Variance estimation of modal parameters from input/output covariance-driven subspace identification

Philippe Mellinger, Michael Döhler, Laurent Mevel

Variance estimation of modal parameters from input/output covariance-driven subspace identification

Philippe Mellinger [1], **Michael Döhler** [2], **Laurent Mevel** [2]
[1] CEA
DAM, DIF, 91297 Arpajon, France

[2] Inria/IFSTTAR, I4S,
Campus de Beaulieu, 35042 Rennes, France
e-mail: **michael.doehler@inria.fr**

Abstract

For Operational Modal Analysis (OMA), the vibration response of a structure from ambient and unknown excitation is measured and used to estimate the modal parameters. For OMA with eXogenous inputs (OMAX), some of the inputs are known in addition, which are considered as realizations of a stochastic process. When identifying the modal parameters from noisy measurement data, the information on their uncertainty is most relevant. Previously, a method for variance estimation has been developed for the output-only case with covariance-driven subspace identification. In this paper, a recent extension of this method for the input/output covariance-driven subspace algorithm is discussed. The resulting variance expressions are easy to evaluate and computationally tractable when using an efficient implementation. Based on Monte Carlo simulations, the quality of identification and the accuracy of variance estimation are evaluated. It is shown how the input information leads to better identification results and lower uncertainties.

1 Introduction

The goal of operational modal analysis (OMA) is the identification of modal parameters (frequencies, damping ratios, mode shapes) from vibration measurements on a structure. The modal parameters are related to the poles and observed eigenvectors of a linear time-invariant system. Subspace-based linear system identification algorithms have been proven efficient for the identification of such systems, fitting a linear model to output-only measurements or input and output measurements taken from a system [1].

Usually, OMA refers to the identification of modal parameters from output-only measurements under unknown ambient excitation (e.g. wind, traffic). Besides such non-measurable excitation sources, in addition some measurable excitation sources (e.g. artificial excitation from hammer impact or shaker, ground motion) may be available for modal parameter identification, which is referred to as Operational Modal Analysis with eXogenous inputs (OMAX). In this case, measured outputs are due to both known and unknown inputs, and may be corrupted by noise (i.e. sensor noise uncorrelated with known inputs). The known inputs are assumed to be a realization of a stochastic process. The use of input information in modal identification usually provides better modal estimates.

Many methods for OMA and OMAX are available in the literature [2]. In this work, we focus on subspace methods [1, 3], in particular on covariance-driven subspace methods. They are based on auto- and cross-covariance matrices of the measurements and can be based on output-only data or input/output data [4–9].

The purpose of this paper is the variance analysis of modal parameters obtained from input/output subspace methods. Using noisy measurement data, subspace algorithms provide parameter estimates that are afflicted with statistical uncertainty due to finite data, unknown inputs and sensor noise properties. In the field of

vibration analysis, explicit expressions for their variance estimation have been proposed for some subspace methods. A successful approach was developed in [10], where an estimated covariance on the measurements is propagated to the desired parameters based on a sensitivity analysis. The required sensitivities are derived analytically through the propagation of a first-order perturbation from the data to the identified parameters. This approach has the advantage of computational convenience: the sample covariance as the starting point is directly linked to the measurements and therefore easy to compute, and the sensitivities are computed using the system identification estimates. In [11], details of this scheme are developed for the covariance computation for output-only covariance-driven subspace identification. This approach offers advantages beyond technical correctness. The exploitation of the stabilization diagram, which is a basic tool for OMA(X), is a crucial step in modal analysis, but computationally costly for the covariance computation. A well-adapted fast and memory efficient implementation in an operational modal analysis context has been proposed in the covariance-driven output-only case [12]. As a result, the computational cost for the variance computation of an entire stabilization diagram analysis is reduced significantly. This was mandatory for realistic applications on large structures like in [13, 14].

While this approach has shown to be working both in theory and practice for the variance computation of modal parameters, it was only available for *output-only covariance-driven subspace identification* [11, 12]. Very recently, the generalization of this approach to a wider class of subspace algorithms in both output-only and input/output frameworks, and for covariance-driven and data-driven methods, has been presented in [15]. Developing the sensitivities of the modal parameters with respect to the correlations of inputs, outputs and cross-correlations between inputs and outputs for the selected subspace method, the uncertainty of the correlation estimates is propagated to the modal parameters in a rigorous way.

In this paper, we recall the new developments on variance estimation for an input/output covariance-driven subspace method. In the theoretical framework of the method, we focus in particular on assumptions about the stochastic nature of the available inputs, which are essential for the proper formulation of the input/output subspace method and the variance analysis. Besides the validation of the new variance computation scheme in Monte-Carlo simulations, we report an application on laboratory data from a vibration test of a composite plate, showing that the estimated variances of natural frequencies and damping ratios with the new method are in good agreement with the sample statistics.

This paper is organized as follows. In Section 2, the vibration model is recalled. Section 3 contains a brief description of the considered covariance-driven input/output subspace method. In Section 4, the variance estimation method is presented and validated in a Monte Carlo simulation and on data from a laboratory experiment in Section 5. Finally, conclusions are given in Section 6.

2 Vibration modeling

2.1 Mechanical and state space models

The vibration behavior of a linear time-invariant mechanical structure, which is observed at some sensor positions, can be described by the equations

$$\begin{cases} \mathcal{M}\ddot{q}(t) + \mathcal{C}\dot{q}(t) + \mathcal{K}q(t) = \mathcal{F}u(t) + \tilde{u}(t) \\ y(t) = L_a\ddot{q}(t) + L_v\dot{q}(t) + L_d q(t) + v(t) \end{cases} \tag{1}$$

where $\mathcal{M}$, $\mathcal{C}$ and $\mathcal{K} \in \mathbb{R}^{m\times m}$ are mass, stiffness and damping matrices, and m is the number of degrees of freedom. The vector $q(t) \in \mathbb{R}^{m\times 1}$ contains the displacements at the degrees of freedom generated by the known inputs $u(t) \in \mathbb{R}^{N_i\times 1}$ and unknown inputs $\tilde{u}(t) \in \mathbb{R}^{m\times 1}$, where N_i is the number of known inputs, and $\mathcal{F} \in \mathbb{R}^{m\times N_i}$ represents how the known inputs are applied to the system. The vector $y(t) \in \mathbb{R}^{N_o\times 1}$ contains the observed outputs, with N_o being the number of sensors. The matrices L_a, L_v and $L_d \in \mathbb{R}^{N_o\times m}$ represent how accelerations, velocities and displacements are obtained from the model degrees of freedom.

