SANDIA REPORT

SAND2015-6896
Unlimited Release
Printed August, 2015

Determining Reduced Order Models for Optimal Stochastic Reduced Order Models

Matthew S. Bonney, 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

SAND2015-6896
Unlimited Release
Printed August, 2015

Determining Reduced Order Models for Optimal Stochastic Reduced Order Models

Matthew S. Bonney, Matthew R.W. Brake
Component Science and Mechanics
Sandia National Laboratories
MS 1070; P.O. Box 5800
Albuquerque, NM 87185-1070

Abstract

The use of parameterized reduced order models(PROMs) within the stochastic reduced order model (SROM) framework is a logical progression for both methods. In this report, five different parameterized reduced order models are selected and critiqued against the other models along with truth model for the example of the Brake-Reuss beam. The models are: a Taylor series using finite difference, a proper orthogonal decomposition of the the output, a Craig-Bampton representation of the model, a method that uses Hyper-Dual numbers to determine the sensitivities, and a Meta-Model method that uses the Hyper-Dual results and constructs a polynomial curve to better represent the output data. The methods are compared against a parameter sweep and a distribution propagation where the first four statistical moments are used as a comparison. Each method produces very accurate results with the Craig-Bampton reduction having the least accurate results. The models are also compared based on time requirements for the evaluation of each model where the Meta-Model requires the least amount of time for computation by a significant amount. Each of the five models provided accurate results in a reasonable time frame. The determination of which model to use is dependent on the availability of the high-fidelity model and how many evaluations can be performed. Analysis of the output distribution is examined by using a large Monte-Carlo simulation along with a reduced simulation using Latin Hypercube and the stochastic reduced order model sampling technique. Both techniques produced accurate results. The stochastic reduced order modeling technique produced less error when compared to an exhaustive sampling for the majority of methods.

Acknowledgment

The authors would like to acknowledge those whose help made this report possible. Dan Kammer from The University of Wisconsin-Madison who designed and developed the Meta-Model method. Brett Robertson who ran the high-fidelity model cases. Jeffery Fike who developed the Hyper-Dual number method and ran the model simulations using Hyper-Dual numbers. Richard Field who helped with the SROM representation and calculations.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Motivation

The implementation of high-fidelity models has recently become the engineering standard in order to produce more accurate results. This increase in model fidelity has emphasized the question of "how accurately do the computational models represent reality?". In order to know this, the practice of uncertainty quantification (UQ) must be utilized and improved in order to tell the accuracy of the high-fidelity models. Uncertainty quantification, however, requires a statistical approach which typically includes distribution sampling that requires multiple evaluations of the model [8]. This raises the problem of computation ability since each high-fidelity model takes a large amount of computational power and man-hours to evaluate.

One of the most efficient way to evaluate the model is to create a Reduced Order Model (ROM) that captures the required information, such as natural frequency or mode shapes, while eliminating the undesired information [5]. The most common method for creating ROMs is to perform a domain shift into the frequency domain. This is utilized in the most common ROM, the Craig-Bampton representation [14] which is discussed in a later section.

One major drawback to using these ROMs is that they do not give any information about a system that has any perturbation in the model. This is usually the case for uncertain material properties or machining tolerances. For each one of these model changes, a brand new ROM must be created, which is very computationally expensive. The goal of this research is to be able to use select information to create a ROM for any value of the model changes.

The method used in this report for implementing the Stochastic Reduced Order Model (SROM) is not the typical use of this technique. Typically, the SROM is used to determine the input of the high-fidelity model. Once the input is determined, then a Taylor series with a finite difference approach to determine sensitivities is used to create a mathematical model instead of the high-fidelity model. With the mathematical model, a brute-force Monte-Carlo sampling method is used to determine the output distribution. The material in this report have a fixed input of the high-fidelity model and then uses the SROM technique to sample the output distribution. This is done in order to show the efficacy of the technique compared to a traditional Monte-Carlo sampling and to another reduced sampling technique, the Latin

Hypercube sampling.

Stochastic Reduced Order Model and Parameterized Reduced Order Model

This research uses two techniques in order to produce a ROM for any given model configuration: SROM and Parameterized Reduced Order Model (PROM). SROM refers to the technique used to optimally select the configurations that the high-fidelity model is evaluated. The uncertain parameter of interest, material property or geometric property, has a specified distribution which is assumed to be known. These high-fidelity calculations are the basis for the generation of the ROM that is used in the determination of the output distribution. The exact procedure is not discussed in depth in this report but is discussed in [6, 7, 8]. Using the SROM technique, in general, uses an optimal subset of the sampling distribution to best match the first several statistical moments of the distribution. The SROM is used to create several ROMs at different configurations with different probabilities that are used to estimate the output distribution of interest. This report discusses what should be done with the information gathered from the high-fidelity calculations. The desired information that is calculated from the models can be any type of information such as: Natural frequencies, mode shapes, fatigue life, ultimate tensile strength, etc. The data used for the calculations in this report are the natural frequencies. This is chosen for the relative simplicity of the calculation which allow for truth data to be calculated in order to validate the methods being compared.

While the previous step focuses on the high-fidelity model and what information is gathered, the next step utilized the previous information to produce a PROM. This parameterization is based on the different configurations that the high-fidelity model took in the previous step. The example utilized in this report has a variation in the material property, the Young's modulus. This is due to the simplicity of the model's ability to change this value. There are two types of parameterizations that are used in this report: Sensitivities and Polynomial fit. The Sensitivities refer to the expression of the output function's derivative to the input change. Multiple derivatives are calculated and utilized. Some of the models will determine these sensitivities by using a finite-difference approach while other calculate these analytically. These derivatives are then used in conjuncture with the Taylor series expansion of the nominal output to get an approximation of the true output. The polynomial fit refers to replacing the output function with a polynomial expression. The coefficients are determined by using the nominal output from the high-fidelity calculations.

The paper is organized as follows: Chapter 2 gives an overview of the five different type of PROMs used in this comparison including some of the requirements for the construction each PROM. In Chapter 3, the example model that is used for these calculations is shown. Chapter 4 shows the results for the case where the input variable is fairly well known and has a small range of values it can take, while Chapter 5 gives the example where very little information is actually known about the input variable which constitutes a large range of

possibilities. Some conclusions and general remarks about each possible model are given in Chapter 6.

Chapter 2

Reduction Models

The choice of what type of reduction model to use is very important. Each method has its own drawbacks and advantages. This research focuses on five different models. Some of these models are physical while some are purely numerical. Each method is discussed along with instructions for the computation so that these results can be replicated if desired. If the data and functionals are desired, please contact the authors. The methods that are used are: A Taylor series expansion using finite difference methods to determine sensitivities, A Proper Orthogonal Decomposition of the data, Using a Craig-Bampton representation, Implementing Hyper-Dual numbers in order to obtain the sensitivities, and using a Meta-Model approach that fits a polynomial to the data.

Taylor Series

The first method is the most straight forward and the most commonly used in the SROM technique. This method uses the output data and a finite difference approach to determine the sensitives. The method for using the finite difference approach can be found in [12]. This computes the sensitivities of different orders as a weighed sum of the functional output at different parameter values. A second order first derivative for a scalar variable is seen as

$$f'(X_o) = \frac{f(X_o + \Delta) - f(X_o - \Delta)}{2\Delta}, \tag{2.1}$$

where $f()$ is the output function $f : R \mapsto R$, $f'()$ is the first derivative with respect to X, X_o is the nominal input parameters, and Δ is the change in the input parameters that the function was evaluated at. This equation assumes that the high-fidelity model was evaluated at at-least three different input parameter values, $X_o - \Delta$, X_o, and $X_o + \Delta$. The more data-points that are available, the more accurate the sensitivities become along with the added benefit that higher order derivatives can also be determined. This research used five data points to generate the sensitivities which can produce up to the fourth derivative of the functional. This model also assumes that the uncertain parameters are independent, which is an important assumption to consider and verify. Otherwise, the finite difference approach can only determine conditional derivatives.

Once the sensitivities are determined, they can be applied to a Taylor series expansion

about the nominal value. This can be seen as

$$f(X_o + \delta) = f(X_o) + \frac{1}{1!}f'(X_o)\delta + \frac{1}{2!}f''(X_o)\delta^2 + \cdots + \frac{1}{n!}f^{(n)}(X_o)\delta^n + \cdots . \tag{2.2}$$

In Eq. 2.2, δ refers to the difference from the nominal parameter value and the superscript (n) is the n^{th} derivative of the function, which is calculated by Eq. 2.1 for the first derivative. This equation is an infinite series but is truncated at the highest order derivative available.

This method is very straight forward; it does not require any change in the high-fidelity code and is accurate for small changes in the input parameters. One problem this method has is that the accuracy is directly related to the amount of times the code is evaluated. The accuracy is also decreased if the data-points that the sensitivities are derived at are far apart. This is the most commonly used method of reduction since the analysis is very simple, but can be time consuming since multiple code evaluations must be used.

Proper Orthogonal Decomposition

The next method is the proper orthogonal decomposition (POD). This method utilizes basis vectors in order to reduce the computational time and storage requirements of propagation. This method is very similar to the Taylor series method in that it uses a finite difference approach to approximate the derivatives. The main difference comes in how the uncertainty parameterization is performed. This method is explained in more detail in [4]. If the data, such as natural frequencies or mass matrix, is sorted into a matrix A, then that matrix can be decomposed via a singular value decomposition. This decomposition takes the form of $A = UDV^T$ where U and V are a collection of orthogonal basis vectors in the form of matrices and D is a diagonal matrix that contain the principal values of the matrix A. It is also noted that this is very similar to an eigenvalue decomposition but does not require A to be a square, symmetric matrix.

