

# Uncertainty quantification in data-driven stochastic subspace identification

Edwin P.B. Reynders\*

*University of Leuven (KU Leuven), Department of Civil Engineering, Kasteelpark Arenberg 40, B-3001 Leuven, Belgium*

---

## Abstract

A crucial aspect in system identification is the assessment of the accuracy of the identified system matrices. Stochastic Subspace Identification (SSI) is a widely used approach for the identification of linear systems from output-only data because it combines a high computational robustness and efficiency with a high estimation accuracy. Practical approaches for estimating the (co)variance of system matrices that are identified using SSI exist for the case where the system matrices are obtained from the shift-invariant structure of the extended observability matrix. However, in data-driven SSI, the system matrices are often obtained in a different way, using identified state sequences. This case is treated in the present work, for three common types of weighting. First, it is shown that the estimated system matrices depend entirely on sample output correlation estimates, the covariance of which can be straightforwardly estimated. Subsequently, a linear sensitivity analysis of the data-driven SSI algorithm is performed, such that the covariance of the identified system matrices can be also computed. A memory efficient implementation is obtained by computing the related Jacobian only implicitly. An extensive numerical validation, covering a range of parameter choices, demonstrates the accuracy of the estimated variance of the identified system description. Finally, the practical use of the method in the context of operational modal analysis is demonstrated in an experimental case study.

*Keywords:* stochastic subspace identification, variance estimation, operational modal analysis

---

## 1. Introduction

Stochastic subspace identification (SSI) is an approach for estimating a linear time-invariant system model from output-only data that relies on the orthogonal projection of a matrix containing the most recent data onto the subspace spanned by the rows of a matrix containing the least recent data [1]. The approach is very widely used for the identification of mechanical systems from response-only data, in particular in the context of operational modal analysis, where the aim is to extract natural frequencies, damping ratios and mode shapes [2, 3, 4]. The main reason for this popularity is that SSI combines a high computational robustness and efficiency with a high estimation accuracy [4]. Several variants of SSI exist, but it has been demonstrated that, as far as the identification of the extended observability matrix is concerned, they boil down to the same basic algorithm with a different choice of weights [5]. This general algorithm is often termed data-driven SSI or SSI-data in the operational modal analysis literature, while covariance-driven SSI or SSI-cov designates a particular choice of weights that reduces the orthogonal projection step in the algorithm to a single data correlation [6]; the SSI-cov weighting is known as the principal component (PC) weighting in the automatic control literature [5]. Once the extended observability matrix has been identified, the system matrices can be identified as well. In SSI-cov, this is usually achieved by exploiting the shift-invariant structure of the extended observability matrix, while in SSI-data, identified Kalman filter state sequences are often employed [6].

In order to use the identified system matrices or related modal characteristics with confidence, it is important that their accuracy can be assessed. In-depth theoretical analyses of the asymptotic statistical properties of SSI are available from the automatic control literature, generally in the setting of Hilbert spaces of random variables. Important theoretical results include proofs that the identified system matrices are asymptotically normally distributed [7, 8, 9], and that they have asymptotically minimum variance (in other words, they are statistically efficient) for a specific type of weighting,

---

\*Corresponding author. Email: edwin.reynders@kuleuven.be. Tel.: +3216321677.

*Postprint submitted to Mechanical Systems & Signal Processing*

*Published version:* E. Reynders. Uncertainty quantification in data-driven stochastic subspace identification. *Mechanical Systems & Signal Processing*, 151(107338):1-18, 2021. <https://doi.org/10.1016/j.ymssp.2020.107338>

the canonical variate analysis (CVA) weighting [10]. Analytical expressions that relate the asymptotic covariance of the identified system matrices to the exact system matrices and to covariances of the unknown inputs and state estimation errors, have also been derived [8, 9, 11, 12], but explicit practical expressions for the actual covariance estimation from experimental data are not available from these works. Such expressions are available from the experimental vibration analysis literature, at least for the special case where the system matrices are identified from the shift-invariant structure of the identified observability matrix, and for unity left projection weighting [13, 14, 15].

The approach that has been taken to arrive at these practical expressions, is to perform a detailed first-order sensitivity analysis of the entire algorithm, resulting in the derivatives (or more precisely, the Jacobian matrix containing the derivatives) of the identified system matrices to the sample output correlations where the algorithm starts from. When combined with an estimate of the covariance of the sample output correlations, these derivatives then result in the covariance of the identified system matrices. The first such analysis was performed for the SSI-cov algorithm [13]; it was later extended to SSI-data for the case where the system matrices are estimated from the shift-invariant structure of the observability matrix, and unweighted principal component (UPC) weighting [15]. A detailed experimental validation of the covariance estimation of the system matrices identified with SSI-cov has been presented in [16].

In the present work, a practical approach is developed for estimating the covariance of the system matrices that are identified with SSI-data using identified Kalman filter state sequences. The analysis comprises three common types of weighting: UPC, PC and CVA. First, it is shown that the estimated system matrices depend entirely on sample output correlation estimates, the covariance of which can be straightforwardly estimated. Subsequently, a linear sensitivity analysis of the data-driven SSI algorithm is performed, such that the covariance of the identified system matrices can be computed. A memory efficient implementation is automatically obtained by computing the related Jacobian only implicitly. In this way, the memory issues that can arise when the Jacobian is explicitly computed [14] are obviated in a natural way.

The remainder of this paper is structured as follows. In section 2, the system models that are identified with stochastic subspace identification are discussed. The identification itself is presented in section 3. The relation between the identified system matrices and the sample output correlations is investigated in section 4. The estimation of the covariances of the identified system matrices is presented in section 5. The computation of modal characteristics and their covariances is outlined in section 6. Numerical verifications and an application involving experimental data are presented in sections 7 and 8, respectively.

## 2. State-space models for linear systems with unknown inputs and measurement errors

The aim of the present section is to introduce the forward innovation model, which is the system model that is identified in the general SSI-data setting, as well as the related notation. The more conventional stochastic state-space model is introduced first, followed by the non-stationary Kalman filter which leads to the forward innovation model.

### 2.1. Conventional stochastic state-space model

The input-output relationship of a linear time-invariant system can, after discretization in time, be described with the following state-space model:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \quad (1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{e}_k^m, \quad (2)$$

where  $\mathbf{x}_k \in \mathbb{R}^n$  is the state vector of the system with  $n$  the system order,  $\mathbf{y}_k \in \mathbb{R}^{n_y}$  is the output vector with  $n_y$  the number of outputs,  $\mathbf{u}_k \in \mathbb{R}^{n_u}$  is the input vector with  $n_u$  the number of inputs, and  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{D}$  are the system matrices. In operational modal analysis, only the outputs  $\mathbf{y}_k$  are observed, with measurement error  $\mathbf{e}_k^m$ , and usually little information on the unobserved loading  $\mathbf{u}_k$  and the unknown measurement error  $\mathbf{e}_k^m$  is available, other than that they take finite values that are centered around zero. For this reason  $\mathbf{u}_k$  and  $\mathbf{e}_k^m$  are assumed to be white noise random processes. This is a good assumption as long as the ambient excitation spectrum is not dominated by specific frequency components, which is for instance the case when strong harmonics are present in the excitation [4]. The white noise model also follows naturally from a Bayesian perspective, because it represents a state of maximum information entropy (or maximum uncertainty)

given the available information [17]. Since  $\mathbf{u}_k$  and  $\mathbf{e}_k^m$  are assumed to be discrete white noise random processes, the state-space model can be reformulated as

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k \quad (3)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k, \quad (4)$$

where  $\mathbf{w}_k := \mathbf{B}\mathbf{u}_k$  and  $\mathbf{v}_k := \mathbf{D}\mathbf{u}_k + \mathbf{e}_k^m$  are called the process noise and output noise, respectively, and

$$\mathbb{E} \left( \begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \begin{bmatrix} \mathbf{w}_{k+l} \\ \mathbf{v}_{k+l} \end{bmatrix}^T \right) = \begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{bmatrix} \delta_1(l). \quad (5)$$

The noise covariance matrices  $\mathbf{Q} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{S} \in \mathbb{R}^{n \times n_y}$  and  $\mathbf{R} \in \mathbb{R}^{n_y \times n_y}$  are time-invariant, and  $\delta_1(l)$  represents the unit impulse function, i.e.,  $\delta_1(0) = 1$  and  $\delta_1(l) = 0$ ,  $l \neq 0$ , with  $l$  denoting a time lag in number of samples.

## 2.2. Reference-based Kalman filter and forward innovation model

The Kalman filter plays an important role in data-driven stochastic subspace identification, in the sense that it enables to transform the system model (3-4) into an alternative form that can be employed for estimating the system matrices  $\mathbf{A}$  and  $\mathbf{C}$  from a sequence of observed responses  $\mathbf{y}_k$ . The essential steps that are necessary for arriving at this alternative system description are summarized here.

The Kalman filter operates in the following setting. Suppose that the system matrices  $\mathbf{A}$  and  $\mathbf{C}$  and the noise covariance matrices  $\mathbf{Q}$ ,  $\mathbf{R}$  and  $\mathbf{S}$  are known, as well as the outputs  $\mathbf{y}_0, \dots, \mathbf{y}_k$ . Given that the stochastic terms  $\mathbf{w}_k$  and  $\mathbf{v}_k$  in (3-4) are not known, how can the next state  $\mathbf{x}_{k+1}$  and the next output  $\mathbf{y}_{k+1}$  (which has not been observed yet) be estimated? Kalman [18] offered a solution to this problem by constructing linear estimates  $\hat{\mathbf{x}}_{k+1}$  and  $\hat{\mathbf{y}}_{k+1}$  that are unbiased and have minimum variance. Reynders and De Roeck [19, App. A] extended the conventional Kalman filter to a reference-based version. In the reference-based Kalman filter, only a subset of the outputs, called reference outputs  $\mathbf{y}_k^{\text{ref}}$ , are employed for the construction of the estimates  $\mathbf{x}_{k+1}$  and  $\mathbf{y}_{k+1}$ . They are defined as

$$\mathbf{y}_k^{\text{ref}} = \mathbf{S}^r \mathbf{y}_k, \quad (6)$$

where  $\mathbf{S}^r \in \mathbb{R}^{n_r \times n_y}$  is a selection matrix and  $n_r$  denotes the number of reference outputs. A selection matrix is a matrix of which each row contains only a single nonzero element, and this element equals unity. The set of reference outputs can be chosen arbitrarily as long as all system modes are observable in this set. However, the best reference-based Kalman filter state estimates are usually obtained when the output channels with the highest ratios of signal to measurement noise are chosen as reference outputs. In multi-setup testing, the set of reference outputs is usually chosen equal to the set of outputs that are common to all setups. If there is no specific reason for working with references, the set of reference outputs can be taken equal to the entire set of outputs, in which case  $\mathbf{S}^r = \mathbf{I}$ .

The construction of the reference-based Kalman filter state and output estimates can be summarized in the following recursive set of equations [19]:

$$\hat{\mathbf{y}}_{k+1} = \mathbf{C}\hat{\mathbf{x}}_{k+1} \quad (7)$$

$$\hat{\mathbf{x}}_{k+1} = (\mathbf{A} - \mathbf{K}_k \mathbf{C}^{\text{ref}}) \hat{\mathbf{x}}_k + \mathbf{K}_k \mathbf{y}_k^{\text{ref}} \quad (8)$$

$$\mathbf{K}_k = (\mathbf{A} \mathbf{P}_k \mathbf{C}^{\text{ref}T} + \mathbf{S}^{\text{ref}}) (\mathbf{R}^{\text{ref}} + \mathbf{C}^{\text{ref}} \mathbf{P}_k \mathbf{C}^{\text{ref}T})^{-1} \quad (9)$$

$$\mathbf{P}_{k+1} = \mathbf{A} \mathbf{P}_k \mathbf{A}^T + \mathbf{Q} - (\mathbf{A} \mathbf{P}_k \mathbf{C}^{\text{ref}T} + \mathbf{S}^{\text{ref}}) (\mathbf{R}^{\text{ref}} + \mathbf{C}^{\text{ref}} \mathbf{P}_k \mathbf{C}^{\text{ref}T})^{-1} (\mathbf{A} \mathbf{P}_k \mathbf{C}^{\text{ref}T} + \mathbf{S}^{\text{ref}})^T \quad (10)$$

where  $\mathbf{C}^{\text{ref}} := \mathbf{S}^r \mathbf{C}$ ,  $\mathbf{R}^{\text{ref}} := \mathbf{S}^r \mathbf{R} \mathbf{S}^{rT}$ ,  $\mathbf{S}^{\text{ref}} := \mathbf{S} \mathbf{S}^{rT}$ ,  $\mathbf{K}_k$  is the non-stationary reference-based Kalman filter gain, and  $\mathbf{P}_k$  is the correlation of the state estimation error:

$$\mathbf{P}_k := \mathbb{E} [\mathbf{e}_k \mathbf{e}_k^T] \quad \text{where} \quad \mathbf{e}_k := \mathbf{x}_k - \hat{\mathbf{x}}_k. \quad (11)$$

As starting values for the Kalman filter,  $\hat{\mathbf{x}}_0 = \mathbf{0}$  and  $\mathbf{P}_0 = \mathbf{0}$  are commonly chosen.