The vector $v(t) \in \mathbb{R}^{N_o \times 1}$ is the sensor noise. Both $\tilde{u}(t)$ and $v(t)$ are assumed to be white noise with finite fourth moments and uncorrelated with the known inputs.

Sampling Eq. (1) at rate τ and assuming zero-order hold for inputs $u(t)$ and $\tilde{u}(t)$ yield the discrete-time state space representation

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + \tilde{B}\tilde{u}_k \\ y_k = Cx_k + Du_k + \tilde{D}\tilde{u}_k + v_k, \end{cases} \tag{2}$$

where $x_k = [\dot{q}(k\tau)^T \;\; q(k\tau)^T]^T \in \mathbb{R}^{n \times 1}$ is the state vector, $n = 2m$ is the model order and

$$A_c = \begin{bmatrix} -\mathcal{M}^{-1}\mathcal{C} & -\mathcal{M}^{-1}\mathcal{K} \\ I_m & 0_{m,m} \end{bmatrix} \in \mathbb{R}^{n \times n},\; A = \exp(A_c\tau),\; B = (A - I)A_c^{-1} \begin{bmatrix} \mathcal{M}^{-1}\mathcal{F} \\ 0_{m,N_i} \end{bmatrix} \in \mathbb{R}^{n \times N_i}, \tag{3}$$

$$C = \begin{bmatrix} L_v - L_a\mathcal{M}^{-1}\mathcal{C} & L_d - L_a\mathcal{M}^{-1}\mathcal{K} \end{bmatrix} \in \mathbb{R}^{N_o \times n},\; D = L_a\mathcal{M}^{-1}\mathcal{F} \in \mathbb{R}^{N_o \times N_i} \tag{4}$$

are the state transition, input, output and feedthrough matrices, respectively. The matrices $\tilde{B}_c$ and $\tilde{D}_c$ are defined similarly with respect to the unknown inputs.

The modal parameters of system (1) are equivalently found in system (2) as follows. Let λ_i and ϕ_i be eigenvalues and eigenvectors of A, for $i = 1, \ldots, n$. Then the eigenvalues μ_i of system (1), modal frequencies f_i, damping ratios ξ_i and mode shapes φ_i are obtained by

$$\mu_i = \frac{\log(\lambda_i)}{\tau},\;\; f_i = \frac{|\mu_i|}{2\pi},\;\; \xi_i = \frac{-\mathrm{Re}(\mu_i)}{|\mu_i|},\;\; \varphi_i = C\phi_i. \tag{5}$$

2.2 Stochastic nature of known inputs

Until now, no assumption has been stated on the known input $\{u_k\}_{k=1,\ldots,N}$ for the case of input/output algorithms. Known inputs are used in what is called deterministic [1] or combined deterministic/stochastic [1, 16] system identification in the literature. In the designation of these methods, there is equivalence between "deterministic" and "the input sequence is known", and between "stochastic" and "there are unknown inputs and/or noise".

However, these designations should not be mistaken with the nature of the *known inputs*. The known input sequence itself may be a *deterministic* signal, or it may be a realization of a *stochastic* process. If it is *deterministic*, it can be defined exactly by a mathematical formula (e.g. a multi-sine excitation), and it could be predicted exactly also for $k > N$. If it is *stochastic*, the properties of the known values $\{u_k\}_{k=1,\ldots,N}$ are defined by the probability distribution of the underlying stochastic process. In this case it is not possible to predict values for $k > N$. The model of the known input – deterministic or stochastic – is important for the subsequent uncertainty analysis of the modal parameters.

For the considered input/output methods in this paper, we assume that known inputs are realizations of a stochastic process satisfying mild ergodicity and mixing properties. This assumption is in particular coherent with the ergodic-algebraic framework of [17] and cross-covariance estimates in [9] for covariance-driven input/output algorithms. In combination with standard white noise assumptions for unknown inputs and measurement noise, this assumption allows the consistent computation of auto-covariances of the inputs, the outputs and of cross-covariances between inputs and outputs, which is a basic step in subspace identification.

These auto- and cross-covariances contribute to the modal parameter estimates. Then, the corresponding uncertainties for these auto- and cross-covariances play together in a non-trivial way to obtain the modal parameter uncertainties, for which the expressions are shown in this paper. The assumption of stochastic inputs is coherent with the techniques for variance analysis that are used in this paper to obtain these expressions. The uncertainty of the modal estimates comes from the fact that only sequences of finite length N are known. Expressions for their variance based on an input/output data sequence of length N are hence also estimates, and the central limit theorem (CLT) ensures that these expressions are asymptotically correct

(i.e. for $N \to \infty$). This requires some stochastic properties of the underlying input and output sequences, for which the assumption of stochastic inputs (and of course also stochastic outputs) is particularly useful.

Note that assuming known inputs as stochastic is no contradiction for the particular case of a system with no unknown inputs and no sensor noise. In this case, the modal parameters can be identified exactly from the known inputs and outputs, and their variance assuming stochastic inputs is indeed zero, shown in [15].

3 Subspace identification

Subspace algorithms are based on projections of output, input or input/output data Hankel matrices with a "past" and "future" time horizon. These projections are designed in a way that the column space of the resulting matrix $\mathcal{H}$ is defined by the observability matrix $\Gamma_p = [C^T \ \ \dots \ \ (CA^{p-1})^T]^T$ of system (2). Matrix $\mathcal{H}$ has the factorization property $\mathcal{H} = \Gamma_p \mathcal{Z}$, where $\mathcal{Z}$ depends on the subspace algorithm [1]. This property leads to the estimation of Γ_p from measurement data and subsequently to the system matrices A and C, from where the desired modal parameters are identified. In this section, the input/output covariance-driven subspace algorithm from [8] is recalled. The three steps of modal identification with subspace algorithms are: 1/ Estimation of auto- and cross-covariance matrices related to "past" and "future" data matrix products, 2/ Computation of the subspace matrix $\mathcal{H}$, and finally 3/ Estimation of the modal parameters from $\mathcal{H}$.

