6th International Conference on Advances in Experimental Structural Engineering
11th International Workshop on Advanced Smart Materials and Smart Structures Technology
August 1-2, 2015, University of Illinois, Urbana-Champaign, United States

Fast Bayesian Model Selection with Application to Large Locally-Nonlinear Dynamic Systems

S. De[1], E.A. Johnson[2], S.F. Wojtkiewicz[3]

1 *Ph.D. Fellow, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA, USA. E-mail: Subhayan.De@usc.edu*
2 *Professor, Sonny Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA, USA. E-mail: JohnsonE@usc.edu*
3 *Associate Professor, Department of Civil and Environmental Engineering, Clarkson University, Potsdam, NY, USA. E-mail: swojtkie@clarkson.edu*

ABSTRACT
Bayesian model selection chooses, based on measured data, using Bayes' theorem, suitable mathematical models from a set of possible models. In structural analysis, linear models are often used to facilitate design and analysis, though they do not always accurately reproduce actual structural responses. When the models also require the inclusion of nonlinearity to improve accuracy, the computation time required for response simulation increases significantly. To reduce this computational burden, this paper proposes incorporating into the model selection process an efficient dynamic response algorithm previously developed by the last two authors for locally nonlinear systems. Additionally, nested sampling, an intelligent sampling algorithm, is used to reduce the number of simulations (using whichever response simulation algorithm) needed for accurate posterior distribution computation. A numerical example of a 20-story three-dimensional building with roof-mounted tuned mass dampers (TMDs), using different linear and nonlinear damping models as the candidates to reproduce the TMD damping, demonstrates that the proposed approach is up to 1000 times faster than traditional Bayesian model selection employing a conventional structural response solver.

KEYWORDS: *Bayesian model selection, nonlinear Volterra integral equation, nested sampling.*

1. INTRODUCTION

Often in the dynamic characterization of structural systems, different models are available to reflect the behavior of certain components of the structure. In structural control, design of the controller is only as good as the modeling of the system; likewise, in structural health monitoring, the accurate modeling of the structure is vital since damage detection is generally performed by identifying changes in the model parameters. The choice of best model(s) must be made after eliminating models and models that can be shown to be incorrect based on experimental measurements. In Bayesian model selection, the plausibility of each model given the measurement data is determined using Bayes' theorem. For structural examples, Bayesian model selection can exploit response measurement data to identify the best possible model(s) for further calculations [1–5]. Smith and Saitta [6] used a global search method in a probabilistic framework for model selection. Yuen [7] applied Bayesian model selection to problems of air quality prediction, hydraulic jump and seismic attenuation relationships. The selection procedure requires the evaluation of a multi-dimensional integral known as the *evidence* or the *marginal likelihood*.

A number of methods have been introduced to reduce the computational requirement of Bayesian model selection. The *posterior harmonic mean estimator* of Newton and Raftery [8] samples from posterior distributions of parameters to evaluate the integral. Kass and Raftery [9] applied importance sampling to improve the efficiency of simple Monte Carlo sampling. Two methods to calculate evidence, closely related to simulated annealing and thermodynamic integration, are *annealed importance sampling* [10] and the *power posterior* method [11]. Ching and Chen [12] introduced the *transitional Markov chain Monte Carlo* method where samples are taken from intermediate distributions ultimately converging to a target distribution. The nested sampling of Skilling [13] increases samples from the high likelihood region. Nevertheless, even with intelligent sampling, the computational cost for evaluating the *evidence* using Monte Carlo sampling becomes burdensome, if not computationally prohibitive, for large nonlinear dynamic systems, which require a computationally-intensive simulation multiplied by the necessary number of realizations for accurate statistical characterization.