The diagonal matrix D can be truncated basis on its values. This can give a reduced order representation of the matrix A which will require less computational time to compute and require less memory storage than the Taylor series method. The vectors U and V are invariant with respect to the change in input parameters [13], so the only value that need to be parameterized is the diagonal matrix D. This invariance comes from the definition of U and V which are direction vectors to the singular value space. The derivatives of D are found using the finite difference approach discussed in the Taylor series method.

This method is advantageous because the matrix D is diagonal and is smaller in size than the raw data matrix A. The disadvantage is that the data matrix must be decomposed which takes more overhead to compute since three matrices are being evaluated instead of just one. This increase in overhead time must be compared to the savings in the propagation step of UQ.

Craig-Bampton

While the previous two methods were numerical based, the Craig-Bampton method is a physical reduction model. This method takes the system matrices that are generated from the model and reduces the size by introducing a hybrid coordinate system that contains both physical interface degrees of freedom (DOF) and fixed-interface modal DOF. This method is explained in [5]. This is useful because the modal DOF are by definition independent so they can be truncated based on the frequency. It is common to choose a truncation frequency based on twice the desired range of importance. For example, if the highest frequency of interest is around $1kHz$, then the fixed-interface modes that are kept are all the modes up to $2kHz$. This is a rule-of-thumb in order to get acceptably accurate results. The more modes that are selected, the more accurate the solution, but that requires more DOF which diminishes the reduction effect.

The first step is to determine the interface of interest. This is usually simple but the size of the interface has a great affect on the reduction of the system since these interface DOF will not be reduced. With this determined, the system matrices are partitioned into interface and interior DOF. This is defined as

$$u = \begin{bmatrix} u_i \\ u_o \end{bmatrix}, \tag{2.3}$$

where u is the physical DOF vector, subscript i is the interior, and subscript o is the interface. In order to transform into this hybrid coordinate system, a transformation matrix is defined as

$$T_{cb} = \begin{bmatrix} \Phi & \Psi \\ 0 & I \end{bmatrix}. \tag{2.4}$$

In Eq. 2.4, Φ is the fixed interface mode shapes, which are mass-normalized, Ψ are the shapes the interior takes due to a unit displacement in each interface DOF individually, 0 is a matrix of zeros, and I is an identity matrix. The matrix Ψ can be computed by using the static shapes which takes the form of $\Psi = -K_{ii}^{-1}K_{io}$ where K is the stiffness matrix.

This transformation matrix is applied to the system so that the system matrices are reduced. The mass, stiffness and forcing vector are shown in Eq. 2.5 as

$$M_{cb} = T_{cb}^T M T_{cb} \quad K_{cb} = T_{cb}^T K T_{cb} \quad F_{cb} = T_{cb}^T F, \tag{2.5}$$

with M as the high-fidelity mass matrix and F is the external forcing vector. Once the Craig-Bampton system matrices are generated, then any typical analysis can be done such as eigenvalue analysis or time integration. It is noted that if all of the fixed-interface modes are kept, then Eq. 2.5 does not change the size of the matrices and results in the exact solution.

This method is different from the rest of the methods in that it is a model based in the physical domain. This also doesn't require a full high-fidelity model code evaluation, just the extraction of the system matrices are required. The sensitivities are computed normally by finite difference, but if available, the derivatives can be computed using Hyper-Dual numbers which will be explained in section of Hyper-Dual reduction. This is not the same as the Hyper-Dual reduction but the underlying math can be used in both methods. This method requires a calculation of the model every time. The other methods so far only required a plug-into a formula but this method requires a full code evaluation which takes longer to compute compared to the other methods but still less than the high-fidelity model. This method is useful for a physical interpretation of the system and can be used in multiple techniques such as sub-structuring.

Hyper-Dual Reduction

The previous models consisted purely in the real domain where the Hyper-Dual reduction consists in the complex domain. The Hyper-Dual reduction uses a generalized complex approach to the reduction. The use of these generalized complex numbers allows for a determination of derivatives in a single code evaluation [9]. The generalized complex number is called a Hyper-Dual number. This is a multi-dimensional expansion of the dual number. This is similar to the use of quaternions for the ordinary imaginary number. The math behind these Hyper-Dual numbers are explained in [1, 9, 10, 11], but a brief explanation is also presented here.

The basis of the Hyper-Dual number is the dual number. The duel number, ϵ is similar to an ordinary complex number but is defined as $\epsilon^2 = 0$. This is very useful since all of the higher-order moments of ϵ are by definition zero, such as ϵ^3. Hyper-Dual numbers take this principal and expands it into the multi-dimensional domain. This report uses a eight dimensional approach, which gives the third derivative of the function. The determination of the derivatives come from the Taylor series expansion for a second derivative of the function is given as

$$f(X_o+h_1\epsilon_1\mathbf{e_i}+h_2\epsilon_2\mathbf{e_j}+0\epsilon_1\epsilon_2) = f(X_o)+\epsilon_1 h_1\frac{\partial f(X_o)}{\partial x_i}+\epsilon_2 h_2\frac{\partial f(X_o)}{\partial x_j}+\epsilon_1\epsilon_2 h_1 h_2\frac{\partial^2 f(X_o)}{\partial x_i \partial x_j}, \quad (2.6)$$

in which $\mathbf{e_i}$ is a zero vector with 1 in the i^{th} position and x_i is the i^{th} component of the input vector of parameters X_o. The form in Eq. 2.6 is the same for the third derivative by the addition of an ϵ_3 term and all the subsequent cross-products. Equation 2.6 assumed that the function is differentiable up to the second derivative. It is also noted that the cross-products are independent of the order of multiplication such as $\epsilon_1\epsilon_2 = \epsilon_2\epsilon_1$ which is not the case for quaternions.

The useful thing about the implementation of Hyper-Dual numbers is that if a class is created, then the high-fidelity code does not have to change by very much. This is however

one of the major drawbacks of using this method. The high-fidelity code must be able to compute Hyper-Dual numbers. This requires some configured code to run with Hyper-Dual numbers or a knowledge of how the code evaluates the model and programming knowledge. SIERRA has already been implemented to utilize Hyper-Dual numbers on select versions. Contact the authors if access to this version is desired. While the special programs required to perform this method is required, there are several helpful aspects of this model. This method only requires a single high-fidelity code evaluation which can significantly reduce the time required in the analysis since the rest of the methods so far required multiple code evaluations. Another benefit is that the derivatives are exact while the other methods only give an approximation. The storage of this data can either be done as a single value with unreal parts or as multiple real values.

Meta-Model

This method is the newest of the five methods. The method is currently being developed jointly between Sandia and The University of Wisconsin-Madison. While no publications have been written as of the time that this report is written, all credit is given to Dr. Dan Kammer. The basic idea of this method is to fit a polynomial curve to the output data. The main contribution of this method is the use of Hyper-Dual numbers for additional data. This method combines the Taylor series and Hyper-Dual reduction methods. Multiple code evaluations of the Hyper-Dual method are performed and a polynomial curve is fit to the function values and the derivatives that are given by the Hyper-Dual code evaluations. The exact polynomial equations vary based on the number of code evaluations and the number of derivatives that are determined. The polynomial fit has the form of

$$f(X) = F_0 + F_1 X + F_2 X^2 + \cdots + F_k X^k, \tag{2.7}$$

where F_i is an unknown coefficient matrix and k is the maximum polynomial power. k is determined by $k = (n+1)m-1$ where n is the number of derivatives calculated at each code evaluation and m is the number of code evaluations. The coefficient matrices are determined by matching the polynomial and the derivative of the polynomial to the code evaluation results.

This method have several advantages to the other methods. This model is the quickest to evaluate in the propagation part of the UQ. The reason for this is due to the fewer multiplications required since it doesn't require a factorial computation that the Taylor series requires. This is also due to the fact that only a polynomial is evaluated which is simple multiplication. This method was developed to take the best of the Taylor series and Hyper-Dual methods. One disadvantage to this method is the determination of the unknown coefficient matrices. This can become very difficult if the output data is large. The polynomial coefficients are determined based on the simple equation $AX = B$ where the vector X is written as $X = [F_o \; F_1 \; \cdots]^T$, which are the unknown matrix coefficients, and B is the output data from the code evaluations including function values and the derivatives. The

matrix A is derived to equate the polynomial value and derivatives at each code evaluation. For the example used in this report, the natural frequencies are used and the coefficient matrices are fairly simple to calculate. The mode shapes was explored but the coefficient matrices were unable to be determined. This was due to the large size.

Chapter 3

Brake-Reuss Beam Example

In order to evaluate these methods, the example is chosen to be the Brake-Reuss beam. This system is chosen due to the relative simplicity and the ability to complicate the system [2, 3]. The material is assumed to be an uncertain temper of Aluminum. This uncertain temper leads to uncertainty in the Young's modulus. The physical beam is constructed by two identical specimens and screwed together through a lap joint. The assembly can be seen in Figure 3.1.