3.1 First step: estimation of auto- and cross-covariances

Data Hankel matrices and the respective auto- and cross-covariances are the basic elements for the computation of $\mathcal{H}$, which are recalled in this section. A data Hankel matrix $\mathcal{A}_{i|j}$ is defined for the samples $a_l \in \mathbb{R}^{r\times 1}$ of a discrete signal a as

$$\mathcal{A}_{i|j} \overset{\text{def}}{=} \begin{bmatrix} a_i & a_{i+1} & \dots & a_{i+N-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_j & a_{j+1} & \dots & a_{j+N-1} \end{bmatrix} \in \mathbb{R}^{(j-i+1)r\times N}. \tag{6}$$

Let $N + p + q$ be the number of available samples for the known inputs u_k and the outputs y_k, where q and p are parameters that define a "past" and "future" time horizon. They are most often equal, and assumed to be large enough to satisfy the condition $\min\{(p-1)N_o, qN_o\} \geq n$. The "past" ($^-$) and "future" ($^+$) data Hankel matrices containing the known data are defined as

$$\mathcal{U}^- \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{U}_{0|q-1},\ \mathcal{U}^+ \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{U}_{q|q+p-1},\ \mathcal{Y}^- \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{Y}_{0|q-1},\ \mathcal{Y}^+ \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{Y}_{q|q+p-1}. \tag{7}$$

Consider similarly the unknown signals, i.e. the unknown input $\tilde{u}$ and the sensor noise v, and define the corresponding matrices $\widetilde{\mathcal{U}}^- \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\widetilde{\mathcal{U}}_{0|q-1}$, $\widetilde{\mathcal{U}}^+ \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\widetilde{\mathcal{U}}_{q|q+p-1}$, $\mathcal{V}^- \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{V}_{0|q-1}$ and $\mathcal{V}^+ \overset{\text{def}}{=} \frac{1}{\sqrt{N}}\mathcal{V}_{q|q+p-1}$, which are needed for deriving the equations of the model.

Known data (outputs and, if available, inputs) are used to estimate auto- and cross-covariances in order to compute the subspace matrix $\mathcal{H}$. Let a and b be placeholders for two of these signals (e.g. u and y, y and y or u and u). Then, their cross-covariance estimate at lag i is defined as

$$R_{[a,b]_i} = \frac{1}{N}\sum_{k=1}^{N} a_{k+i} b_k^T, \tag{8}$$

considering that a and b are ergodic and zero mean random processes. The term auto-covariance is used for the cross-covariance of the same signal, i.e. $a = b$ in (8). Note that due to the finite data length, these auto- and cross-covariance estimates are random variables with a covariance, which will be the starting point of the variance propagation to the modal parameter estimates in Section 4.

Products of "past" and "future" data Hankel matrices as defined in Eq. (7) are closely related to the auto- and cross-covariance estimates. For example, it holds

$$\mathcal{U}^+\mathcal{Y}^{-T} = \begin{bmatrix} R_{[u,y]_q} & R_{[u,y]_{q-1}} & \dots & R_{[u,y]_1} \\ R_{[u,y]_{q+1}} & R_{[u,y]_q} & \dots & R_{[u,y]_2} \\ \vdots & \vdots & \vdots & \vdots \\ R_{[u,y]_{q+p-1}} & R_{[u,y]_{q+p-2}} & \dots & R_{[u,y]_p} \end{bmatrix}, \; \mathcal{U}^+\mathcal{Y}^{+T} = \begin{bmatrix} R_{[u,y]_0} & R_{[u,y]_{-1}} & \dots & R_{[u,y]_{1-p}} \\ R_{[u,y]_1} & R_{[u,y]_0} & \dots & R_{[u,y]_{2-p}} \\ \vdots & \vdots & \ddots & \vdots \\ R_{[u,y]_{p-1}} & R_{[u,y]_{p-2}} & \dots & R_{[u,y]_0} \end{bmatrix} \tag{9}$$

The products $\mathcal{Y}^+\mathcal{Y}^{-T}$ and $\mathcal{U}^+\mathcal{U}^{+T}$ are computed analogously. These products based on auto- and cross-covariances are elementary blocks for the computation of $\mathcal{H}$ in the following section.

Remark 1 (Cross-covariance with unknown inputs and sensor noise) *Products between data Hankel matrices containing outputs or known inputs and noise Hankel matrices are established analogously as above. Some of these cross-covariances tend to zero due to the assumption that unknown inputs and sensor noise are white and uncorrelated with the known inputs. In particular it holds* $\widetilde{\mathcal{U}}^+\mathcal{U}^{+T} = 0$ *and* $\mathcal{V}^+\mathcal{U}^{+T} = 0$ *for an infinite amount of data. Under these assumptions, past outputs are uncorrelated with future unknown inputs and with future sensor noise, and it holds similarly* $\widetilde{\mathcal{U}}^+\mathcal{Y}^{-T} = 0$ *and* $\mathcal{V}^+\mathcal{Y}^{-T} = 0$*. These properties are required for the development of the subspace method in the next section.*

3.2 Second step: computation of subspace matrix $\mathcal{H}$

With the definitions in the previous section, the observation equation in (2) can be extended to [1]

$$\mathcal{Y}^+ = \Gamma_p\mathcal{X}^+ + H_p\mathcal{U}^+ + \tilde{H}_p\widetilde{\mathcal{U}}^+ + \mathcal{V}^+ \tag{10}$$

The matrices H_p and $\tilde{H}_p$ are defined in [1].

Orthogonal projections of the data Hankel matrices lead to a matrix $\mathcal{H}$ whose column space is defined by the observability matrix Γ_p. An orthogonal projection M/P of the row space of M onto the row space of P yields $M/P = MP^T\left(PP^T\right)^\dagger P$, and the projection on the orthogonal complement $P^\perp$ yields $M/P^\perp = M(I - P^T\left(PP^T\right)^\dagger P)$, where † denotes the Moore-Penrose pseudoinverse.

