On the Combination of Random Matrix Theory With Measurements on a Single Structure

Igea, Felipe; Chatzis, Manolis N.; Cicirello, A.

DOI
10.1115/1.4054172

Publication date
2022

Document Version
Final published version

Published in
ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering

Citation (APA)
Igea, F., Chatzis, M. N., & Cicirello, A. (2022). On the Combination of Random Matrix Theory With Measurements on a Single Structure. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *8*(4), Article 041203. https://doi.org/10.1115/1.4054172

On the Combination of Random Matrix Theory With Measurements on a Single Structure

Felipe Igea
Department of Engineering Science,
University of Oxford,
Parks Road,
Oxford OX1 3PJ, UK
e-mail: felipe.igea@eng.ox.ac.uk

Manolis N. Chatzis
Department of Engineering Science,
University of Oxford,
Parks Road,
Oxford OX1 3PJ, UK
e-mail: manolis.chatzis@eng.ox.ac.uk

Alice Cicirello
Faculty of Civil Engineering and Geosciences,
Department of Engineering Structures,
Section of Mechanics and Physics of Structures (MPS),
Delft University of Technology,
Stevinweg 1,
Delft NL 2628, The Netherlands
e-mail: a.cicirello@tudelft.nl

An approach is proposed for the evaluation of the probability density functions (PDFs) of the modal parameters for an ensemble of nominally identical structures when there is only access to a single structure and the dispersion parameter is known. The approach combines the Eigensystem realization algorithm on sets of dynamic data, with an explicit nonparametric probabilistic method. A single structure, either a mathematical model or a prototype, is used to obtain simulated data or measurements that are employed to build a discrete time state-space model description. The dispersion parameter is used to describe the uncertainty due to different sources such as the variability found in the population and the identification errors found in the noisy measurements from the experiments. With this approach, instead of propagating the uncertainties through the governing equations of the system, the distribution of the modal parameters of the whole ensemble is obtained by randomizing the matrices in the state-space model with an efficient procedure. The applicability of the approach is shown through the analysis of a two degrees-of-freedom mass-spring-damper system and a cantilever system. The results show that if the source of uncertainty is unknown and it is possible to specify an overall level of uncertainty, by having access to a single system's measurements, it is possible to evaluate the resulting PDFs on the modal parameters. It was also found that high values of the dispersion parameter may lead to nonphysical results such as negative damping ratios values. [DOI: 10.1115/1.4054172]

1 Introduction

The quantification of the uncertainties on the modal parameters during the design process of structures is of great interest for the assessment of their dynamic performance [1,2]. The measurement of the dynamic response across nominally identical structures can reveal largely different modal parameter estimates due to uncertainties originated by the variability of the manufacturing processes of structural components, boundary conditions, and assemblage [3,4]. Identifying the distribution of the modal parameters across the ensemble of nominally identical structures would enable the selection of designs that are robust to these uncertainties, avoiding extensive modifications of the manufactured product [5]. This paper proposes an approach for the evaluation of the probability density functions (PDFs) of the natural frequencies, damping ratios, and modal shapes of an ensemble of nominally identical linear time-invariant systems, for cases where there is only access to a single structure and additional information is known in the form of the so-called dispersion parameter. This single structure, represented either as a mathematical model or a prototype, is used to obtain simulated data or measurements, respectively, which in turn are employed to build a discrete time state-space model description. This model is then used to efficiently assess the effects on the modal properties of different levels of uncertainties, represented through suitably chosen dispersion parameters. This method is also of interest within the Bayesian Inference framework, as it can be used to build a prior distribution consistent with the available information, for the cases where there would be insufficient information in the data so that resulting likelihood cannot overrule the prior assumptions [6].

At early design stages (before a prototype is being built), when a physics-based model is developed, uncertainties on the model parameters are usually described using probabilistic [7–9] or nonprobabilistic uncertainty descriptions [10]. These uncertainties are then propagated through the equation of motions to yield the corresponding description of the response. This is the so-called parametric model of uncertainty. However, the choice of the uncertain parameters and their description would directly affect the resulting distribution of the resulting modal parameters [1,2,8]. An alternative approach is to employ a nonparametric model of uncertainty by exploiting random matrix theory (RMT) results [1,11]. This is useful as nonparametric methods [11] avoid the need to specify the uncertainties' sources and the description of the parameters model's uncertainty, which are often hard to determine. Moreover, using these methods, modeling errors may also be accounted for [1]. A review of different random matrices and their properties is given in Ref. [12]. The application of these matrices to engineering problems, and in particular, for uncertainty quantification has been the subject of much recent research [13–16]. Broadly speaking, random matrices can be applied using an implicit approach that is based on the derivation of analytical results given a set of assumptions, e.g., by assuming that the physical properties of a structural component are sufficiently random so that the statistical distribution of the natural frequencies and mode shapes tends to a universal distribution associated with the Gaussian orthogonal ensemble of random matrices [17,18], or also by using an explicit approach that is based on the use of Monte Carlo simulations to propagate uncertainty [1,11].

Vishwajeet et al. [19] have recently used results from RMT in combination with the system identification (SI) method Eigensystem realization algorithm (ERA) for the calculation of the analytical expression of the PDFs of the singular values of the Hankel matrix. This was done by assuming that the elements of the Hankel matrix were Gaussian random variables, and therefore, the Hankel matrix times its own conjugate transpose conforms with the noncentral Wishart distribution. This assumption might not be always valid for linear-time-invariant systems, as noise is non-Gaussian in nature, and often only approximated as Gaussian. Moreover, the prediction of the uncertainty of the properties of the modal parameters associated with a priori knowledge of uncertainties related to manufacturing variability, boundary conditions,

Manuscript received October 19, 2021; final manuscript received March 16, 2022; published online April 28, 2022. Assoc. Editor: Ioannis A. Kougioumtzoglou.

and assemblage of other nominally identical structures, using measurements taken from a single structure during the design stage, has not been addressed with current methods. Quantifying the effect of those uncertainties on the untested members of the population becomes paramount if reliable performance assessments of the structures must be produced. This is the focus of the present paper, where no assumption is made on the type of the distribution followed by the elements in the Hankel Matrix, and a nonparametric explicit implementation of RMT is used to build the ensemble employed to evaluate the PDFs of the modal parameters, using the dispersion parameter to control the level of uncertainty without specifying its sources. Relevant uncertainty propagation approaches based on this type of explicit implementation of RMT are briefly reviewed in what follows.

Soize [1,11] pioneered an RMT explicit approach that models the system's matrices (mass, stiffness, and damping) as random matrices. Using the maximum entropy method and the information available (system's matrices are symmetric positive-definite, the second-order moment of their inverse, and their mean matrices exist), an ensemble of random matrices can be used to model the uncertainty of the system's matrices in a nonparametric manner. This explicit implementation of RMT has been successfully applied for building stochastic models to quantify uncertainty in nominally identical systems by defining a single parameter, the so-called dispersion parameter [1,11]. The dispersion parameter is used to control the overall level of uncertainty in the random matrix without the need to specify its origin. This parameter can be used at the design stage to build an ensemble of nominally identical models/structures, and it can be updated once experimental observations become available [1,11]. The dispersion parameter is calculated by assuming that the nominal model/structure is available (or approximated), and that the experimental observations are different realizations of the ensemble of structures [1,20]. The dispersion parameter can be varied to create different ensembles and evaluate the effect on the produced results [21]. An example, where the dispersion parameter that controls the level of the uncertainty in the stiffness matrix, was identified from experimentally obtained frequency response functions of six nominally designed aircraft T-tails, used for the calculation of the experimental modal parameters, is shown in paper [16]. Paper [14] illustrates another recent application using experimental data from booster pumps' thermal units to identify the dispersion parameter. This dispersion parameter is used to build a stochastic computational model that considers mode crossing and veering phenomena through the introduction of an adapted transformation for the calculated modal quantities. Legault et al. [21] applied this nonparametric approach to simple numerical examples and studied its meaning and consequences for the output space (e.g., average modal density, dispersion relation, etc.), especially for large values of the dispersion parameters.

This current paper follows the explicit approach developed by Soize and comprehensively reviewed in Refs. [1] and [11], where the general applicability and proofs of the approach are given. In particular, with the proposed approach, starting from a set of measurements or simulated data obtained from either a mathematical model or a prototype, the Hankel Matrix is calculated. The matrix resulting from the multiplication of the Hankel Matrix times its own conjugate transpose is randomized using the normalized positive definite ensemble defined in Refs. [1] and [11]. For each realization of the ensemble, the ERA [22] is applied to identify the modal parameters (natural frequencies, damping ratios, and modal shapes). ERA has been chosen for its simplicity, although its application is limited to the analysis of structures under some assumptions [22], such as impulsive excitations or free vibration. If the appropriate changes to the proposed method are considered, other loading conditions may be investigated; however, these changes are not explored in this paper. The results of each realization are then used to build the PDFs of the modal parameters of the ensemble. The proposed approach enables the assessment of the effect on the modal parameters' uncertainties of different dispersion parameter values and size of Hankel matrices.

It is worth noting that the PDFs of the modal parameters of an ensemble of nominally identical structures obtained with the proposed approach correspond to those that would be obtained from an ensemble of structures from a production line, such as cars, or by considering manufactured structures, which are realizable in principle, but for which only one sample may be built, such as a bridge. This virtual ensemble accounts for a collection of linear structures (not all of them are necessarily existent but are physically feasible) in which the geometric, mechanical properties, and boundary conditions of all the elements of each structure are uncertain and conform to unknown distributions. The applicability of the proposed approach is shown using numerical applications that represent situations where the system is under free vibration.

This paper is structured as follows: The ERA method is reviewed in Sec. 2. The steps involved on how to combine RMT and ERA for the quantification of the uncertainty of the modal parameters are found in Sec. 3. The numerical results obtained are found in Sec. 4. The physical consequences of the application of RMT are then discussed in Sec. 5.

2 Proposed Approach

The proposed approach combines an SI technique with a nonparametric uncertainty description to quantify the effect of different levels of uncertainty on the modal properties of an ensemble of nominally identical structures, using probability density functions of these modal properties. In particular, by using measurements or simulated data obtained from a single structure, an ensemble able to encompass different sources of uncertainty is built as a product of the Hermitian matrices, which is the Hankel Matrix times its own conjugate transpose. The random matrix is specifically constructed to reflect a given level of uncertainty (expressed in terms of the so-called dispersion parameter) and to retain some properties of the Hermitian matrix (i.e., positive definiteness and the existence of the second-order moment and of its inverse). In particular, the normalized positive-definite random matrix (random matrix ensemble $SG+$) is chosen [11].

A schematic representation of the approach is shown in Fig. 1. Each of the blocks of the approach is discussed in more detail in Secs. 2.1–2.4.

2.1 Measurements on a Single Structure to Construct a Virtual Ensemble. When the physics-based model of a structure is available, and the structure is not yet built, the model can be

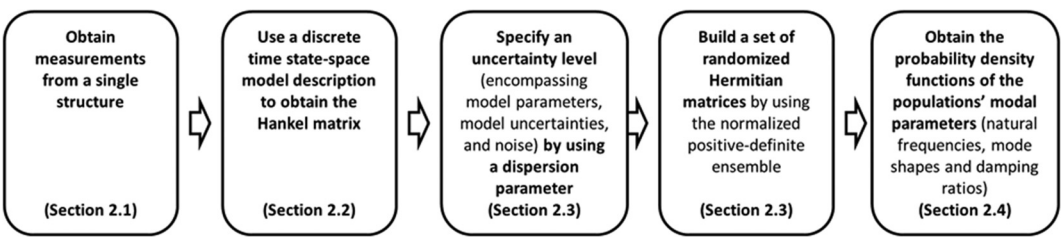


Fig. 1 Schematic representation of the approach

used for simulating its dynamic response, i.e., displacement, velocity, or acceleration signals, using its nominal properties [11]. The outputs signals can then be contaminated with noise. Therefore, a set of response signals, ideal measurements, can be obtained by using the physics-based model. Alternatively, real measurements can be carried out on a prototype of the real structure or on a structure that is a member of an ensemble of structures that would be produced. These simulated data or measurements on a single structure can then be used to create the nominal model, that in conjunction with the dispersion parameters is used to construct the virtual ensemble. In this work, the virtual ensemble consists of a set of linear structures, described by a set of common geometric properties, mechanical properties, and boundary conditions, and where each realization corresponds to a different structure across the ensemble. Therefore, this ensemble is used to account for structure-to-structure response variability caused by different sources of uncertainty that may occur in the modeling and manufacturing process. In addition, it has to be considered that the variability that the virtual ensemble encompasses several sources of uncertainty, such as the noise of the experimental measurements and the uncertainties introduced by the modeling process.

Starting from the simulated or real set of measurements, a discrete time state-space model description can then be readily built or identified, respectively. The number of measurements (in terms of sensors) needed depends on the model order, number of modes, and the complexity of the structure. The measurements used are those obtained from vibration-based experiments such as displacement, velocity, and acceleration readings.

In this paper, the ensemble is constructed starting from the physics-based model in order to ensure reproducibility of the results shown.

2.2 Evaluation of Hankel Matrix. Modal identification methods are used to obtain the modal properties (natural frequencies, mode shapes, and damping ratios) of a system. The approach proposed in this paper requires the evaluation of the Hankel Matrix.

The Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$ of size r by s contains time series data from measurement data [23] and it can be built as shown in Eq. (1). The output vector sequences $y_{\{k,p\}}$ contain the measurements read in p channels at different times t_k. Assuming free vibration conditions, $\boldsymbol{y}_0, \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \ldots, \boldsymbol{y}_k$ are obtained

$$\boldsymbol{H}_{rs}(k-1) = \begin{bmatrix} \boldsymbol{y}_k & \boldsymbol{y}_{k+1} & \cdots & \boldsymbol{y}_{k+s-1} \\ \boldsymbol{y}_{k+1} & \boldsymbol{y}_{k+2} & \cdots & \boldsymbol{y}_{k+s} \\ . & . & & . \\ . & . & & . \\ \boldsymbol{y}_{r+k-1} & \boldsymbol{y}_{r+k} & \cdots & \boldsymbol{y}_{r+k+s-2} \end{bmatrix} \tag{1}$$

In the present analysis, the ERA is chosen because of its simplicity when dealing with systems excited using an impulse force (e.g., hammer strike) or systems in free vibration [22].

