



**HAL**  
open science

# Uncertainty quantification of the Modal Assurance Criterion in Operational Modal Analysis

Szymon Gres, Michael Döhler, Laurent Mevel

► **To cite this version:**

Szymon Gres, Michael Döhler, Laurent Mevel. Uncertainty quantification of the Modal Assurance Criterion in Operational Modal Analysis. *Mechanical Systems and Signal Processing*, 2021, 152, pp.107457. 10.1016/j.ymssp.2020.107457 . hal-03276725

**HAL Id: hal-03276725**

**<https://inria.hal.science/hal-03276725v1>**

Submitted on 2 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Uncertainty quantification of the Modal Assurance Criterion in Operational Modal Analysis

Szymon Greś<sup>a,\*</sup>, Michael Döhler<sup>b</sup>, Laurent Mevel<sup>b</sup>

<sup>a</sup>*Aalborg University, Department of Civil and Structural Engineering, Thomas Manns Vej 23, 9000 Aalborg, Denmark*

<sup>b</sup>*Univ. Gustave Eiffel, Inria, COSYS-SII, I4S, Campus de Beaulieu, 35042 Rennes, France*

---

## Abstract

The Modal Assurance Criterion (MAC) is a modal indicator designed to decide whether the mode shapes used in its computation are corresponding to the same mode. During structural monitoring, it can be applied to evaluate changes in the mode shapes. When the mode shapes are estimated from measurement data, the MAC inherits their statistical properties, thus is afflicted with statistical uncertainty. The evaluation of this uncertainty is particularly relevant when the MAC estimate is close to 1, where 1 indicates equal mode shapes. In structural monitoring, it can be used to assess changes in mode shapes after early damage. While the framework for uncertainty quantification of modal parameters is well-known and developed in the context of subspace-based system identification methods, uncertainty quantification for the MAC has not been developed yet. A particular challenge for its statistical characterization is its boundedness in the interval between 0 and 1. In this paper it is shown that this boundedness yields two different distributions of the MAC estimates. The MAC computed between estimates of different mode shapes is inside the interval  $(0, 1)$ , and a Gaussian approximation of its distribution is obtained. When the MAC is computed between estimates of equal mode shapes, the resultant MAC estimate is close to 1, and the classical Gaussian approximation is inadequate. In this case it is shown that the MAC estimate is linked to a quadratic form of the mode shapes, whose distribution can be approximated by a scaled and shifted  $\chi^2$  distribution. For both cases, uncertainty bounds related to the MAC estimates are established. The proposed frameworks are validated by extensive Monte Carlo simulations and then applied to evaluate mode shape changes due to damage during monitoring of the S101 Bridge.

*Keywords:* Operational modal analysis, Modal Assurance Criterion, Ambient excitation, Uncertainty bounds, Damage detection

---

## 1. Introduction

Modal indicators describe some physical aspects of the mode shape vectors and facilitate the interpretation of modal parameters. The Modal Assurance Criterion (MAC) [1] is an important modal indicator used

---

\*Corresponding author; *E-mail address:* [sg@civil.aau.dk](mailto:sg@civil.aau.dk)

in application to e.g. model updating [2–4], matching mode shape estimates with their analytic counterparts [5–7], clustering mode shape estimates in stabilization diagrams [8], comparison of mode shape estimates between different system identification routines [9], damage detection [10, 11] and many more applications that involve the comparison of mode shape vectors. The MAC is related to the angle between two mode shape vectors and it is bounded between 0 and 1, where the boundaries 0 and 1 indicate orthogonal and collinear vectors, respectively.

In application to structural health monitoring, the MAC is often used as one of the metrics to detect changes of modal parameters in time [12]. Therein, parameter estimates from different data sets are never exactly equal even under no structural change, since they are afflicted with statistical uncertainty due to finite data length, unknown excitation and measurement noise. Quantification of these uncertainties is classical for modal parameter estimates [13–18], however has not been achieved for modal indicators, whose estimates inherit the uncertainty of the mode shapes. Therefore, even if the underlying theoretical mode shapes are equal, the MAC estimate is not exactly one but close to one. Thus, in order to evaluate if the mode shapes are equal up to statistical uncertainty, it is important to quantify the uncertainty in the MAC estimate.

The theoretical goal of this paper is to characterize the statistical distribution of the MAC estimates for their uncertainty quantification, and to infer statistically about the equality of the underlying mode shapes computed from different data sets. This statistical result is used in an engineering application to detect damage-induced changes in mode shapes of the progressively damaged S101 Bridge in Austria [17]. It is assumed that the mode shape estimates are asymptotically Gaussian distributed, which is e.g. the case when using subspace identification [16, 19] or maximum likelihood methods [13], the distribution of the corresponding MAC estimate can be derived. Since the MAC can be viewed as the correlation coefficient between two (asymptotically) Gaussian vectors, one way to infer about the distribution is the analysis of the inner product of two complex-valued unit Gaussian vectors, whose distribution properties were given considerable attention in the past [20]. For example, the distribution of the product of centered Gaussian variables can be generally expressed as an infinite sum of Bessel functions of the second kind [21]. The modified Bessel functions are also used to characterize the exact distribution for the product of two correlated Gaussian vectors of arbitrary variances in [22]. However the resultant probability density function (PDF) cannot be directly computed from data, since it requires the knowledge of the *exact* mean values of both vectors, while in practice only the mode shape *estimates* are available. A characteristic function of the inner product for two independent complex Gaussian vectors whose covariance is a multiple of the identity is derived in [23]. However, an analytic formula for the PDF requires a numerical inversion of this function, and moreover the exact mean values of both Gaussian vectors are required that are not available in practice.

Instead of analyzing the exact distribution properties of the inner product of Gaussian vectors, the distribution of the MAC estimate is analyzed *asymptotically* in this paper, which is coherent with the

aforementioned uncertainty quantification of the (asymptotically Gaussian) modal parameters. The resulting approximation of the MAC distribution is consistent and easy to evaluate on experimental data in practice. For its derivation, two cases are distinguished. The first case considers that the MAC is computed between estimates of different mode shapes, which yields the MAC estimate to be located in the interior of the interval  $(0, 1)$ . Consequently, the distribution of the corresponding MAC estimate is shown to be asymptotically Gaussian. The second case considers that the MAC is computed between estimates of equal mode shapes, which results in the MAC estimate to be approximately 1. As a consequence, the distribution of this MAC estimate will accumulate near 1 and cannot be approximated by a Gaussian distribution. This is in contrast to classical cases, where statistical uncertainty is usually described by symmetric confidence intervals in terms of the standard deviation, which (implicitly) assumes Gaussian distributions when dealing with it. When the distribution differs from a Gaussian, the standard deviation alone is insufficient to characterize the uncertainty. For its analysis, the distribution type and its properties must be known.

The main theoretical contribution of this paper is the characterization of the distribution of the MAC estimate for equal mode shapes by means of the second-order delta method [24], which can be easily computed in practice. Consequently, this allows to establish an uncertainty quantification framework in terms of the second-order Taylor expansion of the MAC, which is linked to a quadratic form of the mode shapes. Its distribution is non-trivial, but can be approximated by a scaled and shifted  $\chi^2$  distribution [25], whose quantile is evaluated for uncertainty quantification. A data-based decision framework is developed based on the derived uncertainties, to evaluate if mode shapes are equal or not. This decision framework is then embedded in a modal parameter-based damage detection algorithm.

This paper is organized as follows. Some background on the MAC indicator and the considered problem is stated in Section 2. The asymptotically Gaussian distribution of the MAC estimate is derived for the case when the mode shapes estimates relate to different mode shapes in Section 3. For the case of MAC computed between estimates of equal mode shapes, the probability distribution is developed in Section 4, leading to an approximation by a scaled and shifted  $\chi^2$  distribution. A data-based decision flow chart between both frameworks is proposed in Section 5 based on the MAC estimate and the related uncertainties. The resulting damage detection scheme is applied to real monitoring data from the S101 Bridge in Section 6.

## 2. Background on MAC indicator

In this section, the vibration model is recalled and the statistical properties of the MAC indicator are introduced. The considered MAC is computed between mode shapes that are both estimates from the structural response data. It is assumed that the underlying method for modal analysis yields asymptotically Gaussian mode shape estimates, which is e.g. the case for subspace identification methods [19]. No further modeling assumptions are made, e.g., towards the type of system damping, which is elaborated in this

section.

### 2.1. Mode shape properties

The dynamic behavior of a linear time-invariant (LTI) structural system with  $d$  degrees of freedom (DOF) is described by the differential equation of motion

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = u(t) \quad (1)$$

where  $t$  denotes continuous time, and matrices  $M, C, K \in \mathbb{R}^{d \times d}$  denote mass, damping and stiffness matrices, respectively. Vectors  $q(t) \in \mathbb{R}^d$  and  $u(t) \in \mathbb{R}^d$  denote the continuous-time displacements and the unknown external forces, respectively.

Assume that system (1) is observed by e.g. acceleration sensors at  $r$  DOF of the structure. A “true” mode shape  $\varphi_* \in \mathbb{C}^r$  of the structure is equal to an eigenvector of system (1) at the observed  $r$  DOF. In output-only system identification, the estimate  $\hat{\varphi}$  of mode shape  $\varphi_*$  is obtained from vibration measurements at discrete time instants  $t_k = k\tau$ ,  $k = 1, \dots, N$ , where  $\tau$  is the time step. Note that the mode shapes of the system are considered complex-valued, as no restrictions on the type of damping are made. In particular, system (1) can be non-proportionally damped, and the strategies to characterize the statistical distribution of the MAC developed in this paper do not require any assumptions on the damping. To characterize the underlying distribution of the mode shapes, it is assumed that system (1) has no repeated eigenvalues [26].

Hereafter assume that the system identification method deployed to estimate the modal parameters yields consistent mode shape estimates, i.e. the estimate converges to the true value  $\hat{\varphi} \rightarrow \varphi_*$  when the data length  $N$  goes to infinity. Moreover, assume that real and imaginary parts of the estimated mode shapes are asymptotically Gaussian, satisfying

$$\begin{bmatrix} \Re(\hat{\varphi}) \\ \Im(\hat{\varphi}) \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \Re(\varphi_*) \\ \Im(\varphi_*) \end{bmatrix}, \frac{1}{N} \Sigma_{\varphi_*} \right),$$

where  $\Re(\cdot)$  and  $\Im(\cdot)$  express the real and imaginary parts of a complex variable,  $\mathcal{N}(\mu, \Sigma)$  denotes a Gaussian distributed variable with mean  $\mu$  and covariance  $\Sigma$ , and  $\Sigma_{\varphi_*} \in \mathbb{R}^{2r \times 2r}$  is the asymptotic mode shape covariance. Multiple system identification methods satisfy the aforementioned criteria, in particular stochastic subspace identification (SSI) methods [19, 27]. The computation of a consistent estimate  $\hat{\Sigma}_{\varphi_*}$  of the mode shape covariance with these methods can be found e.g. in [14–16], and case studies confirming the Gaussian characteristics of modal parameter estimates can be found e.g. in [28].

### 2.2. MAC computation and uncertainties

The MAC between two complex-valued mode shape vectors  $\varphi$  and  $\psi$  is related to the angle between both vectors and writes [1]

$$\text{MAC}(\varphi, \psi) = \frac{|\varphi^H \psi|^2}{\varphi^H \varphi \psi^H \psi} = \frac{\varphi^H \psi \psi^H \varphi}{\varphi^H \varphi \psi^H \psi}. \quad (2)$$

The MAC is bounded between zero and one. An important feature in the design of the MAC indicator is that if and only if both mode shapes are identical, i.e.  $\varphi = k\psi$  for some scalar  $k \in \mathbb{C}$  since mode shapes are defined up to a constant, then  $\text{MAC}(\varphi, \psi) = 1$ . Different mode shapes yield  $\text{MAC}(\varphi, \psi) \neq 1$ , and the lower boundary of  $\text{MAC}(\varphi, \psi) = 0$  is reached for the (in practice unlikely) case of mode shapes with  $\varphi^H \psi = 0$ .

Let  $\hat{\varphi}$  and  $\hat{\psi}$  be two mode shape estimates computed on data of length  $N$ , with their true (unknown) values being  $\varphi_*$  and  $\psi_*$ , respectively. An estimate of the MAC can then be obtained by computing it on the mode shape estimates, namely as  $\text{MAC}(\hat{\varphi}, \hat{\psi})$ . In particular, when both mode shapes are different estimates of the same theoretical mode shape (e.g. from different data sets), then  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  is not exactly one due to estimation errors, but  $\text{MAC}(\hat{\varphi}, \hat{\psi}) \approx 1$  while  $\text{MAC}(\varphi_*, \psi_*) = 1$ .