The forward innovation  $\mathbf{e}_k^f$  is defined as the error between the true system output  $\mathbf{y}_k$  and the system output  $\hat{\mathbf{y}}_k$  that is obtained from the reference-based Kalman filter state estimate  $\hat{\mathbf{x}}_k$ :

$$\mathbf{e}_k^f := \mathbf{y}_k - \hat{\mathbf{y}}_k. \quad (12)$$

The innovations are uncorrelated with the outputs from which the estimate  $\hat{\mathbf{y}}_k$  is constructed [4]. Combining the above definition with eqs. (7-8) leads to

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{K}_k\mathbf{S}^r\mathbf{e}_k^f \quad (13)$$

$$\mathbf{y}_k = \mathbf{C}\hat{\mathbf{x}}_k + \mathbf{e}_k^f \quad (14)$$

The system model (13-14) is termed the (non-stationary) forward innovation model of the system. It is easy to show that, when the starting point for the construction of the Kalman filter is  $k = l > 0$  instead of  $k = 0$  (i.e., the starting values are  $\hat{\mathbf{x}}_l = \mathbf{0}$  and  $\mathbf{P}_l = \mathbf{0}$ ), then (13-14) modifies into

$$\hat{\mathbf{x}}_{k+1}^{(l)} = \mathbf{A}\hat{\mathbf{x}}_k^{(l)} + \mathbf{K}_{k-l}\mathbf{S}^r\mathbf{e}_k^{f(1)}, \quad k \geq l \quad (15)$$

$$\mathbf{y}_k = \mathbf{C}\hat{\mathbf{x}}_k^{(l)} + \mathbf{e}_k^{f(1)}, \quad (16)$$

where  $\hat{\mathbf{x}}_k^{(l)}$  denotes the Kalman filter state sequence for which the starting value was fixed at  $k = l$ , and  $\mathbf{e}_k^{f(1)}$  denotes the corresponding forward innovation.

### 3. Data-driven stochastic subspace identification

The reference-based data-driven stochastic subspace identification (SSI-data) algorithm aims at identifying the system matrices  $\mathbf{A}$  and  $\mathbf{C}$ , and sometimes also the noise covariance matrices  $\mathbf{Q}$ ,  $\mathbf{R}$  and  $\mathbf{S}$ , from a sequence of outputs  $\mathbf{y}_0, \dots, \mathbf{y}_{N-1}$  where  $N$  denotes the total number of output samples. It is assumed that the system of interest obeys the stochastic state-space model or, equivalently, the forward innovation model that was presented in the previous section. The purpose of the present section is to concisely introduce the SSI-data method and the related notations; more detailed treatments and mathematical proofs can be found in e.g. [5, 20].

The method starts by collecting the outputs into a large block Hankel matrix  $\mathbf{H}$  with  $2i$  block rows and  $j$  columns:

$$\mathbf{H} := \begin{bmatrix} \mathbf{Y}_{0|i-1}^{\text{ref}} \\ \mathbf{Y}_{i|2i} \end{bmatrix} := \frac{1}{\sqrt{j}} \begin{bmatrix} \mathbf{y}_0^{\text{ref}} & \mathbf{y}_1^{\text{ref}} & \cdots & \mathbf{y}_{j-1}^{\text{ref}} \\ \mathbf{y}_1^{\text{ref}} & \mathbf{y}_2^{\text{ref}} & \cdots & \mathbf{y}_j^{\text{ref}} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{y}_{i-1}^{\text{ref}} & \mathbf{y}_i^{\text{ref}} & \cdots & \mathbf{y}_{j+i-2}^{\text{ref}} \\ \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{j+i-1} \\ \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \cdots & \mathbf{y}_{j+i} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{y}_{2i} & \mathbf{y}_{2i+1} & \cdots & \mathbf{y}_N \end{bmatrix}. \quad (17)$$

The outputs are scaled with a factor  $1/\sqrt{j}$  for reasons that will be explained further. Subsequently, an LQ factorization of this matrix is performed, i.e., a factorization into a lower triangular matrix and a matrix with orthonormal rows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{Y}_{0|i-1}^{\text{ref}} \\ \mathbf{Y}_{i|i}^{\text{ref}} \\ \mathbf{Y}_{i|i}^{\sim\text{ref}} \\ \mathbf{Y}_{i+1|2i-1} \\ \mathbf{Y}_{2i|2i} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{21} & \mathbf{L}_{22} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{31} & \mathbf{L}_{32} & \mathbf{L}_{33} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{41} & \mathbf{L}_{42} & \mathbf{L}_{43} & \mathbf{L}_{44} & \mathbf{0} \\ \mathbf{L}_{51} & \mathbf{L}_{52} & \mathbf{L}_{53} & \mathbf{L}_{54} & \mathbf{L}_{55} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \\ \mathbf{Q}_3 \\ \mathbf{Q}_4 \\ \mathbf{Q}_5 \end{bmatrix} \quad (18)$$

Note that the bottom part  $\mathbf{Y}_{i|2i}$  of  $\mathbf{H}$  has been divided into smaller parts and that the corresponding definitions will be employed in what follows.  $\mathbf{Y}_{i|i}^{\text{ref}}$  is the part of  $\mathbf{Y}_{i|i}$  containing the reference channels, while  $\mathbf{Y}_{i|i}^{\sim\text{ref}}$  is the part containing the non-reference channels.

Each column of the block Hankel matrix  $\mathbf{H}$  represents a data sequence. The top part,  $\mathbf{Y}_{0|i-1}^{\text{ref}}$ , contains the ‘past’ part of each sequence, that will be employed for constructing the next Kalman filter state. The bottom half of the block Hankel matrix,  $\mathbf{Y}_{i|2i}$  contains the ‘future’ part of each sequence. This part is now elaborated in forward innovation form. As an example, consider the first column of  $\mathbf{Y}_{i|2i-1}$ . By employing the forward innovation model (15-16), it can be elaborated as

$$\begin{bmatrix} \mathbf{y}_i \\ \mathbf{y}_{i+1} \\ \vdots \\ \mathbf{y}_{2i-1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix}}_{:=\mathcal{O}_i} \hat{\mathbf{x}}_i^{(0)} + \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{CK}_i\mathbf{S}^r & \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{K}_{2i-2}\mathbf{S}^r & \mathbf{CA}^{i-3}\mathbf{K}_{2i-3}\mathbf{S}^r & \dots & \mathbf{I} \end{bmatrix}}_{:=\mathcal{F}_i} \begin{bmatrix} \mathbf{e}_i^{f(0)} \\ \mathbf{e}_{i+1}^{f(0)} \\ \vdots \\ \mathbf{e}_{2i-1}^{f(0)} \end{bmatrix} \quad (19)$$

Note that  $\hat{\mathbf{x}}_i^{(0)}$  represents the reference-based Kalman filter state corresponding to the initial values  $\mathbf{K}_0 = \mathbf{0}$  and  $\mathbf{P}_0 = \mathbf{0}$ , and  $\mathbf{e}_i^{f(0)}$  till  $\mathbf{e}_{2i-1}^{f(0)}$  are the corresponding forward innovations. The matrix  $\mathcal{O}_i$  which is defined in Eq. (19) is termed the observability matrix. It contains the unknown system matrices  $\mathbf{A}$  and  $\mathbf{C}$ , and the first step of the SSI-data algorithm is the identification of the observability matrix from the measured outputs. In order to achieve this, a similar decomposition as in (19) is performed for all subsequent columns of  $\mathbf{Y}_{i|2i-1}$ , which leads to

$$\mathbf{Y}_{i|2i-1} = \mathcal{O}_i \mathbf{X}_{i|i} + \mathcal{F}_i \mathbf{E}_{i|2i-1}^f \quad (20)$$

where the matrices  $\mathcal{O}_i$  and  $\mathcal{F}_i$  have been defined in Eq. (19),

$$\mathbf{X}_{i|i} := \frac{1}{\sqrt{j}} \begin{bmatrix} \hat{\mathbf{x}}_i^{(0)} & \hat{\mathbf{x}}_{i+1}^{(1)} & \dots & \hat{\mathbf{x}}_{j+i-1}^{(j-1)} \end{bmatrix} \quad \text{and} \quad \mathbf{E}_{i|2i-1}^f := \frac{1}{\sqrt{j}} \begin{bmatrix} \mathbf{e}_i^{f(0)} & \mathbf{e}_{i+1}^{f(1)} & \dots & \mathbf{e}_{j+i-1}^{f(j-1)} \\ \mathbf{e}_{i+1}^{f(0)} & \mathbf{e}_{i+2}^{f(1)} & \dots & \mathbf{e}_{j+i-2}^{f(j-1)} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{e}_{2i-1}^{f(0)} & \mathbf{e}_{2i}^{f(1)} & \dots & \mathbf{e}_{N-1}^{f(j-1)} \end{bmatrix}. \quad (21)$$

Note that the columns of  $\mathbf{X}_{i|i}$  do not belong to the same Kalman filter state sequence, because each column is estimated from  $i$  past outputs, and therefore the starting index of the filtering process is different. The corresponding matrix with initial values is

$$\mathbf{X}_{0|0} := 1/\sqrt{j} \begin{bmatrix} \hat{\mathbf{x}}_0^{(0)} & \hat{\mathbf{x}}_1^{(1)} & \dots & \hat{\mathbf{x}}_{j-1}^{(j-1)} \end{bmatrix} = \mathbf{0} \quad (22)$$

Eq. (20) can be interpreted as follows: the future output sequence matrix  $\mathbf{Y}_{i|2i-1}$  can be decomposed into a free vibration sequence matrix which depends on the initial condition matrix  $\mathbf{X}_{i|i}$ , and a forced vibration sequence matrix which depends on the excitation sequence matrix  $\mathbf{E}_{i|2i-1}^f$ .

The key step of the subspace identification algorithm is the following. Since the forward innovation vectors in  $\mathbf{E}_{i|2i-1}^f$  are uncorrelated from the corresponding Kalman filter states in  $\mathbf{X}_{i|i}$ , they are also uncorrelated from the corresponding past reference outputs in  $\mathbf{Y}_{0|i-1}^{\text{ref}}$ . The Kalman filter states however are fully correlated with the past outputs. The free vibration part of the future output matrix can therefore be asymptotically recovered by projecting the future output matrix onto the row space of the past reference output matrix:

$$\text{a.s. } \lim_{j \rightarrow \infty} \mathbf{Y}_{i|2i-1} / \mathbf{Y}_{0|i-1}^{\text{ref}} = \mathcal{O}_i \mathbf{X}_{i|i} \quad (23)$$

With the LQ decomposition defined above, the projection can be computed as

$$\mathbf{Y}_{i|2i-1} / \mathbf{Y}_{0|i-1}^{\text{ref}} = \begin{bmatrix} \mathbf{L}_{21} \\ \mathbf{L}_{31} \\ \mathbf{L}_{41} \end{bmatrix} \mathbf{Q}_1. \quad (24)$$

Eq. (23) reveals that the observability and state matrices can be estimated from a singular value decomposition of the projected data matrix  $\mathcal{P}$ :

$$\mathcal{P} := \mathbf{W}_1 \left( \mathbf{Y}_{i|2i-1} / \mathbf{Y}_{0|i-1}^{\text{ref}} \right) \mathbf{W}_2 = \mathbf{USV}^T \quad (25)$$

which leads to

$$\hat{\mathcal{O}}_i = \mathbf{W}_1^{-1} \mathbf{U} \mathbf{S}^{\frac{1}{2}} \quad \text{and} \quad \hat{\mathbf{X}}_{i|i} = \hat{\mathcal{O}}_i^\dagger \mathbf{Y}_{i|2i-1} / \mathbf{Y}_{0|i-1}^{\text{ref}}. \quad (26)$$

The matrices  $\mathbf{W}_1$  and  $\mathbf{W}_2$  which appear in the above expressions are weighting matrices that can be chosen arbitrarily as long as  $\mathbf{W}_1$  has full rank, and  $\mathbf{W}_2$  preserves the rank of the matrices with which it is multiplied. Particular combinations of weights result in specific algorithms, such as [5]

- the unweighted principal component (UPC) algorithm, for which  $\mathbf{W}_1 = \mathbf{I}$  and  $\mathbf{W}_2 = \mathbf{I}$ ;
- the principal component (PC) algorithm, for which  $\mathbf{W}_1 = \mathbf{I}$  and  $\mathbf{W}_2 = \mathbf{Y}_{0|i-1}^{\text{ref}} \left( \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{ T}} \right)^{-\frac{1}{2}} \mathbf{Y}_{0|i-1}^{\text{ref}}$ ;
- the canonical variate algorithm (CVA), for which  $\mathbf{W}_1 = \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}} \right)^{-\frac{1}{2}}$  and  $\mathbf{W}_2 = \mathbf{I}$ .