However, other loading conditions may be considered if the appropriate changes to the proposed approach are implemented. For example, the combination of the natural excitation technique with ERA [24] allows to deal with situations where ambient excitation is observed. These methods are beyond the scope of the present work, and the reader is referred to Ref. [25].

2.3 Defining uncertainty by Using a Dispersion Parameter δ_S and Matrix Randomization. In parametric probabilistic approaches for uncertainty quantification [7,8], the different types of uncertainties (model noise, measurement noise, population uncertainty, etc.) are described by using some assumed PDFs. The uncertainties can then be propagated through the relevant equations describing the behavior of the system, and the effects of the uncertainties on the parameters of interest are assessed.

In this paper, an alternative approach that accounts for the effects of uncertainties without the need to make explicit assumption on the PDF of specific model parameters is considered. In particular, a nonparametric technique based on RMT is used by following the explicit approach proposed by Soize [1,11]. A dispersion parameter δ_S is used to build the random matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ that will be used to create the virtual ensemble.

The dispersion parameter value encompasses the overall level of uncertainty caused by different sources of uncertainty that may be present, such as modeling errors, manufacturing variability, identification errors, errors in the recorded signals, and variability of the boundary conditions among the members of the ensemble.

The prior estimation of the dispersion parameter δ_S, can be classified into three cases:

(i) The first case occurs when no prior data are available, and two approaches can be considered. The nominal value and the mean matrix of the stochastic model are the same. In this case, the dispersion parameter δ_S is a variable that is used for a sensitivity analysis that relates the level of uncertainties to the stochastic solution [1]. The sensitivity analysis is performed by varying the uncertainty levels controlled by the dispersion parameter value δ_S in a predefined range and observing how the PDFs of the modal parameters (natural frequencies, mode shapes and damping ratios) change [1].
Alternatively, Legault et al. [21] define a dispersion parameter δ_S that is set a priori based on an expected level of uncertainty. For example, low, mid, and high values of the dispersion parameter were assigned in Ref. [21] to evaluate the effects on their chosen properties. When high uncertainty is present, higher dispersion parameter values are used.

(ii) The second case occurs when prior data (either measurements taken on the ensemble of structures or an existent computational model with specified parametric uncertainties) are available [1]. For this case, from the observed data, it is possible to update the mean matrix, and using the least-square method or the maximum likelihood method to obtain the value of the dispersion parameter δ_S that optimizes the employed objective function (e.g., minimizing the difference on the resulting coefficient of variation on the first natural frequency). A detailed explanation about both methods may be found in [1].

(iii) The third case occurs when in addition to the parametric prior information (case (ii)), there is also a statistically independent nonparametric prior information (case (i)). For example, the uncertainty in the measurements noise can be specified as a zero mean Gaussian with a given variance, and the modeling error is specified with a dispersion parameter δ_A. The goal is to obtain an overall dispersion parameter $\delta_s = \delta_D + \delta_a$, where δ_D is calculated as shown in case (ii) without accounting for the nonparametric uncertainty.

The three cases above are investigated in the first numerical application (Sec. 4.1). The dispersion parameter quantifies the quadratic distance from the random matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ to the identity matrix I_n as shown by Eq. (2) below [1,11]

$$\delta_S = \left(\frac{1}{n} E\left\{ \| \hat{\boldsymbol{G}}_{\boldsymbol{S}} - \boldsymbol{I}_n \|_F^2 \right\} \right)^{1/2} \tag{2}$$

where $\|\|_F$ denotes the Frobenius norm and n is the number of rows or columns of the matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$. The dispersion parameter takes values within the inequality shown by the Eq. (3) below [11]:

$$0 < \delta_S < \sqrt{(n+1)(n+5)^{-1}} \tag{3}$$

In general, as it is not possible to have multiple elements of the virtual ensemble or prior measurements in the design stage, in this paper, the strategy (i) described in Ref. [21] is followed, and different a priori values of the dispersion parameter δ_S are used to

evaluate the effect of the dispersion value on the resulting PDFs of the modal properties of the virtual ensemble. In particular, low and high values of the dispersion parameters are considered.

For a given full row rank Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$, a Hermitian matrix $\boldsymbol{S}(k-1)$ is calculated [19] using

$$\boldsymbol{S}(k-1) = \boldsymbol{H}_{rs}(k-1)\boldsymbol{H}_{rs}^{*T}(k-1) \tag{4}$$

A Hermitian matrix $\boldsymbol{S}(k-1)$ that is symmetric positive-definite may be Cholesky factorized [1,11] as shown in the following equation:

$$\boldsymbol{S}(k-1) = \boldsymbol{L}_S^{*T}\boldsymbol{L}_S \tag{5}$$

where $\boldsymbol{L}_S$ is an upper triangular matrix.

In order to generate the ensemble of matrices $\hat{\boldsymbol{S}}(k-1)$, the randomization of the matrix $\boldsymbol{S}(k-1)$ is performed according to [1,11]

$$\hat{\boldsymbol{S}}(k-1) = \boldsymbol{L}_S^{*T}\hat{\boldsymbol{G}}_{\boldsymbol{S}}\boldsymbol{L}_S \tag{6}$$

where $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ is a normalized positive-definite matrix [1,11]. The expected value of the $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ matrix is equal to the identity matrix $\boldsymbol{I}_n$. As shown in Eq. (2), the value of the dispersion parameter depends on the size number n of the matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$.

The randomizing matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ is built using [1,11]

$$\hat{\boldsymbol{G}}_{\boldsymbol{S}} = \hat{\boldsymbol{L}}_{\boldsymbol{GS}}^{*T}\hat{\boldsymbol{L}}_{\boldsymbol{GS}} \tag{7}$$

Where the elements ik (row, column) of the random upper triangular matrix $\hat{\boldsymbol{L}}_{\boldsymbol{GS}}$ are defined through the equations below [11]:

$$\text{if } i < k, \quad \hat{\boldsymbol{L}}_{\boldsymbol{GS},ik} = \sigma_S U_{ik} \tag{8}$$

$$\text{if } i = k, \quad \hat{\boldsymbol{L}}_{\boldsymbol{GS},ik} = \sigma_S (2V_i)^{1/2} \tag{9}$$

where

$$\sigma_S = \frac{\delta_S}{(n+1)^{1/2}} \tag{10}$$

U_{ik} is a Gaussian random variable of unit variance and zero mean, and V_i is a Gamma random variable with parameters α, β [1,11] as

$$\alpha = \frac{n+1}{2\delta_S^2} + \frac{1-i}{2} \tag{11}$$

$$\beta = 1 \tag{12}$$

The choice of the normalized positive-definite matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ is motivated by the use of the maximum entropy principle given the information available: there is a mean matrix $\boldsymbol{S}(k-1)$, it is positive definite, and the second-order moment of its inverse exists [1,11]. The theoretical statistical properties of the random matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ can be found in Refs. [1] and [11].

2.4 Probability Density Functions of the Modal Parameters for an Ensemble of Nominally Identical Structures. To calculate a new matrix $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ in each realization of the ensemble, a Monte Carlo simulation in which independent random samples are drawn for the Gaussian random variable U_{ik} and the Gamma random variable V_i, is performed. The number of samples needed to generate the ensemble has to ensure that the PDFs of the uncertain parameters obtained, converge to their true PDF. This is achieved by checking the Kullback–Leibler (KL) divergence - divergence value of the PDFs of each identified parameter with respect to the PDFs for incrementally increasing number of simulations. The modal parameters (natural frequencies, damping ratios, and modal shapes) estimated using ERA from each random realization of the system is used to approximate the PDFs of the virtual ensemble using a kernel density estimate [26].

3 Advantages and Summary of the Steps for the Proposed Method

The main advantages of the proposed approach when compared to forward uncertainty quantification methods with physics-based models are:

(i) The different sources of uncertainties are not required to be specified explicitly.
(ii) The uncertainty in the parameters does not have to be propagated through the equations of motion.
(iii) In some cases, it may be difficult or impractical to model a structure/product based on physics. Consequently, the results obtained would have been significantly affected by the modeling errors. In these occasions, building the ensemble from the measurements of a prototype structure, i.e., a data-driven model, can be advantageous.
(iv) Compared to other SI methods that account for uncertainties, this methodology does not require to make explicit assumptions on the type of the PDFs of specific model parameters. Additionally, the PDFs of the modal parameters may be produced without specifying the origin of such uncertainties, and therefore, no assumptions on the sources of the uncertainties are required.

The results of this method may also be applied to anomaly detection techniques, where the evaluated ensemble's distributions of the modal parameters can be directly used.

The steps proposed in this methodology are:

Step 1: A Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$ of size r by s is built by arranging the outputs $\boldsymbol{y}_k$ (obtained from the measurement signals) as shown on Eq. (1).

Step 2: For a given full row rank Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$, a Hermitian matrix $\boldsymbol{S}(k-1)$ is calculated using Eq. (4).

Step 3: The resulting matrix $\boldsymbol{S}(k-1)$ is symmetric positive-definite and is Cholesky factorized using Eq. (5), where $\boldsymbol{L}_{\boldsymbol{S}}$ is an upper triangular matrix.

Step 4: The dispersion parameter value is defined. A realization $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ of the random matrix $\boldsymbol{G}_{\boldsymbol{S}}$ is calculated using Eq. (7), the theoretical statistical properties of the random matrix $\boldsymbol{G}_{\boldsymbol{S}}$ can be found in Refs. [1] and [11]. The randomization of the matrix $\boldsymbol{S}(k-1)$ is performed according to Eq. (6).

Step 5: The eigenvalues and eigenvectors of the matrix $\hat{\boldsymbol{S}}(k-1)$ are calculated. The eigenvalues and the eigenvectors of the matrix $\hat{\boldsymbol{S}}(k-1)$ are, respectively, the squared singular values and the left singular vectors of the Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$ [23]. The right singular vectors of the Hankel Matrix $\boldsymbol{H}_{rs}(k-1)$ are given by the product matrix $\boldsymbol{H}_{rs}^{*T}(k-1)\boldsymbol{U}\boldsymbol{\Sigma}^{-1}$ [23].

Step 6: During step 5, the singular values of $\boldsymbol{H}_{rs}(k-1)$ are calculated. Physical singular values should be separated from spurious (mathematical) singular values related to the noise in this process that defines the model order of the system [23]. From a practical point of view, the elimination of spurious modes may be performed using stabilization diagrams [27]. In the numerical simulations shown in Secs. 4 and 5, the number of singular values n kept is known.

Step 7: The procedures defined from steps 1–4 are also applied to the shifted Hankel Matrix $\boldsymbol{H}_{rs}(k)$. For every realization, the Gaussian random variable U_{ik} and Gamma random variable V_i are independently resampled to produce different realizations of the random matrices $\hat{\boldsymbol{S}}(k-1)$ and $\hat{\boldsymbol{S}}(k)$.

The randomization of $\hat{\boldsymbol{S}}(k)$ is carried out with the same dispersion parameter value δ_S but a different realization $\hat{\boldsymbol{G}}_{\boldsymbol{S}}$ is used. This is because although the measurements used to construct $\boldsymbol{S}(k-1)$ and $\boldsymbol{S}(k)$ are the same, they are obtained using the Hankel and the shifted Hankel matrices, respectively.

Step 8: Modal characteristics of the system (natural frequencies, modal shapes, and damping defined in the Appendix where further equations and explanations about the referred quantities are provided) are calculated using Eqs. (13)–(15), respectively [22]

$$\omega_i = |\lambda_{c(i)}| \tag{13}$$

$$\boldsymbol{\phi}_i = \tilde{\boldsymbol{C}}\boldsymbol{v}_i \tag{14}$$

$$\zeta_i = -\frac{\text{Real}\left(\lambda_{c(i)}\right)}{|\lambda_{c(i)}|} \tag{15}$$

Each realization of the modal characteristics is stored.

Steps from 4 to 8 are repeated for a prescribed number of times for each realization of the random matrix $\hat{\boldsymbol{S}}(k-1)$ and $\hat{\boldsymbol{S}}(k)$. The modal parameters estimated from each random realization of the system (via a Monte Carlo simulation) are used to approximate the PDFs of the virtual ensemble using a kernel density estimate [26].

It is worth mentioning that the nominal damping matrix of the studied structure is proportional. For the calculation of the damping ratios in Eq. (15), only the real part is taken, this assumes that Rayleigh or proportional damping is present. However, the perturbations introduced by the randomization process introduced by the RMT do not necessarily introduce proportional damping in the system. Therefore, results obtained from Eq. (15) for high values of the dispersion parameter may produce nonphysical values, as shown in Sec. 4.

4 Numerical Results

This section shows the applicability of the proposed approach by considering two different numerical simulations: the first a two degrees-of-freedom (2DOF) mass-spring-damper (Sec. 4.1), and the second a cantilever system (Sec. 4.2). For each simulation, two different dispersion parameter values that translate to a low and high level of uncertainty are considered. In the cantilever system, the physical model, that generated the signal, and the identified model have a different state order. The dispersion parameter is used to calculate the realizations of the random matrices $\hat{\boldsymbol{S}}(k-1)$ and $\hat{\boldsymbol{S}}(k)$. The values of the modal parameters obtained in each realization are used to calculate the probability distributions of the modal parameters of the virtual ensemble of these systems. The effects of the value of the dispersion parameter, the state order, and their probability distributions in this approach are also shown.

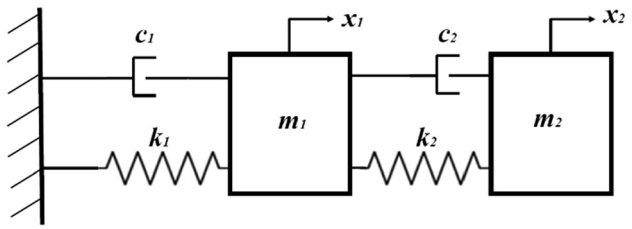


Fig. 2 2DoF mass-spring-damper system

4.1 Two Degrees-of-Freedom System. Let us consider the 2DOF mass-spring-damper system shown in Fig. 2. This is a toy problem that shows the applicability of the algorithm, how it works, and the general trends observed when uncertainty is introduced in the form of a dispersion parameter value.

