SANDIA REPORT

SAND2014-19209
Unlimited Release
Printed October, 2014

Utilizing Soize's Approach to Identify Parameter and Model Uncertainties

Matthew S. Bonney and Matthew R. W. Brake

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2014-19209
Unlimited Release
Printed October, 2014

Utilizing Soize's Approach to Identify Parameter and Model Uncertainties

Matthew S. Bonney[1,2] and Matthew R. W. Brake[1]

[1]: Component Science and Mechanics
Sandia National Laboratories
MS 1070; P.O. Box 5800
Albuquerque, NM 87185-1070

[2]: Engineering Physics
University of Wisconsin-Madison
1500 Engineering Dr.
Madison, WI 53706

Abstract

Quantifying uncertainty in model parameters is a challenging task for analysts. Soize has derived a method that is able to characterize both model and parameter uncertainty independently. This method is explained with the assumption that some experimental data is available, and is divided into seven steps. Monte Carlo analyses are performed to select the optimal dispersion variable to match the experimental data. Along with the nominal approach, an alternative distribution can be used along with corrections that can be utilized to expand the scope of this method. This method is one of a very few methods that can quantify uncertainty in the model form independently of the input parameters. Two examples are provided to illustrate the methodology, and example code is provided in the Appendix.

Acknowledgements

The authors would like to thank Marc Mignolet from Arizona State University for some clarifications and insight on using this approach. The authors would also like to thank Dan Kammer from the University of Wisconsin-Madison for all of his help, which was very useful in the understanding the special considerations such as for rigid body modes.

Contents

Figures

1 Introduction

One major difficulty that exists in reconciling model predictions of a system with experimental measurements is assessing and accounting for the uncertainties in the system. There are several sources of uncertainty in model prediction of physical phenomena, the primary ones being: aleatoric uncertainty (i.e. uncertainty in the model parameters), epistemic uncertainty (i.e. uncertainty in the model itself), and model solution error. These forms of uncertainty can have insidious consequences for modeling if not properly identified and accounted for. In particular, confusion between aleatoric and epistemic uncertainty can lead to a fundamentally incorrect model being inappropriately fit to data such that the model seems to be correct. As a consequence, model predictions may be nonphysical or nonsensical outside of the regime for which the model was calibrated.

In order to approach a problem with some amount of uncertainty, several steps are typically taken:

1. Postulate a mathematical model for the system,
2. Measure values for each of the model parameters,
3. Fit the data to the assumed form of statistical distribution,
4. Propagate the distributions through the model via a process such as a Monte Carlo (MC) analysis,
5. Assess the margins of the predicted distribution with the thresholds set by design requirements.

Alternative approaches may include assuming uncertainty in the distribution parameters and Bayesian methods, where the input distributions are determined from the output.

Multiple variations exist for this process, such as if the analyst has scant information regarding the distribution form for a given model parameter. Variations such as this, though, introduce more assumptions into the modeling process; but, probably the most dangerous assumption is that of the model form.

In general, model form error is a type of epistemic error. The insidious nature of model form error, though, is that it can often be mistaken for aleatoric uncertainty. In this report, a strategy for decoupling the effects of model form error in measurements from true aleatoric uncertainty is discussed. This methodology is originally developed in (Soize 2001; Soize and Ghanem 2004; Soize 2010; Guilleminot et al. 2011), and presented here in a format that is more accessible to most engineers.

2 Soize's Methodology

The methodology created and expanded by Soize over the past decade is a combination of maximum entropy formulations and random matrix theory. Random matrices are used to create the stochastic information from the computational model. Random matrix theory is a mature theory that developed from the field of information theory, but it was applied only recently by Soize to be used in a structural dynamics sense. The main purpose of this method is to treat parameter/aleatoric uncertainty (such as manufacturing tolerances), and model/epistemic uncertainty (such as model form error), independently.

There are two methods to calculate these errors, which are explained in (Soize 2010): with experimental data, and without experimental data. In what follows, the assumption is made that some experimental data is available. In particular, in the context of structural dynamics, experimental data about the natural frequencies and the frequency response functions are assumed to be known. Information such as the mode shapes can also be used in conjunction with the natural frequency by incorporating a metric such as the modal assurance criterion (MAC) value or cross-orthogonality values.

2.1 General Procedure

This methodology requires several MC analyses. A MC is conducted at every combination of possible dispersion values. The dispersion variable is similar to the coefficient of variation, and is a variable that determines the distribution parameter for the gamma distributions used in the following analysis (see Eq. 6). For larger systems, the number of MC samples need to be adjusted due to time and computational limitations. Two different number of MC samples are used for the examples: one to determine the intermittent probability measure and another to calculate the final distribution. Since the examples are computationally inexpensive, the number of samples for each are relatively large, 1,000 and 100,000 respectably.

In what follows, Soize's method is broken into seven steps:

1. Define the model and truth data.
2. Define the parameter and model variables.
3. Sweep the parameter dispersion variable.
4. Determine the maximum likelihood function.
5. Sweep the model dispersion variable.
6. Determine the model dispersion variables.
7. Calculate the final distribution.

Define Model and Truth Data

The first step in the method is to choose a model form. There are no restrictions on the model form: both linear and nonlinear models are suitable. In the example, a mass-spring system is chosen. The model assumes that the spring is massless, which is one source of model form error. The assumed model must be able to quasi-accurately predict the response. If the predicted values deviate too much from the true values (in this context, the distribution of experimentally measured data or the analytical solution), a new model must be used and subsequently analyzed. For the example problem, this states that the mass of the spring must be small compared to the mass of the block.