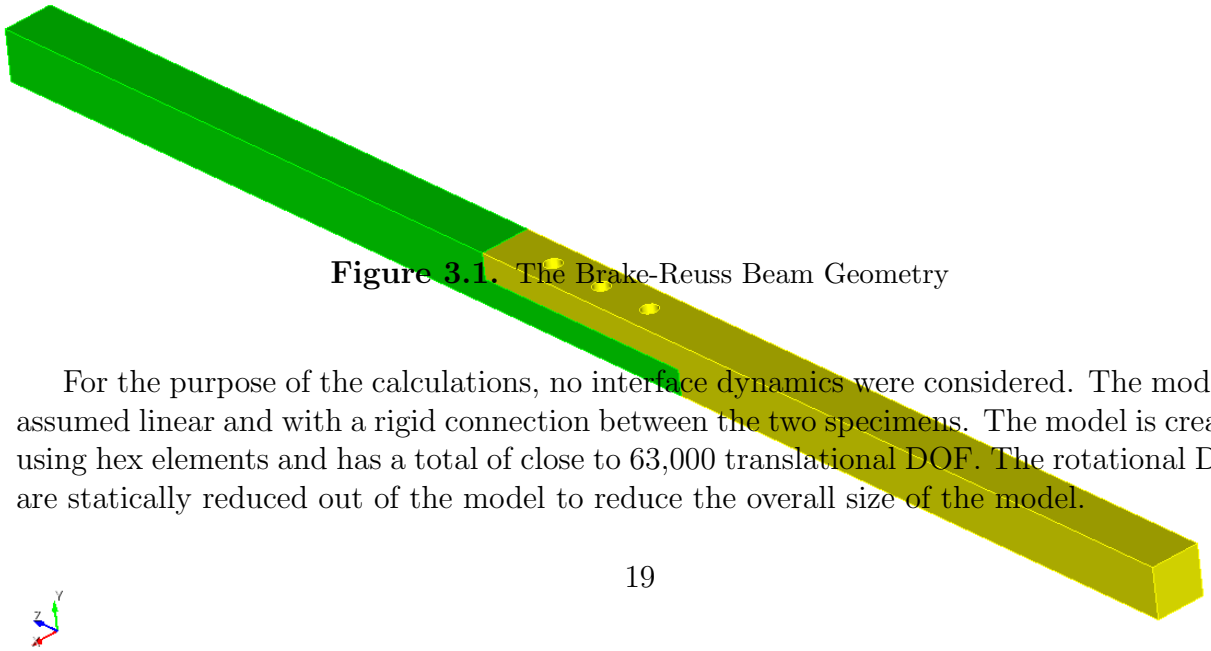

Figure 3.1. The Brake-Reuss Beam Geometry

For the purpose of the calculations, no interface dynamics were considered. The model is assumed linear and with a rigid connection between the two specimens. The model is created using hex elements and has a total of close to 63,000 translational DOF. The rotational DOF are statically reduced out of the model to reduce the overall size of the model.

The comparison between the models will consist of the first fourteen elastic free-free modes. This selection contains all modes up to 4400 Hz. The selection of the first fourteen is decided based on determining the effect of this uncertainty in the low and mid-frequency ranges. This range of frequencies contains bending, torsional, and axial mode shape deformations. Another important aspect to consider is the interface DOF for the Craig-Bampton representation. For this model, one end of a specimen was decided to be the interface DOF. This was chosen for the relative small amount of node, 81 nodes on each end, and the simplicity compared to considering the joint interface that is hard to determine and has a large amount of nodes in that contact area. One important note is that the screws are not physically modeled in this analysis. Since this report considers the interface as rigid, the addition of the screws are unnecessary.

This example can become a very useful example for both model reduction and mechanical joint analysis. For model reduction, there is not a complicated geometry or materials to consider; It is a simple metal beam. The mechanical joints are a very interesting aspect to this example. Since the specimens are manufactured, each can contain slightly different surface finishes at the lap joint which can lead to different dynamics experimentally.

Chapter 4

Small Variability

The first simulation that is presented is for if some knowledge is known about the distribution of the uncertain parameter. This is interpreted as a very small variability. The first result is the parameter sweep. This gives an idea on how accurate the reduction method is compared to the true solution. All modes were computed individually and the first and last natural frequency are presented. The results are presented in Figures 4.1-4.2. These figures show the first order approximation which is the least accurate.

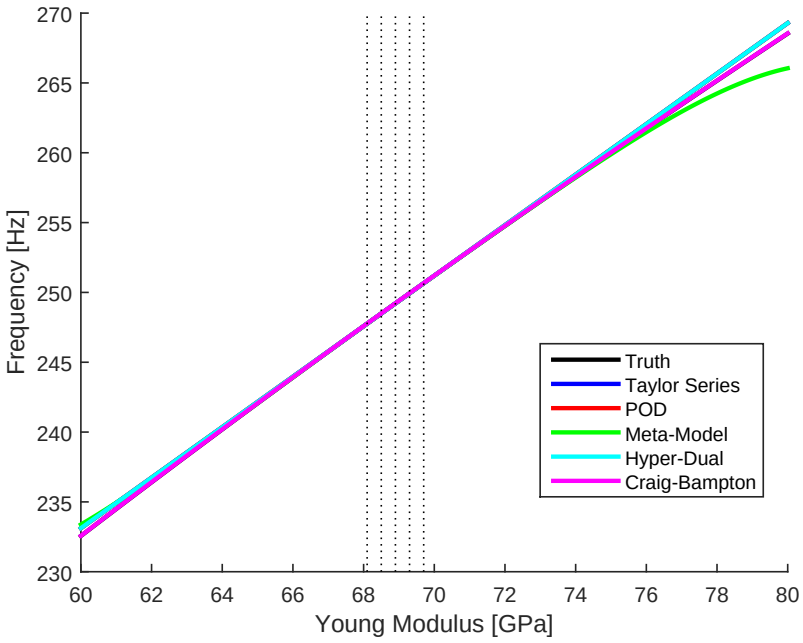

Figure 4.1. First Natural Frequency with Small Variability Parameter Sweep

In Figures 4.1-4.2, the vertical black dotted lines correspond to the data points that the reduction was produced at. The Taylor series, POD, and Craig-Bampton used the data from all five data points. The Meta-Model used the middle and the outer points so that only three points were used in the derivation. The Hyper-Dual method only used the nominal value

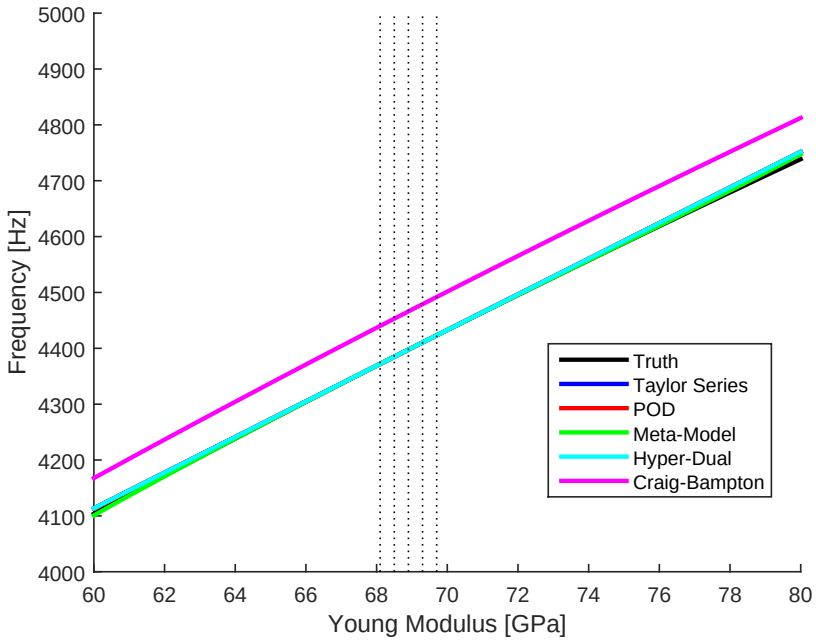

Figure 4.2. Fourteenth Natural Frequency with Small Variability Parameter Sweep

which is the middle vertical line. It can be seen that for the first natural frequency, all the methods do very well for the most part but start to loose accuracy closer to the extremes of the parameter sweep. The fourteenth natural frequency however shows more inaccuracy. The first thing is that the Craig-Bampton model has a frequency offset from the truth data. This is not unexpected since the Craig-Bampton is a physical based reduction method and the higher frequency mode require more information, which is reduced out. This error however can be acceptable. The second thing of notice is that the different methods start to vary from the truth data at more points in the fourteenth natural frequency compared to the first natural frequency.

One of the major motivations for this report is to compare both accuracy and time requirements to use these PROMs. The first comparison is the required time to compute these parameter sweep curves. This utilizes 201 model evaluations. The time required to produce the parameter sweep curve is presented in Table 4.1. Each method uses less than 6 high-fidelity code evaluations while the truth data requires 201 high-fidelity code evaluations.

The order refers to the number of derivatives are used in the parameterization. The Craig-Bampton representation takes the longest to compute by a factor of almost three to the next highest value. This is a major drawback but it still contains a physical meaning. The Meta-Model requires the least amount of time to compute the curve. This is due to the polynomial that is more simple to evaluate than a Taylor series. The Meta-Model however

Table 4.1. Computation Time to Generate Parameter Sweep [s]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	28.78	6.61	16.20	0.21	88.14
2	29.86	6.65	17.48	0.19	85.01
3	30.75	6.64	18.90	0.18	86.07
4	32.08	6.55	N/A	N/A	86.07

takes more time to compute the unknown coefficients than the Taylor series.

