AIAA 2016-1443
AIAA SciTech
4-8 January 2016, San Diego, California, USA
18th AIAA Non-Deterministic Approaches Conference

Compositional Uncertainty Analysis via Importance Weighted Gibbs Sampling for Coupled Multidisciplinary Systems

S. F. Ghoreishi* and D. L. Allaire†

Department of Mechanical Engineering, Texas A&M University, College Station, TX, 77843

This paper presents a novel compositional multidisciplinary uncertainty analysis methodology for systems with feedback couplings, and model discrepancy. Our approach incorporates aspects of importance resampling, density estimation, and Gibbs sampling to ensure that, under mild assumptions, our method is provably convergent in distribution. A key feature of our approach is that disciplinary models can all be executed offline and independently. Offline data is synthesized in an online phase that does not require any further model evaluations or any full coupled system level evaluations. We demonstrate our approach on a simple aerodynamics-structures system.

I. Introduction

In response to progress in science and technology, demands on modern aerospace vehicles to have better performance, higher reliability and robustness, and lower cost and risk are ever increasing [1]. This demand has led to the development of highly coupled systems designed to exploit interactions among disciplines to achieve greater performance. There is typically a great deal of uncertainty associated with such systems due to potential new regimes of system behavior arising from unexpected multi-physics interactions. In traditional design practice, to account for such uncertainties, constraints imposed on the design are often reformulated deterministically, with pre-defined safety factors and margins, to ensure the reliability of a design. This approach will typically result in an overly conservative design that cannot meet the demands placed on today's modern aerospace systems. Thus, there is a critical need for the development of advanced, scalable technologies aimed at rigorously quantifying uncertainty in multi-physics aerospace systems.

Multidisciplinary uncertainty analysis of numerical simulation models entails the propagation of uncertainty from model inputs to model outputs. Sources of uncertainty in such systems, as described in Ref. 2, typically include parametric uncertainty, which involves uncertainty associated with model input parameters, parametric variability, which generally refers to variation that cannot be controlled (e.g., operating conditions), code uncertainty, which refers to the uncertainty associated with interpolating between known system responses, and model discrepancy, which is uncertainty associated with the fact that no model is perfect. Often, multidisciplinary simulation capabilities are composed by integrating pre-existing disciplinary physics-based models. For such composed multi-physics systems, the task of uncertainty analysis can be challenging owing to the disciplinary models being managed by separate entities or housed in separate locations, analysis capabilities running on different computational platforms, models with significantly different analysis run times, and the sheer number of disciplines required for a given analysis. Complex, multi-physics systems also often exhibit feedback coupling (e.g., aeroelastic coupling in an aero-structural analysis of a wing) that can have a substantial impact on system level uncertainty analysis results. Further, the coupled nature of multi-physics systems necessitates an iterative uncertainty analysis procedure in traditional Monte Carlo based system level uncertainty analysis approaches that can be computationally prohibitive.

To alleviate the computational burden of multidisciplinary uncertainty analysis in coupled systems composed of integrated disciplinary models, we propose a probabilistic, sample-based compositional uncertainty

*Ph.D. Candidate, AIAA Student Member
†Assistant Professor, AIAA Member

analysis methodology. This work builds off of the work of Refs. 3 and 4, where a decomposition-based uncertainty analysis methodology is developed for feedforward systems with parametric input uncertainty. Recognizing that the disciplinary models composing the system are imperfect, our method takes into account the model discrepancy associated with coupling variables in the system.

A key feature of our work is an offline/online approach to uncertainty analysis enabled by the compositional nature of an integrated multidisciplinary system. Specifically, our approach is designed to use offline *independent* uncertainty analysis results for each discipline of the composed system in an online synthesis procedure. Offline results can be generated at any time (e.g., from a previous use of a disciplinary model) and do not require knowledge of the specific input probability distributions that will be used in the composed system level analysis. Further, the proposed online synthesis of offline data from the disciplinary models does not require any coupled system level evaluations. The result is that the computational expense of coupled multidisciplinary uncertainty analysis is moved offline, allowing for substantial gains in computational efficiency. The proposed methodology incorporates aspects of density estimation, Radon-Nikodym importance weights, and ensemble Gibbs sampling to ensure convergence in distribution of system level uncertainty analysis results under mild assumptions on the disciplinary models and probabilistic distributions. We demonstrate our methodology on an aero-structural system adapted from Ref. 5.

The rest of the paper is organized as follows. Section II presents background on related work. In Section III, the approach is described, which is followed by the presentation of our algorithm in Section IV. In Section V, we present the results, and conclusions are drawn in Section VI.

II. Background

Previous work on multidisciplinary uncertainty analysis has focused on approximations such as surrogate modeling and simplified representations of system uncertainty. The use of surrogates for disciplinary models in a composed system can provide computational savings, as well as simplify the task of integrating components [6]. Approximate representations of uncertainty, such as using mean and variance information in place of a full probability distribution have been used to avoid the need to propagate uncertainty between disciplines. Such simplifications are commonly used in uncertainty-based multidisciplinary design optimization methods as a way to avoid a system-level uncertainty analysis [1]. These approachs include implicit uncertainty propagation [7], reliability-based design optimization [8], robust moment matching [9–11], advanced mean value method [12], collaborative reliability analysis using most probable point estimation [13], and a multidisciplinary first-order reliability method [14].

Other recent work has focused on exploiting the structure of a given multidisciplinary system. Ref. 15 present a likelihood-based approach to decouple feedback loops, thus reducing the problem to a feed-forward system. Dimension reduction and measure transformation to reduce the dimensionality and propagate the coupling variables between coupled components have been performed in a coupled feedback problem with polynomial chaos expansions [16–18]. Coupling disciplinary models by representing coupling variables with truncated Karhunen-Loève expansions, has been studied for multi-physics systems [19]. A hybrid method that combines Monte Carlo sampling and spectral methods for solving stochastic coupled problems has also been proposed by Refs. 20 and 21.

Our approach builds on the work of Refs. 3 and 4, where the challenges of uncertainty analysis for feed-forward multidisciplinary systems were dealt with using a decomposition-based approach. As shown notionally in Figure 1, the multidisciplinary uncertainty analysis is decomposed into individual discipline level uncertainty analyses. These analyses are then assembled in a provably convergent manner to the desired multidisciplinary uncertainty analysis results. Taking this approach leads to several benefits, such as enabling offline disciplinary analyses to be conducted when suitable, enabling concurrent evaluation of the disciplinary models for online uncertainty analysis, avoiding the challenges of integrating various disciplinary models that may have been created on different computational platforms, and being consistent with many organizational structures. In the work presented here, we extend the methodology of Ref. 4 to handle feedback coupling between disciplines, as well as model discrepancy associated with each discipline. This concept is presented notionally in Figure 2, where a two-discipline system is represented. We note that here we use the term compositional rather than decompositional to stress the concept of integrating a set of information sources rather than beginning with a monolithic system that is to be decomposed.

In this example, each discipline takes as input, an output from the other discipline. The output of each discipline has additive model discrepancy associated with it, denoted by δ_1 and δ_2 for discipline 1 and 2

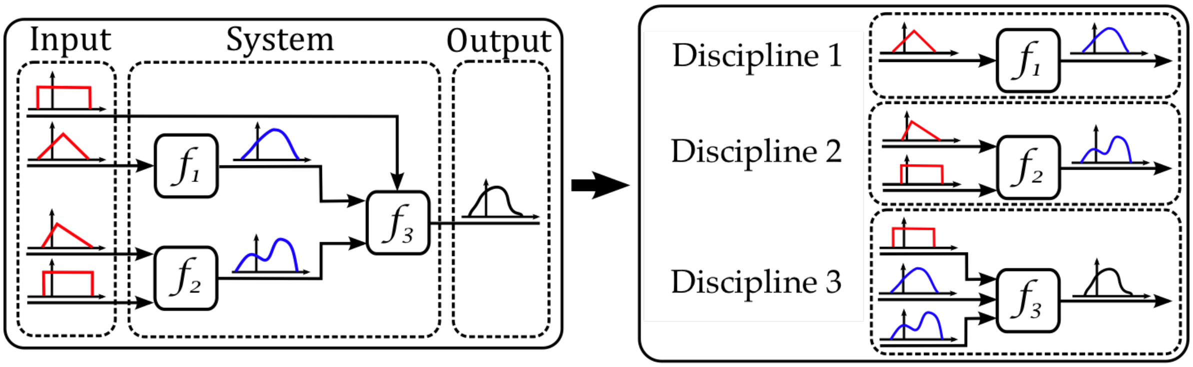

Figure 1: The decomposition-based feed-forward multidisciplinary uncertainty analysis method of Ref. 4. The method decomposes the problem into manageable components and synthesizes the system level uncertainty analysis without needing to evaluate the system in its entirety. f_1, f_2 and f_3 are the input-output functions associated with each component.