Herein, the computational effort of determining probabilistic response metrics of such structures with local nonlinearities is reduced by: (a) exactly reducing the system equations of motion into a set of low-order nonlinear Volterra integral equations that is solved numerically [14] to evaluate the system responses by exploiting the localized nature of the nonlinearities; and (b) using an intelligent Monte Carlo sampling algorithm — *nested sampling*

[13] — to direct the computational simulations to those that make the evidence calculation most effective. The proposed approach is illustrated through its application to a numerical example of a 1623 degree-of-freedom (DOF) model of a three dimensional building with nonlinear tuned mass dampers (TMDs) attached on the roof. Different nonlinear damping models are used as candidates to model the TMD damping; the computational efficiency of the proposed approach is compared to MATLAB's `ode45`, demonstrating three-orders-of-magnitude reductions in computation time.

2. METHODOLOGY

The proposed method consists of two parts: the Bayesian model selection with intelligent sampling, and the efficient response analysis of systems with local nonlinearities as described below.

2.1. Bayesian Model Selection

Let $\mathscr{M}_1$, $\mathscr{M}_2$,..., $\mathscr{M}_{N_\mathrm{m}}$ be the N_m different models for a particular structural problem. The uncertain parameter vector for each model is $\boldsymbol{\theta}^{(k)}$; $k = 1,2,\ldots,N_\mathrm{m}$. Given the measurement data set $\mathfrak{D}$, the posterior probability of models is evaluated for every model $\mathscr{M}_k$,

$$\mathbb{P}(\mathscr{M}_k|\mathfrak{D}) = \frac{\mathrm{p}(\mathfrak{D}|\mathscr{M}_k)\mathbb{P}(\mathscr{M}_k)}{\mathrm{p}(\mathfrak{D})}; \quad k = 1,2,\ldots,N_\mathrm{m}. \tag{2.1}$$

where the denominator probability density $\mathrm{p}(\mathfrak{D}) = \sum_{k=1}^{N_\mathrm{m}} \mathrm{p}(\mathfrak{D}|\mathscr{M}_k)\mathbb{P}(\mathscr{M}_k)$. The numerator probability $\mathbb{P}(\mathscr{M}_k)$ is an *a priori* measure of plausibility assigned to model $\mathscr{M}_k$, where $\sum_{k=1}^{N_\mathrm{m}} \mathbb{P}(\mathscr{M}_k) = 1$, based on the user's past experience and the problem at hand.

However, the main challenge in Bayesian model selection problems lies in evaluating the evidence $\mathscr{E}^{(k)} = \mathrm{p}(\mathfrak{D}|\mathscr{M}_k)$. For a particular model $\mathscr{M}_k$, this evidence can be written as

$$\begin{aligned} \mathscr{E}^{(k)} &= \int_{\boldsymbol{\Theta}^{(k)}} \mathrm{p}(\mathfrak{D}|\boldsymbol{\theta}^{(k)},\mathscr{M}_k)\mathrm{p}(\boldsymbol{\theta}^{(k)}|\mathscr{M}_k)d\boldsymbol{\theta}^{(k)} \\ &= \int_{\boldsymbol{\Theta}^{(k)}} \mathscr{L}(\boldsymbol{\theta}^{(k)},\mathscr{M}_k)\mathrm{p}(\boldsymbol{\theta}^{(k)}|\mathscr{M}_k)d\boldsymbol{\theta}^{(k)} \end{aligned} \tag{2.2}$$

where $\mathscr{L}(\boldsymbol{\theta}^{(k)},\mathscr{M}_k)$ is the likelihood function and $\mathrm{p}(\boldsymbol{\theta}^{(k)}|\mathscr{M}_k)$ is the parameters' prior density function (again based on prior knowledge or expert judgement).

