

A BAYES NETWORK APPROACH TO UNCERTAINTY QUANTIFICATION IN HIERARCHICALLY DEVELOPED COMPUTATIONAL MODELS

Angel Urbina,^{1,*} Sankaran Mahadevan,² & Thomas L. Paez³

¹Sandia National Laboratories, Uncertainty Quantification and Model Validation Dept., P.O. Box 5800, Mail Stop 0828, Albuquerque, NM 87185-0828

²Vanderbilt University, Nashville, TN 37235

³Thomas Paez Consulting, Durango, CO 80301

Original Manuscript Submitted: 07/12/2011; Final Draft Received: 11/9/2011

Performance assessment of complex systems is ideally accomplished through system-level testing, but because they are expensive, such tests are seldom performed. On the other hand, for economic reasons, data from tests on individual components that are parts of complex systems are more readily available. The lack of system-level data leads to a need to build computational models of systems and use them for performance prediction in lieu of experiments. Because their complexity, models are sometimes built in a hierarchical manner, starting with simple components, progressing to collections of components, and finally, to the full system. Quantification of uncertainty in the predicted response of a system model is required in order to establish confidence in the representation of actual system behavior. This paper proposes a framework for the complex, but very practical problem of quantification of uncertainty in system-level model predictions. It is based on Bayes networks and uses the available data at multiple levels of complexity (i.e., components, subsystem, etc.). Because epistemic sources of uncertainty were shown to be secondary, in this application, aleatoric only uncertainty is included in the present uncertainty quantification. An example showing application of the techniques to uncertainty quantification of measures of response of a real, complex aerospace system is included.

KEY WORDS: *uncertainty quantification, Markov chain Monte Carlo, Bayesian inference, hierarchical model development, structural dynamics*

1. INTRODUCTION

The reliability of high-consequence systems, such as aerospace vehicles, has been traditionally established by testing individual systems and verifying that their performance is within some acceptable limits. Although full-scale testing may not be feasible for some large, complex systems under actual use environments, some limited testing is often available for components, subsystems (i.e., groups of components), and a very limited number of tests of the full system in other use environments. Modeling and simulation of a complex system attempts to fill the gap left by the lack of full-scale testing in the actual use environments. Because component level data are usually cheaper and easier to obtain relative to the system data, it is advantageous to use these data when available. This leads to a hierarchical approach for building system-level models. As a result, the uncertainty in the system model is affected by the component level data and by the knowledge not captured in the component- or subsystem-level data. Furthermore, because tests cannot be performed for many actual use environments, the model is required to extrapolate beyond the data from which it was developed. To establish confidence in an extrapolated model prediction, sources of uncertainty must be identified, quantified, and propagated to the response quantity of interest at the system model. Sources of uncertainty can arise

*Correspond to Angel Urbina, E-mail: aurbina@sandia.gov

from (i) physical variability, (ii) the coupling of various components, (iii) data collection, (iv) modeling assumptions, and (v) model approximations. This paper develops a method to treat a difficult but very practical problem, namely, the quantification of uncertainty in full-scale systems based on experimental data from components and subsystems. Particular attention is placed on aleatoric uncertainty, which is also referred to irreducible uncertainty. Typically, the largest sources of aleatoric uncertainty are part-to-part and test-to-test variability.

At Sandia National Laboratories, there has been an emphasis on developing models of components using first principles, calibrating them from simple exploratory experiments, validating them relative to a different set of experiments, and then using them within a more complex model. For example, one could investigate the behavior of mechanical joints using simple experiments, develop a model that explains the behavior, validate its performance based on a laboratory environment, and use it as part of a larger system. Similar steps could be taken for other components, which then will aggregate to form a full system. This is a hierarchical or building-block approach to building a system-level model by incorporating simpler component models and coupling them together. This model building approach was described in several published papers [1, 2] and has been implemented in [3] for the example problem used in this paper. To quantify the uncertainty in the prediction of individual component behavior, multiple tests of these components should be available from which an estimate of the variability in the components could be obtained. Additional contribution to the uncertainty in system-level model prediction is the possibility that the interactions of the various components were never tested; thus, no information on the coupling of components will be available. Or, interactions between components could have been tested at excitation levels that are not comparable to those of the full system. In the implementation described in [3], only information available at the most fundamental level was used to estimate the parameters of the model and data at other levels were used to assess the model's predictive capability for the intended use environment [4, 5] (i.e., model validation). In this paper, we propose using all the available data at different levels to estimate the parameters of a model. By using a Bayes network, it is possible to incorporate all the available data for various components and subsystems and update the model parameters. This concept has been explored in [6] for a system consisting of one type of component. Our paper extends that concept to problems with multiple levels and multiple types of components, and adds the very substantial complications of working with actual experimental data at multiple levels and complex finite element models.

In this paper, the aim is to quantify and propagate uncertainty to the system-level response by including all available data at every level. To demonstrate the methodology developed in this paper, a practical, physical problem developed at Sandia National Laboratories (hence forth, referred to as Sandia) is used, which is described in Section 2. Section 3 gives a description of a Bayesian approach to uncertainty quantification and its implementation to the hierarchical model development problem. Section 4 presents numerical results, and Section 5 summarizes the work.

2. DEMONSTRATION PROBLEM

The complex, physical example problem shown in Fig. 1 was chosen to examine the effect of sources of uncertainty and the data and models from this will be used to develop a Bayes network representation of the full system. This problem framework was created for the purpose of implementing the rigorous uncertainty quantification and model validation methodology developed at Sandia. The problem has the following features:

1. It is a multicomponent problem, where one branch of the problem involves a mechanical joint and the other an encapsulating foam. Both are energy-dissipating mechanisms.
2. It is a multilevel problem, where the phenomena observed at the lowest level is assumed to be present at subsequent levels (e.g., damping in the joints and foam is assumed similar at all levels).
3. The individual component branches converge to a system-level hardware, where the two couple together.
4. Experimental data consist of repeated tests on several, nominally identical hardware systems at the subcomponent and component levels. These are intended to quantify the variability (aleatoric uncertainty) inherent in a physical system.

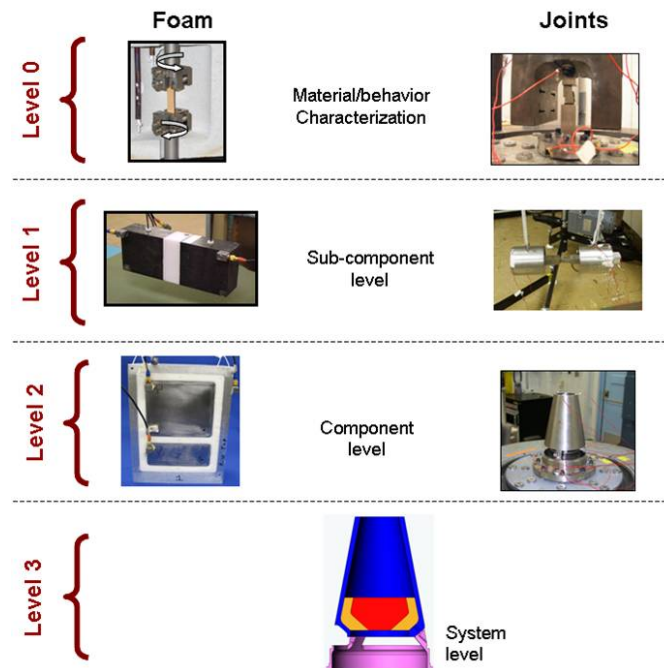


FIG. 1: Pictorial of the example problem.

5. Finite element models are built and calibrated to simulate a particular behavior of the physical hardware. The model parameters have been calibrated from simple, discovery experiments (shown as level 0) aimed at isolating the particular physical phenomenon that the model is trying to represent.

Additional details on the available data and models used are given in the following subsections.

2.1 Available Experimental Data and Models

In this section, the experimental data and models available for this example are briefly described. The basis of on a careful analysis of the experimental data collected, it was determined that no significant systematic or bias errors are present. In addition, it was also determined that measurement errors are much smaller than the uncertainty due to hardware variation; therefore, uncertainty due to experimental measurement is not treated in this study. It is acknowledged that this source of uncertainty should be included in the uncertainty analysis if preliminary analysis shows a significant contribution.

Also described in this section are the various models used in this work. These models range from simple algebraic expression to complex, three-dimensional finite element models comprised of millions of degrees of freedom. Although different types of models to represent the physics of interest were evaluated in this study, uncertainty due to model form is not explicitly included in this work. This is not to say that these are not important but due to time constraints, they were not included. In order to perform uncertainty quantification, one implementation issue is the need to perform many evaluations of the full finite element model in the most efficient manner. A candidate approach is the use of surrogate models, such as a Gaussian process (GP) model. GP models provide a simple and efficient approach to map the small number of parameters to the response quantities of interest. One advantage of a GP model is that it is a nonparametric modeling technique that avoids the need for a functional relationship among data to be known in advance. This provides much flexibility to establishing a relationship between a set of inputs and outputs, but it also creates a limitation as to the applicability of such a model. In general, as long as the GP model is used in a regime (i.e., a set of inputs) that does not extend too far outside the boundaries within which the GP model was created, the interpolations will generally be good. A GP model will not normally produce accurate extrapolations.

