



**HAL**  
open science

# Statistical model-based optimization for damage extent quantification

Szymon Gres, Michael Döhler, Laurent Mevel

► **To cite this version:**

Szymon Gres, Michael Döhler, Laurent Mevel. Statistical model-based optimization for damage extent quantification. *Mechanical Systems and Signal Processing*, 2021, 160, pp.107894. 10.1016/j.ymssp.2021.107894 . hal-03468260

**HAL Id: hal-03468260**

**<https://inria.hal.science/hal-03468260>**

Submitted on 7 Dec 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.

# Statistical model-based optimization for damage extent quantification\*

Szymon Greś\*, Michael Döhler, Laurent Mevel

*Univ. Gustave Eiffel, Inria, COSYS/SII, I4S, Campus de Beaulieu, 35042 Rennes, France*

---

## Abstract

Damage localization and quantification constitute different aspects of structural damage diagnosis, which are of particular interest in the Structural Health Monitoring field. Therein, a classical solution is model updating, where the parameters of a finite element model of the possibly damaged structure are optimized to match with the corresponding parameters estimated from its vibration responses. To avoid ill-posedness of the classical finite element updating problem, damage localization and quantification can be treated separately. First, the information about regions or clusters of possibly damaged elements in the structure is obtained by a damage localization method. Then, this information is used to reduce the number of parameters for damage quantification. A framework combining the advantages of methods for damage localization with model optimization is proposed in this paper. For the exploration of the clustered physical model space, a stochastic optimization algorithm is coupled with the evaluation of the statistical properties of the MAC and frequency differences between the numerical model and the estimated modes for an adequate treatment of the data-based uncertainties. Herein, the development of the statistical properties of the MAC estimate is an important step, which is based on a recent quadratic framework that is adapted to the context of the inner product between an estimated mode shape and a numerical mode shape. This statistical information is used in the formulation of the objective function as well as in a data-driven stopping criterion for the optimization search. The proposed framework is validated on numerical simulations of a beam model, where damage at multiple locations is quantified up to the clustering precision.

*Keywords:* Damage quantification, Stochastic optimization, Modal parameters, Modal assurance criterion, Uncertainty quantification

---

## 1. Introduction

The damage diagnosis problem can be divided into damage detection, localization and quantification [1]. When performing these tasks in a cascade fashion, the full structural diagnosis problem is amenable to a

---

\*©<2021>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

\*Corresponding author; *E-mail address:* [szymon.gres@inria.fr](mailto:szymon.gres@inria.fr)

solution. Whereas the detection of damage from vibration measurements is well established, e.g., in [2–6], the localization and quantification of damage is more complex, and requires additional physical information on the examined structure [7–10]. In this respect, few complete frameworks for damage identification exist, e.g., [11–13]. Some methods consider a specific structural type in their theory, e.g., beams [12], while other methods incorporate the required physical information from a finite element (FE) model, which allows application to more complex structures. In this context, the sensitivities of a damage feature computed from structural responses can be used to infer FE model parameter changes in [11], under the assumption of small damage extents. Contrary to these approaches, FE model updating methods are usually not limited by the structural type or by the damage extent. Model updating is a generic term encompassing a family of direct and indirect methods [14, 15] that are often applied in the damage quantification context [16–19]. While applicable to arbitrary structures, model updating is often poorly conditioned due to the possibly large FE parameter space in comparison to relatively few parameters that can be extracted from data, and is prone to statistical uncertainty errors of the estimated features. The objective of this paper is to develop a cascade framework where the damage is first localized and then quantified on the subset of the damaged parameter space, and which explicitly considers data-based uncertainties. The resultant methodology belongs to the family of indirect methods and relies on an iterative FE model optimization.

The purpose of model optimization is to adjust the model to minimize the misfit between some model features and the corresponding data features. The model is expressed by a full physical parametrization that cannot be directly retrieved from data due to limited sensor instrumentation and limited modal information. A minimization procedure will yield the best model candidate regarding the data, and in the context of the current paper, the best change candidate for damage quantification. In this context, the features of potential interest are derived from the modal parameters [20–23].