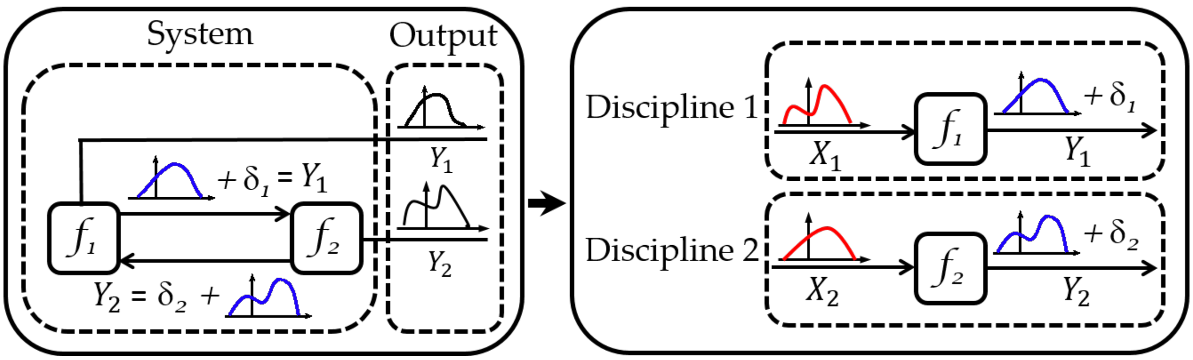

Figure 2: A depiction of the concept of a compositional multidisciplinary uncertainty analysis approach for a two-discipline system with model discrepancy and feedback coupling. Here, f_1 and f_2 are the model functions, X_1, X_2 and Y_1, Y_2 are the inputs and outputs of the respective disciplines, and δ_1 and δ_2 are the additive model discrepancies associated with the outputs. The approach composes disciplinary uncertainty analysis results without needing to evaluate the coupled system.

respectively. We discuss this source of uncertainty in more detail in the following section. We note here that it is this source of uncertainty that necessitates an iterative approach to uncertainty quantification as discussed in Ref. 10, which leads to substantial computational expense for multidisciplinary uncertainty analysis.

III. Approach

We first present how we characterize model discrepancy. We then provide an overview of our compositional multidisciplinary uncertainty analysis, followed by a discussion of the methodological key ingredients of importance resampling and Gibbs sampling. We also discuss the use of effective sample size as a heuristic indicator of the quality of a proposal distribution and as a means of identifying stationarity.

A. Model Discrepancy Characterization

Model discrepancy arises because mathematical models of reality are not perfect, and thus, some aspects of reality may have been omitted, improperly modeled, or contain unrealistic assumptions. Following Ref. 2, we represent model discrepancy as an additive stochastic process. For example, if we have a model that consists of a function $f(\mathbf{x})$, where $\mathbf{x}$ is an input vector to the model, and reality is denoted as $f_r(\mathbf{x})$, then the model discrepancy of the model can be represented as

$$\delta(\mathbf{x}) = f_r(\mathbf{x}) - f(\mathbf{x}), \tag{1}$$

where here we are assuming there are no parameters in the model to be calibrated. Typically, we will have available experimental data of reality (which will contain experimental variability), which can be used to create a stochastic process representation for $\delta(\mathbf{x})$. In this work, we assume that model discrepancy has been quantified previously for all available disciplinary models in the form of Gaussian processes. Thus, we add the Gaussian process model discrepancy term to the output of a disciplinary model, as shown notionally in Figure 2. For example, for Discipline 1 in Figure 2, we have

$$Y_1(\mathbf{x}) = f_1(\mathbf{x}) + \delta_1(\mathbf{x}), \tag{2}$$

where $\delta_1(\mathbf{x})$ is the Gaussian process representation of the model discrepancy of Discipline 1 and $Y_1(\mathbf{x})$ is the estimate of reality from Discipline 1 with quantified model discrepancy. In this work, without loss of generality, we focus on discrepancies that are not a function of an input.

B. Overview

Our proposed compositional multidisciplinary uncertainty analysis approach for systems with feedback coupling, and model discrepancy, consists of an offline and online phase. To make the discussion more concrete but without loss of generality, we develop the aspects of the approach with the system presented in Figure 2 in mind. We begin offline, by performing uncertainty analysis independently for each discipline of the system. To do this, we must propose distributions for inputs for each discipline, since we do not know the distributions of the coupling variables in advance. We define theses distributions as proposals which are represented by π_{X_1} and π_{X_2}. This process is shown notionally in Figure 3. As it is shown, the samples generated from the proposal distributions are propagated through the disciplines with functions f_1 and f_2 to generate the corresponding samples of the discipline outputs. The underlying densities of the output samples are then approximated as π_{Y_1} and π_{Y_2} using a density estimation technique such as kernel density estimation [22].

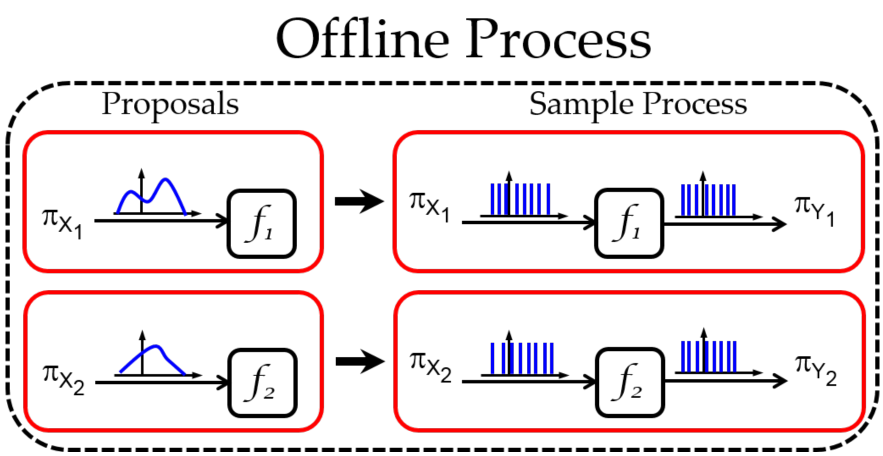

Figure 3: A depiction of the offline uncertainty analysis process adapted from Ref. 4. Here, π_{X_1}, π_{X_2}, π_{Y_1}, and π_{Y_2} are the densities associated with the inputs (proposals) and outputs of the respective disciplines.