Next, the nominal values of the variables must be specified/estimated, and the criterions used to compare the stochastic model (i.e. the model populated with the distribution defined in what follows) to the reference model, which can be the experimental or truth model, must be defined. For the spring stiffness and mass of the block on the spring, the nominal values are specified as the assumed values for each. The mass of the spring does not need to be considered a nominal value since the assumed model does not take it into account.

For this example, two types of information are used to calibrate the dispersion variables (that are used in what follows): the natural frequency and the frequency response function of the system. This problem has one natural frequency and one frequency response at any frequency. For a more complex problem, multiple natural frequencies and the frequency response at different locations can be used. Another possible truth model for multi-degrees of freedom systems that can be used is the mode shapes. These can be compared by using the MAC values or by using an orthogonality value. Instructions for how to incorporate multiple natural frequencies are described in later sections.

Define Parameter and Model Variables

The purpose of this method is to separate the parameter uncertainty and the model uncertainty. There are several methods for realizing this separation. In (Soize 2010), variables that are material properties and physical dimensions are categorized as parameter variables, and parameters such as the mass, damping, and stiffness matrices are categorized as the model variables. For this example, the spring stiffness is treated as the parameter variable and the mass is treated as the model variable.

The random distributions of these variables do not have to be known a priori, in which case this method uses random matrices that have a gamma distribution kernel (Soize 2010). This is done to enforce that the mass and stiffness matrices are positive definite. The choice of the gamma distribution is also related to using the maximum entropy approach to determine the distribution. The gamma distributions used in this method are based on a single parameter, whereas the second parameter (the scale parameter) is assumed to be one based on maximum entropy.

Sweep Parameter Uncertainty Variable

Now that the model and parameters are defined, the next step is to sweep (i.e. uniformly sample) the dispersion variables over their entire range for each of the parameters in the model. During this sweep, it is assumed that the model variables are deterministic and do not vary. The dispersion variable is defined to have a specified range of values. If the variable is outside of this range, the value of the uncertainty is infinite, which violates one of the assumptions made by Soize. In what follows, the uncertainty dispersion variable is δ_g, where g is an index denoting which variable the term is associated with. The range of this value is $[0, \sqrt{\frac{n+1}{n+5}})$ where n is the degrees of freedom of the system. In the example, the range of the dispersion variable is $[0, \sqrt{1/3})$ as there is only one degree of freedom. The relationship between the dispersion variable and the distribution of the variable is presented in a later section.

For each combination of values of the dispersion variable, a MC analysis must be performed. Since there are multiple occurrences of this MC, the number of ensemble members much be chosen such that the analysis can be done efficiently. After each MC, one value needs to be computed: the probability that the truth data happens in this distribution. For most practical cases, a tolerance range needs to be defined. This can be done by using a histogram and determining the number of samples that are in the same bin that the truth data is within or by using the bins adjacent to the same bin that the truth data is within. This is a choice that the analyst must make and should depend on where the truth data falls within the bin. If the data is at the edge of the bin, then combining that bin and the bin it is adjacent to would be advantageous. The nature of this function is conducive to parallel computing for multiple parts if those are available. The dispersion variable sweep is independent and the MC analysis is, by definition, independent. Thus, parallel computing can be used for the sweeps.

The probability can be computed for each truth data, i.e. multiple natural frequencies each have an associated probability. This is approximated in the subsequent analyses by calculating the histograms for a given distribution and dividing by the total number of samples taken. For multiple natural frequencies, it is advantageous to take the natural log of each probability. The reason for this will be explained in the next section.

Determine the Maximum Likelihood Function

In order to select the combination of parameter dispersion variables, the maximum likelihood function is used. By definition of the maximum likelihood function, the parameter dispersion variables that are determined should maximize the joint probability density at the truth data. This is evident in Eq. 1: it is assumed that each probability is independent such that the joint probability density is equal to the product of the individual probability density functions

$$\delta = \text{argmax} \sum_i \ln \left[\rho_i \left(X_{Truth,i} \middle| \delta_j \right) \right]. \tag{1}$$

Since the natural log is a monotonically increasing function, if the argument maximum of a value is found, then the same argument maximum would be found by finding the maximum of the natural

log of that function. The natural log also allows a product to become a summation, which is simpler to do computationally, and is the reason for the format in Eq. 1, where δ_g is the dispersion variable, $\rho_i(X_{Truth,i}|\delta_j)$ is the probability that the i^{th} truth data $X_{Truth,i}$ falls within the distribution of the combination of dispersion variables δ_j.

In the mass-spring example, only one natural frequency is used, so there is no summation needed. Once this dispersion variable is chosen, it remains fixed at that dispersion variable value. The parameters still are random, but the dispersion variable for each parameter is fixed.