Depending on the model complexity, the physical parametrization size may be large. Different values of the parametrization may yield the same numerical frequencies and mode shapes at the limited number of sensor positions, and may be indistinguishable when compared to the data features. Conversely, the estimated data features may react similarly to damage occurring in different parameter components. Thus a change in some parameter components of the physical model can be indistinguishable for the considered feature, i.e., the modal parameters. A pragmatic solution is to group such physical parameter components into a cluster. A cluster contains a subset of parameters, whose sensitivity is *close* and no distinction among different elements of the cluster can be made. Then, only a global evaluation of the change in the parameter value over the cluster is meaningful. A sensitivity-based parameter clustering was first proposed in [24] and developed further in [25–27]. In [7], the limitation of working with a large physical parameter space and a limited set of sensors and modes was addressed by Fisher information-based clustering. The benefit of using the Fisher information-based clusters resides in accounting for the statistical uncertainty of the metric to locate changes in the considered parametrization, which is in contrast to deterministic sensitivity-based

approaches. As such, the Fisher information-based clustering will be used in the current paper to reduce the dimension of the parameter search domain for model optimization.

Model optimization is to be done by means of an objective function. Considering the large size of the parameter search domain, the minimization of the objective function by exhaustive search may not be an option. To circumvent taxing brute-force search, different algorithms have been investigated, like stochastic gradient-type methods [28] which are however limited to convex functions, or global deterministic search methods [29]. Global and stochastic algorithms, like genetic algorithms and evolution strategies, are good candidates for finding the optimum for non-convex and rugged objective functions. Therein the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [30] appears to be suitable for optimization of non-linear, non-convex functions with local minima, and has been applied to damage quantification [31]. A number of convergence studies suggest that CMA-ES is likely to reach the optimum for a large enough number of iterations [32–34], which is usually fixed to an upper bound. As a consequence and drawback, this large number of iterations can lead to an unnecessary computational burden for the resulting optimization. In this context, an improvement will be the development of a stopping criterion able to assess the quality of the solution to achieve an early stop. A good stopping criterion will prevent a waste of resources by ceasing the algorithm when further optimization is not statistically significant, as well as avoiding a premature stopping [35]. Classic criteria are formulated with empirical factors, e.g., relative change in the optimized parameters, relative change in the value of the objective function, or maximum number of iterations. Those factors are usually related to the performance of the optimization algorithm, and not to the quality of the investigated solution, and are defined without any insight on the uncertainty of the estimated parameters. A properly designed stopping criterion should relate to the properties of the investigated parameters [36]. A stopping criterion based on the modal parameter estimates and the MAC estimate is proposed herein and developed to take into account their statistical uncertainty.

The uncertainty in model updating can be due to statistical feature estimation or due to idealization of the model-based parameters [37]. The former originates due to the finite data length and noise [38], and the latter stems from the assumptions and simplifications in modeling. In order to assess the quality of the solution given by the model optimization, the model and measurement uncertainties should be accounted for. Model errors are oftentimes taken into account with stochastic model updating [39–41], which is beyond the scope of this paper. Parameter estimation errors can be accounted for by incorporating their statistical uncertainty to the objective function, which was addressed by Bayesian weighting least-squares problem [23], by model perturbation [42, 43] or by using the variance of estimated model parameters computed empirically over different data sets [15, 21]. An extensive review on this subject can be found in [37], where the necessity of further examining the statistical uncertainty related to estimated parameters was pointed out. Besides the uncertainty errors, both the estimated and the model features can vary due to changes in environmental and operating conditions. Such variations must be properly considered when analyzing the

misfit between both, either by removing the environmental effects from the estimated features [44–47], or by using adequate modeling assumptions for both the numerical model and the identification method [48, 49]. As such, accounting for changes in environmental and operating conditions is important in practice, however it is beyond the scope of this paper.