The online process consists of a Gibbs sampling procedure that is enabled by iteratively re-weighting the samples generated offline for each discipline through a combination of density estimation and sequential importance resampling. The use of importance weights allows us to change underlying input probability distributions to each discipline. These weights can then be used on the output probability distributions of each discipline. The weighted output distributions converge in distribution to the output distributions that would have resulted from the modified input probability distributions sampled from in the first place. The overall online process is presented notionally in Figure 4 for the system presented in Figure 2. Here, we are assuming that each discipline contains model discrepancy, and the output of the systems (e.g., the quantity of interest) is the joint distribution of the coupling variables. Our online process evolves this joint distribution iteratively via Gibbs sampling of the conditional densities. For example, in Figure 4, we begin with the joint distributions $\pi_{X_1|Y_2}, \pi_{Y_1|X_1}, \pi_{X_2|Y_1}$, and $\pi_{Y_2|X_2}$. The initial distribution assumed for each conditional distribution is given by the offline uncertainty analysis. From these proposed distributions (π_{X_1} and π_{X_2}), some samples are drawn to pass through each discipline independently to produce the initial samples from the target distributions ($\pi_{Y_1|X_1}$ and $\pi_{Y_2|X_2}$). In the online process, the joint distribution of each conditional distribution is updated sequentially and iteratively using Radon-Nikodym importance weights. To do so, we re-weight and update the samples from $\pi_{X_1|Y_2}$ and $\pi_{X_2|Y_1}$ followed by a density estimation technique and using the ratio of densities of target and proposal distributions as the Radon-Nikodym weights. By sequential importance resampling, $\pi_{X_1|Y_2}$ and $\pi_{X_2|Y_1}$ match $\pi_{Y_2|X_2}$ and $\pi_{Y_1|X_1}$ respectively, so the target distributions are simulated and propagated without rerunning the model online. This iterative process results in the convergence in distribution of the joint distribution of Y_1 and Y_2 as the number of iterations grows. Hence, the joint distribution of the two sources of uncertainty in the system is recovered without ever performing a coupled system level analysis.

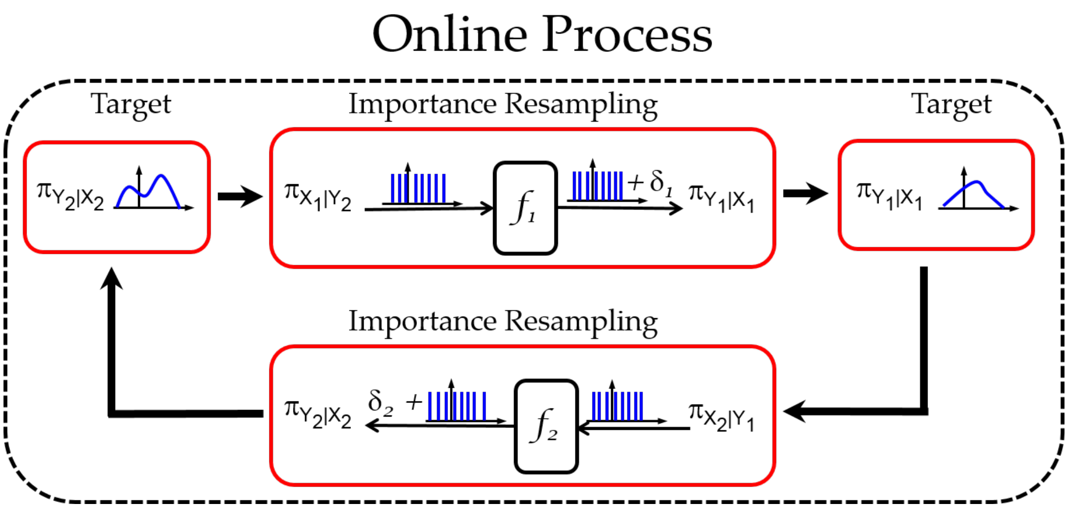

Figure 4: A depiction of the online uncertainty analysis process. We note here that the models f_1 and f_2 are not actually executed in the online process. Instead, the Radon-Nikodym importance weights found for re-weighting the input proposal distribution are passed on to the output distributions.

C. Importance Weights and Density Estimation

A key ingredient of our online approach is the iterative re-weighting of offline sample information. We achieve this using sequential importance resampling [23]. We use this method as follows. Consider a bivariate density of the random variables Y_1 and Y_2 and assume we have sampled from a bivariate normal distribution where $Y_1 \sim \mathcal{N}(0, 2)$, $Y_2 \sim \mathcal{N}(0, 2)$, and the random variables are independent. Contours of this density are shown in red on the left in Figure 5 and samples of this distribution are shown with red dots on the same figure. These samples could then be propagated through a model, say $f(Y_1, Y_2)$ to compute statistics of interest such as the expectation, $\mathbb{E}[f(Y_1, Y_2)]$. Now suppose that instead of the distribution represented by the red contours on the left in Figure 5, we wished to compute the expectation of $f(Y_1, Y_2)$ with Y_1 and Y_2 distributed jointly according to the distribution represented by the blue contours on the left in Figure 5. A traditional approach would require sampling from this distribution and then propagating these samples through the model. However, importance resampling, which is based on the Radon-Nikodynm theorem [24], allows us to compute this information by simply re-weighting the samples from the original distribution, which we can then resample from with replacement. The concept is demonstrated in Figure 5, where Radon-Nikodym importance weights are used to re-weight the samples given by the red dots on the left figure. The result is the blue dots on the right figure, where the size of the dots denotes the relative weights given to the samples. These weights are the Radon-Nikodym derivatives of the desired distribution with respect to the original distribution evaluated at the original sample points. Mathematically, if we wished to compute the expected value of some function $f(Y_1, Y_2)$ using the density represented by the blue contours but had only samples of the density represented by the red contours, we would proceed as follows. Referring to the original distribution as the *proposal* distribution and denoting its density by $p(Y_1, Y_2)$ and referring to the desired distribution as the *target* distribution and denoting its density by $q(Y_1, Y_2)$, the expected value of $f(Y_1, Y_2)$

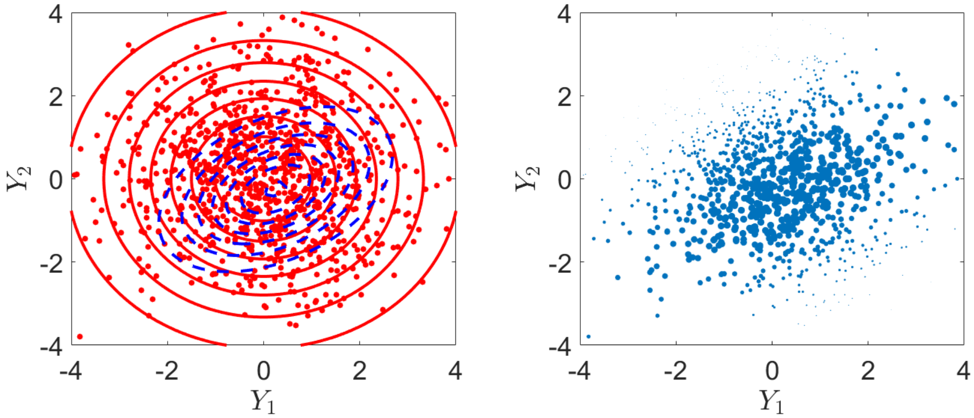

Figure 5: The importance resampling process uses the realizations shown by red dots on left figure, generated from a proposal distribution $P(Y_1, Y_2)$ (corresponding density shown as red solid contour on left figure) to approximate a target distribution $Q(Y_1, Y_2)$ (blue dash contour on left figure), by weighting the proposal realizations shown by blue dots on right figure (adapted from Ref. 4).

with respect to the target distribution is given as

$$\mathbb{E}_Q[f(Y_1, Y_2)] = \int_{\mathcal{P}} \frac{q(Y_1, Y_2)}{p(Y_1, Y_2)} f(Y_1, Y_2) p(Y_1, Y_2) d\boldsymbol{Y}, \tag{3}$$

where $\mathcal{P}$ is the support of the proposal density and the target distribution is absolutely continuous with respect to the proposal distribution. We can evaluate this integral with Monte Carlo simulation as

$$\mathbb{E}_Q[f(Y_1, Y_2)] \approx \frac{1}{N} \sum_{i=1}^{N} f(Y_1^i, Y_2^i) \frac{q(Y_1^i, Y_2^i)}{p(Y_1^i, Y_2^i)}, \tag{4}$$

where the samples $\{Y_1^i, Y_2^i\}$ are drawn from the proposal distribution (i.e., the red contours in Figure 5). This is the key aspect of the method. The proposal samples are propagated through a model offline. Online, we re-weight these samples so as to simulate a target distribution being propagated through a model. Importance resampling allows us to achieve this simulation *without rerunning the model online*. We note here that we do sequential importance resampling (sampling from the discrete distribution given by the Radon-Nikodym importance weights) followed by a density estimation technique to update the joint densities in the online process. This is required because the Radon-Nikodym weights are density ratios.