Sweep Model Uncertainty Variable

The sweep of the model dispersion variable is similar to the parameter dispersion variable sweep but varies in several ways. For most problems, the model variables usually include the mass, stiffness, and damping matrices (in this example, however, the problem is one-dimensional; a higher dimensional example is presented later). These matrices have specific properties (e.g. positive definiteness) that they must maintain. The main property that the variables must maintain is the positive definiteness of the matrix, which is what necessitates the theory of random matrices (Soize 2001). The derivation has two key results that are used in what follows. The first result is that each matrix must be able to be decomposed into a transpose product of an upper triangular matrix, i.e. the Cholesky decomposition. An example of this is given in Eq. 2 for the mass matrix $[M(X)]$, with a distribution of parameters about the mass matrix's nominal value

$$[M(X)] = [L_m(X)]^T [L_m(X)], \tag{2}$$

where $[L_m(X)]$ is the upper triangular matrix representing the mass, and ^{T} is the transpose of a matrix. This form is then expanded to include a random germ that preserves the positive definite attribute of the matrix. The random germ is a stochastic matrix that is strongly diagonal. This random germ (3) must also be decomposed into a transpose product of an upper triangular matrix, in the same manner as the mass matrix in Eq. 2. The random upper triangular matrix does not depend on the parameter uncertainty, but it is dependent on the model dispersion parameter. The off diagonal terms in this upper triangular matrix are standard normally distributed terms multiplied by a scaling factor. Each term of this matrix is, by definition, independent. The diagonal term of this upper triangular matrix is more complicated, which can be seen in Eq.4. The resulting structure of L is the same general form for both the parameter uncertainty and the model uncertainty, where $[G_n]$ is the random germ for the variable n, $[L_{G_n}]$ is the upper triangular matrix of the random germ, δ_n is the dispersion value for the variable n, Z_{ij} is a standard normal variable, V_{ii} is the gamma distribution which can both be seen in Eq. 6

$$[G_n] = [L_n]^T [L_n] \tag{3}$$

$$[L_n]_{ij|i\neq j} = \frac{\delta}{\sqrt{n+1}} Z_{jj} \tag{4}$$

$$[L_n]_{ii} = \frac{\delta}{\sqrt{n+1}}\sqrt{2V_{ii}} \tag{5}$$

$$V_{ii} \approx \Gamma\left(\frac{n+1}{2\delta^2} + \frac{1-i}{2}\right). \tag{6}$$

Once the germ is generated, it needs to be combined into each mass, stiffness, and/or damping matrix. The mass matrix example can be seen in Eq.7. For the first example, these matrices are 1x1 in size. The parameter uncertainty comes into the calculation on the L_m terms

$$[M] = [L_m]^T [G_m] [L_m] \tag{7}$$

Once each one of these random matrices is generated, a MC analysis is conducted for each combination of model dispersion variables. During the parameter dispersion variable sweep, only the distributions of the parameter variables are taken into account. For the model form dispersion variable sweep, both the parameter and model form variables are taken into account. The parameter variable distributions factor into the Cholesky decomposition of the matrices and the model form distributions are accounted for in the random germ.

For the first example, the frequency response function is used as the truth data. This is not compared directly, but is instead transformed into z_j, which is a second order centered random variable. This decomposition was a statistical reduced representations, similar to a Karhunen-Loeve expansion (Park and Cho 1996). Each z_j requires the information of the mean and covariance matrices of the frequency responses calculated from the stochastic analysis; $\mathbb{M}_{\mathbb{Y}}$ is the mean vector of the stochastic data, λ_j and $\mathbb{X}^j$ are the j^{th} eigenvalue and eigenvector of the covariance matrix respectively, and $\langle , \rangle$ is the inner or dot product. The random observed vector $\mathbb{Y}$ is the stochastic data, which is a $n_{obs} \times n_{freq}$ vector where n_{obs} is the number of MC samples used and n_{freq} is the number of frequency points that is done experimentally. This transformation is calculated for both the stochastic data and the reference data. For each z_j, the probability that the truth is correct in each density must be found.

$$z_j = \frac{1}{\sqrt{\lambda_j(\delta_g)}}\langle \mathbb{Y} - \mathbb{M}_{\mathbb{Y}}(\delta_g), \mathbb{X}^j(\delta_g)\rangle \tag{8}$$

Not every value of z_j needs to be computed; a given number of the values can be compared instead. A subset of all of the z_j can be used to decrease computation time since these vectors can be very large. The same transformation as in Eq. 8 can be used for the experimental FRF. This is the way the probability is computed by using the z_j terms as the summation for the comparison.

Determine Model Form Dispersion Variables

Once all of the probabilities have been calculated for each combination of model form dispersion variables, the likelihood function is evaluated. This is done in the same manner as for the parameter uncertainty. The determination of which dispersion variable provides the highest probability is found as in Eq. 1.

Calculate the Final Distribution

The final step is to perform a large MC analysis for the final set of dispersion variables. The dispersion variables for both the parameters and model are found in the previous sections. A final MC analysis is able to get the full uncertainty bands for the computations. A MC can also be done for the dispersion variables of the parameters that can be used as a comparison between the two distributions.

3 Example Applications

Two examples are provided to illustrate the application of Soize's method: a one dimensional example (previously introduced during the formal definition of the method) in which a mass on a spring with a mass is approximated as a mass on a massless spring, and a two dimensional example in which two masses connected by an elastic bar are approximated as two masses connected by a linear spring.