The current paper focuses on the impact of the uncertainty of the estimated modal parameters in the model optimization, and most particularly for the MAC estimate, whose statistical properties in this context were never investigated, but only inferred in Bayesian updating [50]. In general, the MAC can be viewed as the inner product of two Gaussian unit vectors. Considerable attention has been devoted to the derivation of the distribution properties of such a product in the past, which also applies to the MAC. The majority of the work, however, focuses on the case where both vectors are stochastic [51–53]. A possible solution was investigated in [54], where the exact mean values of both vectors are required, which are not available in practice. In the present paper, the uncertainty quantification is achieved by the statistical delta method [55], which has been deployed previously as a first order linearization for evaluating uncertainties of modal parameters [38, 56], modal indicators [57, 58], or for physical parameters of mechanical systems [28], without a priori information on the solution. In [58], a second-order extension of the first-order delta method was used to derive a general scheme for uncertainty quantification of the MAC between two estimated mode shapes. The present paper considers a particular application of MAC, where one mode shape is estimated and the other is deterministic. As such, its uncertainty can be derived based on the general approach from [58] but under different assumptions.

In summary, this paper comprises the following developments on three different levels, namely

- theoretical: the explicit derivation of the statistical uncertainty of the MAC between a mode shape of an FE model and an estimated mode shape,
- algorithmic: the design of an objective function and a stopping criterion incorporating the variance of modal parameters estimated from one data set,
- methodological: the development of a cascade framework for damage quantification, where first the damage is localized among a set of clustered parameters, and then quantified by minimization of the previously elaborated objective function thanks to an evolutionary algorithm.

These developments lead to a coherent optimization approach for damage quantification, whose novelty resides in the consideration of uncertainties for its objective function, its stopping criterion and the definition of the search domain in terms of Fisher information-based parameter clusters.

The plan of this paper is as follows. The considered model optimization problem is stated in Section 2. The effect of the statistical uncertainties on the objective function is illustrated in Sections 3 and 4, where adequate uncertainty quantification approaches are developed and incorporated to its formulation. The

indistinguishability of physical parameters is considered further in Section 4.3. The damage quantification framework based on the damage localization-based parameter clustering and the stochastic optimization scheme is presented in Section 5, before illustrating it on a numerical application in Section 6.

## 2. Problem statement

The overall goal of damage quantification is the estimation of the structural parameter change between a healthy (reference) state and the current damaged state of the structure. Let  $\theta \in \Theta \subset \mathbb{R}^p$  be the parameter vector that contains the damage-sensitive parameters of the structural elements of interest for the considered problem within a bounded parameter space  $\Theta$ . The parametrization is user-defined and adapted to the specific monitoring problem at hand. It is assumed that the parameter vector  $\theta_0$  in the healthy state of the structure is known, and let  $\theta_*$  be the parameter vector in the damaged state. It is the goal to obtain an estimate  $\hat{\theta}$  of  $\theta_*$  from vibration measurements in the damaged states for estimating the damage extent  $\delta = \theta_* - \theta_0$ . Finding the optimal  $\hat{\theta}$  can be formulated as an optimization problem, where an objective function  $F(\theta)$ , also called cost function, is minimized over the parameter space  $\Theta$ .

The objective function is designed to represent the discrepancy between the estimate of a feature vector  $\hat{v}$  computed from measurement data recorded under the (unknown) system parameter  $\theta_*$ , and its counterpart  $v(\theta)$  computed from a parametric model. The optimal solution for  $\theta$  is then obtained as

$$\hat{\theta} = \arg \min_{\theta \in \Theta} F(\theta). \quad (1)$$

The monitored system is modeled as linear time-invariant (LTI) with  $d$  degrees of freedom (DOF) and it is described by the differential equation of motion

$$\mathcal{M}^\theta \ddot{q}(t) + \mathcal{D}^\theta \dot{q}(t) + \mathcal{K}^\theta q(t) = u(t) \quad (2)$$

where  $t$  denotes continuous time, and  $\mathcal{M}^\theta$ ,  $\mathcal{D}^\theta$  and  $\mathcal{K}^\theta \in \mathbb{R}^{d \times d}$  denote the mass, damping and stiffness matrices, respectively, which depend on the system parameter  $\theta$ . The 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.

