literature [5]. The HMC algorithm simulates a Markov chain in which each iteration involves a Metropolis update with a resampled stochastic proposal and a deterministic proposal found using the solution to a discretized Hamiltonian dynamics problem. Cheung and Beck [2] adopted an HMC algorithm to solve a high-dimensional model updating problem in structural dynamics using real time data.

While MALA and HMC potentially improve the mixing time and optimal acceptance rate of the MCMC sampling algorithms, their tuning remains a challenge for highly correlated parameter spaces involved in FE model updating problems. Another improvement to MCMC algorithms is called mMALA [12], a method in which the proposal process is derived using the discretized Itô stochastic differential equation of Langevin diffusion on a Riemann manifold where the metric tensor is the Hessian of the negative logarithm of the target distribution. If the local curvature of the Riemann manifold is assumed to be constant, the algorithm collapses to the simplified mMALA (smMALA) algorithm. The drift term in the proposal process of the smMALA method can be interpreted as a scaled Newton step from an optimization perspective [6] and is used to drive the Markov chain towards the higher probability region. Using the information of the manifold structure of the parameter space provides a systematic manner to design a Markov chain which explores the posterior distribution more efficiently in highly correlated target distributions, provided that the scale factor is small enough.

This study primarily shows how the smMALA algorithm can be adopted to solve challenging model updating problems in structural dynamics. Practical issues are addressed to improve the effectiveness of the method when dealing with high dimensional posterior distributions. The potential applicability of the proposed adaptive algorithm to estimate the posterior distribution of the parameters in a Bayesian FE model updating problem is illustrated using simulated data from a 6DOF mass-spring-damper system with 16 parameters to be calibrated.

2 FRF-BASED BAYESIAN MODEL UPDATING

2.1 Forward problem

Using the FE method, the governing partial differential equations of motion of a linear structure can be transformed into a set of ordinary differential equations as

$$\boldsymbol{M}(\boldsymbol{\theta})\ddot{\boldsymbol{x}}(t) + \boldsymbol{V}(\boldsymbol{\theta})\dot{\boldsymbol{x}}(t) + \boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{x}(t) = \boldsymbol{f}(t) \tag{2}$$

Here $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ is the nodal displacement vector, $\boldsymbol{f}(t)$ is the external load vector which is governed by a Boolean transformation of the input vector, $\boldsymbol{f}(t) = \boldsymbol{P}_{\mathbf{u}}\boldsymbol{u}(t)$. The real, positive-definite, symmetric matrices $\boldsymbol{M}(\boldsymbol{\theta})$, $\boldsymbol{V}(\boldsymbol{\theta})$ and $\boldsymbol{K}(\boldsymbol{\theta}) \in \mathbb{R}^{n_x \times n_x}$ are the parameterized mass, damping and stiffness matrices, respectively, and $\boldsymbol{\theta} \in \mathbb{R}^{n_\mathrm{p}}$ is the normalized vector of model parameters. The Fourier transform of Eq. (2) (assuming a stationary state) gives $\boldsymbol{Y}(\omega;\boldsymbol{\theta}) = \boldsymbol{H}(\omega;\boldsymbol{\theta})\boldsymbol{F}(\omega)$ where $\boldsymbol{Y}(\omega;\boldsymbol{\theta}) = -\omega^2\boldsymbol{X}(\omega;\boldsymbol{\theta})$ denotes the spectrum of nodal accelerations and

$$H^{mn}(\omega;\boldsymbol{\theta}) = -\omega^2[-\omega^2\boldsymbol{M}(\boldsymbol{\theta}) + i\omega\boldsymbol{V}(\boldsymbol{\theta}) + \boldsymbol{K}(\boldsymbol{\theta})]^{-1} \tag{3}$$

denotes the predicted acceleration Frequency Response Function (FRF) of the mth degree of freedom due to an external load at the nth degree of freedom.

