat{R}_j$ in (14) for each dimension (j) is less than 1.2, the posterior samples on the MC chain converged into a stationary distribution. The MAP estimator, $\mathbf{x}_{\text{MAP}}$, is obtained with the maximum posterior PDF. Also, the standard deviation of the posterior samples, $\sigma_{\mathbf{x}(\theta)}$, can be obtained. Then, the algorithm turns into the second step. The gradient and Hessian matrix of the posterior samples within

the boundary of $\mathbf{x}_{\text{MAP}} \mp 3\sigma_{\mathbf{x}(\theta)}$ are calculated at each iteration. The proposed informal likelihood measure as in (11) will be used for the updating of the posterior samples. The estimation procedure will stop till the prescribed precision, ε , is satisfied. The gradient matrix and the Hessian matrix at each sample can be written as

$$\mathbf{G}(\hat{\mathbf{x}}(\theta_j)) = \left[\frac{\partial L}{\partial \hat{x}_1}, \frac{\partial L}{\partial \hat{x}_2}, \dots, \frac{\partial L}{\partial \hat{x}_n} \right],$$

$$\mathbf{H}(\hat{\mathbf{x}}(\theta_j)) = \begin{bmatrix} \frac{\partial^2 L}{\partial \hat{x}_1^2} & \frac{\partial^2 L}{\partial \hat{x}_1 \partial \hat{x}_2} & \dots & \frac{\partial^2 L}{\partial \hat{x}_1 \partial \hat{x}_n} \\ \frac{\partial^2 L}{\partial \hat{x}_2 \partial \hat{x}_1} & \frac{\partial^2 L}{\partial \hat{x}_2^2} & \dots & \frac{\partial^2 L}{\partial \hat{x}_2 \partial \hat{x}_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial \hat{x}_n \partial \hat{x}_1} & \frac{\partial^2 L}{\partial \hat{x}_n \partial \hat{x}_2} & \dots & \frac{\partial^2 L}{\partial \hat{x}_n^2} \end{bmatrix}, \quad (16)$$

where j denotes the j th posterior sample and n is the parametric dimension. The diagonal and off-diagonal elements of the $\mathbf{H}(\hat{\mathbf{x}}(\theta_j))$ can be given by (16) and the following equation:

$$H^{ll}(\hat{\mathbf{x}}(\theta_j)) = \frac{L(\hat{\mathbf{x}}(\theta_j) + \Delta \hat{\mathbf{x}}_l(\theta_j)) - 2L(\hat{\mathbf{x}}(\theta_j)) + L(\hat{\mathbf{x}}(\theta_j) - \Delta \hat{\mathbf{x}}_l(\theta_j))}{(\Delta \hat{x}_l(\theta_j))^2},$$

$$H^{ll'}(\hat{\mathbf{x}}(\theta_j)) = \frac{1}{4\Delta \hat{x}_l(\theta_j) \Delta \hat{x}_{l'}(\theta_j)} \cdot [L(\hat{\mathbf{x}}(\theta_j) + \Delta \hat{\mathbf{x}}_l(\theta_j) + \Delta \hat{\mathbf{x}}_{l'}(\theta_j)) - L(\hat{\mathbf{x}}(\theta_j) + \Delta \hat{\mathbf{x}}_l(\theta_j) - \Delta \hat{\mathbf{x}}_{l'}(\theta_j)) - L(\hat{\mathbf{x}}(\theta_j) - \Delta \hat{\mathbf{x}}_l(\theta_j) + \Delta \hat{\mathbf{x}}_{l'}(\theta_j)) + L(\hat{\mathbf{x}}(\theta_j) - \Delta \hat{\mathbf{x}}_l(\theta_j) - \Delta \hat{\mathbf{x}}_{l'}(\theta_j))]. \quad (17)$$

3.4. Identification Procedures. The procedure of the proposed posterior density estimation is as follows.

Step 1. Use the LHS method to generate N_s sequences for the initial state of MC chains, respecting the prescribed limits of the search space. The likelihood measure of each sample is obtained by (5).

Step 2. Update the posterior sample of the Markov chain by mutation strategy using (13) and by the crossover probability using (14). Calculate the density for the updated samples.

