SAND2016-5656C

Numerical Investigation of Probability Measures Utilized in a Maximum Entropy Approach

M. Bonney [1], **D. Kammer** [1], **M. Brake** [2]
[1] The University of Wisconsin-Madison,
Department of Engineering Physics,
Madison, WI

[2] Sandia National Laboratories,
Albuquerque, NM

e-mail: **msbonney@wisc.edu**

Abstract

The quantification of uncertainty is an important aspect in modern design. One relatively new technique to quantify this uncertainty is the maximum entropy approach. This approach characterizes the physical system in terms of random matrices. Each of the random matrices is characterized by a single dispersion parameter and is derived to maintain positive-definiteness of the system matrix. The random matrices are generated to have a mean value of the nominal system and the variance is controlled by the dispersion parameter. This dispersion parameter characterizes the model-form uncertainty, which is a large area of current research. The maximum entropy approach has the ability to characterize both the model-form and parametric uncertainty independently. The determination of the dispersion parameter is done by creating the maximum likelihood estimate. This estimate requires a numerical probability measure for any real world system since an analytical distribution cannot be determined. This paper investigates the effect of multiple sampling techniques, such as Monte-Carlo and Latin Hypercube, and multiple probability measurements, such as histograms and density kernel estimation. The investigation is performed on a planar frame truss with uncertainty introduced due to a reduced model using different reduction techniques such as free-interface and fixed-interface reductions. The induced uncertainty is compared to the truth data from the non-reduced system using the fixed-interface and free-interface modes. This technique is a systematic method that produces repeatable results that can be used in model simulations to produce a stochastic output range that is verified to the desired truth data.

1 Introduction

Modern engineering design primarily uses the advancement in computational ability to aid and optimize a design. With the use of these techniques, designs are able to be developed that meet the design requirements and can lead to reduced overall cost. One of the main techniques used is the discretization of the system using finite elements. The implementation of this and other techniques introduces several assumptions that lead to uncertainty. These uncertainties are very important and needs to be quantified. One of the important sources of uncertainty is the selection of the model, the model form error. High-fidelity finite elements are used in order to give accurate results, but using this computer model requires a large amount of computational power. In order to perform multiple possible designs, a reduced order model (ROM) is typically used. By using ROMs, additional model form error is introduced into the system. Quantifying this error is important to incorporate a confidence in the model output.

Determination of the model form error is difficult and can be characterized by using different methods. One classical approach to characterizing the model form error is with the use of fuzzy sets. This idea was originally derived in [1]. Fuzzy sets are thought of as a set that are defined by non-distinct bounds. A non-engineering example of a fuzzy set is describing a person's height as tall or short. The basic concept of this technique is simple but the implementation in computational codes is very complex. A review of this approach and a comparison between this and other techniques is presented in [2] and in [3].

Since the implementation of fuzzy sets is complicated, another method is required to simplify the calculation of the model form error. One technique is the maximum entropy, non-parametric approach. This method is originally derived in [4] by incorporating random matrices into the system equation of motion. Originally, this method is used to apply some randomness into the system that is not associated with a specified cause. The theory is generalized and expanded in [5]. Later on, in [6], this technique is expanded to be used to character the model form error and the parametric error independently. A simplified methodology of this technique is presented in [7] with application on a wind turbine in [8]. Expansion of this method to increase repeatability for the application of ROMs is presented in [9].

This paper explains the maximum entropy approach for the application of ROMs, primarily fixed-interface and free-interface ROMs. One of the main focuses of this paper is to quantify the sensitivity of this method to the engineering choices made. The main choices that are made include: what type of distribution sampling technique to use, how many samples are used in the sampling technique, and how the probability measure is calculated. Section 2 describes the maximum entropy approach that is used throughout this research. In Section 3, the different engineering choices are discussed and the underlying theory of each choice is explained. This technique is applied to a specific computer model that is described in Section 4. Two different ROMs are used and discussed: fixed-interface ROM is presented in Section 5 and the results from the free-interface ROM is discussed in Section 6. Some concluding remarks are presented in Section 7.

2 Maximum Entropy Approach

In the system of interest, the equations of motion are based on linear mass and stiffness. The mass matrix is positive definite while the stiffness matrix is either positive definite or semi-positive definite depending if the system is constrained or contains rigid body motion. Due to the definiteness of the matrices, these can be decomposed using Cholesky decomposition similar to

$$[M] = [L_M]^T [L_M], \tag{1}$$

with $[L_M]$ is an upper triangular matrix that is the Cholesky decomposition of the mass matrix $[M]$. For a non-singular matrix, $[L_M]$ is a square, upper triangular matrix, while for a singular matrix, $[L_M]$ is a full, rectangular matrix obtained by using a Cholesky-like decomposition using the covariance of the matrix. The Cholesky decomposition is a function of the input parameters to the FE model. To incorporate model form error, a random germ is added that is also generated to be positive definite so that the resulting matrix is still positive definite or semi-positive definite. The germ is incorporated using the relation