D. Effective Sample Size

As noted in Section IIIB, since we do not know the joint distributions of the coupled variables in advance, we need to propose distributions for the inputs of each discipline. These proposal distributions come from the previous knowledge about the system, and can impact the convergence performance of the compositional multidisciplinary uncertainty analysis approach. In general, we cannot evaluate the quality of the proposal distributions before executing the approach; but after knowing the importance weights, we can do so by computing the effective sample size as

$$n_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} (w(x_i))^2}, \tag{5}$$

where N is the number of samples and $w(x_i)$ is the normalized importance weight assigned to the proposal sample x_i [25–27]. The effective sample size ranges from $n_{\text{eff}} = 1$ to $n_{\text{eff}} = N$. We have $n_{\text{eff}} = N$ when all the samples have the same weights ($\frac{1}{N}$), so the proposal and target distributions are equal to each other, while $n_{\text{eff}} = 1$ shows that the weight of all the samples except one sample is 0. Thus, effective sample size is often used as a heuristic for determining when the proposal distribution adequately captures the target distribution [28]. We take this approach here as discussed in Section IV, where we also consider the use of an unchanging effective sample size as a heuristic indicator of stationarity in a target distribution.

E. Gibbs Sampling

The second key ingredient of our approach is Gibbs sampling. The presence of model discrepancy in a multidisciplinary system with feedback coupling requires the addition of that model discrepancy iteratively. That is, considering the system in Figure 2, if we first evaluate Discipline 1 (with an initial estimate of the input to this discipline from Discipline 2), we have an estimate of the output of Discipline 1 that feeds into Discipline 2. To this output, we add the discrepancy term denoted by δ_1 in the figure. Discipline 2 can then be evaluated and to its output that feeds back into Discipline 1, we must add the discrepancy term denoted by δ_2 in the figure. This process must be repeated until some form of convergence is achieved. Here, the convergence we require is convergence in distribution. That is, we require an iterative process that, when enough iterations have occurred, the joint distribution of the coupling variables is stationary. One method for achieving this is through the use of Gibbs sampling.

Gibbs sampling was first described by Ref. 29 and has also been referred to as successive substitution sampling [30]. It is a Markov chain Monte Carlo based method for generating samples from a joint distribution that cannot be directly sampled. Following Refs. 30 and 31, suppose we have random variables (which can be vector-valued) $Y_1, \dots, Y_k$, and we wish to generate samples from the joint distribution of those random variables, which we denote as $\pi_{Y_1,\dots,Y_k}$. Assume we have a complete set of conditional distributions, $\pi_{Y_i|\{Y_j, j\neq i\}}$ for $i = 1, \dots, k$, available for sampling. Then, starting from some initial arbitrary set of values, $y_1^{(0)}, \dots, y_k^{(0)}$, where the superscript denotes the iteration number of the Gibbs sampler, we draw a sample $y_1^{(1)}$ from $\pi_{Y_1|Y_2=y_2^{(0)},\dots,Y_k=y_k^{(0)}}$. We then draw a sample $y_2^{(1)}$ from $\pi_{Y_2|Y_1=y_1^{(1)},Y_3=y_3^{(0)},\dots,Y_k=y_k^{(0)}}$, and continue in this manner up to y_k from $\pi_{Y_k|Y_1=y_1^{(1)},\dots,Y_{k-1}=y_{k-1}^{(1)}}$, which completes one iteration of the Gibbs sampler. After m iterations we obtain $(y_1^{(m)}, \dots, y_k^{(m)})$. For continuous distributions, Ref. 30 has shown, under mild assumptions, that this k-tuple converges in distribution to a random observation from $\pi_{Y_1,\dots,Y_k}$ as $m \to \infty$.

To demonstrate the applicability of Gibbs sampling to multidisciplinary uncertainty analysis, consider again the system shown in Figure 2. Let $y_1^{(0)}$ be an initial estimate from Discipline 1, where an arbitrary estimate from Discipline 2 has been assumed. We can then propagate this sample through Discipline 2, which generates a sample from the conditional distribution, $\pi_{Y_2|Y_1=y_1^{(0)}}$, which we denote as $y_2^{(1)}$. We can then propagate this sample through Discipline 1 to generate a sample from the conditional distribution $\pi_{Y_1|Y_2=y_2^{(1)}}$, and so on. By repeating this process many times, we ensure that our random sample, $(y_1^{(m)}, y_2^{(m)})$ converges in distribution to a random observation from the true joint distribution of Y_1 and Y_2.

For our compositional multidisciplinary uncertainty analysis approach and the system represented by Figures 2 and 4, we perform Gibbs sampling over the conditional joint distributions of $X_1|Y_2, Y_1|X_1, X_2|Y_1$, and $Y_2|X_2$. We begin with initial distributions, $\pi_{X_1^{(0)}}, \pi_{Y_1^{(0)}}, \pi_{X_2^{(0)}}$, and $\pi_{Y_2^{(0)}}$, which were generated offline. We then use Radon-Nikodym importance weights from the ratio of $\pi_{Y_2^{(0)}}$ to $\pi_{X_1^{(0)}}$ to generate the distribution $\pi_{X_1^{(1)}|Y_2^{(0)}}$ as well as $\pi_{Y_1^{(1)}|X_1^{(1)}}$. We can then use Radon-Nikodym importance weights from the ratio of $\pi_{Y_1^{(1)}}$ to $\pi_{X_2^{(0)}}$ to generate the distribution $\pi_{X_2^{(1)}|Y_1^{(1)}}$ and $\pi_{Y_2^{(1)}|X_2^{(1)}}$, which completes one iteration of the importance weighted Gibbs sampler.

IV. Compositional Multidisciplinary Uncertainty Analysis

We present here our compositional multidisciplinary uncertainty analysis algorithm. We also provide a discussion on the convergence properties of our approach.

A. Algorithm

Our proposed compositional multidisciplinary uncertainty analysis method for coupled systems with model discrepancy is presented in Algorithm 1. The data necessary for this algorithm are the samples generated from the proposal distributions as inputs and the output samples of each discipline obtained independently in the offline process. We represent the number of samples and the number of disciplines as N and N_d respectively. Since this approach is based on samples, to be able to compute the importance weights, we need to estimate the densities of the inputs, π_X, and outputs, π_Y, of the disciplines. The algorithm starts from the first discipline, and computes the unnormalized weights $(w_{i,j}^{un})$ for each sample of disciplines $(\{x_{i,j}\}_{i=1,\dots,N}^{j=1,\dots,N_d})$ as $w_{i,j}^{un} = \frac{\pi_{Y_{j-1}}(x_{i,j})}{\pi_{X_j}(x_{i,j})}$, which is the ratio of the target input density (output from the upstream discipline) and the proposal input density computed at each sample previously simulated from the proposal input distribution. After assigning weights to all the samples, N samples are drawn with replacement from the output samples, and the model discrepancy corresponding to the current discipline is added to the importance resampled samples. From these updated samples, the probability density of the output, which is the new target input distribution for the downstream discipline, is updated. To compute the effective sample size, we need to normalize the weights assigned to the samples to sum to unity. After repeating these steps for all the disciplines, one iteration of the Gibbs sampler is completed, and M, which represents the number of Gibbs iterations and is initialized to 0, is updated to $M+1$. The algorithm stops when the relative difference between the effective sample size in the current Gibbs iteration and the previous one, $|\frac{n_{\text{eff}}^j(M)-n_{\text{eff}}^j(M-1)}{n_{\text{eff}}^j(M)}|$, for all the disciplines is less than a user-defined threshold (ϵ), which shows that the approach has converged to a stationary joint distribution of the coupling variables.

