Modeling and quantification of model-form uncertainties in eigenvalue computations using a stochastic reduced model

Charbel Farhat, Adrien Bos, Philip Avery, Christian Soize

Modeling and Quantification of Model-Form Uncertainties in Eigenvalue Computations Using a Stochastic Reduced Model

Charbel Farhat[*], Adrien Bos[†]
Stanford University, Stanford, CA 94305, USA
Philip Avery[‡]
US Army Research Laboratory, Adelphi, MD
and Stanford University, Stanford, CA 94305, USA
and Christian Soize[§]
Université Paris-Est, 77454 Marne-la-Vallee, France

A feasible, nonparametric, probabilistic approach for modeling and quantifying model-form uncertainties associated with a computational model designed for the solution of a generalized eigenvalue problem is presented. It is based on the construction of a Stochastic, Projection-based, Reduced-Order Model (SPROM) associated with a High-Dimensional Model (HDM) using three innovative ideas: the substitution of the deterministic Reduced-Order Basis (ROB) with a stochastic counterpart (SROB) featuring a reduced number of hyperparameters; the construction of this SROB on a subset of a compact Stiefel manifold in order to guarantee the linear independence of its column vectors and the satisfaction of any constraints of interest; and the formulation and solution of a reduced-order inverse statistical problem to determine the hyperparameters so that the mean value and statistical fluctuations of the eigenvalues predicted using the SPROM match target values obtained from available data. Consequently, the proposed approach for modeling model-form uncertainties can be interpreted as an effective approach for extracting from data fundamental information and/or knowledge that are not captured by a deterministic computational model, and incorporating them in this model. Its potential for quantifying model-form uncertainties in generalized eigencomputations is demonstrated for a natural vibration analysis of a small-scale replica of an X-56 type aircraft made of a composite material for which ground ground vibration test data are available.

Nomenclature

ϱ	Vector of problem/model parameters
ϱ_j	j-th scalar problem/model parameter stored in the vector ϱ
m_ϱ	Number of problem/model parameters and therefore dimension of ϱ
ϱ^i	i-th problem/model parameter vector ϱ sampled in the parameter space
m_ϱ^s	Number of sampled problem/model parameter vectors
N	Number of degrees of freedom of a High-Dimensional Model (HDM) and therefore dimension of this HDM
V	$N \times n$ matrix representing a Reduced-Order Basis (ROB)

[*]Vivian Church Hoff Professor of Aircraft Structures, Department of Aeronautics and Astronautics, William F. Durand Building Room 257, Stanford University, Stanford, CA 94305-3035. AIAA Fellow. Email: cfarhat@stanford.edu.

[†]Graduate Student, Department of Aeronautics and Astronautics, William F. Durand Building Room 028, Stanford University, Stanford, CA 94305-3035. Email: abos@stanford.edu.

[‡]US Army Research Laboratory, Adelphi, MD, USA and Department of Aeronautics and Astronautics, William F. Durand Building Room 028, Stanford University, Stanford, CA 94305-3035. Email: pavery@stanford.edu

[§]Professor, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, Université Paris-Est, 5 bd Descartes, 77454 Marne-la-Vallee, France. E-mail: christian.soize@univ-paris-est.fr.

n	Dimension of the ROB V
M	Symmetric mass matrix associated with a high-dimensional Finite Element (FE) structural model
K	Symmetric stiffness matrix associated with a high-dimensional FE structural model
u	Semi-discrete amplitude vector of an assumed harmonic vibration of a structure
ω	Circular frequency associated with u
u_i	i-th natural mode shape or eigenvector of the pencil (K, M)
ω_i^2	Square of the i-th natural circular frequency or i-th eigenvalue associated with u_i
f_i	Natural frequency associated with ω_i
m_u	Number of eigenvalues of interest of the pencil (K, M)
t	Time
$\dot{()}$	Time derivative of the quantity ()
f^{ext}	Semi-discrete force vector
u_0	Semi-discrete initial displacement condition
X	Snapshot matrix
u_r	Vector of generalized coordinates of dimension n associated with the ROB V
ϖ	Circular frequency associated with u_r
$()^T$	Transpose operation
K_r	Symmetric reduced stiffness matrix associated with the ROB V
M_r	Symmetric reduced mass matrix associated with the ROB V
I_n	Identity matrix of size (dimension) n
Q	$N \times N$ metric (orthonormality) matrix for V ($V^T Q V = I_n$)
u_{r_i}	i-th eigenvector of the pencil (K_r, M_r)
ϖ_i^2	i-th eigenvalue associated with u_{r_i}
N_{bc}	Number of displacement constraints in an HDM (which can be zero)
B^T	$N_{bc} \times N$ matrix of displacement constraints satisfying $B^T B = I_{N_{bc}}$
$\mathbf{V}$	Stochastic counterpart of the ROB V
$\mathbf{u}$	Stochastic counterpart of u
$\mathbf{u}_r$	Stochastic counterpart of u_r
$\boldsymbol{\varpi}$	Stochastic counterpart of ϖ associated with $\mathbf{u}_r$
$\mathbf{K}_r$	Stochastic counterpart of K_r
$\mathbf{M}_r$	Stochastic counterpart of M_r
$\mathbf{u}_{r_i}$	Stochastic counterpart of u_{r_i} and i-th eigenvector of the pencil ($\mathbf{K}_r$, $\mathbf{M}_r$)
$\boldsymbol{\varpi}_i^2$	Stochastic counterpart of ϖ_i^2 and i-th eigenvalue associated with $\mathbf{u}_{r_i}$
E	Mathematical expectation of a random variable
$\mathbf{G}$	$N \times n$ non-Gaussian second-order centered random matrix
β	Hyperparameter of $\mathbf{G}$ and the probability model underlying $\mathbf{V}$
σ	Upper triangular matrix of hyperparameters of the probability model underlying $\mathbf{V}$
s	Real number satisfying $s \geq 0$ and hyperparameter of the probability model underlying $\mathbf{V}$
$s_{_M}$	Real number satisfying $s_{_M} > 0$ and hyperparameter of the probability model underlying $\mathbf{V}$
α	Vector-valued hyperparameter of the probability model underlying $\mathbf{V}$ $\left(\alpha = (s, s_{_M}, \beta, \sigma)^T = (\alpha_1, \cdots, \alpha_{m_\alpha})^T\right)$
m_α	Total number of hyperparameters of the probability model underlying $\mathbf{V}$ and therefore dimension of α
o	Vector-valued quantity of interest
o_i	i-th scalar quantity of interest stored in the vector o
m_o	Number of quantities of interest and therefore dimension of o
$\boldsymbol{o}$	Stochastic counterpart of o
J	Cost function for the reduced-order inverse statistical problem governing the identification of α
J_{mean}	Component of J for controlling the identification of α with respect to the mean value
J_{std}	Component of J for controlling the identification of α with respect to the statistical fluctuations
w_J	Weight satisfying $w_J \geq 0$ and defining $J(\alpha) = w_J\, J_{\text{mean}}\,(\alpha) + (1 - w_J)\, J_{\text{std}}\,(\alpha)$
Y	Material Young's modulus
ν	Material Poisson's ratio
ρ	Material density
$\mathbf{Y}$	Stochastic counterpart of Y
$\boldsymbol{\nu}$	Stochastic counterpart of ν
$\boldsymbol{\rho}$	Stochastic counterpart of ρ

I. Introduction

Parametric, projection-based, model order reduction is a mathematical technique for generating a parametric, Projection-based, Reduced-Order Model (PROM) by first carefully constructing a parameter-independent Reduced-Order Basis (ROB), then projecting a parametric High-Dimensional Model (HDM) of interest onto this ROB. Because it is capable of reducing computational costs by orders of magnitude while maintaining a desired level of accuracy, this technique is essential for any scenario where real-time simulation responses are desired.[5] Hence, it is rapidly becoming indispensable for a large number of applications including, among others, computational-based parametric studies and design optimization, multiscale analysis, statistical analysis, model predictive control, simulation-based decision making, and Uncertainty Quantification (UQ).[1–4]

For UQ applications, a PROM is typically viewed or used as a *tool* for accelerating auxiliary Monte Carlo computations. In this paper however which focuses on *model-form* UQ – that is, the modeling and quantification of the uncertainties about the ability of a computational model to truly represent the physics underlying the problem to be solved – the role of a PROM is elevated to that of enabling the extraction from data of information and/or knowledge that are not captured by a deterministic computational model and their infusion in this model. These data can be nonstatistical or statistical experimental data, in which case the UQ performed by the method proposed in this paper pertains to the HDM of interest as well as the PROM constructed for this HDM. It can also be numerical data generated by the HDM, in which case the UQ performed by this method addresses the effects on the fast predictions performed using the PROM of the various sources of model-form uncertainties associated with its construction. Hence, unlike other approaches for UQ that rely on a PROM for accelerating stochastic computations, the method proposed in this paper for modeling and quantifying the effects of model-form uncertainties on HDM computations of interest accounts for the modeling errors associated with the construction of the auxiliary PROM. Such errors are highlighted below, after a brief overview of a typical procedure for constructing a PROM is outlined.

An HDM is often associated with a physics-based, high-fidelity model. In most engineering applications, this model is parameterized, for example, for the purpose of performing parametric studies, design optimization, or certification by numerical simulations. Hence, a parametric HDM typically leads to CPU-intensive computations. Its parameters are represented here by the vector $\varrho = (\varrho_1, \ldots, \varrho_{m_\varrho})^T$, where m_ϱ denotes the total number of parameters. If the HDM of interest has a size N – that is, if it contains N degrees of freedom (dofs) – its projection on a ROB $V \in \mathbb{R}^{N \times n}$ representing a subspace of dimension $n \ll N$ leads to a PROM that has only n *generalized* dofs. Hence, the dimension of the PROM is in this case $n \ll N$. When the ROB V is properly constructed after some knowledge about the response of the physical system represented by the parametric HDM – or ϱ-HDM – has been gained, the PROM obtained by projecting this ϱ-HDM onto this ROB can retain most of the fidelity of this HDM. In general, knowledge about the system response is obtained during a *training* procedure that is performed *offline*. During this procedure, the problem/model parameters represented here by ϱ are sampled at a relatively small number of points ϱ^i in the problem/model parameter space, $i = 1, \cdots, m^s_\varrho$, and a set of problems related to the main problem of interest are solved at these points using the ϱ-HDM in order to compute a set of solution snapshots. Then, these snapshots are compressed using, for example, the Singular Value Decomposition (SVD) algorithm to construct a *global* ROB V. In general, the sampling strategy is designed so that V is reliable in a large region of the problem/model parameter space. The computational complexity of the construction of a PROM scales with the dimension N of the underlyding HDM, in addition to the dimension n of the global ROB V. Hence, for nonlinear problems where a PROM needs to be reconstructed several times – for example, each time the solution is updated – a global PROM does not necessarily enable fast simulations. For this reason, a PROM is usually equipped with a procedure for approximating the projections that are necessary for its construction, whose computational complexity scales only with the small size n of this PROM.[6–9] Such a procedure is known in the literature as hyperreduction. It transforms a nonlinear PROM into a hyperreduced counterpart that guarantees a significant computational speedup, while maintaining as much as possible the desired level of accuracy. It follows that a PROM or Hyperreduced PROM (HPROM) is typically tainted by two different sources or types of model-form uncertainties or errors:

- The model-form uncertainties inherited from the underlying ϱ-HDM, since a PROM is generated by projecting a ϱ-HDM onto a ROB V.

- The various modeling errors introduced by the finite sampling in the problem/model parameter space, the truncation error induced by the projection process, and the approximation errors introduced by the hyperreduction process. All these errors can be treated as model-form uncertainties.