To efficiently evaluate the evidence, a variety of intelligent sampling algorithms could be used. Herein, the *nested sampling* algorithm [13] is utilized. This algorithm treats $\mathscr{L}$ as a nonnegative random variable; it's cumulative distribution function can be written, omitting superscript $^{(k)}$ and model $\mathscr{M}_k$ for clarity, as

$$\mathrm{P}_\mathscr{L}(\lambda) = \mathbb{P}[\mathscr{L}(\boldsymbol{\theta}) < \lambda] = \int_{\mathscr{L}(\boldsymbol{\theta})<\lambda} \mathrm{p}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \tag{2.3}$$

The *evidence*, which is the expected value of the likelihood, is estimated using the result that the expected value of a nonnegative random variable X is equal to the complementary area under its probability distribution curve; *i.e.*, $\mathbb{E}[X] = \int_0^\infty [1 - \mathrm{P}_X(x)]\mathrm{d}x$. This allows the evidence integral Eq. 2.2 (omitting $\mathscr{M}_k$) to be rewritten as

$$\int_{\boldsymbol{\Theta}} \mathscr{L}(\boldsymbol{\theta})\mathrm{p}(\boldsymbol{\theta})d\boldsymbol{\theta} = \mathbb{E}_{\boldsymbol{\theta}}[\mathscr{L}(\boldsymbol{\theta})] = \int_0^\infty [1 - \mathrm{P}_\mathscr{L}(\lambda)]\mathrm{d}\lambda = \int_0^\infty \chi(\lambda)\mathrm{d}\lambda \tag{2.4}$$

where the $\chi(\lambda)$ is a monotonically decreasing function that is the probability mass enclosed in the subset of parameter space $\boldsymbol{\Theta}$ where likelihood $\mathscr{L}(\boldsymbol{\theta})$ exceeds λ; *i.e.*, $\chi(\lambda) = \int_{\mathscr{L}(\boldsymbol{\theta})>\lambda} \mathrm{p}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$ with $\chi(\infty) = 0$ and $\chi(0) = 1$. Defining the inverse function $\phi(\chi(\lambda)) = \lambda$, the *evidence* can then be approximated using a quadrature rule as

$$\mathscr{E} = \int_0^1 \phi(\chi)\mathrm{d}\chi \approx \sum_i \phi_i w_i \tag{2.5}$$

This numerical integration can be performed by taking $w_i = \chi_i - \chi_{i-1}$. Starting with prior volume $\chi_0 = 1$, N_s

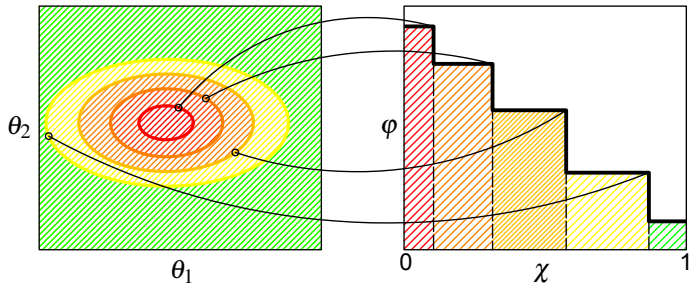

Figure 2.1 The *nested sampling* algorithm with a two-dimensional parameter space.

samples are drawn from the prior $\mathrm{p}(\boldsymbol{\theta})$. At each iteration, a new sample is drawn from the prior with $\mathscr{L}_{\text{new}} > \mathscr{L}_i$, where $\mathscr{L}_i$ is the lowest likelihood value of the current sample pool; *i.e.*, $\mathscr{L}_i = \min_j \mathscr{L}(\boldsymbol{\theta}_j)$. The prior volume corresponding to this new sample is $\chi_{i+1} = \tau_{i+1}\chi_i$, where τ follows the probability density $\mathrm{p}(\tau) = N_\mathrm{s}\tau^{(N_\mathrm{s}-1)}$ for the largest of N_s samples drawn from the standard uniform distribution $\mathscr{U}(0,1)$. The means of τ and $\ln\tau$ are, respectively, $\mathbb{E}[\tau] = N_\mathrm{s}/(N_\mathrm{s}+1)$ and $\mathbb{E}[\ln\tau] = -1/N_\mathrm{s}$. Either $\chi_i = [N_\mathrm{s}/(N_\mathrm{s}+1)]^i$ or $\chi_i = \exp(-i/N_\mathrm{s})$ can be used for simplicity with a deterministic approach [15]. The stopping criterion is when either i exceeds $n_{\max}$ or the change in evidence $\mathscr{E}$ is less than 1%.