The 2DOF mass-spring-damper system has equal spring stiffnesses $k_1 = k_2 = 1\,\text{N}\cdot\text{m}^{-1}$; equal masses $m_1 = m_2 = 0.05\,\text{kg}$ and equal viscous dampers with coefficients of damping $c_1 = c_2 = 0.1\,\text{N}\cdot\text{sm}^{-1}$. The system is experiencing a free-vibrations response with nonzero initial conditions: $x_1(0) = x_2(0) = 0$; $\dot{x}_1(0) = 1\,\text{m/s}$; $\dot{x}_2(0) = 0$, where $\dot{(\cdot)}$ represents the derivative with respect to time.

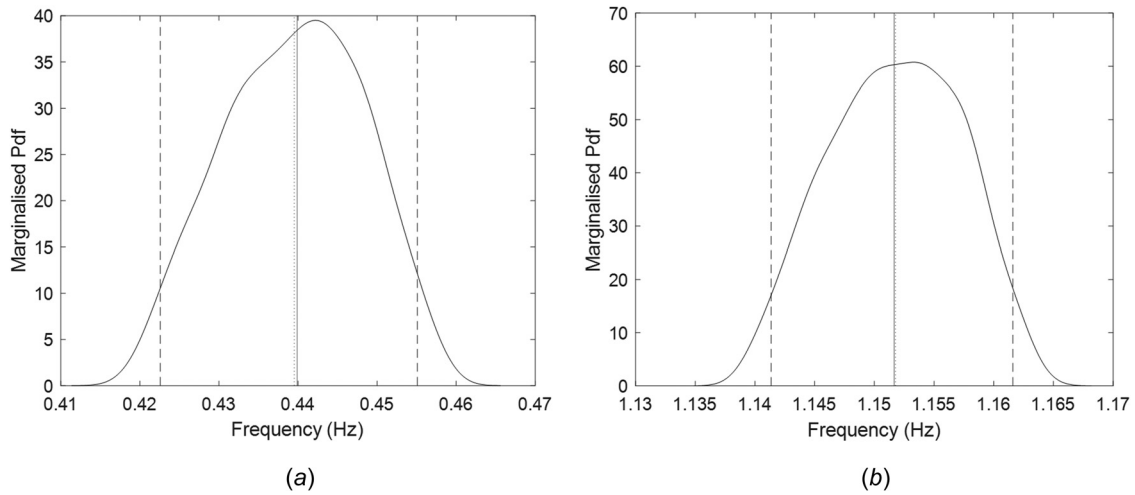
The undamped natural frequencies and damping ratios that correspond to the specified properties of the two-degree-of freedom system are given by: $f_1 = 0.4399\,\text{Hz}$, $f_2 = 1.1517\,\text{Hz}$, $\zeta_1 = 0.1382$, and $\zeta_2 = 0.3618$.

For this example, no measurement noise is considered. Parametric uncertainties are assumed in Sec. 4.1.1. Nonparametric uncertainties are assumed in Sec. 4.1.2. These uncertainties in the system correspond to uncertain modal parameters which are investigated.

This simple example is used to: (i) directly compare the difference between a parametric and nonparametric uncertainty description and (ii) assess the effect of the dispersion parameter value on the resulting PDFs of the modal parameters.

4.1.1 Probability Density Functions Using a Parametric Probabilistic Approach. An example of model parametric uncertainty is shown to illustrate how this affects the uncertainty in the natural frequencies of the 2DOF mass-spring-damper system.

The parametric probabilistic approach requires the assumption of the PDFs of the uncertain input parameters and the propagation of these uncertainties through the equations of motion. Given the unavailability of information, uniform priors were used for both the stiffness, with $k_1 \sim U(0.95, 1.05)$ (N/m) and the mass $m_2 \sim U(0.0475, 0.0525)$ (kg). A Monte Carlo simulation with 10,000 (k_1, m_2) pairs of independent samples was used to obtain the



(*a*) (*b*)

Fig. 3 Probability density functions of modal parameters subject to parametric uncertainty: ····· mean value; ———— 95th percentile; ——— deterministic value: (*a*) first natural frequency and (*b*) second natural frequency

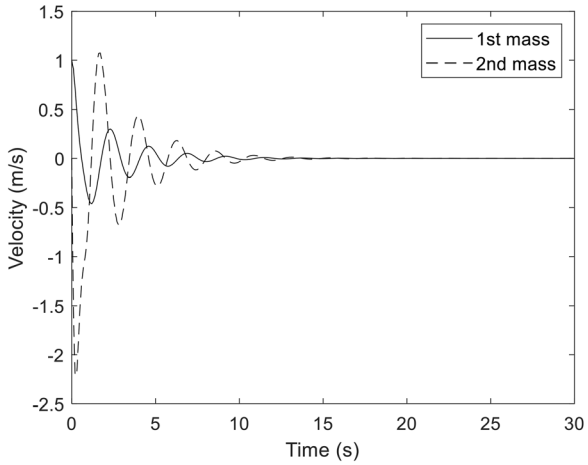


Fig. 4 Velocity of the first and second mass subject to specified initial conditions

PDFs of the natural frequencies of the 2DOF mass-spring-damper system, which are shown in Fig. 3. In this specific example, the computational complexity of the model is low; however, if a more complex model had been used, the computational cost required to obtain estimates of the modal properties would have been substantially higher.

It is well-known that if different parametric distributions would have been used (e.g., Lognormal), the resulting PDFs of the natural frequencies would have been different. Therefore, this approach would require the identification of the uncertainty parameters, the knowledge of the correct parametric uncertainty affecting the system, and the availability of the physics-based model used to propagate the uncertainty in each run of the model to obtain the modal parameters of the system.

4.1.2 Probability Density Functions Using the Proposed Approach. With the proposed approach, no previous knowledge on the uncertainty of specific parameters is needed as described in case (i) of Sec. 2.3. The statistics of the modal parameters (natural frequencies, damping ratios, and modal shapes) are investigated for two cases: low ($\delta_S = 0.001$) and high dispersion parameter ($\delta_S = 0.3$). These two values correspond to a low and a high level of uncertainty, without explicitly identifying the source of this uncertainty. For this numerical case, the state order of the identified system is equal to four.

The free decay of the system is simulated for thirty seconds and the values of $\dot{x}_1$ and $\dot{x}_2$ (that are the velocity of mass 1 and mass 2, respectively) are recorded using a sampling frequency of 10 Hz. The physics-based model is not needed in any of the following steps. These velocity signals shown in Fig. 4 are used to construct the Hankel Matrix $\boldsymbol{H}_{rs}$ described in Eq. (1) for $\boldsymbol{H}_{rs}(k-1)$ and $\boldsymbol{H}_{rs}(k)$. The values of r and s in Eq. (1) have been set to be equal to 50.

The algorithm's steps 3–7 in Sec. 3 are run 10,000 times to produce each time a realization of the random matrices $\hat{\boldsymbol{S}}(k-1)$ and $\hat{\boldsymbol{S}}(k)$. For the modal shapes, the mean value, the deterministic value, and the 95th percentile bounds were calculated. To ensure

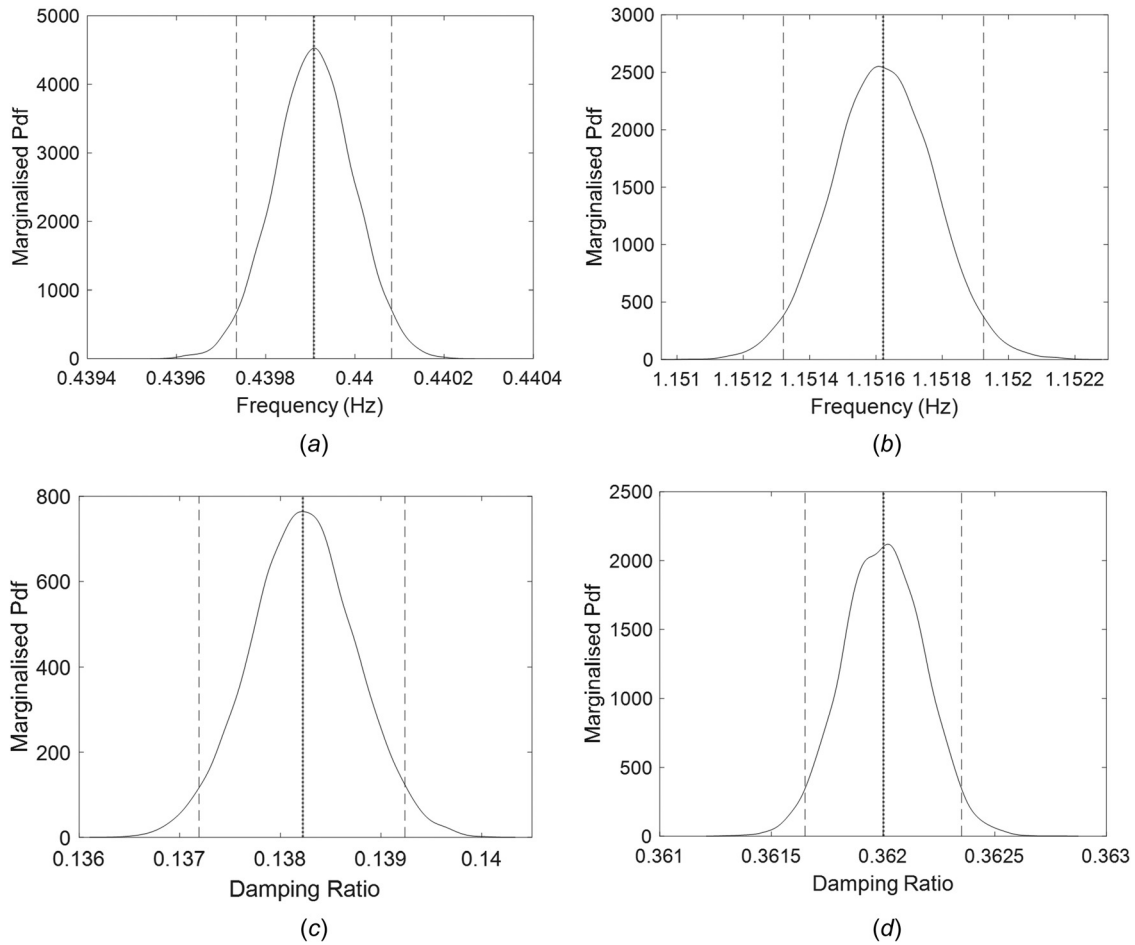


Fig. 5 Probability density functions of modal parameters for a dispersion value: $\delta_S = 0.001$: ····· mean value; ——— 95th percentile; ——— deterministic value: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) first damping ratio, and (*d*) second damping ratio

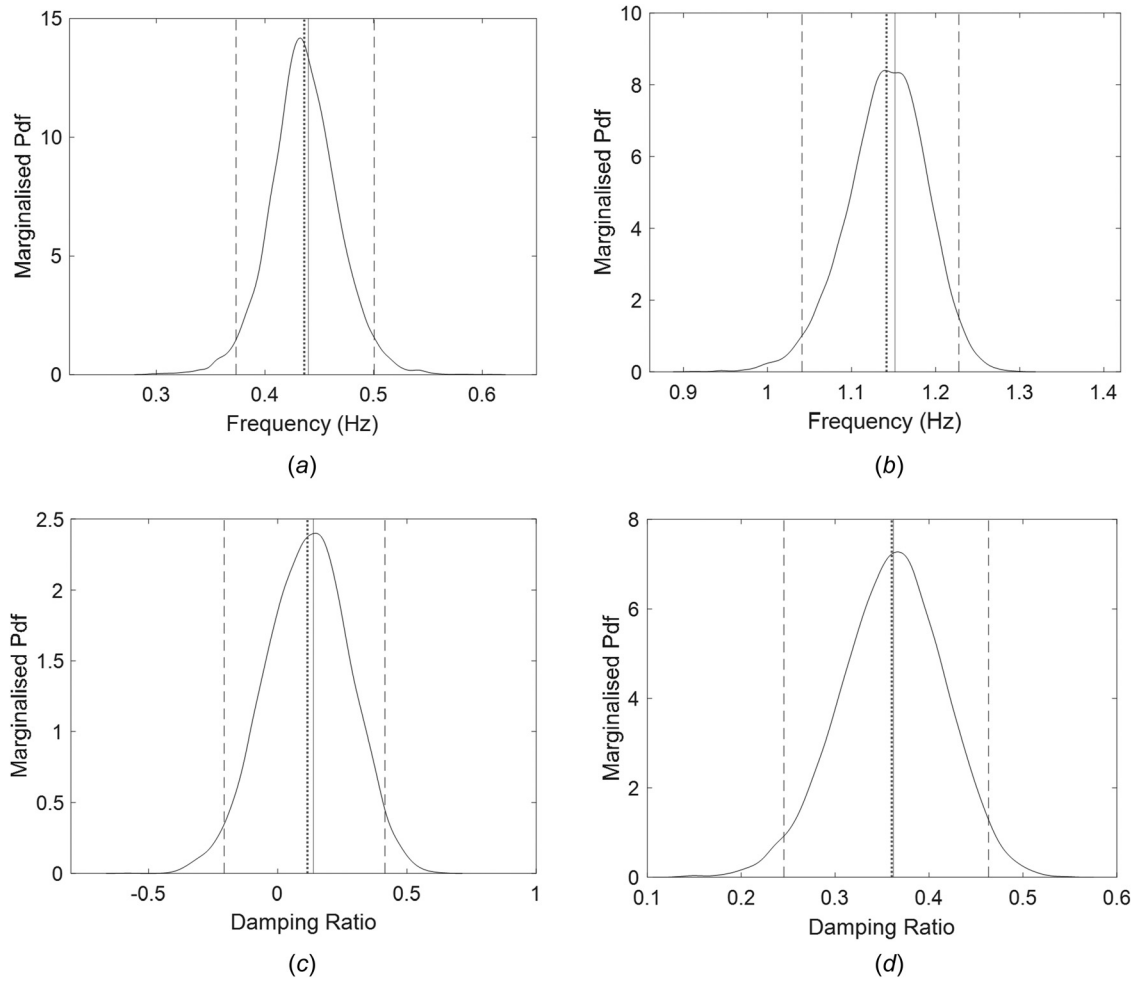


Fig. 6 Probability density functions of modal parameters for a dispersion value: $\delta_S = 0.3$: ····· mean value; ———— 95th percentile; ——— deterministic value: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) first damping ratio, and (*d*) second damping ratio

that convergence was achieved after those 10,000 simulations, the KL-divergence value of the PDFs of each identified parameter with respect to the PDFs for incrementally increasing number of simulations was calculated. It was found that when comparing the distributions obtained from 5000 versus 10,000 simulations, the KL-divergence value was approximately equal to zero for each identified parameter.