The GP model has another significant feature of interest in this research in that it provides a direct representation of the uncertainty associated with its interpolative approximation. As noted in [7], GP modeling is quite powerful but there is a steep learning curve needed to obtain a working understanding of the methodology, and the implementation can lead to erroneous conclusions if the parameters of the model are not selected carefully. A brief discussion of the theory behind GP models is described next.

Consider that one wants to build an approximation to a function of a vector-valued input X , based only on m observations of the inputs and outputs: $Y(x(1)), \dots, Y(x(m))$. As noted above, an appropriate selection of these input/output pairs is critical to the performance of the GP model. The basic idea of the GP interpolation model is that the outputs, Y , are modeled as a Gaussian process that is indexed by the inputs, x . A Gaussian process is simply a set of random variables such that any finite subset has a multivariate Gaussian distribution. A Gaussian process is defined by its mean function and covariance function, which in this case are functions of X . Once the Gaussian process is observed at m locations $x(1), \dots, x(m)$, the conditional distribution of the process can be computed at any new location, x^* , which provides both an expected value and variance (uncertainty) of the underlying function. The key here is that the function describing the covariance among the outputs, Y , is a function of the inputs, X . The covariance function is constructed such that the covariance between two outputs is large when the corresponding inputs are close together, and the covariance between two outputs is small when the corresponding inputs are far apart. As shown below, the conditional expected value of $Y(x^*)$ is a linear combination of the observed outputs, $Y(x(1)), \dots, Y(x(m))$, in which the weights depend on how close x^* is to each of $x(1), \dots, x(m)$. In addition, the conditional variance (uncertainty) of $Y(x^*)$ is small if x^* is close to the training points and large if it is not. Furthermore, the GP model may incorporate a systematic, parametric trend function whose purpose is to capture large-scale variations. This trend function can be, for example, a linear or quadratic regression of the training points. It turns out that this trend function is actually the (unconditional) mean function of the Gaussian process. The effect of the mean function on predictions that interpolate the training data tends to be small, but when the model is used for extrapolation, the predictions will follow the mean function very closely as soon as the correlations with the training data become negligible. Additional references to GP models can be obtained from various published works [8–11].

2.1.1 Level 0: Material/Behavior Characterization

Foam

The first step in the process is to characterize the material model that describes the behavior of epoxy-based foam with a nominal density of 20 pcf (pounds per cubic foot) and also to quantify the variability in the material. This step is represented graphically in Fig. 1 (Level 0/Foam). The material density of the foam samples was estimated through physical measurements and the elastic and shear modulus were estimated using tension/compression and torsional experiments. The elastic modulus versus density data from those experiments is shown in Fig. 2.

A model describing the relationship between density and modulus of elasticity was chosen from Gibson and Ashby [12], and it is given as follows:

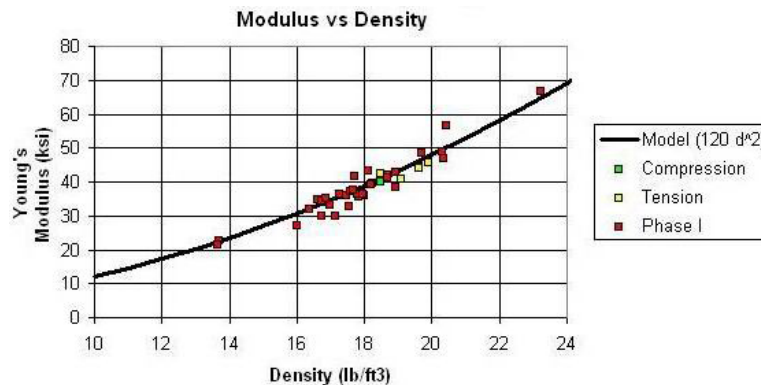


FIG. 2: Calibration data for foam modulus of elasticity.

$$E = c_1 * (\rho)^{\text{pow}}, \quad (1)$$

The model parameters c_1 and pow were calibrated through regression using the data in Fig. 2, and estimated as $c_1 = 120$ and $\text{pow} = 2.0$.

Joints

For the joints, the hardware configuration used to characterize the energy dissipation phenomenon across an interface is shown in Fig. 1 (Level 0/Joints). Experiments consisted of sine sweeps (excited via an electrodynamic shaker) at five load levels (100, 200, 300, 400, and 500 lbs) and acceleration time histories were recorded. The variability in the data arising from sample-to-sample variations was quantified from a total of nine combinations of three top and three bottom pieces of a bolted connection. In addition, the variation in the data due to testing conditions was also obtained by assembling each hardware configuration, testing it, disassembling, and repeating the process again. This was done five times for each of the nine hardware combinations in order to quantify test-to-test variability. Forty-five experimental configurations were obtained. Starting from acceleration time histories recorded from each experiment, the procedure shown in Fig. 3 is used to calculate the energy dissipation as a function of force [13]. These are shown in Fig. 4.

Using the data shown in Fig. 4 and hysteretic-type information also obtained from experiments, a constitutive model relating the energy dissipated per cycle of response to force across the joint can be calibrated. In this case, the following model proposed by Smallwood [14] was used:

$$f_j = k_{\text{lin}}(d_j - d_i) - k_{\text{non}}(d_j - d_i)^{\text{npow}} + f_i, \quad (2)$$

where $f_{i \& j}$, force across the joint at the reversal point in the hysteresis loop (measured during experiment) and $d_{i \& j}$, relative displacement across the faces of the joint at the reversal point in the hysteresis loop (measured during experiment). The model parameters to be calibrated are k_{lin} , linear stiffness component; k_{non} , non-linear stiffness component; and npow , degree of nonlinearity. It is the slope of each experimental curve of energy dissipated, E_d versus force curve shown in Fig. 4 in log-log space. Because there are 45 repeated experiments (nine hardware combinations times five assembly/disassembly combinations), a probability distribution for each of the model parameters can be estimated.

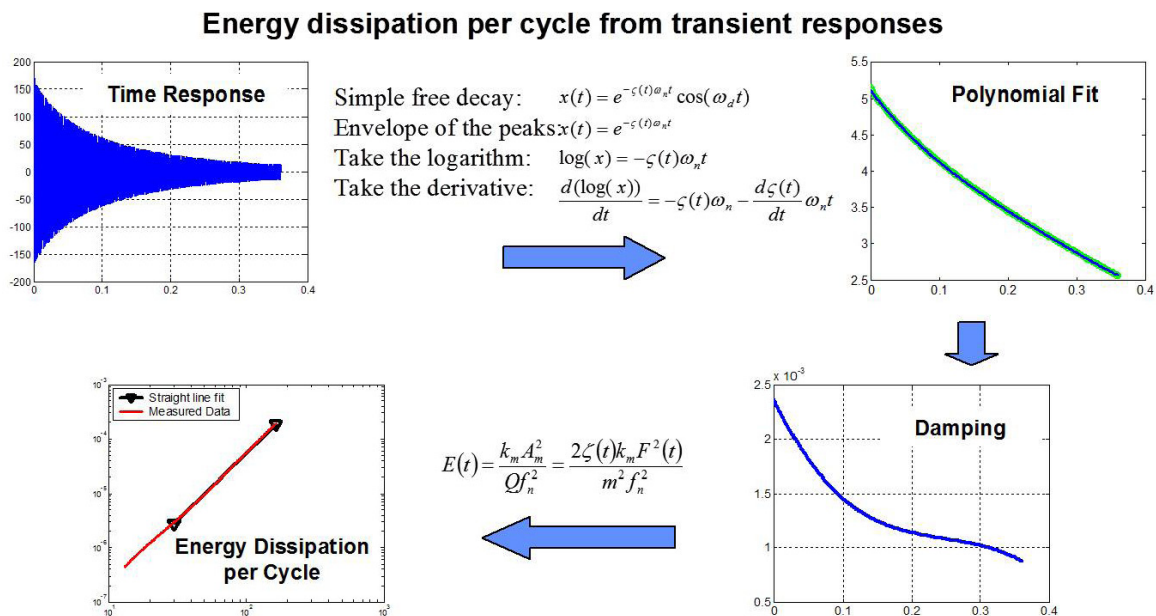


FIG. 3: Procedure to calculate energy dissipation.

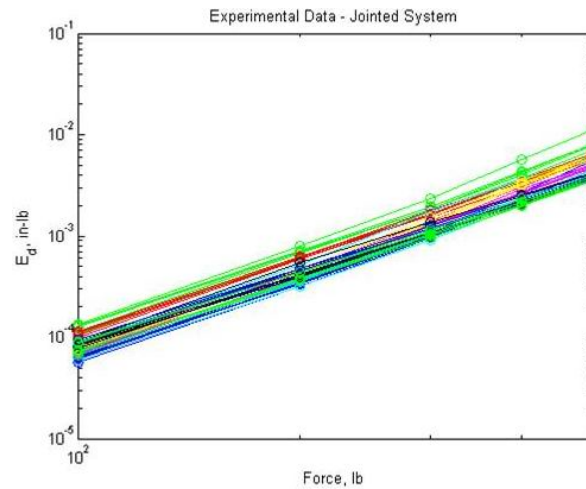


FIG. 4: Forty-five realizations of force versus energy dissipation curves.