$$[M] = [L_M]^T [G(\delta_M)][L_M], \tag{2}$$

where $[G]$ is a random germ with an expected value of I and δ_M is the dispersion parameter for the mass matrix. The variance of the random germ is a function of δ scaled by the size of the matrix. More detail of how to formulate this germ is given in [4, 7, 8, 10]. The value of δ is bounded by the size of the system. This range is given as $\delta \in [0, \sqrt{\frac{n+1}{n+5}})$, where n is the size of the system. As the system becomes very large, this range approaches $[0, 1)$ that leads to the presentation of this value as a percentage, although, performing this analysis of such a large system can become computationally infeasible.

The dispersion parameter is chosen to best fit the truth data, which corresponds to having the highest probability in the probability density function. Truth data can be given as many things such as the frequency response function, mode shapes, or natural frequencies. This research uses the first 11 elastic natural frequencies as the truth data. The probability is best given as the likelihood function, which is the objective function for determining the dispersion variable. The main calculation is the likelihood value for each dispersion variable. This likelihood value is defined as the joint probability that the truth data exists within the model with a given value of the dispersion parameter, and is shown as

$$L(\delta) = \Pr(TruthData|\delta), \tag{3}$$

where $L(\delta)$ is the likelihood value and $Pr()$ is the probability value that the truth data exists in the model. For the simplicity of calculations, each natural frequency is treated as independent. Due to the independence, the joint probability can be rewritten as

$$\Pr(TruthData|\delta) = \Pi_i \Pr(\omega_i|\delta). \tag{4}$$

This calculation yields a value on the range of $[0, 1]$. The optimal dispersion value is defined as the argument maximum of the likelihood function. Since this calculation has a very small range, the differences between each dispersion value can be small that can be affected by the computer resolution error. In order to reduce this error, the log-likelihood function is used. For these calculations, the negative log-likelihood function is used in order to simply expand this method for the use of an optimization technique. The negative log-likelihood is calculated by computing $\ell = \sum_i -log(\Pr(\omega_i|\delta))$. The reason this works is because $-log()$ is a monotonically decreasing function, which means that the argument minimum of the negative log-likelihood is the same as the argument maximum of the likelihood function.

The calculation of the likelihood is based on the probability measure. There are two ways to compute this probability: analytically or numerically. In order to calculate the probability analytically, the distribution of the natural frequency must be known, which is very hard to generate. These simulations calculate the probability numerically based on a distribution sampling simulation. In order to calculate the probability, a histogram can be used. For these calculations, the tolerance band is based on 0.1% of the nominal natural frequency. The probability is calculated as the percentage of the number of times a natural frequency lands within this tolerance band. This determination of the band is very critical in the execution of this analysis. A band that is too small will result in very large variability, while a band that is too large will not be able to differentiate between different dispersion values. The argument minimum of this negative log-likelihood function is the statistical estimate of the dispersion variable.

Using a histogram is one method to compute the probability numerically; kernel estimation is another useful method to compute the probability. This method of determining the probability is a classical statistical approach that is commonly used to generate the probability density function. The calculation of the density function is based off of a finite number of samples and increases the accuracy of the probability compared to using a histogram. An analysis of the underlying calculations can be found in [11]. This describes the math in terms of a generic kernel: a discussion of various kernels is presented in [12]. The comparison between using a kernel estimator and histograms is presented in [13] and the selection of a smoothing parameter that is used in several kernel functions is discussed in [14]. In this research, a Gaussian kernel that is programmed into Matlab is used. Future work includes the results from the selection of different kernel functions.

Determining the probability is very dependent on the sampling methods, and the probability will vary based on what samples are used. In order to reduce this randomness, the empirical likelihood function is applied. The empirical negative log-likelihood is a technique that is able to reduce the randomness due to the sampling technique. This calculation can be determine by using the expression

$$EL(\delta) = \sup_i \ell_i(\delta), \tag{5}$$

in which, i is the independent evaluation of the negative log-likelihood, ℓ_i. This empirical formulation accounts for the randomness in the sampling technique. The optimal dispersion parameter is then chosen based on

$$\delta^{optimal} = \arg min\ EL(\delta). \tag{6}$$

The optimal dispersion parameter is used in a forward propagation of the matrices to determine the distribution of the desired output, such as reliability measures or dynamic characteristics.

3 Theory of Specific Engineering Decisions

The use of the maximum entropy approach requires several assumptions to be made. Each additional assumption or choice can have an effect on the output of the method. One of the main decisions is how the distribution is sampled. This discussion is very dependent on the available computational ability. Two methods of sampling are described in this section. Along with the different sampling methods, the theory of the model reduction techniques is also described in this section.