Step 3. The Metropolis acceptance (4) is used for choosing the accepted posterior samples.

Step 4. Return to Step 2 and to Step 3, after the burn-in period and calculating the convergence criteria using (15) for each

TABLE 1: Structural properties.

Stiffness	(N/m)
Levels 1-5	5.0×10^3
Mass	(kg)
Levels 1-4	50
Levels 5	45
Damping ratio	
$\zeta_{1,2}$	0.05

dimension of the structural parameter. If the convergence criteria of the MC chain are met ($\hat{R}_j < 1.2$), the MAP estimator and the standard deviation of the samples are obtained; otherwise, return to Step 2.

Step 5. Calculate the gradient and Hessian matrix at the point of each sample within the interval of $[\mathbf{x}_{\text{MAP}} \mp 3\sigma_{\mathbf{x}(\theta)}]$ by (16) and (17).

Step 6. The informal likelihood measure as in (12) is used for the transition of the posterior samples till the predefined stop condition is satisfied.

4. Numerical Study

4.1. Identification of a 5-DOF LTI System. Numerical simulation of a 5-DOF LTI system was carried out to verify the proposed method. The structural system is simulated as (6) and the measured signal is as (7). The input was an El-Centro wave lasting 40 s and the sampling frequency was 100 Hz (Figure 3). The measurement noise of the 5th DOF in the 100% noise level scenarios is shown in Figure 3. Table 1 shows the structural properties of the dynamic system.

The influence of the limited availability of measurements on the proposed method is also assessed in this study. In the “full output” scenario, measurements of all DOFs are available, whereas, in the “partial output” case, only data from DOFs 3 and 5 are available. The mass is assumed to be known; hence, an n -DOF system with m -available measurements is described by a model set, of which the interested parameterized vector is as follows:

$$\mathbf{x}(\theta) = \{k_1(\theta), \dots, k_n(\theta), \zeta_1(\theta), \zeta_2(\theta)\}. \quad (18)$$

The parameters of the DREAM algorithm were set as follows: the number of Markov chain samples (N_s) was 20, the crossover probability (CR) was 0.85, and the number of sample pairs (δ) was 5. The search domain was taken to be 0.5–2.0 times the true value. The identified results are exhibited in Table 2.

From Table 2, the conclusion that the proposed method can find the optimal value, $\hat{\mathbf{x}}(\theta^*)$, in the scenarios of both “full outputs” and “partial outputs,” even when considering large noise level, can be drawn. Ignoring the measured noise error, the proposed method can identify the structural parameters with no bias. The coefficient variance (cov.) of posterior samples is identified as zero, which demonstrates that the inverse problem under the noise-free scenario is deterministic. When

FIGURE 3: Input and simulated measured error.

the acceleration of each DOF is available for measure, the maximum relative error ranged from 0.494% in the 30% noise case to 0.726% in the 100% noise case. While considering that the measurements of the 3rd and the 5th DOFs are available, the maximum relative error ranged from 0.701% in the 30% noise case to 1.044% in the 100% noise case. The credible range of the posterior samples is depicted by the coefficient variance of the MC sequences. It is clear that the parametric uncertainty was additive with the increasing of measurement error. In the “full outputs” scenario, the maximum coefficient variance (cov.) of the posterior samples ranged from 0.897% in the 30% noise level to 2.397% in the 100% noise case. Correspondingly increasing with noise level in the partial outputs scenario, the maximum cov. rises from 2.423% to 4.197%.

The convergence progress of the MC samples for each dimensional structural parameter in the scenario of partial output and the consideration 100% noise level are shown in Figure 4. From Figure 4, it is found that the probability density of the posterior samples in Markov chain approached to be stationary after 1500 iterations, which is clearly shown in Figure 4(b). The proposed likelihood measure used for the transition of posterior samples to find the optimal $\hat{\mathbf{x}}(\theta^*)$ works after 2500 iterations. The posterior uncertain range that assures a reliability of 95% can be obtained from the posterior samples of the model class which denotes the plausibility of each I/O system. Figure 5 shows the uncertain range of the response with 95% assurance and with measurement errors at each time interval by incorporating the identified standard deviation of the prediction error (time history of initial 10 s is shown). The percentage of the response considering a 100% measurement error within the uncertain range that considers prediction error is 94.76%.