The estimates  $\hat{\mathcal{O}}_i$  and  $\hat{\mathbf{X}}_{i|i}$  are strongly consistent, which means that they converge to their true values, when the number of columns  $j$  in the data block Hankel matrix  $\mathbf{H}$  becomes larger and larger. This is because the projection of the future output matrix onto the past reference output matrix results in the exact free response for  $j \rightarrow \infty$ , see Eq. (23).

Several options are available for estimating the system matrices  $\mathbf{A}$  and  $\mathbf{C}$ . One option is to make use of the shift structure of the observability matrix  $\hat{\mathcal{O}}_i$ , as is done in covariance-driven stochastic subspace identification (SSI-cov). However, in the SSI-data algorithm as introduced in [6] for modal testing, the system matrices are estimated from the identified Kalman filter state sequence. In order to recover the Kalman filter state matrix  $\mathbf{X}_{i+1|i+1}$  which relates to  $\mathbf{X}_{i|i}$  and the initial condition  $\mathbf{X}_{0|0}$ , an expression similar to (20) can be derived:

$$\mathbf{Y}_{i+1|2i} = \mathcal{O}_i \mathbf{X}_{i+1|i+1} + \mathcal{F}_{i+1} \mathbf{E}_{i+1|2i}^f. \quad (27)$$

A strongly consistent estimate of  $\mathbf{X}_{i+1|i+1}$  is then

$$\hat{\mathbf{X}}_{i+1|i+1} := \hat{\mathcal{O}}_i^\dagger \mathbf{Y}_{i+1|2i} / \mathbf{Y}_{0|i}^{\text{ref}}. \quad (28)$$

The orthogonal projection which appears in this expression can be computed with the LQ decomposition that was defined earlier:

$$\mathbf{Y}_{i+1|2i} / \mathbf{Y}_{0|i}^{\text{ref}} = \begin{bmatrix} \mathbf{L}_{41} & \mathbf{L}_{42} \\ \mathbf{L}_{51} & \mathbf{L}_{52} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix}. \quad (29)$$

Application of the reference-based forward innovation model results in

$$\begin{bmatrix} \hat{\mathbf{X}}_{i+1|i+1} \\ \mathbf{Y}_{i|i} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \hat{\mathbf{X}}_{i|i} + \begin{bmatrix} \mathbf{K}_i \mathbf{S}^r \mathbf{E}_{i|i}^f \\ \mathbf{E}_{i|i}^f \end{bmatrix}. \quad (30)$$

Since the estimation errors  $\mathbf{E}_{i|i}^f$  are uncorrelated from the Kalman filter states  $\mathbf{X}_{i|i}$ , the following estimate of the system matrices is strongly consistent:

$$\begin{bmatrix} \hat{\mathbf{A}} \\ \hat{\mathbf{C}} \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{X}}_{i+1|i+1} \\ \mathbf{Y}_{i|i} \end{bmatrix} \hat{\mathbf{X}}_{i|i}^\dagger \quad (31)$$

where the operator  $\dagger$  denotes the Moore-Penrose pseudo-inverse.

#### 4. Relation between the identified system matrices and sample output correlations

In the previous section, the SSI-data algorithm was presented in its conventional form [5, 6]. In the present section, the algorithm is analyzed in more detail in order to reveal that the identified system matrices are exclusively estimated from output sample correlation matrices. Such a demonstration can also be found in [15] for a different version of the SSI-data algorithm with UPC weighting, in which the system matrices are obtained from the observability matrix instead of the state sequences.

Before starting this analysis, it is helpful to review the following two basic results from linear algebra. Consider any two matrices  $\mathcal{A}$  and  $\mathcal{B}$  with an equal number of columns and suppose that  $\mathcal{B}$  has full row rank. First, under these conditions, the pseudo-inverse of  $\mathcal{B}$  equals

$$\mathcal{B}^\dagger = \mathcal{B}^T (\mathcal{B}\mathcal{B}^T)^{-1}. \quad (32)$$

Second, the projection of  $\mathcal{A}$  onto the row space of  $\mathcal{B}$  equals

$$\mathcal{A}/\mathcal{B} = \mathcal{A}\mathcal{B}^\dagger\mathcal{B}. \quad (33)$$

With the second result, one finds from (26) and (28) that, respectively,

$$\hat{\mathbf{X}}_{i|i} = \hat{\mathbf{O}}_i^\dagger \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \mathbf{Y}_{0|i-1}^{\text{ref}} \quad (34)$$

and

$$\hat{\mathbf{X}}_{i+1|i+1} = \hat{\mathbf{O}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref}\dagger} \mathbf{Y}_{0|i}^{\text{ref}}. \quad (35)$$

Subsequent application of the first result leads to

$$\hat{\mathbf{X}}_{i|i}^\dagger = \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \left( \hat{\mathbf{O}}_i^\dagger \right)^T \mathbf{J}_1^{-1} \quad (36)$$

with

$$\mathbf{J}_1 := \hat{\mathbf{O}}_i^\dagger \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{i|2i-1}^T \left( \hat{\mathbf{O}}_i^\dagger \right)^T. \quad (37)$$

It may be noted that (36) requires that  $\hat{\mathbf{X}}_{i|i}$  has full row rank. This is indeed the case thanks to the persistency of excitation of the white noise input [21]. Substitution of these expressions into (31) results in

$$\begin{bmatrix} \hat{\mathbf{A}} \\ \hat{\mathbf{C}} \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{O}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref}\dagger} \mathbf{Y}_{0|i}^{\text{ref}} \\ \mathbf{Y}_{i|i} \end{bmatrix} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \left( \hat{\mathbf{O}}_i^\dagger \right)^T \mathbf{J}_1^{-1}. \quad (38)$$

This result already reveals that, apart from the observability matrix estimate  $\hat{\mathbf{O}}_i$  that will be analyzed next, the estimated system matrices are obtained exclusively from output sample correlation matrices. This is because the parts of the output Hankel matrix  $\mathbf{H}$  that are involved in this expression do not appear on their own, but always in multiplication with the transpose of another part. For example, the first block column of  $\mathbf{Y}_{i|i} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger}$  equals

$$\mathbf{Y}_{i|i} \mathbf{Y}_{0|0}^{\text{ref}\dagger} = \frac{1}{j} \sum_{k=0}^{j-1} \mathbf{y}_{k+i} \mathbf{y}_k^{\text{ref}\dagger} \quad \text{such that} \quad \text{a.s. } \lim_{j \rightarrow \infty} \mathbf{Y}_{i|i} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} = \mathbf{E} \left( \mathbf{y}_{k+i} \mathbf{y}_k^{\text{ref}\dagger} \right). \quad (39)$$

The second equality states that the sample correlation converges almost surely to the true correlation.

In order to obtain the link between the estimate of the extended observability matrix,  $\hat{\mathbf{O}}_i$ , and the sample output correlations, one can proceed as follows [15]. Let us first define the matrix  $\mathbf{P}$  as

$$\mathbf{P} := \mathcal{P}\mathcal{P}^T. \quad (40)$$

For the UPC and CVA algorithms, one has that

$$\mathbf{P} = \mathbf{W}_1 \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \left( \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^{-1} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \mathbf{W}_1^T \quad (41)$$

while for the PC algorithm, a simpler expression is obtained:

$$\mathbf{P} = \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger} \right)^T \quad (42)$$

In both cases  $\mathbf{P}$  is seen to consist entirely of sample output correlations. With (25), one has that

$$\mathbf{P} = \mathbf{U}\mathbf{S}^2\mathbf{U}^T \quad (43)$$

This means that the matrices  $\mathbf{U}$  and  $\mathbf{S}$  of which  $\hat{\mathbf{O}}_i$  is composed according to (26), can be obtained as the matrix with eigenvectors and squared eigenvalues of  $\mathbf{P}$ , respectively.

## 5. Covariance of the estimated system matrices

### 5.1. General approach

As discussed previously and illustrated in (39), the elements of  $\mathbf{H}\mathbf{H}^T$  represent sample output correlations. As they are estimated from a finite number of samples  $j$ , they differ from the true output correlations. In order to simplify the notation, let us define the statistical estimator of the output correlations as  $\hat{\boldsymbol{\theta}}$ , and the matrix with the true values as  $\boldsymbol{\theta}$ :

$$\hat{\boldsymbol{\theta}} := \text{vec}(\mathbf{H}\mathbf{H}^T) \quad \text{and} \quad \boldsymbol{\theta} := \mathbb{E}(\hat{\boldsymbol{\theta}}), \quad (44)$$

where  $\text{vec}$  is a vectorization operator that stacks the columns of a matrix on top of each other. The estimation error is then defined as

$$\Delta\boldsymbol{\theta} := \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}. \quad (45)$$

The analysis in the previous section has revealed that the estimated system matrices  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  are an implicit function of the elements of  $\mathbf{H}\mathbf{H}^T$ :

$$\begin{bmatrix} \text{vec}(\hat{\mathbf{A}}) \\ \text{vec}(\hat{\mathbf{C}}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\hat{\boldsymbol{\theta}}) \\ \mathbf{g}(\hat{\boldsymbol{\theta}}) \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \\ \mathbf{g}(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \end{bmatrix}, \quad (46)$$

where  $\mathbf{f}$  and  $\mathbf{g}$  are vector functions. As a result, the matrices  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  are (non-exact) estimates of the system matrices  $\mathbf{A}$  and  $\mathbf{C}$ , even when the true system perfectly obeys all assumptions that have been outlined in Sec. 2. If the errors  $\Delta\boldsymbol{\theta}$  are sufficiently small, then an approximate linear relation between the system matrix errors and the output correlation errors can be obtained with a first-order Taylor expansion:

$$\begin{bmatrix} \text{vec}(\Delta\mathbf{A}) \\ \text{vec}(\Delta\mathbf{C}) \end{bmatrix} \approx \begin{bmatrix} \mathbf{J}_f \\ \mathbf{J}_g \end{bmatrix} \Delta\boldsymbol{\theta}, \quad (47)$$

where  $\mathbf{J}_f$  and  $\mathbf{J}_g$  are Jacobian matrices, i.e., their elements are first order derivatives:

$$J_{f,kl} := \frac{\partial f_k(\boldsymbol{\theta})}{\theta_l} \quad \text{and} \quad J_{g,kl} := \frac{\partial g_k(\boldsymbol{\theta})}{\theta_l}. \quad (48)$$

The covariance of the system matrix estimates is then approximately obtained as

$$\text{Cov}_{AC} := \mathbb{E} \left( \begin{bmatrix} \text{vec}(\Delta\mathbf{A}) \\ \text{vec}(\Delta\mathbf{C}) \end{bmatrix} \begin{bmatrix} \text{vec}(\Delta\mathbf{A}) \\ \text{vec}(\Delta\mathbf{C}) \end{bmatrix}^T \right) \approx \begin{bmatrix} \mathbf{J}_f \\ \mathbf{J}_g \end{bmatrix} \mathbb{E}(\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^T) \begin{bmatrix} \mathbf{J}_f \\ \mathbf{J}_g \end{bmatrix}^T \quad (49)$$

The covariance matrix of the output sample correlations can be obtained by splitting up the data Hankel matrix  $\mathbf{H}$  into  $n_b$  blocks of  $j_b$  columns each:

$$\mathbf{H} = [ \mathbf{H}_1 \mid \mathbf{H}_2 \mid \dots \mid \mathbf{H}_{n_b} ] = \begin{bmatrix} \mathbf{Y}_{0|i-1,1}^{\text{ref}} & \mathbf{Y}_{0|i-1,2}^{\text{ref}} & \dots & \mathbf{Y}_{0|i-1,n_b}^{\text{ref}} \\ \mathbf{Y}_{i|2i,1} & \mathbf{Y}_{i|2i,2} & \dots & \mathbf{Y}_{i|2i,n_b} \end{bmatrix}. \quad (50)$$

The output sample correlations are then computed as

$$\mathbf{H}\mathbf{H}^T = \sum_{b=1}^{n_b} \mathbf{H}_b \mathbf{H}_b^T \quad (51)$$

If the block length  $j_b$  is sufficiently large, the data blocks are approximately statistically independent. If this is indeed the case, the covariance of the output sample correlations can be estimated as

$$\mathbb{E}(\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^T) = \mathbb{E} \left( \text{vec}(\Delta(\mathbf{H}\mathbf{H}^T)) \text{vec}(\Delta(\mathbf{H}\mathbf{H}^T))^T \right) \quad (52)$$

$$\approx n_b \mathbb{E} \left( \text{vec}(\Delta(\mathbf{H}_b \mathbf{H}_b^T)) \text{vec}(\Delta(\mathbf{H}_b \mathbf{H}_b^T))^T \right) \quad (53)$$