3.1 Sampling Methods

One of the major decisions in the maximum entropy approach is how the distribution should be sampled in order to determine the probability. Since the analytical solution is not known, the probability needs to be computed numerically. This requires the input distribution to be sampled and then the deterministic process is evaluated at each sampled point. Two types of sampling methods are presented here: Monte-Carlo (MC) and Latin Hypercube (LHC). The MC method is a simple technique but requires a large amount of samples. One method to reduce the number of samples is the LHC method that divides the distribution into sections and samples each section. Each method is described in detail in this section.

3.1.1 Monte-Carlo Sampling

The first sampling technique is the most simple and straightforward of the techniques, MC sampling. This idea was originally proposed in [15] and named after the famous casino in Monte-Carlo, France. This technique is based on the pseudo-random number generator that is programmed in computer software. These pseudo-random numbers require a distribution in order to generate the values that will be used in the deterministic simulations.

The basic concept of this method is to perform a large number of simulations with varied input parameters then to describe the output of the simulation in terms of a probability distribution. In order to use this method, several engineering decisions must be made. The first decision is the input distribution of the parameters. This must be chosen based on the physical characteristics and previous knowledge of these parameters. One example of this could be a machining

tolerance that is specified, setting the bounds of the distribution, or using previous test data for material properties.

Another decision that must be made is the number of simulations that can be performed. This is a very important decision. The method requires a large number of simulations in order to produce accurate results, but the computational burden must also be considered. This number of simulations can be increased by using a less accurate model such as ROMs or surrogate models. One study, presented in [16], shows that the convergence of a MC simulation is exponential but very slow so a large number of samples are required for the required accuracy.

In this type of analysis, each simulation is treated as equally probable and each simulation is treated as independent. This leads to a very simple continuation into parallel processing. The output distribution is generated based on the fact that the more probable input parameter will have more simulations than the less probable parameter values. There are several variations on this method. This particular method involves random numbers generated independently from the computer list of quasi-random numbers.

The major strength of this method is its simplicity that requires almost zero code modification in order to perform. This allows an engineer without much Uncertainty Quantification (UQ) experience to perform this type of analysis. One of the major weaknesses is the computational burden of the method. This burden can be reduced via parallel computing or surrogate modeling.

3.1.2 Latin Hypercube

The LHC sampling technique is a stratification of the traditional MC technique. This was originally introduced in [17] to improve on the number of required samples to converge to a solution. This is done by dividing the probability density function into equally probable sections. Within each of the sections, one design point is selected. Then the deterministic simulations are performed at each of the selected values, and are treated as equally probable. Extensive use of LHC for various distributions and data sets is studied in [18].

The selection of the value within the equally probable section can be determined via two different ways: deterministically or randomly. The deterministic approach uses the mid-point within the section as the selected value. A random approach is to assign a uniform distribution to the section and choose a value based on this distribution. As the number of samples increase, the difference between these two selection methods decreases. This is due to the range of each section decreasing, and the difference between the mid-point and a random point becomes small.

This technique is very effective on bounded distributions such as beta or uniform distributions. There are several numerical issues that can arise for semi-bounded or non-bounded distributions such as normal or exponential distributions. The main numerical instability involves the tails of the distribution. As the number of samples increases, the bounds on the tail section increases and produce a very large number that can create bit resolution errors in the computer code. Another issue is the memory requirements to use this sampling technique. All the values have to be determined before the simulations are performed. These values must be stored in global

memory. This issue is not a large problem for projects that only have low dimensionality for the stochastic variables but can become important for higher dimensionality.

This technique is very useful for simulations that require a very large amount of computational power. The LHC is able to give an accurate representation with far less number of samples. This difference is about an order of magnitude less that traditional MC sampling as studied in [19].

3.2 Reduction Techniques

One commonly used method to reduce the computational burden is to design and analyze components individually and then assemble the full system using a substructuring technique. Typically each substructure is further reduced in order to lessen the computational burden. The reduction is usually done by transforming the non-interface degrees of freedoms (DOFs) into modal DOFs and then truncated, typically based on frequency. Mode shapes can also be used to determine the truncation [20]. There are two main types of reduction used in this research, free-interface and fixed-interface reductions. This refers to the modal DOF that are used and the reduced. The fixed-interface reduction is called the Craig-Bampton (CB) reduction. This was originally proposed in [21] and reviewed along with other methods in [22]. The theory of this reduction is further explained in Section 3.2.1. The other commonly used reduction uses free-interface modes. This reduction is used in the Craig-Chang (CC) synthesis technique originally proposed in [23]. The theory of this reduction and synthesis is given in detail in Section 3.2.2.