The *true* modal parameters of system (2) are obtained from the eigenvalues  $\mu_i^\theta \in \mathbb{C}$  and eigenvectors  $\Psi_i^\theta \in \mathbb{C}^d$ , which are the solution of the eigenvalue problem

$$((\mu_i^\theta)^2 \mathcal{M}^\theta + \mu_i^\theta \mathcal{D}^\theta + \mathcal{K}^\theta) \Psi_i^\theta = 0, \quad (3)$$

and yield the natural frequencies  $f_i^\theta = \frac{|\mu_i^\theta|}{2\pi}$  and the mode shapes  $\psi_i^\theta = L \Psi_i^\theta \in \mathbb{C}^r$  at the observed DOF for a given sensor type, where  $L \in \mathbb{C}^{r \times d}$  maps the eigenvectors to the mode shapes that can be obtained from measurements at  $r$  sensors, respectively. Denote  $\hat{f}_i$  and  $\hat{\varphi}_i$  the respective estimates of the  $i$ -th natural frequency and the  $i$ -th mode shape, and assume that the deployed system identification method is consistent,

i.e.,  $\hat{f}_i$  and  $\hat{\varphi}_i$  converge to their true values  $f_{i,*}$  and  $\varphi_{i,*}$  when the data length  $N$  goes to infinity. This is satisfied by multiple system identification methods, for example stochastic subspace identification (SSI) methods [56, 59, 60], where these parameters can be identified, e.g., from acceleration, velocity, displacement or strain measurements.

A classic feature vector  $\hat{v}$  for the design of the objective function  $F(\theta)$  in (1) is based on  $\hat{f}_i$  and  $\hat{\varphi}_i$  for  $i = 1, \dots, m$  where  $m$  denotes the number of considered modes, and  $v(\theta)$  is its counterpart computed from the model. The main issue for the solution of the resulting optimization problem is its possible ill-posedness or ill-conditionedness due to the fact that different parameter vectors  $\theta$  can yield the same or close feature vectors  $v(\theta)$ . The problem is aggravated by the presence of uncertainties in the estimate  $\hat{v}$ , which are unavoidable due to unknown excitation and measurement noise. Indeed, the modal parameter estimates are random variables that have been proved to be asymptotically Gaussian for SSI methods [61, 62], satisfying

$$\hat{f}_i \approx \mathcal{N}\left(f_{i,*}, \frac{1}{N}\sigma_{f_i}^2\right), \text{ and } \begin{bmatrix} \Re(\hat{\varphi}_i) \\ \Im(\hat{\varphi}_i) \end{bmatrix} \approx \mathcal{N}\left(\begin{bmatrix} \Re(\varphi_{i,*}) \\ \Im(\varphi_{i,*}) \end{bmatrix}, \frac{1}{N}\Sigma_{\varphi_i}\right), \quad (4)$$

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  $\sigma_{f_i}^2 \in \mathbb{R}$  and  $\Sigma_{\varphi_i} \in \mathbb{R}^{2r \times 2r}$  are the asymptotic covariance of  $\hat{f}_i$  and  $\hat{\varphi}_i$ , respectively. The computation of the respective covariance estimates  $\hat{\sigma}_{f_i}^2$  and  $\hat{\Sigma}_{\varphi_i}$  from the same data set as the modal parameter estimates can be found, e.g., in [38, 56, 62].

In Section 2.1, an academic beam example is defined, which will serve as a supporting proof of concept for the different methodological steps introduced throughout the paper, as well as for the application of the developed quantification method in Section 6. In Section 2.2 a standard modal parameter-based objective function is recalled, which will be further developed by incorporating the estimation uncertainties in Sections 3 and 4.

### 2.1. Illustrative example

The considered Euler-Bernoulli beam has 95 elements and consequently 192 DOF. The beam is modeled in free-free boundary conditions. For the sake of simulation, damping is assumed to be proportional, where the damping matrix is defined such that each mode has a damping ratio of 1%. The system is excited by a white noise signal acting at all translational DOF. The acceleration responses are recorded by six sensors at a sampling frequency of 5000 Hz with a length of 200 s, and white Gaussian measurement noise with a standard deviation of 5% of the standard deviation of the output is added to the each response measurement. The beam model is illustrated in Figure 1.