2.2 Uncertainty in FRF measurements

Let represent the discrete time history of a structure's response, in terms of acceleration, by $\{\hat{\boldsymbol{y}}_j \in \mathbb{R}^{N_o}, j = 1, \dots, N\}$ where N_o is the number of measurement channels and N is the number of time samples. In the context of Bayesian model updating, the observed acceleration data can be modeled as follows

$$\hat{\boldsymbol{y}}_j = \boldsymbol{y}_j(\boldsymbol{\theta}) + \boldsymbol{\epsilon}_j^y, j = 1, \dots, N \tag{4}$$

where $\boldsymbol{y}_j(\boldsymbol{\theta})$ denotes the accelerations predicted by the structural model and $\boldsymbol{\epsilon}_j^y$ is the prediction error. This formulation thus states that there is a difference between measured response and the model response due to measurement noise. The prediction error is modeled as zero-mean, independent and identically distributed (*i.i.d*) Gaussian noise with variance σ^2. It is also assumed that the prediction error in each measurement channel is independent of other channels. For the sake of simplicity, we furthermore assume that the structure is excited through a single, noise free, input channel.

Using the Fast Fourier Transform (FFT), the measured spectra can be defined as

$$\hat{\boldsymbol{Y}}_l = \boldsymbol{Y}_l(\boldsymbol{\theta}) + \boldsymbol{\epsilon}_l^Y, l = 1, \dots, N \tag{5}$$

where the FFT of $\hat{\boldsymbol{y}}$ is defined as follows

$$\begin{aligned}\hat{\boldsymbol{Y}}_l &= Re(\hat{\boldsymbol{Y}}_l) + i\, Im(\hat{\boldsymbol{Y}}_l) \\ &= \frac{1}{\sqrt{N}}\sum_{m=1}^{N} \hat{\boldsymbol{y}}_m \exp\left(-2\pi i \frac{ml}{N}\right), l = 1, \dots, N\end{aligned} \tag{6}$$

Here, $l = 1, \dots, N$ is corresponding to the frequencies $f_l = (l-1)/(N\Delta t)$, where Δt is the sampling time interval. A similar formulation is used to define the FFT of the model response and the prediction error. The measurement noise $\boldsymbol{\epsilon}_l^Y$ is an asymptotically circular normal variable and independent over the frequency lines f_l [14]. This means that the real and imaginary parts of $\boldsymbol{\epsilon}_l^Y$ are independent and identically distributed as $N(\mathbf{0}, \frac{\sigma^2}{2}\boldsymbol{I}_{N_o})$, where $\boldsymbol{I}_{N_o}$ is the $N_o \times N_o$ identity matrix. Let the measured spectrum of the input which corresponds to the frequency f_l be denoted by $\boldsymbol{U}_l$. Then, the measured FRFs, $\hat{\boldsymbol{H}}_l \in \mathbb{C}^{N_o \times 1}$, can be written as

$$\hat{\boldsymbol{H}}_l = \frac{\hat{\boldsymbol{Y}}_l}{\boldsymbol{U}_l} = \frac{\boldsymbol{Y}_l(\boldsymbol{\theta})}{\boldsymbol{U}_l} + \frac{\boldsymbol{\epsilon}_l^Y}{\boldsymbol{U}_l} = \boldsymbol{H}_l(\boldsymbol{\theta}) + \boldsymbol{\epsilon}_l^H \tag{7}$$

For Bayesian model updating, the FRF values corresponding to frequencies $\{f_l,\ l = 2, \dots, N_{nyq}\}$ contain the required information, where $N_{nyq} = \mathrm{INT}(N/2)$ is the Nyquist frequency. The data at $l = 1$ and $l = N_{nyq} + 1, \dots, N$ is omitted since the value for $l = 1$ is the scaled average value of the signal and is often contaminated by the offset error of the measurement channel, and the values for $l = N_{nyq} + 1, \dots, N$ are the conjugate mirror images of the values for $l = 2, \dots, N_{nyq}$. Hence, they add no more information.

where $\boldsymbol{B}_t$ is the n-dimensional Brownian motion. The Langevin diffusion process converges to the stationary distribution $\pi(\boldsymbol{\theta}|D)$. The actual implementation of the process is to discretize Eq. (12) using the Euler-Maruyama scheme as

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \frac{\varepsilon^2}{2}\nabla \log \pi(\boldsymbol{\theta}_t|D) + \varepsilon \boldsymbol{z}_t \tag{13}$$

where $\boldsymbol{z}_t \sim N(\boldsymbol{0}, \boldsymbol{I}_{n_\mathrm{p}})$ and ε is the discretization step. Due to discretization error the Markov chain of Eq. (13) behaves different from the diffusion process of Eq. (12) and may not converge to $\pi(\boldsymbol{\theta}|D)$ as its invariant measure [12]. A MH update step can be adopted after each iteration of the discretization algorithm to avoid this discrepancy and guarantee the reversibility of the chain with respect to $\pi(\boldsymbol{\theta}|D)$. Algorithm 1 represents a step-by-step implementation of the random walk Metropolis-Hastings algorithm.