Algorithm 1: Compositional Multidisciplinary Uncertainty Analysis.

Data: Offline sample sets of inputs $\{\mathbf{x}_{i,j}\}_{i=1,\ldots,N}^{j=1,\ldots,N_d}$ and outputs $\{\mathbf{y}_{i,j}\}_{i=1,\ldots,N}^{j=1,\ldots,N_d}$ obtained from offline Monte Carlo based individual discipline analyses.
Result: The joint distribution of outputs of all the disciplines $\{Y_j\}_{j=1,\ldots,N_d}$ and their statistics.
Parameters: N = number of samples; N_d = number of disciplines; M = number of Gibbs iterations; δ_j = additive model discrepancy of discipline j; n_{eff} = effective sample size; ϵ = user-defined threshold for the iteration stopping criterion; $w_{i,j}^{un}$ = unnormalized weights; $w_{i,j}$ = normalized weights

Initialize input and output distributions of $\{X_j\}_{j=1,\ldots,N_d}$ (π_{X_j}) and $\{Y_j\}_{j=1,\ldots,N_d}$ (π_{Y_j}) using offline sample sets, and $M=0$.
while $\max_{j\in\{1,\ldots,N_d\}}(\epsilon_j) > \epsilon$ **do**
  for $j=1:N_d$ **do**
    for $i=1:N$ **do**
      if $j=1$ **then**
        $w_{i,j}^{un} = \frac{\pi_{Y_{N_d}}(x_{i,j})}{\pi_{X_j}(x_{i,j})}$
      else
        $w_{i,j}^{un} = \frac{\pi_{Y_{j-1}}(x_{i,j})}{\pi_{X_j}(x_{i,j})}$
      end
    end
    end
    Importance resample from the discrete distribution given by the computed weights.
    Add the model discrepancy associated with the current discipline (δ_j) to the respective importance resampled samples.
    Update the probability density of Y_j (π_{Y_j}) from the updated samples.
    Normalize the unnormalized weights $\{w_{i,j}^{un}\}$ to obtain the normalized weights $\{w_{i,j}\}_{i=1,\ldots,N}$.
    $n_{\text{eff}}^j(M) = \frac{1}{\sum_{i=1}^N (w_{i,j})^2}$
    $\epsilon_j = |\frac{n_{\text{eff}}^j(M)-n_{\text{eff}}^j(M-1)}{n_{\text{eff}}^j(M)}|$
  end
  $M = M+1$
end

B. Convergence Analysis

In this section we discuss the convergence properties of our proposed compositional multidisciplinary uncertainty analysis method. Convergence results from applications of the law of large numbers, Skorokhod's representation theorem, and the convergence of kernel density estimation, importance weighted empirical distributions, and the Gibbs sampler.

In the offline process of the compositional multidisciplinary uncertainty analysis method, we perform

Monte Carlo analysis and generate samples from the input proposal distributions of each discipline. The convergence of these input proposal samples to their respective input proposal distributions follows from the law of large numbers [32]. From the generated input samples, we obtain the output samples independently for each discipline, so the output variances and other distributional quantities can similarly be estimated using Monte Carlo simulation results. By assuming that the function model of each discipline, f_j, is bounded and continuous, then as an application of Skorokhod's representation theorem, the output empirical distribution converges to the true output proposal distribution of each discipline [32].

In the online process of the compositional multidisciplinary uncertainty analysis method, we use importance sampling to weight the proposal samples, so as to approximate the target input distribution, using the samples previously simulated from the proposal input distributions. The convergence of the sampling and importance resampling (SIR) algorithm is discussed in Ref. 23. To apply SIR, we need to estimate the proposal and target densities from the corresponding samples to compute the importance weights. By using a kernel density estimation method that is strongly uniformly convergent [33, 34], we obtain pointwise estimates of the densities where we are ensured to converge pointwise as the number of samples increases to the true density [35]. Then, by an application of Skorokhod's representation theorem [32], the weighted empirical output proposal distribution converges to the true output target distribution. As can be seen in Figure 4, these target distributions are conditional distributions for each coupling variable conditioned on all other coupling variables. Hence, iteratively sampling via the process of re-weighting offline samples constitutes a Gibbs sampling process, which has been shown to converge under certain conditions (i.e., compatible conditional distributions) to samples from the true joint distribution of the coupling variables [36].

Thus, our proposed compositional multidisciplinary uncertainty analysis approach can perform the uncertainty analysis for coupled systems with model discrepancy in a provably convergent manner under mild assumptions on the disciplinary functions and the model discrepancy terms.

V. Application

We present here a demonstration of the effectiveness of our proposed compositional multidisciplinary uncertainty analysis approach. The application system of interest is adapted from Ref. 5. The application is discussed followed by the description of the results.

A. A Simple Aerodynamics-Structures System

The example problem we use to demonstrate our method is a two-dimensional airfoil in airflow from Ref. 5 and shown in Figure 6. As described in Ref. 5, the airfoil is supported by two linear springs attached to a

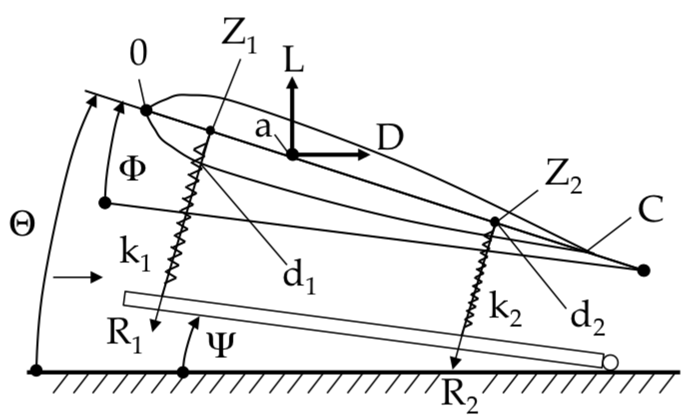

Figure 6: Simple coupled aerodynamics-structures system adapted from Ref. 5.

ramp. The airfoil is permitted to pitch and plunge. The lift, L, and the elastic pitch angle, ϕ, are the coupling variables and also the outputs in this system. A complete description of the problem can be found in Ref. 5 and for the sake of completeness, the equations and variable values are presented in the appendix. A block diagram of the system is shown in Figure 7, where the model discrepancies we consider are highlighted. For this demonstration, we assume $\delta_1 \sim \mathcal{N}(0, 250)$ (N) and $\delta_2 \sim \mathcal{N}(0, 1e-6)$ (rad), where $\mathcal{N}(\mu, \sigma^2)$ represents a normal distribution with mean μ and variance σ^2. Our objective then is to correctly estimate the joint distribution of L and ϕ using our compositional multidisciplinary uncertainty analysis methodology.

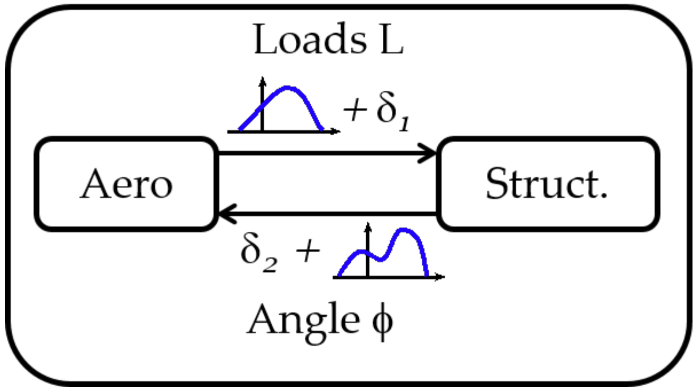

Figure 7: Block diagram of the simple coupled aerodynamics-structures system adapted from Ref. 5 showing the model discrepancies considered. L is the lift and ϕ is the elastic pitch angle.