3.1 One Dimensional Example

In the one dimensional model of a mass on a spring that has mass, the block has mass M_m, and the spring has mass M_s and stiffness K. There are two main truth data points that are used for the comparison: The natural frequency ω_n of the system and the frequency response of the system. The natural frequency of this system is (Stokey 1976)

$$\omega_n = \sqrt{\frac{K}{M_m + M_s/3}}. \tag{9}$$

It is assumed that for the stochastic analysis, $M_s = 0$ and M_m and K are random variables in independent probability spaces. The other truth information used is the frequency response of the system. For this, it is assumed that the effective mass is equal to the mass of the block and a third of the mass of the spring taking after the same form as Eq. 9. This effective mass is taken as the mass in the linear equation of motion. The system is excited by a sinusoidal forcing function acting on the block. The analytical solution for the frequency response of the true system (where $M_S \neq 0$) is

$$Y(\omega) = \frac{1}{K - \omega^2 (M_m + M_s/3)}. \tag{10}$$

For the purpose of this method, this response is taken at several frequencies ω and compared to the stochastic data that is generated at those frequencies.

Results and Analysis of the One Dimensional System

When the mass of the spring is neglected, the nominal natural frequency $\omega_n = 10$ rad/s (with $M = 1$ and $K = 100$). For the actual system, $M_s = 0.1M_m$, which yields $\omega_n = 9.84$ rad/s. Two sets of data are used for the comparison and calibration of the dispersion variables: the true natural frequency and the frequency response. This example assumes that the system is linear, the block is rigid, there is no damping, and the effective mass can be used in place of separately considering the mass of the block and the mass of the spring. The frequency response is compared to the stochastic analysis at five frequencies evenly spaced and centered about the resonance. These points are chosen based

on the results of a convergence study that demonstrated that the solution had converged for this comparison of frequency responses with five frequency points.

The distribution of the predicted natural frequencies (calculated using the procedure of §2) is of particular interest in this example. The natural frequencies are calculated twice: once for considering only the parameter uncertainty of the spring, and once for considering both the parameter uncertainty and the model uncertainty. A noticeable difference is observed between the two distributions, as shown in Figs. 1-2. Thus, accounting for model uncertainty, in this case, has a better representation of the true nature of the system.

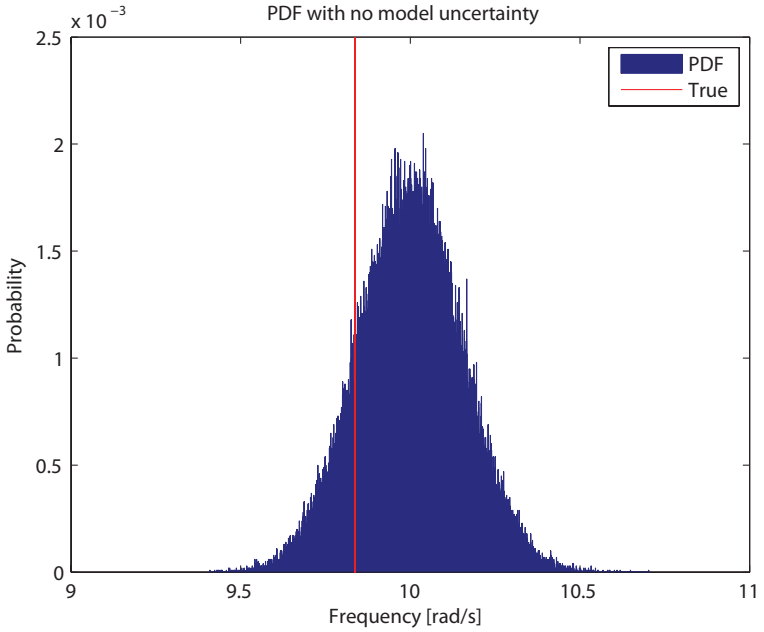

Figure 1. Histogram considering only parameter uncertainty.

3.2 Two Dimensional Example

A two dimensional example is presented in order to verify that the matrix generator is well posed and to demonstrate application to a multi-dimensional problem. Two masses are connected together by an elastic bar; in order to not have any rigid body modes, each one of these masses are connected to a rigid wall by a spring. Rigid body modes are possible with this methodology but require some special conditioning, as is further explained in the next section.

For simplicity, the elastic bar is treated as a massless spring in the model used in the stochastic analysis. The truth model for this example, though, incorporates both the stiffness and mass of the elastic bar, which gives a more accurate representation of the system. This example uses the natural frequencies along with the mode shapes (using the MAC) to calibrate the dispersion parameters.

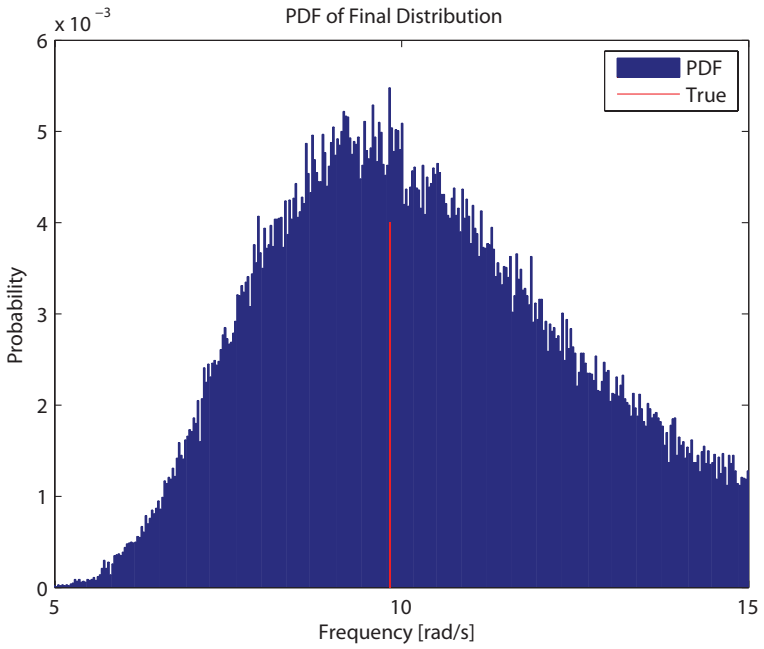

Figure 2. Histogram containing parameter and model uncertainty.