The next result to report is the root mean square (RMS) error of the parameter sweep. Model reduction error can be quantitatively measured using this value. The model reduction error measure takes the RMS of the difference between each model and the truth data at each natural frequency across the different Young's modulus values for each order. This produces a singular value that can be used for comparison. The first and fourteenth natural frequency are presented in Table 4.2-4.3 respectively.

Table 4.2. RMS Error for Fundamental Frequency [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.119	0.119	0.119	0.225	0.004
2	0.008	0.008	0.008	41.96	0.004
3	0.001	0.001	0.001	Large	0.004
4	0.000	0.000	N/A	N/A	0.004

Table 4.3. RMS Error for Fourteenth Natural Frequency [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.119	0.119	0.119	0.045	1.556
2	0.008	0.008	0.008	8.422	1.556
3	0.001	0.001	0.001	Large	1.556
4	0.000	0.000	N/A	N/A	1.556

The first trend to note is that the Taylor series, POD, and Hyper-Dual methods have the same accuracy. The POD and Taylor are expected to have the same accuracy but the Hyper-Dual gives the same accuracy because the points used for the finite difference are very closely spaced which gives the same derivative as the Hyper-Dual. The Craig-Bampton model is the only one that will have different accuracies for different natural frequencies since it is physical based while the other methods are purely numerical based. It is also the only model that is able to calculate additional natural frequencies, or another desired output, without have to be completely reformulated.

Another important trend is for the Meta-Model. The first order matches the value and the first derivative at each of the data-points. This provides a very accurate representation of the data. When the extra derivatives are also matched, the error increases over the range. This is believed to occur due to added power on the polynomial curve. The value of perturbation gets large quicker which allows for larger error to occur. This can easily be seen in Figure 4.3 near the extremes where the Meta-Model creates very large error. Figure 4.3 show the second order approximation for the fundamental frequency where all the models are overlay-ed on each other besides the Meta-Model curve.

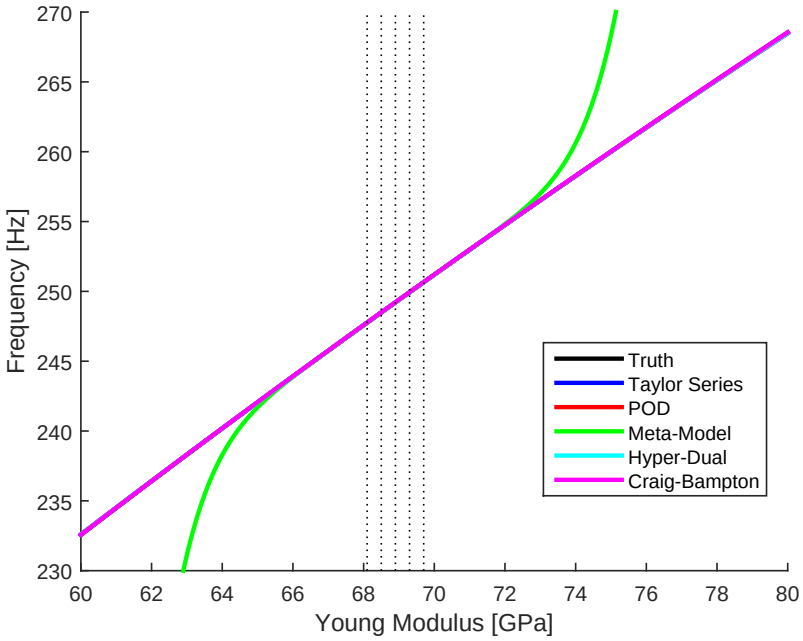

Figure 4.3. First Natural Frequency with Small Variability Parameter Sweep using 2nd Order Approximations

Chapter 5

Large Variability

The simulations the only contained small variations is unable to differentiate between the Taylor series, POD, and Hyper-Dual methods based on accuracy. There was a difference in computation time but the computations resulted in the exact same values. In order to determine if there is a difference, the variation in the Young's Modulus is increased.

Parameter Sweep

A parameter sweep is the first simulation of this larger variability. The main difference is the range of the sweep; The previous sweep ran between $60 - 80$ GPa, while this sweep is between $50 - 300$ GPa. Each model is reconstructed at different nominal values. The results of the sweep is shown in Figures 5.1-5.2.

It is easy to see that the accuracy has decreased due to the large variability. The first natural frequency doesn't show much decrease in accuracy, which shows the first order analysis. These are linear with the change in the Young's modulus due to the Taylor series. There are only two non-linear method, at this order, the Craig-Bampton and Meta-Model which follow the truth data almost perfectly. The fourteenth natural frequency shows the most differences between the methods that can accurately predict the first natural frequency.

The accuracy of the models is very dependent on the output relationship to the input parameters. The more linear the relationship, the more accurate the reduction is. The Young's modulus is fairly linear with respect to the natural frequencies. Other types of variability, such as Poisson ratio or a geometric length, decreases the accuracy of these models reduction techniques. These other possible uncertainties are currently being researched. The first quantitative measurable difference between the methods is the computational time to compute the parameter sweep curves. These computation times are shown in Table 5.1 that show the same trends as in Table 4.1. The Meta-Model is the fastest by a large amount and the Craig-Bampton model takes the most time.

The next measure is the RMS error. This measure is the same as in Chapter 4. The results are presented in Tables 5.2-5.3. One important trend is that the Taylor series and POD gives the exactly same solution. This is a good comparison because the time required to evaluate the POD model is significantly less than Taylor but gives the same accuracy

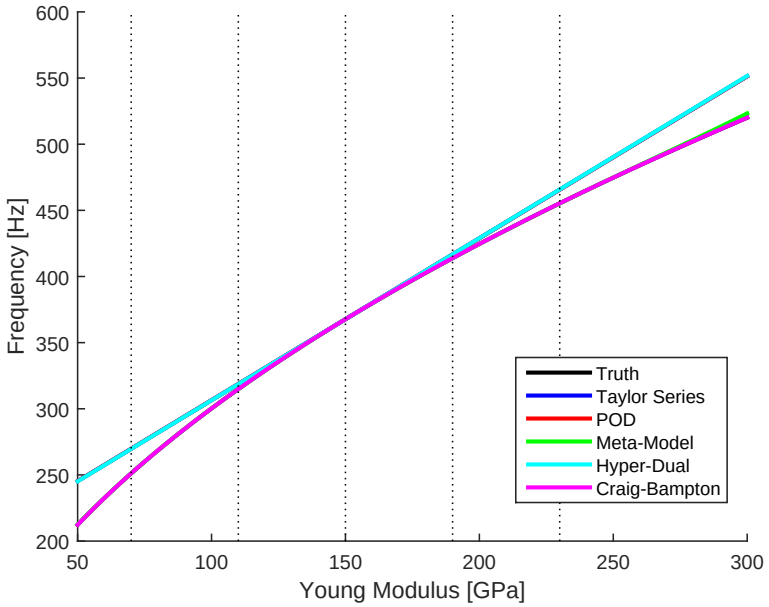

Figure 5.1. First Natural Frequency with Large Variability Parameter Sweep

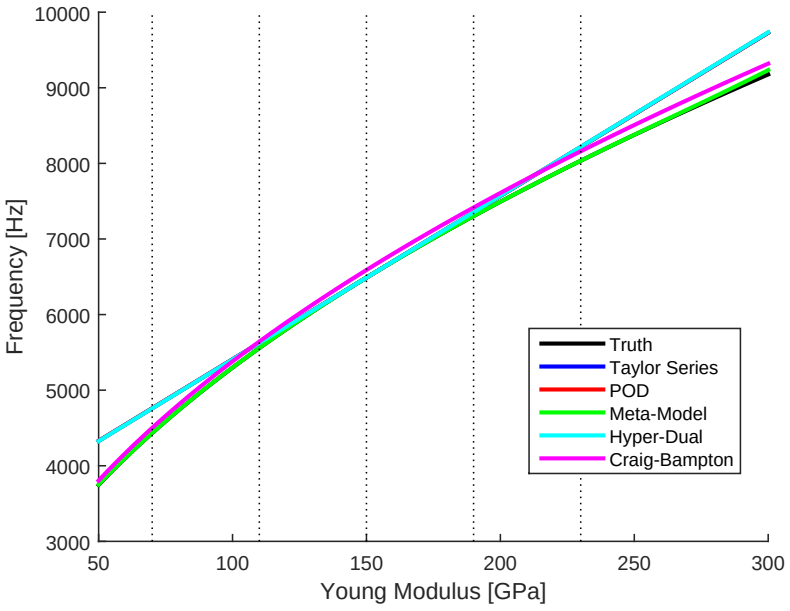

Figure 5.2. Fourteenth Natural Frequency with Large Variability Parameter Sweep

Table 5.1. Computation Time to Generate Parameter Sweep [s]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	33.75	8.00	20.56	0.24	107.75
2	35.11	8.36	21.40	0.23	107.84
3	35.41	8.16	22.75	0.23	109.87
4	37.04	8.06	N/A	N/A	109.87

measurement. Another important trend is the comparison between the Taylor series and the Hyper-Dual representation. The Hyper-Dual only takes data from a single data-point and it provides a more accurate solution than the Taylor series method which uses five data-points. This shows the advantage of using the Hyper-Dual method.