TABLE 2: Identified results of structural parameters.

		Full outputs			Partial outputs		
		No noise	30% noise	100% noise	No noise	30% noise	100% noise
k_1	Error	0.000	0.224	0.561	0.000	0.241	0.932
	Cov.	0.000	0.484	1.657	0.000	1.244	3.339
k_2	Error	0.000	0.239	0.487	0.000	0.395	0.498
	Cov.	0.000	0.358	1.137	0.000	1.852	3.039
k_3	Error	0.000	0.383	0.603	0.000	0.516	0.705
	Cov.	0.000	0.628	2.234	0.000	1.402	3.394
k_4	Error	0.000	0.353	0.472	0.000	0.447	0.819
	Cov.	0.000	0.417	1.322	0.000	0.921	2.898
k_5	Error	0.000	0.473	0.681	0.000	0.486	0.691
	Cov.	0.000	0.443	1.475	0.000	1.893	2.323
ζ_1	Error	0.000	0.398	0.581	0.000	0.695	1.044
	Cov.	0.000	0.319	2.322	0.000	2.145	3.549
ζ_2	Error	0.000	0.494	0.726	0.000	0.701	0.993
	Cov.	0.000	0.897	2.397	0.000	2.423	4.197

The “Error” is in %; and the “Cov.” (the ratio of standard deviation to the mean) is in %.

FIGURE 4: Identification progress for stiffness of the 5th floor (partial output: 100% noise).

FIGURE 5: 95% uncertainty ranges for acceleration of the 5th DOF (100% noise: partial output).

4.2. The Identification of a 10-DOF LTI System. To show the advantage of the proposal, the identification results of a 10-DOF LTI system using the improved method, as is proposed in (12), are compared with the estimation solutions using the formal log-likelihood measure, as is in (5). For simplicity, the mass of each floor is assumed to be 50 kg, the stiffness of each DOF is equal to 5000 N/m, and the first two-mode shape damping ratio is assumed to be 0.05. The input excitation is the same as that used in the simulation of the 5-DOF LTI system. The white noise is added to the measured response, simulated by (11), where the noise level is assumed to be 10%, 30%, and 100%. When considering the scenario of partial measurements, only the even DOFs (2nd, 4th, 6th, 8th, and 10th) are assumed to be available for Bayesian updating. The results are shown in Figure 6.

In the case of “full outputs,” the maximum relative error of $\hat{x}_{\text{MAP}}$ ranges from zero to 0.617% in 10% noise level case,

FIGURE 6: Identified results using the proposed and traditional likelihood measures (gray bars: traditional likelihood measure; blue bars: proposed likelihood measure).

1.175% in 30% noise level, and 1.766% in 100% noise level. Correspondingly, the maximum relative error of $\hat{x}_{MAP}$ using the traditional likelihood measure increases from zero in the case of ignoring measurement noise to 0.754% in 10% noise level, 1.999% in 30% noise level, and 6.491% in 100% noise level. The improvement is also clear in the scenario of partial outputs. The maximum relative error of $\hat{x}_{MAP}$ using the proposed method increases from zero in noise-free case to 1.479% in 10% noise level, 2.175% in 30% noise level, and 3.543% in 100% noise level. While the maximum relative error of $\hat{x}_{MAP}$ using the traditional method rises from zero in no-noise case to 2.621% in 10% noise level, 5.974% in 30% noise level, and 8.451% in 100% noise level, it can be found that the minimum and maximum relative errors of mean posterior samples in model set are all reduced using the proposed method. Moreover, it can be found from Figure 6 that using (12) for the parametric uncertainty (the cov.) becomes smaller than that obtained by the formal log-likelihood measure as in (5). For instance, when considering the case of 100% noise level and partial outputs are available, the maximum coefficient variance of the MC samples using the proposed method is 6.198% compared with that obtained by traditional

method which is 11.09%. Comparing the identification results shown in Figure 6, it is clear that the accuracy of the estimator using the proposed likelihood measure is improved.