3.2.1 Fixed-Interface/Craig-Bampton Reduction

The CB representation is a mixed coordinate system that contains fixed-interface modal DOF and physical interface DOF together [21]. This formulation is also called the primal formulation in [24]. There are two different approaches to forming the ROMs. In [24], Rixen uses a Lagrange multiplier technique and in [25], Craig uses a more physical approach. This section uses the more physical approach taken by Craig. The displacement vector takes the form of $u = [\eta \;\; u_j]^T$ where η is the fixed-interface modal DOF and u_j are the physical displacement interface DOF. This representation allows for a truncation of the fixed-interface modal DOF according to frequency or desired mode shapes [20] that can greatly decrease the size of the subsystem if there are only a few interface DOF. The transformation of the displacement from the full subsystem to the CB representation is given by

$$\begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} \hat{\Phi} & \Psi_c \\ 0 & I \end{bmatrix} \begin{bmatrix} \eta \\ u_j \end{bmatrix} = T_{cb}\, u_{cb} \tag{7}$$

where u_i are the interior physical DOF, and $\hat{\Phi}$ are the kept fixed-interface modes that are mass normalized. Ψ_c is the constraint modes, which is the displacement of the interior DOF due to a unit deflection at one of the interface DOF and maintaining zero deflection at the other interface DOF, 0 is a matrix of zeros, and I is an identity matrix. The constraint modes are generated geometrically by

$$\Psi_c = -K_{ii}^{-1} K_{ij} \tag{8}$$

with the subscript i corresponds to the interior DOF and the subscript j corresponds to the interface DOF. This transformation matrix, T_{cb}, is a non-square transformation that can be applied to the equations of motion by the mass and stiffness matrix, given by

$$\hat{M} = T_{cb}^T \, M \, T_{cb}. \tag{9}$$

3.2.2 Free-Interface/Craig-Chang Reduction

The CB method uses fixed-interface modes as the DOF that are reduced. In the CC formulation, free-interface modes are the modal DOF that are used. Using free-interface modal DOF can also be used for method such as the ones used by MacNeal [26] and Rubin [27]. These use physical displacement interface DOF while the CC method [23] uses the forces at the interface. The same DOF are used in the dual formulation of CB substructures introduced by Rixen in [28]. The dual CB only enforces force equilibrium at the interface while CC enforces this and displacement compatibility. This difference does not affect the substructure itself, but does effect the synthesized system. The main affect is that the system matrices are no longer semi-positive definite, which prevents the use of the maximum entropy approach.

This representation is produced with the free-interface modes along with residual attachment modes as presented in [23]. The transformation from physical DOF to a CC DOF set is given by

$$\begin{bmatrix} u_i \\ u_j \end{bmatrix} = \begin{bmatrix} \Phi_k & \Psi_d \end{bmatrix} \begin{bmatrix} P_i \\ P_j \end{bmatrix} = T_{cc} \, u_{cc} \tag{10}$$

where Φ_k are the kept free-interface modes that are mass normalized, Ψ_d are the residual attachment modes, P_i are the free-interface modal coordinates, and P_j are the interface generalized coordinates that correspond to the force at each DOF on the interface. The free-interface modes are the normal modes when the interface is unconstrained and they contain the rigid body motion of the system. The residual attachment modes are the response to a unit force at a single interface DOF and are linearly independent of the free-interface modes. To ensure linear independence, the residual attachment modes are derived based on the residual flexibility matrix, which can be expressed as

$$G^d = K^{-1} - \Phi_k \Lambda_k^{-1} \Phi_k^T \tag{11}$$

where G^d is the residual flexibility matrix, K is the physical stiffness matrix, and Λ_k^{-1} is a diagonal matrix containing the inverse of the eigenvalues of the free-interface modes on the diagonal. One complication of this method is the inverse of the stiffness matrix. For a system that contains rigid body motion, this stiffness matrix is singular, which implies that the inverse of the matrix does not exist. One method to eliminate this effect is with the use of oblique

projectors [29]. These project the stiffness matrix onto a non-singular domain. This implies that the inverse exists, which is mathematically expressed in [30] and the implementation for structural dynamics is expressed in [31]. The residual attachment modes can be written in terms of the residual flexibility matrix as

$$\Psi_d = \begin{bmatrix} G^d_{ij} \\ \\ G^d_{jj} \end{bmatrix}. \tag{12}$$

With this transformation matrix, the equations of motion can be transformed into the new coordinates by

$$\hat{M} = T_{cc}^T \, M \, T_{cc}. \tag{13}$$

4 Example System

The system used to illustrate these techniques is a planar fame with an appendage. This system is shown in Figure 1, and contains two subsystems that are combined to form the full system separated at the red nodes at $x = 0$. The subsystem to the right is called system "B" and the subsystem to the left with the appendage is called system "D". These subsystems are reduced using the desired ROM technique. The red connection nodes are shared between the two subsystems and are defined as the interface. The nodes shown in Figure 1 are connected using Euler beam elements. Each of the two subsystems has 129 DOF. This system is relatively small but large enough to show the complexity of the techniques and how they can be applied to larger systems.