The standard Monte Carlo (MC) estimate for the evidence is $\mathscr{E}^{(k)} = \frac{1}{N_\mathrm{s}}\sum_{i=1}^{N_\mathrm{s}} \mathscr{L}(\boldsymbol{\theta}_i)$ where $\{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{N_\mathrm{s}}\}$ are samples from prior $\mathrm{p}(\boldsymbol{\theta})$. Since the high likelihood region is generally very different compared to the supports of the prior, the variance of the standard MC estimator will be high unless a large number of samples are used [12]. On the other hand, nested sampling samples more from the high likelihood region. Thus, less variance can be expected for this estimator compared to MC if same number of samples are used.

2.2. Efficient Analysis of Systems with Local Nonlinearities

Let the nonlinear model follows the following state space equation,

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}\mathbf{g}(\overline{\mathbf{X}}(t)), \qquad \mathbf{X}(0) = \mathbf{x}_0 \tag{2.6}$$

where $\mathbf{X}(t)$ is the $n \times 1$ state vector, $\mathbf{A}$ is the $n \times n$ state matrix, $\mathbf{u}$ is a $m \times 1$ external excitation, $\mathbf{B}$ is the $n \times m$ influence matrix, $\mathbf{g}(\cdot)$ is a $n_\mathrm{g} \times 1$ nonlinear function of an $n_\mathrm{o} \times 1$ subset of states $\overline{\mathbf{X}}(t) = \mathbf{G}\mathbf{X}(t)$, $\mathbf{L}$ is a $n \times n_\mathrm{g}$ influence matrix which maps the nonlinear force vector to all the states, and $\mathbf{x}_0$ is the initial condition.

The nominal linear system corresponding to the nonlinear system Eq. 2.6 is

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0 \tag{2.7}$$

The response of nonlinear system Eq. 2.6 is calculated by the superposition of $\mathbf{x}(t)$, the solution of linear system Eq. 2.7, and $\mathbf{x}^{(\mathrm{nl})}(t)$ due to the nonlinear forcing function $\mathbf{g}(\overline{\mathbf{X}}(t))$. Hence,

$$\begin{aligned} \mathbf{x}(t) &= \exp(\mathbf{A}t)\mathbf{x}_0 + \int_0^t \mathbf{H}_\mathrm{B}(t-s)\mathbf{u}(s)\mathrm{d}s, \\ \mathbf{x}^{(\mathrm{nl})}(t) &= \int_0^t \mathbf{H}_\mathrm{L}(t-s)\mathbf{g}(\overline{\mathbf{X}}(s))\mathrm{d}s \end{aligned} \tag{2.8}$$

where $\mathbf{H}_\mathrm{B}(t) = \mathrm{e}^{\mathbf{A}t}\mathbf{B}$ and $\mathbf{H}_\mathrm{L}(t) = \mathrm{e}^{\mathbf{A}t}\mathbf{L}$ are impulse responses of the nominal system; and total response is $\mathbf{X}(t) = \mathbf{x}(t) + \mathbf{x}^{(\mathrm{nl})}(t)$. The solution of $\mathbf{x}^{(\mathrm{nl})}$ can be done efficiently with the following:

$$\begin{aligned} \mathbf{p}(t) &= \mathbf{g}(\overline{\mathbf{X}}(t)); \\ \overline{\mathbf{X}}(t) &= \overline{\mathbf{x}}(t) + \int_0^t \overline{\mathbf{H}}_\mathrm{L}(t-s)\mathbf{p}(s)\mathrm{d}s \end{aligned} \tag{2.9}$$