Table 5.2. RMS Error for Fundamental Frequency [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	4.311	4.311	4.304	0.127	0.004
2	1.519	1.519	1.488	0.052	0.004
3	0.808	0.808	0.671	0.269	0.004
4	0.221	0.221	N/A	N/A	0.004

Table 5.3. RMS Error for Fourteenth Natural Frequency [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	4.311	4.311	4.304	0.127	1.556
2	1.519	1.519	1.488	0.053	1.556
3	0.808	0.808	0.671	0.270	1.556
4	0.221	0.221	N/A	N/A	1.556

Another interesting trend is between the Meta-Model and Hyper-Dual methods. The Meta-Model provides better accuracy than the Hyper-Dual method at about 1% of the computational time. The determination of the model takes more time but can overall be a cost savings. The last trend of interest is the Craig-Bampton model. This model produced the same accuracy for both the scenarios. The results for the Craig-Bampton model in Table 4.3 and Table 5.3 are the same values.

Distribution Comparison

The accuracy of the reduction models are not purely based on a point-by-point comparison from a parameter sweep. The accuracy is also dependent on how the model correctly generates the output distribution. This is the ultimate information that is desired when using a ROM. This allows for statistical correlations and reliability calculations to be performed. In order to generate this distribution, a Monte-Carlo (MC) sampling analysis is performed. This is a traditional MC analysis which takes random values from the input parameter and propagates it into the final distribution. The first important determination is the distribution of the input parameter. For some parameters, this is a simple determination based on large amount of data while some others are slightly more up to interpretation. One major consideration is the bounds of the distribution must remain physical. For example, the mass density must be positive. For the Young's modulus, a log-normal distribution was selected as the distribution. This was chosen based on the bounds. The log-normal distribution has a support of $(0, \infty)$. This distribution never reaches zero but is always positive. It is chosen that this distribution has a mean of 150 GPa and a coefficient of variation of 14%. This was chosen to fully test the range of the parameter sweep. The input probability density distribution can be seen in Figure 5.3 with $100,000$ samples.

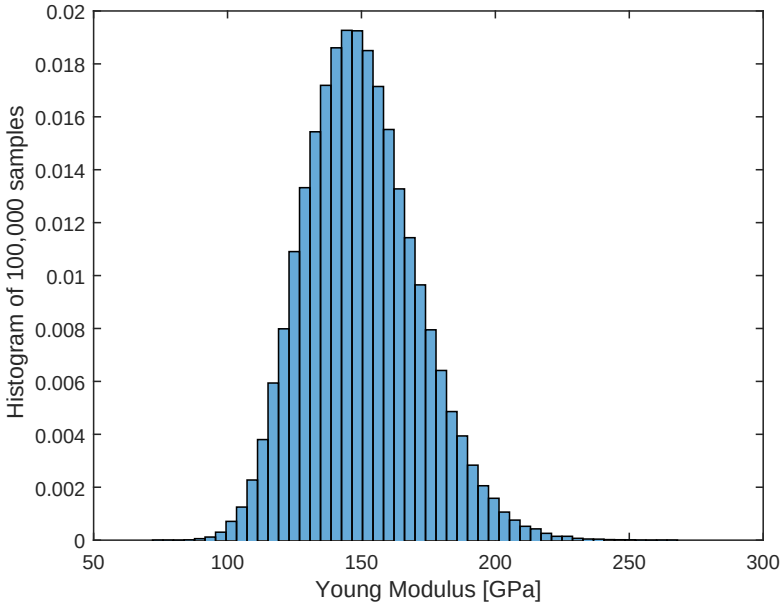

Figure 5.3. Histogram of 100,000 Samples of Young's Modulus in Monte-Carlo analysis

The next major consideration is how to compare the distributions. There are several ways to compare distributions: visually, box-plot, statistical moments, etc. Statistical moments were chosen to be the comparative value, in particular, the first four statistical moments

and the lower fractile were chosen. The probability distributions are also presented visually to compare qualitatively. This comparison between statistical moments requires some truth data to be compared to. The output distribution is not analytically known so an approximation must be used. In order to approximate the real solution, the data from the parameter sweep is utilized. For each value of the Young's modulus, the tabular data of the high-fidelity model is used as the truth data. If the value is between two data-points in the tabular data, a linear interpolation is used to determine the true value.

The first result is a qualitative analysis of the distribution. Figure 5.4 show the distributions for each type of model reduction and the truth data. The truth data in Fig. 5.4 is hard to see but it is overlay-ed with the Meta-Model. By a visual examination, the distributions look very similar. The first trend is that the Craig-Bampton method is shifted to the right of the other curves. The rest of the curves are very close to the truth data. A point-by-point difference of the distribution is shown in Figure 5.5. One of the main things to see in Figure 5.5 is that the Meta-Model has very little difference in probabilities. The other trend to note is that the Taylor series, POD, and Hyper-Dual methods are close to the same amount of error as seen by the overlay of the lines.

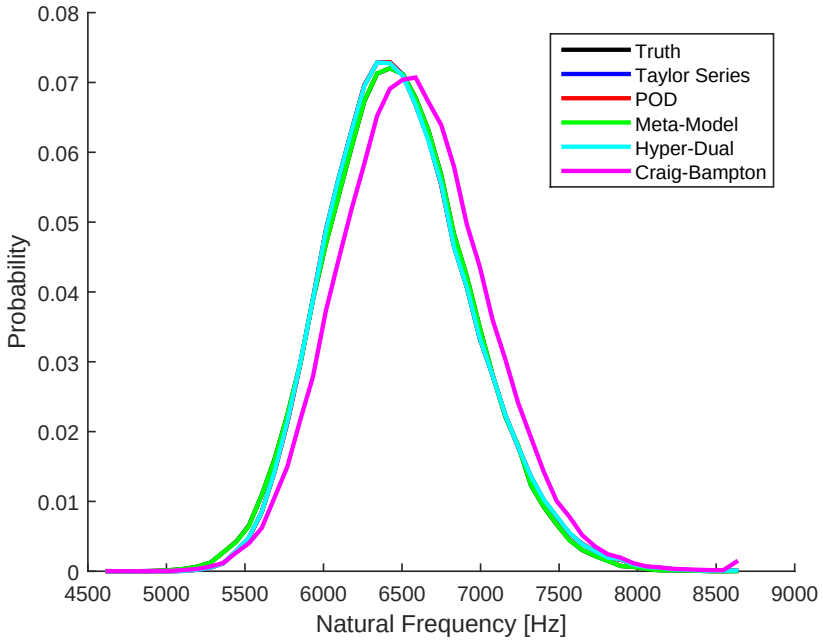

Figure 5.4. Output Distributions for Each Method

The qualitative analysis does not show many differences between the distributions besides the Craig-Bampton and Meta-Model. A quantitative analysis is required. The first analysis is the time requirement to compute the distribution. These time trend follow the same as the parameter sweep time trends. For $100,000$ samples, the Meta-Model takes 5 seconds and the Craig-Bampton takes 9 minutes on 128 cores with the help of parallel processing. The

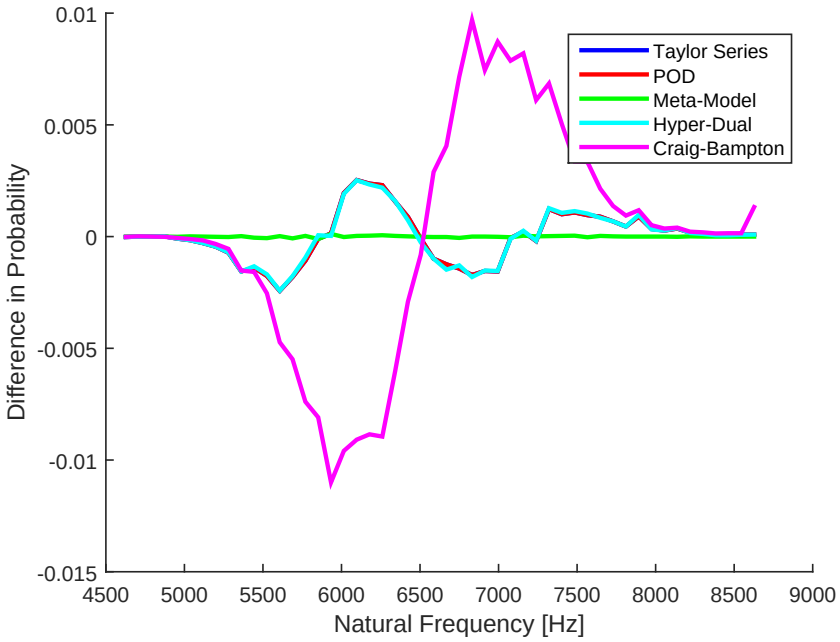

Figure 5.5. Difference in Output Distributions for Each Method

actual values are repetitive and are not reported.

The next quantitative analysis is the difference of the statistical moments. For this analysis the first four statistical moments are compared, which are related to the mean, variance, skewness, and kurtosis. The error is taken as the percent difference from the truth model. The error presented in Table 5.4 are root mean squared through all fourteen natural frequencies and all four moments.