Hence, the MAC estimate is subject to uncertainties that are originating from the mode shape estimates  $\hat{\varphi}$  and  $\hat{\psi}$ . In order to evaluate if a MAC estimate actually results from equal mode shapes and thus is close enough to one in a statistical sense, it is important to know its statistical uncertainties. However, the fact that the MAC indicator is bounded between 0 and 1 makes its uncertainty assessment difficult, especially when MAC estimates are close to the boundaries of this interval.

### 2.3. Considered context

The mode shape estimates  $\hat{\varphi}$  and  $\hat{\psi}$  for the MAC computation depend on the application of interest. For example, a classical application of the MAC indicator is the comparison of mode shapes estimated from different data sets, used e.g. for modal tracking and structural health monitoring. Therein, the MAC can be used as a criterion for mode comparison. It can also be used to detect if there is a change in the mode shape of a monitored mode, where  $\hat{\varphi}$  and  $\hat{\psi}$  are mode shape estimates corresponding to the same mode before and after a possible structural change. The computation of  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  is considered in this context, where  $\hat{\varphi}$  and  $\hat{\psi}$  are mode shape estimates from two different data sets. The following cases are distinguished:

1.  $\hat{\varphi}$  and  $\hat{\psi}$  are estimates of two different mode shapes, i.e.  $\text{MAC}(\varphi_*, \psi_*) \neq 1$ ,
2.  $\hat{\varphi}$  and  $\hat{\psi}$  are estimates of equal mode shapes, i.e.  $\text{MAC}(\varphi_*, \psi_*) = 1$ .

### 2.4. Distribution of MAC

To appreciate the statistical distribution of a random variable, a histogram based on multiple estimates of the variable can give a first insight into its empirical distribution. The MAC distributions corresponding to the cases outlined in the previous section are illustrated in the following on a numerical example. Note that the purpose of the following numerical exercise is to validate the proposed theory for uncertainty quantification of MAC. Application of this methodology for damage detection is enclosed in Section 6.

Consider a 6 DOF chain-like system as illustrated in Figure 1 that, for any consistent set of units, is modeled with spring stiffnesses  $k_1 = k_3 = k_5 = 100$  and  $k_2 = k_4 = k_6 = 200$ , mass  $m_i = 1/20$  and a

proportional damping matrix such that each mode has a damping ratio of 2%. The exact modal parameters of the system are depicted in Table 1. The system is excited by a white noise signal in all DOFs and sampled with a frequency of 50 Hz for a duration of 2000 seconds. The responses are measured only at DOFs 1 and 2, with the purpose of analyzing mode shapes that may be close but not equal due to poor spatial resolution. To emulate measurement noise, Gaussian white noise with 5% of the standard deviation of the output is added to the response for each channel.

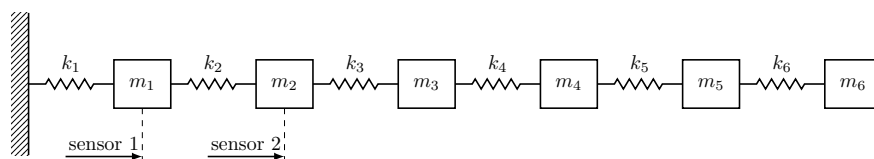


Figure 1: Illustration of 6 DOF chain system used for Monte Carlo simulation.

Table 1: Exact modal parameters of the chain system.

Natural frequency (Hz)						Damping ratio (%)					
$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$\zeta_1$	$\zeta_2$	$\zeta_3$	$\zeta_4$	$\zeta_5$	$\zeta_6$
1.936	5.618	8.682	14.494	15.798	17.007	2	2	2	2	2	2

A Monte Carlo experiment is performed, where 2000 realizations of the described signal are computed. The mode shapes of the chain system are estimated with output-only data-driven subspace-based system identification using the unweighted principal component method (SSI-UPC) [27], where the estimates are obtained at system order 12, using 15 time lags for the computation of the data Hankel matrix. Six modes of the chain system are estimated from each realization of the signal. To emulate the environment of structural monitoring where modal parameter estimates are compared between different data sets, the obtained estimates are divided into two groups of  $m = 1000$  estimates each, and the MAC between the corresponding mode shape estimates is evaluated.

An example of MAC values between mode shape estimates from two of those estimates is depicted in Figure 2. Since the MAC is approximately one for estimates of equal mode shapes, the diagonal of the MAC matrix is close to one, but not exactly one. Due to a small amount of sensors used in the simulation, the mode shapes have poor spatial resolution and the MAC between estimates of different mode shapes is high, e.g. 0.97 between mode shapes 5 and 6.

Consider now independent estimates of the MAC indicator in a structural monitoring setting, between both groups of mode shapes with each 1000 estimates in the Monte Carlo experiment. First, the MAC is computed for the case of mode shape estimates of different modes, namely between estimates of  $\varphi_5$  and  $\varphi_6$ ,

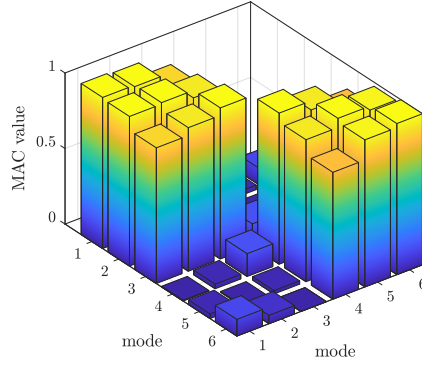


Figure 2: MAC values computed for one simulation.

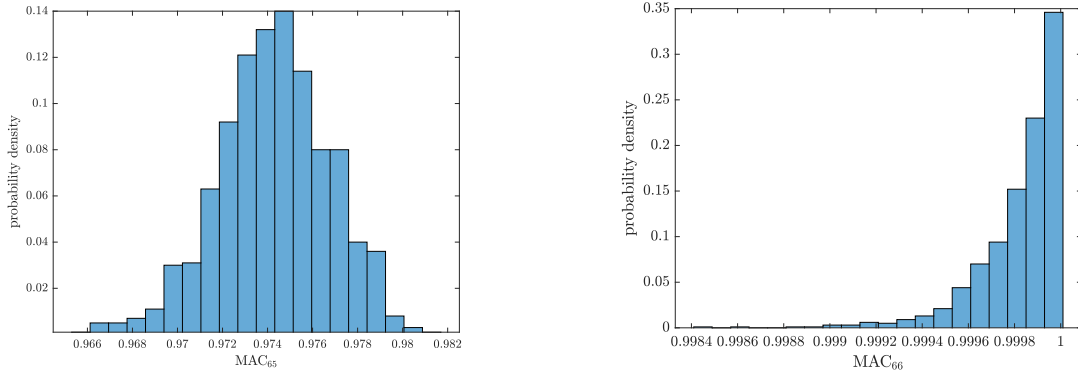


Figure 3: Histogram of MAC computed between estimates of different mode shapes (left), and between estimates of equal mode shapes (right). In both cases, the mode shapes are estimated from different data sets.

where the former is computed from the first group of 1000 estimates, and the latter from the second group. Second, the MAC is computed for the case of estimates of equal mode shapes, namely for  $\varphi_6$  estimated from both groups of estimates. Their Monte Carlo histograms are shown in Figure 3, illustrating the empirical distributions of the MAC. One can clearly see that both distributions illustrate two distinctively different cases. The distribution appears to be Gaussian in the case when the MAC is computed between estimates of different mode shapes, and it tends to a skewed distribution approaching the boundary of the MAC interval at one when the MAC is computed between estimates of equal mode shapes from different data sets.

These properties indicate some practical and theoretical problems, and highlight an application of MAC uncertainty for structural monitoring. Firstly, the MAC distribution cannot be approximated by the same kind of distribution in both cases, and it is not even evident which one to choose from simply looking at the data. Secondly, for uncertainty computation, this choice can only be made based on one data set in practice, without the support of any distribution histogram. Thirdly, this choice and the evaluated MAC uncertainty are helpful to detect whether a mode shape has changed or not after a possible damage.



These problems are addressed in the remainder of this paper, for which the empirical distributions from Figure 3 are base-case examples.

### 2.5. Summary of the proposed methods

The purpose of this paper is twofold. First, is to examine the statistical properties of the MAC between two independent mode shapes estimated from different data sets. Second, is to use those properties for deciding about changes due to damage in the components of a mode shape, which can be applied for structural monitoring as illustrated in Section 6. Two cases corresponding to the different distributions of MAC estimates are distinguished as illustrated in Figure 3:

- The MAC is computed between estimates of different mode shapes. For this case it is shown that its distribution can be approximated by a Gaussian in Section 3, and the procedure is summarized in Algorithm 1.
- The MAC is computed between estimates of equal mode shapes. It is shown that its distribution can be approximated by a scaled and shifted  $\chi^2$  distribution in Section 4, and the procedure is summarized in Algorithm 2.

Based on the resulting uncertainty computation, a decision is inferred about the theoretical MAC value in Section 5.

## 3. MAC computed from estimates of different mode shapes

In this section a classical Gaussian approximation of the MAC distribution is developed, for the case when its theoretical value is neither 0 nor 1, but inside of the interval  $(0, 1)$ . This corresponds to the case where estimates of two different mode shapes are compared, see Figure 3 (left). Based on the Gaussian approximation, the covariance of the MAC estimate and the corresponding confidence intervals are evaluated.

### 3.1. Gaussian approximation of the MAC indicator

The MAC is computed between mode shape estimates  $\hat{\varphi}$  and  $\hat{\psi}$ . Their theoretical values are  $\varphi_*$  and  $\psi_*$  and asymptotic covariances  $\Sigma_{\varphi_*}$  and  $\Sigma_{\psi_*}$ . Since the mode shapes are obtained from different data sets, they are statistically independent. To obtain the distribution properties of the MAC, which is a function of these Gaussian variables, the first-order delta method is used. The delta method is a statistical tool that allows to characterize the probability distribution of a function of a Gaussian vector as also Gaussian [24]. Moreover, it provides an expression for the covariance of this function, thus allowing to estimate the covariance of the MAC and to derive confidence intervals.

The first-order delta method is based on linearization through the first-order Taylor expansion of the MAC estimate

$$\text{MAC}(\hat{\varphi}, \hat{\psi}) \approx \text{MAC}(\varphi_*, \psi_*) + \mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} \hat{X}, \quad (3)$$

where  $\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}}$  is the derivative of the MAC with respect to the real and imaginary parts of the mode shapes

$$\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} = \begin{bmatrix} \frac{\partial \text{MAC}}{\partial \Re(\varphi)}(\varphi_*, \psi_*) & \frac{\partial \text{MAC}}{\partial \Im(\varphi)}(\varphi_*, \psi_*) & \frac{\partial \text{MAC}}{\partial \Re(\psi)}(\varphi_*, \psi_*) & \frac{\partial \text{MAC}}{\partial \Im(\psi)}(\varphi_*, \psi_*) \end{bmatrix}, \quad (4)$$

with the partial derivatives of  $\text{MAC}(\varphi, \psi)$  based on (2) for any  $\varphi$  and  $\psi$  being

$$\frac{\partial \text{MAC}}{\partial \Re(\varphi)}(\varphi, \psi) = \frac{2\Re(\psi\psi^H\varphi)^T}{\varphi^H\varphi\psi^H\psi} - \frac{2\Re(\varphi)^T\text{MAC}(\varphi, \psi)}{\varphi^H\varphi}, \quad (5)$$

$$\frac{\partial \text{MAC}}{\partial \Im(\varphi)}(\varphi, \psi) = \frac{2\Im(\psi\psi^H\varphi)^T}{\varphi^H\varphi\psi^H\psi} - \frac{2\Im(\varphi)^T\text{MAC}(\varphi, \psi)}{\varphi^H\varphi}, \quad (6)$$

$$\frac{\partial \text{MAC}}{\partial \Re(\psi)}(\varphi, \psi) = \frac{2\Re(\varphi\varphi^H\psi)^T}{\varphi^H\varphi\psi^H\psi} - \frac{2\Re(\psi)^T\text{MAC}(\varphi, \psi)}{\psi^H\psi}, \quad (7)$$

$$\frac{\partial \text{MAC}}{\partial \Im(\psi)}(\varphi, \psi) = \frac{2\Im(\varphi\varphi^H\psi)^T}{\varphi^H\varphi\psi^H\psi} - \frac{2\Im(\psi)^T\text{MAC}(\varphi, \psi)}{\psi^H\psi}, \quad (8)$$

and

$$\hat{X} = \begin{bmatrix} \Re(\hat{\varphi} - \varphi_*) \\ \Im(\hat{\varphi} - \varphi_*) \\ \Re(\hat{\psi} - \psi_*) \\ \Im(\hat{\psi} - \psi_*) \end{bmatrix}. \quad (9)$$