Damage is emulated as changes in the mass properties of the structure. In the reference state, the density of each beam element is  $\rho_j = 2700$ , where  $j = 1 \dots 95$  is the element number. Two damage cases are considered, namely

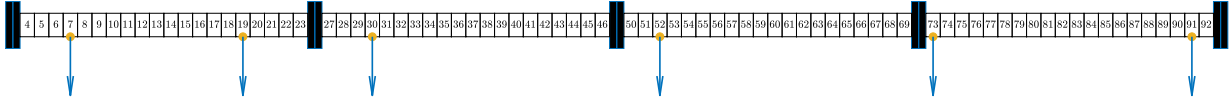


Figure 1: Numerical beam model with sensors.

Table 1: Exact natural frequencies  $f$  (Hz) of the beam model.

State	$f_1$	$f_2$	$f_3$	$f_4$	$f_5$	$f_6$	$f_7$	$f_8$	$f_9$	$f_{10}$	$f_{11}$	$f_{12}$
Healthy	39.98	110.56	215.91	371.27	532.01	713.03	958.19	980.86	1275.5	1300.5	2119.7	2130.8
Case 1	39.95	110.56	215.90	371.18	531.73	712.65	956.57	979.61	1275.5	1300.5	2112.6	2127.6
Case 2	39.93	110.42	215.90	370.76	530.99	712.56	956.08	979.31	1274.8	1299.7	2112.3	2127.6

- Case 1: 15% increase in  $\rho_9$ ,
- Case 2: 15% increase in  $\rho_9$  and  $\rho_{33}$ .

Hereafter let  $\theta$  be the collection of the element densities  $\rho_j$ . Albeit the inflicted damage relates to changes in the density of beam elements, the damage quantification framework proposed herein is general and can be formulated to quantify changes in any parameter of the numerical model for which the modal parameter sensitivities are non-zero.

Both damage scenarios modify the exact natural frequencies of the FE model. The chosen modes for the analysis comprise the first twelve flexible modes of the beam and are presented in Table 1. It can be viewed that the reduction in the natural frequencies is small and no more than 0.08% for the first mode in the second damage scenario. The applied damage also perturbs the mode shape vectors. The MAC between two complex-valued mode shapes vectors  $\varphi$  and  $\psi$  is defined as [63]

$$\text{MAC}(\varphi, \psi) = \frac{|\varphi^H \psi|^2}{\varphi^H \varphi \psi^H \psi} = \frac{\varphi^H \psi \psi^H \varphi}{\varphi^H \varphi \psi^H \psi}, \quad (5)$$

and the change in mode shapes can be observed when comparing the mode shapes of the healthy and damaged models through the corresponding MAC values. The maximum decrease in the MAC occurs for the twelfth mode shape vector at up to 14.8% for the MAC between the mode shape of the healthy model and the same mode shape in damage case 2.

## 2.2. Modal parameter-based objective function

A classic objective function for model optimization is expressed by the difference between the estimated and the numerical modal parameters [37]. For this purpose, the estimated modal parameters must be paired with their counterparts from the numerical model to correspond to the same mode. It is hereafter assumed that such a pairing is accomplished, e.g., with modal parameter-based criteria [64]. Denote  $\Delta_{f_i}^\theta$  as the



normalized difference between the estimated natural frequency  $\hat{f}_i$  and the natural frequency  $f_i^\theta$  from the model under parameter  $\theta$ , with

$$\Delta_{f_i}^\theta = 1 - \frac{f_i^\theta}{\hat{f}_i}. \quad (6)$$

For the comparison of the mode shape estimate  $\hat{\varphi}_i$  (computed from measurements under unknown system parameter  $\theta_*$ ) to the numerical counterpart  $\psi_i^\theta$  from the model, the MAC is used. If  $\theta = \theta_*$ , the MAC between both mode shapes is close to 1 and tends to 1 when the data length  $N \rightarrow \infty$ , since  $\hat{\varphi}_i$  is a consistent estimate. Conversely, if the MAC between both mode shapes tends to 1, the parameter vector  $\theta$  is a valid candidate for  $\theta_*$ , up to some question of model identifiability that is discussed later in the paper.