2.1.2 Level 1: Sub-Component Level

Foam

For this level, the structure shown in Fig. 1 (level 1/foam) is an element of foam (shown as a white block in the center of the test hardware) bonded to steel end masses. Modal tests were performed on the samples. Structural response was measured in the time domain, and the fundamental frequency of the response for a particular mode of vibration was calculated. In addition, the density of the foam material was measured. Six of the nominally identical test hardware were constructed and tested.

Joints

This configuration places the single leg joint hardware used for the behavior characterization done in level 0 in a different loading configuration. Two 30 lbs masses are bolted at the ends of the single leg creating a dumbbell hardware. This is shown in Fig. 1 (level 1/joints). This configuration is then supported by bungee cords to simulate a free-free environment and is subjected to an impulse excitation provided by an instrumented hammer. The free decay acceleration time history response of the dumbbell on the side opposite to the excitation is recorded and used to estimate the energy dissipation of the system at a particular force level. A total of 45 experiments were conducted that consisted of five repetitions of the experiment for each of the nine leg configurations.

2.1.3 Level 2: Component Level

Foam

This hardware uses a set of steel masses encapsulated in foam and contained within an aluminum outer shell. The hardware was instrumented with four triaxial accelerometers in order to minimize mass loading effects and additional damping introduced by the cables. The instrumented hardware is shown in Fig. 1 (level 2/foam). Six different test specimens were fabricated in order to establish part-to-part variability. The test fixture was mounted on an electrodynamic shaker and excited with a shock waveform at levels that would not strain the foam to failure. All testing was done at room temperature and acceleration time histories at each accelerometer location were recorded. Similarly to level 1/foam, the fundamental frequency of the test hardware in the axial direction was obtained and density measurements were performed.

Joints

The experimental system is a truncated conic shell supported on legs at three approximately equidistant locations around the circumference. The support structure beneath the legs is a cylindrical shell relatively thin on its top and

transitioning into a thicker section. The conic shell is attached to the support structure via three screws, each of which passes through a hole in a thin, flat plate at the top of a leg. Three nominally identical replicates of the conic shell were fabricated, along with three nominally identical support structures. They are shown in Fig. 1 (level 2/joints). The nine combinations of shells and support structures were tested in environments generated using an electrodynamic shaker. Each of the nine test structures was excited by either an accurate realization of the decaying oscillatory shock input or a wavelet-type input. Each shell-base combination was assembled-disassembled-reassembled three times, and tested each time. The average acceleration structural responses at the tops of the shells were recorded and yielded 27 time histories, nine structures times three tests each. From this, the energy dissipated was calculated similarly to the level 1/joints data.

2.1.4 Level 3: System Level

The final step in the process merges the two parallel paths, foam and joints, into a single system and is shown in Fig. 1 (level 3). In this paper, it is assumed that there are no experimental data available and only a computational model can be used to make prediction of the system behavior of interest. A finite element model of the physical system was constructed using the Sandia-developed CUBIT meshing tool and analyzed using Salinas. The finite element model uses the Smallwood model in Eq. (2) to represent the nonlinear energy dissipation behavior of the bolted connection between the conic part and the lower assembly, and the linear-elastic model for the encapsulating foam component. All solid pieces [the conic part, the bottom piece, and the internal encapsulated mass, shown in Fig. 1 (level 3)] are made of stainless steel. The encapsulating foam is shown as the region between the top conical section's outer shell and the interior frustum in Fig. 1 (level 3) and uses the same type of foam which was used in levels 0, 1, and 2. Full adhesion is assumed between the epoxy foam and the inside of the conic section and between the epoxy foam and the encapsulated mass. The model consists of 8052, 20-node, hexagonal-type elements, which yields approximately 42×10^3 nodes in the model and was verified by doing a Richardson extrapolation on the natural frequencies. Nonlinear, transient analyses were performed using Salinas to predict structural response. An estimate of the uncertainty in the peak acceleration response of the top of the conical section is sought. Section 3 describes a Bayesian approach to integrate all the available experimental data to estimate the uncertainty at the system level for the response quantity of interest.

3. BAYESIAN APPROACH TO UNCERTAINTY QUANTIFICATION IN A HIERARCHICAL SYSTEM MODEL

3.1 Background

This paper proposes the use of a Bayes network as a tool to integrate observed data and prior knowledge within a hierarchically built system-level model. Bayesian updating propagates uncertainty through this network up to the system-level response of interest, which could be used to estimate confidence bounds for the system level prediction. The approach aims to include all the available data to update the uncertainty in the various component model parameters and propagates it to the system-level model prediction. This is a major difference in the approach taken in this research relative to previous approaches [3], where only the data at the lowest level were used to quantify uncertainty in the parameters and no further updating of the parameters was made using higher level information.

Recent use of Bayesian updating for a complex system is found in [7], where the use of Bayesian updating to calibrate the parameters of a model in a multilevel problem was investigated. In that study, data from the highest level of complexity were used to calibrate a model's parameters that described the controlling physics of the problem and predictions were made at the highest level. One of the examples deals with the same bolted connection as in the current paper and was used to demonstrate the methodology. To expedite the updating of the model parameters, Gaussian process models were used in lieu of the full finite element model. Although some runs of the finite element model were still necessary to calibrate the Gaussian process model, they numbered on the order of 10–20 versus the thousands needed for Bayesian updating, thus offering a major computational efficiency. In this paper, data from different levels of complexity, not just from the highest level, are used. This incorporates all available information, which in some cases is more plentiful at the component level than at the full system level. In other words, instead of

using only bottom-level data [3] or only top-level data [7], we propose using data at all levels, whenever available, to quantify the uncertainty in system-level prediction.

A recent text on Bayesian analysis [15] concludes that the principles of Bayesian inference are now well established, and the only issue left to address is the efficient implementation of Bayesian inference to real-world problems. This speaks to the maturity of the Bayesian methods and now the need to fully incorporate them into the application domain. Published papers [16, 17] demonstrate that Bayesian methods have made headway in research laboratories, such as Sandia and Los Alamos National Laboratories, which deal with high-consequence systems. In these studies, complex systems are analyzed via Bayesian methods and useful results are presented, but most importantly, it is shown that Bayesian methods can be computationally tractable. Although great strides have been made to make Bayesian analysis more appealing for real applications, some challenges still remain. Issues of scalability when hundreds or thousands of parameters are present is a challenge. The representation of error, such as measurement error and model form error and how they should be incorporated is another issue. In this paper, we offer a way to quantify this but the problem of how to separate the various contributors to this overall error is a subject of research. Finally, assessment of convergence of Markov chain Monte Carlo (MCMC) chains is very difficult, especially if the posteriors are multimodal. Despite these challenges, we believe that Bayesian analysis and, in particular, Bayes networks provide an ideal framework to handle problems where data are available at different levels of complexity and there is a need to quantify uncertainty given the available data.

In this paper, we propose the use of a Bayes networks to address the problem of multiple levels of complexity and incorporation of any available data for the purpose of propagating and quantifying uncertainty. At the heart of a Bayes network is establishing a relationship between the various variables involved in the model. Furthermore, a Bayes network consists of the following [18]:

1. A set of variables (nodes) and a set of directed edges (arcs) between variables
2. The variables together with the directed edges form a directed acyclic graph (DAG). DAGs do not allow circular causality.
3. To each variable B with parents A_1, A_n , there is an associated conditional probability $P(B|A_1, A_n)$.

The definition of a Bayes network does not refer to causality, and there is no requirement that the links represent causal effect. The requirement is that the nodes of interest in a Bayes network are d -separation when evidence is introduced into the network. An important consideration is that d -separation implied by the Bayes network structure hold true relative to the physical interpretation of the network. The d in d -separation stands for dependence; therefore, d -separation basically refers to the conditional independence of the nodes relative to their connectivity in the network. Thus, if two variables are d -separated relative to a set of variables V in a directed graph, then they are independent, conditional on V in all probability distributions such a graph can represent. Roughly, two variables A and B are independent conditional on V if knowledge about A gives you no extra information about B once you have knowledge of V . In other words, once you know V , A adds nothing to what you know about B [19]. For completeness sake, the formal definition of d -separation is as follows: Two distinct variables A and B are d -separated if, for all paths between A and B , there is an intermediate variable V (distinct from A and B) such that either the connection is serial or diverging and V is instantiated or the connection is converging and neither V nor any V descendants have received evidence. Instantiation of a node occurs when data (or evidence) is available for that particular node. For a more comprehensive explanation on d -separation (see [18]). With this background, we now present the implementation of a Bayes network to the demonstration problem.