Hence, performing UQ for a PROM can serve two different purposes: certifying the simulation-based decisions it can enable; and, when experimental data are available, enabling, albeit indirectly, a *computationally feasible and reliable* UQ for its underlying HDM.

The focus of this paper is set on structural dynamics problems, for which the Galerkin method is the preferred projection approach, and specifically, on generalized eigenvalue problems arising from structural vibrations. For structural dynamics problems in general, several methods are currently available for analyzing various types of model uncertainties. These include the standard output-predictive error method introduced in Reference,[10] whose drawback is that it does not enable the computational model of interest, in this case a PROM of the underlying ϱ-HDM, to learn from data. They also include the family of parametric probabilistic methods for UQ that is relatively well developed, at least for a reasonably small number of problem/model parameters, for parameter uncertainty – that is, uncertainty associated with inputs to the computational model of interest whose exact values cannot be controlled in physical experiments, or whose values cannot be exactly inferred by statistical methods. Typically, such methods construct prior and posterior stochastic models of uncertain problem/model parameters pertaining, for example, to geometry, boundary conditions and material properties.[11–14] Many of them have been shown to be computationally efficient for both a ϱ-HDM and a class of associated PROMs (for example, see[15,16]). Unfortunately, these methods do not account for either source of model-form uncertainties outlined above.

Alternatively, nonparametric probabilistic approaches for modeling uncertainties induced by more general modeling errors have been developed in References,[17,18] in the context of linear structural dynamics. They have been extended in Reference[19] to a limited class of geometrically nonlinear structural analysis problems. Furthermore, a new nonparametric probabilistic approach based on the construction of a Stochastic PROM (SPROM) was recently introduced in Reference[3] for addressing specifically model-form uncertainties, the broader case of arbitrarily nonlinear structural dynamics problems, and the context of HDMs for which PROMs can be constructed. However, the applicability of this approach to the UQ of generalized eigencomputations is not straightforward. Furthermore, its potential for model-form UQ was highlighted in Reference[3] primarily for modeling the truncation errors inherent to a PROM.

Hence, each of the main objective and main contribution of this paper is two-fold. First, to extend the scope of the nonparametric probabilistic approach introduced in Reference[3] for modeling model-form uncertainties to generalized eigencomputations, as these are important in many fields including structural dynamics. Second, to demonstrate the intrinsic ability of this approach to model and quantify model-form uncertainties associated with an HDM for an application for which experimental data are available. For this purpose, the focus of this paper is set specifically on the discrete, generalized EigenValue Problem (EVP) associated with a given Finite Element (FE) structural dynamics model, and on the modeling and quantification of the model-form uncertainties in such a model for which a PROM can be constructed. In principle, even though the discrete, generalized EVP considered here is a linear EVP, its hyperreduction rather than mere reduction can be beneficiary. This is because similarly to nonlinear deterministic computations, the Monte Carlo computations underlying the proposed method for UQ – and for this matter, most if not all other methods for UQ – incur the repeated constructions of projection operators whose computational costs scale not only with the small dimension n of the PROM, but also with the large dimension N of the underlying ϱ-HDM. However, in order to keep the focus of this paper on its main objective stated above, and because the HDM considered in this paper for demonstration is of a small size, attention is focused on a PROM rather than an HPROM, and the complex discussion of hyperreduction is reported to a followup paper. Furthermore, because the main objective of this work stated above is independent of the parametric aspect of a PROM, and because the proposed approach for model-form UQ is related to that introduced in Reference[3] which was successfully demonstrated for parametric problems, the parametric aspect of a PROM is not illustrated again in the examples considered in this work.

To this effect, the remainder of this paper is organized as follows. In Section II, the discrete, generalized EVP considered in this work is briefly described. In Section III, the construction of a deterministic PROM for this problem is discussed. In Section IV, the mathematical approach introduced in Reference[3] for constructing a Stochastic ROB (SROB) for modeling model-form uncertainties in structural dynamics computations is tailored to the specific case of generalized eigencomputations. The process tailoring involves the introduction in the probability distribution of the SROB of one additional hyperparameter that is ex-

plained and justified. In Section V, the performance of the resulting nonparametric probabilistic approach for performing model-form UQ in generalized eigencomputations is illustrated for a natural vibration analysis problem associated with the mAEWing1 flying wing for which ground vibration test data are available.[23] Finally, conclusions are offered in Section VI.

II. Generalized Eigenvalue Problem in Structural Dynamics

Given two symmetric, FE, mass and stiffness matrices $M \in \mathbb{R}^{N\times N}$ and $K \in \mathbb{R}^{N\times N}$ associated with a high-dimensional FE structural dynamics model and embedding the displacement boundary conditions of this HDM, the corresponding discrete, generalized EVP can be written as

$$Ku = \omega^2 Mu \tag{II.1}$$

where $u \in \mathbb{R}^N$ denotes the semi-discrete amplitude of an assumed harmonic vibration of the structure, and ω^2 is the square of the corresponding circular frequency. For large-scale FE models – that is, for large values of N – the solution of this problem is computationally intensive, particularly when this problem is parameterized and the total number of problem/model parameters m_ϱ is large. Furthermore, interest is typically focused in such cases on the subset of m_u eigenvalues and eigenvectors at the lower end of the spectrum, as the higher end of the spectrum is often associated with mesh frequencies rather than physical ones. Even in such cases, reducing the dimensionality of this problem before solving it in the problem/model parameter space can be attractive for some applications. However, as stated in the introduction of this paper, the reduction of the above problem is considered here primarily for the purpose of modeling and quantifying the model-form uncertainties affecting M, K, and the m_u pairs of eigensolutions (ω_i^2, u_i), $i = 1, \cdots, m_u < N$. These *and the associated natural frequencies* $f_i = \omega_i/2\pi$, $i = 1, \cdots, m_u < N$, are defined in this paper as the Quantities of Interest (QoIs) for problem (II.1).

III. Projection-Based Reduced-Order Model

When N is large, the eigenvectors u_i, $i = 1, \cdots, m_u < N$ obtained from the numerical solution of the discrete, generalized EVP (II.1) are also often used to construct a ROB $W \in \mathbb{R}^{N\times m_u}$ for the reduction of the dimensionality of the same FE model that gave rise to M and K in the first place. However, because the main interest here is in the modeling and quantification of the effect of model-form uncertainties on the computation of the eigensolution pairs (ω_i^2, u_i), $i = 1, \cdots, m_u$, such a ROB W cannot be considered in this case for reducing the dimensionality of problem (II.1). Instead, a global ROB $V \in \mathbb{R}^{N\times n}$, where $m_u \leq n \ll N$, can be built for the fast approximate solution of this problem in the problem/model parameter space using, for example, the Proper Orthogonal Decomposition (POD) and the method of snapshots[20] in the time domain. In this case, the time domain counterpart of problem (II.1)

$$\begin{aligned} M\ddot{u} + Ku &= f^{\text{ext}}(t) \\ u(0) &= u_0 \\ \dot{u}(0) &= v_0 \end{aligned} \tag{III.1}$$

is considered. In Equation (III.1) above, t denotes time, a dot designates a time derivative, and $f^{\text{ext}}(t)$ and u_0 denote the semi-discrete time-dependent force vector and semi-discrete initial condition, respectively. Furthermore, $f^{\text{ext}}(t)$ and/or $\{u_0, v_0\}$ are chosen so that they generate a broadband excitation of the structure. If M and K are parameterized by ϱ, the problem/model parameter space is sampled at m_ϱ^s points ϱ^i using a preferred sampling strategy, problem (III.1) is solved for each sampled parameter point $\varrho^i = (\varrho_1^i, \cdots, \varrho_{m_\varrho}^i)^T$ using a preferred time-integrator, and the computed solutions of problem (III.1) are collected at a chosen sampling time-step Δs and stored in a snapshot matrix $X \in \mathbb{R}^{N\times N_s}$, where N_s denotes the total number of collected solution snapshots. Then, X is compressed using, for example, the SVD method, in which case the left singular vectors associated with the dominant singular values are selected to form the global ROB V.

Next, performing a Galerkin approximation of u – that is, approximating u as

$$u \approx V u_r \tag{III.2}$$

substituting this approximation in problem (II.1) and premultiplying both sides of the algebraic equation defining this problem by V^T, where the superscript T designates the transpose operation, delivers the PROM

for the discrete, generalized EVP

$$\begin{aligned} K_r u_r &= \varpi^2 M_r u_r \\ &\text{where} \\ K_r = V^T K V \qquad &\text{and} \qquad M_r = V^T M V \end{aligned} \tag{III.3}$$

The notation ϖ is used here to differentiate the circular frequencies governed by problem (III.3) from their counterparts governed by the HDM (II.1).

When using SVD for compressing the snapshot matrix X, V can be constructed to be either an orthonormal matrix – that is, satisfying $V^T V = I_n$, where I_n denotes the identity matrix of dimension n – or an M-orthonormal matrix – that is, satisfying $V^T M V = I_n$. In general, the second choice is preferred for structural dynamics applications,[8] unless M is singular, as it is the case when some dofs of the high-dimensional FE model do not have masses. In this case, the first choice ($V^T V = I_n$) is more practical. For all these reasons, both cases of orthonormal and M-orthonormal ROB matrices V are considered in this paper. They are captured by the general property

$$V^T Q V = I_n \tag{III.4}$$

where $Q = M$ for some applications, and $Q = I_N$ for others.

Because it is formulated using a PROM, problem (III.3) can be solved in real time. If V is sufficiently accurate (which in practice means well-trained), the reconstructed eigensolutions

$$(\omega_i^2, u_i) \approx (\varpi_i^2, V u_{r_i}), \qquad i = 1, \cdots, m_u \leq n \ll N$$

where (ϖ_i^2, u_{r_i}) is the i-th eigenpair solution of problem (III.3) can be expected to be very accurate.

IV. Nonparametric Probabilistic Method for Modeling and Quantifying Model-Form Uncertainties

A. Main Idea

Model-form uncertainties, also known as structural uncertainties, typically result from an imperfect modeling of the problem of interest. This can be due, for example, to the lack of knowledge of the true physics underlying this problem or how to model them, or to the lack of attention to modeling details such as bolts, rivets, freeplay, $\cdots$. Such uncertainties affect the ability of the HDM (II.1) to deliver predictive QoIs – that is, values of the QoIs that match reasonably well their experimental counterparts (which themselves can also be tainted by other types of uncertainties). To model and quantify the effects of these model-form uncertainties, an innovative nonparametric probabilistic approach based on projection-based model reduction was developed in Reference.[3] Simply put, the main idea behind this method is two-fold:

- To extract the information missing in a ϱ-HDM from experimental data, when available, and infuse it in this ϱ-HDM or any other computational model derived from it.
- To formulate and implement this process of extraction and infusion using a PROM in the linear case, and an HPROM in the nonlinear case, for the sake of computational feasibility.

This idea addresses both types of model-form uncertainties outlined in Section I, namely, the model-form uncertainties inherent to the ϱ-HDM, and the modeling errors introduced on the way by the model reduction process. If however experimental data are not available, the main objective is reduced to that of modeling and quantifying the modeling errors due to model reduction. In that case, the main idea presented in Reference[3] simplifies to extracting the information missing in the PROM/HPROM from the HDM data (which can always be made available), and infusing it in the constructed PROM/HPROM to adapt it and improve its predictive capability. *In the remainder of this entire section, the word "data" is used to refer to either nonstatistical or statistical experimental or HDM data*, so that both cases where experimental data are available or missing can be discussed simultaneously.