where $\overline{\mathbf{x}}(t) = \mathbf{G}\mathbf{x}(t)$, $\overline{\mathbf{H}}_{\mathrm{L}}(t) = \mathbf{G}\mathbf{H}_{\mathrm{L}}(t)$. Equations Eq. 2.9 can be combined to give

$$\mathbf{p}(t) - \mathbf{g}\left(\overline{\mathbf{x}}(t) + \int_0^t \overline{\mathbf{H}}_{\mathrm{L}}(t-s)\mathbf{p}(s)\mathrm{d}s\right) = \mathbf{0} \tag{2.10}$$

The system of equations Eq. 2.10 is a vector nonlinear Volterra type integral equation (NVIE) written in non-standard form. Hence, the nonlinear system in Eq. 2.6 is exactly reduced to low-order NVIE Eq. 2.10. In the algorithm proposed in [14], a combination of a Newton-Gregory integration scheme along with fast Fourier transforms (FFTs) is applied to compute the required convolutions in Eq. 2.10, which dramatically reduces the required computational cost to solve first for the nonlinear force vector time history $\mathbf{p}(t)$ and, subsequently, for the desired system outputs using any standard linear system solution (*e.g.*, time marching, FFTs, etc.).

3. NUMERICAL EXAMPLE: THREE-DIMENSIONAL WIND-EXCITED STRUCTURE

The 20-story moment-resisting frame building model with three nonlinear TMDs, shown in Figure 3.1, adapted from Wojtkiewicz and Johnson [16], is considered here to evaluate the efficacy of the proposed method. The building has a height of 80 m, and has a 40 m $\times$ 30 m rectangular 5-bay $\times$ 3-bay plan in the first five stories, then 3-bay $\times$ 2-bay in stories 6–10, and 2-bay $\times$ 2-bay in the top half of the building. Cross braces carrying axial loads provide additional stiffness for lateral bending and torsion. Euler-Bernoulli beams are used to model columns and floor beams; the beam-column joints are assumed rigid. Additional in-plane stiffness of the floors is provided by additional cross elements on the floor.

Without the TMDs, the structure model with 1,620 DOF has its first six modal frequencies at 0.5718 Hz (first y-direction mode), 0.5893 Hz (first x-direction mode), 0.9363 Hz (first torsional mode), 1.3632 Hz (second y-direction mode), 1.5346 Hz (second x-direction mode), and 2.0292 Hz (second torsional mode). Two TMDs are placed in the y-direction (each about 0.55% of the building mass) and one TMD in the x-direction (about 1.1% of the building mass) which splits the first y mode into two modes with frequencies at 0.5062 and 0.6282 Hz, the first x mode into two modes with frequencies at 0.5214 and 0.6475 Hz, and the first torsional mode into two modes with frequencies at 0.5615 and 0.9506 Hz.

The building is subjected to one-directional wind excitation (oriented toward the east-northeast, at a 30° angle from

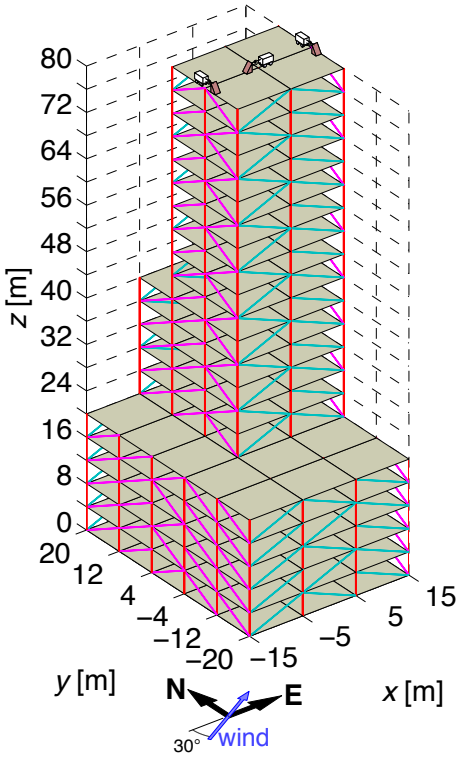

Figure 3.1 Complex three-dimensional wind-excited structure.