In Algorithm 1, if the instrumental density is replaced by $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_t) = N(\boldsymbol{\theta}_t + \frac{\varepsilon^2}{2}\nabla \log \pi(\boldsymbol{\theta}_t|D), \varepsilon \boldsymbol{I}_{n_\mathrm{p}})$, the resulting MCMC algorithm is called the MALA algorithm. As was discussed in the introduction, proposals obtained by the MALA algorithm consist of two terms, where the drift term, which is the scaled gradient of the negative logarithm of the target distribution, drives the chain towards the high probability region. As a consequence, the MALA algorithm explores the invariant distribution faster, and outperforms outperforming RWHM. Since the second term of the MALA algorithm is an isotropic Brownian motion, highly correlated parameter structures place a constraint on the discretization step size to be smaller than the standard deviation in the most constrained direction. This issue leads the MALA algorithm to explore such parameter spaces slowly. The performance of the MALA algorithm is improved by exploiting the geometric structure of the parameter space [12]. Thus, Girolami *et al.* [12] defined the Itô stochastic differential equation of the Langevin diffusion on a Riemann manifold with the metric tensor $\boldsymbol{G}(\boldsymbol{\theta})$ as

$$d\boldsymbol{\theta}_t = \sqrt{\boldsymbol{G}^{-1}(\boldsymbol{\theta}_t)}d\boldsymbol{B}_t + \boldsymbol{b}(\boldsymbol{\theta}_t)dt \tag{14}$$

where

$$\begin{aligned} b_i &= \frac{1}{2}\sum_{j=1}^{n} \boldsymbol{G}_{ij}^{-1}(\boldsymbol{\theta}_t)\frac{\partial}{\partial \boldsymbol{\theta}_j}\log \pi(\boldsymbol{\theta}_t|D) \\ &+ |\boldsymbol{G}(\boldsymbol{\theta})|^{-1/2}\sum_{j=1}^{n}\frac{\partial}{\partial \boldsymbol{\theta}_j}\left(\boldsymbol{G}_{ij}^{-1}(\boldsymbol{\theta}_t)|\boldsymbol{G}(\boldsymbol{\theta})|^{1/2}\right) \end{aligned} \tag{15}$$

and the metric tensor $G(\boldsymbol{\theta})$ is the Hessian of the negative logarithm of the invariant PDF, *i.e.* $-\nabla^2 \log(\pi(\boldsymbol{\theta}|D))$. This Itô stochastic equation has $\pi(\boldsymbol{\theta}|D)$ as its unique stationary distribution. The first term on the right hand side of Eq. (14) transforms the isotropic Brownian motion into a new coordinate system that is aligned according to the local curvature of the Riemann manifold. On the other hand, the second term in Eq. (15) involves the derivative of the local curvature of the manifold. This term is expensive to evaluate and reduces to zero if one assumes that the manifold curvature is locally constant. For a flat manifold with a constant curvature, Eq. (14) can be discretized using the Euler-Maruyama integrator such that

$$\begin{aligned} \boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t &+ \frac{\varepsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta}_t)\nabla \log \pi(\boldsymbol{\theta}_t|D) \\ &+ \varepsilon\sqrt{\boldsymbol{G}^{-1}(\boldsymbol{\theta}_t)}\boldsymbol{z}_t \end{aligned} \tag{16}$$