To implement the two-fold idea outlined above, it was proposed in Reference[3] to randomize the ROB V in order to transform the deterministic solution of the problem of interest into a non deterministic one that can adapt to the available data so that it can learn from it. To this effect, it is noted here that throughout

the remainder of this paper, the bold font is used to designate a stochastic quantity. For example, $\mathbf{V}$ denotes the SROB associated with the deterministic ROB V.

However, the randomization of V cannot be performed arbitrarily, in the sense that $\mathbf{V}$ must remain an admissible basis of global modes. Hence, in all events, its columns must remain linearly independent and must satisfy any displacement constraints underlying the ϱ-HDM from which V is constructed. To this effect, it is first pointed out that any set of semi-discrete displacement constraints associated with a high-dimensional FE structural model can always be written as

$$B^T u = 0, \quad B \in \mathbb{R}^{N \times N_{bc}} \quad \text{and} \quad B^T B = I_{N_{bc}} \tag{IV.1}$$

where N_{bc} denotes the total number of semi-discrete displacement constraints. Then, it is also noted that a simple approach for enforcing almost surely the linear independency of the columns of the SROB $\mathbf{V}$ is to preserve almost surely the orthonormality property satisfied by the deterministic ROB V.

Therefore, if $V \in \mathbb{R}^{N \times n}$ is constructed to be a Q-orthonormal deterministic ROB (see Section III and Equation (III.4)) whose columns satisfy the displacement constraints of interest – that is, a matrix which satisfies

$$V^T Q V = I_n \quad \text{and} \quad B^T V = 0 \tag{IV.2}$$

the SROB $\mathbf{V}$ is constructed so that:

- $\mathbf{V} \in \mathbb{R}^{N \times n}$ is a random matrix whose probability distribution is constructed using the Maximum Entropy principle of Information Theory and a stochastic representation of random fields.
- The support of this probability distribution is the subset of real $N \times n$ matrices satisfying the constraints

$$\mathbf{V}^T Q \, \mathbf{V} = I_n \quad \text{and} \quad B^T \mathbf{V} = 0 \tag{IV.3}$$

  almost surely.
- The probability distribution of $\mathbf{V}$ depends on a vector-valued hyperparameter $\alpha = (\alpha_1, \ldots, \alpha_{m_\alpha})$, where the dimension m_α is sufficiently small to enable the identification of α from the solution of a reduced-order inverse statistical problem. This problem is designed to minimize the discrepancies between the mean value and statistical fluctuations of the QoIs predicted using the SPROM constructed using $\mathbf{V}$, and their counterparts based on the available data.
- The statistical fluctuations of $\mathbf{V}$ happen around its deterministic counterpart V.

The general procedure for constructing an SROB $\mathbf{V}$ that meets the above requirements was first presented in Reference.[3] It is described in simplified form in Section IV.C.1 of this paper in order to keep it self-contained. The selection of the hyperparameters governing the probability distribution of this SROB is tailored in Section IV.C.2 to the context of the discrete, generalized EVP (II.1). The approach for identifying the vector-valued hyperparameter $\alpha = (\alpha_1, \ldots, \alpha_{m_\alpha})^T$ so that the SROB constructed in Section IV.C.1 can reproduce the mean value and statistical fluctuations of the QoIs computed from the available data is presented in Section IV.D. However, the SPROM to be built using the constructed SROB $\mathbf{V}$ and the Galerkin projection procedure is first presented in Section IV.B below.

B. Stochastic Projection-Based Reduced-Order Model

The stochastic counterpart of the approximation (III.2) can be written as

$$\mathbf{u} \approx \mathbf{V}\mathbf{u}_r$$

where $\mathbf{u}$ and $\mathbf{u}_r$ are the stochastic counterparts of u and u_r, respectively, and therefore represent the stochastic solution and its generalized coordinates in the basis $\mathbf{V}$, respectively. Substituting this approximation in problem (II.1) and premultiplying both sides of the algebraic equation defining this problem by $\mathbf{V}^T$ leads to the following SPROM for the discrete, generalized EVP of interest

$$\begin{gathered} \mathbf{K}_r \mathbf{u}_r = \boldsymbol{\varpi}^2 \mathbf{M}_r \mathbf{u}_r \\ \text{where} \\ \mathbf{K}_r = \mathbf{V}^T K \mathbf{V} \quad \text{and} \quad \mathbf{M}_r = \mathbf{V}^T M \mathbf{V} \end{gathered} \tag{IV.4}$$

Because it is formulated using an SPROM of dimension $n \ll N$, the stochastic problem (IV.4) can be rapidly solved to compute the realizations of the stochastic QoIs

$$(\boldsymbol{\omega}_i^2, \mathbf{u}_i) \approx (\boldsymbol{\varpi}_i^2, \mathbf{V}\mathbf{u}_{r_i}), \qquad i = 1, \cdots, m_u \leq n \ll N$$

where $(\boldsymbol{\varpi}_i^2, \mathbf{u}_{r_i})$ is the i-th eigenpair solution of problem (IV.4).

C. Stochastic Reduced-Order Basis

1. *Construction Procedure*

From the first constraint equation in (IV.3), it follows that the $N \times n$ SROB $\mathbf{V}$ must be constructed on the compact Stiefel manifold denoted here by $\mathbb{S}_{N,n}$. From the second constraint equation in (IV.3), it follows that this SROB must be constructed more precisely on the subset of $\mathbb{S}_{N,n}$ associated with the constraint $B^T\mathbf{V} = 0$. This can be done using the procedure summarized below and graphically depicted in Figure 1.

- Generate a centered random matrix $\mathbf{U} \in \mathbb{R}^{N\times n}$ (see the corresponding numerical procedure outlined below)
- Transform $\mathbf{U}$ into a random matrix $\mathbf{A}$ that satisfies the second constraint equation in (IV.3)

  $$\mathbf{A} = \mathbf{U} - BB^T\mathbf{U} \tag{IV.5}$$

  Indeed, from the above definition of $\mathbf{A}$ and (IV.1), it follows that

  $$B^T\mathbf{A} = B^T\mathbf{U} - B^TBB^T\mathbf{U} = B^T\mathbf{U} - B^T\mathbf{U} = 0 \tag{IV.6}$$

- Transform $\mathbf{A}$ into a random matrix $\mathbf{Z}$ that satisfies

  $$B^T\mathbf{Z} = 0 \tag{IV.7}$$

  and lies on the tangent space to $\mathbb{S}_{N,n}$ at V

  $$\mathbf{Z} = \mathbf{A} - \frac{1}{2}V\left(V^TQ\mathbf{A} + \mathbf{A}^TQV\right) \tag{IV.8}$$

  Indeed, from the above definition of $\mathbf{Z}$, the result (IV.6) and the second constraint equation in (IV.2), it follows that

  $$B^T\mathbf{Z} = B^T\mathbf{A} - \frac{1}{2}B^TV\left(V^TQ\mathbf{A} + \mathbf{A}^TQV\right) = 0 - 0 = 0$$

  Also, from (IV.8), it follows that $\mathbf{Z}$ can also be written as

  $$\mathbf{Z} = \frac{1}{2}V(V^TQ\mathbf{A} - \mathbf{A}^TQV) + (I_N - VV^TQ)\mathbf{A} \tag{IV.9}$$

  where, from the first constraint equation in (IV.2), it follows that $V^TQ(I_N - VV^TQ)\mathbf{A} = V^TQ\mathbf{A} - V^TQ\mathbf{A} = 0$. This shows that (IV.9) expresses $\mathbf{Z}$ as the sum of one component in the subspace generated by V and one component in the subspace that is Q-orthogonal to V, and therefore $\mathbf{Z}$ lies on the tangent space to $\mathbb{S}_{N,n}$ at V.

- Transform $\mathbf{Z}$ into the matrix $\mathbf{V}$ which satisfies both constraint equations in (IV.3) and therefore lies on $\mathbb{S}_{N,n}$

  $$\mathbf{V} = (V + s\mathbf{Z})\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2} \tag{IV.10}$$

  where $0 \leq s \leq 1$ can be treated as a hyperparameter of the probability model underlying $\mathbf{V}$.

  Indeed, from the symmetry property of the mass matrix Q, the first constraint equation in (IV.2) and the definition of $\mathbf{Z}$ given in (IV.8), it follows that

  $$\begin{aligned}
  \mathbf{V}^TQ\,\mathbf{V} &= \left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2}(V^T + s\mathbf{Z}^T)Q\,(V + s\mathbf{Z})\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2} \\
  &= \left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2}\left(I_n + s\underbrace{(V^TQ\mathbf{Z} + \mathbf{Z}^TQV)}_{0} + s^2\mathbf{Z}^TQ\mathbf{Z}\right)\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2} \\
  &= I_n
  \end{aligned}$$

Also, from the definition of $\mathbf{V}$ given in (IV.10), the second constraint equation in (IV.2) and (IV.7), it follows that

$$B^T\mathbf{V} = (B^TV + sB^T\mathbf{Z})\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2} = (0+0)\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2} = 0$$

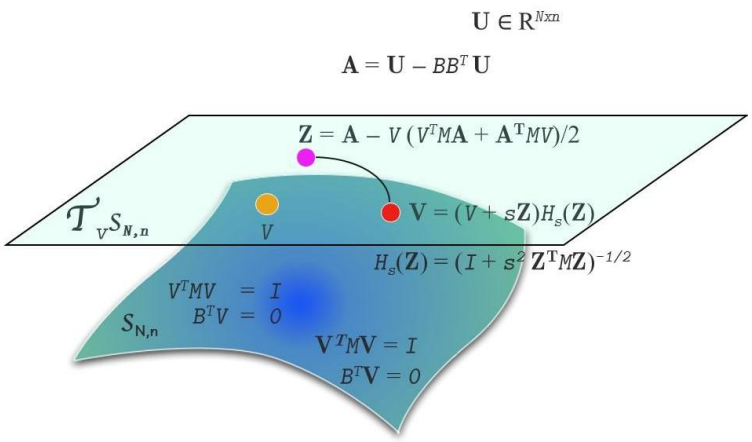

Figure 1. Construction of the SROB $\mathbf{V}$ on a subset of a compact Stiefel manifold $\mathbb{S}_{N,n}$ (construction procedure shown with $Q = M$).

Note that the definition (IV.10) given above shows that:

- The statistical fluctuations of $\mathbf{V}$ happen around its deterministic counterpart V, as desired. In particular, for $s = 0$, $\mathbf{V} = V$. However, it is noted that in general, $E(\mathbf{V}) \neq V$ (see[3]).
- The real number s controls the level of fluctuations of $\mathbf{V}$ in the tangent plane to $\mathbb{S}_{N,n}$, around $V \in \mathbb{S}_{N,n}$.

Note also that whereas the matrices $\mathbf{U}$, $\mathbf{A}$, and $\mathbf{Z}$ leading to the construction of $\mathbf{V}$ are random matrices, the construction process of $\mathbf{V}$ itself is deterministic. In particular, no matter how large is the excursion of $V + s\mathbf{Z}$ in the tangent plane to $\mathbb{S}_{N,n}$ around $V \in \mathbb{S}_{N,n}$, the resulting matrix $\mathbf{V} = (V + s\mathbf{Z})\left(I_n + s^2\mathbf{Z}^TQ\mathbf{Z}\right)^{-1/2}$ lies on the manifold $\mathbb{S}_{N,n}$. Hence, for all these reasons, the SROB $\mathbf{V}$ defined in (IV.10) satisfies all requirements formulated in Section IV.A.