The linearization in (3) is only meaningful if the derivative  $\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}}$  is non-zero. A necessary and sufficient condition for  $\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} \neq 0$  is

$$\text{MAC}(\varphi_*, \psi_*) \notin \{0, 1\} \iff \mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} \neq 0, \quad (10)$$

which is shown in Appendix A. Hence, the derivative is always non-zero in the considered case, where the MAC is computed from estimates of different mode shapes. Thus, the aforementioned linearization can be applied, and the first-order delta method states that the distribution of the MAC estimate is asymptotically Gaussian with

$$\sqrt{N} \left( \text{MAC}(\hat{\varphi}, \hat{\psi}) - \text{MAC}(\varphi_*, \psi_*) \right) \longrightarrow \mathcal{N} \left( 0, \sigma_{\text{MAC}}^2 \right), \quad (11)$$

where  $\sigma_{\text{MAC}}^2 = \mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} \Sigma_{\varphi_*, \psi_*} (\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}})^T$  with the joint asymptotic mode shape covariance

$$\Sigma_{\varphi_*, \psi_*} = \begin{bmatrix} \Sigma_{\varphi_*} & 0 \\ 0 & \Sigma_{\psi_*} \end{bmatrix}.$$

Notice that (11) shows two important results, namely first stating that the distribution of the MAC estimates is Gaussian for large  $N$  when  $\text{MAC}(\varphi_*, \psi_*) \notin \{0, 1\}$ , and second providing a practical formula for

the asymptotic variance of MAC. An estimate of the variance of  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  is then directly obtained as

$$\text{var}(\text{MAC}(\hat{\varphi}, \hat{\psi})) \approx \frac{1}{N} \hat{\sigma}_{\text{MAC}}^2 = \frac{1}{N} \mathcal{J}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}} \hat{\Sigma}_{\varphi_*, \psi_*} (\mathcal{J}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}})^T, \quad (12)$$

where  $\mathcal{J}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}}$  is an estimate of the MAC derivative that can be directly obtained by evaluating (5)–(8) with  $\hat{\varphi}$  and  $\hat{\psi}$ , and  $\hat{\Sigma}_{\varphi_*, \psi_*}$  contains the mode shape covariance estimates  $\hat{\Sigma}_{\varphi_*}$  and  $\hat{\Sigma}_{\psi_*}$  obtained from the system identification method.

### 3.2. Confidence intervals

The confidence intervals computed based on this Gaussian approximation are symmetric and centered around the computed estimate. They cover a range of plausible values that include the true value of the MAC indicator with some given confidence level  $\gamma$ , e.g.  $\gamma = 0.95$ . As such, the true value  $\text{MAC}(\varphi_*, \psi_*)$  is situated for 95% of the experiments in the interval [29]

$$\left( \text{MAC}(\hat{\varphi}, \hat{\psi}) - 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{\text{MAC}}, \text{MAC}(\hat{\varphi}, \hat{\psi}) + 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{\text{MAC}} \right). \quad (13)$$

The procedure for the Gaussian characterization of the MAC estimate is summarized in Algorithm 1.

---

**Algorithm 1:** MAC uncertainty computation for estimates of *different* mode shapes

---

**Input** : mode shape estimates  $\hat{\varphi}$ ,  $\hat{\psi}$  and the corresponding covariance estimates  $\hat{\Sigma}_{\varphi_*}$ ,  $\hat{\Sigma}_{\psi_*}$  e.g. from subspace identification [14–16, 18]

**Output:** confidence interval of the MAC estimate

- 1 compute  $\mathcal{J}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}}$  based on (4), (5)–(8) using  $\hat{\varphi}$  and  $\hat{\psi}$  ;
  - 2 compute  $\hat{\sigma}_{\text{MAC}}^2$  in (12) ;
  - 3 compute confidence interval of MAC estimate for the desired confidence level, e.g. as in (13)
- 

Note that the computed Gaussian confidence interval may overlap the theoretical boundary of MAC values at 1, since it is centered at the MAC estimate that itself may be drawn from the right tail of the MAC distribution as illustrated in Figure 4. While this is still a valid confidence interval, the question may arise if the underlying assumption of  $\text{MAC}(\varphi_*, \psi_*) \neq 1$  actually holds. The Gaussian confidence interval does not allow to validate or invalidate this assumption. In such a case, it may be helpful to analyze the confidence interval for an increasing data length  $N$ . As shown in Figure 4 (right), a larger  $N$  yields smaller confidence intervals and a better separation of the Gaussian approximation from the boundary at 1 if  $\text{MAC}(\varphi_*, \psi_*) \neq 1$ , unlike in Figure 4 (left), where  $N$  is smaller.

### 3.3. Validation of Gaussian approximation

In this section, a numerical validation of the Gaussian approximation is carried out based on the Monte Carlo experiment of the chain system from Section 2.4. For each pair  $j = 1, \dots, m$  of mode shape estimates

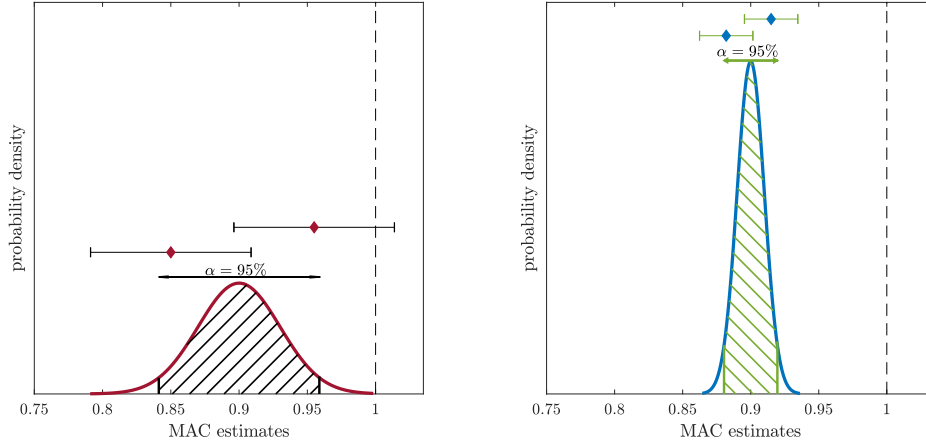


Figure 4: Illustration of confidence intervals of two cases of Gaussian approximations, each with example estimates from the left and the right tail of the distribution. Low data length (left), high data length (right).

from two data sets, where  $m = 1000$ , denote the estimated MAC value between mode shapes 6 (from first data set) and 5 (from second data set) as  $\text{MAC}_{65,j}$ . Based on these values, it is straightforward to compute their sample mean  $\mu_{\text{MC}}$  and their sample standard deviation  $s_{\text{MC}}$  of the Monte Carlo experiment, as a reference for comparison.

The standard deviation computed based on the first-order delta method requires only one data set for its computation, and it should be directly comparable with  $s_{\text{MC}}$ . Its validation is carried out as follows. Denote  $\sigma_{\text{DM},j}$  as the computed standard deviation of the  $j$ -th MAC estimate. Their mean  $\bar{\sigma}_{\text{DM}}$  should match the sample standard deviation  $s_{\text{MC}}$ . The sample standard deviation and the mean of the computed standard deviations with the developed method are shown together with their spread ( $s_{\sigma_{\text{DM}}}$ ) in Table 2 for different data lengths of the simulated data. It can be seen that  $s_{\text{MC}}$  and  $\bar{\sigma}_{\text{DM}}$  match very well, which suggests that the proposed method estimates well the standard deviation of the MAC. The spread of the estimated standard deviations is quite low with about 10%, except for the smallest data lengths. The standard deviations are decreasing with the data length, as expected.

The respective histograms of  $\text{MAC}_{65,j}$  are shown in Figure 5 together with their computed Gaussian fits from one arbitrarily selected estimate each that are centered at the true mean. It can be seen that the Monte Carlo histograms fit more and more a Gaussian distribution as the data length  $N$  increases. The estimates get closer to the true value and their variance decreases. For the selected estimates, Figure 6 shows the computed confidence intervals. In all cases, the true value is included in the interval. It can also be seen that the corresponding confidence interval may overlap the maximal possible MAC value of 1.

In Table 2, the computation of the standard deviation of the MAC with the proposed method has been validated, even for small data lengths, and in Figure 5 it has been visualized how delta method-based Gaussian fits from selected estimates match the Monte Carlo histogram. For a quantitative evaluation of

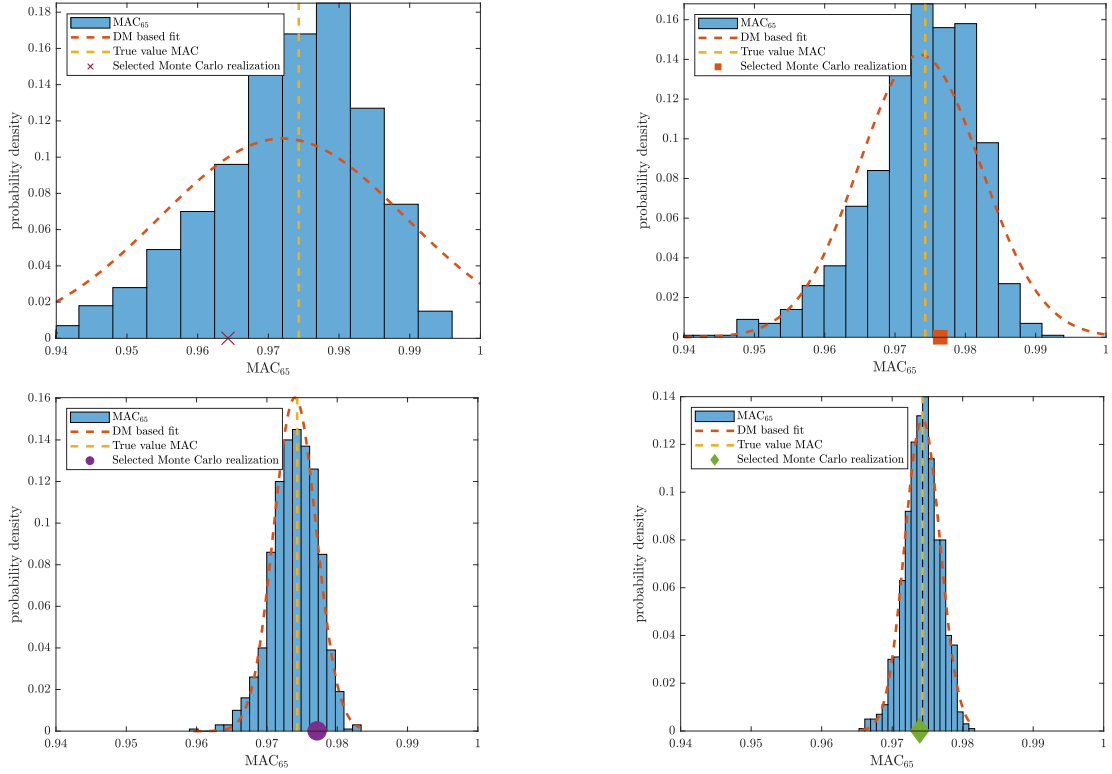


Figure 5: Monte Carlo histograms of  $MAC_{65}$  with delta method based Gaussian fit of one selected estimate, for data lengths  $N = 5,000$  (top left),  $N = 10,000$  (top right),  $N = 50,000$  (bottom left),  $N = 100,000$  (bottom right).

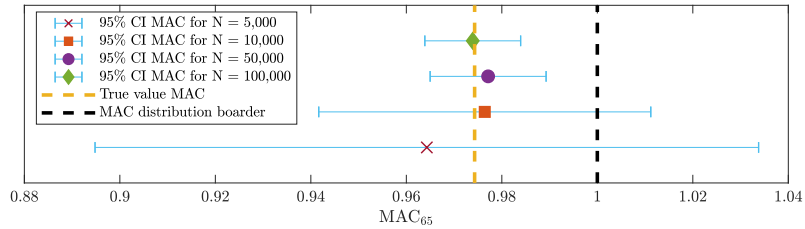


Figure 6: 95% Gaussian confidence intervals of  $MAC_{65}$  estimates for different data lengths.

the Gaussian approximation quality, the goodness of fit can be evaluated with the Pearson  $\chi^2$  test. First, the Gaussian shape of the empirical MAC distribution is evaluated for reference. The resultant goodness of fit test-statistic between the Monte Carlo histogram and its Gaussian approximation based on its sample mean  $\mu_{MC}$  and sample standard deviation  $s_{MC}$  is shown in the second column of Table 3 for the different data lengths. From there it can be seen that the Pearson goodness of fit test values for  $N = 5,000$  and  $N = 10,000$  are above the significance threshold at a 95% confidence level, indicating some differences to the fit. For  $N = 50,000$  and onwards test values are below the threshold, indicating that the Gaussian fit of the empirical MAC distribution is indeed adequate. Second, the delta method-based Gaussian approximations

Table 2: Comparison of the Monte Carlo and delta method-based standard deviations of  $\text{MAC}_{65}$ .