3.2 Methodology

To begin, the Bayes network for the example problem depicted in Fig. 1 is constructed as shown in Fig. 5. The definition of the nodes in the Bayes network shown in Fig. 5 is given as follows:

$\theta^j = [k_{\text{lin}}, k_{\text{non}}, \text{npow}]$ —Joints parameters (stochastic node)

$\theta^f = [E]$ —Foam parameter (stochastic nodes)

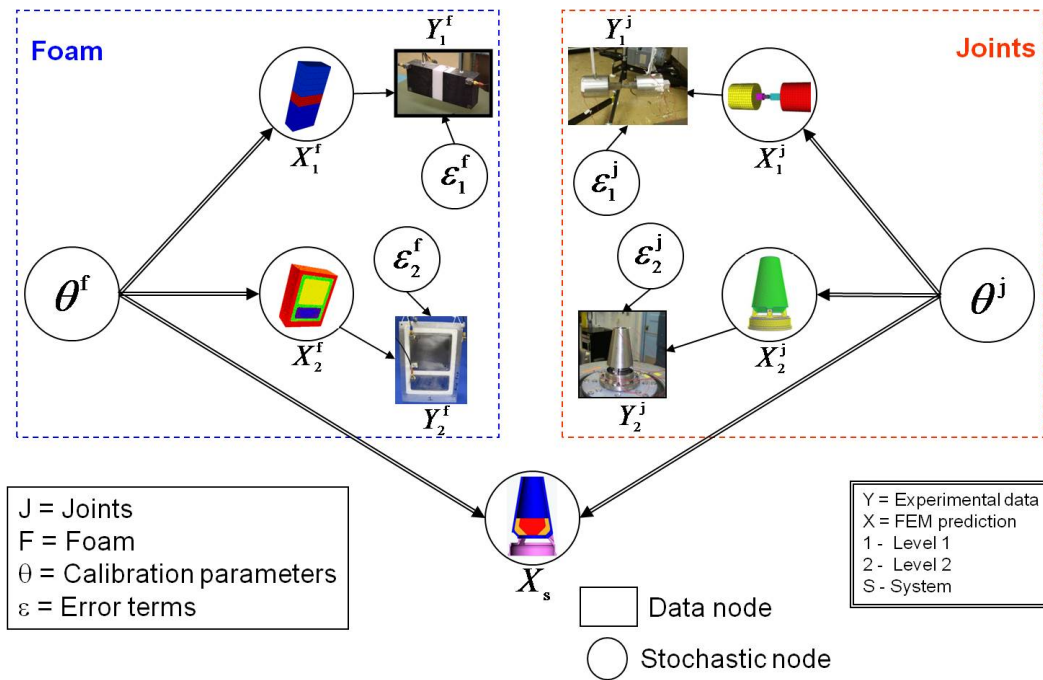


FIG. 5: Bayes network representation of example problem.

$Y_{(1\&2)}^f$ —Experimental data: modal frequencies from (1) subcomponent- and (2) component-level experiments (data nodes)

$Y_{(1\&2)}^j$ —Experimental data: energy dissipated at 300 lbs. from (1) subcomponent- and (2) component-level experiments (data nodes)

$X_{(1\&2)}^f$ —Model evaluations: modal frequencies from (1) subcomponent- and (2) component-level model simulations (logical nodes)

$X_{(1\&2)}^j$ —Model evaluations: energy dissipated at 300 lbs. from (1) subcomponent- and (2) component-level model simulations (logical nodes)

ϵ 's—Error terms associated with discrepancy between experiments and simulations at each particular level of complexity and for the joints and foam (stochastic nodes)

In the above definitions, the term “stochastic nodes” refers to nodes that are updated with data. “Data nodes” refer to nodes where data (in this case, experimental data) is introduced into the network. This data are used to update the stochastic nodes. Finally, “logical nodes” are basically nodes where a function evaluation is performed. In this example, model parameters are passed into a surrogate model and a function response is evaluated. This serves as a transfer function between the available type of data and the model parameters that are being updated. Stochastic, data, and logical nodes are terms used in the context of WinBUGS [20] which is the software used in this study.

There is a conditional probability that relates the model at each level back to a set of model parameters denoted by θ . These model parameters are updated based on the observed data at various levels. It is expected that the model and its parameters can only account for some of the behavior resulting in the observed data (i.e., no model is a perfect representation of the underlying phenomenon). To account for the remaining differences, an error term is included in the formulations. The error nodes shown in Fig. 5 are associated with the discrepancy in the model predictions relative to the observed data at each level. The formulation used in this paper is a modification of that proposed

by Kennedy and O'Hagan [21]. The formulation presented in this paper has been implemented in various practical examples [7, 16, 17] and it is represented by

$$Y = X(\theta) + \epsilon, \quad (3)$$

where Y is the experimental data, $X(\theta)$ is the model predictions, and ϵ is the error term.

This error term is assumed to have a normal distribution with zero mean and some standard deviation, which will be updated. In the context of this research, these error terms will be interpreted as sources of uncertainty whose origin is not known. Furthermore, the error terms are placeholders for uncertainty that is not related to the input parameters. These error terms include, among others, measurement errors, model representation of the actual hardware (i.e., dimensions, boundary conditions, contact surfaces, etc.), model form errors (related to the form of the model representing the physics of interest), and solution approximation errors.

A few points are important to note at this point as follows:

1. The Bayes network framework presented in this paper relies on the fact that the same model parameters are present in all the different levels of the hierarchy. For example, the foam parameter, E , is present in all the models and in all the levels and likewise for the joint parameters.
2. The data from level 0 are used to determine the prior knowledge of the joint and foam parameters. This prior knowledge comes from the work done in [3], and these are the parameters that will be updated using the data from levels 1 and 2.
3. As noted in the previous section, there are no requirement that the links in the Bayes network represent a causal effect. Only the d -separation properties need to be satisfied. By analyzing the Bayes network shown in Fig. 5, we can conclude that there are two distinct diverging branches and noting the nodes that are instantiated with evidence, it can be concluded that the proposed Bayes network does indeed satisfy the d -separation properties for the nodes labeled θ^f and θ^j . From a physical standpoint, this is exactly what we want.

In the intended use of the Bayes network, one would like to calculate the probability density function (PDF) associated with the stochastic nodes. To do this, it is known that each parent node has a PDF associated with it and each child node has a conditional probability density function, given the value of the parent node. The entire network can be represented using a joint probability density function which is given by the general expression [18]

$$P(U) = \prod_i P[X_i | \text{parent}(X_i)], \quad (4)$$

where $P(U)$ is the joint probability of the Bayes network and X_i are the i th stochastic nodes of the Bayes network. From Fig. 5 and Eq. (4), one formulates the joint probability density function of the entire network, U as

$$\begin{aligned} f(U) = & f(k_{\text{lin}}) * f(k_{\text{non}}) * f(\text{npow}) * f(E) * f(\epsilon_1^f) * f(\epsilon_2^f) * f(\epsilon_1^j) * f(\epsilon_2^j) * f(Y_1^f | X_1^f, \epsilon_1^f) * f(Y_2^f | X_2^f, \epsilon_2^f) \\ & * f(Y_1^j | X_1^j, \epsilon_1^j) * f(Y_2^j | X_1^j, \epsilon_2^j) \end{aligned} \quad (5)$$

where

$$X_1^f = g_1(E) \quad (6)$$

$$X_2^f = g_2(E) \quad (7)$$

$$X_1^j = h_1(k_{\text{lin}}, k_{\text{non}}, \text{npow}) \quad (8)$$

$$X_2^j = h_2(k_{\text{lin}}, k_{\text{non}}, \text{npow}). \quad (9)$$

In the previous equations, g_1 , g_2 , h_1 , and h_2 represent functional relationships between the random variables and the response quantities of interest. These could be finite element models or, in the case of this study, GP models.

The first line in Eq. 5 represents the prior distributions of the joint and foam parameters. These are the parent nodes in the Bayes network shown in Fig. 5, denoted as θ^j and θ^f , and have no conditional probabilities associated with them. The prior distributions of these parameters represent the current knowledge about these parameters expressed as probability distributions with statistics, mean, μ , and standard deviation, σ . The statistics of these parameters can be obtained by fitting these parameters to each of the data obtained from the level 0 joints and foam experiments. On the basis of these fits, a Normal distribution for all the parameters was deemed adequate. The prior for the foam model parameter is

$$f(E) \sim N(\mu_E, \sigma_E), \quad (10)$$

the priors for the joint model parameters are

$$\begin{aligned} f(k_{\text{lin}}) &\sim N(\mu_{\text{klin}}, \sigma_{\text{klin}}) \\ f(k_{\text{non}}) &\sim N(\mu_{\text{knon}}, \sigma_{\text{knon}}) \\ f(\text{npow}) &\sim N(\mu_{\text{npow}}, \sigma_{\text{npow}}), \end{aligned} \quad (11)$$

and the priors for the error terms are

$$\begin{aligned} f(\epsilon_1^f) &\sim N(0, \sigma_1^f) \\ f(\epsilon_2^f) &\sim N(0, \sigma_2^f) \\ f(\epsilon_1^j) &\sim N(0, \sigma_1^j) \\ f(\epsilon_2^j) &\sim N(0, \sigma_2^j). \end{aligned} \quad (12)$$

The second through fourth lines in Eq. (5) are the likelihood functions of the foam and joints, respectively. These functions relate any available data back to the parameters of the foam and joint models. A normal distribution is used for all the likelihood functions. Note that when comparing the Bayes network shown in Fig. 5 and the resulting Eq. (5), there is no term involving the node labeled X_s . Because no data are available for updating at this node, this node is a logical node and, thus, it only performs a function evaluation. In the context of uncertainty, this node represents the forward propagation of the uncertainty in the model parameters to the system-level response.