B. Results

In this section, we present the results of our compositional multidisciplinary uncertainty analysis (CUA) approach applied to the simple aerodynamics-structures system. We present a comparison to brute force Monte Carlo simulation results, the evolution of the joint distribution of L and ϕ using our compositional multidisciplinary uncertainty analysis methodology, and convergence results for both increasing Gibbs iterations and increasing sample size for statistics of interest.

The Monte Carlo results that we compare our methodology to consisted of generating 100,000 samples from the joint distribution of L and ϕ using Gibbs sampling. The Gibbs sampler underwent 15 iterations for each sample, thus the models were evaluated 1,500,000 times each. Our compositional multidisciplinary uncertainty analysis approach used 200,000 offline evaluations of each model independently. For the aerodynamics model of the example problem, an input distribution of $\phi \sim \mathcal{N}(0.017, 4e-6)$ was used as the proposal. For the structures model, the independent input distribution, $L \sim \mathcal{N}(500, 400)$, was used as the proposal. We stress here that these proposals were completely independent (that is, all off-diagonal terms in the covariance matrix of the joint distribution of the variables were zero), and that the offline evaluations of the aerodynamics and structures models were run in isolation of each other. That is, there was never any coupled system evaluation conducted offline.

Statistics of the joint distribution of L and ϕ are presented in Table 1 for both the full Monte Carlo simulation approach (MCS) and the compositional multidisciplinary uncertainty analysis (CUA) approach proposed here. The results demonstrate the mean, variance, covariance, and correlation coefficient of the joint distribution are being estimated well by our approach. The joint distribution of the variables using

Table 1: Quantities of interest computed in MCS and CUA methods

Variable	MCS	CUA
μ_L	502.0265	502.0563
σ_L^2	352.2087	353.0435
μ_ϕ	0.0176	0.0176
σ_ϕ^2	$1.4240e-06$	$1.4318e-06$
$cov(L, \phi)$	0.0122	0.0123
$cor(L, \phi)$	0.5447	0.5476

each method is shown in Figure 8 as contour plots. The left side of the plot is the MCS result whereas the right side is the result of our approach. Graphically it is clear that our method has captured the joint distribution of interest. This was achieved without any online evaluations of either model separately or as a coupled system. Thus, the online cost of using our method to propagate uncertainty through this system was negligible, whereas the MCS approach incurred substantial computational cost.

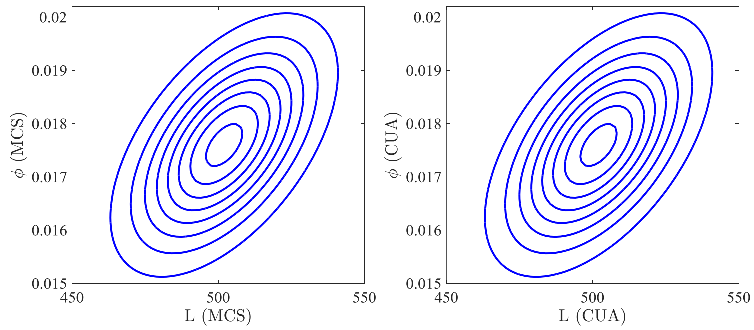

Figure 8: Joint distributions represented with contour plots for both the MCS and CUA methods. The MCS approach used 1,500,000 samples and the CUA approach used 200,000.

Figure 9 shows the evolution of the joint distribution of L and ϕ using our proposed approach after one, three, and five Gibbs iterations. Red samples are generated from the proposal distributions shown as red solid contours which are re-weighted to capture the evolving target distribution, shown as blue dashed contours, and the weighted samples are shown in blue dots in the plots below the distributions.

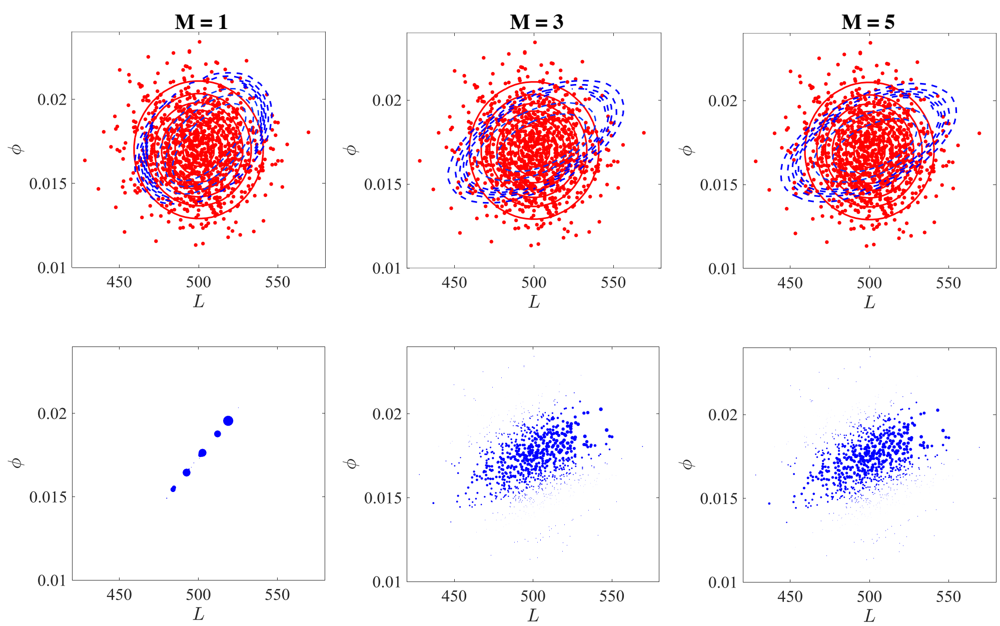

Figure 9: Evolution of the joint distribution of L and ϕ using our proposed approach after one, three, and five Gibbs iterations. Red samples are generated from a proposal distribution shown as red solid contours which are re-weighted (weighted blue dots in below plots) to capture the evolving target distribution, shown as blue dashed contours.

To demonstrate the convergence of the joint distribution of L and ϕ obtained using our compositional methodology, we compute the Cramer von-Mises criterion [37] between our empirical joint distribution and that recovered by brute force Monte Carlo simulation. The Cramer von-Mises criterion is defined as

$$\omega = \int_{-\infty}^{+\infty} [F_n(x) - F^*(x)]^2 \, dF^*(x), \tag{6}$$

where F_n and F^* are the cumulative distribution functions of the empirical distribution and a desired distribution, respectively. Figure 10 presents the Cramer von-Mises criterion as a function of the number of offline samples used in our approach averaged over 100 independent simulations. Also, the effective sample size is shown as a function of the number of offline samples averaged over the 100 independent simulations. The results show that the Cramer von-Mises criterion converges with the number of samples, and the rate of convergence is $1/N$, where N is the number of samples, which is expected based on the dependence of our convergence analysis on the law of large numbers.

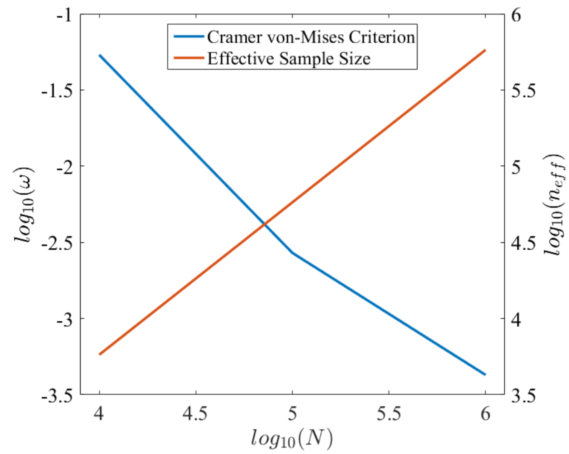

Figure 10: Average Cramer von-Mises criterion and effective sample size over 100 runs of the CUA approach as a function of the number of offline samples.