The considered input/output covariance-driven subspace algorithm is defined by the projection [8]

$$\begin{aligned} \mathcal{H} &\stackrel{\text{def}}{=} (\mathcal{Y}^+/\mathcal{U}^{+\perp})\mathcal{Y}^{-T} \\ &= \mathcal{Y}^+\mathcal{Y}^{-T} - \mathcal{Y}^+\mathcal{U}^{+T}\left(\mathcal{U}^+\mathcal{U}^{+T}\right)^\dagger\mathcal{U}^+\mathcal{Y}^{-T}. \end{aligned} \tag{11}$$

The factorization property of $\mathcal{H}$ to obtain the observability matrix can be shown as follows. Plugging the extended state-space equation (10) for $\mathcal{Y}^+$ into (11), it follows

$$\mathcal{H} = \left((\Gamma_p\mathcal{X}^+ + H_p\mathcal{U}^+ + \tilde{H}_p\widetilde{\mathcal{U}}^+ + \mathcal{V}^+)/\mathcal{U}^{+\perp}\right)\mathcal{Y}^{-T}. \tag{12}$$

By definition, it holds $\mathcal{U}^+/\mathcal{U}^{+\perp} = 0$, and since known inputs are assumed uncorrelated with unknown inputs and with sensor noise (Remark 1), it holds

$$\mathcal{H} = \left(\Gamma_p\mathcal{X}^+/\mathcal{U}^{+\perp} + \tilde{H}_p\widetilde{\mathcal{U}}^+ + \mathcal{V}^+\right)\mathcal{Y}^{-T}. \tag{13}$$

Finally, using the assumptions that past outputs are uncorrelated with future unknown inputs and with future sensor noise (Remark 1), it follows

$$\mathcal{H} = \Gamma_p\left(\mathcal{X}^+/\mathcal{U}^{+\perp}\right)\mathcal{Y}^{-T} = \Gamma_p\mathcal{Z}. \tag{14}$$

3.3 Third step: estimation of the modal parameters from $\mathcal{H}$

For the estimation of the modal parameters, the observability matrix Γ_p is obtained from $\mathcal{H}$, which contains the system matrices A and C. Thanks to the factorization property (14), $\mathcal{H}$ and Γ_p share the same column space. Thus, Γ_p can be obtained by the singular value decomposition (SVD) [18]

$$\mathcal{H} = USV^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{15}$$

as $\Gamma_p = U_1 S_1^{\frac{1}{2}}$, where the SVD is truncated at the desired model order n. The output matrix C is then obtained by direct extraction of the first block row of Γ_p as $C = \begin{bmatrix} I_{N_o} & 0_{N_o,(p-1)N_o} \end{bmatrix} \Gamma_p$. The state transition matrix A is obtained from the shift invariance property of Γ_p as the least squares solution $A = \Gamma_{up}^{\dagger} \Gamma_{dw}$ [4], where Γ_{dw} and Γ_{up} are obtained from Γ_p by removing the first and last block row, respectively, as

$$\Gamma_{dw} = \begin{bmatrix} CA \\ \vdots \\ CA^{p-1} \end{bmatrix}, \quad \Gamma_{up} = \begin{bmatrix} C \\ \vdots \\ CA^{p-2} \end{bmatrix}. \tag{16}$$

Finally, the modal parameters are obtained from matrices A and C as stated in Section 2.1.

4 Variance estimation

In [11, 12], a method for the variance estimation of modal parameters was developed for the output-only covariance-driven subspace algorithm. In [15], this method was extended for an input/output covariance-driven subspace algorithm, amongst other data-driven and input/output algorithms, which is presented in the following.

The computation of the modal parameter covariance results from the propagation of sample covariances on auto- or cross-covariance estimates (involving measured outputs and known inputs) through all steps of the modal identification algorithm. These sample covariances reflect in particular the unknown inputs due to non-measurable excitation sources and the sensor noise, and they contribute in a non-trivial way to the covariance of the modal parameter estimates. The propagation to the modal parameter estimates is based on the delta method [19], where the analytical sensitivity matrices are obtained from perturbation theory [10, 11].

Let ΔX be a first-order perturbation of a matrix-valued variable X. Then, for a function $Y = f(X)$ it holds $\mathrm{vec}(\Delta Y) = \mathcal{J}_{Y,X} \, \mathrm{vec}(\Delta X)$, where $\mathcal{J}_{Y,X} = \partial \mathrm{vec}(f(X)) / \partial \mathrm{vec}(X)$, where $\mathrm{vec}(\cdot)$ denotes the column stacking vectorization operator. Subsequently, covariance expressions for the estimates satisfy

$$\mathrm{cov}(\mathrm{vec}(\widehat{Y})) \approx \widehat{\mathcal{J}}_{Y,X} \mathrm{cov}(\mathrm{vec}(\widehat{X})) \widehat{\mathcal{J}}_{Y,X}^T. \tag{17}$$

For simplicity of notation we dismiss the notation $\widehat{\ }$ for an estimate in the following.

4.1 Sample covariance estimation

The starting point of the variance propagation to the modal parameters is the covariance of the auto- and cross-covariance matrices that are involved in the subspace algorithm. In Eq. (11), this concerns the covariance of and between the matrices

$$\mathcal{R}_1 = \mathcal{Y}^+ \mathcal{Y}^{-^T}, \quad \mathcal{R}_2 = \mathcal{Y}^+ \mathcal{U}^{+^T}, \quad \mathcal{R}_3 = \mathcal{U}^+ \mathcal{U}^{+^T}, \quad \mathcal{R}_4 = \mathcal{Y}^- \mathcal{U}^{+^T}, \tag{18}$$

i.e. the computation of $\mathrm{cov}(\mathrm{vec}(\mathcal{R}_i), \mathrm{vec}(\mathcal{R}_j))$ between any of the matrices $\mathcal{R}_i$ and $\mathcal{R}_j$ (including $i = j$) is required. This computation follows the lines of the output-only covariance-driven algorithm, i.e. for $\mathrm{cov}(\mathrm{vec}(\mathcal{R}_1))$, as described in detail in [11, 12], and is generalized in [15] as follows.

Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ be placeholders for any of the data Hankel matrices $\mathcal{U}^-$, $\mathcal{U}^+$, $\mathcal{Y}^-$ or $\mathcal{Y}^+$, such that $\mathcal{R}_i = \mathcal{A}\mathcal{B}^T$ and $\mathcal{R}_j = \mathcal{C}\mathcal{D}^T$ are the respective auto- or cross-covariance matrices. Divide the data matrices into n_b blocks and normalize them with respect to their length, such that

$$\begin{aligned} \sqrt{N}\,\mathcal{A} &= \sqrt{N_b}\begin{bmatrix}\mathcal{A}^1 & \mathcal{A}^2 & \dots & \mathcal{A}^{n_b}\end{bmatrix}, & \sqrt{N}\,\mathcal{B} &= \sqrt{N_b}\begin{bmatrix}\mathcal{B}^1 & \mathcal{B}^2 & \dots & \mathcal{B}^{n_b}\end{bmatrix}, \\ \sqrt{N}\,\mathcal{C} &= \sqrt{N_b}\begin{bmatrix}\mathcal{C}^1 & \mathcal{C}^2 & \dots & \mathcal{C}^{n_b}\end{bmatrix}, & \sqrt{N}\,\mathcal{D} &= \sqrt{N_b}\begin{bmatrix}\mathcal{D}^1 & \mathcal{D}^2 & \dots & \mathcal{D}^{n_b}\end{bmatrix}, \end{aligned} \tag{19}$$

where each block $\mathcal{A}^k$, $\mathcal{B}^k$, $\mathcal{C}^k$ and $\mathcal{D}^k$ may have the same length N_b, with $n_b \cdot N_b = N$ for simplicity. Each block may be long enough to assume statistical independence between the blocks. On each of these blocks, the respective auto- or cross-covariance estimate can be computed as $\mathcal{R}_i^k = \mathcal{A}^k\mathcal{B}^{kT}$ and $\mathcal{R}_j^k = \mathcal{C}^k\mathcal{D}^{kT}$, which can be assumed to be i.i.d., yielding

$$\mathcal{R}_i = \frac{1}{n_b}\sum_{k=1}^{n_b}\mathcal{R}_i^k, \quad \mathcal{R}_j = \frac{1}{n_b}\sum_{k=1}^{n_b}\mathcal{R}_j^k\,. \tag{20}$$

It follows $\mathrm{cov}\left(\mathrm{vec}(\mathcal{R}_*)\right) = \frac{1}{n_b}\mathrm{cov}\left(\mathrm{vec}(\mathcal{R}_*^k)\right)$, and the covariance between the auto- or cross-covariance matrices can be computed from the usual sample covariance as

$$\mathrm{cov}(\mathrm{vec}(\mathcal{R}_i), \mathrm{vec}(\mathcal{R}_j)) = \frac{1}{n_b(n_b-1)}\sum_{k=1}^{n_b}\Big(\mathrm{vec}(\mathcal{R}_i^k) - \mathrm{vec}(\mathcal{R}_i)\Big)\Big(\mathrm{vec}(\mathcal{R}_j^k) - \mathrm{vec}(\mathcal{R}_j)\Big)^T. \tag{21}$$

4.2 Sensitivity and covariance of $\mathcal{H}$

In this section, the sensitivity of $\mathcal{H}$ is developed with respect to the underlying auto- and cross-covariance matrices in their computation. It holds, from Eq. (11), $\mathcal{H} = \mathcal{R}_1 - \mathcal{R}_2\mathcal{R}_3^\dagger\mathcal{R}_4^T$, and thus [15]

$$\Delta\mathcal{H} = \Delta\mathcal{R}_1 - \Delta\mathcal{R}_2\mathcal{R}_3^\dagger\mathcal{R}_4^T + \mathcal{R}_2\mathcal{R}_3^\dagger\Delta\mathcal{R}_3\mathcal{R}_3^\dagger\mathcal{R}_4^T - \mathcal{R}_2\mathcal{R}_3^\dagger\Delta\mathcal{R}_4^T, \tag{22}$$

$$\mathrm{vec}(\Delta\mathcal{H}) = \mathcal{J}_{\mathcal{H},\mathcal{R}}\begin{bmatrix}\mathrm{vec}(\Delta\mathcal{R}_1)\\ \mathrm{vec}(\Delta\mathcal{R}_2)\\ \mathrm{vec}(\Delta\mathcal{R}_3)\\ \mathrm{vec}(\Delta\mathcal{R}_4)\end{bmatrix}, \tag{23}$$

where $\mathcal{J}_{\mathcal{H},\mathcal{R}} = \begin{bmatrix} I_{pN_oqN_o} & -\mathcal{R}_4\mathcal{R}_3^\dagger \otimes I_{pN_o} & \mathcal{R}_4\mathcal{R}_3^\dagger \otimes \mathcal{R}_2\mathcal{R}_3^\dagger & -(I_{qN_o} \otimes \mathcal{R}_2\mathcal{R}_3^\dagger)\mathcal{P}_{qN_o,pN_i}\end{bmatrix}$, $\otimes$ denotes the Kronecker product and $\mathcal{P}_{qN_o,pN_i}$ is a permutation matrix with the property $\mathrm{vec}(\mathcal{R}_4^T) = \mathcal{P}_{qN_o,pN_i}\mathrm{vec}(\mathcal{R}_4)$. Then, it follows for the covariance

$$\mathrm{cov}(\mathrm{vec}(\mathcal{H})) = \mathcal{J}_{\mathcal{H},\mathcal{R}}\,\mathrm{cov}\left(\begin{bmatrix}\mathrm{vec}(\mathcal{R}_1)\\ \mathrm{vec}(\mathcal{R}_2)\\ \mathrm{vec}(\mathcal{R}_3)\\ \mathrm{vec}(\mathcal{R}_4)\end{bmatrix}\right)\mathcal{J}_{\mathcal{H},\mathcal{R}}^T, \tag{24}$$

where the blocks for the covariance matrix on the right side are computed in Eq. (21).