This proposal mechanism defines a new instrumental density $q(\boldsymbol{\theta}^*|\boldsymbol{\theta}_t) = N(\boldsymbol{\theta}_t + \frac{\varepsilon^2}{2}\boldsymbol{G}^{-1}(\boldsymbol{\theta}_t)\nabla \log \pi(\boldsymbol{\theta}_t), \varepsilon\sqrt{\boldsymbol{G}^{-1}(\boldsymbol{\theta}_t)}\boldsymbol{I}_{n_\mathrm{p}})$ which can be employed in algorithm 1 in conjunction with the MH step to ensure that the new MCMC method, denoted the simplified manifold MALA, smMALA, converges to the invariant distribution $\pi(\boldsymbol{\theta}|D)$. One can expect that the smMALA explores the target distribution faster than the MALA algorithm, provided that the Hessian of the target distribution is positive definite and the scale factor ε^2 is sufficiently small. Thus, the appropriate tuning of the scaling factor is of particular interest in practice.

3.2 Starting from high probability region

smMALA is a local MCMC algorithm, meaning that it is not able to explore posterior PDFs with well-separated regions of high probability. However, in practice, there is often one important region with high probability content. Thus, initiating the chain in this region leads to efficient exploration of the target distribution. To this end, a simulated annealing algorithm is employed in this study to find an appropriate starting point $\boldsymbol{\theta}_0$ in the high probability region.

3.3 Numerical approximation of Hessian

The smMALA algorithm requires computation of the full Hessian of the negative log posterior, *i.e.*,

$$\begin{aligned} \phi(\boldsymbol{\theta}) = \frac{1}{2}\left\|\mathrm{vect}(\widehat{\boldsymbol{\vartheta}}) - \mathrm{vect}(\boldsymbol{\vartheta}(\boldsymbol{\theta}))\right\|^2_{\boldsymbol{C}^{-1}_{\mathrm{noise}}} \\ + \frac{1}{2}\left\|\boldsymbol{\theta} - \overline{\boldsymbol{\theta}}\right\|^2_{\boldsymbol{C}^{-1}_{\mathrm{prior}}} + \mathrm{constant} \end{aligned} \tag{17}$$

In this formulation, the noise covariance function can be decomposed as $\boldsymbol{C}_{\mathrm{noise}} = \boldsymbol{C}\boldsymbol{C}^T$ using either the symmetric square root or Cholesky factorization. Let the first term in Eq. (17) be called the misfit. This function can be rewritten as

$$\frac{1}{2}\left\|\boldsymbol{C}^{-1}\left(\mathrm{vect}(\widehat{\boldsymbol{\vartheta}}) - \mathrm{vect}(\boldsymbol{\vartheta}(\boldsymbol{\theta}))\right)\right\|^2 = \frac{1}{2}\|\boldsymbol{r}(\boldsymbol{\theta})\|^2 \tag{18}$$

We assume that the predicted FRFs are twice continuously Fréchet differentiable functions, and we denote the Jacobian of the misfit function by $\boldsymbol{J}(\boldsymbol{\theta})$. Thus, the gradient and Hessian of $\phi(\boldsymbol{\theta})$ can be written as

$$\nabla\phi(\boldsymbol{\theta}) = \boldsymbol{J}^T(\boldsymbol{\theta})\boldsymbol{r}(\boldsymbol{\theta}) + \boldsymbol{C}^{-1}_{\mathrm{prior}}(\boldsymbol{\theta} - \overline{\boldsymbol{\theta}}) \tag{19}$$

$$\nabla^2\phi(\boldsymbol{\theta}) = \boldsymbol{J}^T(\boldsymbol{\theta})\boldsymbol{J}(\boldsymbol{\theta}) + \boldsymbol{Q}(\boldsymbol{\theta}) + \boldsymbol{C}^{-1}_{\mathrm{prior}} \tag{20}$$

where $\mathbf{H}_{\mathrm{misfit}} = \boldsymbol{J}^T(\boldsymbol{\theta})\boldsymbol{J}(\boldsymbol{\theta}) + \boldsymbol{Q}(\boldsymbol{\theta})$ is the Hessian of the misfit function, and $\boldsymbol{Q}(\boldsymbol{\theta}) = \sum_{i=1}^{m}\boldsymbol{r}_i(\boldsymbol{\theta})\nabla^2\boldsymbol{r}_i(\boldsymbol{\theta})$ denotes second order terms of the Hessian of the misfit function. In general, computation of the full Hessian of the misfit function,

$$\mathrm{var}[\bar{h}] = E\left[\left(\frac{1}{N_s}\sum_{t=1}^{N_s} h(\boldsymbol{\theta}_t) - E[h(\boldsymbol{\theta})]\right)^2\right] \tag{27}$$

$$= \frac{1}{{N_s}^2}\sum_{k=1}^{N_s}\sum_{t=1}^{N_s} E[(h(\boldsymbol{\theta}_k) - E[h(\boldsymbol{\theta})])(h(\boldsymbol{\theta}_t) - E[h(\boldsymbol{\theta})])]$$