Data length	$s_{\text{MC}} \times 10^{-4}$	$\bar{\sigma}_{\text{DM}} \times 10^{-4}$	$s_{\sigma_{\text{DM}}} \times 10^{-4}$
$N = 5,000$	124.79	127.79	51.65
$N = 10,000$	79.91	81.26	21.17
$N = 50,000$	32.21	34.51	3.88
$N = 100,000$	24.22	23.56	2.17

Table 3: Pearson goodness of fit test-statistic between Gaussian approximations and Monte Carlo histogram for 95% significance level, threshold at 27.59.

Data length	$s_{\text{MC}}$ -based fit	average $\sigma_{\text{DM}}$ -based fit	poor $\sigma_{\text{DM}}$ -based fit
$N = 5,000$	74.74	120.08	577.76
$N = 10,000$	45.1	57.91	282.55
$N = 50,000$	18.96	24.17	74.36
$N = 100,000$	25.78	19.36	34.21

of the MAC distribution are analyzed. They are performed for each of the Monte Carlo estimates and depend on the computed standard deviations  $\sigma_{\text{DM},j}$ . To quantify the quality of the respective Gaussian approximations, two representative values are chosen from the goodness of fit test-statistics among all the Monte Carlo estimates, namely at the 50% quantile that represents an average fit, as well as at the 95% quantile that represents a poor fit. Their goodness of fit test-statistics are shown in the last two columns of Table 3 for the different data lengths. It can be observed that the test-statistic decreases when the number of samples increases, indicating better Gaussian fits. The average Gaussian  $\sigma_{\text{DM}}$ -based fits match well the Monte Carlo histogram for  $N = 50,000$  and onwards, which is coherent with the  $s_{\text{MC}}$ -based fits. For the poor Gaussian  $\sigma_{\text{DM}}$ -based fits, the test-statistic decreases strongly with the data length and is nearly at the significance threshold for  $N = 100,000$ , which indicates that the delta method-based Gaussian approximation can become adequate even for the few data sets whose estimates are in the tail of the distribution if the data length is long enough.

These results illustrate that the distribution of the MAC estimates is well approximated by a Gaussian with the delta method if the data length is long enough. The practical application of the proposed Gaussian approximation is however not to establish a distribution function, but to estimate the corresponding confidence intervals based on the Gaussian premise as given in (13). Therefore to finalize this validation section, the merit of the computed confidence intervals is studied with the Monte Carlo experiment. Based on the empirical MAC distribution, it is first verified that 95% of the Monte Carlo estimates are in the  $\pm 2s_{\text{MC}}$  interval around the true MAC value for the chosen confidence level of  $\gamma = 0.95$ . The respective ratios are shown in the second column of Table 4 for the different data lengths. Indeed, all values are very close to 95%,

Table 4: Percentage of Monte Carlo estimates in  $\pm 2s_{\text{MC}}$  confidence interval, and percentage of delta method-based  $\pm 2\sigma_{\text{DM},j}$  confidence intervals containing the true MAC value

Data length	$s_{\text{MC}}$ -based CI	$\sigma_{\text{DM},j}$ -based CI
$N = 5,000$	95.4%	95.1%
$N = 10,000$	95.0%	96.6%
$N = 50,000$	95.8%	96.5%
$N = 100,000$	95.9%	93.0%

which indicates that the Gaussian distribution assumption is adequate for the confidence interval computation. Second, the quality of the delta-method based confidence intervals is examined, which are computed on each of the Monte Carlo estimates in (13). According to the theory [29], 95% of them should contain the true MAC value. The respective ratios are shown in the last column of Table 4. All values are again very close to 95%. This shows that even if the empirical MAC distribution is not perfectly matching a Gaussian distribution for short data lengths, the computed confidence intervals based on the Gaussian premise are quite accurate for all considered data lengths, which validates the proposed uncertainty quantification.

#### 4. MAC computed between estimates of equal mode shapes

In this section, the case is analyzed where the theoretical MAC value is 1, i.e.  $\text{MAC}(\varphi_*, \psi_*) = 1$ , with the estimate  $\text{MAC}(\hat{\varphi}, \hat{\psi}) \approx 1$ . This case is represented by  $\text{MAC}_{66}$  whose empirical distribution is very close to the right boundary of the interval  $(0, 1)$  and cannot be approximated by a Gaussian distribution anymore, as illustrated in Figure 3.

In this case, the statistical analysis of the MAC estimate is carried out under the assumption of knowing the true MAC value, namely at 1. Then, this assumption can be accepted or rejected based on the value of the MAC estimate and the obtained uncertainties. This is slightly different to the Gaussian case, where the confidence interval is required to infer the true (unknown) MAC value. To achieve this goal, an appropriate approximation of the MAC distribution is derived for  $\text{MAC}(\varphi_*, \psi_*) = 1$ . It is shown successively:

1. The MAC estimate can be characterized as a quadratic form of the mode shapes based on a second-order Taylor expansion, as shown in Section 4.1.
2. The distribution of the quadratic form can be approximated by a scaled and shifted  $\chi^2$  distribution, whose parameters are estimated in Section 4.2.
3. Based on this distribution, the quantile for a desired confidence level is provided to express a plausible range for MAC estimates in Section 4.3, and thus to validate the assumption  $\text{MAC}(\varphi_*, \psi_*) = 1$ .

#### 4.1. Development of the quadratic form

When both mode shapes are independent estimates of equal mode shapes and  $\text{MAC}(\varphi_*, \psi_*) = 1$ , then the MAC derivative  $\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}}$  in (3) equals to zero, as indicated by (10). In this case, the previously used scheme for approximating the distribution and computing the covariance cannot be used, since the linearization by the first-order Taylor expansion from (3) is insufficient. Following that, a second-order Taylor expansion yields

$$\text{MAC}(\hat{\varphi}, \hat{\psi}) \approx \underbrace{\text{MAC}(\varphi_*, \psi_*)}_{=1} + \underbrace{\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}}}_{=0} \hat{X} + \frac{1}{2} \hat{X}^T \mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}} \hat{X}, \quad (14)$$

where  $\mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}} \in \mathbb{R}^{4r \times 4r}$  is the Hessian, i.e. the second derivative of  $\text{MAC}(\varphi, \psi)$  in  $\varphi_*$  and  $\psi_*$ , and  $\hat{X}$  contains the real and imaginary mode shape components as defined in (9). The computation of the Hessian is detailed in the next section. The asymptotic properties of  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  follow from (14) as

$$N(1 - \text{MAC}(\hat{\varphi}, \hat{\psi})) \approx \hat{X}_{\mathcal{N}}^T \bar{\mathbf{H}}_{\varphi_*, \psi_*}^{\text{MAC}} \hat{X}_{\mathcal{N}} = Q(\hat{X}_{\mathcal{N}}), \quad (15)$$

where  $\bar{\mathbf{H}}_{\varphi_*, \psi_*}^{\text{MAC}} = -\frac{1}{2} \mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}}$ , and  $\hat{X}_{\mathcal{N}} = \sqrt{N} \hat{X}$  converges to a Gaussian random variable whose distribution is  $\mathcal{N}(0, \Sigma_{\varphi_*, \psi_*})$  as  $N$  increases.

Equation (15) describes a quadratic form of a Gaussian variable that is directly linked to the MAC estimate. The exact distribution of such a quadratic form is difficult to characterize. It is clearly non-Gaussian, and thus cannot be defined solely by its mean and variance. However, an approximation of the distribution using a scaled and shifted  $\chi^2$  distribution [25] can be applied when the matrix  $\bar{\mathbf{H}}_{\varphi_*, \psi_*}^{\text{MAC}}$  is positive semi-definite. With this approximation, the statistical properties of the quadratic form can be related to the asymptotic covariance  $\Sigma_{\varphi_*, \psi_*}$  of the mode shape estimates, and subsequently the statistical properties of the MAC estimate can be quantified through relation (15) as described in the next section.

#### 4.2. Approximation of the quadratic form

Many studies have been devoted to the approximation of the distribution of the inner product of Gaussian vectors [30–32], however due to the required knowledge of the mean value of the underlying vectors they cannot be directly applied for the characterization of the MAC distribution. Instead, the MAC distribution will be developed based on the quadratic form (15), which requires only the estimates of the Hessian of the MAC and of the covariance of the mode shapes, but no knowledge on the mode shape means. The resulting confidence bounds can be derived using only quantities computed from the data of a single data set, under a framework unifying both Gaussian and quadratic approximations. Moreover, this framework is an extension of the methodology previously used to derive confidence intervals for modal parameters and shown to be particularly efficient in practice [14, 16].



As stated before, the quadratic form is characterized by its inner matrix  $\bar{H}_{\varphi_*, \psi_*}^{\text{MAC}}$ , which is related to the Hessian  $H_{\varphi_*, \psi_*}^{\text{MAC}}$  of the MAC. Since the MAC distribution is developed for the assumption that  $\varphi_*$  and  $\psi_*$  are corresponding to equal mode shapes, these mode shapes are collinear with

$$\varphi_* = k \cdot \psi_* \quad (16)$$

for some  $k \in \mathbb{C}$ . After a tedious derivation that is outlined in Appendix B, the Hessian writes

$$H_{\varphi_*, \psi_*}^{\text{MAC}} = \frac{2}{|k|^2 d_*^2} \begin{bmatrix} -I_{2r} & 0 \\ 0 & H_1^T \end{bmatrix} \begin{bmatrix} H_2 & H_2 \\ H_2 & H_2 \end{bmatrix} \begin{bmatrix} -I_{2r} & 0 \\ 0 & H_1 \end{bmatrix}, \quad (17)$$

where

$$H_1 = \begin{bmatrix} \Re(k)I_r & -\Im(k)I_r \\ \Im(k)I_r & \Re(k)I_r \end{bmatrix}, \quad H_2 = \begin{bmatrix} M_{xx}^* + M_{yy}^* - d_* I_r & M_{xy}^* - M_{yx}^* \\ M_{yx}^* - M_{xy}^* & M_{xx}^* + M_{yy}^* - d_* I_r \end{bmatrix}, \quad (18)$$

with  $M_{xx}^* = \Re(\psi_*)\Re(\psi_*)^T$ ,  $M_{yy}^* = \Im(\psi_*)\Im(\psi_*)^T$ ,  $M_{xy}^* = \Re(\psi_*)\Im(\psi_*)^T$ ,  $M_{yx}^* = \Im(\psi_*)\Re(\psi_*)^T$ , and  $d_* = \psi_*^H \psi_*$ . To approximate the distribution of the quadratic form (15) based on [25], the inner matrix  $\bar{H}_{\varphi_*, \psi_*}^{\text{MAC}}$  needs to be positive semidefinite. It can be easily shown that  $H_2$  is negative semidefinite, and since  $H_1$  is invertible, the Hessian  $H_{\varphi_*, \psi_*}^{\text{MAC}}$  in (17) is also negative semidefinite. Thus, the inner matrix  $\bar{H}_{\varphi_*, \psi_*}^{\text{MAC}} = -\frac{1}{2} H_{\varphi_*, \psi_*}^{\text{MAC}}$  of the quadratic form is indeed positive semidefinite, and the approximation of the distribution of the quadratic form (15) from [25] can be applied as follows. It is based on the quantities  $c_1$ ,  $c_2$  and  $c_3$  that are directly related to the cumulants of the quadratic form, defined as

$$c_j = \text{tr} \left( (\bar{H}_{\varphi_*, \psi_*}^{\text{MAC}} \Sigma_{\varphi_*, \psi_*})^j \right), \quad j = 1, 2, 3. \quad (19)$$

The distribution of the quadratic form is approximated by a scaled and shifted  $\chi^2$  distribution with probability density function (pdf)

$$f_Q(x) = \frac{1}{\alpha} f_{\chi_l^2} \left( \frac{x - \beta}{\alpha} \right), \quad x \in [\beta, +\infty), \quad (20)$$

and  $f_Q(x) = 0$  for  $x < \beta$ , whose parameters are defined as:

- The number of degrees of freedom  $l$  is given by

$$l = c_2^3 / c_3^2. \quad (21)$$

Note that  $l$  is in general not an integer. Strictly speaking,  $f_{\chi_l^2}$  is always the density function of a Gamma distribution, generalizing the  $\chi^2$  distribution family for non-integer values of  $l$ . For simplicity, this distribution is still denoted as  $\chi^2$ .