The elastic modulus of the elastic bar is used as the parameter variable, with the effective spring stiffness calculated based off of the elastic modulus. The model variables are the mass and stiffness matrices. The nominal model of the bar between the masses is treated as aluminum in density (for the truth model) and nominal elastic modulus.

Due to the mathematics behind the eigenvalue analysis, it is determined that the elastic modulus does not affect the first natural frequency as the first mode involves both masses moving together. This can lead to inaccurate representation of the data. If nothing is known about the material between the bars, then an analyst could conclude that the first natural frequency is deterministic while in reality it is not. Accounting for model error, however, can yield a better representation of the data. Plots of the resulting histograms are shown in Figs. 3-6. Figure 3 shows the first natural frequency while only considering the parameter uncertainty. The values calculated are deterministic and have 0% probability of reflecting the true data. In contrast, the first natural frequency, taking into account both the model uncertainty and parameter uncertainty, is shown in Fig. 4.

Figures 5-6 show the second natural frequency, which is dependent on the parameter uncertainty. While only Fig. 6 considers model uncertainty, these two histograms are qualitatively similar but have significantly different standard deviations (note the scales of the x-axes). The distribution used for the results of Fig. 6 accounts for both model (epistemic) and variable (aleatoric) uncertainty; this resulting distribution is thus more likely to be an accurate estimate of the real system than when epistemic uncertainty is not accounted for (Fig. 5).

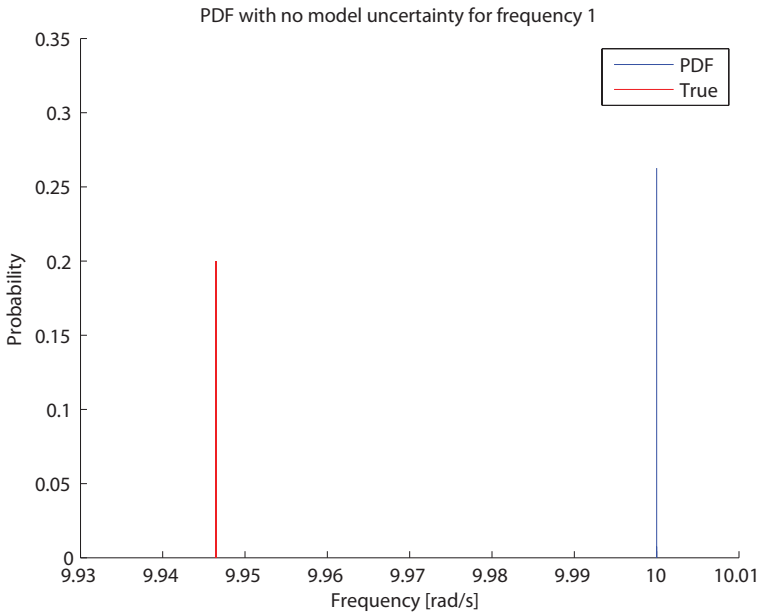

Figure 3. First natural frequency with only parameter uncertainty.

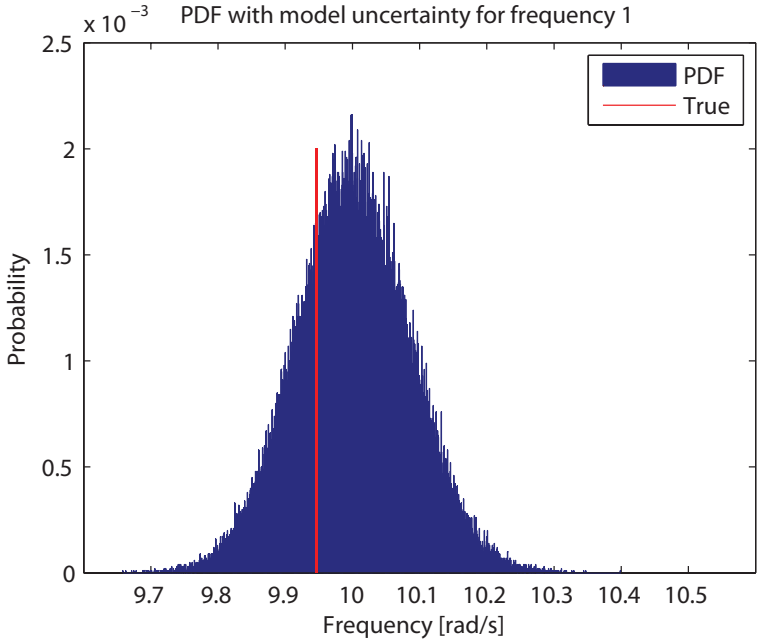

Figure 4. First natural frequency with parameter and model uncertainty.