The Bayes network facilitates the inclusion of nodes that represent the observed data, and thus, the updated densities can be obtained for all the nodes. The joint probability density function for the network can be updated using Bayes theorem when data are available. The expression for Bayes theorem for this Bayes network is

$$f(k_{\text{lin}}, k_{\text{non}}, \text{npow}, E, \epsilon_{1\&2}^f, \epsilon_{1\&2}^j | Y_{1\&2}^j, Y_{1\&2}^f) = \frac{f(U)}{\int f(U) d(k_{\text{lin}}) d(k_{\text{non}}) d(\text{npow}) d(E) d(\epsilon_1^f) d(\epsilon_2^f) d(\epsilon_1^j) d(\epsilon_2^j)}, \quad (13)$$

where $f(U)$ is given by Eq. (5). The integration in the denominator of Eq. (13) can be difficult to perform especially in a complicated expression such as the one shown in Eq. (5); therefore, to get the posterior marginal distributions of each of the parameters, MCMC techniques are employed [22]. The MCMC solution for one of the parameters in the demonstration problem is shown next.

Let us assume that we want the marginal posterior distribution of the parameter, E which has a prior distribution given by Eq. (10). We will use a specific MCMC technique, called Gibbs sampling, to obtain the marginal posterior distribution of E given experimental data at levels 1 and 2. Using Eq. (13), we write the following for the parameter, E

$$f_E(E|Y) = \frac{f(U)}{\int f(U) d(E)}, \quad (14)$$

where $f(U)$ is given by Eq. (5). If we observe that the denominator in Eq. (14) evaluates to a constant (which might not be easy to perform), we can then say that the left-hand side of Eq. (14) can be known up to a proportionally constant (the integral in the denominator) and, when substituting Eq. (5) can then be written as

$$f_E(E|Y) \propto f(k_{\text{lin}}) * f(k_{\text{non}}) * f(\text{npow}) * f(E) * f(\epsilon_1^f) * f(\epsilon_2^f) * f(\epsilon_1^j) * f(\epsilon_2^j) * f(Y_1^f | g_1(E), \epsilon_1^f) \\ * f(Y_2^f | g_2(E), \epsilon_2^f) * f(Y_1^j | h_1(k_{\text{lin}}, k_{\text{non}}, \text{npow}), \epsilon_1^j) * f(Y_2^j | h_2(k_{\text{lin}}, k_{\text{non}}, \text{npow}), \epsilon_2^j). \quad (15)$$

To implement Gibbs sampling for multivariable problems, we sample each variable from the distribution of that variable conditioned on all other variables, making use of the most recent values and updating the variable with its new value as soon as it has been sampled. The procedure to sample the i th iteration of the variable E is

1. Pick an initial value for $k_{\text{lin}}, k_{\text{non}}, \text{npow}, \epsilon_1^f, \epsilon_2^f, \epsilon_1^j, \epsilon_2^j$
2. For each sample i , sample from the univariate conditional distribution

$$E^{(i)} \sim f(E^{(i)} | k_{\text{lin}}^{(i)}, k_{\text{non}}^{(i)}, \text{npow}^{(i)}, \epsilon_1^{f(i)}, \epsilon_2^{f(i)}, \epsilon_1^{j(i)}, \epsilon_2^{j(i)}), \quad (16)$$

Because the conditional distribution of one variable given all others is proportional to the joint distribution, Eq. (15) can be used for sampling. For the random variable E , the other variables become constants. A similar procedure is then repeated in a componentwise fashion for each of the parameters that need to be updated. For this research, a Gibbs sample for all the variables was found using the software WinBUGS [20], a free Bayesian inference software.

To simplify and expedite the implementation of the Bayes network for the sample problem, GP models were developed and used in lieu of the full three-dimensional finite element models that represent the hardware at each of the levels. The GP software used in this study was coded by McFarland and Bichon [7, 23]. One of the main tasks to implement this was to make code written in MATLAB and Fortran available to WinBUGS, which does not allow external calls to other programs; therefore, the GP code was implemented as a function written in Component Pascal [24] and compiled directly into the WinBUGS software. This allowed the software to make the necessary function evaluations which related the model prediction to the calibration parameters and allowed for the updating of the parameters within a Bayesian framework. The GP predictions relative to the training data was evaluated and found to be within 5% and deemed adequate. This is a source of uncertainty that should be included in the final uncertainty estimate, and it is the subject of future research.

4. RESULTS OF UNCERTAINTY QUANTIFICATION USING BAYES NETWORK

4.1 Model Parameters Results

Some selected results of the Bayes network implementation are shown in this section. For the estimation of probability density functions, 5000 samples were used. These were in addition to up to 10,000 samples discarded as burn-in samples to allow the Markov chain to become stationary. Convergence of the Markov chain is assessed by considering the samples created after the burn-in. Figure 6 show the samples for the linear stiffness, K_{lin} , nonlinear stiffness, K_{non} , nonlinear exponent, npow, and modulus of elasticity, E . Convergence is checked by plotting in Fig. 7, the moving average of the data from Fig. 6, which shows converged mean values for K_{lin} , K_{non} , npow, and E .

The first set of results show a comparison of the kernel density estimators (KDE) [25] for the updated parameters of the joint model, K_{lin} , K_{non} , and npow as they compare to the prior distributions. A KDE is an approximation to the PDF of the source of values of s , and it is computed from n data realizations, $s(\cdot)_j, j = 1 \dots n$. The form of the KDE used here is

$$\hat{f}_s(\alpha) = \frac{1}{n} \sum_{j=1}^n \frac{1}{\sqrt{2\pi}h} \exp \left[-\frac{1}{2h^2}(\alpha - s_j)^2 \right] \quad -\infty < \alpha < \infty, \quad (17)$$

where h is the width of a Gaussian kernel and is given by

$$h = k * \sigma_{\text{data}} * \text{npts}^{-0.2}, \quad (18)$$

where $k = 1.2$ and it is chosen to provide an undersmoothing of the data, σ_{data} is the standard deviation of the data and npts is the number of data points to fit the KDE to.

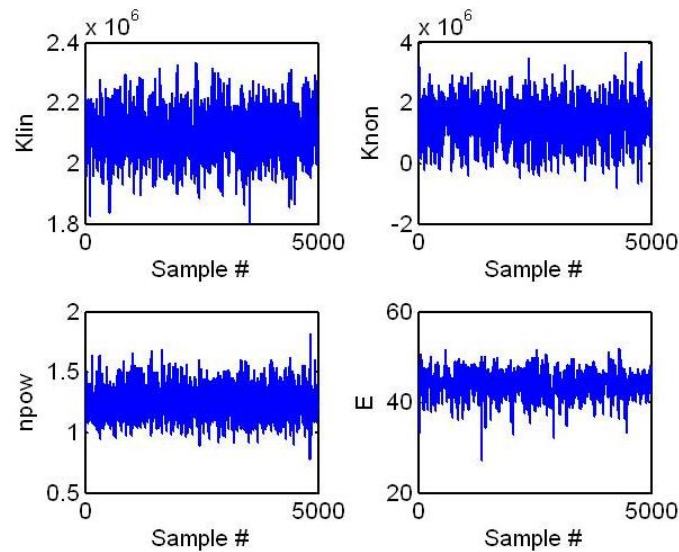


FIG. 6: Samples of K_{lin} , K_{non} , $npow$, and E used to demonstrate convergence of Markov chain to constant mean and variance.

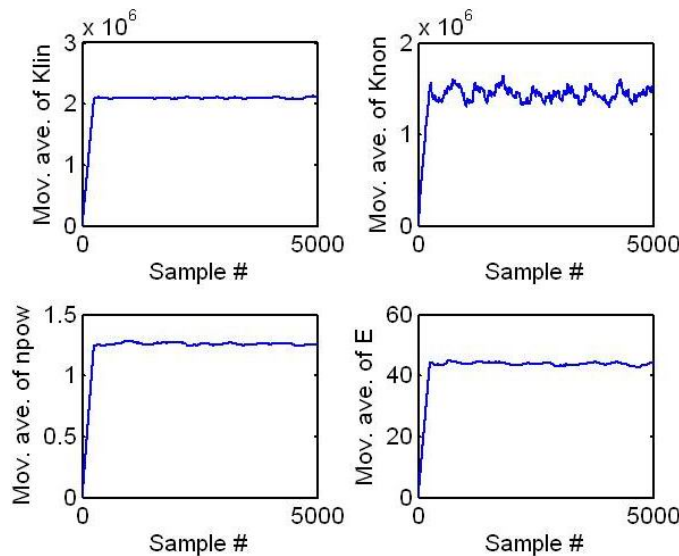


FIG. 7: Moving average to assess Markov chain convergence.