Using the definition of the autocovariance function, *i.e.* $\gamma(\tau) = E[(h(\boldsymbol{\theta}_k) - E[h(\boldsymbol{\theta})])(h(\boldsymbol{\theta}_{k+\tau}) - E[h(\boldsymbol{\theta})])]$, the above formula can be rewritten as

$$\mathrm{var}[\bar{h}] = \frac{\gamma(0)}{N_s}(1+\eta) \tag{28}$$

where $\eta = 2\sum_{\tau=1}^{N_s}(1-\tau/N_s)\frac{\gamma(\tau)}{\gamma(0)}$ and the autocovariance function can be approximated using the following expression:

$$\gamma(\tau) \approx \frac{1}{N_s-\tau}\sum_{k=1}^{N_s-\tau}\big(h(\boldsymbol{\theta}_{k+\tau}) - \bar{h}\big)\big(h(\boldsymbol{\theta}_k) - \bar{h}\big) \tag{29}$$

From Eq. (28) one can show that $\gamma(0)/N_s \le \mathrm{var}[\bar{h}] \le \gamma(0)$. The lower bound of this inequality is achieved if the samples are independent while the upper bound is achieved for perfectly correlated samples. Thus, the effective sample size, which is the effective number of independent samples, can be computed by $ESS = N_s/(1+\eta)$.

4 APPLICATION TO FE MODEL UPDATING

This section presents a numerical example to demonstrate the efficiency of the smMALA algorithm to solve the FE model updating problem with high dimensional parameter spaces.

4.1 6DOF mass-spring-damper system

Figure 1 illustrates a 6DOF mass-spring system. A relative modal damping of 1% is assigned to all modes of the model. Suppose that the system is excited through the sixth degree of freedom and the noisy acceleration data, simulated in this study, is measured at all degrees of freedom. The acceleration data is contaminated with a small amount of noise, 8% RMS noise-to-signal ratio, not to violate the small-noise-level assumption made for the Gauss-Newton approximation of Hessian. FE model updating is to be performed to estimate all the masses, m_i, and the stiffnesses, k_i, of the system. Thus, we need to estimate 16 parameters in this problem.