the x-axis as shown in Figure 3.1), modeled as a narrowband filtered Gaussian white noise process shaped vertically along the building height. The filter is a 16th order Butterworth band-pass filter with cutoff frequencies 1.2 times smaller and larger than the fundamental structural natural frequency, resulting in most of the energy in the range of 0.35–1.5 Hz, exciting primarily the fundamental mode in the east-west (E-W) x-direction and, secondarily, the fundamental modes in the north-south (N-S) y-direction and in torsion. The vertical power-law shaping of the wind excitation is proportional to the height to the 0.3 power [17]. The force is assumed for simplicity to be fully correlated at all heights along the building.

The *true* model, used to generate response measurements, has in each of the three TMDs a power-law damper that exerts a force

$$f_{\mathrm{d,P}} = c_{\mathrm{d,P}} |\dot{u}|^{\beta_{\mathrm{P}}} \operatorname{sgn}(\dot{u}) + c_{\mathrm{nom,P}} \dot{u} \tag{3.1}$$

where $\beta_{\mathrm{P}} = 0.8$. For the one TMD that moves in the in x direction, the damper coefficients are $c^x_{\mathrm{d,P}} = 200\ \mathrm{kN{\cdot}(s/m)}^{\beta_{\mathrm{P}}}$ and $c^x_{\mathrm{nom,P}} = 30\ \mathrm{kN{\cdot}(s/m)}$; for each of the two TMDs that move in the y direction, the coefficients are $c^y_{\mathrm{d,P}} = 100\ \mathrm{kN{\cdot}(s/m)}^{\beta}$ and $c^y_{\mathrm{nom,P}} = 15\ \mathrm{kN{\cdot}(s/m)}$. This true model is simulated with the wind excitation to produce output acceleration responses at the top-floor center-of-mass in the x and y directions and the torsional acceleration (which could be approximated with two non-collocated accelerometers). These measurements are sampled at 20 Hz and corrupted with 20% additive Gaussian white sensor noise.

3.1. Models for Damping of the TMDs

One linear and two nonlinear models are postulated to represent the nonlinear damping of the TMDs. Their details are given in Table 3.1. The models are assumed *a priori* to be equally likely. The priors for the parameters of different models are given in Table 3.2. The linear viscous damping model assumes that the damping force f_{d}, exerted by each TMD, is proportional to the relative velocity between the TMD and its roof attachment point. The power-law viscous damping model has the damping force proportional to the relative velocity raised to power β_{P} as well as a linear viscous damping term. In the cubic polynomial damping model, a cubic term, similar to the power law damping with $\beta_{\mathrm{P}} = 3$, is added to the linear term.

The result of the model selection, given in Table 3.3, clearly shows that the power law damping model emerges as the most probable, as expected, by achieving a posterior model probability of 0.98, whereas the next two best models are the linear and cubic polynomial damping. This is also expected as the *true* power law model ($\beta_{\mathrm{P}} = 0.8$) is close to a linear one.

Table 3.1 Nonlinear damping models, where u is a TMD displacement relative to its roof attachment point.