- The scaling and shift parameters  $\alpha$  and  $\beta$  are defined as

$$\alpha = \sigma_Q / \sigma_{\chi^2}, \quad \beta = \mu_Q - (\mu_{\chi^2} \sigma_Q) / \sigma_{\chi^2}, \quad (22)$$

where  $\mu_{\chi^2}$  and  $\sigma_{\chi^2}$  are the mean and standard deviation of the  $\chi^2$  distribution, and  $\mu_Q$  and  $\sigma_Q$  are the mean and standard deviation of the quadratic form, which yield

$$\mu_{\chi^2} = l, \quad \sigma_{\chi^2} = \sqrt{2l}, \quad \mu_Q = c_1, \quad \sigma_Q = \sqrt{2c_2}. \quad (23)$$

With the approximation of the distribution of the quadratic form in (20), the distribution of  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  follows immediately based on relation (15), having the pdf

$$f_{\text{MAC}}(x) = \frac{N}{\alpha} f_{\chi^2} \left( \frac{N(1-x) - \beta}{\alpha} \right), \quad x \in (-\infty, 1 - \frac{\beta}{N}]. \quad (24)$$

The MAC distribution is hence approximated by a scaled and shifted  $\chi^2$  distribution when the MAC is computed between estimates of equal mode shapes, assuming the theoretical MAC value being 1. The related uncertainty can be expressed by a quantile of the distribution for a desired confidence level, as detailed in the following section.

For the computation of the distribution parameters  $l$ ,  $\alpha$  and  $\beta$  of the  $\chi^2$  approximation, estimates of the mode shape covariance  $\Sigma_{\varphi_*, \psi_*}$  and of the Hessian  $\mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}}$  are required for (19). Mode shape covariance estimates  $\hat{\Sigma}_{\varphi_*}$  and  $\hat{\Sigma}_{\psi_*}$  are obtained from the system identification method, and an estimate  $\mathbf{H}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}}$  of the Hessian is obtained by evaluating (17) with  $\hat{\varphi}$ ,  $\hat{\psi}$  and  $\hat{k} = \hat{\psi}^H \hat{\varphi} / \hat{\psi}^H \hat{\psi}$  based on (16).

#### 4.3. Quantiles of the scaled and shifted $\chi^2$ approximation

In the case of the Gaussian approximation of the MAC distribution in Section 3, the computed confidence interval in (13) was centered around the estimate and described a plausible range for the *true MAC value* with some confidence level. The case of the scaled and shifted  $\chi^2$  approximation of the MAC distribution in this section is based on the assumption that the MAC is computed between estimates of equal mode shapes, i.e. the true value of MAC is already known as  $\text{MAC}(\varphi_*, \psi_*) = 1$ . Hence it is not useful to similarly define a confidence interval around the MAC estimate for a plausible range in which the *true MAC value* should lie, already being assumed as 1. Instead, it should be assessed if the assumption of  $\text{MAC}(\varphi_*, \psi_*) = 1$  is valid or not based on the obtained distribution properties. Indeed, a plausible range of *MAC estimates* can be obtained for a chosen confidence level from these properties. Then it is checked if the actual MAC estimate  $\text{MAC}(\hat{\varphi}, \hat{\psi})$  lies within this range or not for a decision if the assumption  $\text{MAC}(\varphi_*, \psi_*) = 1$  is plausible or not.

This range can be obtained based on a quantile  $t_{\text{MAC}}$  of the pdf of the approximated MAC distribution in (24) for the chosen confidence level  $\gamma$ , with e.g.  $\gamma = 0.95$ . Since the true MAC value is at the maximum of 1,  $t_{\text{MAC}}$  is chosen as the lower bound of the range of plausible MAC estimates. Since  $t_{\text{MAC}}$  is a quantile of the scaled and shifted  $\chi^2$  distribution in (24), it is easily computed as follows. First, the quantile  $t_{\chi^2}$  of the  $\chi^2$  distribution with  $l$  degrees of freedom is obtained such that

$$\int_0^{t_{\chi^2}} f_{\chi^2_l}(x) dx = \gamma. \quad (25)$$

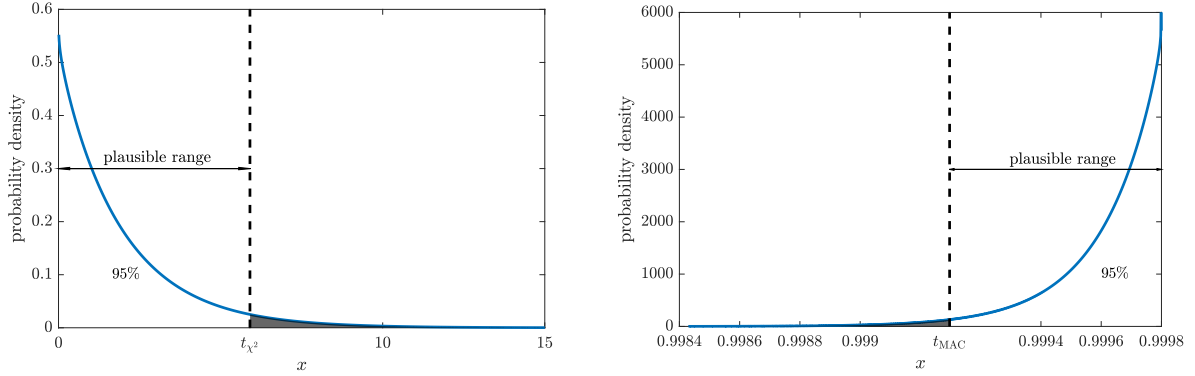


Figure 7: Classic  $\chi^2$  distribution function with 2 degrees of freedom, and the corresponding 95% quantile interval (left). Scaled and shifted  $\chi^2$  approximation of MAC distribution, and the corresponding 95% quantile (right).

Then, due to (24) the quantile for the MAC follows as

$$t_{\text{MAC}} = 1 - \frac{\beta}{N} - \frac{\alpha}{N} t_{\chi^2}. \quad (26)$$

Based on this quantile it can be inferred that MAC estimates with  $\text{MAC}(\hat{\varphi}, \hat{\psi}) \geq t_{\text{MAC}}$  are plausible, whereas estimates satisfying  $\text{MAC}(\hat{\varphi}, \hat{\psi}) < t_{\text{MAC}}$  are not. In the latter case it can be inferred that  $\text{MAC}(\varphi_*, \psi_*) \neq 1$  for the considered confidence level, and then that the MAC is computed between estimates of different mode shapes. The 95% quantile of the  $\chi^2$  distribution and the 95% quantile of the corresponding scaled and shifted  $\chi^2$  distribution (24) approximating the MAC are illustrated in Figure 7.

The procedure for the scaled and shifted  $\chi^2$  approximation of the MAC estimate and the computation of the associated quantile for the chosen confidence level  $\gamma$ , e.g. with  $\gamma = 0.95$ , is summarized in Algorithm 2.

---

**Algorithm 2:** MAC uncertainty computation for estimates of *equal* mode shapes

---

**Input** : mode shape estimates  $\hat{\varphi}$ ,  $\hat{\psi}$  and the corresponding covariance estimates  $\hat{\Sigma}_{\varphi_*}$ ,  $\hat{\Sigma}_{\psi_*}$  e.g. from subspace identification [14–16, 18]

**Output:** quantile of the MAC distribution approximation

- 1 compute Hessian  $\mathbf{H}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}}$  based on (17) using  $\hat{\varphi}$ ,  $\hat{\psi}$  and  $\hat{k} = \hat{\psi}^H \hat{\varphi} / \hat{\psi}^H \hat{\psi}$ , and  $\bar{\mathbf{H}}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}} = -\frac{1}{2} \mathbf{H}_{\hat{\varphi}, \hat{\psi}}^{\text{MAC}}$ ;
  - 2 compute  $c_1$ ,  $c_2$  and  $c_3$  in (19) ;
  - 3 compute number of degrees of freedom  $l$  in (21) ;
  - 4 compute scaling and shift parameters  $\alpha$ ,  $\beta$  in (22)–(23) ;
  - 5 compute quantile of scaled and shifted  $\chi^2$  approximation for the desired confidence level in (25)–(26)
- 

#### 4.4. Validation of scaled and shifted $\chi^2$ approximation

The MAC distribution is not only characterized by its mean and standard deviation, but also by further cumulants since it is clearly non-Gaussian. The scaled and shifted  $\chi^2$  distribution in (24) that approximates

the MAC distribution is based on the first three cumulants of the related quadratic form [25]. They can be obtained empirically from the Monte Carlo estimates of the MAC based on (15), here for  $\text{MAC}_{66}$ , as a reference for comparison.

With the second-order delta method, these cumulants are computed on only one estimate in (19), and are then the basis for the approximation of the MAC distribution with the scaled and shifted  $\chi^2$  distribution. The comparison of the empirical cumulants of the Monte Carlo histogram of  $\text{MAC}_{66}$  and their corresponding mean values computed with the second-order delta method together with their sample standard deviations ( $s_{c_j}$ ) is presented in Table 5. It can be viewed that the second-order delta method-based cumulants are indeed in the vicinity of their empirical counterparts, validating the cumulant computation with the proposed method for the subsequent characterization of the MAC estimates with the scaled and shifted  $\chi^2$  distribution from [25].

The Monte Carlo histograms of  $\text{MAC}_{66}$  are shown together with their computed fits with the scaled and shifted  $\chi^2$  distribution based on one arbitrarily selected estimate in Figure 8 for different data lengths  $N$ . It can be seen that the histograms seem to be correctly fitted with the proposed method. With larger  $N$ , the histogram accumulates more near the theoretical value 1, and the fits are qualitatively better as well. In Figure 9, the impact of the data length  $N$  on the 95% quantile  $t_{\text{MAC}}$  is shown, illustrating the plausible range of values for the MAC estimates. This range indeed contains the MAC estimate that was used to develop the distribution approximation, and it can be seen that the length of the range decreases with  $N$ , as expected.

In the next step, the quality of the scaled and shifted  $\chi^2$  distribution fits to the Monte Carlo histogram is evaluated. First, the fit based on the empirical MAC distribution is evaluated for reference. The empirical cumulants of the Monte Carlo histogram are computed based on (21), (22) and (24) for a fit based on all available Monte Carlo estimates. The goodness of fit to the histogram is presented in the second column of Table 6 for the different data lengths. Similar to the Gaussian case, it can be observed that the goodness of fit test values are below the significance threshold for the data length  $N = 50,000$  and onwards, indicating a good approximation of the empirical MAC distribution with the scaled and shifted  $\chi^2$  distribution. Second, the delta method-based approximations with the scaled and shifted  $\chi^2$  distribution are analyzed as in the Gaussian case. For each of the Monte Carlo estimates, the goodness of fit test-statistic is computed. From

Table 5: Validation of the cumulant computation for  $N = 100,000$

$\text{MAC}_{66}$	$c_1$	$c_2$	$c_3$
Monte Carlo	17.56	148.42	1227.6
Second-order delta method	17.82	161.09	1478.5
Sample standard deviation $s_{c_j}$	1.41	27.23	379.8

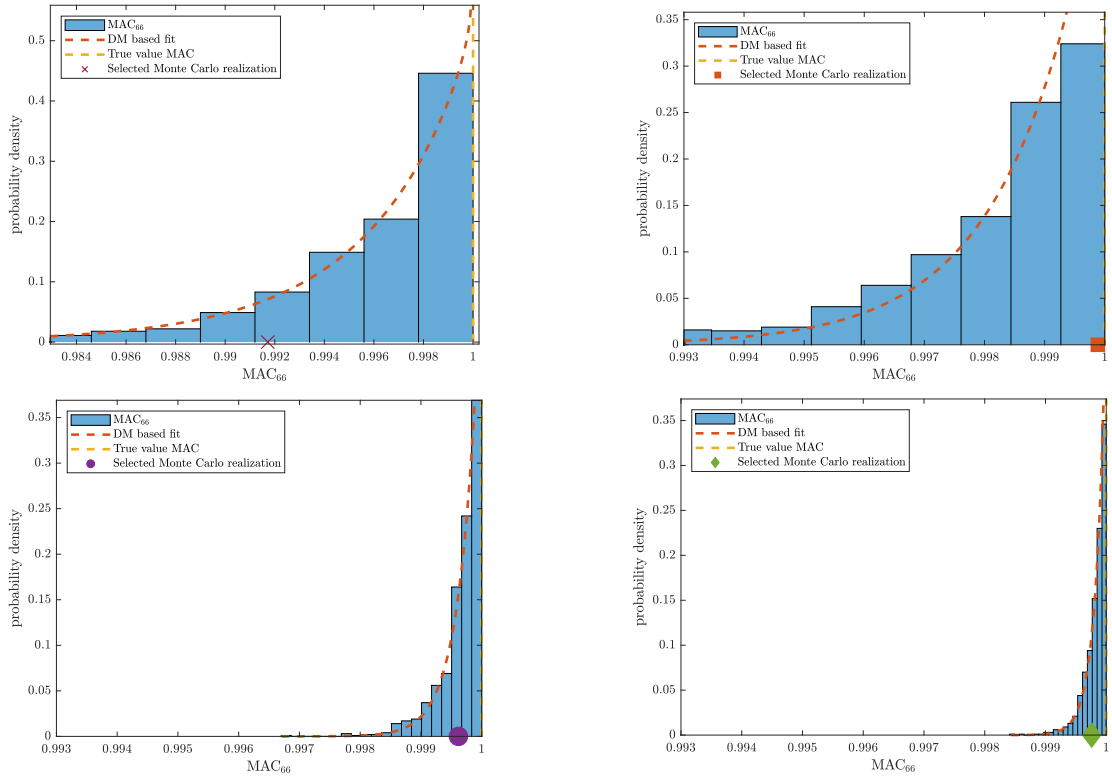


Figure 8: Delta method-based scaled and shifted  $\chi^2$  fits to Monte Carlo histogram of  $MAC_{66}$ .  $N = 5,000$  (top left),  $N = 10,000$  (top right),  $N = 50,000$  (bottom left),  $N = 100,000$  (bottom right).