4.3 Covariance of modal parameters

From $\mathcal{H}$, the modal parameters are obtained as described in Section 3.3. The respective propagation of the covariance of $\mathrm{vec}(\mathcal{H})$ to the modal parameters has been described in detail in [11, 12].

5 Application

5.1 Simulation study

In [15], a simulation study was made on a damped mass-spring chain model of six degrees of freedom (see Figure 1) to validate the variance computation for the output-only and input/output subspace methods in an academic example that is summarized here. The mass of all elements is $m_i = 1/20$, stiffnesses are $k_1 = k_3 = k_5 = 100, k_2 = k_4 = k_6 = 200$ and damping is defined such that all modes have a damping ratio of $\xi_i = 0.020$.

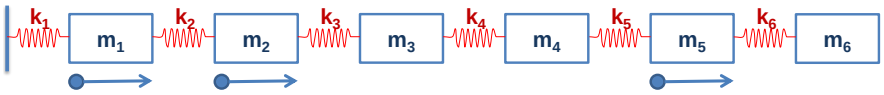

Figure 1: Mass-spring chain (with modal damping), three sensors, for simulations.

The covariance-driven output-only algorithm is compared with the covariance-driven input/output algorithm considered in this paper, using the same output realizations. The input/output algorithm uses the simulated Gaussian white noise excitation force at element 1 as the known input, besides further unknown inputs at all elements with 10% of the standard deviation of the known input. 10,000 acceleration output time series each of length $N = 10{,}000$ are simulated with time step $\tau = 0.02$ s at elements 1, 2 and 5, where Gaussian white sensor noise with 5% of the standard deviation of each output is added.

Each time the modal parameters and their variances are estimated with the covariance-driven output-only and input/output subspace algorithms. The means of the identified modal parameters with both subspace algorithms are shown together with the theoretical modal parameters from the mass-spring chain model in Table 1. The identified values are more precise with the input/output algorithm than with the output-only algorithm. In particular, the first mode in the output-only method deviates from the theoretical value, which would have required longer datasets for more precision.

In each simulation, the standard deviations of the modal parameters are computed for the output-only covariance-driven subspace algorithm with the method described in [12], and for the input/output covariance-driven algorithm with the method described in Section 4. They are obtained from the square root of the respective modal parameter variances. The mean $\overline{\sigma}_{\mathrm{est}}$ of these 10,000 estimated standard deviations from each of the Monte Carlo simulations is shown for each modal parameter and algorithm in Table 2. These values are compared to the sample standard deviations σ_{MC} of the 10,000 estimated modal parameters in Table 2. An excellent agreement between the estimated standard deviations and the sample statistics from the Monte Carlo simulations can be found for both methods, validating the variance computation from a theoretical point of view. Note that the standard deviations are significantly lower with the input/output method, compared to the output-only method, roughly by factor 4 in this application.

Table 2 also shows the one sigma sample standard deviation of the estimated standard deviations during the Monte Carlo simulations. It is mostly smaller than 15% of their mean value (except for the badly estimated

	Frequency (Hz)						Damping ratio (‰)					
Algorithm	1	2	3	4	5	6	1	2	3	4	5	6
Actual model	1.930	5.618	8.682	14.49	15.85	17.01	20.0	20.0	20.0	20.0	20.0	20.0
OO	1.937	5.618	8.683	14.49	15.85	17.01	29.4	20.1	20.1	20.0	20.0	20.0
IO	1.930	5.618	8.682	14.49	15.85	17.01	20.2	20.0	20.0	20.0	20.0	20.0

Table 1: Means for frequencies and damping ratios from modal identification in 10,000 Monte Carlo simulations. OO: output-only, IO: input/output algorithm.

Algorithm		Standard deviation of frequency (Hz $\cdot 10^{-2}$) 1	2	3	4	5	6
OO	$\overline{\sigma}_{\mathrm{est}}$	2.02 ± 1.80	1.10 ± 0.15	1.55 ± 0.21	2.60 ± 0.36	2.64 ± 0.38	3.51 ± 0.48
	σ_{MC}	1.99	1.14	1.59	2.63	2.65	3.48
IO	$\overline{\sigma}_{\mathrm{est}}$	0.42 ± 0.06	0.24 ± 0.03	0.29 ± 0.03	0.66 ± 0.08	0.49 ± 0.05	0.87 ± 0.10
	σ_{MC}	0.41	0.25	0.29	0.68	0.50	0.87

Algorithm		Standard deviation of damping ratio (‰) 1	2	3	4	5	6
OO	$\overline{\sigma}_{\mathrm{est}}$	16.46 ± 26.64	1.98 ± 0.28	1.80 ± 0.25	1.70 ± 0.22	1.67 ± 0.23	2.09 ± 0.28
	σ_{MC}	17.12	2.07	1.85	1.71	1.71	2.09
IO	$\overline{\sigma}_{\mathrm{est}}$	2.13 ± 0.34	0.43 ± 0.05	0.33 ± 0.04	0.45 ± 0.05	0.31 ± 0.04	0.52 ± 0.06
	σ_{MC}	2.07	0.44	0.34	0.45	0.31	0.53

Table 2: Mean of estimated standard deviations ($\overline{\sigma}_{\mathrm{est}}$) and sample standard deviations (σ_{MC}) for frequencies and damping ratios of the six modes in 10,000 Monte Carlo simulations. OO: output-only, IO: input/output algorithm.