The PDFs in Figs. 5 and 6 are obtained using a 200 points kernel smoothing function with the ks density function in MATLAB [28] on the samples (f, ζ, ϕ) produced by the combined ERA and RMT method.

In this section, the general qualitative trends on these results are stated. In Sec. 5, the physical significance and limitations of these

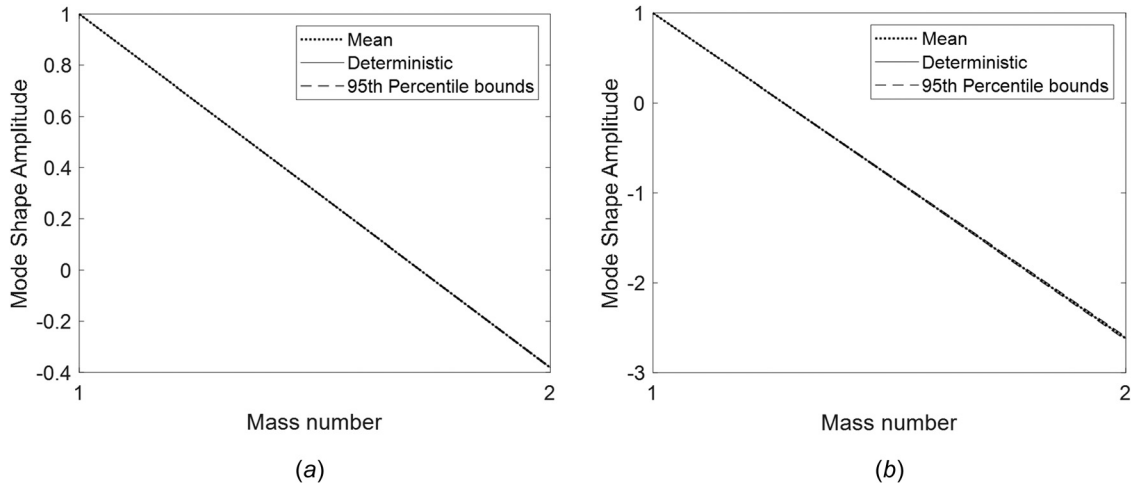


Fig. 7 Modal shapes for a dispersion value $\delta_S = 0.001$: (*a*) first modal shape and (*b*) second modal shape

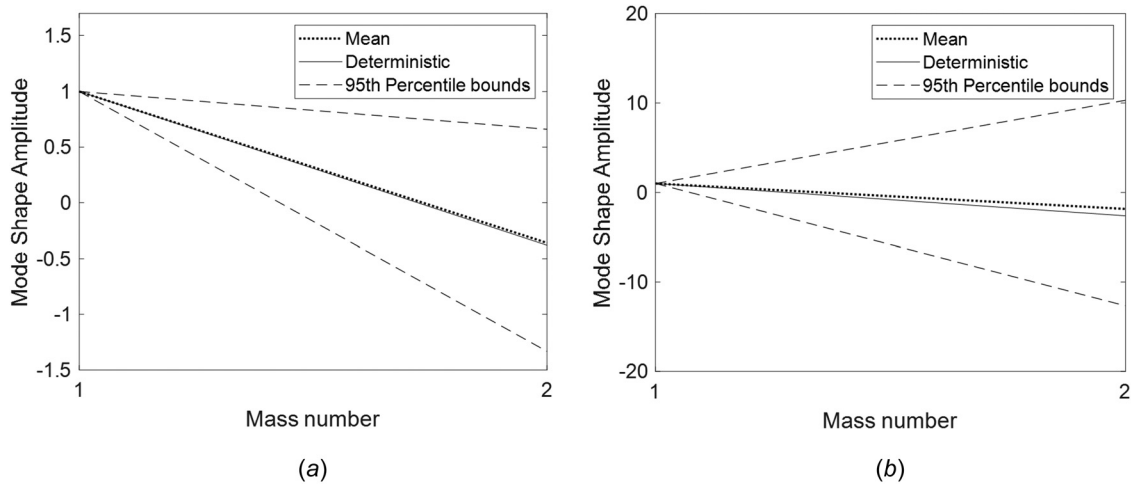


Fig. 8 Modal shapes for a dispersion value: $\delta_S = 0.3$: (*a*) first modal shape and (*b*) second modal shape

results obtained are explained by using new simulations where the dispersion parameter value and the Hankel Matrix size is varied.

The modal parameter f, ζ PDFs shown in Fig. 5 and the modal shapes ϕ in Fig. 7 are calculated for a given dispersion parameter value $\delta_S = 0.001$. The dispersion parameter determines the randomization level of the matrices $\hat{\boldsymbol{S}}(k-1)$ and $\hat{\boldsymbol{S}}(k)$, for this case a low level of uncertainty is introduced. As it was expected, for this level of randomization, the deterministic and mean values of the uncertain modal parameters are found to be approximately equal. The PDFs are also found to be fairly symmetrical.

For the case with higher dispersion value ($\delta_S = 0.3$), which corresponds to higher uncertainties for the identified parameters (Figs. 6 and 8), the PDFs of the identified modal parameters are found to be slightly asymmetrical. Negative values in the support of the PDF are also found for the first damping ratio. The mean and deterministic values for the modal shapes

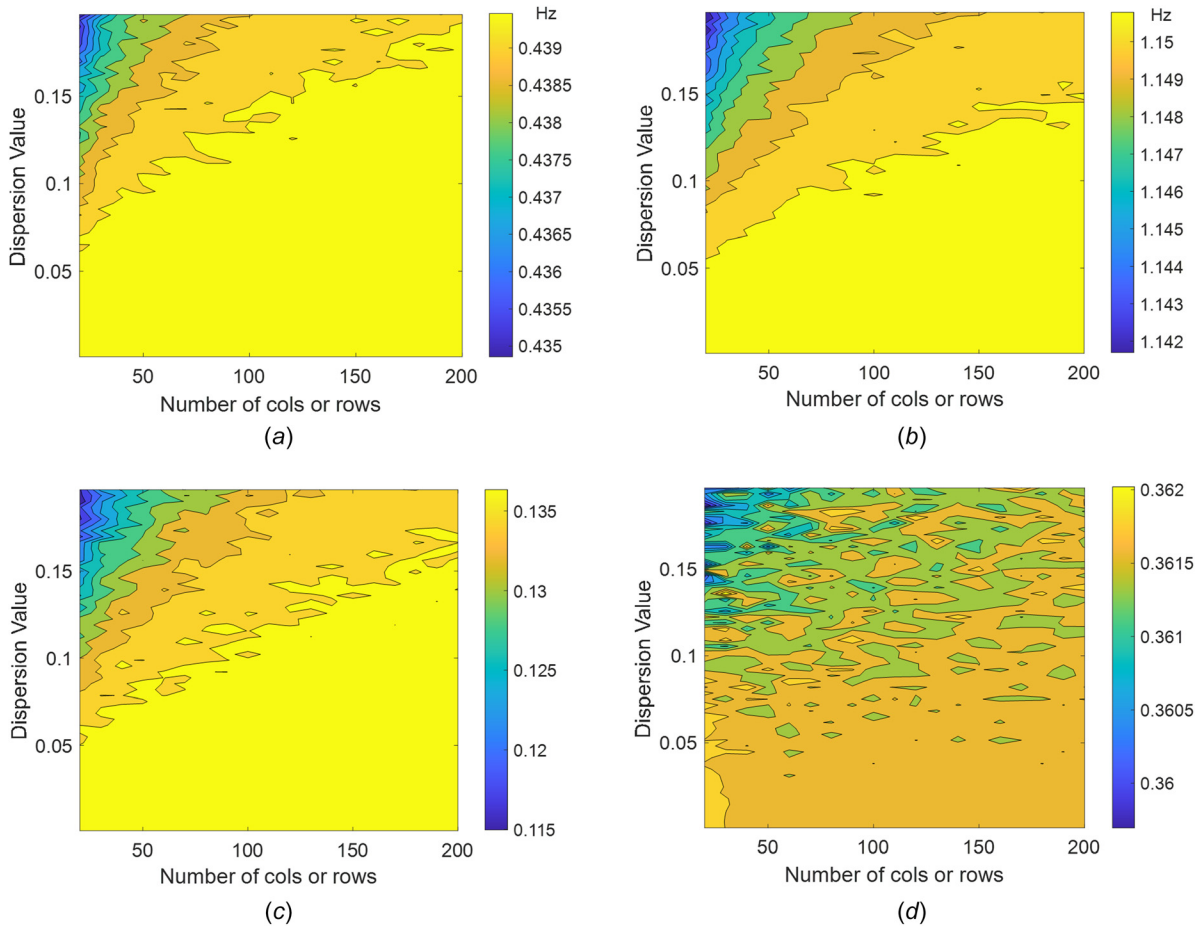


Fig. 9 Contour plots of natural frequencies and damping ratios of their mean values: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) first damping ratio, and (*d*) second damping ratio

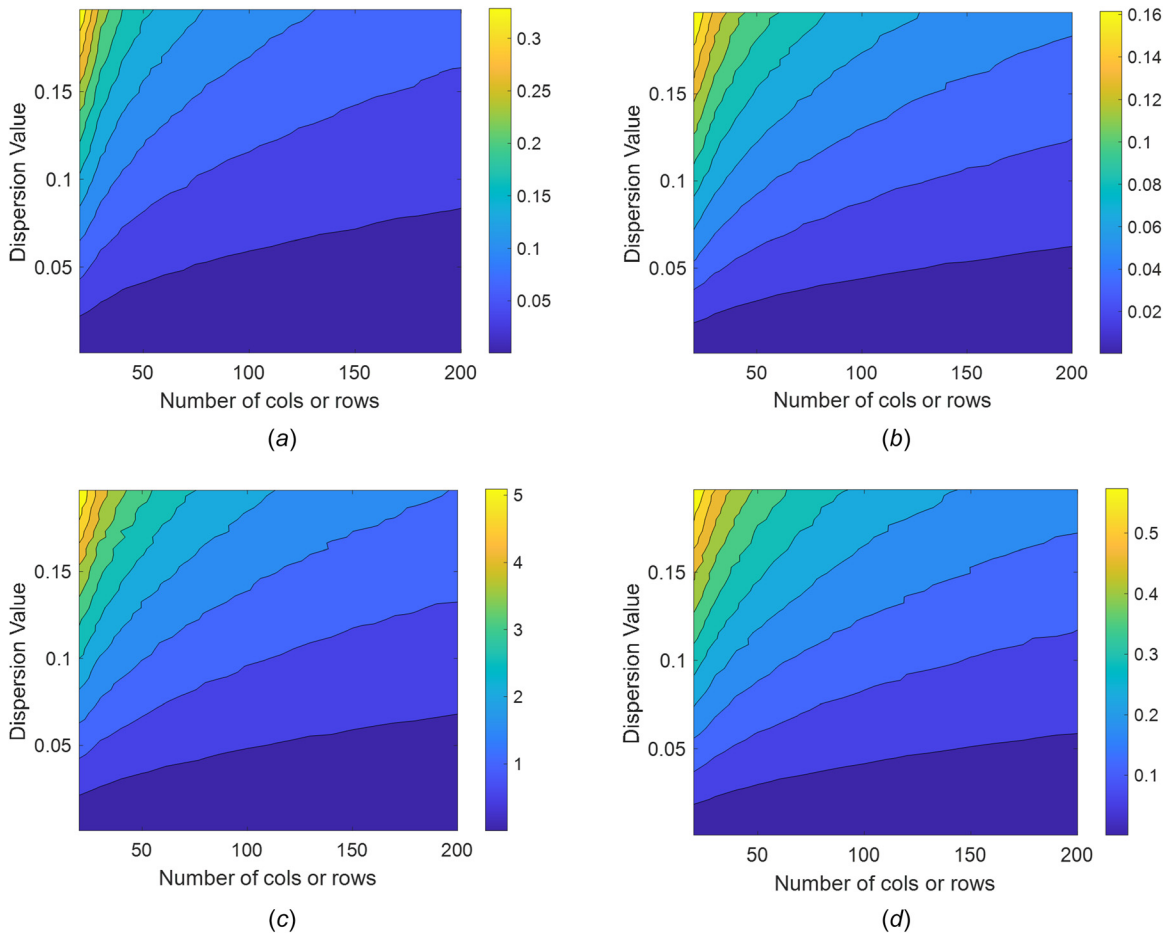


Fig. 10 Contour plots of natural frequencies and damping ratios of their 95% confidence interval divided by their mean values: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) first damping ratio, and (*d*) second damping ratio

are shown to be quite similar with a slightly higher deviation between them for the second mode. The 95% percentile bounds increase as the mode number increases. These effects are also shown in Figs. 9 and 10 and are explained in detail in Sec. 5.

As expected, higher levels of uncertainty are seen for the case with higher dispersion value. In this case, there is a higher deviation between the mean and deterministic values, and the range of the 95th percentile bounds increases compared to the case with low dispersion value.