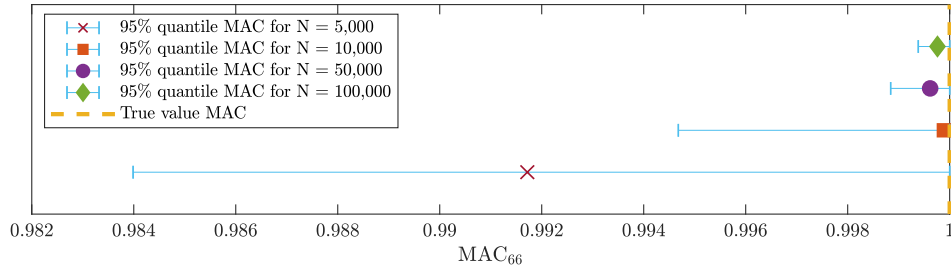


Figure 9: 95% support intervals of the  $\chi^2$  approximation of  $MAC_{66}$  estimated for different data lengths.

the test values, an estimate is selected at the 50% quantile that represents an average fit, as well as an estimate at the 95% quantile that represents a poor fit. Their goodness of fit test-statistics are shown in the last two columns of Table 6 for the different data lengths. It can be observed that the average delta method-based fits also match well the Monte Carlo histogram for  $N = 50,000$  and onwards. For the poor fits, the test-statistic decreases strongly with the data length and is nearly at the significance threshold for  $N = 100,000$ , similar to the Gaussian case, indicating adequate fits if the data length is long enough even for the few data sets with poor estimates.

Table 6: Pearson goodness of fit test-statistic for shifted and scaled  $\chi^2$  distribution fits to MC histogram for 95% significance level, threshold at 30.14

Data length	MC cumulant-based fit	average DM cumulant-based fit	poor DM cumulant-based fit
$N = 5,000$	38.01	129.67	476.52
$N = 10,000$	38.19	71.71	244.81
$N = 50,000$	26.92	28.33	61.67
$N = 100,000$	27.21	14.53	32.07

Table 7: Percentage of MAC estimates contained in the 95% support intervals for shifted and scaled  $\chi^2$  distribution approximations based on Monte Carlo histogram and on the second-order delta method

Data length	Monte Carlo-based interval	Delta method-based interval
$N = 5,000$	94.7%	97.5%
$N = 10,000$	94.5%	96.1%
$N = 50,000$	94.4%	95.7%
$N = 100,000$	95.4%	95.4%

The practical application of the proposed scaled and shifted  $\chi^2$  approximation is to estimate the quantile  $t_{\text{MAC}}$  of this distribution in (26) to infer if the MAC estimate is significantly close to one with the chosen confidence level of  $\gamma = 0.95$ . Based on the empirical MAC distribution, the quantile  $t_{\text{MAC}}$  can be computed from the distribution fit based on the cumulants from the Monte Carlo histogram as a reference. Then, 95% of the Monte Carlo estimates should be in the interval  $[t_{\text{MAC}}, 1]$ . The actual ratios that are contained in this interval are shown in the second column of Table 7 for the different data lengths. Indeed, all values are very close to 95%, which indicates that the proposed scaled and shifted  $\chi^2$  distribution assumption is adequate for the uncertainty quantification. Second, the delta method-based quantiles  $t_{\text{MAC},j}$  are examined, which are computed on each of the Monte Carlo estimates in (26). According to the theory, 95% of the intervals  $[t_{\text{MAC},j}, 1]$  should contain the respective MAC estimate. The ratios of occurrences where the estimates are in the interval are shown in the last column of Table 4. All values are again very close to 95%, showing that the proposed uncertainty quantification is accurate even if the empirical MAC distribution and the shifted and scaled  $\chi^2$  distribution approximation do not match perfectly for short data lengths.

These results illustrate that the developed characterization of the MAC distribution based on one estimate is accurate. Then the quantile of the distribution can be determined well, quantifying the uncertainty of the MAC estimate. In practice, where usually no histograms but only one MAC estimate is available, this will help to confirm or reject the underlying assumption that the MAC is computed between estimates of equal mode shapes.

## 5. Data-based decision framework for mode shape equality and damage detection

In Section 3, the uncertainty quantification of the MAC estimate has been carried out under the assumption that the true value  $\text{MAC}(\varphi_*, \psi_*) \neq 1$ , i.e. the MAC is computed from estimates of different mode shapes. In Section 4, the developed uncertainty quantification assumed  $\text{MAC}(\varphi_*, \psi_*) = 1$ , i.e. the mode shape estimates correspond to equal mode shapes. Both assumptions are exclusive in theory. In practice, no a priori information may be available on the true MAC value, and the choice between the assumptions needs to be made based on the mode shape estimates and their computed uncertainties.

For the case when the true MAC value is 1, the Gaussian approximation from Section 3 is invalid since  $\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} = 0$ . On the other side, the approximation with the scaled and shifted  $\chi^2$  distribution and its quantile  $t_{\text{MAC}}$  can always be computed in Section 4, and the underlying assumption that the true MAC value is 1 can be rejected if the actual MAC estimate is smaller than  $t_{\text{MAC}}$ . Therefore the quantile  $t_{\text{MAC}}$  is employed for a decision between both approximation frameworks: If the true MAC value is 1, the scaled and shifted  $\chi^2$  approximation is indeed adequate, yielding  $\text{MAC}(\hat{\varphi}, \hat{\psi}) \geq t_{\text{MAC}}$  for the chosen confidence level  $\gamma$ . If the true MAC value is smaller than 1, the MAC estimate will not be plausible for this distribution, yielding  $\text{MAC}(\hat{\varphi}, \hat{\psi}) < t_{\text{MAC}}$ , but then it is assured that the Gaussian approximation will be valid and can actually be computed. Based on these properties, a decision between both assumptions can be taken as follows:

- If  $\text{MAC}(\hat{\varphi}, \hat{\psi}) < t_{\text{MAC}}$ , then the MAC estimate is not plausible for the assumption  $\text{MAC}(\varphi_*, \psi_*) = 1$ , and thus a decision for  $\text{MAC}(\varphi_*, \psi_*) \neq 1$  is made. Then, the confidence interval of the Gaussian approximation can be computed with Algorithm 1.
- If  $\text{MAC}(\hat{\varphi}, \hat{\psi}) \geq t_{\text{MAC}}$ , then the assumption  $\text{MAC}(\varphi_*, \psi_*) = 1$  is indeed plausible, but the assumption  $\text{MAC}(\varphi_*, \psi_*) \neq 1$  could also be valid. In this case, a decision is made for  $\text{MAC}(\varphi_*, \psi_*) = 1$ . This choice is motivated by the fact that the quantile  $t_{\text{MAC}}$  approaches 1 with order  $\frac{1}{N}$  in (26), whereas the Gaussian interval shrinks only with order  $\frac{1}{\sqrt{N}}$  in (13). Thus the assumption  $\text{MAC}(\varphi_*, \psi_*) = 1$  is harder to satisfy in practice.

The decision between both assumptions depends on the chosen confidence level  $\gamma$  for the computation of the quantile  $t_{\text{MAC}}$ . The corresponding user-defined type I error for deciding  $\text{MAC}(\varphi_*, \psi_*) = 1$  is thus  $1 - \gamma$ . The resulting decision framework is summarized in Figure 10.

Note that any data-based decision in OMA is limited by the sensor instrumentation. In case of insufficient instrumentation, the mode shapes of changed (or different) modes could be classified as equal by the decision framework in Figure 10. Nonetheless, the proposed flow chart can be used in the damage detection context to monitor mode shape changes, which is detailed in the next section.

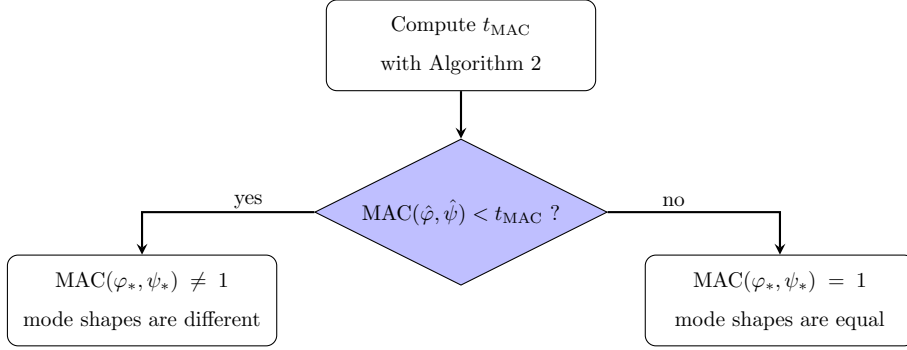


Figure 10: Flow chart of the decision framework for mode shape equality based on MAC estimates.

### 5.1. Usage in damage detection

In a monitoring context, the proposed framework can be used to assess if modes have changed over time, which will be illustrated on real data in Section 6. The basic assumption for vibration-based damage detection is that damage induces changes in the modal parameters, in particular in the natural frequencies and mode shapes. Early damage may induce only small changes in these values, thus the knowledge on their statistical uncertainties is particularly helpful to discriminate estimation errors from actual changes.

To assess possible changes over time, let a set of reference modal parameters  $(f_i^0, \hat{\varphi}_i)$ ,  $i = 1, \dots, m$  be computed from a data set of length  $N$  in the (healthy) reference state together with the asymptotic variances of the frequencies  $\hat{\sigma}_{f_i^0}^2$  and mode shapes  $\hat{\Sigma}_{\varphi_i}$ , where  $m$  is the number of analyzed modes. Then, based on a test data set recorded at a later point in time, the corresponding modal parameters  $(\hat{f}_i, \hat{\psi}_i)$ ,  $i = 1, \dots, m$ , are computed together with their covariances. To check if they correspond to the reference parameters with a chosen confidence level  $\gamma$ , it is verified for each mode  $i$  that the frequency  $\hat{f}_i$  is within the respective confidence bound of the reference value, e.g.  $\hat{f}_i \in [f_i^0 - 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{f_i^0}, f_i^0 + 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{f_i^0}]$  for  $\gamma = 0.95$ , and that the mode shapes  $\hat{\varphi}_i$  and  $\hat{\psi}_i$  are classified as equal with the proposed method. Based on each mismatch, a

---

#### Algorithm 3: Outline of damage detection based on modal parameters and their uncertainties

---

**Input** : Frequencies  $f_i^0$ , mode shapes  $\hat{\varphi}_i$  of reference state, with covariances  $\hat{\sigma}_{f_i^0}^2$ ,  $\hat{\Sigma}_{\varphi_i}$ ,  $i = 1, \dots, m$ ;  
 Corresponding frequencies  $\hat{f}_i$  and modes shapes  $\hat{\psi}_i$  in test state, with covariance  $\hat{\Sigma}_{\psi_i}$ ;  
 Confidence level  $\gamma$

**Output:** Set of modal parameters labelled as damaged

```

1 for  $i = 1 : m$  do
2   | infer if  $\hat{f}_i$  is inside  $\gamma$ -confidence interval of  $f_i^0$ , otherwise label  $\hat{f}_i$  as changed ;
3   | infer if  $\hat{\varphi}_i$  and  $\hat{\psi}_i$  are equal with decision framework in Figure 10, otherwise label  $\hat{\psi}_i$  as changed ;
4 end
  
```

---



damage indicator can be defined. The procedure is outlined in Algorithm 3.

The pertinence of the MAC uncertainty for detection of early damage becomes apparent when there is no significant change in the monitored frequencies. Firstly, it allows to confirm that the monitored system is unchanged if the mode shapes are classified as equal. Secondly, it allows damage detection when damage affects only the mode shapes but not the frequencies. In both cases, the proposed method avoids empirical thresholds for a classification, and instead a statistical decision based on the chosen confidence level is made.

## 6. Application

This section presents an application of the damage detection algorithm from Section 5. Herein, the developed statistical properties of the MAC are used to assess if mode shapes have changed or not. The application is carried out on measurement data from the monitoring campaign of the S101 Bridge in Austria, shown in Figure 11, before its demolition, during which different damages were inflicted on the structure [17].



Figure 11: S101 Bridge before demolition.