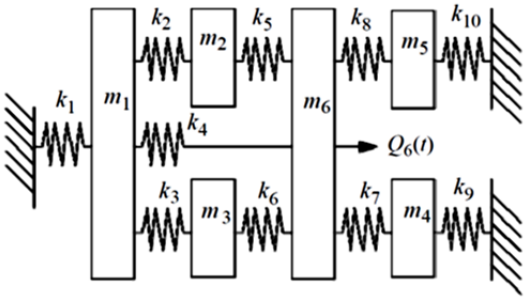

Figure 1. 6DOF mass-spring system with 16 parameters

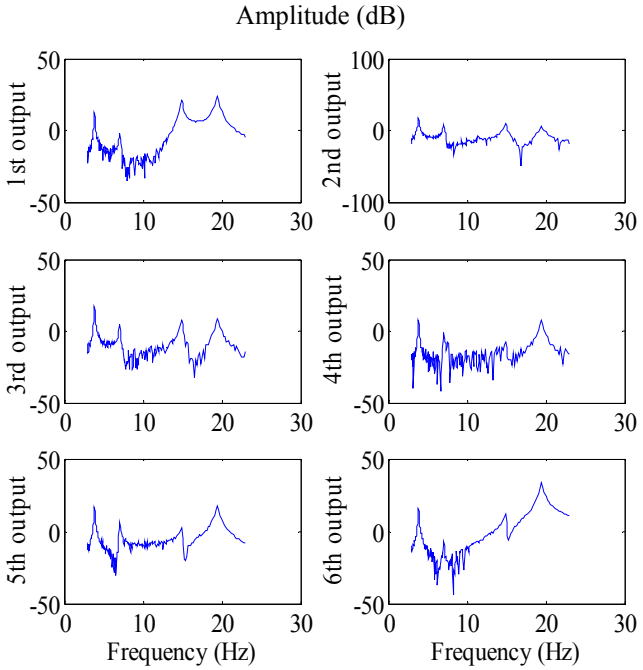

Figure 2. Magnitude of the simulated FRFs (acceleration as output) for 8% RMS noise-to-signal ratio.

Figure 2 depicts the amplitude of the measured FRF data within the frequency range [3, 23] Hz. To ensure that all resonance frequencies within this frequency range are measured with the same level of accuracy, the same number of frequency lines should be selected in the half-band-width around each resonance frequency. This implies taking constant steps in a logarithmic frequency scale, resulting in 206 frequency lines with the frequency resolution of 0.03 Hz in order to have 3 frequency lines within each half-band-width.

The likelihood function of Eq. (10) can be rewritten for this problem as follows

$$\pi(D|\boldsymbol{\theta}) = \frac{1}{(\pi\sigma^2)^{N_q N_o}} \times \exp\left(-\frac{1}{\sigma^2}\left\|\mathrm{vect}(\hat{\boldsymbol{\vartheta}}) - \mathrm{vect}(\boldsymbol{\vartheta}(\boldsymbol{\theta}))\right\|^2\right) \tag{30}$$

Here, $N_o = 6$, $N_q = 206$, $\sigma = 0.057$ and $\|U_l\|^2 = 1, l = 1, \dots, N_q$. It should be noted that this model updating problem is globally identifiable in the sense that there is exactly one Maximum Likelihood Estimate in the parameter space. The prior PDF for the parameter vector $\boldsymbol{\theta}$ is selected as independent Gaussian distributions with means equal to the nominal values $m_{i0}, i = 1, \dots, 6$, and $k_{i0}, i = 1, \dots, 10$, see Table 1, and coefficient of variation (C.o.V) of 10%. The physical parameters, *i.e.*, $m_i, i = 1, \dots, 6$, and $k_i, i = 1, \dots, 10$ are related to the normalized parameter vector $\boldsymbol{\theta}$ and their nominal values such that $m_i = m_{i0}(1+\theta_i), i = 1, \dots, 6$ and $k_i = k_{i0}(1+\theta_i), i = 1, \dots, 10$.

The smMALA chain is initiated in the high probability region of the posterior PDF, $\pi(\boldsymbol{\theta}|D)$, using the simulated annealing algorithm Global Optimization Toolbox of MATLAB by 10000 evaluations of $\pi(\boldsymbol{\theta}|D)$. Then the smMALA algorithm is applied to explore the posterior PDF.

Table 1. Statistical results for structural parameters estimates for 8% RMS noise-to-signal ratio.

Parameter	α_i = nominal value	τ_i = true value	μ_i = mean value of estimates	σ_i/μ_i = c.o.v. estimate of parameters	Error = $\frac{\lvert\tau_i-m_i\rvert}{\tau_i}$
k_1	3810	3600	3665.3	1.29%	1.81%
k_2	1825	1725	1738.9	3.86%	0.80%
k_3	1100	1200	1198.1	4.58%	0.15%
k_4	2000	2200	2214.8	0.60%	0.67%
k_5	1620	1320	1309.8	3.98%	0.77%
k_6	1130	1330	1300.9	4.36%	2.18%
k_7	1000	1500	1413.3	3.33%	5.78%
k_8	5950	5250	5326.6	0.95%	1.45%
k_9	3200	3600	3416.0	3.31%	5.11%
k_{10}	650	850	871.2	3.37%	2.49%
m_1	1.5	1	1.01	0.74%	1.01%
m_2	1.2	1.4	1.41	3.79%	0.71%
m_3	1.5	1.2	1.18	4.46%	1.66%
m_4	2.1	2.2	2.07	3.27%	5.90%
m_5	2.1	2.5	2.53	1.17%	1.20%
m_6	0.7	0.9	0.89	0.10%	1.11%

Figure 4 demonstrates the autocorrelation function for m_1 and k_1 where the first 3000 samples are discarded as the burn-ins. This figure also confirms that the correlation between neighboring samples generated by smMALA is small.