Let us now denote the output Hankel matrix that is constructed from the observed data in a specific experiment as  $\bar{\mathbf{H}}$ , i.e.,  $\bar{\mathbf{H}}$  is an observation of  $\mathbf{H}$ . The expected value which appears in the previous expression can then be obtained as the sample mean estimate that is obtained from the observed block samples:

$$\mathbb{E} \left( \text{vec} \left( \Delta \left( \mathbf{H}_b \mathbf{H}_b^T \right) \right) \text{vec} \left( \Delta \left( \mathbf{H}_b \mathbf{H}_b^T \right) \right)^T \right) \approx \frac{1}{n_b} \sum_{b=1}^{n_b} \Delta \bar{\theta}_b \Delta \bar{\theta}_b^T \quad (54)$$

where

$$\Delta \bar{\theta}_b := \text{vec} \left( \frac{j}{j_b} \bar{\mathbf{H}}_b \bar{\mathbf{H}}_b^T - \bar{\mathbf{H}} \bar{\mathbf{H}}^T \right). \quad (55)$$

This derivation illustrates that it is possible to estimate the covariance of the identified system matrices as follows: estimate the errors for each data block as in (55), estimate subsequently the covariance of the sample output correlations from (53) and (54), compute the Jacobian matrices  $\mathbf{J}_f$  and  $\mathbf{J}_g$ , and finally apply (49). This is the approach that was followed in previous works on practical covariance estimation for SSI algorithms that make use of the shift-invariant structure of the observability matrix [13, 15]. However, the derivation of explicit expressions for the Jacobian matrices involves the extensive use of Kronecker algebra, and the matrices involved tend to have large dimensions, leading to substantial memory usage unless an optimized implementation is developed, as in [14].

Therefore, a different approach is proposed in the present work. The main idea is not to derive explicit expressions for the Jacobians anymore, but rather to directly estimate the first-order errors  $\Delta \mathbf{A}_b$  and  $\Delta \mathbf{C}_b$  that are caused by the errors on the sample output correlation estimates for a specific block  $b$ ,  $\Delta \theta_b$ . Once  $\Delta \mathbf{A}_b$  and  $\Delta \mathbf{C}_b$  are obtained for every block  $b$ , the covariance of the system matrix estimates,  $\mathbf{Cov}_{AC}$ , follows directly from these estimated errors. Both the explicit and the implicit approaches yield the same final result, as can be seen from the substitution of (53) and (54) into (49):

$$\mathbf{Cov}_{AC} \approx \sum_{b=1}^{n_b} \begin{bmatrix} \text{vec} \left( \Delta \bar{\mathbf{A}}_b \right) \\ \text{vec} \left( \Delta \bar{\mathbf{C}}_b \right) \end{bmatrix} \begin{bmatrix} \text{vec} \left( \Delta \bar{\mathbf{A}}_b \right) \\ \text{vec} \left( \Delta \bar{\mathbf{C}}_b \right) \end{bmatrix}^T \quad \text{where} \quad \begin{bmatrix} \text{vec} \left( \Delta \bar{\mathbf{A}}_b \right) \\ \text{vec} \left( \Delta \bar{\mathbf{C}}_b \right) \end{bmatrix} \approx \begin{bmatrix} \mathbf{J}_f \\ \mathbf{J}_g \end{bmatrix} \Delta \bar{\theta}_b \quad (56)$$

The mathematical derivation is however less complex for the implicit approach, and memory issues are also automatically avoided since the dimensions of the matrices involved do not exceed those that appear in the original algorithm.

## 5.2. Computation of first-order system matrix errors

### 5.2.1. General first-order sensitivity results

Before starting the detailed analysis, it is again helpful to review some results from matrix algebra. If a matrix  $\mathcal{A}$  is square and has full rank, then it follows immediately from the identities

$$\mathcal{A} \mathcal{A}^{-1} = \mathbf{I} \quad \text{and} \quad (\mathcal{A} + \Delta \mathcal{A}) (\mathcal{A}^{-1} + \Delta \mathcal{A}^{-1}) = \mathbf{I} \quad (57)$$

that, up to first-order accuracy,

$$\Delta \mathcal{A}^{-1} \approx -\mathcal{A}^{-1} \Delta \mathcal{A} \mathcal{A}^{-1}. \quad (58)$$

Combining this result with (32), one finds that, if  $\mathcal{A}$  is not necessarily square but has full row rank,

$$\Delta \mathcal{A}^\dagger \approx \left( \mathbf{I} - \mathcal{A}^\dagger \mathcal{A} \right) \Delta \mathcal{A}^T \left( \mathcal{A} \mathcal{A}^T \right)^{-1} - \mathcal{A}^\dagger \Delta \mathcal{A} \mathcal{A}^\dagger. \quad (59)$$

Similarly, when  $\mathcal{A}$  has full column rank,

$$\mathcal{A}^\dagger = \left( \mathcal{A}^T \mathcal{A} \right)^{-1} \mathcal{A}^T \quad \text{and} \quad \Delta \mathcal{A}^\dagger \approx \left( \mathcal{A}^T \mathcal{A} \right)^{-1} \Delta \mathcal{A}^T \left( \mathbf{I} - \mathcal{A} \mathcal{A}^\dagger \right) - \mathcal{A}^\dagger \Delta \mathcal{A} \mathcal{A}^\dagger. \quad (60)$$

If  $\mathcal{A}$  is a real symmetric matrix of full rank and with unique eigenvalues, then it admits an orthonormal eigenvalue decomposition of the form

$$\mathcal{A} = \mathbf{U} \Lambda \mathbf{U}^T \quad \text{with} \quad \mathbf{U}^T \mathbf{U} = \mathbf{U} \mathbf{U}^T = \mathbf{I}, \quad (61)$$

where  $\Lambda$  is a diagonal matrix, the nonzero diagonal elements of which are denoted as  $\lambda_k$ . The corresponding columns of  $\mathbf{U}$  are the orthonormal eigenvectors  $\mathbf{u}_k$ , such that

$$\mathcal{A}\mathbf{u}_k = \lambda_k\mathbf{u}_k \quad \text{and} \quad \mathbf{u}_k^T \mathcal{A} = \lambda_k\mathbf{u}_k^T \quad (62)$$

from which it directly follows that

$$\Delta\lambda_k \approx \mathbf{u}_k^T \Delta\mathcal{A}\mathbf{u}_k \quad (63)$$

and

$$\Delta\mathbf{u}_k \approx \left( \mathbf{I} - \frac{\mathcal{A}}{\lambda_k} \right)^\dagger \left( \Delta\mathcal{A} \frac{\mathbf{u}_k}{\lambda_k} - \frac{\mathcal{A}\mathbf{u}_k}{\lambda_k^2} \Delta\lambda_k \right). \quad (64)$$

### 5.2.2. First-order error on the observability matrix estimate $\hat{\mathbf{O}}_i$

In this section, the error on the estimated observability matrix due to errors in the sample output correlation matrices of a particular data block  $b$  is derived. Application of the chain rule to (26) leads to

$$\Delta\hat{\mathbf{O}}_{i,b} = \Delta(\mathbf{W}_1^{-1})_b \mathbf{U}\mathbf{S}^{\frac{1}{2}} + \mathbf{W}_1^{-1} \Delta\mathbf{U}_b \mathbf{S}^{\frac{1}{2}} + \mathbf{W}_1^{-1} \mathbf{U} \Delta(\mathbf{S}^{\frac{1}{2}})_b \quad (65)$$

The error on the weighting matrix is zero for the UPC and PC algorithms. For the CVA weighting, this error will be quantified in section 5.2.5.

From (43), it is clear that the columns  $\mathbf{u}_k$  of  $\mathbf{U}$  are the orthonormal eigenvectors of  $\mathbf{P}$ , and the squares of the diagonal elements  $s_k$  of  $\mathbf{S}$  are the eigenvalues of  $\mathbf{P}$ . One can then invoke the general results (63) and (64) to arrive at

$$\Delta\mathbf{u}_{k,b} = \left( \mathbf{I} - \frac{\mathbf{P}}{s_k^2} \right)^\dagger \left( \Delta\mathbf{P}_b \frac{\mathbf{u}_k}{s_k^2} - \frac{\mathbf{P}\mathbf{u}_k}{s_k^4} \Delta s_{k,b}^2 \right) \quad (66)$$

and

$$\Delta s_{k,b}^2 = \mathbf{u}^T \Delta\mathbf{P}_b \mathbf{u} \quad \text{such that} \quad \Delta s_{k,b}^{1/2} = \frac{1}{4s_k^{3/4}} \mathbf{u}^T \Delta\mathbf{P}_b \mathbf{u}. \quad (67)$$

The remaining question is then, how errors on  $\mathbf{P}$  are related to errors on output covariance matrices. For the UPC and CVA algorithms, the chain rule can be applied to (41), leading to

$$\Delta\mathbf{P}_b \approx \Delta\mathbf{P}_{1,b} + (\Delta\mathbf{P}_{1,b})^T + \Delta\mathbf{P}_{2,b} + (\Delta\mathbf{P}_{2,b})^T + \Delta\mathbf{P}_{3,b} \quad (68)$$

with

$$\Delta\mathbf{P}_{1,b} = \Delta\mathbf{W}_{1,b} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}T} \right) \mathbf{J}_P^T \quad \text{with} \quad \mathbf{J}_P := \mathbf{W}_1 \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}T} \right)^\dagger \quad (69)$$

$$\Delta\mathbf{P}_{2,b} = \mathbf{W}_1 \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{0|i-1,b}^{\text{ref}T} \right) \mathbf{J}_P^T \quad (70)$$

$$\Delta\mathbf{P}_{3,b} = -\mathbf{J}_P \Delta \left( \mathbf{Y}_{0|i-1,b}^{\text{ref}} \mathbf{Y}_{0|i-1,b}^{\text{ref}T} \right) \mathbf{J}_P^T \quad (71)$$

where in the final equality, the general result (58) has been employed. For the PC algorithm, one can easily derive from (42) that

$$\Delta\mathbf{P}_b = \Delta\mathbf{P}_{4,b} + \Delta\mathbf{P}_{4,b}^T \quad \text{with} \quad \Delta\mathbf{P}_{4,b} := \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{0|i-1,b}^{\text{ref}T} \right) \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}T} \right)^\dagger. \quad (72)$$

### 5.2.3. First-order error on the output matrix estimate $\hat{\mathbf{C}}$

The previous results are now employed for deriving the error on the estimated system matrices due to errors in the sample output correlation matrices of a particular data block  $b$ . Application of the chain rule in the expression for  $\hat{\mathbf{C}}$  as available from (38) leads to

$$\Delta\mathbf{C}_b \approx \Delta\mathbf{C}_{1b} + \Delta\mathbf{C}_{2b} + \Delta\mathbf{C}_{3b} + \Delta\mathbf{C}_{4b} \quad (73)$$

where

$$\Delta \mathbf{C}_{1b} := \Delta \left( \mathbf{Y}_{i|b} \mathbf{Y}_{0|i-1,b}^{\text{ref} \dagger \text{T}} \right) \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (74)$$

$$\Delta \mathbf{C}_{2b} := \left( \mathbf{Y}_{i|b} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right) \Delta \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)_b^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (75)$$

$$\Delta \mathbf{C}_{3b} := \left( \mathbf{Y}_{i|b} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right) \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \Delta \hat{\boldsymbol{\theta}}_{i,b}^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (76)$$

$$\Delta \mathbf{C}_{4b} := - \left( \mathbf{Y}_{i|b} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right) \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \Delta \mathbf{J}_{1,b} \mathbf{J}_1^{-1}. \quad (77)$$

The evaluation of  $\Delta \mathbf{C}_{1b}$  is straightforward.  $\Delta \mathbf{C}_{2b}$  can be computed by noting that

$$\begin{aligned} \Delta \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)_b &= \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) \left( \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right)^{-1} \\ &\quad - \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right) \Delta \left( \mathbf{Y}_{0|i-1,b}^{\text{ref}} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) \left( \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right)^{-1} \\ &= \left( \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) - \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \Delta \left( \mathbf{Y}_{0|i-1,b}^{\text{ref}} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) \right) \left( \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \right)^{-1}. \end{aligned} \quad (78)$$