The responses of the bridge to ambient load were recorded using 15 acceleration sensors mounted on its deck, which is a realistic instrumentation example. Measurements were sampled with a frequency of 500 Hz for the period of three days. During this time the bridge was subjected to artificial damage, whose type and extent were controlled. From these measurements, 30 data sets of length  $N = 165,000$  are selected. Therein the first ten sets correspond to the healthy state of the bridge and the other data sets correspond to four different damage scenarios, namely the start of the first cut through the left pier, second cut through the left pier, settlement of the left pier by 1 cm and settlement of the left pier by 2 cm. Those scenarios are respectively denoted as  $d_1$ ,  $d_2$ ,  $d_3$  and  $d_4$ . Prior to the modal parameter estimation and the statistical analysis of MAC, data are decimated to 25 Hz. Subsequently, the modal parameters and their covariances are estimated at system order 80 using the data-driven SSI-UPC algorithm. The covariances are computed using the method developed in [16]. Three fundamental modes of the bridge with natural frequencies at around 4.05 Hz, 6.28 Hz and 9.69 Hz are estimated. The corresponding mode shapes from the first healthy data set are used as reference mode shapes, to which the mode shapes from the other data sets are compared

in the MAC computation.

Let the MAC and the natural frequencies estimated in the reference state be used to track sets of corresponding modal parameters throughout the measurement campaign. The resultant sequences of MAC estimates are respectively denoted as  $MAC_{11}$ ,  $MAC_{22}$  and  $MAC_{33}$ , and are shown together with the corresponding natural frequency estimates in Figure 12.

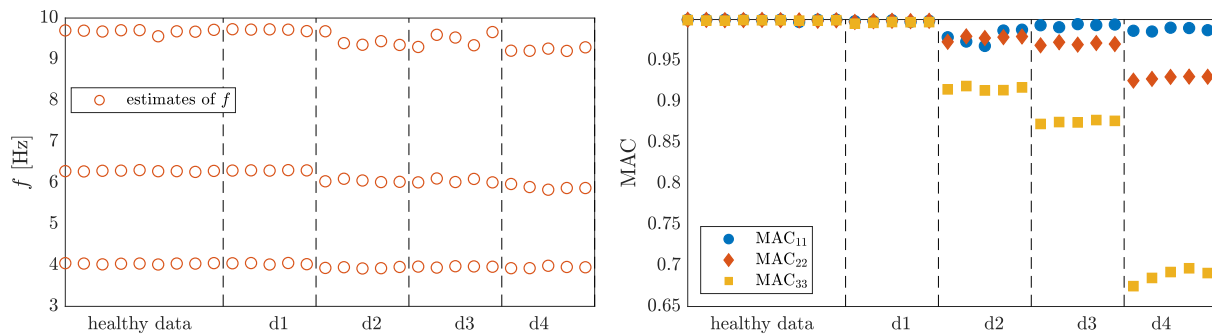


Figure 12: Natural frequencies estimated from the tested data sets (left). MAC values corresponding to the three modes (right).

It can be observed that the estimated natural frequencies are quite stable for the healthy and  $d_1$  data sets, and start to decrease as more damage is inflicted on the structure in  $d_2 - d_4$ . Throughout all damage cases, the corresponding MAC values are in the range of  $0.92 - 1.00$  for the first two modes, suggesting little or even no change in these mode shapes. For the third mode, they decrease to  $0.70$  in damage cases  $d_2 - d_4$ , suggesting a significant change in this mode shape as the damage progresses. To evaluate whether the respective mode shapes estimated throughout the measurement campaign correspond to the reference mode shapes from the first data set, the quantile  $t_{MAC}$  of the second-order approach is computed first based on each MAC estimate, and checked whether the MAC estimate is below or above the quantile, as outlined in Section 5. Subsequently, this information is used to infer about the damage, as sketched in Algorithm 3. This is particularly interesting for the first damage scenario  $d_1$ , since the MAC estimates for all the corresponding mode shapes are very close to 1, and the corresponding natural frequency estimates are statistically indistinguishable.

In Figure 13, the MAC estimates between the reference (first data set) and the second data set in the healthy state are shown. Their respective Gaussian confidence intervals and the quantile  $t_{MAC}$  of the scaled and shifted  $\chi^2$  approximation are computed with Algorithm 1 and Algorithm 2, each for a confidence level of  $\gamma = 0.95$ . Since the MAC estimate is within the respective quantile interval for all modes, the assumption of equal mode shapes is correctly chosen, and the Gaussian intervals are invalid. Similarly, the MAC estimates are shown in Figure 14 for the first damage case  $d_1$ . For this small damage, the MAC estimates are outside the quantile for all three modes, thus the assumption of equal mode shapes is rejected. In this case, the Gaussian confidence intervals are valid, and it can be seen that they are either separated or just slightly

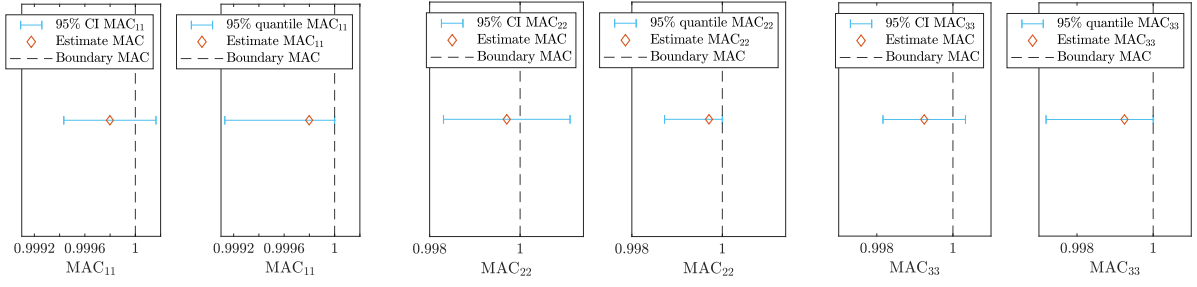


Figure 13: Estimates of  $MAC_{11}$  (left),  $MAC_{22}$  (middle) and  $MAC_{33}$  (right) with corresponding 95% confidence intervals and quantiles in the healthy state (data set 2).

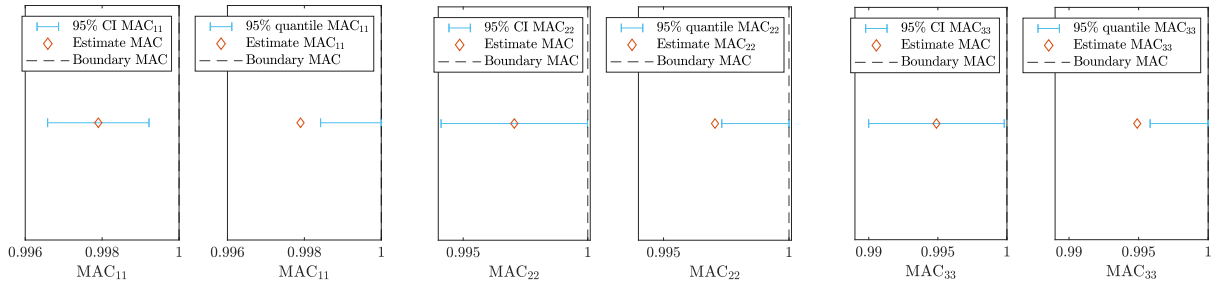


Figure 14: Estimates of  $MAC_{11}$  (left),  $MAC_{22}$  (middle) and  $MAC_{33}$  (right) with corresponding 95% confidence intervals and quantiles in the damage state  $d_1$  (data set 11).

overlapping the boundary at 1. Note that the choice of the confidence level is related to the application at hand, and it is defined by the user. It is a trade-off between the detection of small changes and a low false alarm rate. The chosen confidence level of  $\gamma = 0.95$  is used in many modal parameter uncertainty related applications [13, 14, 33], but higher confidence levels can be evaluated analogously with the proposed methodology. In such a case, it is possible that the small change of the mode shapes in damage scenario  $d_1$  will not be considered as significant anymore. The final choice about the value of  $\gamma$  is thus a trade-off, and depends on the end-user preference between an early detection and a sure decision.

The MAC estimates for all tested data sets with their corresponding  $\chi^2$  quantile intervals are shown in Figure 15. In the healthy state, all MAC estimates are located inside the intervals, which confirms the plausibility of the related assumption, i.e. that the mode shape estimates are statistically equal and that no change happened for that mode. Starting with damage case  $d_1$ , none of the MAC estimates are inside the  $\chi^2$  quantile intervals, showing that the assumption of equal mode shapes does not hold anymore, with the exception of one outlier for the third mode. Thus, a change in the mode shapes is detected already for this small damage thanks to the uncertainty quantification of the MAC estimate. Hence, it is correctly verified that the mode shapes are different between healthy and damaged states based on the MAC and its uncertainty.

Figure 16 displays the time evolution of the estimated natural frequencies, which correspond to  $MAC_{11}$ ,

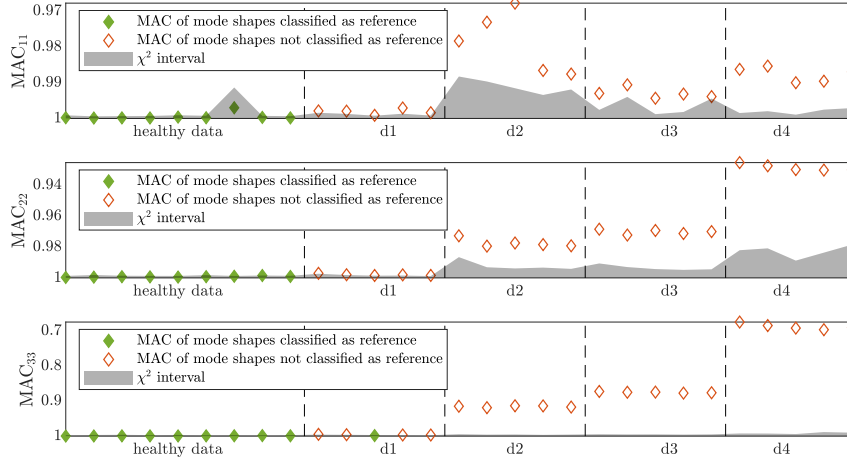


Figure 15: Estimates of  $MAC_{11}$  (top),  $MAC_{22}$  (middle) and  $MAC_{33}$  (bottom) with corresponding 95%  $\chi^2$  quantile intervals.

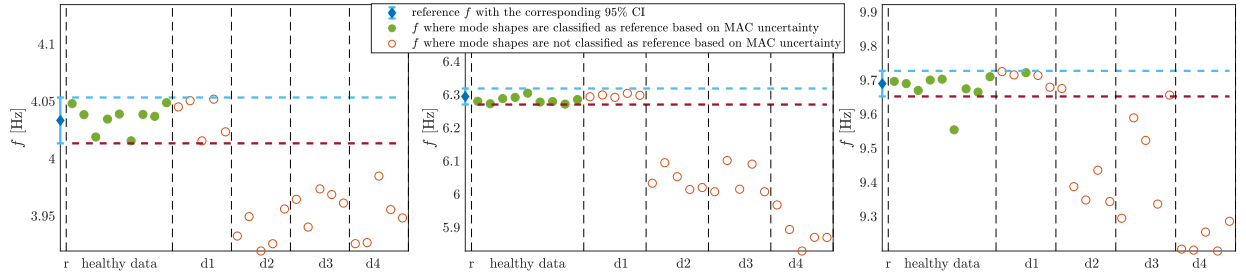


Figure 16: Time histories of the first three natural frequency estimates. The dashed horizontal lines represent the 95% confidence interval of the reference natural frequency.

$MAC_{22}$  and  $MAC_{33}$ . For the healthy period of the bridge, 96% of the estimates were within the confidence intervals defined by their respective reference frequency estimates, and only one outlier was present in the third mode. This is in line with the chosen confidence level. During the  $d_1$  period, damage is very small, and all the frequencies reside within the reference confidence intervals. Thus, based on the natural frequency estimates alone,  $d_1$  is not detected as damaged. The occurrence of more severe damages present in  $d_2 - d_4$ , however, can be simply detected solely based on the natural frequency estimates.

To detect both the early damages of  $d_1$  and the larger damages of  $d_2 - d_4$ , Algorithm 3 is used. All mode shapes in the healthy period are correctly labeled as unchanged. For  $d_1$ , 93% of the mode shapes were reported to be different from the reference mode shape. The uncertainty information of the reference natural frequency estimate is used in a par with the MAC uncertainty to label the modes in  $d_2 - d_4$  as damaged. In all the considered cases, the uncertainty information of MAC has either confirmed the frequency-based detection of damages, or improved upon it for the early damage period.