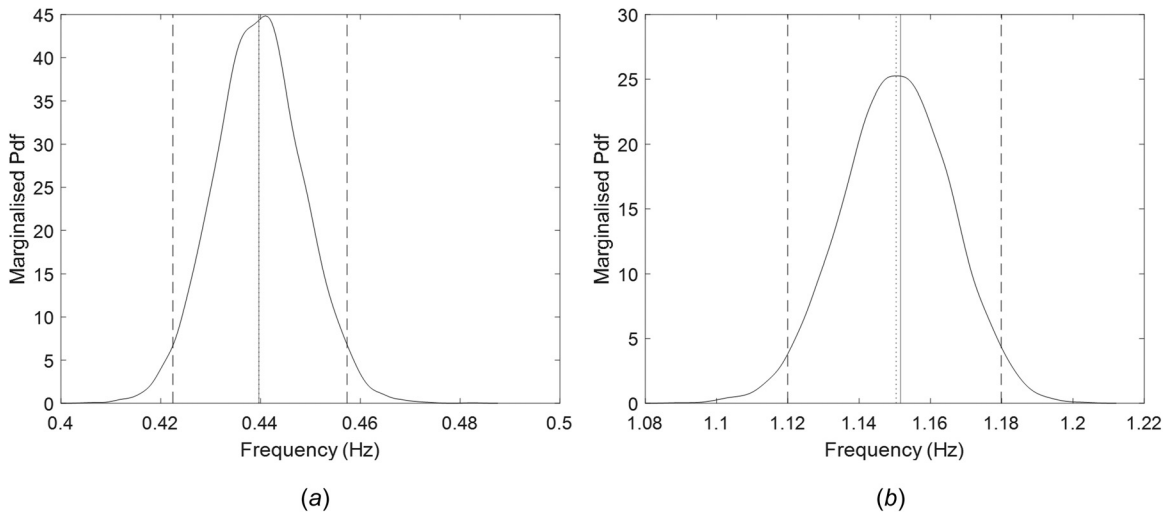


Fig. 11 Probability density functions of modal parameters given equivalent dispersion parameter value $\delta_S = 0.098$: ····· mean value; ——— 95th percentile; —— deterministic value: (*a*) first natural frequency and (*b*) second natural frequency

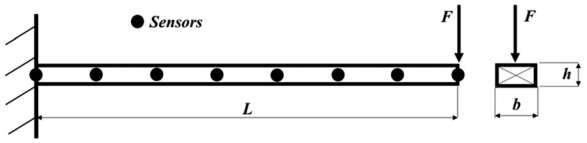


Fig. 12 Schematic representation of the cantilever system

It should be noted that slight discrepancies between the mean identified values and the deterministic values of the system appear. These discrepancies are observed in the form of a shift in the mean identified values. This shift is shown in Fig. 9 where the mean identified values are affected by the dispersion parameter and the size of the Hankel Matrix. The shapes of PDFs of the identified parameters change significantly, and it has to be noted that although the mean is usually a good descriptor, it may not be the best for cases with high dispersion values.

It has been found, that as expected, when a small dispersion parameter δ_S is considered, results are close to the deterministic values.

4.1.3 Comparison of Results. For the parametric uncertainty example in Sec. 4.1.1, the resulting densities of the modal parameters are obtained from the prior knowledge of the uncertain physical parameters k_1 and m_2, by propagation of uncertainties through the physical model. Therefore, for this parametric approach, access to the model is required, and the computational cost involved in the propagation of that parametric uncertainty is dependent on the complexity of that model.

For the nonparametric uncertainty method in Sec. 4.1.2, a physical model is not required for calculating the PDFs of the modal parameters. Calculations are based on the knowledge of the dispersion parameter value, and measurements obtained from either a prototype or numerical 'measurements' from the nominal model.

The main difference between the parametric and nonparametric method is how the source of uncertainty is defined [1]. In the parametric method, each source of uncertainty needs to be specified; however, for the nonparametric approach, it is not specified. As shown in Example 4.1.2, the overall uncertainty of the system is encompassed in the dispersion parameter.

The dispersion parameter value that yields the same coefficient of variation (2%) of the first natural frequency for the parametric case, is calculated using the maximum likelihood method [1] as described in case (ii) in Sec. 2.3.

This equivalent dispersion parameter is $\delta_S = 0.098$. By comparing Figs. 4 and 11, it is possible to observe that the PDF of the first natural frequency is similar to the one obtained using the parametric probabilistic approach. However, the second natural frequency obtained from the proposed approach has a higher coefficient of variation compared to that obtained from the parametric probabilistic approach. This is expected, since the dispersion parameter would inherently account for additional model-parameter uncertainties. The parametric probabilistic approach is not capable of considering those uncertainties. Even though the above is only shown for the parametric case of model parameters (stiffness and mass), the same methodology can be applied for other cases such as the case of known measurement noise.

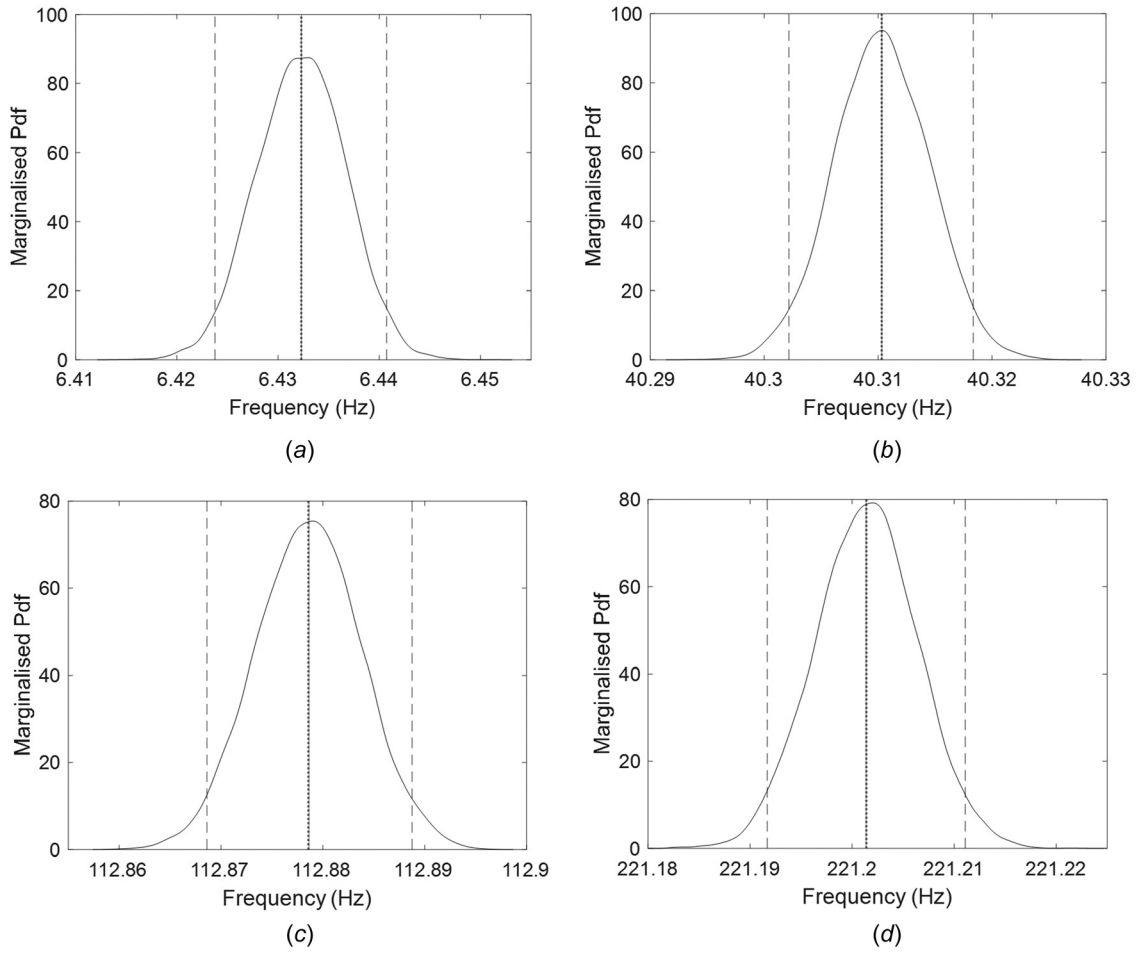


Fig. 13 Probability density functions of natural frequencies for a dispersion value: $\delta_S = 0.001$: ····· mean value; ———— 95th percentile; —— deterministic value: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) third natural frequency, and (*d*) fourth natural frequency

However, the user may believe that substantial modeling errors are still present and would be dominating the overall uncertainty. Therefore, the user may represent the additional level of uncertainty with $\delta_A = 0.202$ to yield an overall dispersion parameter of $\delta_S = 0.3$ as shown in the example in Sec. 4.1.2.

As a result, the sources of uncertainty for the parametric and the nonparametric methods are different. Therefore, the overall results obtained from these two methods are not directly comparable. It should be noted that the method to be chosen is dependent on the information available on the system and the sources of uncertainty.

4.2 Cantilever System. The cantilever system shown on Fig. 12 is the continuous system under investigation in this second example.

A force $\boldsymbol{F}$ (triangular pulse of length 4 ms and maximum amplitude of 2000 N) is used to excite the tip of a cantilever beam modeled as a rectangular Euler–Bernoulli beam with uniform density. The following geometric and material properties were used: L (length) = 1.5 m; b (base) = 0.05 m; h (height) = 0.03 m; ρ(density) = 7850 kg/m^3; E(Young's modulus) = 70 GPa. The cantilever has eight velocity sensors that are at equally spaced intervals (Fig. 12). The velocity values are recorded over time with a sampling frequency of 1000 Hz in order to capture the frequency content well-excited by the triangular pulse. As a rule of thumb, the modes of the system that are below a maximum frequency given by 1.5 times the inverse of the pulse duration are used in the simulation. Therefore, for this system, the number of modes that are used are those with frequencies below 375 Hz.

The response of the continuous system to the force $\boldsymbol{F}$ is numerically simulated using only the first five modes. These free vibration responses are the signals used to construct the Hankel Matrices $\boldsymbol{H}_{\text{rs}}$ described in Eq. (1) for $\boldsymbol{H}_{\text{rs}}$ (k–1) and $\boldsymbol{H}_{\text{rs}}$ (k). The values of r and s on the Hankel matrices have been set to be equal to 100.

The natural frequencies and damping ratios that correspond to the specified properties of the cantilever system are given by: $f_1 = 6.4$ Hz, $f_2 = 40.3$ Hz, $f_3 = 112.9$ Hz, $f_4 = 221.2$ Hz, $f_5 = 365.6$ Hz, $\zeta_1 = \zeta_2 = \zeta_3 = \zeta_4 = \zeta_5 = 0.1$.

For this numerical case, the order of the identified system is assumed to be equal to eight. This corresponds to the assumption that the measured signals are given by the contribution of a model that is described entirely by four modes. A lower order model, than the one that generated the data, has been chosen to observe if the identified modal parameters distributions are affected by this choice.

The statistics of the modal parameters (natural frequencies, damping ratios and modal shapes) are obtained for two cases: low, $\delta_S = 0.001$, and high dispersion parameter, $\delta_S = 0.075$. The same KL-divergence checks and PDF estimations performed in Sec. 4.1 were also carried out for this system.

The PDFs of the modal parameters shown in Figs. 13–15 are calculated using the method described in Sec. 3 for a dispersion parameter value $\delta_S = 0.001$. With this small dispersion value, it is shown that for the identified modal parameters the deterministic value is approximately equal to the mean value and the PDFs are also fairly symmetrical. The range of the 95% percentile bounds

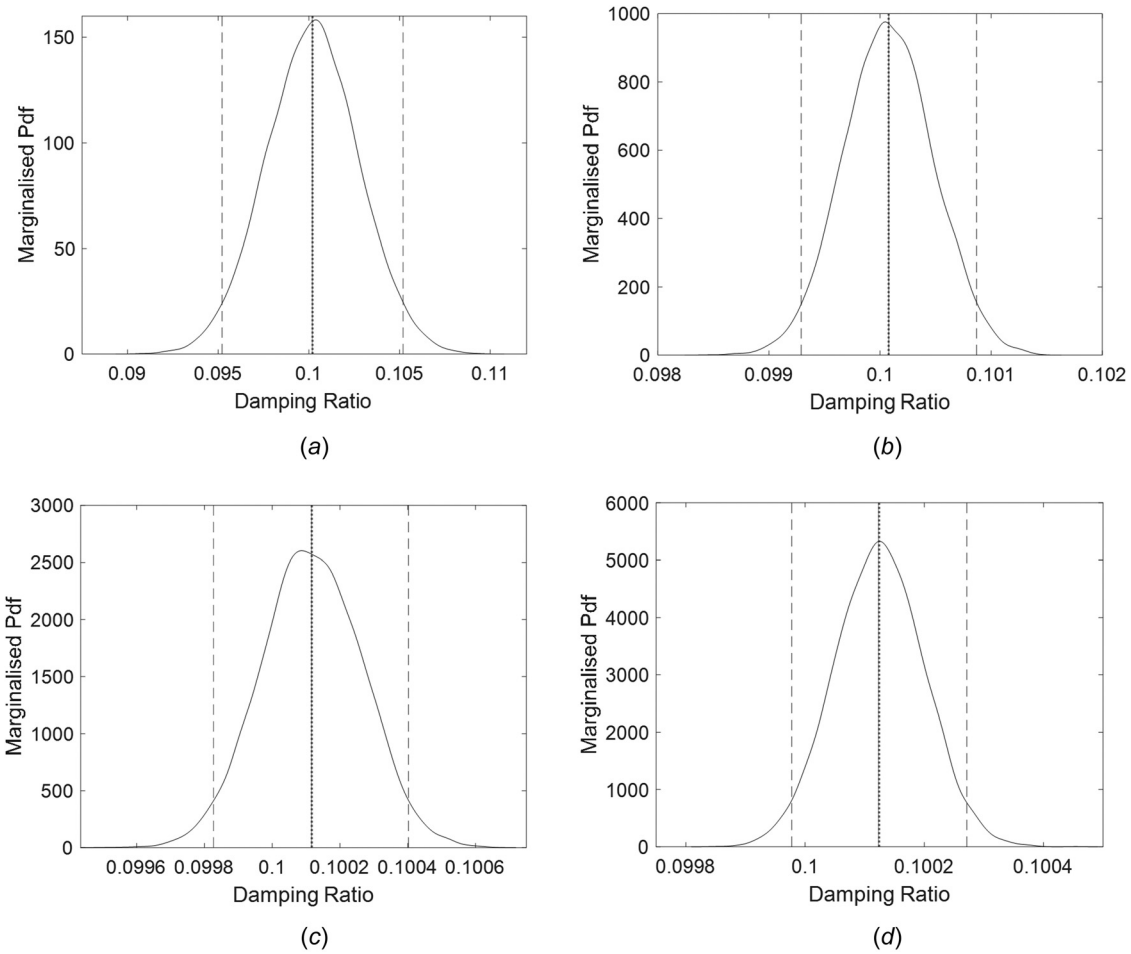


Fig. 14 Probability density functions of damping ratios for a dispersion value: $\delta_S = 0.001$: ····· mean value; ———— 95th percentile; ——— deterministic value: (*a*) first damping ratio, (*b*) second damping ratio, (*c*) third damping ratio, and (*d*) fourth damping ratio

also appears to be small but it is increased as the mode number increases. It was found that these results follow the same general trend described for the previous case in Sec. 4.1.