The simulations for the CB reductions are based on matching the system frequencies for the first 11 elastic free-free modes. For subsystem B, the 11^{th} natural frequency is 4656.4 Hz and subsystem D has an eleventh natural frequency at 5511.5 Hz. This selection of the first 11 elastic free-free modes was chosen based on the idea that the higher frequency modes are more subject to model form error.

For the CC simulations, the use of free-free modes for the substructure truth data is redundant since that information is used in the reduction. So in order to test the CC reduction, the fixed interface modes are used for the truth data. These simulations still use the first 11 elastic modes.

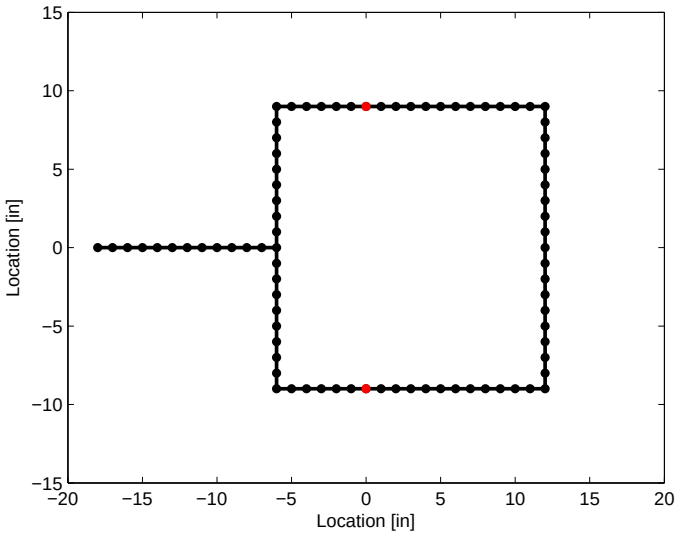

Figure 1: Planar Frame Example System

5 Fixed-Interface Results

The first results for this system are when the fixed-interface/CB reductions are used for each substructure. One consideration in using this maximum entropy approach is the selection of the truth data. For both of the model reduction techniques, the truth data is determined from the non-reduced system. This is the data that is used to determine the probability. Another important truth data is the true dispersion parameter. Since the true probability density function is not known, an approximation needs to be performed. For this approximation, a very large scale MC sampling is used. This MC analysis uses $100,000$ samples and is performed 30 times for the empirical likelihood. In addition to the large number of samples, the Gaussian kernel estimator is also used to generate the truth data. This decision is based on the available time and computational resources.

There are four different analysis compared in this section: two different sampling and two different methods of determining probability. For the MC sampling method, $10,000$ samples are used, and for the LHC sampling method, 200 samples are used. All of analyses are performed with 30 evaluations for the empirical likelihood. These numbers of samples are chosen based on the computational time to perform the analysis. The total time for each analysis takes approximately the same amount of time on the author's computer.

The main result to report is the optimal dispersion parameter. This was determined based on the mass matrix, while treating the stiffness as deterministic. One of the main reasons for this judgment is the topology of the stiffness matrix. When the stiffness matrix is converted into the CB reduction, the reduced matrix is block-diagonal while the mass matrix is a full matrix. In order to maintain this block diagonal form, the stiffness is chosen to be deterministic and the

mass to be stochastic.

The first system to be analyzed is system B. The optimal dispersion parameter for the mass matrix can be seen in Figure 2, where the blue lines represent using the Gaussian kernel estimator and the red lines represent using a histogram to determine the probability measure. The solid line represents using a MC analysis and the dashed line represents using a reduced sampling method, LHC. Truth data in Figure 2 is shown in the solid black line.