Figures 8–10 show these KDEs. Again, normal distributions were used as prior distributions for these parameters, with statistics (mean and standard deviation) obtained from the level 0 experiments. These are plotted as the solid lines in Figs. 8–10. These priors are updated with the available energy dissipation data measured in levels 1 and 2. The effect of the available experimental data is reflected in each posterior distribution relative to the prior. It is important to note that data from both levels 1 and 2 are simultaneously used to update these parameters, and thus, the effect of this data is reflected in the posterior of the parameters. It is interesting to observe in Figs. 8–10 that the posterior distributions of the joint parameters are not significantly different from the priors. This is due to the fact that the information that was used to generate the prior distributions for the joint parameters is consistent with the data obtained from levels 1 and 2; that is, the phenomena represented by the model parameters is very similar for all levels in the hierarchy, as far as the joint parameters are concerned.

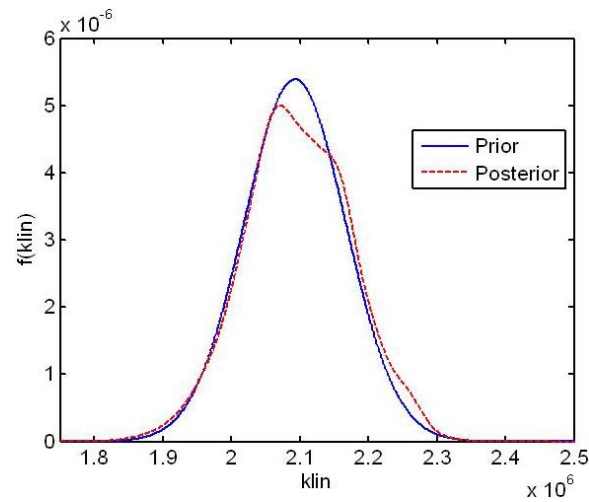


FIG. 8: KDE of linear stiffness from the joint model.

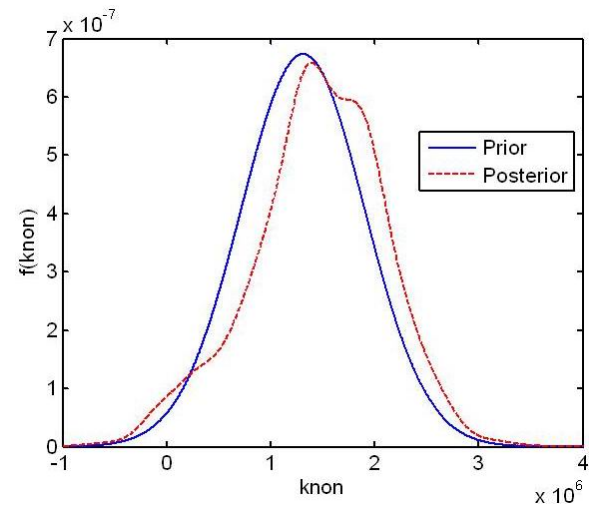


FIG. 9: KDE of nonlinear stiffness from the joint model.

Figure 11 shows the KDE for the modulus of elasticity which is the parameter of interest describing the foam. For the foam, the data used to update the parameter is the first natural frequency of the axial deformation mode measured at both levels 1 and 2. From Fig. 11, it can be seen that after updating, the posterior distribution of the modulus of elasticity has a smaller variance when compared to the prior. This is a reflection of having more information, and thus, reducing the uncertainty about a given parameter. It also gives an indication of the possible range of values, where this parameter might lie when both levels of complexity are included.

Figures 8–11 showed the effects of the experimental data at several input parameter nodes in the Bayes network when updating the parameters. Next, we will see the effect on the predicted model responses relative to the actual data.

4.2 Model Errors Results

In this section, the errors between the model predictions and the experimental data at each level are analyzed. These errors are denoted as corresponding to the joints level 1 and 2 and to the foam level 1 and 2, respectively, and are

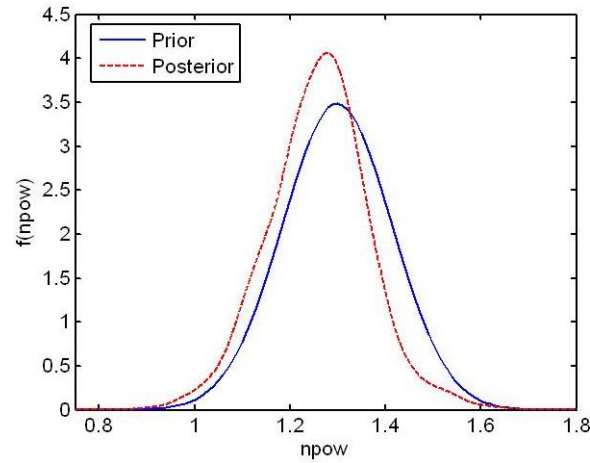


FIG. 10: KDE of degree of nonlinearity from the joint model.

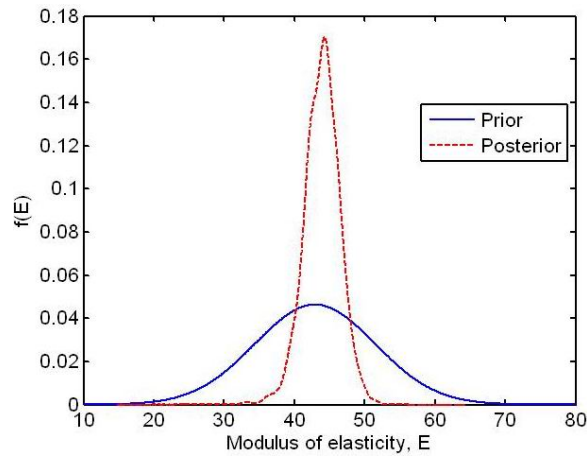


FIG. 11: KDE of modulus of elasticity of foam.

included in the Bayes network. The errors are assumed to follow a normal distribution with zero mean and standard deviation, σ . The standard deviation is updated with the observed data. This formulation is shown as follows:

$$\epsilon \sim N(0, \sigma), \quad (19)$$

where the hyperparameter σ is given a prior distribution of

$$\sigma \sim \gamma(0.0001, 0.0001), \quad (20)$$

and is subsequently updated in the Bayes network. In Eq. (20), the first parameter of the γ distribution refers to the shape of the distribution and the second parameter is the scale. All of the error terms are given the same prior for σ and were set to a very large range. Although it is acknowledged that this is not the most efficient way to select these parameters, we chose to start with an uninformed prior so that after the updating process, these parameters reflect the error contribution given the various data used for updating. The error terms estimated with this approach gave us an estimate that included various sources of error. Our ultimate goal is to determine what the various contributors to this error are, but this is the subject of further research and not discussed here. After updating the Bayes network, the posterior distribution of all the error terms are calculated and plotted in Figs. 12–15.

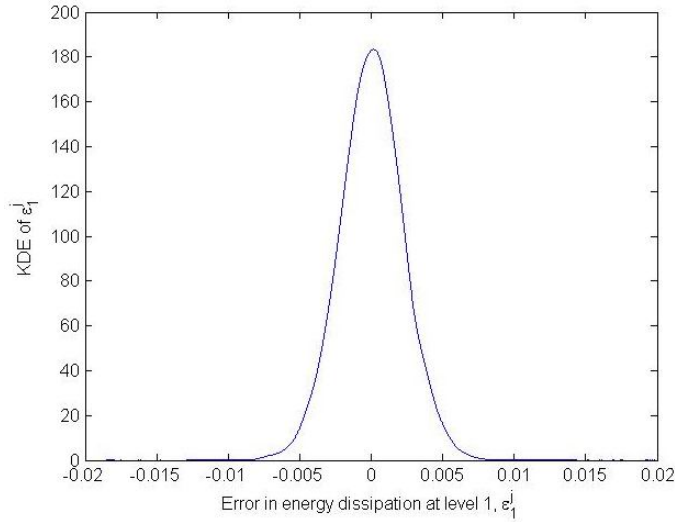


FIG. 12: KDE of error term ϵ_1^j for joints at level 1.

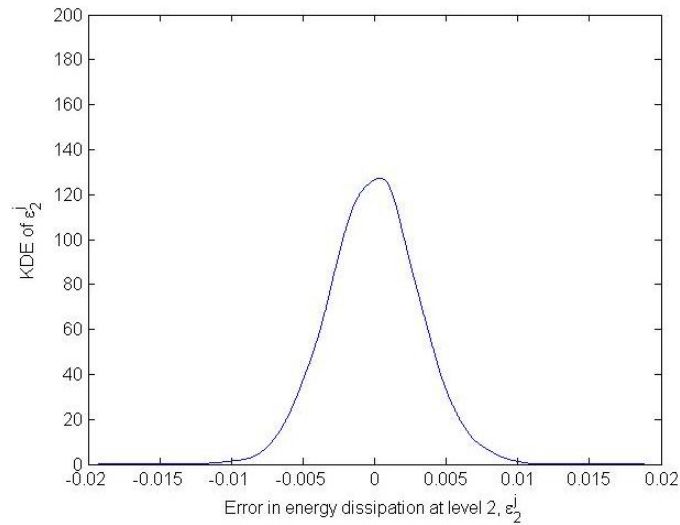


FIG. 13: KDE of error term ϵ_2^j for joints at level 2.