For the case with higher dispersion value ($\delta_S = 0.075$), which corresponds to a case with higher uncertainties for the identified parameters (Figs. 16–18), the natural frequencies PDFs are asymmetrical. The first and second damping ratios show negative values in the support of the PDF. As explained in Sec. 3, this phenomenon occurs as a consequence of the randomization process of the method. The appearance of negative values may occur because the realizations of the ensemble may produce damping matrices that are nonproportional. The mean and deterministic values of the natural frequencies show minor discrepancies. However, for the modal shapes, larger discrepancies are found between the mean and deterministic values. Assuming that classical damping/proportional damping is observed, the modal shapes obtained are normalized by defining a unit modal amplitude for the eighth node of the system. In some rare occasions, damping may be potentially nonproportional and therefore, considering real mode shapes may not be optimal [29]. The 95% percentile bounds increase as the mode number increases. The results of this case are also found to follow the general trend found in Sec. 4.1, and their significance is also discussed in Sec. 5.

Also, in this numerical application, higher deviation between the mean and deterministic values is seen for the case with higher dispersion value, and the range of the 95th percentile bounds increases compared to the case with low dispersion value. The PDF shapes of the identified parameters change significantly with respect to the case with lower dispersion values.

In this example, the physical model that generated the signal, and the identified model have a different state order. The results show that for the chosen assumption of the state order of the model identified, the obtained results are not affected by the truncation order.

5 Analysis of the Effect of the Dispersion Value and the Hankel Matrix Size on the Modal Parameters

This section further investigates the two-degree-of-freedom system (Sec. 4.1) and the cantilever system (Sec. 4.2) models.

The examples were chosen in such a manner that the similarities and common trends between them, when the method described in Sec. 3 is applied, can be observed. For the 2DOF system the state order chosen for identification purposes corresponds to the same order of the system that generated the data. However, for the cantilever system, the state order of the identified model (eight) is lower than the order of the physical model that generated the signal. The numerical results obtained when the state order is different to the state order that generated the signal in Sec. 4.2 show that the identified modal parameters are still of significance and have not been largely affected either by the truncation order or the randomization process in this example. This is because the loading was such that primarily the first four modes were excited.

For both simulations, it is observed that for low levels of dispersion parameter values, the variability introduced in the modal parameters is low. The mean identified modal parameters, the identified parameters from the unperturbed Hankel matrix and the modal parameters used to simulate the velocity signal are

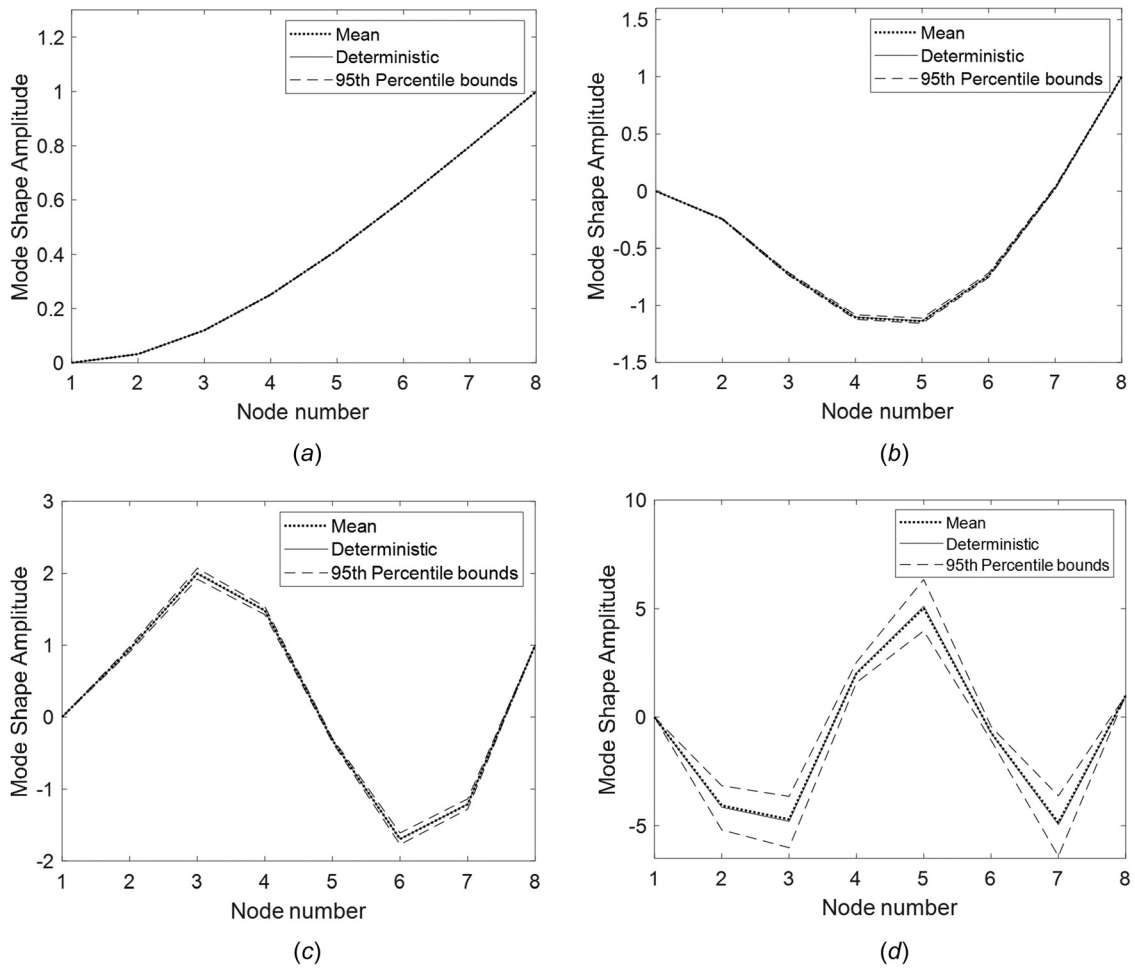


Fig. 15 Mode shape uncertainties for a dispersion value: $\delta_S = 0.001$: (*a*) mode 1, (*b*) mode 2, (*c*) mode 3, and (*d*) mode 4

approximately equal. However, both examples demonstrate that high levels of dispersion value may lead to a shift between the mean of the occurring distribution of natural frequencies and damping ratios and the underlying nominal values. The uncertainty introduced on the modal parameters is higher than the one introduced by lower levels of dispersion value.

Therefore, further simulations of the two-degree-of-freedom system are used to investigate the physical significance and limitations of the results obtained from the application of the method described in Sec. 3. Contour plots for the two-degree-of-freedom system have been created for varying values of dispersion parameter δ_S and numbers of columns/rows used to construct the Hankel Matrix and are presented in the Fig. 9. In Fig. 9, the same number of rows and columns was used, and it corresponds to the number value in the horizontal axis of the figure and the data used are taken from the beginning of the signal. The mean contour plots (Fig. 9) show that there are regions where the mean values of the uncertain parameters identified are the same as the deterministic value of the system. However, if the dispersion value is significant and the number of columns/rows used in the Hankel is low, a bias on the average value of the parameters identified can be observed.

For the particular case in Fig. 9, the mean values of the identified parameters decrease as the dispersion value increases and as the number of columns/rows decreases. This is because the RMT nonparametric method introduces a nonphysical coupling when constructing the random matrices used in steps 3 and 6 of Sec. 3. In this case, the nonphysical coupling may be due to the coupling of the measured signals at the different observed positions. This nonphysical coupling was observed before in Ref. [21] when a nonparametric approach was used for the modeling of uncertainties of mechanical systems. RMT was applied in Ref. [21] to the mass and stiffness matrix, resulting in remote coupling of nonadjacent degrees-of-freedom as a result of the randomization process.

Therefore, it is found that high values of dispersion parameter affect the underlying physical system to be identified.

It can be noted that in principle, the dispersion parameter should be specified as a function of the Hankel Matrix size and cannot be defined independently. That is, the dispersion parameter for a given Hankel matrix size should be defined. To explain the shape of the 95% Confidence Interval/Mean contour plots shown in Fig. 10, two different effects should be considered. First, as the number of points used to construct the Hankel Matrix increases (i.e., more time steps of the measured signals are used), and for a fixed level of noise, the identified values will tend toward the deterministic value up to a limit where the number of data points will not increase the accuracy of the identified value. Second, the effect of keeping the dispersion parameter fixed and increasing the number of columns/rows *rs* of the Hankel Matrix, is that the uncertainty of the system is reduced, as shown in Fig. 10. This double effect is more significant for an identified system with greater dispersion value and a smaller number of columns/rows *rs* of the Hankel matrix than for a system with higher number of columns/rows *rs* of the Hankel matrix and the same level of dispersion value. This shows that the 95% Confidence Interval/Mean values are less sensitive to changes in the dispersion value when Hankel matrices of large size are

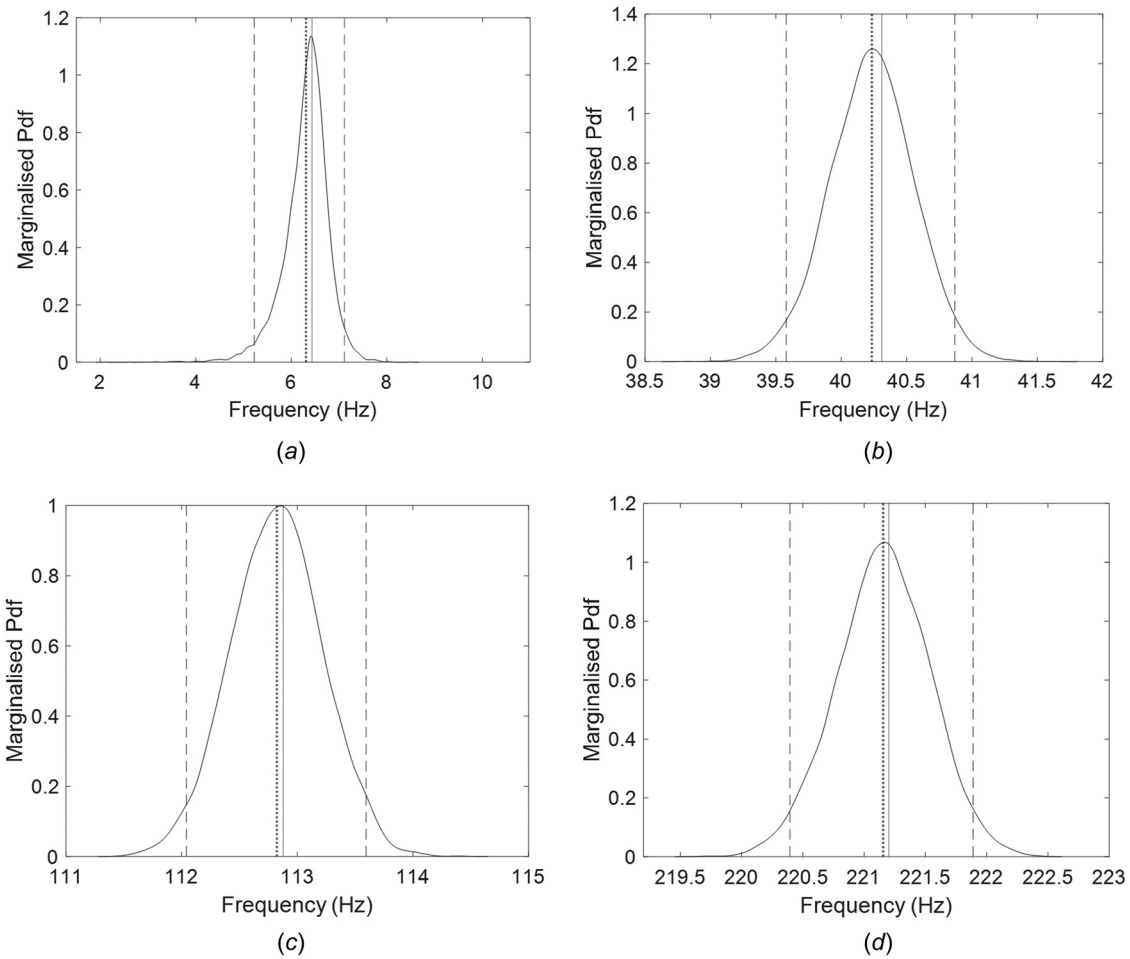


Fig. 16 Probability density functions of natural frequencies for a dispersion value: $\delta_S = 0.075$: ····· mean value; ———— 95th percentile; —— deterministic value: (*a*) first natural frequency, (*b*) second natural frequency, (*c*) third natural frequency, and (*d*) fourth natural frequency

used. This is in fact the reason that the 95% confidence interval/mean contour plots have this specific shape and both effects are observed.