Table 5.4. RMS Error for the Statistical Moments for all Natural Frequencies [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.685	0.685	0.690	0.009	1.909
2	0.032	0.032	0.030	0.000	1.933
3	0.021	0.021	0.014	0.000	1.944
4	0.001	0.001	NA	NA	2.037

The results in Table 5.4 show that every method used has relatively small error. Craig-Bampton representation had the largest error of 2% which is still relatively small. POD and Taylor again result in the exact same results but the time difference between the two methods are significant. The Hyper-Dual and Taylor series produce very similar results but

the Hyper-Dual only required one high-fidelity code evaluations. The Meta-Model gives the best results out of all the methods and gives the results in the least amount of time.

Distribution Determination using Reduced Sampling

The previous section utilized $100,000$ sample in a traditional MC analysis. That large of a sample based is not usually available for more difficult or complex problems even with the use of ROMs. Some ROMs can still take a large amount of time depending on what type of ROM is used. Traditionally, a reduced sample size is required due to time and computation restrictions. Two methods are examined to test the accuracy of each method compared to the full MC analysis: SROM and Latin Hypercube. The SROM technique is presented in [6, 7]. This method takes a subset of the large sampling that matches the statistical moments by using different probability at each data-point. The Latin Hypercube sampling technique is presented in [15]. Latin Hypercube sampling is essentially dividing the distribution into equal probable sections then selects a random point in each section to use for the sampling. For each sampling technique, two different number of reduction points are used: 50 and 100 data-points. This is a dramatic reduction from $100,000$ but shows the effectiveness of these different sampling techniques.

Latin Hypercube Sampling

The Latin Hypercube sampling technique splits the input domains into equally probable sections and uses a data-point from each section. Latin Hypercube is implemented in MatLab for a normal distribution, but not a log-normal distribution. A custom function is created to use this sampling technique for a log-normal distribution. Latin Hypercube sampling can become more complicated to implement in multiple dimensions but this example problem is only using a univariate distribution of the Young's modulus. Using this method requires more memory to store the data-points than was required with the traditional MC analysis, which selects each point at random. This memory increase is typically acceptable since the required number of data-points is small, less than 0.1% in this example.

The first case of this sampling technique is the results for 50 data-points. Qualitative results are shown in Figures 5.6-5.7. This shows that the Meta-Model and the truth data are almost identical. Figure 5.6 shows the distribution of the fourteenth natural frequency. The point-by-point difference between the distributions is shown in Figure 5.7.

These distributions only compare the methods to each other and do not show to ability to accurately predict the full MC analysis. The error of the moments are shown in Table 5.5. This shows the RMS error compared to the full MC analysis statistical moments. The first interesting trend is that the Meta-Model gives very accurate results even in the first order approximation. The second trend is how bad the Craig-Bampton is in a reduced framework. This is partially due to the inaccuracy of the Craig-Bampton on an individual

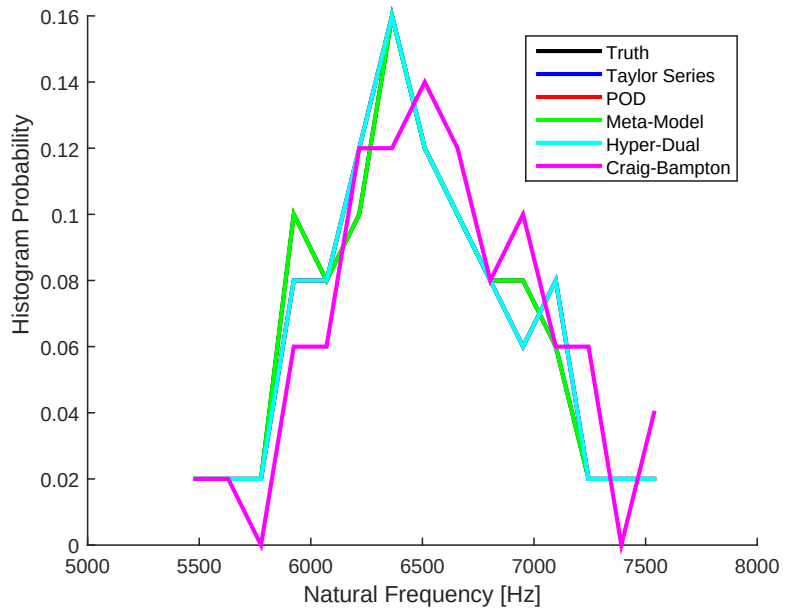

Figure 5.6. Output Distributions for Each Method with 50 Samples LHS

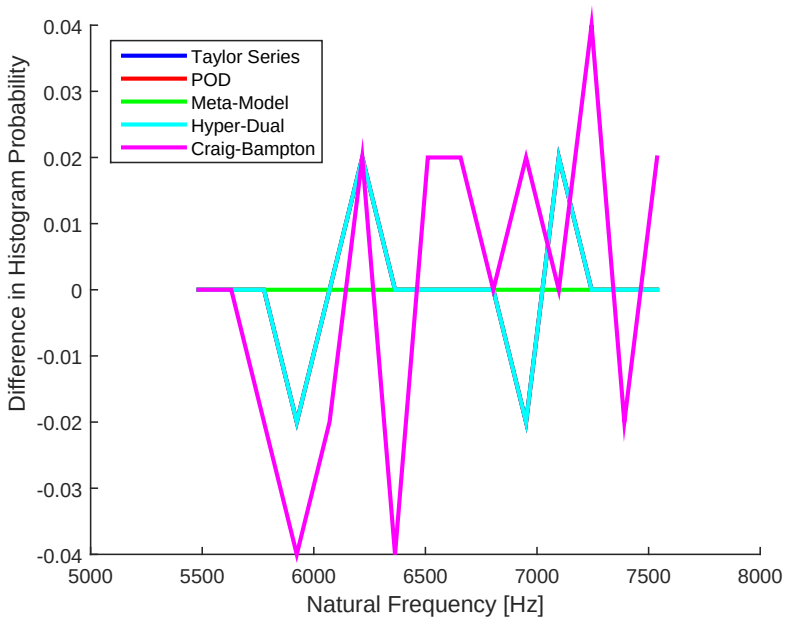

Figure 5.7. Difference in Output Distributions for Each Method with 50 Samples LHS

code evaluation. The inaccuracy is also noted in Figure 5.4 with the shift in the peak.

Table 5.5. RMS Error for the Statistical Moments for all Natural Frequencies with 50 Samples LHS [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.601	0.601	0.605	0.061	20.334
2	0.095	0.095	0.094	0.069	1.774
3	0.051	0.051	0.057	0.069	25.688
4	0.068	0.068	NA	NA	39.710

While the accuracy of using 50 samples are very decent, the increase in number of samples can increase the accuracy. The accuracy of using 100 samples is also determined. These distributions can be seen in Figures 5.8-5.9. They show the same trends as the 50 sample results.

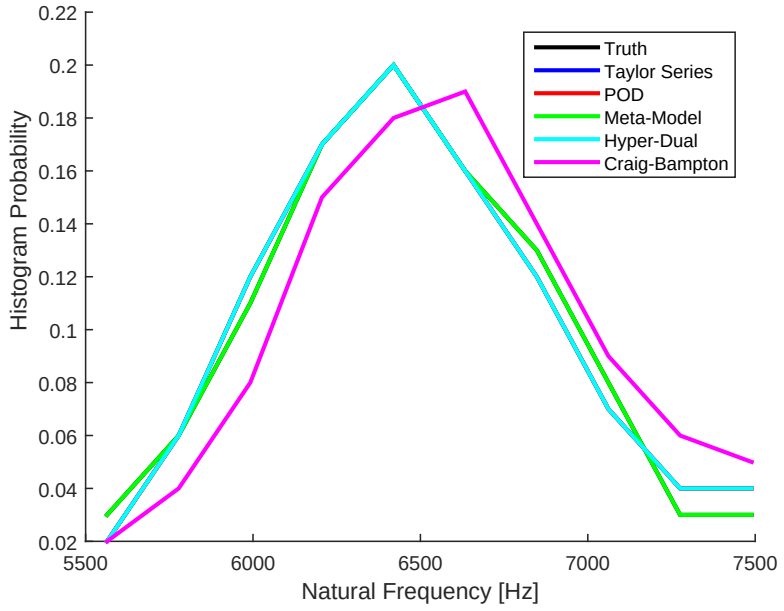

Figure 5.8. Output Distributions for Each Method with 100 Samples LHS

The more interesting aspect is the accuracy increase for 100 samples. These results are presented in Table 5.6. The first comment is the error for almost all the methods decrease at most orders. An exception is the Craig-Bampton method that had a decrease in error for first order but an increase in error for the other orders. The next trend is that the double in samples does not correspond to a halving of the error. The Taylor series, POD, and Hyper-Dual methods only have a small change in the error.