Certain quantities of interest, such as means, variances, and the covariance of the joint distribution of L and ϕ also converge to their true values as shown in the following numerical results. We compute the expected value of the relative error of the mean and variance estimates as, $\mathbb{E}[\frac{|\hat{\mu}_q - \mu_q|}{\hat{\mu}_q}]$ and $\mathbb{E}[\frac{|\hat{\sigma}_q^2 - \sigma_q^2|}{\hat{\sigma}_q^2}]$ respectively, in which q refers to L or ϕ. The expected value of the relative error of the covariance of L and ϕ is computed as, $\mathbb{E}[\frac{|c\hat{o}v(L,\phi) - cov(L,\phi)|}{c\hat{o}v(L,\phi)}]$. The expectation is computed as the average of 100 independent uncertainty analysis trials. The mean and variance, $\hat{\mu}_q$ and $\hat{\sigma}_q^2$, are obtained using the full Monte Carlo simulation uncertainty analysis results, and μ_q and σ_q^2, are the mean and variance estimates (for L and ϕ) obtained from the compositional multidisciplinary uncertainty analysis approach. To assess the quality of the proposal distributions, we compute the effective sample size once the target distribution is known after each Gibbs iteration. Figure 11 shows the expected error in estimating the covariance of L and ϕ, and also the effective sample size as the number of Gibbs iterations increases. As it is seen, in this particular example,

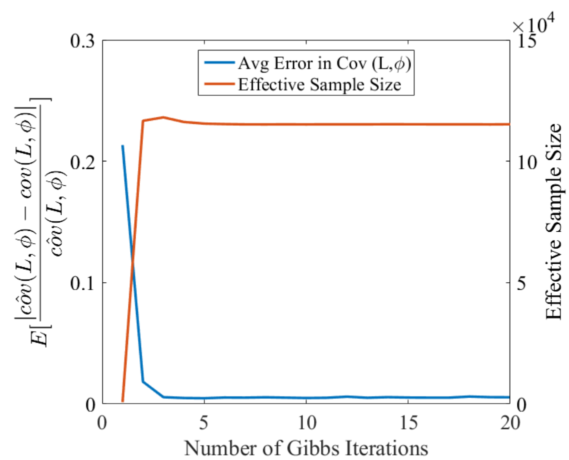

Figure 11: Average effective sample size and the expected value of the relative error of the covariance of L and ϕ using the CUA method with 200,000 samples and averaged over 100 runs versus the number of Gibbs iterations.

the joint distribution is captured after a few Gibbs iterations. As shown in the figure, the effective sample size and relative error in the covariance estimate level off around the same number of Gibbs iterations. As discussed in Section IV, we use the leveling off of the effective sample size as a heuristic to indicate that our approximation of the target distribution is stationary. In this situation, the only mechanism for obtaining a better estimate of the target is the introduction of more samples, which is a topic of future work.

Figure 12 shows the expected value of the relative error of the means and variances of L and ϕ, as well as the effective sample size as the number of offline samples increases. The results show that our proposed approach is effectively estimating the statistics of L and ϕ.

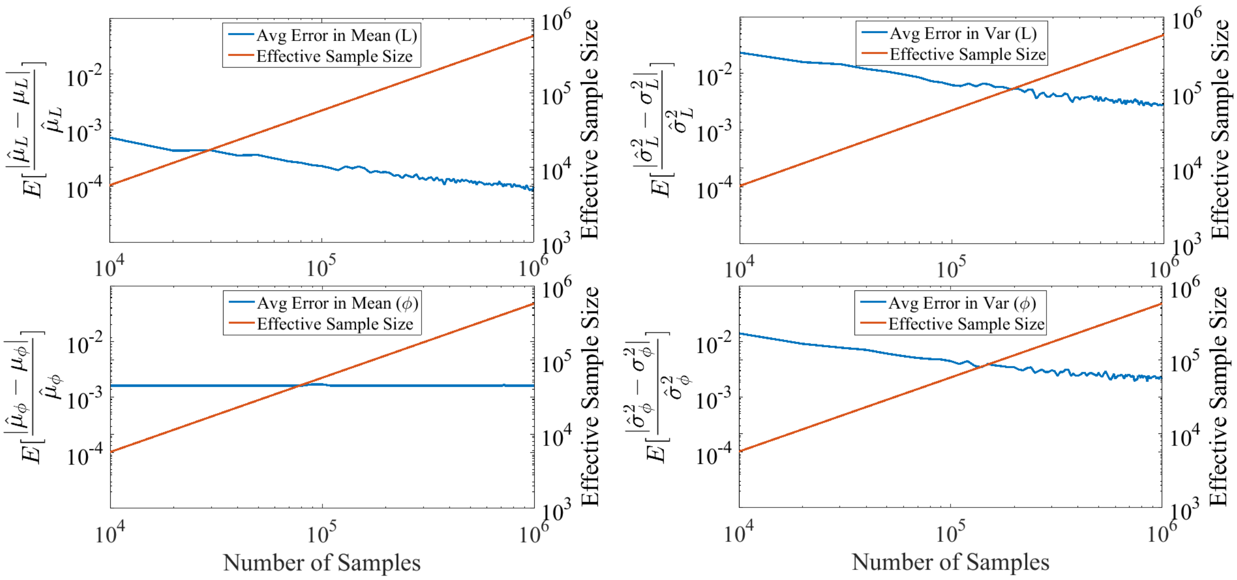

Figure 12: Average effective sample size and expected value of the relative error of the mean and variance of L and ϕ computed using our approach averaged over 100 runs versus number of samples on a logarithmic scale

VI. Conclusion

This paper has presented a compositional multidisciplinary uncertainty analysis methodology. The approach was motivated by the fact that often, multidisciplinary simulation capabilities are composed by integrating pre-existing disciplinary physics-based models. For such composed multi-physics systems, the task of uncertainty analysis can be challenging owing to the disciplinary models being managed by separate entities or housed in separate locations, analysis capabilities running on different computational platforms, models with significantly different analysis run times, and the sheer number of disciplines required for a given analysis. Further, it is also often the case for such systems to exhibit feedback couplings, which can lead to computationally prohibitive expense for uncertainty analysis tasks. Thus, there is a need to alleviate the computational burden of multidisciplinary uncertainty analysis in coupled systems composed of integrated disciplinary models. We have achieved this by moving most of the computational expense of multidisciplinary uncertainty analysis to an offline phase, that can be conducted whenever there is the opportunity to do so. The key advantage of our approach is that there is no need to do system level uncertainty analysis, and all the disciplinary analyses can be performed independently in the offline process. Offline data is used in an online process that does not require any further model evaluations or any system level analysis. We demonstrated our approach on a coupled aerodynamics-structures system with model discrepancies. In future research, we will extend our approach to systems with parametric uncertainty. We will also investigate methods for adaptively introducing new samples online, for situations where proposals are inadequately capturing targets.

Acknowledgment

This work was supported by the AFOSR MURI on multi-information sources of multi-physics systems under Award Number FA9550-15-1-0038, program manager Jean-Luc Cambier.

References

[1] Yao, W., Chen, X., Luo, W., van Tooren, M., and Guo, J., "Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles," *Progress in Aerospace Sciences*, Vol. 47, No. 6, 2011, pp. 450–479.

[2] Kennedy, M. C. and O'Hagan, A., "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 63, No. 3, 2001, pp. 425–464.

[3] Amaral, S., Allaire, D., and Willcox, K., "A decomposition approach to uncertainty analysis of multidisciplinary systems," *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2012.

[4] Amaral, S., Allaire, D., and Willcox, K., "A decomposition-based approach to uncertainty analysis of feed-forward multicomponent systems," *International Journal for Numerical Methods in Engineering*, Vol. 100, No. 13, 2014, pp. 982–1005.