2. Random Matrix $\mathbf{U}$ *and Associated Hyperparameters*

From (IV.8) and (IV.10), it follows that if $\mathbf{A} = 0$, then $\mathbf{Z} = 0$ and $\mathbf{V} = V$. This suggests that, consistently with the idea that the statistical fluctuations of $\mathbf{V}$ should occur around its deterministic counterpart V, $\mathbf{A}$ should be a centered random matrix ($E[\mathbf{A}] = 0$, where E designates the mathematical expectation of a random variable). From (IV.5), it follows that $\mathbf{U}$ should also be a centered random matrix – that is,

$$E[\mathbf{U}] = 0 \tag{IV.11}$$

The covariance tensor $\mathbb{C}_{jkj'k'} = E[\mathbf{U}_{jk}\mathbf{U}_{j'k'}]$ is a fourth-order symmetric tensor with $Nn(Nn+1)/2 = \mathcal{O}(N^2n^2)$ coefficients that can be considered as parameterization coefficients. As already mentioned in Section IV.A, part of the idea here is to use such coefficients as hyperparameters and determine them so that the mean value and statistical fluctuations of the eigenvalue solutions of the SPROM (IV.4) match target values obtained from available data. However, working for this purpose with $\mathcal{O}(N^2n^2)$ hyperparameters is not computationally feasible. Hence, in order to reduce its number of hyperparameters, the random matrix $\mathbf{U}$ is constructed so that

$$\mathbb{C}_{jkj'k'} = E[\mathbf{U}_{jk}\mathbf{U}_{j'k'}] = C_N(\beta)_{jj'}c_{n_{kk'}} \tag{IV.12}$$

where β is a hyperparameter, $C_N(\beta) \in \mathbb{R}^{N\times N}$ controls the correlations between the components of each column vector of $\mathbf{U}$, and the matrix $c_n \in \mathbb{R}^{n\times n}$, which is a Symmetric Positive Definite (SPD) matrix, controls the correlations between the column vectors of $\mathbf{U}$. For a given application, the construction of

$C_N(\beta)$ is given by a mathematical process whose description can be found in Appendix D of Reference[3]. According to the Cholesky decomposition, c_n can be written as

$$c_n = \sigma^T \sigma \tag{IV.13}$$

where $\sigma \in \mathbb{R}^{n \times n}$ is an upper triangular matrix. Indeed, this approach for constructing $\mathbf{U}$ reduces the number of hyperparameters from $Nn(Nn+1)/2$ to $1 + n(n+1)/2 = \mathcal{O}(n^2)$.

To this effect, it is shown in Reference[3] that both objectives $E[\mathbf{U}] = 0$ and (IV.12) can be achieved by constructing a random matrix $\mathbf{G} \in \mathbb{R}^{N \times n}$ that: is centered ($E[\mathbf{G}] = 0$); is computed by semi-discretizing a random field defined at the continuous level of the governing equations of the problem of interest; verifies $E[\mathbf{G}_{jk}\mathbf{G}_{j'k'}] = C_N(\beta)_{jj'}\delta_{kk'}$; and generates the random matrix $\mathbf{U}$ as follows

$$\mathbf{U} = \mathbf{G}(\beta)\sigma \tag{IV.14}$$

Given that s in (*IV*.10) is also a hyperparameter, it follows that the probability model underlying $\mathbf{V}$ defined so far has $2 + n(n+1)/2$ hyperparameters: s, β, and the coefficients of the upper triangular matrix σ. In the remainder of this paper, these parameters of prior distribution are referred to as the *main* hyperparameters.

3. Additional Hyperparameter for Reducing the Bias of the Mean Value

For computing the lowest eigenvalues, it is sometimes difficult to identify optimal values of the main hyperparameters described in Section IV.C.2 – that is, the values of these parameters that sufficiently reduce the bias between the mean value $E\{\boldsymbol{\varpi}_j^2\}$ estimated using the SPROM (IV.4) and the corresponding value extracted from the available data. To address this issue, an additional, strictly positive hyperparameter $s_{_M} > 0$ belonging to a real interval in the neighborhood of 1 is introduced here as follows. Specifically, each stochastic QoI $\boldsymbol{\varpi}_j^2$ in the SPROM (IV.4) is re-written as

$$\boldsymbol{\varpi}_j^2 = s_{_M}\widehat{\boldsymbol{\varpi}}_j^2, \qquad j = 1, \ldots, m_u \tag{IV.15}$$

and the statistics of the random eigenvalues are constructed for $\widehat{\boldsymbol{\varpi}}_j^2$ instead of $\boldsymbol{\varpi}_j^2$. Noting that $E\{\boldsymbol{\varpi}_j^2\} = s_{_M}E\{\widehat{\boldsymbol{\varpi}}_j^2\}$, it follows that the purpose of the hyperparameter $s_{_M}$ is to enable a better global fit of the mean values $\{E\{\boldsymbol{\varpi}_j^2\}, j = 1, \cdots, m_u\}$.

Hence, for the class of eigenvalue problems considered in this paper, the hyperparameters governing the probability distribution of the SROB $\mathbf{V}$ are:

- The hyperparameter s which is a real number and controls the level of fluctuations of $\mathbf{V}$ in the tangent plane to $\mathbb{S}_{N,n}$, around $V \in \mathbb{S}_{N,n}$ (see (IV.10)).
- The hyperparameter β controlling the matrix $C_N(\beta)_{jj'}$ (see (IV.12)).
- The $n(n+1)/2$ hyperparameters of the matrix σ (see (IV.14)).
- The hyperparameter $s_{_M}$ introduced above (see (IV.15)).

These parameters of prior distribution are collected here in the vector-valued hyperparameter

$$\alpha = (s, s_{_M}, \beta, \sigma)^T \tag{IV.16}$$

In principle, the upper triangular matrix σ (IV.13) can be sparsified. Thus, an upper bound for the total number of hyperparameters described above is $m_\alpha^{ub} = 3 + n(n+1)/2$, and a lower bound is $m_\alpha^{lb} = 4$. Consequently, the dimension of the vector-valued hyperparameter (IV.16) satisfies in general

$$4 \leq m_\alpha \leq 3 + n(n+1)/2$$

D. Identification of the Hyperparameters

As already stated in Section IV.A, the identification of the vector-valued hyperparameter α is to be achieved so that the mean value and statistical fluctuations of the QoIs predicted using the SPROM match target values obtained using the available data. This can be performed using the maximum likelihood method, or a nonlinear Least-Squares (LS) method for the QoIs. For example, a nonlinear LS method can be formulated as described below, for both types of model-form uncertainties outlined in Section I.

Let

$$o(\varrho) = (o_1(\varrho), \ldots, o_j(\varrho), \ldots, o_{m_o}(\varrho))^T$$

denote the vector QoI (or system observation). For example, for problem (II.1), this vector can be $o(\varrho) = (\omega_1^2(\varrho), \ldots, \omega_j^2(\varrho), \ldots, \omega_{m_u}^2(\varrho))^T$ ($m_o = m_u \leq n \ll N$). For problem (III.3), it can be $o(\varrho) = (\varpi_1^2(\varrho), \ldots, \varpi_j^2(\varrho), \ldots, \varpi_{m_u}^2(\varrho))^T$ ($m_o = m_u \leq n \ll N$), and for problem (IV.4), it can be $\boldsymbol{o}(\varrho) = (\boldsymbol{\varpi}_1^2(\varrho), \ldots, \boldsymbol{\varpi}_j^2(\varrho), \ldots, \boldsymbol{\varpi}_{m_u}^2(\varrho))^T$ ($m_o = m_u \leq n \ll N$).

Let also $J(\alpha)$ be the cost function defined as

$$J(\alpha) = w_J \, J_{\text{mean}}\,(\alpha) + (1 - w_J)\, J_{\text{std}}\,(\alpha) \tag{IV.17}$$

where w_J is a weight satisfying $0 \leq w_J \leq 1$, and $J_{\text{mean}}\,(\alpha)$ and $J_{\text{std}}\,(\alpha)$ are two functions constructed for controlling the identification of α with respect to the mean value and statistical fluctuations, respectively. Specifically, $J_{\text{mean}}\,(\alpha)$ and $J_{\text{std}}\,(\alpha)$ are defined as

$$\begin{aligned} J_{\text{mean}}\,(\alpha) &= \frac{1}{c_{\text{mean}}(\varrho^1, \ldots, \varrho^{m_\varrho^s})} \sum_{i=1}^{m_\varrho^s} ||o^{\text{ref}}(\varrho^i) - E[\boldsymbol{o}(\varrho^i, \alpha)]||^2 \\ \text{and} & \\ J_{\text{std}}\,(\alpha) &= \frac{1}{c_{\text{std}}(\varrho^1, \ldots, \varrho^{m_\varrho^s})} \sum_{i=1}^{m_\varrho^s} ||v^{\text{ref}}(\varrho^i) - \mathbf{v}(\varrho^i, \alpha)||^2 \end{aligned} \tag{IV.18}$$

where:

- m_ϱ^s denotes as before the number of sampled parameter points ϱ^i in the parameter space.
- E denotes as before the mathematical expectation.
- The definition of o^{ref} depends on the type of available data, which dictates the types of model-form uncertainties that can be accounted for by the proposed approach for modeling model-form uncertainties:
  - If only high-dimensional data – that is, data generated using a ϱ-HDM – is available, o^{ref} corresponds to the QoIs (eigenvalue solutions) of the ϱ-parametric problem (II.1) formulated using an HDM. In this case, only the errors/uncertainties due to model reduction are taken into account by the proposed approach for modeling model-form uncertainties.
  - If experimental data are available:
    - If these data are nonstatistical experimental data, then

      $$o^{\text{ref}} = o^{\text{exp}} \tag{IV.19}$$

      where o^{exp} is extracted from these available data.
    - If these data are statistical experimental data, then

      $$o^{\text{ref}} = E[\boldsymbol{o}^{\text{exp}}] \tag{IV.20}$$

      where $\boldsymbol{o}^{\text{exp}}$ is extracted from these available data.

    In either case of experimental data described above, both of the errors/uncertainties due to model reduction and the uncertainties inherited from the ϱ-parametric HDM are taken into account by the proposed approach for modeling model-form uncertainties.
- $v^{\text{ref}}(\varrho^i) = (v_1^{\text{ref}}(\varrho^i), \ldots, v_j^{\text{ref}}(\varrho^i), \ldots, v_{m_o}^{\text{ref}}(\varrho^i))^T$, $i = 1, \ \cdots, \ m_\varrho^s$, is defined as follows:

– If only high-dimensional data is available, then

$$v_j^{\text{ref}}(\varrho^i) = \gamma\,|o_j^{\text{ref}}(\varrho^i) - o_j(\varrho^i)|, \qquad j = 1, \dots m_o$$

where $\gamma > 0$ is a tunable control parameter, $o_j^{\text{ref}}(\varrho^i) = \omega_j^2(\varrho^i)$ and $o_j(\varrho^i) = \varpi_j^2(\varrho^i)$.

– If nonstatistical experimental data are available, then

$$v_j^{\text{ref}}(\varrho^i) = \gamma\,|o_j^{\text{ref}}(\varrho^i) - o_j(\varrho^i)|, \qquad j = 1, \dots m_o \tag{IV.21}$$

where, again, $\gamma > 0$ is a tunable control parameter and $o_j(\varrho^i) = \varpi_j^2(\varrho^i)$, but $o_j^{\text{ref}}(\varrho^i) = o_j^{\text{exp}}(\varrho^i)$ (from (IV.19)).