A general observation from these figures is that the model predictions are not perfect and some degree of error is present for all the measures of importance (i.e., energy dissipated and first axial frequency). In this example, the problem is well posed and the “apportionment” of uncertainty between parametric uncertainty and the error terms seems reasonable. This is a function of the form of the error and the choice of distribution attributed to this error. An inappropriate choice of the error form can lead to an overestimation or underestimation of the parametric uncertainty. This is an important consideration because one goal of uncertainty quantification is to properly quantify the different sources of uncertainty and their relative contribution to the total uncertainty in the system. Of importance in this study is the fact that the error term, which is a source of uncertainty, can be quantified and treated separately from the parametric uncertainty (due to the model parameters only). Although not treated in this study, some possible sources of this error are measurement error, model form error (arising from the choice of model selected to represent the physics of interest), and solution approximation error (e.g., mesh discretization).

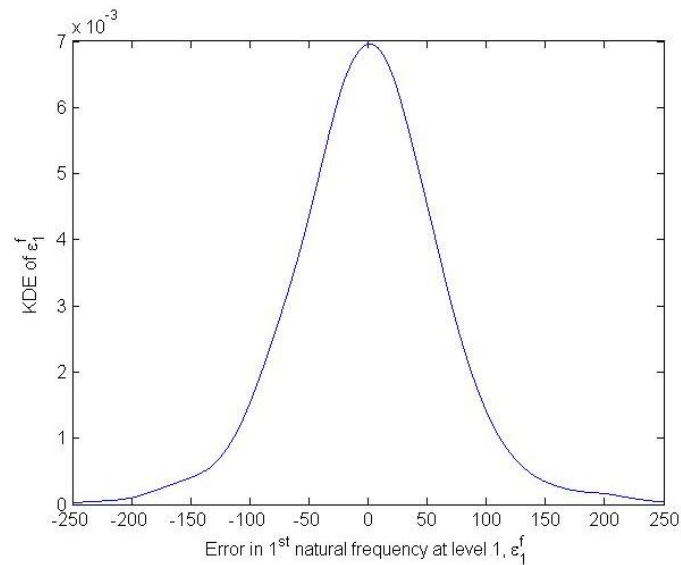


FIG. 14: KDE of error term ϵ_1^f for foam at level 1.

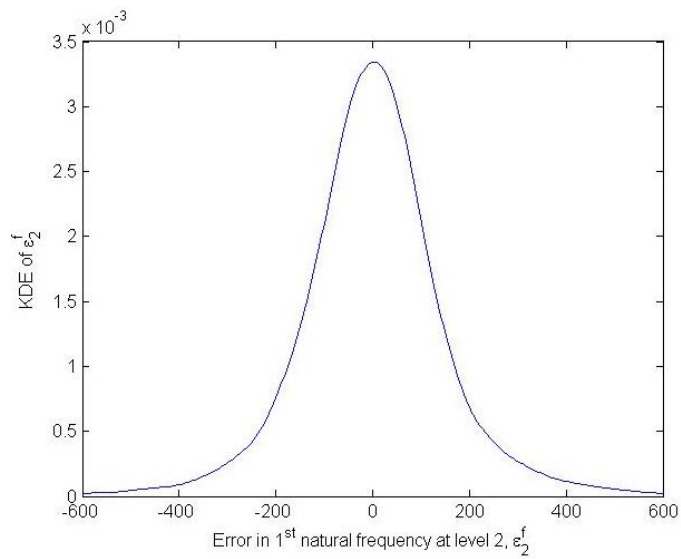


FIG. 15: KDE of error term ϵ_2^f for foam at level 2.

An interesting topic for further research would be to examine the individual contribution of each error source listed above to the overall error in the system level prediction. This will help identify possible areas of improvement and target those that have the greatest effect on the system-level prediction error.

4.3 Quantities of Interest Results

In this section, we look at the effect of updating the model parameters with data on the various quantities of interest for this system. It is important to point out that these plots are not intended to show validation of the model. They are merely for comparison purposes. A more formal validation could be carried out, but it is not the focus of this paper. For the joints, the quantity of interest is the energy dissipated at a particular force level. In this case, the force was chosen

to be 300 lbs, mainly because it is at the midpoint in the calibration data (level 0), which span 100–500 lbs. With this specification, Fig. 16 shows the response quantity energy dissipated per cycle of response at a mechanical joint force level of 300 lbs—for the level 1 hardware when computed using the prior distribution of the input parameters and after updating the parameters (shown as the posterior distributions). A comparison is also made to measured data. Figure 16 shows that the response of the model, when evaluated using the updated parameters, is a good representation of the mean behavior of the hardware at level 1 when compared to the KDE of the experimental data. It can be observed that a large part of the PDF is located roughly around 1.3×10^3 , which is close to the mean of the test data, and the variance of the prediction is smaller than the variance of the test data. The effect of the experimental data on the prior prediction is also observed from Fig. 16. The net effect is a reduction in the variance of the posterior distribution of this parameter relative to the prior distribution.

Next consider the measure of response of the foam at level 2, which, in this case, is the natural frequency of the axial deformation mode of the hardware. The KDEs of the prior and posterior predictions and the posterior prediction including error are shown in Fig. 17. The prediction refers to the model response for the foam at level 2 using the

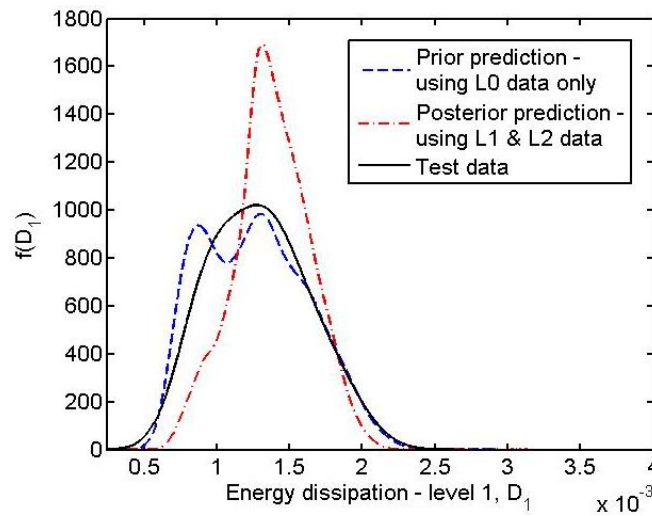


FIG. 16: KDE of energy dissipation at level 1.

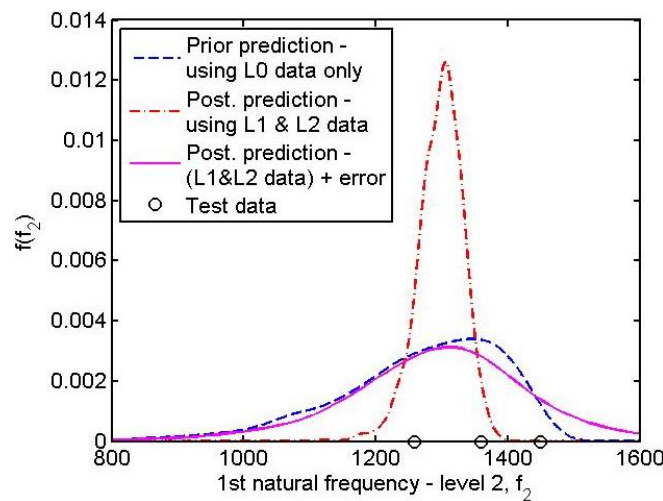


FIG. 17: KDE of natural frequency of the axial mode at foam level 2 and test data (shown as o).

updated parameter (i.e., the posterior distribution of the modulus of elasticity), and the prior prediction is calculated using the prior distributions of the parameters. The error is sampled from the KDE shown in Fig. 15 and added to the prediction as shown in Eq. (3). Also shown are the actual test data used to update the prior distributions. These are shown as small circles plotted on the abscissa. In this case, the KDE of the predicted behavior encompasses two of the three data points. When the KDE of the prediction plus the error term is plotted, it now captures all available data. This indicates that there are other sources of uncertainty that come into play to explain the discrepancy between model prediction and experimental data, which are not fully explained by the model input parameters themselves. In this case, the error term captures this discrepancy. This is an encouraging result because one of the hopes for this formulation was to properly apportion the uncertainty in model prediction to various sources. In other words, if a source of uncertainty is something other than parametric uncertainty, this should be captured as a separate error term, not rolled into the parameters themselves. In this research, no attempt is made (nor is it possible with the information available) to establish what the source of the error is. An understanding of the degree to which this error is present and whether it is a minor or major contributor to the overall uncertainty in the model prediction is desired. In this case, for level 2 foam, it shows some contribution. When observing Fig. 17, it is clear that the KDE of the prediction plus the error spans the available data, whereas the prediction by itself does not. Similarly to the energy dissipation at level 1, the prior prediction shows a larger variance when compared to the posterior distribution. This demonstrates one of the key features of a Bayes network; that is, to incorporate data that have the effect of reducing the uncertainty of a quantity of interest, in this case, the natural frequency of the axial deformation.