Table 1 shows the nominal value, exact value, sample mean, sample C.o.V and the estimation error of the uncertain parameters. Two observations are immediate. First, the uncertainty in the parameters is reduced in comparison to the prior uncertainty inasmuch as the measured data provides information about these parameters. Second, the estimation error is small: $0.15-5.78\%$ for stiffness parameters and $0.71-5.9\%$ for mass parameters.

5 CONCLUSION

This paper presented a simplified manifold Metropolis adjusted Langevin algorithm to solve Bayesian FE model updating problems with large scale parameter spaces. The proposal density in the smMALA involves two parts. The deterministic part, which is a scaled Newton step in deterministic optimization and takes the current state of the chain towards the high probability region of the target density given that the Hessian is positive definite. The stochastic term is distributed using an *n*-dimensional Gaussian distribution fitted to the local structure of the target density using the Hessian of the negative log posterior as the inverse covariance matrix. The Gauss-Newton approximation of the Hessian matrix guarantees that the geometric tensor is positive semidefinite and that the drift term finds its way towards the high probability region. It also reduces the computational time of Hessian matrix.

The smMALA algorithm was presented and its features were discussed in details. An illustrative example demonstrated that the presented MCMC method is capable of generating samples from a high dimensional posterior distribution when the structural model class is identifiable. One possible extension of this work is to improve the smMALA to make it an enabling methodology to solve Bayesian model updating problems with unidentifiable parameters.

REFERENCES

[1] Gelman A., Carlin J., Stern H., Rubin D., Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science), Chapman and Hall/CRC, 2003.

[2] Cheung S., Beck J., Bayesian Model Updating Using Hybrid Monte Carlo Simulation with Application to Structural Dynamic Models with Many Uncertain Parameters, Journal of Engineering Mechanics, 135(4), 243, 2009.

[3] Ching J., Chen Y.C., Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, Journal of Engineering Mechanics, 133(7), 816-832, 2007.

[4] Ching J., Muto M., Beck J.L., Structural model updating and health monitoring with incomplete modal data using Gibbs sampler, Computer-Aided Civil and Infrastructure Engineering, 21(4), 242-257, 2006.

[5] Duane S., Kennedy A.D., Pendleton B.J., Roweth D., Hybrid Monte Carlo, Physics Letters B, 195(2), 216-222, 1987.

[6] Martin J., Wilcox L., Burstedde C., Ghattas O., A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion, SIAM Journal on Scientific Computing, 34(3), A1460-A1487, 2012.

[7] Neal R.M., MCMC using Hamiltonian dynamics, 2010.

[8] Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E., Equation of state calculations by fast computing machines, The Journal of Chemical Physics, 21(6), 1087-1092, 1953.

[9] Hastings W.K., Monte carlo sampling methods using Markov chains and their applications, Biometrika, 57(1), 97-109, 1970.

[10] Robert C., Casella G., Monte Carlo Statistical Methods, Springer-Verlag, 1999.

[11] Roberts G., Rosenthal J., Examples of adaptive MCMC, Stat. and Comput, 18(2), 349-367, 2009.

[12] Girolami M., Calderhead B., Riemann manifold Langevin and Hamiltonian Monte Carlo methods, Journal of the Royal Statistical Society. Series B: Statistical Methodology, 73(2), 123-214, 2011.

[13] Roberts G.O., Stramer O., Langevin Diffusions and Metropolis-Hastings Algorithms, Methodology and Computing in Applied Probability, 4(4), 337-357, 2002.

[14] Pintelon R., Schoukens J., System Identification: A Frequency Domain Approach, Second ed., Wiley-IEEE Press, 2012.

[15] Jaynes E.T., Information Theory and Statistical Mechanics, Physical Review, 106(4), 620-630, 1957.

[16] Beck J., Au S.-K., Bayesian Updating of Structural Models and Reliability using Markov Chain Monte Carlo Simulation, Journal of Engineering Mechanics, 128(4), 380-391, 2002.