– If statistical experimental data are available, then

$$v_j^{\text{ref}}(\varrho^i) = \sqrt{E\left(\boldsymbol{o}_j^{\text{exp}}(\varrho^i)^2\right) - \left[E\left(\boldsymbol{o}_j^{\text{exp}}(\varrho^i)\right)\right]^2}, \qquad j = 1, \dots m_o \tag{IV.22}$$

- $\mathbf{v}(\varrho^i, \alpha) = \left(\mathbf{v}_1(\varrho^i, \alpha), \dots, \mathbf{v}_{m_o}(\varrho^i, \alpha)\right)^T$, where

$$\mathbf{v}_j(\varrho^i, \alpha) = \sqrt{E\left(\boldsymbol{o}_j^2(\varrho^i, \alpha)\right) - \left[E\left(\boldsymbol{o}_j(\varrho^i, \alpha)\right)\right]^2}, \qquad j = 1, \dots m_o$$

- The positive constants $c_{\text{mean}}(\varrho^1, \dots, \varrho^{m_\varrho^s})$ and $c_{\text{std}}(\varrho^1, \dots, \varrho^{m_\varrho})$ are defined as

$$c_{\text{mean}}(\varrho^1, \dots, \varrho^{m_\varrho^s}) = \sum_{i=1}^{m_\varrho^s} \|o^{\text{ref}}(\varrho^i)\|^2, \qquad c_{\text{std}}(\varrho^1, \dots, \varrho^{m_\varrho^s}) = \sum_{i=1}^{m_\varrho^s} \|v_j^{\text{ref}}(\varrho^i)\|^2 \tag{IV.23}$$

It follows that in the definitions (IV.17) and (IV.18) of the cost function $J(\alpha)$, o^{ref} is the target for the mean value with a weight w_J, and v^{ref} is the target for the standard deviation with a weight $1 - w_J$.

The identification of the vector-valued hyperparameter α consists in calculating α^{opt} such that

$$\alpha^{\text{opt}} = \arg\min_{\alpha} J(\alpha) \tag{IV.24}$$

V. Application

The main purpose of this section is to demonstrate the intrinsic ability of the nonparametric probabilistic method described in this paper to model and quantify the model-form uncertainties associated with generalized eigencomputations performed using a FE HDM for which a PROM can be constructed, or the PROM itself. Specifically, the problem of reliably predicting the first few natural modes of the mAEWing1 flying wing made of a composite material and fabricated at the University of Minnesota[24] is considered.

For this purpose, a nonparametric HDM is constructed because the mAEWing1 aircraft has already been built. The performance of the proposed nonparametric probabilistic method is assessed for two different scenarios:

- A hypothetical scenario designed for demonstrating the ability of the nonparametric probabilistic method proposed in this paper for modeling and quantifying the effects of model-form uncertainties to perform well, even in the presence of problem/model parameter errors.
- A real scenario focused on ground vibration test data to demonstrate the ability of this proposed method to perform well with real experimental data.

All offline computations reported herein are performed using the FE structural analyzer AERO-S.[21,22] All online computations based on PROMs and SROMs are performed using MATLAB©.

A. The mAEWing1 Flying Wing

mAEWing1 is the name of a small-scale replica of an X-56 type aircraft made of a composite material and fabricated at the University of Minnesota (see Figure 2). It is a 10-ft span, low-speed flying wing mainly composed of a solid spar and a wing form. The solid spar has a foam core surrounded by composite laminated reinforcements. The foam core is made of extruded polystyrene (XPS). The reinforcement is a 3-layer carbon fiber laminate with the ply orientations $[45/0/0]_T$ in the form of a plane weave with 3K tow. The composite matrix is made of a laminating epoxy.

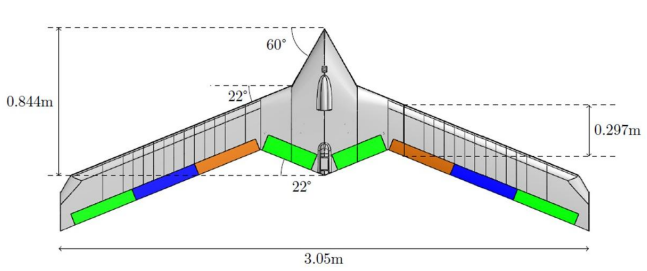

Figure 2. mAEWing1 aircraft: a 10-ft span, low-speed flying wing aircraft (courtesy of the University of Minnesota).

B. Finite Element Structural Model

Based on a NASTRAN FE model originally developed by Virginia Tech, a deterministic, undamped, *unrestrained*, AERO-S, FE, structural dynamics, "stick" model of the mAEWing1 aircraft is constructed using an assembly of: flexible and rigid beam elements with homogenized, isotropic, linear, elastic material properties; revolute-joint-spring combination elements with 6 dofs/node; discrete masses; and rigid, massless, "phantom" shell elements for the purpose of transferring the flow-induced loads from a CFD (Computational Fluid Dynamics) model to this FE structural model when it is used for aeroelastic computations. This FE model, which is graphically depicted in Figure 3, contains 4,146 dofs. It is riddled with model-form uncertainties due to, among other things, the stick and lumped mass modeling, the homogenized representation of the composite material, and the lack of accounting for damping.

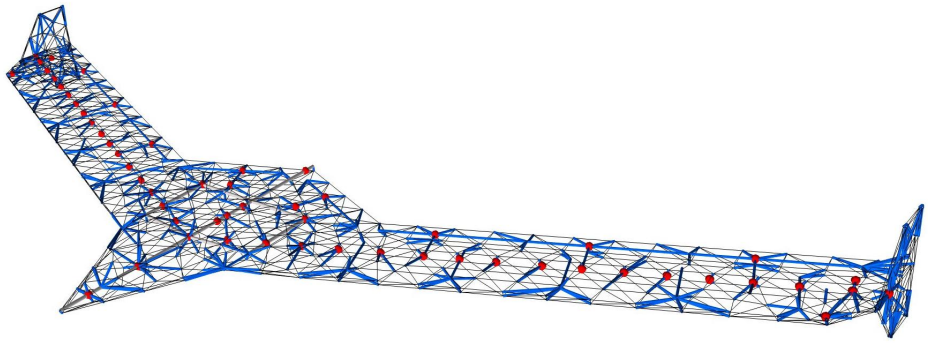

Figure 3. Undamped, FE, structural dynamics, stick model of the mAEWing1 aircraft: blue tubes designate multipoint constraints associated with rigid beam elements and red spheres designate locations of discrete masses and moments of inertia.

Five flexible natural modes of two supposedly identical mAEWing1 flying wings were identified during Ground Vibration Tests (GVTs). These are discussed in Reference[23] and in Section V.D of this paper. Here, the first seven nonzero natural frequencies of the deterministic FE structural model described above are computed ($m_u = 7$) and reported in Table 1. Five of these can be compared with the experimental data reported in Table 2 and organized in two cases associated with the two supposedly identical flying wings. The other two natural frequencies illustrate the need for predictive eigenmode computations given that they were not identified during GVT.

Table 1. mAEWing1 aircraft: natural frequencies predicted using the deterministic FE structural model (mode numbering starts at the first flexible mode).

Flexible mode	1	2	3	4	5	6	7
Frequency f (Hz)	7.94	10.29	15.21	19.71	30.16	32.54	39.21

From the results reported in Table 1 and Table 3, the reader can observe relative discrepancies between

the deterministically predicted natural frequencies and their experimentally identified counterparts. Some of these relative discrepancies are as high as 26%, which highlights the effects of model-form uncertainties. The reader can also observe that the GVTs appear to have missed in one case the third and fifth flexible modes whose natural frequencies are estimated at 15.21 Hz and 30.16 Hz, respectively, and in the other case the second and fifth flexible modes whose natural frequencies are estimated at 10.29 Hz and 30.16 Hz, respectively. These misses stress the importance of complementing tests with predictive analyses.

C. Surrogate Experimental Data

As stated above, a hypothetical scenario is first considered for this application in order to demonstrate the ability of the proposed method, which is designed to model and quantify the effects of model-form uncertainties, to perform well even for problem/model parameter uncertainties. For this purpose, surrogate experimental data are generated here by introducing perturbations in the material properties of the FE structural model outlined in Section V.B. Specifically, the material Young's modulus Y, Poisson's ratio ν, and density ρ are perturbed to enable a statistical analysis of the QoIs.

From the deterministic viewpoint, Y and ν are two independent material properties. However, it was shown in Reference[25] that for a homogeneous, isotropic, linear, elastic material:

- The random variables $\mathbf{Y}$ and $\boldsymbol{\nu}$ associated with Y and ν must be considered as dependent variables. Hence, given that no information on their statistical dependence is usually available *a priori*, Information Theory cannot be applied to construct their probability models, as it would incorrectly lead to two independent random variables $\mathbf{Y}$ and $\boldsymbol{\nu}$.
- However, the Lamé coefficients λ and μ can be modeled using two independent random variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, as long as their probability distributions are constructed using Information Theory as appropriate gamma distributions[26] – that is, as long as the fluctuations in $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are generated as

$$\begin{aligned} \boldsymbol{\lambda} &\sim \Gamma(k_\lambda, \theta_\lambda), \text{ where } k_\lambda = 1 - \zeta_{Lame} \text{ and } \theta_\lambda = \frac{E[\boldsymbol{\lambda}]}{1-\zeta_{Lame}} \\ \boldsymbol{\mu} &\sim \Gamma(k_\mu, \theta_\mu), \text{ where } k_\mu = 1 - \zeta_{Lame} \text{ and } \theta_\mu = \frac{E[\boldsymbol{\mu}]}{1-\zeta_{Lame}} \end{aligned}$$

  and $\zeta_{Lame} \in]-\infty; 1/5[$ is a parameter introduced to control the level of the statistical fluctuations of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$.

Noting that Y and ν are related to λ and μ by

$$Y = \frac{2\mu(3\lambda - 2\mu)}{(\lambda - 2\mu)} \quad \text{and} \quad \nu = \frac{\lambda - 2\mu}{2\lambda} \quad \Longrightarrow \quad \mathbf{Y} = \frac{2\boldsymbol{\mu}(3\boldsymbol{\lambda} - 2\boldsymbol{\mu})}{(\boldsymbol{\lambda} - 2\boldsymbol{\mu})} \quad \text{and} \quad \boldsymbol{\nu} = \frac{\boldsymbol{\lambda} - 2\boldsymbol{\mu}}{2\boldsymbol{\lambda}} \quad \text{(V.1)}$$

it follows that the mean values and statistical fluctuations of the two dependent random variables $\mathbf{Y}$ and $\boldsymbol{\nu}$ can be inferred from the mean values and statistical fluctuations of the two independent random variables λ and μ as follows. For each sampled pair of values of the independent random variables $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, the relations (V.1) can be used to obtain the corresponding values of the pair of dependent random variables $(\mathbf{Y}, \boldsymbol{\nu})$. Then, the mean values of $\mathbf{Y}$ and $\boldsymbol{\nu}$ and their statistical fluctuations can be obtained by post-processing the inferred sampled values of $\mathbf{Y}$ and $\boldsymbol{\nu}$.

The material density ρ is also randomized using a gamma distribution around its mean value as follows

$$\boldsymbol{\rho} \sim \Gamma(k_\rho, \theta_\rho) \text{ where } k_\rho = 1 - \zeta_\rho \text{ and } \theta_\rho = \frac{E[\boldsymbol{\rho}]}{1-\zeta_\rho}$$

and $\zeta_\rho \in]-\infty; 1[$ is another parameter introduced to control the level of statistical fluctuations of $\boldsymbol{\rho}$.