For the evaluation of  $\Delta \mathbf{C}_{3b}$ , one can first invoke (60) and derive that

$$\Delta \hat{\boldsymbol{\theta}}_{i,b}^\dagger = \left( \hat{\boldsymbol{\theta}}_i^{\text{T}} \hat{\boldsymbol{\theta}}_i \right)^{-1} \Delta \hat{\boldsymbol{\theta}}_{i,b}^{\text{T}} \left( \mathbf{I} - \hat{\boldsymbol{\theta}}_i \hat{\boldsymbol{\theta}}_i^\dagger \right) - \hat{\boldsymbol{\theta}}_i^\dagger \Delta \hat{\boldsymbol{\theta}}_{i,b} \hat{\boldsymbol{\theta}}_i^\dagger \quad (80)$$

The computation of  $\Delta \hat{\boldsymbol{\theta}}_{i,b}$  has been treated in section 5.2.2. The computation of  $\Delta \mathbf{C}_{4b}$  requires a first-order sensitivity analysis of  $\mathbf{J}_1$ . Application of the chain rule to (37) leads to

$$\Delta \mathbf{J}_{1,b} := \Delta \mathbf{J}_{11,b} + \Delta \mathbf{J}_{11,b}^{\text{T}} + \Delta \mathbf{J}_{12,b} + \Delta \mathbf{J}_{12,b}^{\text{T}} + \Delta \mathbf{J}_{13,b} \quad (81)$$

where

$$\Delta \mathbf{J}_{11,b} := \Delta \hat{\boldsymbol{\theta}}_{i,b}^\dagger \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \mathbf{Y}_{0|i-1}^{\text{ref}} \mathbf{Y}_{i|2i-1}^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \quad (82)$$

$$\Delta \mathbf{J}_{12,b} := \hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \Delta \left( \mathbf{Y}_{0|i-1,b}^{\text{ref}} \mathbf{Y}_{i|2i-1,b}^{\text{T}} \right) \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \quad (83)$$

$$\Delta \mathbf{J}_{13,b} := -\mathbf{J}_{131} \Delta \left( \mathbf{Y}_{0|i-1,b}^{\text{ref}} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) \mathbf{J}_{131}^{\text{T}} \quad \text{with} \quad \mathbf{J}_{131} := \hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger}. \quad (84)$$

#### 5.2.4. First-order error on the system matrix estimate $\hat{\mathbf{A}}$

Application of the chain rule in the expression for  $\hat{\mathbf{A}}$  as available from (39) leads to

$$\Delta \hat{\mathbf{A}}_b = \Delta \hat{\mathbf{A}}_{1,b} + \Delta \hat{\mathbf{A}}_{2,b} + \Delta \hat{\mathbf{A}}_{3,b} + \Delta \hat{\mathbf{A}}_{4,b} + \Delta \hat{\mathbf{A}}_{5,b} + \Delta \hat{\mathbf{A}}_{6,b} \quad (85)$$

where

$$\Delta \hat{\mathbf{A}}_{1,b} = \Delta \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)_b \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (86)$$

$$\Delta \hat{\mathbf{A}}_{2,b} = \hat{\boldsymbol{\theta}}_i^\dagger \Delta \left( \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \right)_b \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (87)$$

$$\Delta \hat{\mathbf{A}}_{3,b} = \hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \Delta \left( \mathbf{Y}_{0|i,b}^{\text{ref}} \mathbf{Y}_{0|i-1,b}^{\text{ref} \text{T}} \right) \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (88)$$

$$\Delta \hat{\mathbf{A}}_{4,b} = \hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \Delta \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)_b^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (89)$$

$$\Delta \hat{\mathbf{A}}_{5,b} = \hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \Delta \hat{\boldsymbol{\theta}}_{i,b}^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \quad (90)$$

$$\Delta \hat{\mathbf{A}}_{6,b} = -\hat{\boldsymbol{\theta}}_i^\dagger \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref} \dagger} \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i-1}^{\text{ref} \text{T}} \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref} \dagger} \right)^{\text{T}} \left( \hat{\boldsymbol{\theta}}_i^\dagger \right)^{\text{T}} \mathbf{J}_1^{-1} \Delta \mathbf{J}_{1,b} \mathbf{J}_1^{-1}. \quad (91)$$

Following the same lines as in the derivation of (79), one has that

$$\Delta \left( \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref}\dagger} \right)_b = \left( \Delta \left( \mathbf{Y}_{i+1|2i,b} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}} \right) - \mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref}\dagger} \Delta \left( \mathbf{Y}_{0|i,b}^{\text{ref}} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}} \right) \right) \left( \mathbf{Y}_{0|i}^{\text{ref}} \mathbf{Y}_{0|i}^{\text{ref}\text{T}} \right)^{-1}. \quad (92)$$

### 5.2.5. First-order error on the weighting matrix $\mathbf{W}_1$ in the CVA algorithm

Since the matrix  $\mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}}$  is symmetric and of full rank, it admits an eigenvalue decomposition with orthonormal eigenvectors:

$$\mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}} = \mathbf{U}_c \mathbf{\Lambda}_c \mathbf{U}_c^{\text{T}} \quad \text{with} \quad \mathbf{U}_c \mathbf{U}_c^{\text{T}} = \mathbf{U}_c^{\text{T}} \mathbf{U}_c = \mathbf{I} \quad (93)$$

such that, for the CVA algorithm,

$$\mathbf{W}_1^{-1} = \left( \mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}} \right)^{\frac{1}{2}} = \mathbf{U}_c \mathbf{\Lambda}_c^{1/2} \mathbf{U}_c^{\text{T}}. \quad (94)$$

Application of the chain rule then leads to

$$\Delta \left( \mathbf{W}_1^{-1} \right)_b \approx \Delta \mathbf{U}_{c,b} \mathbf{\Lambda}_c^{1/2} \mathbf{U}_c^{\text{T}} + \mathbf{U}_c \Delta \left( \mathbf{\Lambda}_c^{1/2} \right) \mathbf{U}_c^{\text{T}} + \mathbf{U}_c \mathbf{\Lambda}_c^{1/2} \Delta \mathbf{U}_{c,b}^{\text{T}}. \quad (95)$$

Using the general results (63) and (64), one has for the eigenvalues  $\lambda_{c,k}$ :

$$\Delta \left( \lambda_{c,k}^{\frac{1}{2}} \right)_b \approx \frac{1}{2\lambda_{c,k}^{\frac{1}{2}}} \mathbf{u}_{c,k}^{\text{T}} \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{i|2i-1,b}^{\text{T}} \right) \mathbf{u}_{c,k} \quad (96)$$

and for corresponding eigenvectors  $\mathbf{u}_{c,k}$

$$\Delta \mathbf{u}_{c,k,b} = \left( \mathbf{I} - \frac{\mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}}}{\lambda_{c,k}} \right)^{\dagger} \left( \Delta \left( \mathbf{Y}_{i|2i-1,b} \mathbf{Y}_{i|2i-1,b}^{\text{T}} \right) \frac{\mathbf{u}_{c,k}}{\lambda_{c,k}} - \frac{\mathbf{Y}_{i|2i-1} \mathbf{Y}_{i|2i-1}^{\text{T}} \mathbf{u}_{c,k}}{\lambda_{c,k}^2} \Delta \lambda_{c,k} \right). \quad (97)$$

Finally, (58) results in

$$\Delta \mathbf{W}_{1,b} \approx -\mathbf{W}_1 \Delta \left( \mathbf{W}_1^{-1} \right)_b \mathbf{W}_1. \quad (98)$$

### 5.3. Algorithmic summary

The procedure for estimating the covariance of the identified system matrices  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  can be summarized as follows:

1. For the chosen system order  $n$  and half number of block rows  $i$ , estimate the system matrices  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  as detailed in Section 3. Keep the matrices  $\hat{\mathbf{O}}_i$ ,  $\mathbf{U}$ ,  $\mathbf{S}$  and  $\mathbf{W}_1$  which appear in the intermediate steps (25-26).
2. Compute the sample output correlation matrix  $\mathbf{Y}_{0|2i} \mathbf{Y}_{0|i}^{\text{ref}\text{T}}$ . All sample correlation matrices that are needed for the variance calculation can be straightforwardly obtained as sub-matrices by appropriate block indexing.
3. Compute the matrices  $\mathbf{Y}_{i|2i-1} \mathbf{Y}_{0|i-1}^{\text{ref}\dagger}$  and  $\mathbf{Y}_{i+1|2i} \mathbf{Y}_{0|i}^{\text{ref}\dagger}$ . They can be straightforwardly obtained from the LQ-factorization (18).
4. Compute  $\mathbf{J}_1$  from (37),  $\mathbf{J}_p$  from (69) and  $\mathbf{J}_{131}$  from (84).
5. For each data block  $b$ 
  - (a) Compute the sample output correlation matrix  $\mathbf{Y}_{0|2i,b} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}}$  by using the data of block  $b$  only. Estimate  $\Delta \mathbf{Y}_{0|2i,b} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}}$  as  $\Delta \mathbf{Y}_{0|2i,b} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}} \approx \mathbf{Y}_{0|2i,b} \mathbf{Y}_{0|i,b}^{\text{ref}\text{T}} - \mathbf{Y}_{0|2i} \mathbf{Y}_{0|i}^{\text{ref}\text{T}}$ . All sample output correlation error estimates that are needed for the variance calculation can be straightforwardly obtained as sub-matrices by appropriate block indexing.
  - (b) Estimate the first-order error on the observability matrix,  $\Delta \hat{\mathbf{O}}_{i,b}$ , as detailed in Section 5.2.2. In case the CVA algorithm is employed, estimate the error  $\Delta \mathbf{W}_{1,b}$  first as detailed in Section 5.2.5.
  - (c) Estimate the first-order error on the system matrices,  $\Delta \hat{\mathbf{C}}_b$  and  $\Delta \hat{\mathbf{A}}_b$ , as detailed in sections 5.2.3 and 5.2.4, respectively.
6. Using the estimates  $\Delta \hat{\mathbf{C}}_b$  and  $\Delta \hat{\mathbf{A}}_b$  for all blocks, apply (56) for computing the covariance of the system matrix estimates  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$ .

## 6. Operational modal analysis

In this section, the link between the identified state-space matrices and the modal characteristics (natural frequencies, damping ratios and mode shapes) of the system is highlighted, such that the relevance of the above sections on SSI-data and its uncertainty quantification for operational modal analysis is clarified.

The prime goal in operational modal analysis is the identification of the system matrices  $\mathbf{A}$  and  $\mathbf{C}$  from the observed output sequence  $\mathbf{y}_k$ , because the knowledge of  $\mathbf{A}$  and  $\mathbf{C}$  is sufficient for estimating the modal characteristics. Indeed, an eigenvalue analysis of the matrix  $\mathbf{A}$ ,

$$\mathbf{A}\boldsymbol{\psi}_k = \lambda_k\boldsymbol{\psi}_k, \quad k = 1, \dots, n, \quad (99)$$

where  $\lambda_k$  denotes an eigenvalue and  $\boldsymbol{\psi}_k$  the corresponding eigenvector, yields the continuous-time eigenvalues  $\lambda_{ck}$ , eigenfrequencies  $f_k$ , and damping ratios  $\xi_k$  (in % of critical) of the considered system [4]:

$$\lambda_{ck} = f_s \ln \lambda_k, \quad f_k = \frac{|\lambda_{ck}|}{2\pi}, \quad \text{and} \quad \xi_k = -100 \frac{\mathcal{R}(\lambda_{ck})}{|\lambda_{ck}|}, \quad (100)$$

where  $\mathcal{R}(\cdot)$  denotes the real part of a variable, and  $f_s$  denotes the sampling frequency. Combining the eigenvectors of  $\mathbf{A}$  with  $\mathbf{C}$  yields the mode shapes  $\boldsymbol{\phi}_k$  in terms of the measured quantities:

$$\boldsymbol{\phi}_k = \mathbf{C}\boldsymbol{\psi}_k. \quad (101)$$

The system matrices  $\mathbf{A}$  and  $\mathbf{C}$  are not exactly known, but they are identified from output-only data using the SSI-data method, which results in the estimates  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$ . For each mode  $k$ , a first-order sensitivity analysis of the above relations between the system matrices and the modal characteristics can be performed, leading to the related Jacobian:

$$\Delta \begin{bmatrix} f_k \\ \xi_k \\ \mathcal{R}(\boldsymbol{\phi}_k) \\ \mathcal{I}(\boldsymbol{\phi}_k) \end{bmatrix} \approx \mathbf{J}_k \begin{bmatrix} \text{vec}(\Delta\mathbf{A}) \\ \text{vec}(\Delta\mathbf{C}) \end{bmatrix}, \quad (102)$$

where  $\mathcal{R}(\cdot)$  denotes, as before, the real part of a variable, vector or matrix, and  $\mathcal{I}(\cdot)$  denotes the imaginary part. Expressions for  $\mathbf{J}_k$  are available from the literature; see, e.g., [13] or [16, App. A]. A first-order Taylor approximation of the covariance matrix between the identified characteristics of modes  $k$  and  $l$  can then be computed from

$$\text{Cov} \left( \begin{bmatrix} \hat{f}_k \\ \hat{\xi}_k \\ \mathcal{R}(\hat{\boldsymbol{\phi}}_k) \\ \mathcal{I}(\hat{\boldsymbol{\phi}}_k) \end{bmatrix}, \begin{bmatrix} \hat{f}_l \\ \hat{\xi}_l \\ \mathcal{R}(\hat{\boldsymbol{\phi}}_l) \\ \mathcal{I}(\hat{\boldsymbol{\phi}}_l) \end{bmatrix} \right) \approx \mathbf{J}_k \text{Cov} \left( \begin{bmatrix} \text{vec}(\hat{\mathbf{A}}) \\ \text{vec}(\hat{\mathbf{C}}) \end{bmatrix} \right) \mathbf{J}_l^T. \quad (103)$$

For large systems, the covariance matrix of  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  can have very large dimensions. The related memory issues can be circumvented by application of (102) to each block estimate of  $\Delta\mathbf{A}_b$  and  $\Delta\mathbf{C}_b$ , and by computing the sample covariance matrix of the modal characteristics directly from the related block values. This leads to exactly the same result as the direct application of (103).