A complementary investigation is performed as shown in Fig. 19, considering that the signal used to construct the Hankel matrix is contaminated by white Gaussian noise. A Monte Carlo simulation with 10,000 samples is used to build the contour plot on Fig. 19, that has as its *y*-axis varying levels of standard deviation of a white Gaussian noise contaminated signal and as the *x*-axis the number of columns/rows used to construct the Hankel matrix. For a small number of columns/rows the contour lines have a positive gradient. It is also seen that in Fig. 19, after the number of rows/columns increase to a specified value that depends on the standard deviation of the white Gaussian noise, the contour lines become horizontal. In Fig. 19, the first effect described previously on the value of the 95% confidence interval/mean in Fig. 10 is also clearly shown.

The uncertainty introduced by the dispersion parameter affects the modal parameters' uncertainty differently. As the dispersion value is increased, the uncertainty for both the natural frequency and the damping ratio estimation is shown to decrease if the mode number is also increased, as shown in Fig. 5 in Sec. 4.1 and also in Sec. 4.2 for both Figs. 16 and 17. Therefore, the highest levels of uncertainty are found on the lower mode number natural frequencies and damping ratios as shown in Figs. 5, 6, 13, 14, 16, and 17. For the cases studied, the modal shape uncertainty is observed to increase as the mode number is increased as shown in Figs. 8, 15, and 18.

6 Conclusions

This paper presents a procedure for the evaluation of the PDFs of the modal parameters for an ensemble of nominally identical structures when there is only access to a single member of the ensemble, and additional knowledge on the statistical properties of the population expressed through the value of a known dispersion parameter is also available. Given a set of measurements or simulated data, acquired from either a prototype or mathematical model, respectively, a discrete time state-space model description is built. The normalized positive definite ensemble [1,11] is used to randomize the matrix calculated by the multiplication of the Hankel matrix by its own conjugate transpose. ERA is applied to identify the modal parameters (natural frequencies, damping ratios, and modal shapes), for each realization of the random matrix. The results of each realization are used to build the PDFs of the modal parameters of the ensemble.

This method is developed to evaluate the modal parameter distributions using the dispersion parameter to account for the uncertain dynamic response across nominally identical structures originated by uncertainties in the manufacturing processes of structural components, modeling errors, boundary conditions, and assemblage. Different a priori values of the dispersion parameter are used to evaluate the effect of dispersion value on the resulting PDFs. These distributions are important for the performance assessment of the designed structures, as this would enable the selection of designs that are robust to these uncertainties, avoiding extensive

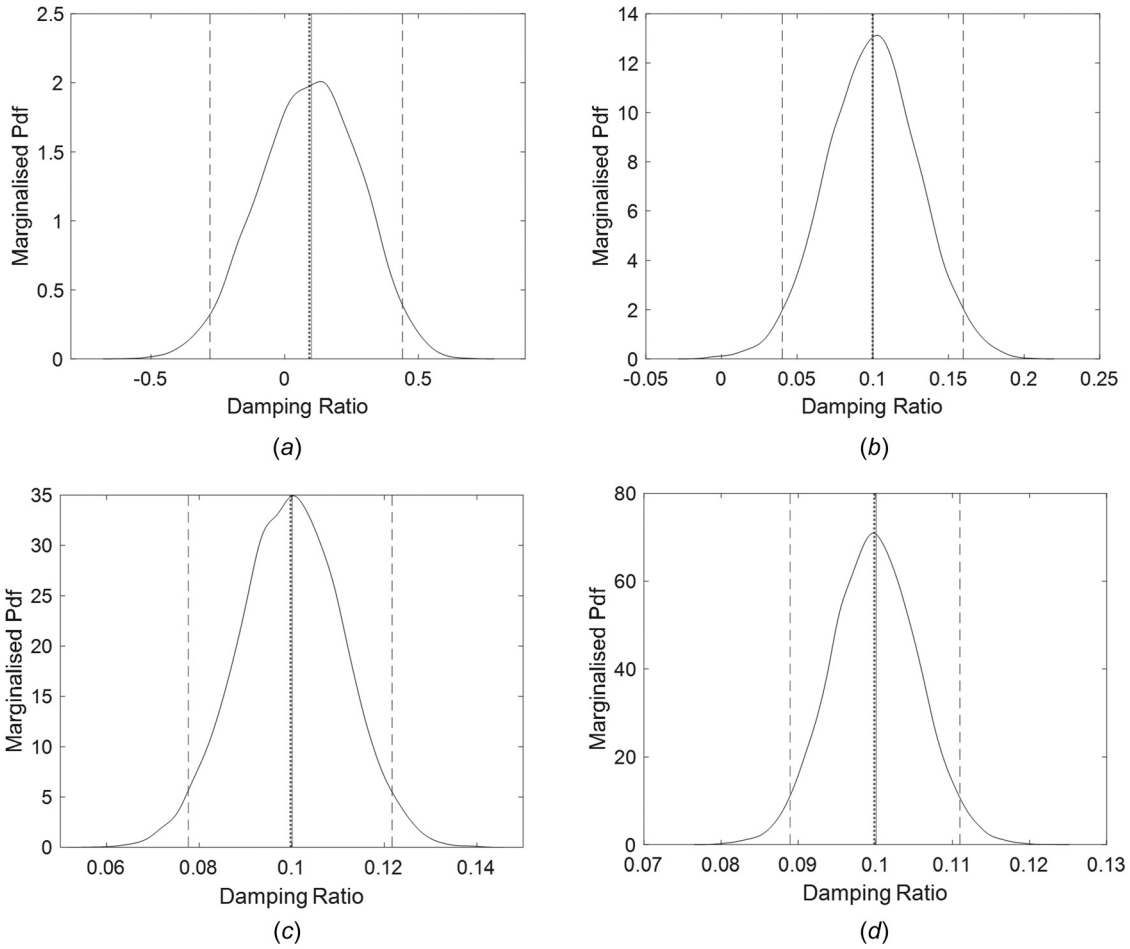


Fig. 17 Probability density functions of damping ratio for a dispersion value $\delta_S = 0.075$: ····· mean value; ——— 95th percentile; ——— deterministic value: (*a*) damping ratio of the first resonant peak, (*b*) damping ratio of the second resonant peak, (*c*) damping ratio of third resonant peak, and (*d*) damping ratio of the fourth resonant peak

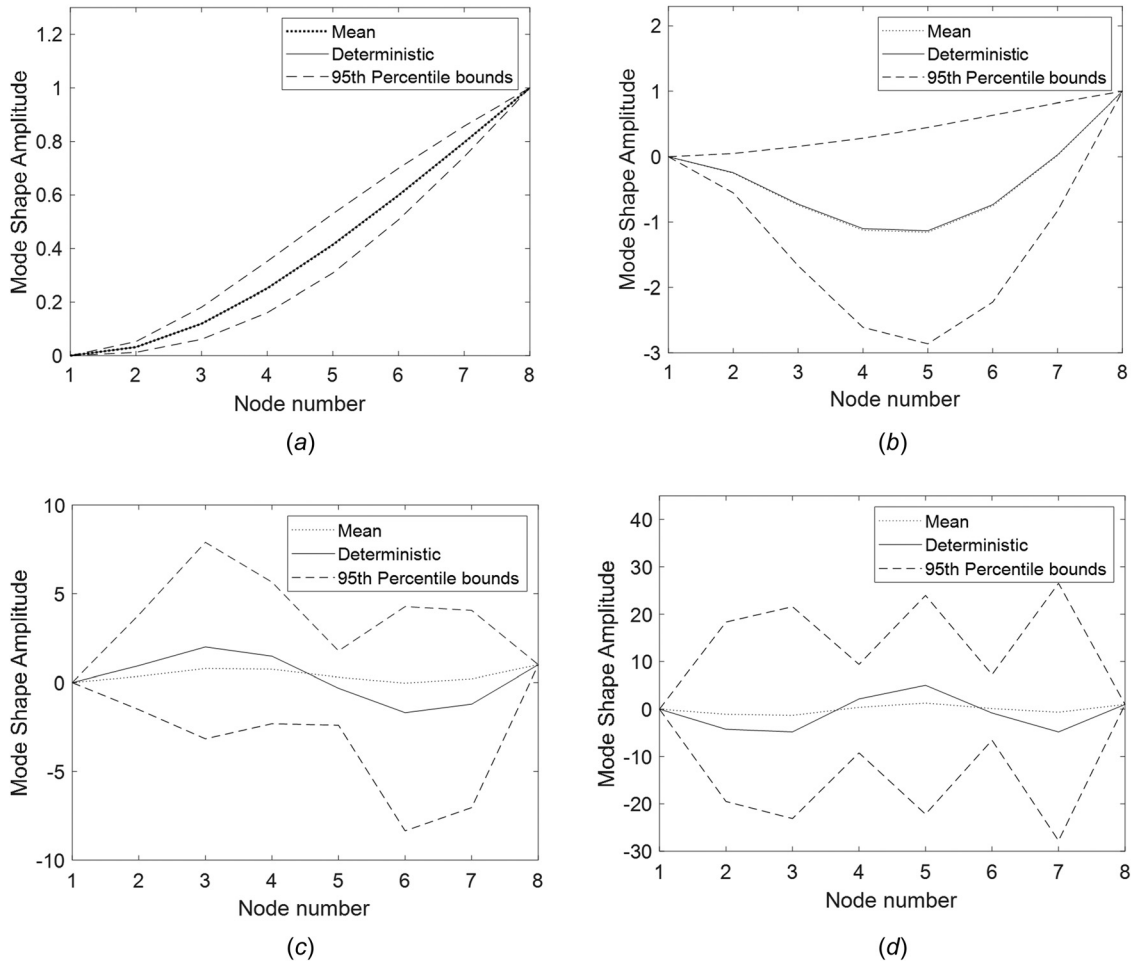


Fig. 18 Mode shape uncertainties for a dispersion value: $\delta_S = 0.075$: (*a*) mode 1, (*b*) mode 2, (*c*) mode 3, and (*d*) mode 4

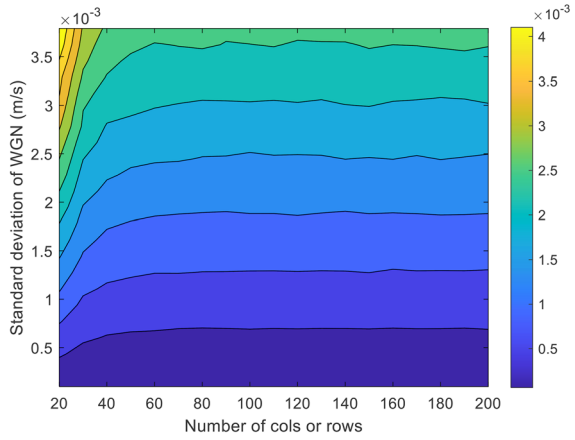


Fig. 19 Contour plot of the 95% confidence interval divided by the mean value for the first natural frequency as a function of the number of the number of columns or rows and the standard deviation of the white Gaussian noise

modifications of the manufactured product. The modal parameter distributions may also be applied in anomaly detection cases.

The main advantage of this approach is that it does not require to specify the different sources of uncertainties in advance. The uncertainty of the parameters does not have to be propagated through the equations of motion, becoming this of great advantage when the model of the system is highly complex or unknown, a circumstance that may introduce modeling errors that affect the overall results. In these circumstances, it may be more convenient to build the ensemble from the measurements of a prototype structure.

Numerical studies were conducted on a two mass-spring-damper and a cantilever system to evaluate the PDFs of the natural frequencies and the damping ratios. The mean and 95th percentile bounds of the modal shapes are also calculated. The study illustrated the performance of the methodology for a wide range of dispersion parameter values. The results obtained from this method generalize well independently of the chosen system (discrete versus continuous system). The physical consequences of the introduction of high values of dispersion were highlighted. It was found that the use of RMT introduces in the calculation a nonphysical behavior, which affects the PDFs of the modal parameters identified. An example of the nonphysical behavior introduced by the method is the appearance of negative values in the support of the PDF of the first damping ratio in the mass-spring-damper system. It was also shown that the state order of the identified system does not affect the identified modal parameters in the cantilever system.

Contour plots were produced to show the effect in the identified parameters' uncertainties of the number of rows/columns used in the Hankel Matrices and the dispersion parameter values. When the dispersion value is significant, and the number of columns/rows used in the Hankel is relatively low, the contour plots of the identified mean values show a systematic shift. This systematic bias increases when either the dispersion value is increased or the number of columns/rows is decreased. Two different effects can

be observed on the 95% Confidence Interval/Mean contour plots. The first one is that increasing number of points used in the Hankel Matrix, the identified values will tend toward the deterministic value, and the second is that when fixing the dispersion parameter value and using larger values of columns/rows, the uncertainty of the system is being reduced. Therefore, the uncertainty introduced in the modal parameters is both dependent on the size of the Hankel Matrix and on the dispersion parameter value.

It is also shown that the uncertainty associated with the natural frequencies and damping ratios is decreased as the mode number increases and/or the dispersion parameter value is increased.

For the case of parametric uncertainty in Sec. 4.1.1, the PDFs of the modal parameters are calculated propagating the prior knowledge of those modal parameters through a physical model. For the nonparametric case in Sec. 4.1.2, only measurements (numerically obtained) from the structure and the value of a dispersion parameter are available. The main difference between the methods is that they depend on how the source of uncertainty is defined.

The prior estimation of the dispersion parameter δ_S is illustrated for three cases. The first case assumes that no prior data are available; then, the dispersion parameter δ_S is used as variable in a sensitivity analysis to assess the stochastic response of the model. The second and third cases assume that prior data are available. Given the observed data, the value of the dispersion parameter value can be calculated. However, for the third case, there is also an additional level of uncertainty present that is represented by and additional term that may be added to the dispersion parameter value calculated from the observed data.

Future work will focus on the estimation of the dispersion parameter value for cases when experimental measurements of the ensemble are available. Another potential direction of interest is the use of different SI methods to determine the compatibility of different experimental conditions of the system tested with the approach proposed in this paper.