The coefficients ζ_{Lame} and ζ_ρ are set to $\zeta_{Lame} = -10$ and $\zeta_\rho = -10$. For each homogenized, isotropic, linear, elastic material used in the FE stick model described above, 100 realizations of $\mathbf{Y}$, $\boldsymbol{\nu}$, and $\boldsymbol{\rho}$ are generated. For each realization, the discrete, *high-dimensional*, generalized eigenvalue problem (II.1) is re-solved using the newly generated FE stiffness matrix $K(\rho, Y, \nu)$ and FE mass matrix $M(\rho, Y, \nu)$.

The resulting probability density functions of $\mathbf{Y}$, $\boldsymbol{\nu}$, and $\boldsymbol{\rho}$ for one of the homogenized, isotropic, linear, elastic materials used for modeling the mAEWing1 aircraft are shown in Figure 4. The corresponding

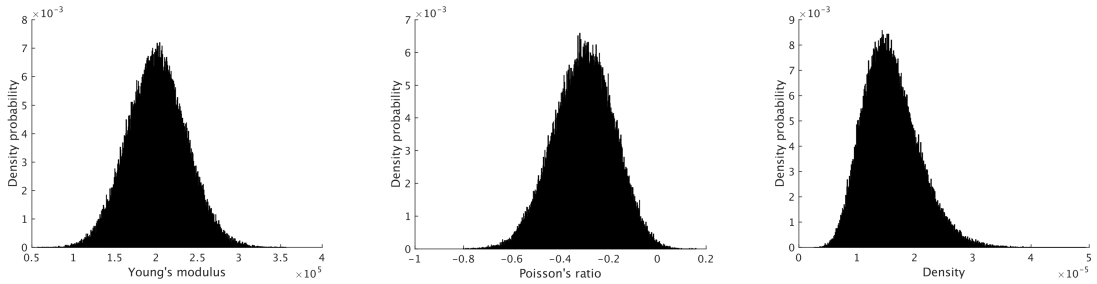

Figure 4. Probability density functions of Young's modulus (left), Poisson's ratio (middle), and the density of one of the homogenized, isotropic, linear, elastic materials used for modeling the mAEWing1 aircraft.

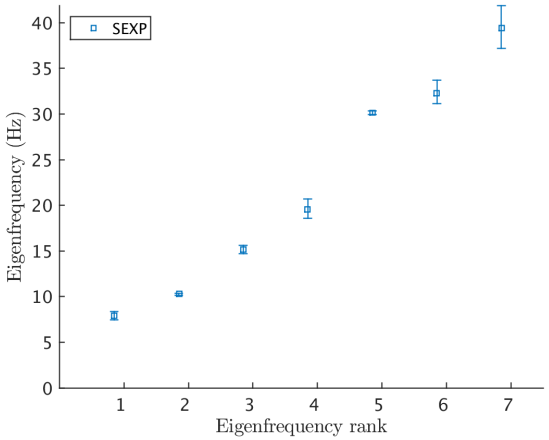

Figure 5. Surrogate statistical experimental values $\boldsymbol{f}_j^{\text{sexp}} = \boldsymbol{\omega}_j^{\text{sexp}}/2\pi$ of the QoIs $\boldsymbol{f}_j = \boldsymbol{\omega}_j/2\pi$, $j = 1, \ldots, 7$, of the mAEWing1 aircraft.

variations of the first 7 nonzero natural frequencies associated with the eigenvalue solutions of problem (II.1) are graphically depicted in Figure 5, adopted as surrogate statistical experimental data, and referred to as the stochastic QoIs $\boldsymbol{f}_j$, $j = 1, \cdots, 7$.

In Figure 4, the reader can observe that Poisson's ratio has negative values: these are due to the modeling approach adopted for constructing the FE stick model. In Figure 5, the reader can observe that the first, second, third, and fifth modes exhibit a low level of fluctuations compared to the fourth, sixth, and seventh modes.

D. Real Ground Vibration Test Data

Two instances of the mAEWing1 aircraft referred to in Reference[23] as "Sköll" and "Hati" have been built. Their dimensions and shapes, and therefore their aerodynamic properties, are supposedly identical. However, their structural properties slightly differ due to differences in their construction processes. GVTs were recently conducted for both aircraft by suspending them from a spring and applying to them an input force via an electrodynamic shaker.[23] In each case, accelerometer measurements were recorded at twenty points along the aircraft wings, and the measured force and acceleration data were post-processed using two different methods to identify natural frequencies and mode shapes. The obtained natural frequencies are reported in Table 2 below, as documented originally in Reference.[23] In this table, IDM refers to the identification method.

Three observations are noteworthy:

- There are only minor discrepancies between the results obtained using the first identification method and their counterparts obtained using the second identification method.

Table 2. mAEWing1 aircraft: ground vibration test results using two different identification methods, as documented in Reference[23] (mode numbering starts at the first flexible mode).

Aircraft	f_1 (Hz)		f_2 (Hz)		f_3 (Hz)		f_4 (Hz)		f_6 (Hz)	
	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2
Sköll	7.23	7.23	8.17	8.14	15.58	15.58	26.10	26.02	–	–
Hati	7.95	7.96	–	13.83	15.96	15.97	19.22	19.07	32.0	31.9

- There are however important discrepancies occasionally between the natural frequencies identified for Sköll, and their counterparts identified for Hati.
- From the narrative and graphical descriptions of the experimentally identified natural mode shapes of both flying wing aircraft that can be found in Reference,[23] it follows that the large discrepancies between the natural frequencies of the second and fourth flexible modes of Sköll and their counterparts for Hati are due to inconsistent numberings of the natural modes of these aircraft – or in other words, inconsistent frequency-mode shape associations.

To fix the inconsisteny issue highlighted in the third bullet above, the mode numbering and natural mode shapes of the FE model described in Section V.B are adopted here as references for all GVT results obtained for the Sköll and Hati aircraft (due to the limited number of accelerometer sensors that can be used in GVT, the eigenmodes of a FE model are typically better resolved than experimentally identified counterparts; therefore, they usually constitute a more reliable reference for performing frequency-mode shape associations and comparing mode shapes). This transforms Table 2 into Table 3 where the results suggest that, as already stated in Section V.B, the GVT of Sköll missed its third and fifth natural flexible modes, and the GVT of Hati missed its second and fifth natural flexible modes. Note that whereas in Table 3 the FE-identified third mode frequency aligns better with the experimentally-identified fourth mode frequencies, the FE-identified third mode shape aligns better with the experimentally-identified third mode shape.

Hence, the target values of the second SPROM (real scenario) described in Section V.E are set to those extracted from Table 3. However, it is noted that even with the different natural frequency-mode shape associations reported in Table 3, the first two observations outlined above still hold. More importantly, it is also noted that using the results of Table 3 instead of their counterparts of Table 2 to define the target values needed for constructing the SPROM (IV.4) does not affect the performance of the model-form UQ method proposed in this paper for generalized eigencomputations.

Table 3. mAEWing1 aircraft: re-ordered ground vibration test results (see,[23] and Table 2 and Table 1) based on a consistent frequency-mode shape association (mode numbering starts at the first flexible mode).

Aircraft	f_1 (Hz)		f_2 (Hz)		f_3 (Hz)		f_4 (Hz)		f_5 (Hz)		f_6 (Hz)	
	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2	IDM-1	IDM-2
Sköll	7.23	7.23	8.17	8.14	–	–	15.58	15.58	–	–	26.10	26.02
Hati	7.95	7.96	–	–	–	13.83	15.96	15.97	–	–	32.0	31.9
FE model	7.94		10.29		15.21		19.71		30.16		32.54	

E. Construction of a PROM and an SPROM

A ROB of dimension $n = 10$ and the corresponding Galerkin PROM of the same dimension are constructed for the discrete, generalized EVP (II.1) associated with the FE structural model of the mAEWing1 aircraft as described in Section III. Specifically, problem (III.1) is initialized with $f^{\text{ext}}(t) = 0$, $u_0 = 0$, and a broadband initial velocity v_0 that excites all m_u natural modes of interest. It is solved in the time-interval [0, 1 seconds] using the midpoint rule. A total number of 1,000 solution snapshots are collected in this time-interval at the sampling time-step $\Delta s = 0.001$ second, and compressed into a ROB $V \in \mathbb{R}^{4,146\times 10}$. Using this ROB, an SPROM of dimension $n = 10$ is constructed based on a stochastic ROB $\mathbf{V}$ with $m_\alpha = 58$ hyperparameters.

For the hypothetical scenario associated with the surrogate statistical experimental data described in Section C, these hyperparameters are identified by solving the optimization problem (IV.24) with:

- $m_\varrho^s = 1$ and $m_u = m_o = 7$.

- $o^{\text{ref}} = E[\boldsymbol{o}^{\text{exp}}]$ as in (IV.20), and $\boldsymbol{o}^{\text{exp}} = (\boldsymbol{f}_1^{\text{sexp}}, \boldsymbol{f}_2^{\text{sexp}}, \boldsymbol{f}_3^{\text{sexp}}, \cdots, \boldsymbol{f}_{m_o}^{\text{sexp}})^T$, where $\boldsymbol{f}_j^{\text{sexp}}$, $j = 1, \cdots, m_o$, denotes the j-th surrogate stochastic experimental natural frequency computed as described in Section C – that is, $\boldsymbol{f}_j^{\text{sexp}} = \boldsymbol{\omega}_j^{\text{sexp}}/2\pi$, where $\boldsymbol{\omega}_j^{\text{sexp}}$ denotes the j-th surrogate stochastic experimental natural circular frequency.
- v_j^{ref}, $j = 1, \cdots, m_o$, defined as in (IV.22), but using the surrogate statistical experimental data described in Section C – that is, $\boldsymbol{o}^{\text{exp}} = (\boldsymbol{f}_1^{\text{sexp}}, \boldsymbol{f}_2^{\text{sexp}}, \boldsymbol{f}_3^{\text{sexp}}, \cdots, \boldsymbol{f}_{m_o}^{\text{sexp}})^T$.

For the real scenario based on the GVTs discussed in Section V.D, the hyperparameters are identified by solving the same optimization problem (IV.24) but with:

- $m_\varrho^s = 1$ and $m_u = m_o = 5$.
- $o^{\text{ref}} = o^{\text{exp}}$ as in (IV.19), and $o^{\text{exp}} = (f_1^{\text{exp}}, f_2^{\text{exp}}, f_3^{\text{exp}}, f_4^{\text{exp}}, f_6^{\text{exp}})^T$, where f_j^{exp} denotes the natural frequency associated with the average measured frequency $f_j^{\text{exp}} = (f_j^{\text{Sköll}} + f_j^{\text{Hati}})/2$, and $f_j^{\text{Sköll}}$ and f_j^{Hati} are the natural frequencies of the j-th flexible mode shapes of the Sköll and Hati mAEWing1 aircraft, respectively, measured using the second identification method IDM-2 (see Table 3 and note that $f_2^{\text{exp}} = f_2^{\text{Sköll}}$ and $f_3^{\text{exp}} = f_3^{\text{Hati}}$).
- v_j^{ref}, $j = 1, \cdots, m_o$, defined by adapting (IV.21) as follows

$$v_j^{\text{ref}} = \gamma \max_{\text{Sköll,Hati}} |o_j^{\text{ref}} - o_j|, \qquad j = 1, \dots m_o$$

  where $o_j = \varpi_j/2\pi$, $j = 1, \cdots, m_o$, and γ is set to $\gamma = 1.2$.