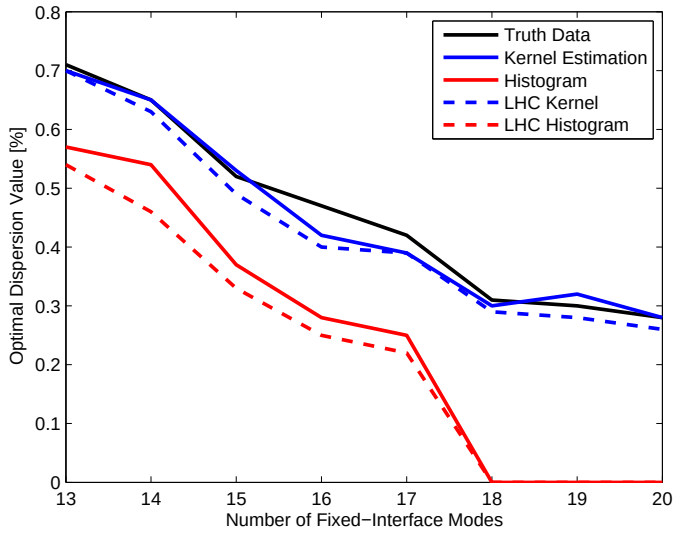

Figure 2: System B using CB reduction

There are a couple trends in Figure 2 of interest. The first trend is the difference between using a histogram and the kernel estimator. Using the kernel estimator gives a larger dispersion parameter that corresponds to a more conservative analysis. The second trend is that using the LHC compared to the MC produced smaller optimal dispersion parameters. This leads to the trend that the MC analysis is more conservative than the LHC sampling techniques. The final trend is the accuracy of each analysis compared to the truth data. As can be seen in Figure 2, the kernel estimated data produces more accurate results. One possible reason for this trend is that the truth data is generated using the kernel estimator. The truth data using a histogram is not calculated due to computational and time restrictions. Taking the difference between the truth data and the kernel estimator data does not yield a trend. At some values, the truth data is more conservative while some other points have the kernel estimator as a more conservative value. The same analysis is also performed on system D. This analysis can be seen in Figure 3.

There are several differences between the two systems. For system D, all four analyses have very similar values. Figure 3 shows that visually, the methods are difficult to differentiate. The only visual trend is for the LHC sampling always produced an optimal dispersion parameter smaller than the truth data. This inability to differentiate between the analysis leads to the necessity of a quantitative difference between the analyses. The quantitative result used is the root-mean-

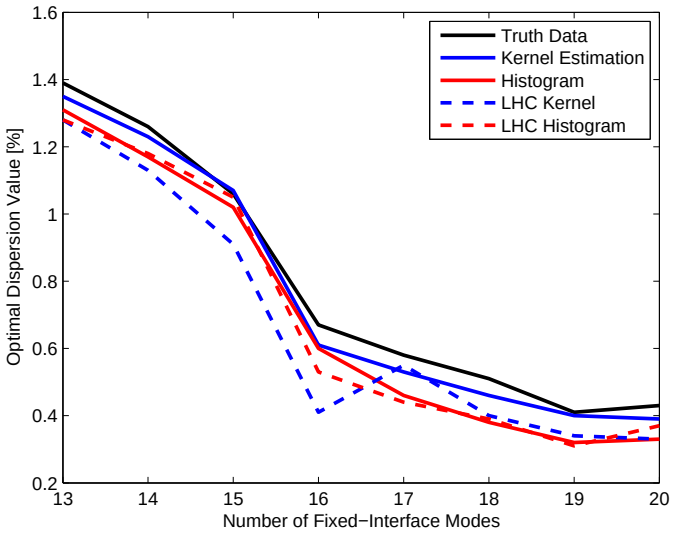

Figure 3: System D using CB reduction

square (RMS) of the percent difference between the analysis and the truth data and is shown in Table 1.

Table 1: RMS error of optimal dispersion parameter for fixed-interface reduction

Analysis	System B Error [%]	System D Error [%]
Kernel	5.30	6.69
Histogram	65.94	16.99
LHC Kernel	7.54	20.00
LHC Histogram	68.21	17.56

The first note about Table 1 is that the MC analysis gives a more accurate result than the LHC for both the kernel estimator and the histogram. Table 1 also gives the same quantitative results for system B. This system has large differences between the analyses both visually and quantitatively.

Both systems have similar trends. In general, the kernel estimator gives more accurate results compared to the histogram with 0.1% tolerance bands. This value for the tolerance band is chosen as a singular value to work for both systems over the entire range of interest. For systems with a smaller optimal dispersion parameter, a smaller tolerance band should be able to get more accurate results, but is dependent on the engineering judgment of this determination. Using a kernel estimator eliminates the engineering choice of this tolerance band and is more accurate.

6 Free-Interface Results

The same analysis that is performed on the fixed-interface/CC reduction is also performed on the free-interface reduction. This analysis results in some similar trends and some new trends as well. The first result is for system B and is presented in Figure 4.