5. Conclusions

To improve the accuracy of the MAP estimator obtained by the traditional likelihood measure using the DREAM based structural identification, the gradient and Hessian of the log-likelihood measure are proposed to formulate the generalized likelihood measure for the density transition of Markov chains. Compared with the formal likelihood function, the relative error of the MAP estimator and the uncertain range of the posterior samples using the proposed method becomes smaller. Numerical simulations of a 5-DOF and 10-DOF LTI system demonstrated their effectiveness in solving identification problems with a high noise level and loss of measurement data. In conclusion, DREAM based Bayesian estimation using the proposed improvement has the potential to solve the problem of “equifinality” especially when considering large level of measurement error.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported in part by a Grant for the project “Implementation of Shelter Guidance System for Commuters Who Are Unable to Return Home Based on Structural Health Monitoring of Tall Buildings after Large-Scale Earthquake” (FY2013–2016, PI: A. Mita) from the Japan Science and Technology Agency.

References

- [1] E. Simoen, G. De Roeck, and G. Lombaert, “Dealing with uncertainty in model updating for damage assessment: a review,” *Journal of Mechanical Systems and Signal Processing*, vol. 56–57, pp. 123–149, 2015.
- [2] J. L. Beck and S.-K. Au, “Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation,” *Journal of Engineering Mechanics*, vol. 128, no. 4, pp. 380–391, 2002.
- [3] M. Muto and J. L. Beck, “Bayesian updating and model class selection for hysteretic structural models using stochastic simulation,” *Journal of Vibration and Control*, vol. 14, no. 1–2, pp. 7–34, 2008.
- [4] J. Ching and Y.-C. Chen, “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, 2007.
- [5] S. H. Cheung and J. L. Beck, “Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters,” *Journal of Engineering Mechanics*, vol. 135, no. 4, pp. 243–255, 2009.

- [6] A. Huhtala and S. Bossuyt, "A Bayesian approach to vibration based structural health monitoring with experimental verification," *Journal of Structural Mechanics*, vol. 44, no. 4, pp. 330–344, 2011.
- [7] D. An, J.-H. Choi, and N. H. Kim, "Identification of correlated damage parameters under noise and bias using Bayesian inference," *Structural Health Monitoring*, vol. 11, no. 3, pp. 293–303, 2012.
- [8] P. L. Green, "Bayesian system identification of a nonlinear dynamical system using a novel variant of Simulated Annealing," *Mechanical Systems and Signal Processing*, vol. 52-53, pp. 133–146, 2015.
- [9] K. J. Beven, "A manifesto for the equifinality thesis," *Journal of Hydrology*, vol. 320, no. 1-2, pp. 18–36, 2006.
- [10] J. R. Stedinger, R. M. Vogel, S. U. Lee et al., "Appraisal of the generalized likelihood uncertainty estimation (GLUE) method," *Water Resources Research*, vol. 44, no. 12, pp. 1–17, 2008.
- [11] M. Sadegh and J. A. Vrugt, "Approximate Bayesian computation in hydrologic modeling: equifinality of formal and informal approaches," *Hydrology and Earth System Sciences Discussions*, vol. 10, pp. 4739–4797, 2013.
- [12] K. Beven and J. Freer, "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology," *Journal of Hydrology*, vol. 249, no. 1–4, pp. 11–29, 2001.
- [13] Y. Zhang, H. H. Liu, and J. Houseworth, "Modified generalized likelihood uncertainty estimation (GLUE) methodology for considering the subjectivity of likelihood measure selection," *Journal of Hydrologic Engineering*, vol. 16, no. 6, pp. 558–561, 2011.
- [14] Z. Li, Q. Chen, Q. Xu, and K. Blanckaert, "Generalized likelihood uncertainty estimation method in uncertainty analysis of numerical Eutrophication models: take bloom as an example," *Mathematical Problems in Engineering*, vol. 2013, Article ID 701923, 9 pages, 2013.
- [15] A. Mita, *Structural Dynamics for Health Monitoring*, Sankeisha, Nagoya, Japan, 2003.
- [16] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequence," *Statistical Science*, vol. 7, no. 4, pp. 457–472, 1992.