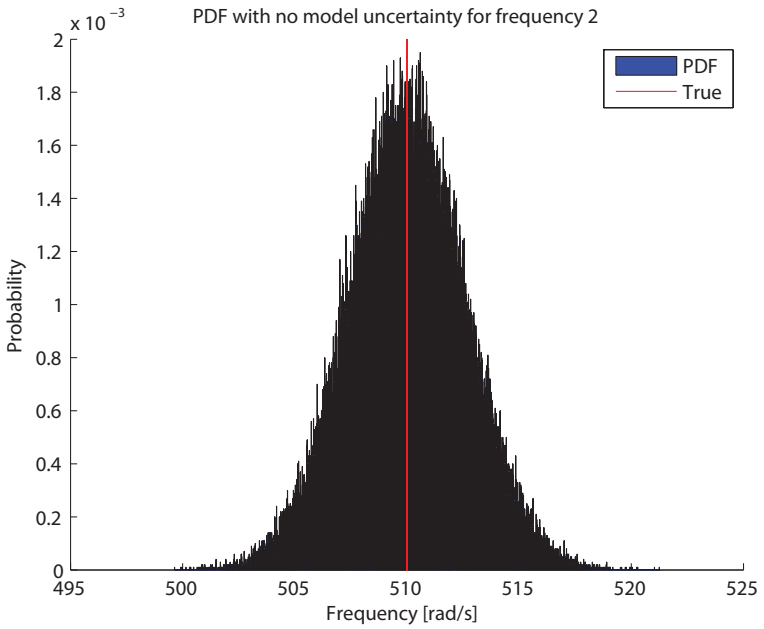

Figure 5. Second natural frequency with only parameter uncertainty.

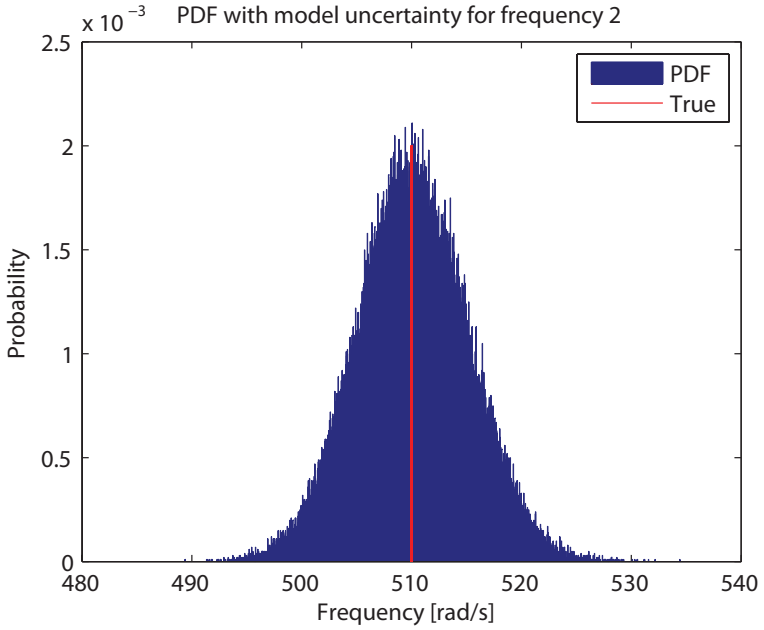

Figure 6. Second natural frequency with parameter and model uncertainty.

4 Alternative Matrix Generation

The random matrices that are used in this approach are based on a gamma distribution on the diagonal and normal distribution on the off diagonal. This approach can be implemented for any size application. It is common that as the size of the problem becomes large, determining every term of the upper triangular matrix becomes computationally expensive. In (Soize 2001), an alternative approach to calculating the random germ matrix is proposed. This alternative approach uses a vector of standard normal variables and generates the germ by taking the transpose product of that vector. The germ can be expressed as

$$[G_n] = \frac{[X][X]^T}{m} \tag{11}$$

$$m = \frac{n+1}{\delta_n^2} \tag{12}$$

where n is the degree of freedom, X is the vector of standard normal variables, $[G_n]$ is the random germ, and δ_n is the dispersion variable. For this alternative approach, Eq. 12 is used to ensure that the value of m is an integer, which results in the size of X being n by m. This is an efficient approach when small amounts of MC trials can be done, which makes the vector smaller and therefor faster to compute than the previous approach at the tradeoff of introducing a small amount of error due to the rounding of Eq. 12.

This approach is thus best used for large systems because it is computationally faster with fewer amounts of trials per MC analysis. This approach is explained briefly in (Soize 2001). While this alternative approach makes analyzing a larger system possible, the formulation of this method is based on using a reduced order model or modal coordinates. A high fidelity model is not under the assumptions made in the derivation of this method. This method may still be able to work for higher fidelity models, but high fidelity models are not within the verified range of operation (Soize 2001).

5 Special Corrections

There are two categories of applications that can benefit from corrections:

- A system that has rigid body modes.
- A high fidelity model that is too large to perform a MC analysis within a reasonable time frame.

The first correction is for a system that contains rigid body modes, which turn the stiffness matrix from positive definite to semi-positive definite since the rigid body modes have zero stiffness. In order to make the stiffness matrix positive definite again, the analyst must remove the rigid body modes from the system. This is done by ordering the eigenvalues then removing the eigenvalues with rounded zero value and removing the corresponding eigenvectors. Once this is done, then the stiffness matrix must be reconstructed. This can be done as a singular value decomposition of the stiffness matrix. This reconstruction is

$$L_k = \Lambda^{1/2}\Phi^T \tag{13}$$

$$K = L_k^T L_k \tag{14}$$

with diagonal matrix Λ containing the non-zero eigenvalues of the system, mass normalized elastic mode shapes Φ, and the reconstructed stiffness matrix K, which is the same size as the original stiffness matrix but does not contain the rigid body mode shapes (which results in it being rank deficient). The square root of the eigenvalues is a term by term square root.