Model ($\mathscr{M}_k$)	Damping force	Parameters
$\mathscr{M}_1$ (Linear)	$f_{\mathrm{d,L}} = c_{\mathrm{d,L}} \dot{u}$	$c_{\mathrm{d,L}}$
$\mathscr{M}_2$ (Power law)	$f_{\mathrm{d,P}} = c_{\mathrm{d,P}} \lvert\dot{u}\rvert^{\beta_{\mathrm{P}}} \operatorname{sgn}(\dot{u}) + c_{\mathrm{nom,P}} \dot{u}$	$\begin{bmatrix} c_{\mathrm{d,P}} & \beta_{\mathrm{P}} & c_{\mathrm{nom,P}} \end{bmatrix}$
$\mathscr{M}_3$ (Cubic polynomial)	$f_{\mathrm{d,C}} = c_{\mathrm{d,C}} \dot{u}^3 + c_{\mathrm{nom,C}} \dot{u}$	$\begin{bmatrix} c_{\mathrm{d,C}} & c_{\mathrm{nom,C}} \end{bmatrix}$

Table 3.2 Prior distributions for model parameters for different damping models for 1623 DOF building structure subjected to wind load.

Model ($\mathscr{M}_k$)	Parameter	Distribution	Mean	Std. dev.
Linear ($\mathscr{M}_1$)	$c^x_{\mathrm{d,L}}$	Normal	350 kN·s/m	35.0 kN·s/m
	$c^y_{\mathrm{d,L}}$	Normal	225 kN·s/m	22.5 kN·s/m
Power law ($\mathscr{M}_2$)	$c^x_{\mathrm{d,P}}$	Normal	225 kN·(s/m)$^{\beta_{\mathrm{P}}}$	20 kN·(s/m)$^{\beta_{\mathrm{P}}}$
	$c^y_{\mathrm{d,P}}$	Normal	120 kN·(s/m)$^{\beta_{\mathrm{P}}}$	10 kN·(s/m)$^{\beta_{\mathrm{P}}}$
	$c^x_{\mathrm{nom,P}}$	Lognormal	27.5 kN·s/m	2.5 kN·s/m
	$c^y_{\mathrm{nom,P}}$	Lognormal	20.0 kN·s/m	2.0 kN·s/m
	β_{P}	Lognormal	0.85	0.05
Cubic polynomial ($\mathscr{M}_3$)	$c^x_{\mathrm{d,C}}$	Lognormal	75 kN·(s/m)^3	7.5 kN·(s/m)^3
	$c^y_{\mathrm{d,C}}$	Lognormal	20 kN·(s/m)^3	2.0 kN·(s/m)^3
	$c^x_{\mathrm{nom,C}}$	Normal	350 kN·s/m	35.0 kN·s/m
	$c^y_{\mathrm{nom,C}}$	Normal	225 kN·s/m	22.5 kN·s/m

Table 3.3 Posterior model probabilities $\mathbb{P}(\mathscr{M}_k|\mathfrak{D})$ with priors $\mathbb{P}(\mathscr{M}_k) = 1/3$.

Model ($\mathscr{M}_k$)	$\ln p(\mathfrak{D}|\mathscr{M}_k)$	$\mathbb{P}(\mathscr{M}_k|\mathfrak{D})$
$\mathscr{M}_1$ (Linear)	4920.9835	0.0093
$\mathscr{M}_2$ (Power law)	4925.6425	0.9828
$\mathscr{M}_3$ (Cubic polynomial)	4920.8230	0.0079

3.2. Computational Gain

The proposed method is implemented with 2^{14} time steps of $\Delta t = 1.83$ ms duration each, which results in acceleration responses that differ from that computed via MATLAB's `ode45` (with default tolerances) by $\mathscr{O}(10^{-3})$ on a relative RMS basis; since `ode45`'s default relative tolerance is also $\mathscr{O}(10^{-3})$, the two methods are of comparable accuracy. The simulations are run on a MacBook Pro, with a 2.3 GHz Intel Core i7 processor and 16 GB RAM, running MATLAB R2014b; `cputime` is used to calculate the required time of the solvers. The proposed method takes 562.830 s to execute the necessary one-time calculations and only 2.059 s for the calculations that must be repeated for each random sample; in contrast, `ode45` requires 47.66 min. per realization. The result is that the proposed method provides a computational gain of 5.06 for one simulation but far greater efficiency for problems such as this where responses must be evaluated for many random samples of the model parameters. For example, to evaluate the power-law damping model, $\sim$550 simulations are required, resulting in a speedup of 927.74.