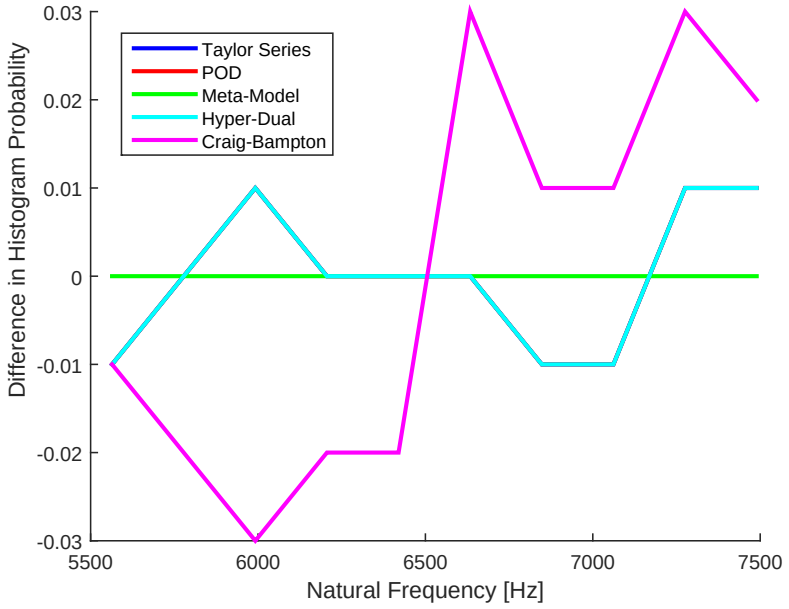

Figure 5.9. Difference in Output Distributions for Each Method with 100 Samples LHS

Table 5.6. RMS Error for the Statistical Moments for all Natural Frequencies with 100 Samples LHS [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.599	0.599	0.603	0.052	14.772
2	0.086	0.086	0.085	0.060	5.879
3	0.045	0.045	0.050	0.060	28.814
4	0.059	0.059	NA	NA	166.520

SROM Sampling

The other technique of reducing the required sampling number is the SROM technique. This method uses a subset of a large sample size and selects the optimal subset to match the first six statistical moments of the large sample. One of the main keys to this method is the use of non-equal probabilities to better represent the distribution curve. Another key is the flexibility of the method since it is defined by minimizing an objective function that the user defines based on the important information for the problem, such as statistical moments. This technique is used on the log-normal distribution of the Young's modulus to be used in the propagation of uncertainty.

The first results are for 50 samples in the distribution. The distributions are shown in Figures 5.10-5.11. These distributions follow the same trend as the Latin Hypercube method, where the Meta-Model is almost perfect with the true value while the Taylor series, POD, and Hyper-Dual methods all have close to the true value but noticeable error. The Craig-Bampton method also has a very large error.

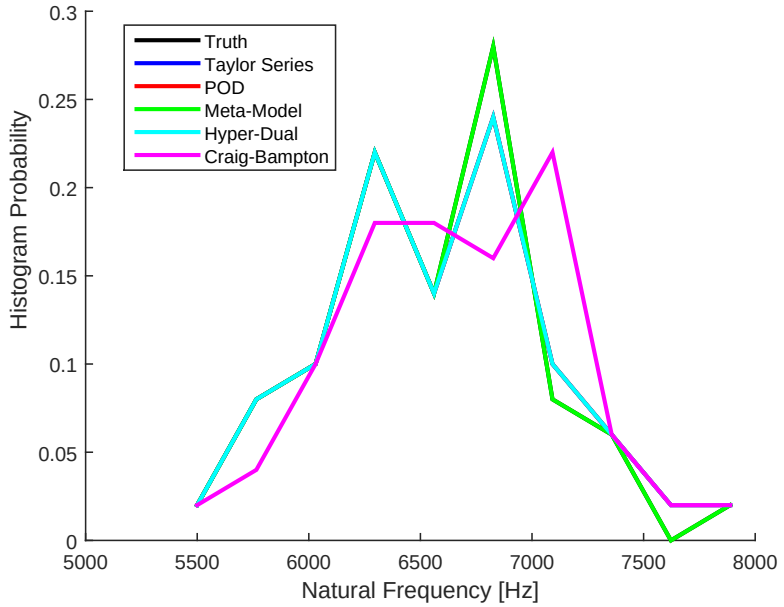

Figure 5.10. Output Distributions for Each Method with 50 Samples SROM

The statistical moments of the distribution is presented in Table 5.7. For the Taylor series, POD, and Hyper-Dual methods, the 50 samples techniques are close to the same for both Latin Hypercube and SROM. For a first order approximation, the Latin Hypercube gives a better representation, and for higher order approximation, the SROM gives less error. One interesting trend is the Meta-Model which has consistency less error in the SROM technique compared to the Latin Hypercube technique. The Craig-Bampton method has large error.

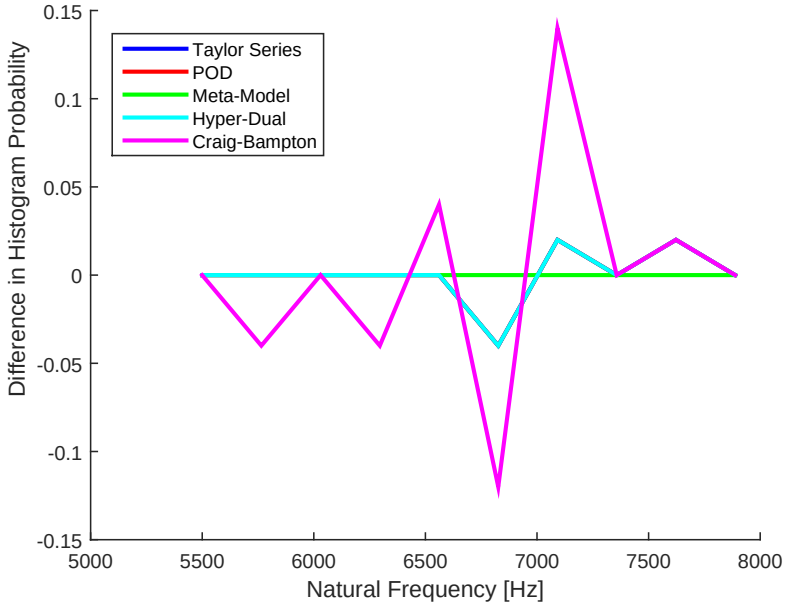

Figure 5.11. Difference in Output Distributions for Each Method with 50 Samples SROM

This error is expected to decrease with the addition of more fixed-interface modes.

Table 5.7. RMS Error for the Statistical Moments for all Natural Frequencies with 50 Samples SROM [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.626	0.626	0.630	0.039	1.789
2	0.078	0.078	0.077	0.047	13.778
3	0.027	0.027	0.034	0.047	60.779
4	0.046	0.046	NA	NA	18.679

The next result is to increase the number of samples to 100. The distributions are presented in Figures 5.12-5.13. These figures show the same trend as the previous ones. The only noticeable difference in these distribution from the others are around 7500 [Hz] where the truth data is not overlay with another curve.

The comparison between these distributions and the true distribution is presented in Table 5.8. One interesting trend compared to 50 samples is that the Taylor series, POD, and Hyper-Dual actually had an increase in error for the first order approximation. The higher order approximations were decreased for those three methods. Another trend that is followed is that the Meta-Model increased in accuracy with the increase in samples. The Craig-Bampton method effectively failed this simulation.

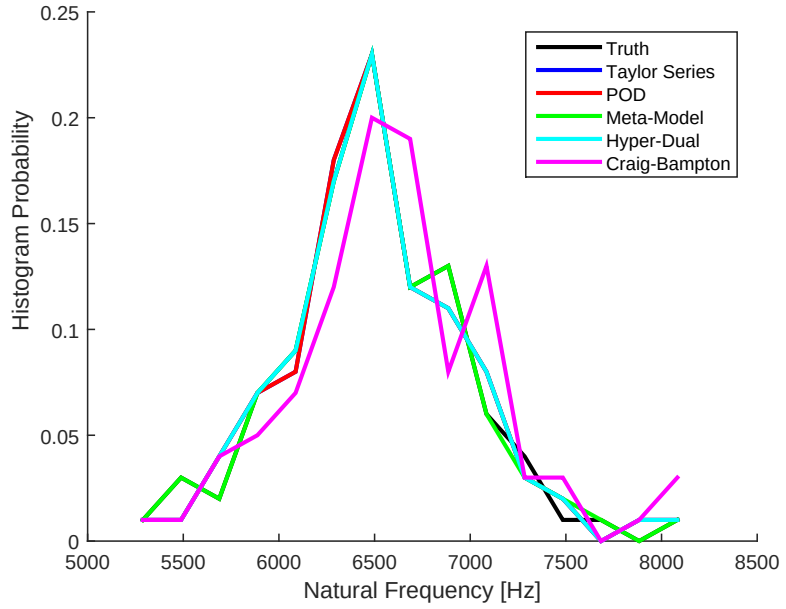

Figure 5.12. Output Distributions for Each Method with 100 Samples SROM

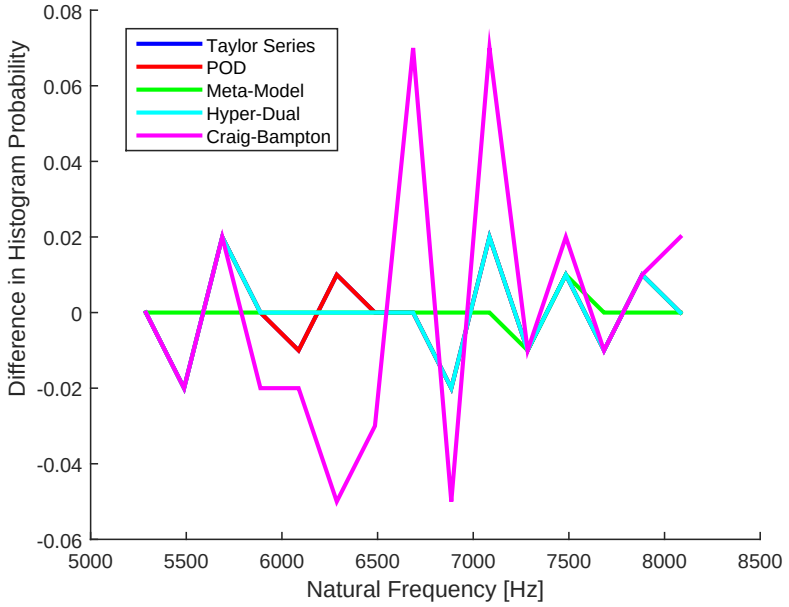

Figure 5.13. Difference in Output Distributions for Each Method with 100 Samples SROM