The second consideration that needs to be made is for very large models. As the size of the model increase, a random variable must be inserted into every position of the mass and stiffness matrices, which can be time consuming especially in a MC analysis. In order to efficiently apply this methodology, a reduced order model should be used. This can be as simple as using the modal matrices, which contain only as many degrees of freedom as the number of modes that are included. It is also important to maintain more mode shapes than the ones of interest; this ensures a sufficient accuracy for the highest frequency mode of interest. A typical choice is to maintain modes up to twice the highest natural frequency of interest.

6 Conclusions

The method proposed in (Soize 2010) allows for considering aleatoric and epistemic uncertainty independently. This method is based on two theories being combined: random matrix theory and the maximum entropy approach. This is a powerful tool in the quantification of uncertainty. This combination allows for a generation of a random matrix that maintains positive definiteness, which is important for mass and stiffness matrices. The process for using this approach consists of seven steps that involve several Monte Carlo analyses for each combination of dispersion variables. These seven steps require some experimental or truth data in order to be used. Alternative approaches exist that do not require experimental data; however, these are not discussed in the present work. A most likely criterion is used to choose the best option of dispersion variables. Two examples are used to demonstrate the usefulness of this method, which verified the random matrix generator and the approach. Along with the nominal approach, several considerations for abnormal systems are given. This method is one of a very few ways of being able to quantify model error independently of any parameter error. Sample code to reproduce this method is included in the Appendices.

References

Guilleminot, J. et al. (2011). "A Probabilistic Model for Bounded Elasticity Tensor Random Fields with Application to Polycrystalline Microstructures". In: *Computer Methods in Applied Mechanics and Engineering* 200.

Park, H. M. and D. H. Cho (1996). "The Use of the Karhunen-Loeve Decomposition for the Modeling of Distributed Parameters". In: *Chemical Engineering Science* 51, pp. 81–98.

Soize, C. (2001). "Maximum Entropy Approach for Modeling Random Uncertainties in Transient Elastodynamics". In: *The Journal of the Acoustical Society of America* 109.

— (2010). "Generalized Probabilistic Approach of Uncertainties in Computational Dynamics Using Random Matrices and Polynomial Chaos Decompositions". In: *International Journal for Numerical Methods in Engineering* 81.

Soize, C. and R. Ghanem (2004). "Physical Systems with Random Uncertainties: Chaos Representation with Arbitrary Probability Measure". In: *SIAM Journal on Scientific Computing* 26.

Stokey, W. F. (1976). *Vibration of Systems Having Distributed Mass and Elasticity*. Shock and Vibration Handbook.