## 7. Numerical simulations

### 7.1. Verification of the proposed variance estimation approach

The performance of the proposed variance estimation approach is verified here, using a numerical model of the four degree of freedom shear frame that is pictured in Fig. 1. A simple mass-spring-damper model is used to represent the dynamic behavior, where the mass  $m_j$  assigned to each floor ( $j = 1, \dots, 4$ ) equals 2 kg and the stiffness assigned to the columns  $k_j$  equals 2500 N/m. The overall equations of the system are then, in continuous time,

$$\mathbf{M} \frac{d^2 \mathbf{y}_d(t)}{dt^2} + \mathbf{C}_d \frac{d\mathbf{y}_d}{dt} + \mathbf{K} \mathbf{y}_d(t) = \mathbf{u}(t), \quad (104)$$

where  $\mathbf{u}$  is the vector containing the horizontal forces applied at each floor, and  $\mathbf{y}_d$  is the vector collecting the horizontal floor displacements. The model is first converted into continuous-time state-space form, with the horizontal force applied at each of the 4 floors as inputs, and the horizontal acceleration of each floor as outputs, such that  $\mathbf{y}(t) = \frac{d^2 \mathbf{y}_d(t)}{dt^2}$ . Subsequently, the continuous-time state-space form is converted into a discrete-time state-space form (1-2), using a zero-order hold assumption on the input force and a sampling frequency of 50 Hz. The conversion procedures are described in detail in [4, Sec. 2.2.4-2.2.5].

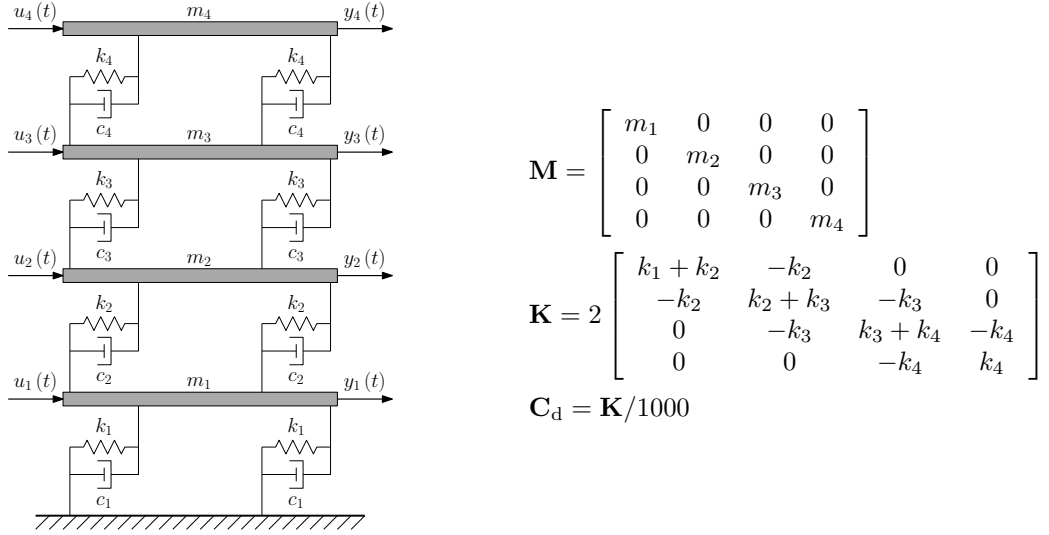


Figure 1: Four-story shear building model (left) and its mass, damping, and stiffness matrices (right). Input forces are denoted as  $u(t)$ , output accelerations as  $y(t)$ .

The simulated inputs are uncorrelated discrete white noise force signals, applied at each of the floors. The force samples are drawn from a zero-mean normal distribution with a standard deviation of 1 N. Measurement noise, as represented by  $\mathbf{e}_k^m$  in Eq. (2), is also added to each output channel. The measurement noise sequences are uncorrelated discrete white noise signals, the samples of which are drawn from a zero-mean normal distribution with a standard deviation of  $0.05 \text{ m/s}^2$ . This results in an output signal-to-noise ratio of approximately 30 dB. With the simulated force-acceleration data, the entries of the system matrices  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{C}}$  and their covariance matrix are computed with the SSI-data/ref algorithm. Subsequently, the identified modal characteristics and their (co)variances are computed. This process is repeated in a Monte Carlo analysis with 1000 runs, where in each run, a new realization of the force and measurement noise sequences are employed. The identified modal characteristics and their covariances are then compared with the corresponding sample statistics from the Monte Carlo simulation.

Such a comparison is displayed in Table 1, for the following fixed parameters:  $i = 20$ ,  $N = 2^{17}$ ,  $n_b = 32$ , UPC weighting, and all outputs as reference outputs. Nearly identical results, not reproduced here, have been obtained for PC and CVA weighting. The empirical mean values of the identified natural frequencies and damping ratios, obtained by sample averaging over the Monte Carlo runs, are very close to the true values. The empirical standard deviations of the natural frequencies and damping ratios are obtained as sample standard deviations over all Monte Carlo runs. In each Monte Carlo run, the standard deviation of the natural frequencies and damping ratios are also estimated using the method that was proposed in Sec. 5, and the mean of these estimates together with the 95% confidence interval, obtained by computing the sample mean and sample standard deviation over all Monte Carlo runs, are also displayed. These estimated standard deviations agree very well with the empirical values. For the mode shapes, a similar comparison is performed in Fig. 2. Again, a very good agreement between the empirical sample standard deviations of the modal displacements and the estimated standard deviations is observed.

## 7.2. Influence of the block length $j_b$

Next, the influence of the block length  $j_b$  on the accuracy of the estimated standard deviations of the modal characteristics is investigated, also by means of Monte Carlo simulations with 1000 runs each. The block length is varied by

	$f_j$ [Hz]	$\mu_{\hat{f}_j}$ [Hz]	$\sigma_{\hat{f}_j}$ [Hz]	$\hat{\sigma}_{\hat{f}_j}$ [Hz]	$\xi_j$ [%]	$\mu_{\hat{\xi}_j}$ [%]	$\sigma_{\hat{\xi}_j}$ [%]	$\hat{\sigma}_{\hat{\xi}_j}$ [%]
Mode 1	2.7637	2.7637	0.0012	$0.0012 \pm 0.0003$	0.868	0.868	0.047	$0.045 \pm 0.013$
Mode 2	7.9577	7.9576	0.0041	$0.0041 \pm 0.0010$	2.500	2.497	0.051	$0.052 \pm 0.014$
Mode 3	12.1920	12.1917	0.0075	$0.0073 \pm 0.0019$	3.830	3.825	0.062	$0.060 \pm 0.015$
Mode 4	14.9557	14.9555	0.0096	$0.0099 \pm 0.0025$	4.699	4.695	0.065	$0.064 \pm 0.016$

Table 1: Eigenfrequencies and modal damping ratios obtained from repeated system identification (Monte Carlo simulation) using SSI-data with  $i = 20$ ,  $N = 2^{17}$ ,  $n_b = 32$  and UPC weighting, and comparison to the true values.  $j$ : identified mode no.,  $f_j$  and  $\xi_j$ : exact values,  $\mu_{\hat{f}_j}$  and  $\mu_{\hat{\xi}_j}$ : empirical sample mean obtained from the Monte Carlo simulation,  $\sigma_{\hat{f}_j}$  and  $\sigma_{\hat{\xi}_j}$ : empirical sample standard deviation obtained from the Monte Carlo simulation, and  $\hat{\sigma}_{\hat{f}_j}$  and  $\hat{\sigma}_{\hat{\xi}_j}$ : mean of the estimated standard deviation with 95% ( $\pm 2\sigma$ ) confidence interval.

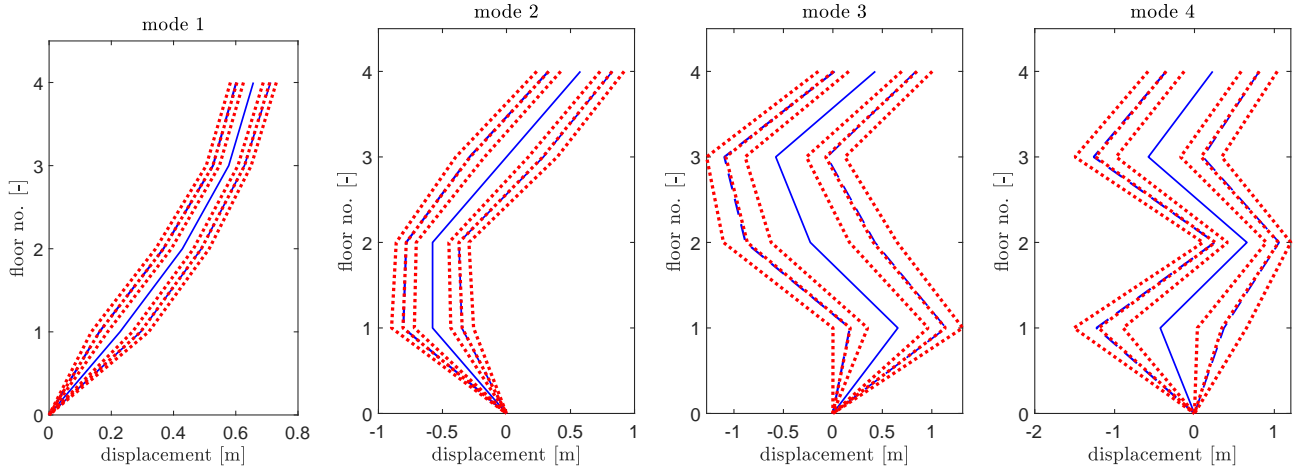


Figure 2: Identified mode shapes of the shear frame. The empirical sample mean and sample standard deviation ( $\times 300$ ) of the estimated mode shapes as computed from the Monte Carlo ensemble are displayed with solid and dashed lines, respectively. The mean of the estimated standard deviations ( $\times 300$ ) and their 95% confidence intervals are displayed with dotted lines.

keeping the number of blocks  $n_b$  constant and changing the total number of samples  $N$ . The results are displayed in Fig. 3 for  $i = 10$ ,  $n_b = 32$ , and UPC weighting. Nearly identical results, not reproduced here, have been obtained for PC and CVA weighting. The empirical coefficient of variation (i.e., the standard deviation divided by the mean value) as computed over all Monte Carlo runs is compared with the mean and the 95% confidence interval of the coefficient of variation that is estimated using the method that was proposed in Sec. 5.

A number of important observations can be made. First, the standard deviation of any quantity decreases roughly as  $1/\sqrt{N}$ , which implies that if the required accuracy need be doubled, the total measurement duration need be multiplied by four. Second, the mean of the estimated standard deviations is close to the empirical value. Third, the relative width of the confidence interval of the estimated standard deviations (i.e., the width divided by the center value) is independent of the number of samples  $N$ . This is due to the fact that  $n_b$  is kept constant in the present analysis.

### 7.3. Influence of the number of data blocks $n_b$

Finally, the influence of the number of blocks  $n_b$  on the accuracy of the estimated standard deviations of the modal characteristics is investigated, again by means of Monte Carlo simulations with 1000 runs. The results are displayed in Fig. 4 for  $i = 10$ ,  $N = 2^{13}$ , and UPC weighting. Nearly identical results, not reproduced here, have been obtained for PC and CVA weighting. Again, a number of important observations can be made. First, the empirical standard deviation is invariant to the number of data blocks  $n_b$ . This is to be expected, since  $n_b$  does not play a role in the SSI-data algorithm itself, only in the variance estimation procedure. Second, the mean of the estimated standard deviations is close to the empirical value. This indicates that the assumption that the different data blocks are statistically independent is a mild assumption, even for a large number of data blocks and consequently a relatively short block length. Third, the width of the confidence interval of the estimated standard deviations decreases with increasing number of blocks  $n_b$ . This is due to the fact that the accuracy of the sample approximation in Eq. (56) increases when the number of data blocks is increased.