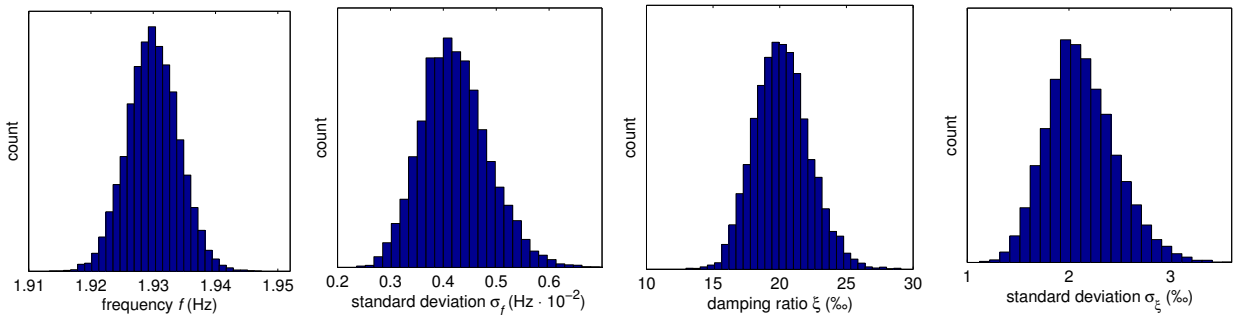

Figure 2: Histograms for first mode identified with the input/output method from Monte Carlo simulations: estimated frequencies, estimated standard deviations on frequencies, estimated damping ratios and estimated standard deviations on damping ratios.

modes), both for frequencies and damping ratios. This means that the estimation error of the standard deviation on a single dataset is in most cases less than 15% with the proposed methods, which is very reasonable for a second-order statistics. As examples, the histograms of the identified frequencies and damping ratios of the first mode, and the histograms of their estimated standard deviations with the methods from this paper are shown in Figure 2 for the Monte Carlo simulations with the input/output method.

5.2 Validation on experimental data

In a second study, experimental input/output data from a lab experiment on a composite plate was considered [20]. The considered vibration test was performed with one input and the outputs were measured at 14 sensors. Input/output time series of length 1,600,000 at a sampling rate of 4096 Hz were recorded. For the experimental validation of the output-only and input/output algorithms, they are separated into 32 smaller datasets of length 50,000.

Stabilization diagrams from both identification methods using one of the datasets is shown in Figure 3, where seven modes could be clearly identified. The means of the identified frequencies and damping ratios are shown in Table 3, where a stable range between model orders 25 and 50 has been selected. It can be observed that some frequencies show a significant difference between both algorithms, which may suggest

that the inputs were actually not exactly white noise. The damping ratios identified with the input/output algorithm are lower than with the output-only algorithm for most of the modes.

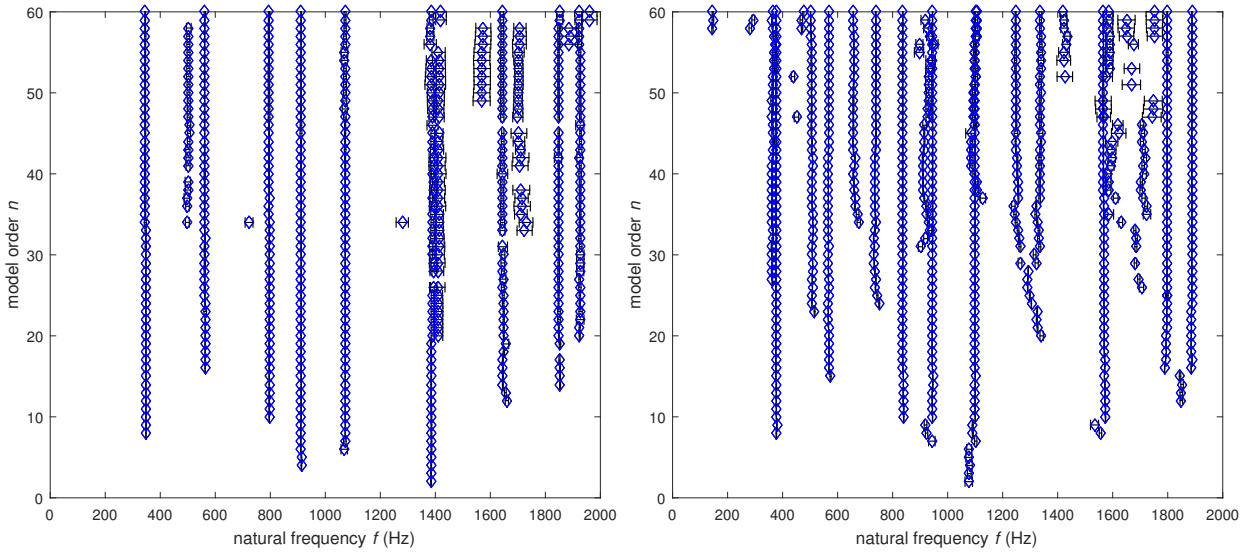

Figure 3: Stabilization diagram of one dataset with output-only (left) and input/output subspace identification (right), including the $\pm 1\sigma$ uncertainty bound.

	mode	1	2	3	4	5	6	7
f (Hz)	OO	344	561	794	912	1073	1849	1925
	IO	376	567	837	946	1099	1798	1888
ξ (‰)	OO	27.4	11.4	16.3	7.4	9.7	12.4	23.7
	IO	9.4	9.4	6.0	9.8	8.0	19.4	13.5

Table 3: Means of frequencies and damping ratios of the seven modes from 32 datasets of the composite plate. OO: output-only, IO: input/output algorithm.

		Standard deviation of frequency (Hz)						
Algorithm		1	2	3	4	5	6	7
OO	$\overline{\sigma}_{\mathrm{est}}$	1.72 ± 1.23	0.54 ± 0.11	1.59 ± 1.04	0.41 ± 0.15	0.56 ± 0.14	1.54 ± 0.41	4.47 ± 3.76
	σ_{MC}	1.61	0.43	1.49	0.39	0.54	1.25	3.37
IO	$\overline{\sigma}_{\mathrm{est}}$	0.33 ± 0.13	0.07 ± 0.02	0.08 ± 0.02	0.18 ± 0.09	0.13 ± 0.05	0.55 ± 0.67	0.84 ± 3.50
	σ_{MC}	0.26	0.06	0.11	0.17	0.13	0.55	0.75

		Standard deviation of damping ratio (‰)						
Algorithm		1	2	3	4	5	6	7
OO	$\overline{\sigma}_{\mathrm{est}}$	5.3 ± 4.2	1.0 ± 0.2	2.1 ± 2.1	0.5 ± 0.1	0.6 ± 0.3	0.9 ± 0.2	5.4 ± 7.9
	σ_{MC}	4.0	0.7	1.7	0.5	0.5	0.7	3.1
IO	$\overline{\sigma}_{\mathrm{est}}$	1.2 ± 0.3	0.1 ± 0.04	0.1 ± 0.03	0.2 ± 0.1	0.1 ± 0.04	0.2 ± 0.2	0.5 ± 2.0
	σ_{MC}	0.9	0.1	0.1	0.2	0.1	0.2	0.2

Table 4: Mean of estimated standard deviations ($\overline{\sigma}_{\mathrm{est}}$) and sample standard deviations (σ_{MC}) for frequencies and damping ratios of the seven modes from 32 datasets of the composite plate. OO: output-only, IO: input/output algorithm.