Finally, the kernel density estimator of the (extrapolated) predictions at the system level is shown in Fig. 18. This is the response at the top center of the internal mass. The extrapolation is both in terms of the input to the system (excitation with a blast load) and in the system that is used (cone with encapsulated mass). Several plots are shown in Fig. 18. The prior prediction refers to the system-level response evaluated using the prior distributions of the input parameters. This KDE shows a slight bimodality. The next three curves shown in Fig. 18 refer to the particular data used to perform the update. The posterior prediction using L1 data only means that only data from level 1 is used, whereas posterior prediction using L2 data only refers to level 2 data being used for updating. The last curve represents the case where all the available data at all levels are used for updating the Bayes network. Among the observations that can be made from Fig. 18, it is noted that, in general, all the posterior predictions show a decrease in variance relative to the prior, which means that any available data reduce the uncertainty present in the system-level response. In addition, and not surprisingly, when all the data are used, the variance of the posterior distribution of the system response is decreased the most relative to the other cases. The effect of the inclusion of data from various levels can be observed in Fig. 18. It is seen from this figure that when removing the level 2 data, the variance of the response is

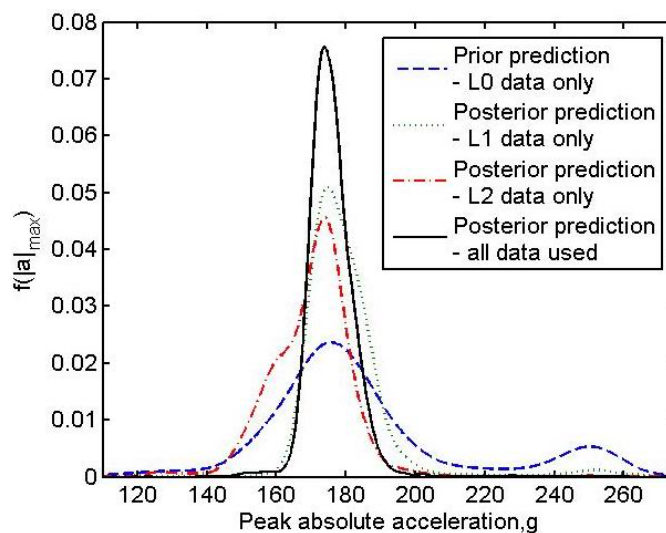


FIG. 18: KDE of peak absolute acceleration at the system level.

similar to the case where all data are used for updating. In contrast, when level 1 data are not included, the variance increases relative to the case where all data are included. This is a type of sensitivity analysis that shows the effect of the data from the different levels on the variance of the system-level response.

5. CONCLUSIONS

This paper presents an approach to quantify and propagate uncertainty in a complex system model that is built in a hierarchical manner. The analysis incorporates sources of aleatoric uncertainty (input variables and model parameters) within the context of a real, physical, complex, two-component problem using a Bayes network as the main analysis framework. The Bayes network is chosen because it conveniently allows the modeling of statistical dependencies between sets of data within a probability framework, based on conditional probabilities between the various nodes of the network. This framework also allows any available experimental data at any level to be incorporated into the analysis and calculates the posterior probabilities of all nodes in the network. A key issue that is addressed is the actual implementation of this methodology to perform fast and efficient computation within a Bayesian framework. This is done by using surrogate models (specifically Gaussian process models) in lieu of the finite element analysis, which can be expensive to run. Results show that using a Bayes network approach is a reasonable way to model a multilevel problem and available experimental data can be easily incorporated into the analysis. Error terms were included in the prediction at each level to account for model errors in addition to parametric uncertainty alone. These error terms were quantified but their sources were not determined at this stage.

The next step is to incorporate epistemic uncertainty into the analysis via the model parameters. This uncertainty can arise from lack of knowledge about a parameter of interest and will need to be treated probabilistically in order to be included in the Bayes network. Following this, the contribution of different sources of uncertainty to the overall uncertainty in the system-level prediction could be investigated.

ACKNOWLEDGMENTS

The research reported in this paper was supported by funds from Sandia National Laboratories, Albuquerque, New Mexico (Contract No. BG-7732), under the Sandia NSF life-cycle engineering program and Sandia National Laboratories Doctoral Study Program in support of the first author. The support is gratefully acknowledged. Sandia National Laboratories is a multiprogram 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 No. DE-AC04-94AL85000.

REFERENCES

1. Oberkampf, W. L. and Trucano, T. G., Validation methodology in computational fluid dynamics, Tech. Rep. No. SAND 2000-1656C, Sandia National Laboratories, 2000.
2. Sindir, M. M., Barson, S. L., Chan, D. C., and Lin, W. H., On the development and demonstration of a code validation process for industrial applications, *Proc. of 27th AIAA Fluid Dynamics Conference*, AIAA Paper No. 96-2032, 1996.
3. Urbina, A., Paez, T. L., Gregory, D., Resor, B., Hinnerichs, T. D., and O'Gorman, C. C., Validation of a combined non-linear joint and viscoelastic encapsulating foam, *Proc. of 2006 Society for Experimental Mechanics Conference*, AIAA Paper No. 96-2032, 2006.
4. U.S. Department of Energy, Advanced simulation and computing (ASCI) program plan, Tech. Rep. No. 01-ASCI-Prog-01, 2000.
5. AIAA, Guide for the verification and validation of computational fluid dynamic simulations, Tech. Rep. No. AIAA-G-077-1998, American Institute of Aeronautics and Astronautics, 1998.
6. Rebba, R. and Mahadevan, S., Model predictive capability assessment under uncertainty, *AIAA J.*, 44(10):2376–2384, 2006.
7. McFarland, J. M., Uncertainty analysis for computer simulations through validation and calibration, PhD thesis, Vanderbilt University, 2008.

8. Rasmussen, C., Evaluation of gaussian processes and other methods for non-linear regression, PhD thesis, University of Toronto, 1996.
9. Martin, J. and Simpson, T., Use of kriging models to approximate deterministic computer models, *AIAA J.*, 43(4):853–863, 2005.
10. Mardia, K. and Marshall, R., Maximum likelihood estimation of models for residual covariance in spatial regression, *Biometrika*, 71:853–863, 1984.
11. Santner, T. J., Williams, B. J., and Noltz, W. I., *The Design and Analysis of Computer Experiments*, Springer: Verlag, Berlin, 2003.
12. Gibson, L. J. and Ashby, M. F., *Cellular Solids: Structure and Properties*, Cambridge University Press, Cambridge, UK, 1999.
13. Paz, M., *Structural Dynamics: Theory and Computation*, Van Nostrand Reinhold, New York, 1991.
14. Smallwood, O. D., Gregory, D., and Coleman, R., Damping investigations of a simplified frictional shear joint, *Proc. of 71st Shock and Vibration Symposium*, Shock and Vibration Information Analysis Center, Arvonion, VA, 2000.
15. Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., *Bayesian Data Analysis*, Chapman and Hall/CRC, Boca Raton, 2004.
16. Robinson, D., A hierarchical bayes approach to system reliability analysis, Tech. Rep. No. SAND2001-3513, Sandia National Laboratories, 2001.
17. Williams, B., Higdon, D., Gattiker, J., Moore, L., McKay, M., and Keller-McNulty, S., Combining experimental data and computer simulations, with an application to flyer plate experiments, *Bayesian Anal.*, 1(4):765–792, 2006.
18. Jensen, F. V., *Bayesian Networks and decision graphs*, Springer: Verlag, Berlin, 2001.
19. Glymour, C., Scheines, R., Spirtes, P., and Kelly, K., *Discovering Causal Structure*, Academic Press, New York, 1987.
20. Spiegelhalter, D. J., Thomas, J. A., Best, N. G., and Lunn, D., WinBUGS User Manual Version 1.4, MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK, 2003.
21. Kennedy, M. C. and O'Hagan, A., Predicting the output from a complex computer code when fast approximation are available, *Biometrika*, 87(1):1–13, 2000.
22. Gilks, G. R., Richardson, S., and Spiegelhalter, D. J., *Markov Chain Monte Carlo in Practice*, Chapman and Hall/CRC, Boca Raton, 1996.
23. Bichon, B. J., Eldred, M. S., Swiler, L. P., Mahadevan, S., and McFarland, J. M., Efficient global reliability analysis for nonlinear implicit performance functions, *AIAA J.*, 46(10):2459–2468, 2008.
24. Oberon Microsystems Inc., *Component Pascal Language Report*, 2006.
25. Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, Boca Raton, Florida, 1986.