Table 5.8. RMS Error for the Statistical Moments for all Natural Frequencies with 100 Samples SROM [%]

Order	Taylor	POD	Hyper-Dual	Meta-Model	CB
1	0.649	0.649	0.653	0.022	319.05
2	0.062	0.062	0.060	0.030	429.42
3	0.010	0.010	0.017	0.030	56.560
4	0.029	0.029	NA	NA	36.730

The next comparison is between Latin Hypercube and SROM with 100 samples each. POD, Taylor series, and Hyper-Dual have relatively the same accuracy for the first order approximation with slightly higher error for the SROM. The higher order approximations were consistently more accurate for the SROM technique. Meta-Model method while using the SROM technique has approximately half the error for all order of approximations. The Craig-Bampton model was not very accurate for any of these simulations.

While the distributions are a good measure of accuracy, the upper fractile is a more useful for reliability engineers. In this analysis, the upper fractile is defined as

$$x \ni Pr(X \leq x) = 0.1\%, \tag{5.1}$$

where x is the location of the upper fractile. The probability is calculated based on a histogram of the sampled data. This value is presented in Table 5.9. The reduced sampling techniques are not considered since there are only 50 samples used in the distribution, which does not have enough resolution to measure 0.1% probability.

Table 5.9. Fundamental Frequency Value of the Upper Fractile Location

	Frequency [Hz]
True Data	295.24
Taylor Series	302.47
POD	302.47
Hyper-Dual	302.38
Meta-Model	295.29
Craig-Bampton	295.24

The data in Table 5.9 shows a very similar trend as the parameter sweep, where the Meta-Model gives very high accuracy and the Taylor series proves a decent accuracy. The trend also follows that the Craig-Bampton model is the most accurate for the fundamental frequency. This accuracy decreases for higher frequencies but is still more accurate than Taylor series, POD, and Hyper-Dual.

Chapter 6

Conclusions

The methods discussed in this report are: Taylor series with finite difference, Proper orthogonal decomposition, Craig-Bampton reduction, Hyper-Dual formulation, and a Meta-Model method. Each method provided has specific advantages and disadvantages. The Taylor series is straight forward but can be computationally expensive. POD reduces the size of the Taylor series but requires singular value decomposition of the output data and still uses a finite difference. Craig-Bampton representation is a physical reduction model but still requires a full code evaluation on the reduced system. Hyper-Dual method provides the same accuracy with only a single high-fidelity code evaluation but requires special programming in order to extract the extra information. The Meta-Model method fits a polynomial curve to the output of multiple Hyper-Dual calculations but the determination of the unknown coefficients can become computationally infeasible.

The Craig-Bampton reduction provided the least accurate results and required the most computation time to evaluate. This method, however, is the most feasible if the output is not completely known. That is to say, if the analysis technique is not quite known or decided, the Craig-Bampton reduction doesn't perform any analysis, just reduces the size of the model. It is able to evaluate new information without a complete redefinition although the errors might become high. Craig-Bampton is also very accurate is predicting the upper fracture of the distribution for all frequencies, although, it is most accurate at the low natural frequencies. The rest of the models were based purely on a numerical approximation and contains no physical information while the Craig-Bampton contains all the physical information. Each model is able to accurately predict both the parametric sweep curve along with the output distribution by using the statistical moments and the upper fractile as an error measurement.

The results in this report are for this one example system as an illustrative example. Each system is different and the results will very, but the general trends will be followed. The selection of the reduction model is dependent on the computational ability. If only a single code evaluation is possible, the Hyper-Dual method provides very accurate results if the code if able to evaluate Hyper-Dual numbers. If only a few evaluations are available and the code is able to evaluate Hyper-Dual numbers, the Meta-Model method is very good since it can provide accurate results and very low computational cost once the unknown coefficients are determined. It is important to select points near the extremes of the input parameters to not have numerical issues for large difference in values compared to the data-points used in the derivation. If no high-fidelity code evaluations are possible, a Craig-Bampton reduction

is very useful. This method does not require a full high-fidelity code evaluation and provides a reasonable error measure. If many code evaluations are available, POD or Taylor series can be used. Both of these require multiple high-fidelity code evaluations in order to provide high accuracy results. The POD method is preferred if the output data is manageable in a matrix since it requires less computational power to produce the distribution and uses less memory for storage.

The determination of the output distribution using a large Monte-Carlo analysis can become very computationally expensive. Two techniques for reducing the number of required samples are evaluated and compared: Latin Hypercube sampling and SROM sampling. Both techniques are able to represent the output distribution moments accurately while using a very small fraction of the number of samples. As a general trend the SROM sampling produced less error in the calculations compared to Latin Hypercube.

References

[1] M.R.W. Brake, J.A. Fike, and S.D. Topping. Parameterized reduced order models from a single mesh using hyper-dual numbers. *Journal of Sound and Vibration*, under review.

[2] M.R.W. Brake, P. Reuss, C.W. Schwigshackl, L. Salles, M.E. Negus, D.E. Peebles, R.L. Mayes, M.S. Bilbao-Ludena, J.C. Bonney, S. Catalfamo, C. Gastaldi, J. Gross, R.M. Lacayo, B.A. Robertson, S. Smith, C. Swacek, and M. Tiedemann. The 2014 Sandia nonlinear mechanics and dynamics summer research institute. Technical Report SAND2015-1876, Sandia National Laboratories Livermore, CA; Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2015.

[3] M.R.W Brake, P. Reuss, D.J. Segalman, and L. Gaul. Variability and repeatability of jointed structures with frictional interfaces. In *IMAC XXXII A Conference and Exposition on Structural Dyanimcs*, February 2014.

[4] Anindya Chatterjee. An introduction to the proper orthogonal decomposition. *Current science*, 78(7):808–817, 2000.

[5] R Craig and M Bampton. Coupling of substructures for dynamic analysis. *AIAA Journal*, 6(7):1313–1319, 1968.

[6] John M Emery, Richard V Field, James W Foulk, and Mircea Dan Grigoriu. Stochastic reduced-order models for multi-scale simulation of laser weld failure. Technical report, Sandia National Laboratories Livermore, CA; Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2013.

[7] John M Emery, Richard V Field, James W Foulk, Kyle N Karlson, and Mircea D Grigoriu. Predicting laser weld reliability with stochastic reduced-order models. *International Journal for Numerical Methods in Engineering*, 2015.

[8] RV Field, M Grigoriu, and JM Emery. On the efficacy of stochastic collocation, stochastic galerkin, and stochastic reduced order models for solving stochastic problems. *Probabilistic Engineering Mechanics*, 2015.

[9] J. A. Fike and J. J. Alonso. The development of hyper-dual numbers for exact second-derivative calculations. In *AIAA paper 2011-886, 49th AIAA Aerospace Sciences Meeting*, 2011.

[10] J. A. Fike and J. J. Alonso. Automatic differentiation through the use of hyper-dual numbers for second derivatives. In S. Forth, P. Hovland, E. Phipps, J. Utke, A. Walther, T. J. Barth, M. Griebel, D. E. Keyes, R. M. Nieminen, D. Roose, and T. Schlick, editors, *Recent Advances in Algorithmic Differentiation*, volume 87 of *Lecture Notes*

in Computational Science and Engineering, pages 163–173. Springer Berlin Heidelberg, 2012.

[11] Jeffrey A. Fike and Matthew R.W. Brake. Parameterized reduced-order models using hyper-dual numbers. Technical report, Sandia National Laboratories Livermore, CA; Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2013.

[12] B Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computations*, 51:699–706, 1988.

[13] S.W. Gilbert. Perturbation theory for the singular value decomposition. 1998.

[14] Daniel J. Rixen. A dual craigbampton method for dynamic substructuring. *Journal of Computational and Applied Mathematics*, 168(12):383 – 391, 2004. Selected Papers from the Second International Conference on Advanced Computational Methods in Engineering (ACOMEN 2002).

[15] Michael Stein. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.

DISTRIBUTION:

1 msbonney@wisc.edu
1 kammer@engr.wisc.edu

1	MS 0346	Richard Field, 1523
1	MS 0346	Diane Peebles, 1526
1	MS 0346	Brett Robertson, 1526
1	MS 0557	C. Dennis Croessmann, 1520
1	MS 0840	Joe Bishop, 1554
1	MS 0840	John Emery, 1526
1	MS 1070	Matthew Brake, 1526
1	MS 1070	Jeffrey Fike, 1526
		,
1	MS 0899	Technical Library, 9536 (electronic copy)