A comparison between the sample standard deviations from the 32 identification results and the means of the estimated standard deviations with the method described in this paper is shown in Table 4. There is a good agreement between both statistics for each mode and for each methods, which validates the proposed uncertainty quantification method for the input/output covariance-driven subspace algorithm on experimental data. For both frequencies and damping ratios, the standard deviation of the estimates is significantly lower with the input/output algorithm than with the output-only algorithm. Thus, taking into account the input information indeed leads to a higher precision of the estimates in terms of the standard deviation.

6 Conclusions

In this paper, we have recalled a recently developed method for the variance estimation of modal parameter estimates from an input/output covariance-driven subspace identification method, and we have validated it on experimental data from a laboratory vibration test. The results show a good agreement between the sample statistics and the variance estimates from the proposed method, which validates the method not only in Monte Carlo simulations, but also on experimental data. Note that often only one dataset is available for modal analysis, which makes a Monte Carlo approach unfeasible in practice. The presented method extends variance estimation for subspace identification to input/output data, which has been shown to give more precise identification results and lower uncertainty bounds than the output-only approach. In this work, inputs have been considered as the realization of a stochastic process. Future work includes the analysis of both stochastic and deterministic inputs in the same framework.

Acknowledgments

We thank M. Luczak and B. Peeters from LMS International for providing the data of the composite panel [20].

References

[1] P. Van Overschee, B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer, Dordrecht, The Netherlands, 1996.

[2] E. Reynders. *System identification methods for (operational) modal analysis: review and comparison*. Archives of Computational Methods in Engineering, 19(1):51–124, 2012.

[3] M. Döhler, L. Mevel. *Fast multi-order computation of system matrices in subspace-based system identification*. Control Engineering Practice, 20(9):882–894, 2012.

[4] B. Peeters, G. De Roeck. *Reference-based stochastic subspace identification for output-only modal analysis*. Mechanical Systems and Signal Processing, 13(6):855–878, 1999.

[5] A. Guyader, L. Mevel. *Covariance-driven subspace methods: input/output vs. output-only*. In Proceedings of the 21st International Modal Analysis Conference, Kissimmee, FL, USA, 2003.

[6] L. Mevel, A. Benveniste, M. Basseville, M. Goursat, B. Peeters, H. Van der Auweraer, A. Vecchio. *Input/output versus output-only data processing for structural identification - application to in-flight data analysis*. Journal of Sound and Vibration, 295(3):531–552, 2006.

[7] M. Döhler, L. Mevel. *Modular subspace-based system identification from multi-setup measurements*. IEEE Transactions on Automatic Control, 57(11):2951–2956, 2012.

[8] A. Esna Ashari, L. Mevel. *Input-output subspace-based fault detection*. In Proc. 8th IFAC Symposium on fault detection, diagnosis and safety of technical processes (SAFEPROCESS), pages 204–209, Mexico City, Mexico, 2012.

[9] E. Gandino, L. Garibaldi, S. Marchesiello. *Covariance-driven subspace identification: A complete input–output approach*. Journal of Sound and Vibration, 332(26):7000–7017, 2013.

[10] R. Pintelon, P. Guillaume, J. Schoukens. *Uncertainty calculation in (operational) modal analysis*. Mechanical Systems and Signal Processing, 21(6):2359–2373, 2007.

[11] E. Reynders, R. Pintelon, G. De Roeck. *Uncertainty bounds on modal parameters obtained from stochastic subspace identification*. Mechanical Systems and Signal Processing, 22(4):948–969, 2008.

[12] M. Döhler, L. Mevel. *Efficient multi-order uncertainty computation for stochastic subspace identification*. Mechanical Systems and Signal Processing, 38(2):346–366, 2013.

[13] M. Döhler, X.-B. Lam, L. Mevel. *Uncertainty quantification for modal parameters from stochastic subspace identification on multi-setup measurements*. Mechanical Systems and Signal Processing, 36 (2):562–581, 2013.

[14] M. Döhler, F. Hille, L. Mevel, W. Rücker. *Structural health monitoring with statistical methods during progressive damage test of S101 Bridge*. Engineering Structures, 69:183–193, 2014.

[15] P. Mellinger, M. Döhler, L. Mevel. *Variance estimation of modal parameters from output-only and input/output subspace-based system identification*. Journal of Sound and Vibration, 379:1–27, 2016.

[16] E. Reynders, G. De Roeck. *Reference-based combined deterministic-stochastic subspace identification for experimental and operational modal analysis*. Mechanical Systems and Signal Processing, 22(3): 617–637, 2008.

[17] M. Verhaegen. *Identification of the deterministic part of MIMO state space models given in innovations form from input-output data*. Automatica, 30(1):61–74, 1994. Special issue on statistical signal processing and control.

[18] G.H. Golub, C.F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, USA, 3rd edition, 1996.

[19] G. Casella, R.L. Berger. *Statistical inference*. Duxbury Press, Pacific Grove, CA, USA, 2002.

[20] M. Luczak, B. Peeters, M. Döhler, L. Mevel, W. Ostachowicz, P. Malinowski, T. Wandowski, K. Branner. *Damage detection in wind turbine blade panels using three different SHM techniques*. In Proc. 28th International Modal Analysis Conference, Jacksonville, FL, USA, 2010.