Appendix A: 1D Example MATLAB Code

```
%This program is used to evaluate the decomposition and analysis of both
%model and parameter uncertainty in a simple spring-mass system where the
%spring is not massless but is modeled as such and thus creates the
%uncertainty.
%
%Created by Matt Bonney 7/3/2014 for Sandia National Laboratories
%POC: Matthew Brake 1526
%
%This program is based on the paper written by Soize in 2009 named
%Generalized probabilistic approach of uncertainties in computational
%dynamics using random matrices and polynomial chaos decompositions
%
%True Model
%==================
%w^2=K/(Mm+Ms/3)
%
%Assumed Model
%==================
%w^2=K/Mm
%
%=====================================================
%% Initialize
clear all; clc; close all
%Setup parallel computing
Parallel;
%Set tolerances
Nmc=3000; m_max=100; k_max=100; incm=.01; inck=.01;
wn_tol=.1; Nmc_final=100000; n_obs=Nmc; Z_tol=.1;
%Nominal values
K=10000; %N/m
Mm=100; %kg
Ms=10; %kg
n=1;
w=[9 9.5 10 10.5 11]'; %Set w's for comparison
%True value of wn
wn_true=sqrt(K/(Mm+Ms/3));
%% Find Parameter of uncertainty for K
%1 uncertain parameters K
%Sweep to find uncertainty parameter
W=cell(k_max,1);
Pc=zeros(k_max,1);
tic
```

```
parfor int_k=1:k_max
    delta_k=inck*int_k;
    %Check limits
    if delta_k<sqrt((n+1)/(n+5))
        %MC for each delta
        Wint_k=zeros(Nmc,1);
        k=K*genrand(Nmc,delta_k);
        for mc=1:Nmc
            Wint_k(mc)=sqrt(k(mc)/Mm);
        end
        %Calculate Probability that the wn is within tol to true
        temp=histc(Wint_k,[wn_true-wn_tol wn_true+wn_tol]);
        Pc(int_k)=temp(1)/Nmc;
    else
        Pc(int_k)=0;
    end
end
toc
%Calculate maximum probability and set index to distribution
[~,ii]=max(Pc);
delta_K=inck*ii;
%% Generate only Parameter uncertainty dist
%Get pdf of final results
wn=zeros(Nmc_final,1);
k=K*genrand(Nmc_final,delta_K);
tic
parfor mc=1:Nmc_final
    wn(mc)=sqrt(k(mc)/Mm);
end
[N,X]=hist(wn,Nmc_final/50);
toc
figure
bar(X,N/Nmc_final)
hold on
plot(wn_true*ones(201),(0:1.25e-5:.0025),'r-')
xlabel('Frequency [rad/s]')
ylabel('Probability')
legend('PDF','True')
title('PDF with no model uncertainty')
xlim([9 11])
hold off

%% Generate FRF as Reference
n_freq=length(w);
Y=zeros(1,n_freq);
```

```
parfor i=1:n_freq
   Y(i)=1/(K-w(i)^2*(Mm+Ms/3));
end

%% Find Uncertainty Parameter of Mass
%Sweep model uncertainty of Mm
W=cell(m_max,1);
Pc=zeros(m_max,1);
Y_r=cell(m_max,1);
tic
parfor int_m=1:m_max
   delta_m=incm*int_m;
   if delta_m<sqrt((n+1)/(n+5))
      %Allocate
      Y_rint_m=zeros(n_obs,n_freq);
      s=n_obs*n_freq;
      %Monte Carlo
      k=K*genrand(s,delta_K);
      m=Mm*genrand(s,delta_m);
      for jj=1:n_freq
         for ii=1:n_obs
            %use 1D array instead of 2D for k m
            in=sub2ind([n_obs n_freq],ii,jj);
            Y_rint_m(ii,jj)=1/(k(in)-w(jj)^2*m(in));
         end
      end
      %Calculate Mean and COV
      myint_m=mean(Y_rint_m);
      Cyint_m=cov(Y_rint_m);
   else
      %Do nothing
   end
end
toc
%% Compare and Find Parameter for M
%Calculate Zj for random data
Z=cell(length(Cy),1);
Z_ref=zeros(length(Cy),n_freq);
tic
parfor int_m=1:length(Cy)
   [PHI,Lamda]=eig(Cyint_m);
   Zint_m=zeros(1,n_freq);
   for jj=1:n_freq
      for ii=1:n_obs
         Zint_m(ii,jj)=1/sqrt(Lamda(jj,jj))*(Y_rint_m(ii,:)-myint_m)*PHI(:,jj);
```

```
        end
        Z_ref(int_m,jj)=1/sqrt(Lamda(jj,jj))*(Y-myint_m)*PHI(:,jj);
    end
end
toc
%Do histogram on Z and get probability that Z is within Z_ref
Pm=zeros(length(Cy),1);
for int_m=1:length(Cy)
    for jj=1:n_freq
        temp=histc(Zint_m(:,jj),[Z_ref(int_m,jj)-Z_tol Z_ref(int_m,jj)+Z_tol]);
        Temp(jj)=log(temp(1)/n_obs);
    end
    Pm(int_m)=sum(Temp);
end
%Pick delta_m so that probability is max
[~,ii]=max(Pm);
delta_M=incm*ii;
%% Create Final pdf
%Get pdf of final results
wn=zeros(Nmc_final,1);
k=K*genrand(Nmc_final,delta_K);
m=Mm*genrand(Nmc_final,delta_M);
tic
parfor mc=1:Nmc_final
    wn(mc)=sqrt(k(mc)/m(mc));
end
[N,X]=hist(wn,Nmc_final/50);
toc
figure
bar(X,N/Nmc_final)
hold on
plot(wn_true*ones(201),(0:1.25e-5:.0025),'r-')
xlabel('Frequency [rad/s]')
ylabel('Probability')
legend('PDF','True')
title('PDF of Final Distribution')
xlim([7 13])
hold off
```

Appendix B: Random Matrix Generator

```
function G = genrand(num,delta,size)
%This function will output the random germ
%This is within the SG+ ensemble
%
%Created by Matt Bonney 7/18/2014 For Sandia National Labs
%
%POC: Matthew Brake 1526
%
% Inputs
%=====================================================
%
% num = Number of samples wanted [scalar]
% delta = The distribution parameter [scalar]
% size = The size of the Germ Matrix. Must be square [Size x Size]
%
% Outputs
%=====================================================
%
% G = an array with the uncertain parameter sampling [Size x Size x num]
%

%Pre-Allocate Space
G=zeros(size,size,num); V=zeros(size,num); L=zeros(size,size,num);
%check input
if delta>=0 && delta<sqrt((size+1)/(size+5))
else
   error('Delta is outside of acceptable range')
end
%Generate scalling parameter
sigma_n=delta/sqrt(size+1);
%Create Diagonal Terms
for j=1:size
   k=(size+1)/(2*delta^2)+(1-j)/2;
   theta=1;
   V(j,:)=gamrnd(k,theta,num,1);
   L(j,j,:)=sigma_n*sqrt(2*V(j,:));
end
%Create Off-Diagional Terms
U=randn(size,size,num); %Generates full matrix of random, not all are used
for j=1:size
   for i=j+1:size
      L(j,i,:)=sigma_n*U(j,i,:);
```

```
    end
end
%Generate Germ Matrix
for p=1:num
    G(:,:,p)=L(:,:,p)'*L(:,:,p);
end
```

DISTRIBUTION:

1 Matthew S. Bonney (electronic)
University of Wisconsin-Madison
Engineering Physics
Madison, WI 53706

1 Daniel C. Kammer (electronic)
University of Wisconsin-Madison
Engineering Physics
Madison, WI 53706

1 Marc P. Mignolet (electronic)
Arizona State University
Mechanical and Aerospace Engineering
Tempe, AZ 85287

1 MS 0346 Organization 1526, 1526 (electronic)
1 MS 0351 C. Dennis Croessmann, 1520 (electronic)
1 MS 0899 Technical Library, 9536 (electronic copy)

Sandia National Laboratories