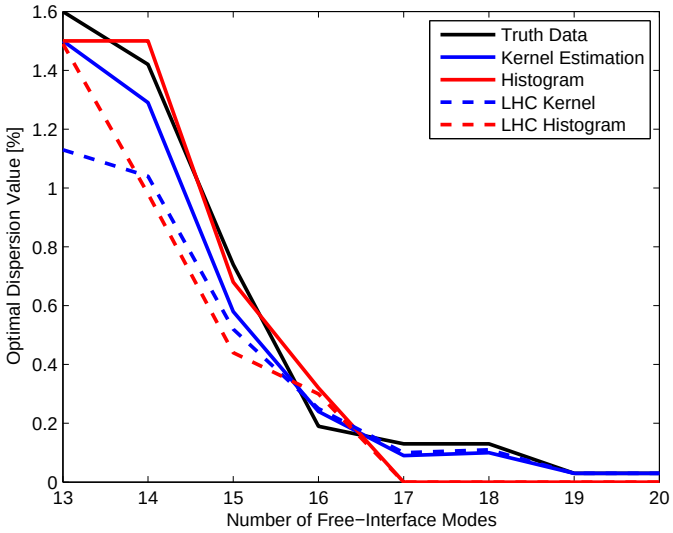

Figure 4: System B using CC reduction

One new trend by using the free-interface reduction is that the conservative trend is no longer followed. At some data points, the LHC sampling is more conservative while at other points, the MC sampling is more conservative. This occurs when the optimal dispersion parameter is less than 0.3%. One reason for this change in trend is believed to be due to the resolution of the sweep. The dispersion parameter is swept up from zero with a resolution of 0.01% to a value larger than the optimal value. As the optimal dispersion parameter becomes a small number, the differences between the probability measures at each value of the dispersion parameter becomes small and more subject to the natural randomness of the sampling method. This trend is more apparent in system D that is shown in Figure 5.

One important aspect to notice in Figure 5 is the value of the optimal dispersion parameter. The largest optimal dispersion parameter is 0.24% and thus is more subject to the randomness of the distribution sampling. Future work will look into a finer resolution in a sweep and an investigation into using an optimization technique. Another interesting trend is the histogram produced a zero optimal dispersion parameter. This is due to the tolerance band being too large for the optimal dispersion values that are small.

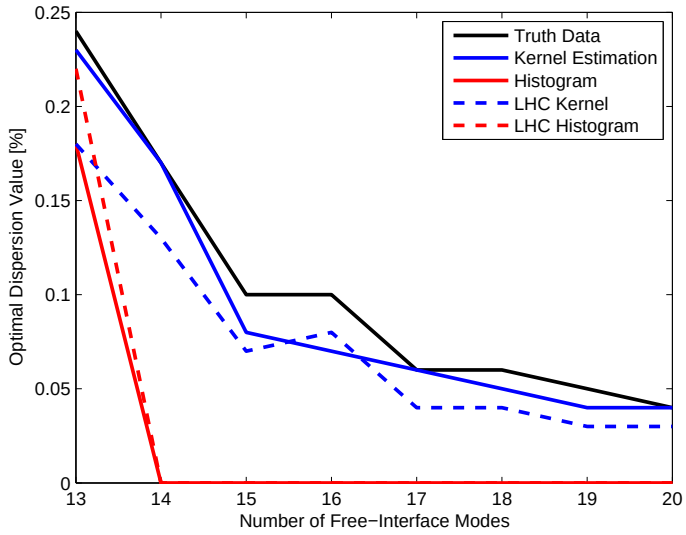

Figure 5: System D using CC reduction

7 Conclusions

The study of model form error is an important consideration for uncertainty quantification of a system. This is particularly important for systems that are unable to have a large amount of experimental data available. The maximum entropy approach is one method for the determination of model form error. This is done with the combination of using the maximum entropy to determine a distribution and the use of random matrices. One major advantages of this method is that the variability can be expressed by a single bounded parameter. This method requires two major engineering choices: how to sample the distribution along with how many samples to use, and how to determine the joint probability. The research presented in this paper looks at two different techniques for each of those choices. Monte-Carlo and Latin hypercube sampling techniques are studied and the probability is determined using a Gaussian kernel estimator and a histogram. These methods are studied on the reduction of two systems using both a free-interface and a fixed-interface reduction techniques commonly used in substructuring. This error represents the model form error of using a reduced order model.

The fixed-interface reduction technique shows some interesting trends. Both systems show that the use of the reduced sampling technique yielded optimal dispersion parameters less than the truth values. For one system, using a histogram gives a very large difference compared to the truth data while the other system shows accuracy for all the methods to be very similar. This is believed to be due to tolerance band and how it related to the true optimal dispersion parameter. A system that has a small optimal dispersion parameter requires a smaller tolerance band in order to account for all the sampled values to lie within the tolerance band.

The free-interface reduction technique yielded similar results to the fixed-interface results with

two main differences. One of the main differences is the general size of the optimal dispersion values. The CB method has larger values compared to the CC method. One of the ramifications of this difference is that the resolution in the parameter sweep needs to be more refined, or an optimization technique should be implemented. The second main difference is that the conservative trends are not followed. This is believed to be related to the resolution of the parameter sweep as well.

The use of the maximum entropy approach is one way to characterize the model form error of a system. One major advantage is that this method describes the uncertainty in the system as a bounded scalar value. While this method only produces a single value, several engineering decisions are required and can have an affect of the output. A few of the major decisions are discussed in this paper and the effect of each decision is studied on the test systems. The method proposed in this paper is a straight foreword methodology that can be applied to most of the engineering decisions made in an analysis, and demonstrates the use on three different decisions.