## 7. Conclusions

In this paper, the uncertainty quantification of the Modal Assurance Criterion between mode shapes estimated from different data sets has been investigated and used to decide if the underlying mode shapes are equal or not. When the MAC is computed between estimates of different mode shapes, then it can be characterized by a Gaussian distribution, for which confidence intervals are easily obtained. Conversely, when the MAC is computed between estimates of equal mode shapes, the resulting MAC estimate is close to the boundary of possible MAC values at one, which prevents the application of a classical Gaussian approximation. In this case, it has been shown that the MAC estimate can be characterized by a scaled and shifted  $\chi^2$  distribution. Then the relevant uncertainty information is obtained by the quantile of this distribution for a desired confidence level. Both approaches have been validated numerically.

If the MAC estimate is close to one, and it is unknown whether the mode shape estimates are actually equal up to statistical uncertainty or not, then it has been shown how the developed MAC uncertainty quantification frameworks can be used to evaluate their equality. This is in particular pertinent for structural health monitoring, which has been illustrated in an engineering application to data of the progressively damaged S101 Bridge. Therein, even very small changes in the mode shapes between the healthy state and an early damage state could be detected. This allowed to distinguish modal parameters corresponding to the healthy state of the bridge from those of the damaged state. While the pertinence of the proposed MAC uncertainty quantification has been shown in the context of structural damage detection in this paper, its relevance for other problems like model updating, modal tracking, or analysis of stabilization diagrams should be explored in future works.

## Acknowledgements

We thank VCE for providing the data from S101 Bridge.

## Appendix A. Proof of Property (10)

First it is proven that the partial derivatives (5)-(8) are equal to zero when  $\text{MAC}(\varphi_*, \psi_*) \in \{0, 1\}$ . Since the MAC measures the degree of linearity between two vectors, the case  $\text{MAC}(\varphi_*, \psi_*) = 1$  indicates that the vectors  $\psi_*$  and  $\varphi_*$  are collinear. The opposite  $\text{MAC}(\varphi_*, \psi_*) = 0$  indicates that two vectors are orthogonal. The first case can be encompassed by  $\varphi_* = k \cdot \psi_*$  where  $k \in \mathbb{C}$  is some constant scalar. Plugging it to the derivatives (5)-(8) yields them be equal to zero, which is trivial to show. Considering the case when two mode shape vectors are orthogonal  $\varphi_*^H \psi_* = 0$  similarly cancels the derivatives (5)-(8), and concludes the first part of the proof. Second we prove

$$\mathcal{J}_{\varphi_*, \psi_*}^{\text{MAC}} = 0 \Rightarrow \text{MAC}(\varphi_*, \psi_*) \in \{0, 1\}.$$

This implies in particular that the partial derivatives (5)-(6) are zero, yielding

$$\frac{\partial \text{MAC}}{\partial \Re(\varphi)}(\varphi_*, \psi_*) + \frac{\partial \text{MAC}}{\partial \Im(\varphi)}(\varphi_*, \psi_*) = \frac{2(\psi_* \psi_*^H \varphi_*)^T}{\varphi_*^H \varphi_* \psi_*^H \psi_*} - \frac{2\varphi_*^T \text{MAC}(\varphi_*, \psi_*)}{\varphi_*^H \varphi_*} = 0. \quad (\text{A.1})$$

Here two cases are distinguished, namely first  $\psi_*^H \varphi_* \neq 0$ , and second  $\psi_*^H \varphi_* = 0$ . Concerning the first case, transposing (A.1) and multiplying with  $\psi_*^H$  from the left yields

$$\psi_*^H \psi_* \psi_*^H \varphi_* = \text{MAC}(\varphi_*, \psi_*) \psi_*^H \varphi_* \psi_*^H \psi_* \Rightarrow \text{MAC}(\varphi_*, \psi_*) = \frac{\psi_*^H \psi_* \psi_*^H \varphi_*}{\psi_*^H \varphi_* \psi_*^H \psi_*} = 1.$$

As for the second case, when  $\psi_*^H \varphi_* = 0$  (A.1) writes as

$$\psi_* \psi_*^H \varphi_* - \text{MAC}(\varphi_*, \psi_*) \varphi_* \psi_*^H \psi_* = 0, \quad \text{MAC}(\varphi_*, \psi_*) \varphi_* = 0 \Rightarrow \text{MAC}(\varphi_*, \psi_*) = 0,$$

leading to the assertion in (10).

## Appendix B. Derivation of the Hessian

The Hessian  $\mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}}$  is a symmetric matrix of shape

$$\mathbf{H}_{\varphi_*, \psi_*}^{\text{MAC}} = \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} \\ T_{21} & T_{22} & T_{23} & T_{24} \\ T_{31} & T_{32} & T_{33} & T_{34} \\ T_{41} & T_{42} & T_{43} & T_{44} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Re(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Re(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Re(\psi) \partial \Re(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\psi) \partial \Re(\varphi)} \\ \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Im(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Im(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Re(\psi) \partial \Im(\varphi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\psi) \partial \Im(\varphi)} \\ \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Re(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Re(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Re(\psi) \partial \Re(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\psi) \partial \Re(\psi)} \\ \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Im(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Im(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Re(\psi) \partial \Im(\psi)} & \frac{\partial^2 \text{MAC}}{\partial \Im(\psi) \partial \Im(\psi)} \end{bmatrix}_{(\varphi, \psi) = (\varphi_*, \psi_*)}$$

To obtain the second-order derivatives, the first-order derivatives (5)-(8) are derived with respect to the respective mode shape parts, and evaluated using the asymptotic relation in (16). The top left 2x2 block of the Hessian writes

$$\begin{aligned} T_{11} &= \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Re(\varphi)}(\varphi_*, \psi_*) = \frac{2\Re(\psi_*)\Re(\psi_*)^T + 2\Im(\psi_*)\Im(\psi_*)^T - 2\psi_*^H \psi_* I_r}{|k|^2 (\psi_*^H \psi_*)^2}, \\ T_{12} &= \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Re(\varphi)}(\varphi_*, \psi_*) = \frac{2\Re(\psi_*)\Im(\psi_*)^T - 2\Im(\psi_*)\Re(\psi_*)^T}{|k|^2 (\psi_*^H \psi_*)^2}, \\ T_{21} &= \frac{\partial^2 \text{MAC}}{\partial \Re(\varphi) \partial \Im(\varphi)}(\varphi_*, \psi_*) = T_{12}^T = -T_{12}, \\ T_{22} &= \frac{\partial^2 \text{MAC}}{\partial \Im(\varphi) \partial \Im(\varphi)}(\varphi_*, \psi_*) = \frac{2\Re(\psi_*)\Re(\psi_*)^T + 2\Im(\psi_*)\Im(\psi_*)^T - 2I_r \psi_*^H \psi_*}{|k|^2 (\psi_*^H \psi_*)^2} = T_{11}, \end{aligned}$$

leading to

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \frac{2}{|k|^2 d_*^2} H_2$$

with  $H_2$  defined in (18) and  $d_* = \psi_*^H \psi_*$ . The remaining blocks of the Hessian are derived analogously, and they can all be tied back to the top left 2x2 block. After straightforward simplifications they yield

$$\begin{aligned} T_{13} &= -\Re(k)T_{11} - \Im(k)T_{12}, & T_{14} &= \Im(k)T_{11} - \Re(k)T_{12}, \\ T_{23} &= -\Re(k)T_{21} - \Im(k)T_{22}, & T_{24} &= \Im(k)T_{21} - \Re(k)T_{22}, \end{aligned}$$

and  $T_{33} = |k|^2 T_{11}$ ,  $T_{34} = |k|^2 T_{12}$ ,  $T_{43} = |k|^2 T_{22}$ ,  $T_{44} = |k|^2 T_{22}$ . With these relations between the blocks of the Hessian and its symmetry, the factorization of the Hessian in (17) follows.

## References

- [1] R. J. Allemang, The modal assurance criterion (MAC): Twenty years of use and abuse, *Sound and Vibration* 37 (8) (2003) 14–21.
- [2] W. Heylen, T. Janter, Applications of the modal assurance criterion in dynamic model updating, in: *Proceedings of the ASME Design Technical Conference*, Vol. 18, 1989, pp. 289–294.
- [3] D. Ribeiro, R. Calada, R. Delgado, M. Brehm, V. Zabel, Finite element model updating of a bowstring-arch railway bridge based on experimental modal parameters, *Engineering Structures* 40 (2012) 413 – 435.
- [4] E. Reynders, A. Teughels, G. De Roeck, Finite element model updating and structural damage identification using OMAX data, *Mechanical Systems and Signal Processing* 24 (5) (2010) 1306 – 1323.
- [5] T. Ting, T. Chen, W. Twomey, Correlating mode shapes based on the modal assurance criterion, *Finite Elements in Analysis and Design* 14 (4) (1993) 353 – 360.
- [6] M. Brehm, V. Zabel, C. Bucher, An automatic mode pairing strategy using an enhanced modal assurance criterion based on modal strain energies, *Journal of Sound and Vibration* 329 (25) (2010) 5375 – 5392.
- [7] L. F. Ramos, R. Aguilar, P. B. Lourenço, S. Moreira, Dynamic structural health monitoring of Saint Torcato church, *Mechanical Systems and Signal Processing* 35 (1) (2013) 1 – 15.
- [8] E. Reynders, J. Houbrechts, G. De Roeck, Fully automated (operational) modal analysis, *Mechanical Systems and Signal Processing* 29 (2012) 228 – 250.
- [9] M. Pastor, M. Binda, T. Harčarik, Modal assurance criterion, *Procedia Engineering* 48 (2012) 543 – 548, *Modelling of Mechanical and Mechatronics Systems*.
- [10] W. M. West, Illustration of the use of modal assurance criterion to detect structural changes in an orbiter test specimen, 1986.
- [11] C. H. J. Fox, The location of defects in structures - A comparison of the use of natural frequency and mode shape data, in: *10th International Modal Analysis Conference*, Vol. 1, 1992, pp. 522–528.
- [12] P. Verboven, E. Parloo, P. Guillaume, M. Van Overmeire, Autonomous structural health monitoring - Part I: Modal parameter estimation and tracking, *Mechanical Systems and Signal Processing* 16 (4) (2002) 637 – 657.
- [13] R. Pintelon, P. Guillaume, J. Schoukens, Uncertainty calculation in (operational) modal analysis, *Mechanical Systems and Signal Processing* 21 (6) (2007) 2359 – 2373.
- [14] 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.
- [15] M. Döhler, L. Mevel, Efficient multi-order uncertainty computation for stochastic subspace identification, *Mechanical Systems and Signal Processing* 38 (2) (2013) 346–366.
- [16] 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 (Supplement C) (2016) 1 – 27.
- [17] M. Döhler, F. Hille, L. Mevel, W. Rücker, Structural health monitoring with statistical methods during progressive damage test of S101 Bridge, *Engineering Structures* 69 (2014) 183 – 193.
- [18] M. Döhler, P. Andersen, L. Mevel, Variance computation of modal parameter estimates from UPC subspace identification, in: *Proceedings of 7th International Operational Modal Analysis Conference (IOMAC)*, Ingolstadt, Germany, 2017.
- [19] 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.

- [20] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*, Springer-Verlag, Berlin, Heidelberg, 2006.
- [21] J. Wishart, M. S. Bartlett, The distribution of second order moment statistics in a normal system, *Mathematical Proceedings of the Cambridge Philosophical Society* 28 (4) (1932) 454-459.
- [22] G. Cui, X. Yu, S. Iommelli, L. Kong, Exact distribution for the product of two correlated Gaussian random variables, *IEEE Signal Processing Letters* 23 (11) (2016) 1662-1666.
- [23] R. K. Mallik, N. C. Sagias, Distribution of inner product of complex Gaussian random vectors and its applications, *IEEE Transactions on Communications* 59 (12) (2011) 3353-3362.
- [24] G. Casella, R. L. Berger, *Statistical inference*, 2nd Edition, Cengage Learning, 2001.
- [25] H. Liu, Y. Tang, H. H. Zhang, A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables, *Computational Statistics and Data Analysis* 53 (4) (2009) 853 - 856.
- [26] M. I. Friswell, The Derivatives of Repeated Eigenvalues and Their Associated Eigenvectors, *Journal of Vibration and Acoustics* 118 (3) (1996) 390-397.
- [27] 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.
- [28] 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.
- [29] J. Neyman, H. Jeffreys, Outline of a theory of statistical estimation based on the classical theory of probability, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 236 (767) (1937) 333-380.
- [30] L. A. Aroian, The probability function of the product of two normally distributed variables, *Ann. Math. Statist.* 18 (2) (1947) 265-271.
- [31] J. C. Hayya, W. L. Ferrara, On normal approximations of the frequency functions of standard forms where the main variables are normally distributed, *Management Science* 19 (2) (1972) 173-186.
- [32] R. Ware, F. Lad, Approximating the distribution for sums of products of normal variables, *Tech. rep.* (2003).
- [33] J. Cara, Modal identification of structures from input/output data using the expectation-maximization algorithm and uncertainty quantification by mean of the bootstrap, *Structural Control and Health Monitoring* 26 (1) (2019) e2272.