4. CONCLUSIONS

Bayesian model selection helps in selecting the *best* models among competing ones based on measured data. However, the main obstacle still remains: the evaluation of the multi-dimensional integral, known as *evidence*; its computational burden increases rapidly, becoming possibly prohibitive, for large systems with nonlinearities. A fast computational methodology is proposed herein to reduce this computational burden by incorporating a NVIE approach [14] along with nested sampling. The efficacy of the proposed method for Bayesian model selection involving a large locally nonlinear model is demonstrated through its application to a 1623 DOF numerical structural example, resulting in a gain of three orders of magnitude in computational efficiency.

ACKNOWLEDGMENT

The authors gratefully acknowledge the partial support of this work by the National Science Foundation, through awards CMMI 13-44937 and 14-36018, and of the first author by a Provost Ph.D. Fellowship from the University of Southern California. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or the University of Southern California.

REFERENCES

1. Beck, J. L., and Yuen, K.-V. (2004). Model selection using response measurements: Bayesian probabilistic approach. *ASCE Journal of Engineering Mechanics*. **130:2**, 192–203.
2. Muto, M., and Beck, J. L. (2008). Bayesian updating and model class selection for hysteretic structural models using stochastic simulation. *Journal of Vibration and Control*. **14:1–2**, 7–34.
3. Beck, J. L. (2010). Bayesian system identification based on probability logic. *Structural Control and Health Monitoring*. **17:7**, 825–847.
4. Cheung, S. H., and Beck, J. L. (2010). Calculation of posterior probabilities for Bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Computer-Aided Civil and Infrastructure Engineering*. **25:5**, 304–321.
5. Saito, T., and Beck, J. L. (2010). Bayesian model selection for ARX models and its application to structural health monitoring. *Earthquake Engineering and Structural Dynamics*. **39:15**, 1737–1759.
6. Smith, I. F., and Saitta, S. (2008). Improving knowledge of structural system behavior through multiple models. *ASCE Journal of Structural Engineering*. **134:4**, 553–561.
7. Yuen, K.-V. (2010). *Bayesian Methods for Structural Dynamics and Civil Engineering*, John Wiley and Sons, Singapore.
8. Newton, M. A., and Raftery, A. E. (1994). Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*. **56:1**, 3–48.

9. Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*. **90:430**, 773–795.
10. Neal, R. M. (2001). Annealed importance sampling. *Statistics and Computing*. **11:2**, 125–139.
11. Friel, N., and Pettitt, A. N. (2008). Marginal likelihood estimation via power posteriors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. **70:3**, 589–607.
12. Ching, J., and Chen, Y.-C. (2007). Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *ASCE Journal of Engineering Mechanics*. **133:7**, 816–832.
13. Skilling, J. (2006). Nested sampling for general Bayesian computation. *Bayesian Analysis*. **1:4**, 833–859.
14. Gaurav, Wojtkiewicz, S. F., and Johnson, E. A. (2011). Efficient uncertainty quantification of dynamical systems with local nonlinearities and uncertainties. *Probabilistic Engineering Mechanics*. **26:4**, 561–569.
15. Chopin, N., and Robert, C. P. (2010). Properties of nested sampling. *Biometrika*. **97:3**, 741–755.
16. Wojtkiewicz, S. F., and Johnson, E. A. (2014). Efficient sensitivity analysis of structures with local modifications — Part I: Time domain responses. *ASCE Journal of Engineering Mechanics*. **140:9**, 04014067.
17. Holmes, J. D. (1996). Along wind response of lattice towers—III. Effective load distributions. *Engineering Structures*. **18:7**, 489–494.