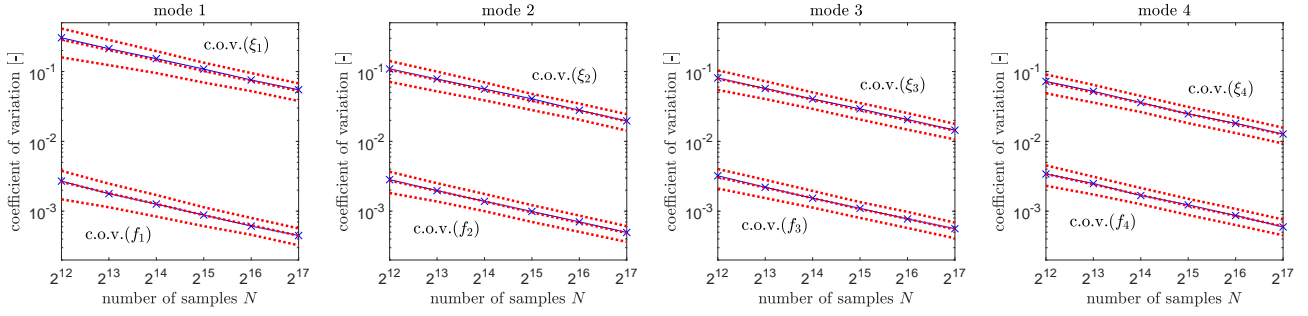


Figure 3: Coefficients of variation of the natural frequency and damping ratio estimates, obtained using SSI-data with  $i = 10$ ,  $n_b = 32$  and UPC weighting, as computed from the Monte Carlo ensemble. Solid lines: empirical values, dotted lines: mean and 95% confidence interval of the estimated values. Upper curves: damping ratio estimates, lower curves: natural frequency estimates.

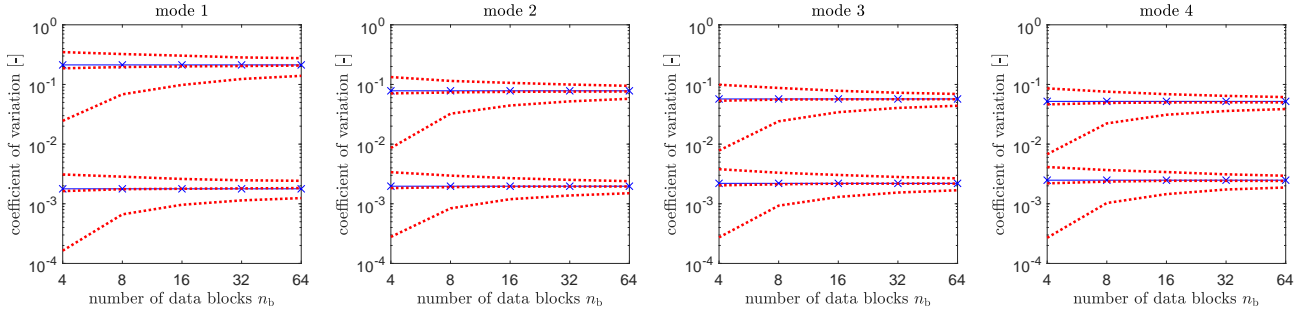


Figure 4: Coefficients of variation of the natural frequency and damping ratio estimates, obtained using SSI-data with  $i = 10$ ,  $N = 2^{13}$  and UPC weighting, as computed from the Monte Carlo ensemble. Solid lines: empirical values, dotted lines: mean and 95% confidence interval of the estimated values. Upper curves: damping ratio estimates, lower curves: natural frequency estimates.

## 8. Application: progressive damage test on a prestressed concrete beam

In vibration-based structural health monitoring, operational modal analysis is often employed for monitoring the modal characteristics of a structure, as changes in modal characteristics may indicate the presence of structural damage. However, the sensitivity of a particular modal characteristic to a particular type of damage may be relatively low, so it is important to assess whether an observed change is significant or not with respect of the identification accuracy.

In a recent study [22], a prestressed concrete beam with a width of 24 cm, a height of 20 cm and a length of 500 cm was subjected to a series of loading-unloading cycles in a four-point bending test as illustrated in Fig. 5a. This progressive damage test was performed at KU Leuven's Structural Mechanics Laboratory. After each loading-unloading cycle, the beam was lifted from its supports using inflatable tires, as illustrated in Fig. 5b. The tires are very flexible, such that they create dynamic free-free boundary conditions for all beam modes. In this lifted state, the beam was subjected to a vibration test in which 20 accelerations were measured on top of the beam, as illustrated in Fig. 5b; however, when the maximal load amplitude  $P$  during a loading-unloading cycle exceeded 55 kN, all accelerometers were removed from the beam, except the ones at positions V1 and V11. The beam was excited with a hammer in the vertical direction at location V11. The sampling frequency was 1000 Hz and the duration of each vibration test was about 30 s. The data that are employed here are used for illustrating the uncertainty quantification procedure of section 5. They are part of a larger data set collected on the same beam; the interested reader is referred to [22] for more information.

In this particular application, the changes in modal characteristics that are induced by the progressive damage test are expected to be very small. This is because cracks that may appear in the concrete during a loading-unloading cycle, are to a very large extent closed by the prestressing forces after unloading, and the modal characteristics are identified in the unloaded state. In this respect, vibration-based damage identification is more challenging for the beam than for full-sized prestressed concrete structures such as bridges, as the latter have a much higher ratio between the self weight



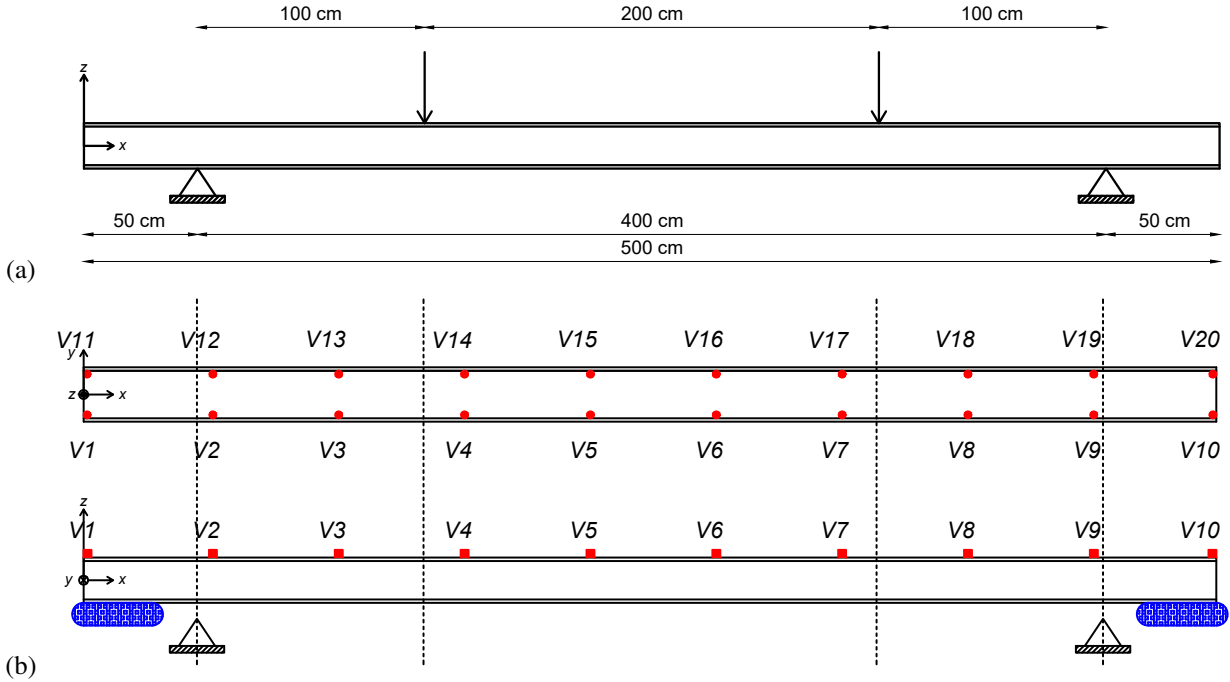


Figure 5: Overview of the two types of test performed on the prestressed concrete beam: (a) quasi-static four-point bending test with hinged supports and loads  $P$  provided by hydraulic jacks (side view), and (b) vibration test in which the beam is lifted from its supports by pneumatic tyres, creating dynamic free-free boundary conditions; the vibrations are captured with accelerometers at positions V1 through V20.

and the externally applied loading [23]. Of course, in-situ testing involves additional challenges such as the elimination of temperature influences [24, 25], which are not present in laboratory conditions.

After removing the DC offset from the vibration signals, they were processed with the SSI-data method with  $i = 30$ ,  $n_b = 16$ , CVA weighting, and all outputs as reference outputs. A stabilization diagram was constructed for all even model orders between 2 and 50 in order to remove as much as possible bias errors (which can arise, for example, from the fact that the excitation and the measurement errors are not perfectly white) from the identified system description [13]. Six modes can be identified below the Nyquist frequency of 500 Hz. Fig. 6 displays the natural frequencies and the corresponding 95% confidence intervals. The first few loading cycles induce only minor changes in the natural frequencies of most modes. This is in agreement with the fact the first cracks were only visually observed when the load reached a value of  $P \approx 27.5$  kN [22]. Nevertheless, the zooms in Fig. 6 show that even below this cracking load, statistically significant changes in the eigenfrequencies of all but the first mode can be observed. They reflect the small loss of stiffness due to micro-cracking of the concrete that remains after unloading. Due to the increasing loss of stiffness, the natural frequencies are expected to decay monotonically with the loading amplitude. This is indeed observed, except for mode 3 at 50 and 55 kN. This exception may be due to a slightly different closure of the cracks after unloading.

The identified mode shapes that have been obtained before quasi-static loading, and after the loading-unloading cycle with amplitude  $P = 55$  kN, which was the last unloading cycle for which the beam was fully instrumented, are plotted in Fig. 7. All modes are bending modes, except for mode 4 which is a torsional mode. In order to have a better visualization of the identification accuracy, the  $\pm 25\sigma$  confidence intervals are plotted (instead of the more conventional  $\pm 2\sigma$ ). From the magnified uncertainty plot, it can be observed that there are no statistically significant changes in mode shape. This can be attributed to the closure of cracks after unloading as discussed earlier, as well as to the fact that cracks were distributed along the loading region, rather than concentrated at a single location. Other modal characteristics, such as modal macro-strains obtained from fiber optic data, are more successful in identifying this type of damage, as recently demonstrated in [22, 23, 26].

A nice feature of the presented uncertainty quantification, besides the detection of statistically significant changes, is