For both scenarios, w_J is set to $w_J = 0.9$ so that the target of J_{mean} is preponderant during the identification of the vector-valued hyperparameter α (IV.16). The solution of the optimization problem (IV.24) is computed using a genetic algorithm and the following constraints:

$$\begin{aligned} a_s \le s \le b_s, \quad &\text{where} \quad a_s = 10^{-5} \text{ and } b_s = 10^{-2} \\ a_\beta \le \beta \le b_\beta, \quad &\text{where} \quad a_\beta = 0.001 \text{ and } b_\beta = 0.1 \\ 0 \le \sigma_{jj} \le 10,000,\ 1 \le j \le n \quad &\text{and} \quad -1,000 \le \sigma_{ij} \le 1,000,\ 1 \le i < j \le n \\ a_{s_{_M}} \le s_{_M} \le b_{s_{_M}}, \quad &\text{where} \quad a_{s_{_M}} = 0.9999 \text{ and } b_{s_{_M}} = 2.0 \end{aligned} \tag{V.2}$$

The genetic optimization algorithm is initialized with $\alpha_0 = (s_0, \beta_0, s_{_{M_0}}, \sigma_0)$, where $s_0 = 4 \times 10^{-5}$, $\beta_0 = 4 \times 10^{-2}$, $s_{_{M_0}} = 1$, and σ_0 is computed using the scheme proposed in Appendix A. It requires an initial population of individuals, each defined by a value of the hyperparameter vector α. The efficiency of such an algorithm depends on the chosen number of individuals in the population, and the distribution of the values of these individuals in the admissible set defined by the constraints (V.2). In both scenarios, the number of individuals in the population is set to $200 \times m_\alpha$. Then, the distribution of the values of α in its admissible set is constructed by generating $200 \times m_\alpha$ independent samples of the random vector $\boldsymbol{\alpha} = (\mathbf{s}, \boldsymbol{\beta}, \boldsymbol{\sigma}, \mathbf{s}_{_M})^T \in \mathbb{R}^{m_\alpha}$ such that $E\{\boldsymbol{\alpha}\} = \alpha_0$. The probability distributions of the random variables $\mathbf{s}$, $\boldsymbol{\beta}$, and $\mathbf{s}_{_M}$ are chosen as uniform distributions on the supports defined by their constraints. The probability model of the random matrix $\boldsymbol{\sigma}$ is constructed as follows. For $1 \le i < j \le n$, $\boldsymbol{\sigma}_{ij} = \sigma_{0ij}(1 + \mathcal{N}_{ij})$, and for $i = 1, \dots, n$, $\boldsymbol{\sigma}_{ii} = \sigma_{0ii}|0.728 + \mathcal{N}_{ii}|$, where $\{\mathcal{N}_{ij}\}_{1 \le i \le j \le n}$ is a family of $n(n+1)/2$ independent variables. Each of these variables is a Gaussian, centered variable with a unit variance. Note that the absolute value is introduced in the expression of $\boldsymbol{\sigma}_{ii}$ in order to guarantee the required positiveness of the diagonal terms of the matrix σ.

In both scenarios too, the Monte Carlo method is used as the stochastic solver for estimating the cost function for every value of α proposed by the genetic optimization algorithm. For such an estimate, 10 independent realizations of the QoIs are performed for estimating the cost function for any given value of α.

For the hypothetical scenario, the value of the cost function at convergence of the genetic optimization algorithm is $J(\alpha^{\text{opt}}) = 4.119 \times 10^{-5}$, and the optimal value of the vector-valued hyperparameter α is $\alpha^{\text{opt}} = (s^{\text{opt}}, s_{_M}^{\text{opt}}, \beta^{\text{opt}}, \sigma^{\text{opt}})^T$ where, $s^{\text{opt}} = 1.6988 \times 10^{-5}$, $s_{_M}^{\text{opt}} = 1.1211$, $\beta^{\text{opt}} = 9.9652 \times 10^{-2}$, and the diagonal entries of σ^{opt} are 0.42372, 0.30974, 0.35313, 0.39247, 0.16401, 1.66568, 0.77585, 0.80936, 0.33892, and 0.07803 (the non-diagonal entries of the matrix σ^{opt} are not reported here because of space limitation).

For the real scenario characterized by the availability of GVT data, $J(\alpha^{\text{opt}}) = 3.025 \times 10^{-3}$, $s^{\text{opt}} = 4.8477 \times 10^{-5}$, $s_{_M}^{\text{opt}} = 1.9426$, $\beta^{\text{opt}} = 9.9764 \times 10^{-2}$, and the diagonal entries of σ^{opt} are 0.89235, 0.28583, 0.00587, 0.11031, 0.08320, 0.36996, 0.25034, 0.07680, 0.50636 and 4.50132 (again, the non-diagonal entries of the matrix σ^{opt} are not reported here because of space limitation).

F. Performance of the SPROM in the Presence of Modeling Errors in Both the HDM and PROM

Figures 6 and 7 report in various forms the different results obtained for the QoIs associated with the hypothetical scenario of the mAEWing1 application defined by the surrogate statistical experimental data described in Section C. The reader can observe that:

- For this scenario, the constructed deterministic PROM is very accurate: it reproduces well the mean values of the realizations performed using the random HDM (FE model).
- For each considered QoI, the realizations generated by the SPROM and the resulting confidence interval corresponding to the quantiles 98% and 2% encompass the deterministic value of this QoI predicted using the deterministic PROM. They also capture relatively well the realizations of this QoI generated using the random FE model and its resulting same-level confidence interval.

Essentially, the results reported in Figures 6 and 7 illustrate the ability of the proposed nonparametric probabilistic method for UQ, which is designed to model and quantify the effects of *model-form* uncertainties, to perform well even in the presence of problem/model *parameter* uncertainties.

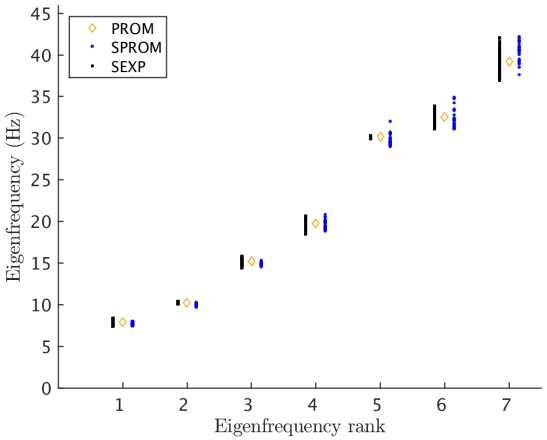

Figure 6. Hypothetical scenario for the mAEWing1 application where surrogate statistical experimental data are generated: QoIs predicted using the deterministic PROM; counterparts realized by the SPROM; counterparts realized by the random FE model (SEXP).

Figures 8 and 9 report in appropriate forms the various results obtained for the QoIs associated with the real scenario of the mAEWing1 application based on the averaged nonstatistical GVT data described in Section V.D ($f_j^{\text{exp}} = (f_j^{\text{Sköll}} + f_j^{\text{Hati}})/2$). These results show that:

- In this scenario for which experimental data are available, the deterministic PROM appears to suffer from the effect of model-form uncertainties inherited from the HDM – that is, the FE model. The values of the QoIs it predicts are in some cases off by up to 24% with respect to the corresponding measured values. The mismatches are exacerbated by the fact that occasionally, two different values are identified for the same QoI by the two different identification methods used during the GVTs.
- The GVT results are not perfect either as they appear to miss the fifth natural flexible mode of the aircraft.
- On the other hand, for all QoIs for which experimental data are available, the realizations and resulting confidence intervals corresponding to the quantiles 98% and 2% computed by the SPROM encompass these data – including when two different experimental values are reported for the same QoI.
- For the fifth natural flexible mode that was missed by the GVTs and therefore for which no experimental data are available, the same-level confidence interval based on the SPROM predictions includes the deterministic PROM prediction. For the seventh flexible mode for which no experimental data are available, the same-level confidence interval does not contain the deterministic PROM prediction.

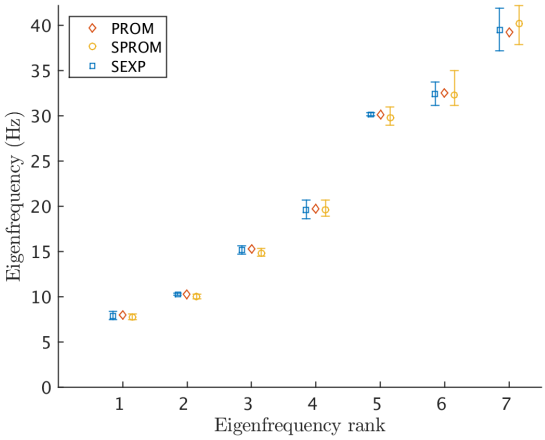

Figure 7. Hypothetical scenario for the mAEWing1 application where surrogate statistical experimental data are generated: QoIs predicted using the deterministic PROM; confidence intervals corresponding to the quantiles 0.98 and 0.02, constructed with 100 QoI samples generated by the SPROM (each circle within a confidence interval represents the statistical mean value of the corresponding QoI); counterpart confidence intervals generated by the random FE model (SEXP).

In summary, the results reported in Figure 8 and 9 demonstrate for the first time the ability of the nonparametric probabilistic method for model-form UQ presented in this paper to perform well in the presence of experimental data. Where such data are not available the method may to lead to slightly larger confidence intervals, most likely because in this case, the cost function underlying the optimization problem (IV.24) governing the identification of the vector-valued hyperparameter lacks the minimum amount of target information needed for producing sharp quantifications.

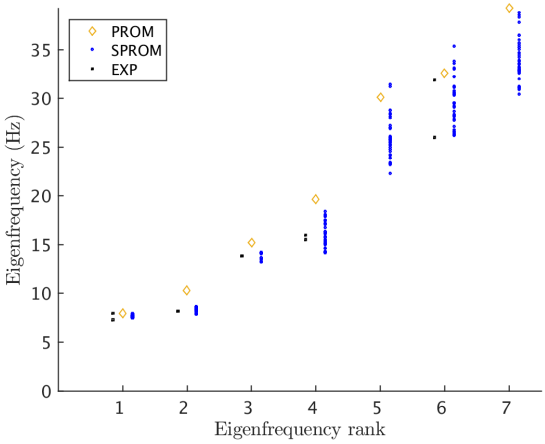

Figure 8. Real scenario of the mAEWing1 application with nonstatistical GVT data: QoIs predicted using the deterministic PROM; counterparts realized by the SPROM; averaged nonstatistical experimental values of the QoIs (EXP).

VI. Summary and Conclusions

A nonparametric, probabilistic method for modeling and quantifying model-form uncertainties associated with a High-Dimensional computational Model (HDM) built for generalized eigencomputations and for which a Projection-based Reduced-Order Model (PROM) can be constructed is presented. This method is feasible because its computational complexity scales with the size of a low-dimensional, Stochastic PROM (SPROM). The construction of this SPROM is based on three innovative ideas: the substitution of the deterministic

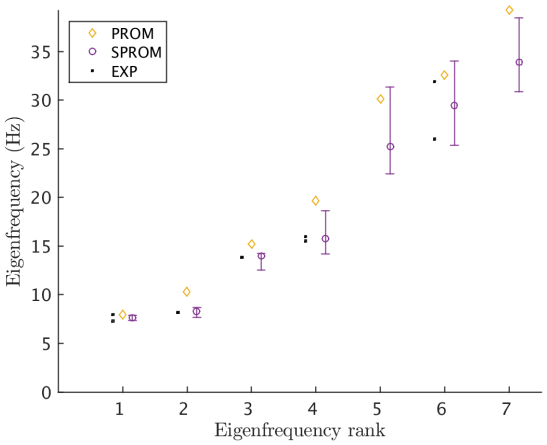

Figure 9. Real scenario of the mAEWing1 application with nonstatistical GVT data: QoIs predicted using the deterministic PROM; confidence intervals corresponding to the quantiles 0.98 **and** 0.02**, constructed with** 100 **QoI samples generated by the SPROM (each circle within a confidence interval represents the statistical mean value of the corresponding QoI); averaged nonstatistical experimental values of the QoIs (EXP).**

Reduced-Order Basis (ROB) underlying the PROM with a stochastic counterpart (SROB) featuring a small number of hyperparameters; the construction of this SROB on a subset of a compact Stiefel manifold in order to guarantee the linear independence of its column vectors and the satisfaction of the boundary conditions; and the formulation and solution of a small-scale inverse statistical problem to determine the aforementioned hyperparameters so that the mean value and statistical fluctuations of the predicted stochastic eigenvalues match target values obtained from available data. These data can be high-dimensional numerical data, in which case the proposed method models and quantifies the effects on the solution of the various sources of model-form uncertainties associated with projection-based model order reduction. It can also be non-statistical or statistical data, in which case the proposed method models and quantifies the effects on the solution of the model-form uncertainties associated with both the PROM and its underlying HDM. In the latter case, the proposed method distinguishes itself from alternative approaches for performing model-form Uncertainty Quantification (UQ) in HDM-based computations that rely on a PROM for accelerating the process by the fact that it accounts for the modeling errors associated with the construction of the auxiliary PROM. In either case, the proposed method can be interpreted as an effective approach for extracting fundamental information or knowledge from data that are not captured by a deterministic HDM and therefore its associated PROM, and incorporating it in this PROM.