Let  $\Delta_{\text{MAC}_i}^\theta$  be the difference between 1 and the respective MAC, defined as

$$\Delta_{\text{MAC}_i}^\theta = 1 - \text{MAC}(\hat{\varphi}_i, \psi_i^\theta). \quad (7)$$

The objective function  $F(\theta)$  indicates the distance between the estimated modal parameters and their model counterparts, and it can be expressed as the sum of the respective differences  $\Delta_{f_i}^\theta$  and  $\Delta_{\text{MAC}_i}^\theta$  for all considered modes  $i = 1, \dots, m$ , as

$$F(\theta) = \sum_{i=1}^m \left| \Delta_{f_i}^\theta \right| + \sum_{i=1}^m \Delta_{\text{MAC}_i}^\theta. \quad (8)$$

The function  $F(\theta)$  in (8) maps the  $p$ -dimensional parameter space to a multidimensional hyperplane whose shape can easily be illustrated for  $p = 2$ . Consider the estimates of  $f_i^{\theta_*}$  and  $\varphi_i^{\theta_*}$  corresponding to the beam data simulated for the damage case 1 as described in Section 2.1, and the beam model parametrized with element densities  $\rho_j$  as illustrated in Figure 1. The function  $F(\theta)$  is displayed for the parameter pair  $(\rho_8, \rho_9) \in [2600, 3400]^2$  in Figure 2 (left). It can be viewed that  $F(\theta)$  is convex in the considered parameter region. The global minimum of  $F(\theta)$  is at 0.022, which is misaligned with the value  $F(\theta_*) = 0.034$  at the exact parameter value pair  $(2700, 3105)$  for the parameters  $\rho_8$  and  $\rho_9$ . This is a consequence of the statistical

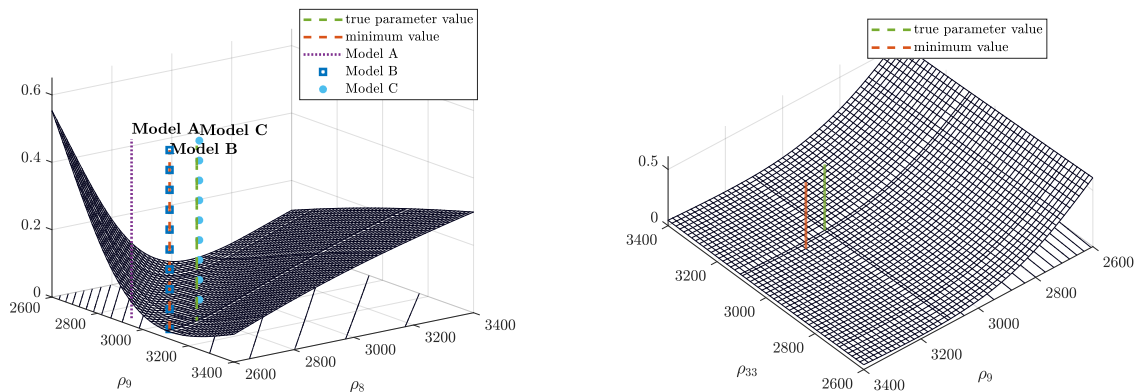


Figure 2: Objective function  $F(\theta)$  for the parameter pair  $\rho_8$  and  $\rho_9$  in the single fault case (left) and for the parameter pair  $\rho_9$  and  $\rho_{33}$  in the double fault case (right).

Table 2: Different model states and corresponding pairs of  $\rho_8$  and  $\rho_9$  in damage case 1.

Model	$\rho_8$	$\rho_9$	$F(\theta)$
A	2600	2950	0.066
B	2616	3096	0.022
C	2735	3070	0.029
Exact	2700	3105	0.034

uncertainty introduced by the modal parameter estimates in  $F(\theta)$ . As a result, the true model can reach higher values of  $F(\theta)$  than other slightly different but *close* models, as illustrated by three parametrization cases in Table 2. The values of  $F(\theta)$  for these models are very close, and to answer which one can be accepted as a solution cannot be solely based on the shape of the discussed objective function  $F(\theta)$ , but should also take into account the underlying modal parameter uncertainties. Similarly, the objective function for damage case 2 is shown in Figure 2 (right), having also low values for a wide range of parameters.