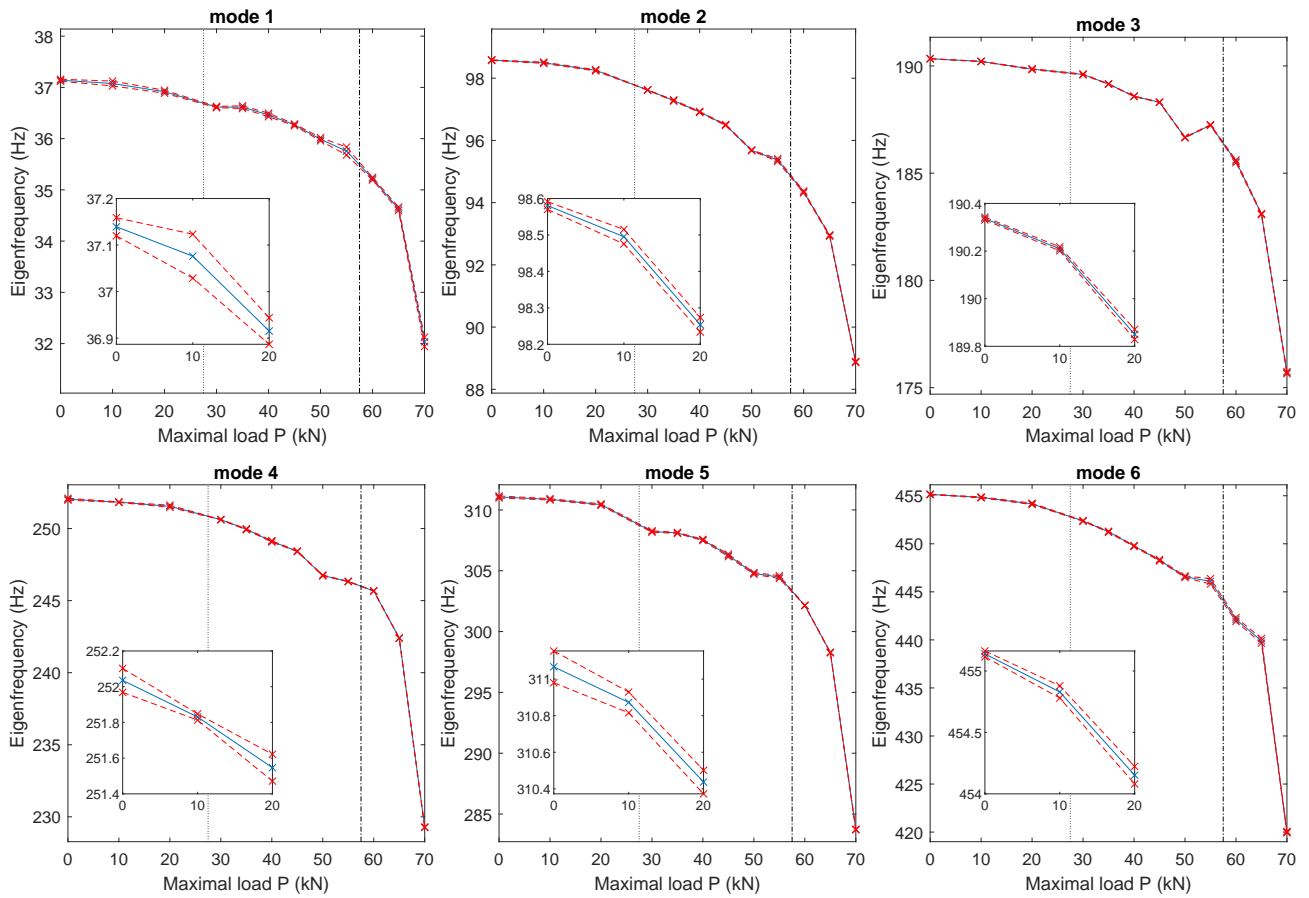


Figure 6: Identified eigenfrequencies (solid lines) and corresponding 95 % confidence intervals (dashed lines) of the prestressed concrete beam, as a function of the maximum amplitude of the load  $P$  during the preceding quasi-static loading-unloading cycle. The dotted vertical line indicates the load amplitude at which cracking was first visually observed, while the dash-dotted vertical line indicates the removal of all but two accelerometers from the beam.

that it allows for comparing the accuracy of the identified modal characteristics with respect to each other. For example, from Fig. 7, it is clear that the identification accuracy of the second and third bending mode shapes is higher than that of the others. In the same Figure, it can also be observed that all mode shapes are identified with higher accuracy in the undamaged situation than in the damaged one, except for mode 5, where the accuracy is similar in both cases. These differences in identification accuracy may be due to the fact that not all modes are always equally well excited [16].

Finally, the influence of the weighting scheme (UPC, PC or CVA) on the identification results has been investigated. In the numerical examples of section 7, in which the model order was known, there were no significant differences. However, in the present measurement example, the singular value decomposition (25) is truncated at a range of model orders when constructing the stabilization diagram, and this truncation is influenced by the weighting matrices. As a result, the stabilization diagrams were different for the three different weighting schemes, and the variance estimates were also different, indicating a different accuracy of the modal estimates. In general, the UPC and CVA weightings resulted in variance estimates of similar magnitude, but the UPC variance values were slightly lower at higher model orders. The PC weighting resulted in significantly higher variance values at higher model orders and for less well excited modes. Further research is necessary to investigate whether these observations are limited to this particular case study, or whether they are more generally valid.

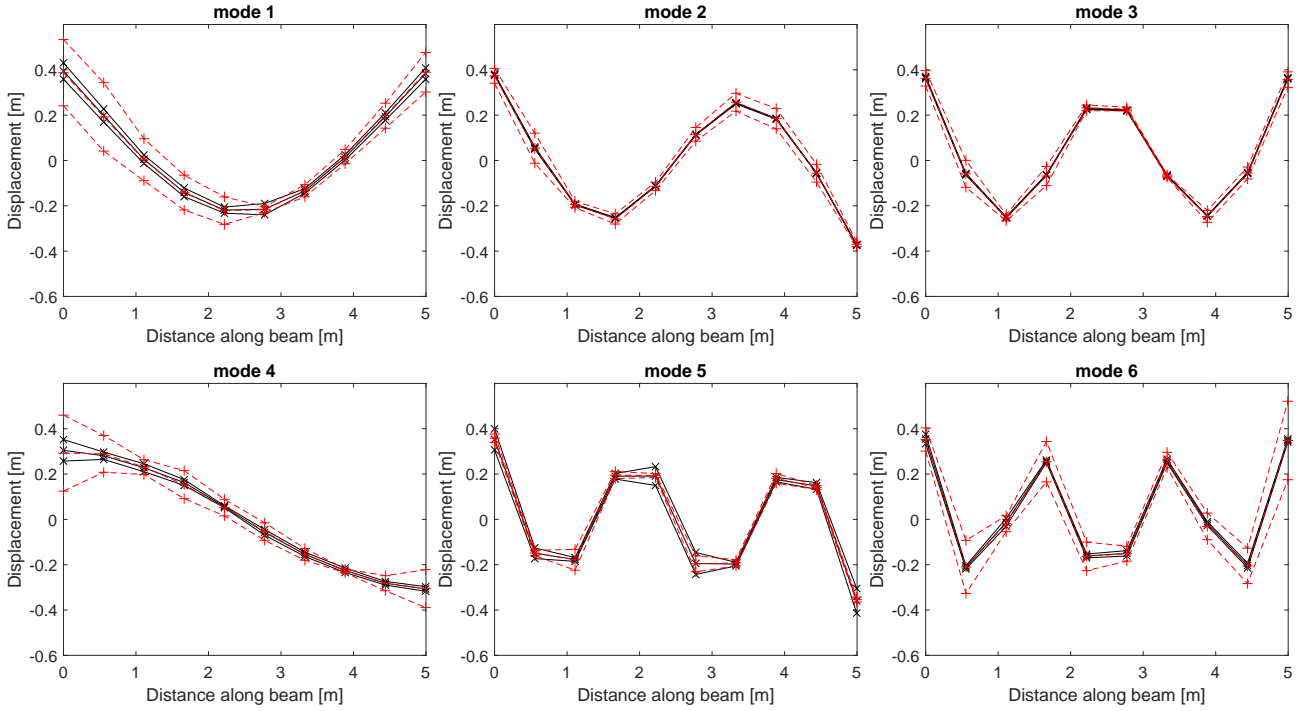


Figure 7: Identified mode shapes (for locations V1 till V10) and corresponding  $\pm 25\sigma$  confidence intervals of the prestressed concrete beam, before quasi-static loading (black solid lines) and after the loading-unloading cycle with amplitude  $P = 55$  kN (red dashed lines).

## 9. Conclusions

In this work, a method was presented for estimating the variance of the system description that is identified with data-driven stochastic subspace identification, or more precisely, the full covariance matrix between any two entries of the identified system matrices. The analysis focused on the case where the system matrices are identified from the state sequences. Three common types of weighting choices were treated. The Jacobian matrices of the linear sensitivity analysis that underlies the presented method are computed only implicitly, such that a memory efficient implementation can be achieved in a straightforward way. When compared to the SSI-data algorithm itself, the variance estimation procedure introduces an additional parameter  $n_b$ , which represents the number of blocks into which the data matrix is divided for computing the variance of the output and state correlation matrices.

An extensive numerical validation has been performed. It demonstrates that the estimated variance of the identified system description is highly accurate for a wide range of parameter values and that the uncertainty on the variance estimate decreases with an increasing number of data blocks  $n_b$ . Finally, the practical use of the method in the context of operational modal analysis and vibration-based structural health monitoring was demonstrated in an application involving experimental vibration data from a prestressed concrete beam for increasing levels of damage.

## Acknowledgments

The research presented in this paper has been performed within the framework of the projects G099014N “Identification and modeling of structural damage” and G005220N “Vibration-based post-fire assessment of concrete structures using Bayesian updating techniques”, funded by the Research Foundation Flanders (FWO), Belgium. Their financial support is gratefully acknowledged. The author also thanks his colleagues Wannes Stalmans and Dimitrios Anastopoulos for their help by creating Figs. 1 and 5, respectively.

## References

- [1] P. Van Overschee, B. De Moor, Subspace algorithms for the stochastic identification problem, in: Proceedings of the 30th IEEE Conference on Decision and Control, Brighton, UK, 1991, pp. 1321–1326.
- [2] B. Peeters, G. De Roeck, Stochastic system identification for operational modal analysis: A review, *ASME Journal of Dynamic Systems, Measurement, and Control* 123 (4) (2001) 659–667.
- [3] F. Magalhães, A. Cunha, Explaining operational modal analysis with data from an arch bridge, *Mechanical Systems and Signal Processing* 25 (5) (2011) 1431–1450.
- [4] E. Reynders, System identification methods for (operational) modal analysis: review and comparison, *Archives of Computational Methods in Engineering* 19 (1) (2012) 51–124.
- [5] P. Van Overschee, B. De Moor, Subspace identification for linear systems, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
- [6] B. Peeters, G. De Roeck, Reference-based stochastic subspace identification for output-only modal analysis, *Mechanical Systems and Signal Processing* 13 (6) (1999) 855–878.
- [7] D. Bauer, M. Deistler, W. Scherrer, Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs, *Automatica* 35 (7) (1999) 1243–1254.
- [8] M. Jansson, Asymptotic variance analysis of subspace identification methods, in: Proceedings of the 12th IFAC symposium on system identification SYSID 2000, Santa Barbara, CA, 2000, pp. 705–714.
- [9] A. Chiuso, G. Picci, The asymptotic variance of subspace estimates, *Journal of Econometrics* 118 (1–2) (2004) 257–291.
- [10] D. Bauer, Comparing the CCA subspace method to pseudo maximum likelihood methods in the case of no exogenous inputs, *Journal of Time Series Analysis* 26 (5) (2005) 631–668.
- [11] D. Bauer, M. Jansson, Analysis of the asymptotic properties of the MOESP type of subspace algorithms, *Automatica* 36 (4) (2000) 497–509.
- [12] D. Bauer, L. Ljung, Some facts about the choice of the weighting matrices in Larimore type of subspace algorithms, *Automatica* 38 (5) (2002) 763–773.
- [13] E. Reynders, R. Pintelon, G. De Roeck, Uncertainty bounds on modal parameters obtained from Stochastic Subspace Identification, *Mechanical Systems and Signal Processing* 22 (4) (2008) 948–969.
- [14] M. Döhler, L. Mevel, Efficient multi-order uncertainty computation for stochastic subspace identification, *Mechanical Systems and Signal Processing* 38 (2) (2013) 346–366.
- [15] P. Mellinger, M. Döhler, L. Mevel, Variance estimation of modal parameters from output-only and input/output subpace-based system identification, *Journal of Sound and Vibration* 379 (2016) 1–27.
- [16] E. Reynders, K. Maes, G. Lombaert, G. De Roeck, Uncertainty quantification in operational modal analysis with stochastic subspace identification: validation and applications, *Mechanical Systems and Signal Processing* 66–67 (2016) 13–30.
- [17] E. Jaynes, *Probability Theory. The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
- [18] R. Kalman, A new approach to linear filtering and prediction problems, *Journal of Basic Engineering, Transactions of the ASME* 82D (1960) 35–45.
- [19] 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) (2008) 617–637.
- [20] T. Katayama, *Subspace methods for system identification*, Springer-Verlag, London, U.K., 2005.
- [21] J. Willems, P. Rapisarda, I. Markovsky, B. De Moor, A note on persistency of excitation, *Systems and Control Letters* 54 (4) (2005) 325–329.
- [22] D. Anastasopoulos, G. De Roeck, E. Reynders, Damage identification on a prestressed concrete beam using modal strains identified from FBG data, in: K. Pitilakis (Ed.), Proceedings of the 16th European Conference on Earthquake Engineering, 16ECEE, Thessaloniki, Greece, 2018.
- [23] D. Anastasopoulos, G. De Roeck, E. Reynders, Influence of damage versus temperature on modal strains and neutral axis positions of beam-like structures, *Mechanical Systems and Signal Processing* 134 (2019) 106311.
- [24] B. Peeters, G. De Roeck, One-year monitoring of the Z24-bridge: environmental effects versus damage events, *Earthquake Engineering and Structural Dynamics* 30 (2) (2001) 149–171.
- [25] E. Reynders, G. Wursten, G. De Roeck, Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification, *Structural Health Monitoring* 13 (1) (2014) 82–93.
- [26] D. Anastasopoulos, M. De Smedt, L. Vandewalle, G. De Roeck, E. Reynders, Damage identification using modal strains identified from operational fiber-optic Bragg grating data, *Structural Health Monitoring* 17 (6) (2018) 1441–1459.