Funding Data

- EPSRC and Schlumberger for an Industrial Case Postgraduate Scholarship (Grant No. EP/T517653/1; Funder ID: 10.13039/501100000266)
- EPSRC for the Overseas Travel Grants (Grant No. EP/R008949/1).

Appendix: Eigensystem Realization Algorithm

The time evolution of a linear, time invariant, dynamic system can be described using a discrete-time state-space representation

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k \tag{A1}$$

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{D}\boldsymbol{u}_k \tag{A2}$$

Where $\boldsymbol{x}_k \in \mathbb{R}^n$, $\boldsymbol{y}_k \in \mathbb{R}^p$, and $\boldsymbol{u}_k \in \mathbb{R}^m$ are, respectively, the state vector, collected output and the value of the control input which remains constant between times t_k and t_{k+1}. Matrices $\boldsymbol{A} \in \mathbb{R}^{nxn}$, $\boldsymbol{B} \in \mathbb{R}^{nxm}$, $\boldsymbol{C} \in \mathbb{R}^{pxn}$, and $\boldsymbol{D} \in \mathbb{R}^{pxm}$ are the system, input, output and feedthrough matrices, respectively.

The ERA [22] is a system identification method that is used to estimate the modal parameters of a system by evaluating the so-called transformed matrices to the modal-space. To evaluate the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, ERA uses an output vector $\boldsymbol{y}_k$ that contains the measurements read in p sensors at the different times t_k for a system excited by an impulse at time zero and assuming zero initial conditions. ERA obtains a transformation of Eqs. (A1) and (A2) that has as its matrices $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{B}}$ and $\tilde{\boldsymbol{C}}$ which correspond to the state vector $\tilde{\boldsymbol{x}}_k$. The eigenvalues of the matrix $\tilde{\boldsymbol{A}}$ and the sensor-based modal shapes in Eq (A9) are transformation invariant and can therefore be used to obtain the eigenvalues and the sensor-based modal shapes of the original system that generated the data. Subsequently, the eigenvalue decomposition of the matrix $\tilde{\boldsymbol{A}}$ (Eq. (A3)) is used to determine the discrete time eigenvalues λ_i and the corresponding set of eigenvectors $\boldsymbol{v}_i$ from [22]

$$\tilde{\boldsymbol{A}}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i \tag{A3}$$

And the following matrix $\boldsymbol{\Lambda}$ and vector $\boldsymbol{V}$ can then be obtained [22]:

$$\boldsymbol{\Lambda} = \mathrm{diag}(\,\lambda_1, \quad \lambda_2, \quad \ldots \quad \lambda_n\,) \tag{A4}$$

$$\boldsymbol{V} = [\,\boldsymbol{v}_1, \quad \boldsymbol{v}_2, \quad \ldots \quad \boldsymbol{v}_n\,] \tag{A5}$$

The eigenvalues of the matrix $\tilde{\boldsymbol{A}}$ can then be used to calculate the eigenvalues $\lambda_{c(i)}$, natural frequencies and damping factors of the equivalent continuous in time system using Eqs. (A6), (13), and (15), respectively, [22]

$$\lambda_{c(i)} = \frac{\ln(\lambda_i)}{\Delta t} \tag{A6}$$

where Δt is the sampling period. The sensor-based modal shapes are then obtained using Eq. (14) [22].

The steps required to obtain the $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{B}}$, and $\tilde{\boldsymbol{C}}$ matrices are:

Step 1: Build the Hankel matrix $\boldsymbol{H}_{rs}(k-1)$ of size r by s is built that contains time series data from measurement data using Eq. (1) [23].

Step 2: Build the matrix $\boldsymbol{H}_{rs}(k)$, also called the shifted Hankel matrix which is one-time-step into the future of the matrix $\boldsymbol{H}_{rs}(k-1)$

$$\boldsymbol{H}_{rs}(k) = \begin{bmatrix} \boldsymbol{y}_{k+1} & \boldsymbol{y}_{k+2} & \cdots & \boldsymbol{y}_{k+s} \\ \boldsymbol{y}_{k+2} & \boldsymbol{y}_{k+3} & \cdots & \boldsymbol{y}_{k+s+1} \\ . & . & & . \\ . & . & & . \\ \boldsymbol{y}_{r+k} & \boldsymbol{y}_{r+k+1} & \cdots & \boldsymbol{y}_{r+k+s-1} \end{bmatrix} \tag{A7}$$

Step 3: Obtain the approximate expression of the $\boldsymbol{H}_{rs}(k-1)$. By using the singular value decomposition, the Hankel matrix $\boldsymbol{H}_{rs}(k-1)$ can be expressed as [23]

$$\boldsymbol{H}_{rs}(k-1) = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{*T} \tag{A8}$$

In Eq. (A8), ^{*T} denotes complex conjugate transpose, matrices $\boldsymbol{U} \in \mathbb{C}^{rxr}$ and $\boldsymbol{V} \in \mathbb{C}^{sxs}$ are unitary, have orthonormal columns, and matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{rxs}$ has at most s nonzero elements in its diagonal. The exact expression of Eq. (A8) can also be written as [23]

$$\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{*T} = [\,\tilde{\boldsymbol{U}} \quad \boldsymbol{U}_t\,] \begin{bmatrix} \tilde{\boldsymbol{\Sigma}} & 0 \\ 0 & \boldsymbol{\Sigma}_t \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{V}}^{*T} \\ \boldsymbol{V}_t^{*T} \end{bmatrix} \tag{A9}$$

Where $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{\Sigma}}$, and $\tilde{\boldsymbol{V}}^{*T}$ denote truncated matrices containing the dominant singular values and vectors, and where $\boldsymbol{U}_t$, $\boldsymbol{\Sigma}_t$ and $\boldsymbol{V}_t^{*T}$ contain the nondominant singular values and vectors that are discarded in the approximation [23]

$$\boldsymbol{H}_{rs}(k-1) \approx \tilde{\boldsymbol{U}}\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{V}}^{*T} \tag{A10}$$

If the Hankel matrix $\boldsymbol{H}_{rs}(k-1)$ does not have full rank, then $\tilde{\boldsymbol{\Sigma}}$ may contain some zero singular values and the Eq. (A10) will be exact. However, if $\tilde{\boldsymbol{\Sigma}}$ contains a number of nonzero singular values smaller than the rank of the Hankel Matrix, then the expression of Eq. (A10) is only approximate.

Step 4: In this step the realization matrices of the system are estimated using the $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{\Sigma}}$, and $\tilde{\boldsymbol{V}}^{*T}$ matrices as [23]

$$\tilde{\boldsymbol{A}} \approx \tilde{\boldsymbol{\Sigma}}^{-1/2}\tilde{\boldsymbol{U}}^{*T}\boldsymbol{H}_{rs}(k)\tilde{\boldsymbol{V}}\tilde{\boldsymbol{\Sigma}}^{-1/2} \tag{A11}$$

$$\tilde{\boldsymbol{B}} \approx \tilde{\boldsymbol{\Sigma}}^{1/2}\tilde{\boldsymbol{V}}^{*T}\begin{bmatrix} \boldsymbol{I_m} & 0_m \\ 0_m & 0_m \end{bmatrix} \tag{A12}$$

$$\tilde{\boldsymbol{C}} \approx \begin{bmatrix} \boldsymbol{I_p} & 0_p \\ 0_p & 0_p \end{bmatrix}\tilde{\boldsymbol{U}}\tilde{\boldsymbol{\Sigma}}^{1/2} \tag{A13}$$

where 0_p and $\boldsymbol{I_p}$ are, respectively, matrices zero and identity of size $p \times p$.

When the matrices $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{B}}$, and $\tilde{\boldsymbol{C}}$ are used in a system of equations similar to Eqs. (A1) and (A2) become the discrete time modal-space representation of the dynamic system [23].

Using ERA, and the matrices $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{C}}$, deterministic values of the modal parameters of a system are obtained for a given set of measurements. However, these measurements might be affected by different sources of uncertainties such as sensor noise. The existence of those uncertainties is considered in ERA by averaging out the noise of the signals using several measurements to obtain the deterministic values of the modal parameters.

References

[1] Soize, C., 2017, *Uncertainty Quantification -An Accelerated Course With Advanced Applications in Computational Engineering*, Springer International Publishing, Cham, Switzerland.

[2] Wright, M., and Weave, R., 2010, *New Directions in Linear Acoustics and Vibration: Quantum Chaos, Random Matrix Theory, and Complexity*, Cambridge University Press, Cambridge, UK.

[3] Kompella, M. S., and Bernhard, R. J., 1993, "Measurement of the Statistical Variation of Structural-Acoustic Characteristics of Automotive Vehicles," SAE Paper No. 931272.

[4] Durand, J.-F., Soize, C., and Gagliardini, L., 2008, "Structural-Acoustic Modeling of Automotive Vehicles in Presence of Uncertainties and Experimental Identification and Validation," J. Acoust. Soc. Am., **124**(3), pp. 1513–1525.

[5] Zhang, J., Shen, G., Du, Y., and Hu, P., 2013, "Modal Analysis of a Lightweight Engine Hood Design Considering Stamping Effects," Applied Mechanics and Materials, Trans Tech Publications, pp. 364–369.

[6] Reich, S., and Cotter, C., 2015, "Probabilistic Forecasting and Bayesian Data Assimilation," Probab. Forecasting Bayesian Data Assimilation, pp. 1–297.

[7] Langley, R. S., 2000, "Unified Approach to Probabilistic and Possibilistic Analysis of Uncertain Systems," J. Eng. Mech., **126**(11), pp. 1163–1172.

[8] Soize, C., 2013, "Stochastic Modeling of Uncertainties in Computational Structural Dynamics - Recent Theoretical Advances," J. Sound Vib., **332**(10), pp. 2379–2395.

[9] Simoen, E., de Roeck, G., and Lombaert, G., 2015, "Dealing With Uncertainty in Model Updating for Damage Assessment: A Review," Mech. Syst. Signal Process., **56**, pp. 123–149.

[10] Faes, M., and Moens, D., 2020, "Recent Trends in the Modeling and Quantification of Non-Probabilistic Uncertainty," Arch. Comput. Methods Eng., **27**(3), pp. 633–671.

[11] Soize, C., 2005, "A Comprehensive Overview of a Non-Parametric Probabilistic Approach of Model Uncertainties for Predictive Models in Structural Dynamics," J. Sound Vib., **288**(3), pp. 623–652.

[12] Mehta, M. L., 2004, *Random Matrices, Elsevier, Amsterdam, The Netherlands*.

[13] Arnst, M., and Soize, C., 2019, "Identification and Sampling of Bayesian Posteriors of High-Dimensional Symmetric Positive-Definite Matrices for Data-Driven Updating of Computational Models," Comput. Methods Appl. Mech. Eng., **352**, pp. 300–323.

[14] Batou, A., Soize, C., and Audebert, S., 2015, "Model Identification in Computational Stochastic Dynamics Using Experimental Modal Data," Mech. Syst. Signal Process., **50–51**, pp. 307–322.

[15] Ezvan, O., Batou, A., Soize, C., and Gagliardini, L., 2017, "Multilevel Model Reduction for Uncertainty Quantification in Computational Structural Dynamics," Comput. Mech., **59**(2), pp. 219–246.

[16] Vishwanathan, A., and Vio, G. A., 2019, "Numerical and Experimental Assessment of Random Matrix Theory to Quantify Uncertainty in Aerospace Structures," Mech. Syst. Signal Process., **118**, pp. 408–422.

[17] Langley, R. S., and Cotoni, V., 2008, "Response Variance Prediction for Uncertain Vibro-Acoustic Systems Using a Hybrid Deterministic-Statistical Method," J. Acoust. Soc. Am., **122**(6), p. 3445.

[18] Langley, R. S., Cicirello, A., and Deckers, E., 2018, "The Effect of Generalised Force Correlations on the Response Statistics of a Harmonically Driven Random System," J. Sound Vib., **413**, pp. 456–466.

[19] Vishwajeet, K., Singla, P., and Majji, M., 2017, "Random Matrix-Based Approach for Uncertainty Analysis of the Eigensystem Realization Algorithm," J. Guid. Control, Dyn., **40**(8), pp. 1877–1891.

[20] Soize, C., 2005, "Random Matrix Theory for Modeling Uncertainties in Computational Mechanics," Comput. Methods Appl. Mech. Eng., **194**(12–16), pp. 1333–1366.

[21] Legault, J., Langley, R. S., and Woodhouse, J., 2012, "Physical Consequences of a Nonparametric Uncertainty Model in Structural Dynamics," J. Sound Vib., **331**(25), pp. 5469–5487.

[22] Juang, J. N., and Pappa, R. S., 1985, "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction," J. Guid. Control, Dyn., **8**(5), pp. 620–627.

[23] Brunton, S. L., and Kutz, J. N., 2019, *Data-Driven Science and Engineering*, Cambridge University Press, Cambridge, UK.

[24] Caicedo, J. M., Dyke, S. J., and Johnson, E. A., 2004, "Natural Excitation Technique and Eigensystem Realization Algorithm for Phase I of the IASC-ASCE Benchmark Problem: Simulated Data," J. Eng. Mech., **130**(1), pp. 49–60.

[25] Reynders, E., 2012, "System Identification Methods for (Operational) Modal Analysis: Review and Comparison," Arch. Comput. Methods Eng., **19**(1), pp. 51–124.

[26] Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.

[27] Wu, C., Liu, H., Qin, X., and Wang, J., 2017, "Stabilization Diagrams to Distinguish Physical Modes and Spurious Modes for Structural Parameter Identification," J. Vibroeng., **19**(4), pp. 2777–2794.

[28] The MathWorks Inc., 2020, "MATLAB, 9.9.0.1467703 (R2020b)," The MathWorks Inc., Natick, MA.

[29] Sinha, S., 2005, "Techniques for Real Normalization of Complex Modal Parameters for Updating and Correlation with FEM Models," Master's thesis, University of Cincinnati, Cincinnati, OH.