Acknowledgment

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.

References

[1] L.A. Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.

[2] T.K. Hasselman, J.D. Chrostowski, and T.J. Ross. Propagation of modeling uncertainty through structural dynamic models. In *35th Structures, Structural Dynamics, and Materials Conference, Hilton Head, SC*, pages 72–83, 1994.

[3] Y. Li, J. Chen, and L. Feng. Dealing with uncertainty: a survey of theories and practices. *Knowledge and Data Engineering, IEEE Transactions on*, 25(11):2463–2482, 2013.

[4] C. Soize. Maximum entropy approach for modeling random uncertainties in transient elastodynamics. *Journal of Acoustical Society of America*, 5:1979–1994, 2001.

[5] C. Soize. Random matrix theory for modeling uncertainties in computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 194:1333–1366, 2004.

[6] C. Soize. Generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions. *International Journal for Numerical Methods in Engineering*, pages 1–32, 2009.

[7] M.S. Bonney and M.R.W. Brake. Utilizing soize's approach to idetify parameter and model uncertainties. Technical Report SAND2014-19209, Sandia National Laboratories, Albuquerque, NM, 2014.

[8] B.A. Robertson, M.S. Bonney, C. Gastaldi, and M.R.W Brake. Quantifying epistemic and aleatoric uncertainty in the ampair 600 wind turbine. *Proceedings of the 33rd International Modal Analysis Conference*, 2015.

[9] M.S. Bonney, D.C. Kammer, and M.R.W. Brake. Determining model form uncertainty of reduced order models. *Proceedings of the 34rd International Modal Analysis Conference*, 2016.

[10] A. Batou, C. Soize, and S. Audebert. Model identification in computational stochastic dynamics using experimental modal data. *Mechanical Systems and Signal Processing*, 50-51:Pages: 307–322, 2014.

[11] M.S Bartlett. Statistical estimation of density functions. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 245–254, 1963.

[12] M.C. Jones. The performance of kernel density functions in kernel distribution function estimation. *Statistics & Probability Letters*, 9(2):129–132, 1990.

[13] M. Rudemo. Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, pages 65–78, 1982.

[14] S.J. Sheather and M.C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 683–690, 1991.

[15] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.

[16] A. Shapiro and T. Homem-de Mello. On the rate of convergence of optimal solutions of monte carlo approximations of stochastic programs. *SIAM Journal on Optimization*, 11(1):70–86, 2000.

[17] M.D. McKay, R.J. Beckman, and W.J. Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[18] J.C. Helton, J.D. Johnson, C.J. Sallaberry, and C.B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91(10):1175–1209, 2006.

[19] J.C. Helton and F.J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, 2003.

[20] D.C. Kammer and M.J. Triller. Selection of component modes for craig-bampton substructure representations. *ASME J. Vib. Acoust.*, 188(2):264–270, 1996.

[21] R.R. Craig and M.C.C. Bampton. Coupling of substructures for dynamic analysis. *AIAA Journal 6*, pages 1313–1319, 1968.

[22] R.R. Craig. Coupling of substructures for dynamic analyses - an overview. *41st Structures, Structural Dynamics, and Materials Conference and Exhibit, Structures, Structural Dynamics, and Materials and Co-located Conferences*, 2000.

[23] R.R. Craig and C.J Chang. On the use of attachment modes in substructure coupling for dynamic analysis. *18th Structural Dynamics and Materials Conference, Structures, Structural Dynamics, and Materials and Co-located Conferences*, 1977.

[24] D. Klerk, D.J. Rixen, and S.N. Voormeeren. General framework for dynamic substructuring: history, review and classification of techniques. *AIAA journal*, 46(5):1169–1181, 2008.

[25] R.R. Craig. Substructure methods in vibration. *Journal of Vibration and Acoustics*, 117(B):207–213, 1995.

[26] R.H. MacNeal. A hybrid method of component mode synthesis. *Computers & Structures*, 1(4):581–601, 1971.

[27] S. Rubin. Improved component-mode representation for structural dynamic analysis. *AiAA Journal*, 13(8):995–1006, 1975.

[28] D.J. Rixen. A dual craig–bampton method for dynamic substructuring. *Journal of Computational and applied mathematics*, 168(1):383–391, 2004.

[29] R.R. Craig and A.J. Kurdila. *Fundamentals of Structural Dynamics*. John Wiley and Sons, Inc., 2nd edition, 2006.

[30] S.N. Afriat. Orthogonal and oblique projectors and the characteristics of pairs of vector spaces. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 53, pages 800–816. Cambridge Univ Press, 1957.

[31] D.C. Kammer. A hybrid approach to test-analysis-model development for large space structures. *Journal of Vibration and Acoustics*, 113(3):325–332, 1991.