The successful application of the presented method to the modeling and quantification of model-form uncertainties in generalized eigencomputations associated with a natural vibration analysis of an mAEWing1 flying wing for which ground ground vibration test data are available demonstrates its potential for the solution of realistic UQ problems. The mAEWing1 aircraft is a small-scale replica of an X-56 type aircraft made of a composite material. Its finite element HDM built by structural dyamics practitioners is riddled with model-form uncertainties. These are due to, among other things, the adopted stick and lumped mass modeling approaches, the homogenized representation of the composite material, and the lack of a structural damping representation. It turns out that these model-form uncertainties have a detrimental effect on the deterministic PROM built for this HDM using the Proper Orthogonal Decomposition and method of snapshots. Indeed, this PROM predicts natural frequencies for the mAEWing1 flying wing that are in some cases off by up to 24% with respect to their measured values. The mismatches are exacerbated by the fact that occasionally, two different values were identified during Ground Vibration Tests (GVTs) for the same natural frequency. On the other hand, the SPROM constructed for this aircraft using the proposed method delivers for all measured natural frequencies realizations and resulting confidence intervals corresponding to the quantiles 98% and 2% that encompass both the values predicted by the deterministic PROM and the available GVT data. This demonstrates the ability of the proposed nonparametric probabilistic method for model-form UQ to perform well in the presence of experimental data.

Acknowledgements

The US authors acknowledge partial support by the Army Research Laboratory through the Army High Performance Computing Research Center under Cooperative Agreement W911NF-07-2-0027, and partial support by DARPA under the Enabling Quantification of Uncertainty in Physical Systems (EQUiPS) program. The mAEWing1 flying wing and its variants were built and tested for the NASA Performance Adaptive Aeroelastic Wing (PAAW) program. This document does not necessarily reflect the position of any of these institutions, and no official endorsement should be inferred.

A. Initialization of the Matrix σ for the mAEWing1 Applications

Regarding the solution of the optimization problem (IV.24), the initialization of the hyperparameters s, β, and s_{M_0} is described in Section V.E. Here, an algorithm for initializing the matrix σ (IV.13, IV.14) is proposed.

Let $U_r \in \mathbb{R}^{n\times n}$ denote the matrix whose columns are the n eigenvectors u_{r_j} of the PROM (III.3). From (III.3), it follows that

$$K_r U_r - M_r U_r \Omega^2 = 0 \tag{A.1}$$

where Ω^2 is the diagonal matrix whose entries are the QoIs ϖ_j^2.

Let $U_r^{\text{ref}} \in \mathbb{R}^{n\times n}$ denote the matrix where each j-th column corresponds to a certain mean value of the random reduced eigenvector $\mathbf{u}_{r_j}$ associated with the target random QoI $\boldsymbol{\varpi}_j^2$, and let $\boldsymbol{\Omega}^{2^{\text{ref}}}$ denote the random diagonal matrix whose j-th diagonal entry is the random QoI $\boldsymbol{\varpi}_j^2$. From (IV.4), it follows that

$$K_r U_r^{\text{ref}} - M_r U_r^{\text{ref}} \boldsymbol{\Omega}^{2^{\text{ref}}} = \boldsymbol{\mathcal{E}}_r \tag{A.2}$$

where $\boldsymbol{\mathcal{E}}_r \in \mathbb{R}^{n\times n}$ is a random matrix that represents the residual error.

An error for the mean value solution can be defined as

$$U_r^e = U_r^{\text{ref}} - U_r \Rightarrow U_r^{\text{ref}} = U_r + U_r^e \tag{A.3}$$

Hence, subtracting Equation (A.1) from Equation (A.2) and using the definition (A.3) gives

$$K_r U_r^e - M_r U_r^e \boldsymbol{\Omega}^{2^{\text{ref}}} + M_r U_r (\Omega^2 - \boldsymbol{\Omega}^{2^{\text{ref}}}) = \boldsymbol{\mathcal{E}}_r \tag{A.4}$$

Assuming that $Q = I_N$ (see (III.4)), it follows that $U_r^{\text{ref}^T} U_r^{\text{ref}} = U_r^T U_r = I_n$, and U_r^e must belong to the admissible set of $n \times n$ real matrices

$$\mathcal{C} = \{U \mid \quad U^T U + U^T U_r + U_r^T U = 0\} \tag{A.5}$$

Then, the error matrix U_r^e can be computed by minimizing the mean-square error, $E\{\|\boldsymbol{\mathcal{E}}_r\|_F^2\}$, over the admissible set $\mathcal{C}$ – that is, by solving the optimization problem

$$U_r^e = \underset{U\in\mathcal{C}}{\operatorname{argmin}}\, E\big\{\|K_r U - M_r U \boldsymbol{\Omega}^{2^{\text{ref}}} + M_r U_r (\Omega^2 - \boldsymbol{\Omega}^{2^{\text{ref}}})\|_F^2\big\} \tag{A.6}$$

This problem can be rapidly solved by an interior-point algorithm with nonlinear constraints such as that implemented in Matlab©'s subroutine *fmincon*.

Noting that $U_r^{e^T} U_r^e$ is an SPD matrix, let $L \in \mathbb{R}^{n\times n}$ be the matrix defining its Cholesky factorization $U_r^{e^T} U_r^e = L^T L$. Recalling the properties (IV.11), (IV.12) and (IV.13), it follows that a good initial value σ_0 of the matrix σ is

$$\sigma_0 = \frac{1}{c} L, \quad \text{where} \quad c = \max_{j=1,\ldots,n} L_{jj} \tag{A.7}$$

References

[1]Zahr M, Farhat C. Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *International Journal for Numerical Methods in Engineering* 2015; 102(5):1077-1110.

[2]Amsallem D, Zahr M, Choi Y, Farhat C. Design optimization using hyper-reduced-order models. *Structural and Multidisciplinary Optimization* 2015; 51(4):919-940.

[3]Soize C, Farhat C. A nonparametric probabilistic approach for quantifying uncertainties in low- and high-dimensional nonlinear models. *International Journal for Numerical Methods in Engineering* 2016; DOI: 10.1002/nme.5312.

[4]Amsallem D, Deolalikar S, Gurrola F, Farhat, C. Model predictive control under coupled fluid-structure constraints using a database of reduced-order models on a tablet. *AIAA-2013-2588*, 31st AIAA Applied Aerodynamics Conference, San Diego, California, June 24-27, 2013.

[5]Washabaugh K, Zahr M, Farhat C. On the use of discrete nonlinear reduced-order models for the prediction of steady-state flows past parametrically deformed complex geometries. *AIAA-2016-1814*, AIAA SciTech 2016, San Diego, CA, January 4-8, 2016.

[6]Ryckelynck D. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics* 2005; 202:346-366.

[7]Grepl MA, Maday Y, Nguyen NC, Patera A. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis* 2007; 41(03):575-605.

[8]Farhat C, Avery P, Chapman T, Cortial J. Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. *International Journal for Numerical Methods in Engineering* 2014; 98(9):625-662.

[9]Farhat C, Chapman T, Avery P. Structure-preserving, stability, and accuracy properties of the Energy-Conserving Sampling and Weighting (ECSW) method for the hyperreduction of nonlinear finite element dynamic models. *International Journal for Numerical Methods in Engineering* 2015; 102(5):1077-1110.

[10]Beck JL, Katafygiotis LS. Updating models and their uncertainties - I: Bayesian statistical framework. *Journal of Engineering Mechanics* 1998; 124(4):455-461.

[11]Ghanem R, Spanos PD. *Stochastic Finite Elements: A Spectral Approach (revised edition)*. Dover Publications: New York, 2003.

[12]Mace R, Worden W, Manson G. Uncertainty in structural dynamics.*Special issue of the Journal of Sound and Vibration* 2005; 288(3):431-790.

[13]Schueller GI, Pradlwarter HJ. Uncertain linear systems in dynamics: Retrospective and recent developments by stochastic approaches. *Engineering Structures* 2009; 31(11):2507-2517.

[14]Soize C. *Uncertainty Quantification. An Accelerated Course with Advanced Applications in Computational Engineering.* Springer: New York, 2017.

[15]Bui-Thanh T, Willcox K, Ghattas O. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamic applications. *AIAA Journal* 2008; 46(10):2520-2529.

[16]Degroote J, Virendeels J, Willcox K. Interpolation among reduced-order matrices to obtain parameterized models for design, optimization and probabilistic analysis, *International Journal for Numerical Methods in Fluids* 2010; 63:207-230.

[17]Soize C. A nonparametric model of random uncertainties for reduced matrix models in structural dynamics. *Probabilistic Engineering Mechanics* 2000; 15(3):277-294.

[18]Soize C. Random matrix theory for modeling random uncertainties in computational mechanics. *Computer Methods in Applied Mechanics and Engineering* 2005; 194(12-16):1333-1366.

[19]Soize C. Nonparametric probabilistic approach of uncertainties for elliptic boundary value problem.*International Journal for Numerical Methods in Engineering* 2009; 80(6-7):673-688.

[20]Sirovich L. Turbulence and the dynamics of coherent structures. I-Coherent structures. II-Symmetries and transformations. III-Dynamics and scaling. *Quarterly of Applied Mathematics* 1987; 45:561-571.

[21]Geuzaine P, Brown G, Harris C, Farhat, C. Aeroelastic dynamic analysis of a full F-16 configuration for various flight conditions. *AIAA Journal* 2003; 41:363-371.

[22]Farhat C, Geuzaine P, Brown, G. Application of a three-field nonlinear fluid-structure formulation to the prediction of the aeroelastic parameters of an F-16 fighter. *Computers and Fluids* 2003; 32:3-29.

[23]Gupta A, Seiler P, Danowsky, B. Ground vibration tests on a flexible flying wing aircraft. *AIAA-2016-1753*, AIAA SciTech 2016, San Diego, CA, January 4-8, 2016.

[24]Regan, C D and Taylor, B R. mAEWing1: design, build, test. *AIAA-2016-1747*, AIAA SciTech 2016, San Diego, CA, January 4-8, 2016.

[25]Guilleminot J, Soize C. On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties. *Journal of Elasticity* 2013; 111(2):109-130.

[26]Hogg RV, Craig AT. Introduction to Mathematical Statistics, 4th edition. New York: Macmillan, 1978. (See Section 3.3)