[5] Sobieszczanski-Sobieski, J., "Sensitivity of complex, internally coupled systems," *AIAA journal*, Vol. 28, No. 1, 1990, pp. 153–160.

[6] Martin, J. and Simpson, T., "A methodology to manage system-level uncertainty during conceptual design," *Journal of Mechanical Design*, Vol. 128, No. 4, 2006, pp. 959–968.

[7] Gu, X. S., Renaud, J., and Penninger, C., "Implicit uncertainty propagation for robust collaborative optimization," *Journal of Mechanical Design*, Vol. 128, No. 4, 2006, pp. 1001–1013.

[8] Chiralaksanakul, A. and Mahadevan, S., "Decoupled approach to multidisciplinary design optimization under uncertainty," *Optimization and Engineering*, Vol. 8, No. 1, 2007, pp. 21–42.

[9] McDonald, M., Zaman, K., and Mahadevan, S., "Uncertainty quantification and propagation for multidisciplinary system analysis," *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Paper 2008-6038*, 2008.

[10] Du, X. and Chen, W., "An efficient approach to probabilistic uncertainty analysis in simulation-based multidisciplinary design," *Proceedings of the 8th AIAA Symposium on Multidisciplinary Analysis and Optimization*, 2000.

[11] Jiang, Z., Li, W., Apley, D. W., and Chen, W., "A System Uncertainty Propagation Approach With Model Uncertainty Quantification in Multidisciplinary Design," *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, 2014, pp. V02BT03A029–V02BT03A029.

[12] Kokkolaras, M., Mourelatos, Z. P., and Papalambros, P. Y., "Design optimization of hierarchically decomposed multilevel systems under uncertainty," *Journal of Mechanical Design*, Vol. 128, No. 2, 2006, pp. 503–508.

[13] Du, X. and Chen, W., "Collaborative reliability analysis under the framework of multidisciplinary systems design," *Optimization and Engineering*, Vol. 6, No. 1, 2005, pp. 63–84.

[14] Mahadevan, S. and Smith, N., "Efficient first-order reliability analysis of multidisciplinary systems," *International Journal of Reliability and Safety*, Vol. 1, No. 1, 2006, pp. 137–154.

[15] Sankararaman, S. and Mahadevan, S., "Likelihood-Based Approach to Multidisciplinary Analysis Under Uncertainty," *Journal of Mechanical Design*, Vol. 134, No. 3, 2012, pp. 031008, 1–12.

[16] Arnst, M., Ghanem, R., Phipps, E., and Red-Horse, J., "Dimension reduction in stochastic modeling of coupled problems," *International Journal for Numerical Methods in Engineering*, Vol. 92, No. 11, 2012, pp. 940–968.

[17] Arnst, M., Ghanem, R., Phipps, E., and Red-Horse, J., "Measure transformation and efficient quadrature in reduced-dimensional stochastic modeling of coupled problems," *International Journal for Numerical Methods in Engineering*, Vol. 92, No. 12, 2012, pp. 1044–1080.

[18] Arnst, M., Ghanem, R., Phipps, E., and Red-Horse, J., "Reduced chaos expansions with random coefficientsin reduced-dimensional stochastic modeling of coupled problems," *International Journal for Numerical Methods in Engineering*, Vol. 97, No. 5, 2014, pp. 352–376.

[19] Constantine, P. G., Phipps, E. T., and Wildey, T. M., "Efficient uncertainty propagation for network multiphysics systems," *International Journal for Numerical Methods in Engineering*, Vol. 99, No. 3, 2014, pp. 183–202.

[20] Arnst, M., Craig, S., and Ghanem, R., "Hybrid sampling/spectral method for solving stochastic coupled problems," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 1, No. 1, 2013, pp. 218–243.

[21] Chen, X., Ng, B., Sun, Y., and Tong, C., "A flexible uncertainty quantification method for linearly coupled multi-physics systems," *Journal of Computational Physics*, Vol. 248, No. 1, 2013, pp. 383–401.

[22] Parzen, E., "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, Vol. 33, No. 3, 1962, pp. 1065–1076.

[23] Smith, A. F. and Gelfand, A. E., "Bayesian statistics without tears: a sampling–resampling perspective," *The American Statistician*, Vol. 46, No. 2, 1992, pp. 84–88.

[24] Glynn, P. W. and Iglehart, D. L., "Importance sampling for stochastic simulations," *Management Science*, Vol. 35, No. 11, 1989, pp. 1367–1392.

[25] Kong, A., Liu, J. S., and Wong, W. H., "Sequential imputations and Bayesian missing data problems," *Journal of the American statistical association*, Vol. 89, No. 425, 1994, pp. 278–288.

[26] Liu, J. S., "Metropolized independent sampling with comparisons to rejection sampling and importance sampling," *Statistics and Computing*, Vol. 6, No. 2, 1996, pp. 113–119.

[27] Doucet, A., Godsill, S., and Andrieu, C., "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and computing*, Vol. 10, No. 3, 2000, pp. 197–208.

[28] Lenth, R. V., "Some practical guidelines for effective sample size determination," *The American Statistician*, Vol. 55, No. 3, 2001, pp. 187–193.

[29] Geman, S. and Geman, D., "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, , No. 6, 1984, pp. 721–741.

[30] Schervish, M. J. and Carlin, B. P., "On the convergence of successive substitution sampling," *Journal of Computational and Graphical Statistics*, Vol. 1, No. 2, 1992, pp. 111–127.

[31] Gelfand, A. E. and Smith, A. F., "Sampling-based approaches to calculating marginal densities," *Journal of the American statistical association*, Vol. 85, No. 410, 1990, pp. 398–409.

[32] Grimmett, G. and Stirzaker, D., *Probability and Random Processes*, Third Edition, Oxford University Press, 2001.

[33] Devroye, L. and Penrod, C. S., "The strong uniform convergence of multivariate variable kernel estimates," *Canadian Journal of Statistics*, Vol. 14, No. 3, 1986, pp. 211–220.

[34] Scott, D. W., *Multivariate density estimation: theory, practice, and visualization*, John Wiley & Sons, 2015.

[35] Rubin, D. B. et al., "Using the SIR algorithm to simulate posterior distributions," *Bayesian statistics*, Vol. 3, No. 1, 1988, pp. 395–402.

[36] Casella, G. and George, E. I., "Explaining the Gibbs sampler," *The American Statistician*, Vol. 46, No. 3, 1992, pp. 167–174.

[37] Darling, D. A., "The kolmogorov-smirnov, cramer-von mises tests," *The Annals of Mathematical Statistics*, 1957, pp. 823–838.

Appendix

Aerodynamics-Structures System

Aerodynamics Model

$L = q.S.C_L$
$\theta = \phi + \psi$
$C_L = u\theta + r[1 - \cos[(\pi/2)(\theta/\theta_0)]]$

Structures Model

$R_1 = L/(1+p) \quad R_2 = Lp/(1+p)$
$d_1 = R_1/k_1 \quad d_2 = R_2/k_2$
$\phi = (d_1 - d_2)/[C.(\bar{z}_2 - \bar{z}_1)]$

Data

$\bar{z}_1 = z_1/C \quad \bar{z}_2 = z_2/C \quad \bar{a} = a/C$
$\bar{h}_1 = \bar{a} - \bar{z}_1 \quad \bar{h}_2 = \bar{z}_1 - \bar{a} \quad p = \bar{h}_1/\bar{h}_2$
$S = B.C$
$B = 100\,cm; \quad C = 10\,cm; \quad \bar{z}_1 = 0.2; \quad \bar{z}_2 = 0.7$
$k_1 = 4000\,N/cm; \quad k_2 = 2000\,N/cm$
$\bar{a} = 0.25; \quad q = 1\,N/cm^2; \quad \theta_0 = 0.26\,rad; \quad \psi = 0.05\,rad$
$u = 2\pi; \quad r = 0.9425$