### 3. Objective function considering the statistical uncertainty of the natural frequency estimates

First, the statistical uncertainties of the frequency estimates are introduced into the objective function, before considering the statistical properties of the MAC in the next section. Notice that only the statistical uncertainty is considered here, under the assumption that there exists a model parameter vector  $\theta = \theta_*$  such that the modal parameter estimates can converge to their model counterparts if the data length is large enough. Misspecified modeling or model errors are not taken into account in this paper.

#### 3.1. Confidence interval of the natural frequency estimates

Since the frequency estimates are approximately Gaussian distributed, their confidence intervals are symmetric and centered around the computed estimates. They cover a range of plausible values that include the true value of the parameter with some given confidence level  $\gamma$ , e.g.,  $\gamma = 0.9544$  for the  $\pm 2\sigma$  bound. At this value, the corresponding confidence interval

$$c_{\hat{f}_i} = [\hat{f}_i^{\min}, \hat{f}_i^{\max}] \quad (9)$$

encompasses the true value  $f_{i,*}$ , where  $\hat{f}_i^{\min} = \hat{f}_i - 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{f_i}$  and  $\hat{f}_i^{\max} = \hat{f}_i + 2 \cdot \frac{1}{\sqrt{N}} \hat{\sigma}_{f_i}$  with 95.44% probability, which is hereafter denoted as the 95% confidence interval. Assume any model parameter  $\theta$  and the data are generated under the parameter  $\theta_*$ . If the natural frequency of the model  $f_i^\theta$  is inside  $c_{\hat{f}_i}$ , then it can be inferred that both  $\theta$  and  $\theta_*$  are statistically close. Conversely, when  $\theta$  differs significantly from  $\theta_*$ , the corresponding model frequency  $f_i^\theta$  is likely to be outside the confidence interval of the estimated frequency. This can be illustrated based on a Monte Carlo simulation of the beam from Section 2.1, where 1000 data sets are simulated for damage case 1. For each data set, the modal parameters and their variances

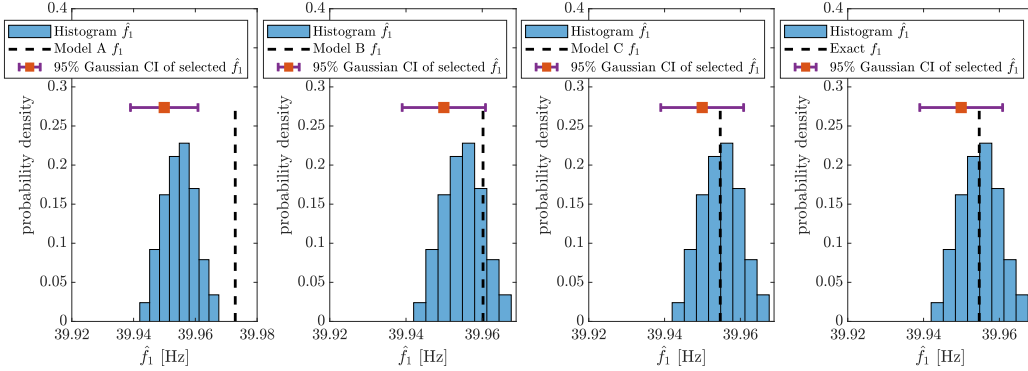


Figure 3: Illustration of Gaussian confidence intervals for frequency estimates.

are estimated with data-driven subspace-based system identification [62]. The estimated modal parameters are then compared to the numerical modal parameters of the models under different parameter vectors  $\theta$  as depicted in Table 2. Histograms of the first natural frequency estimates together with a 95% confidence interval for one selected estimate are illustrated in Figure 3.

It can be viewed in the first histogram that Model A is statistically different from  $\theta_*$ , where  $\theta_*$  corresponds to damage case 1. This is reflected by the fact that its first natural frequency (dashed line) is strictly outside the confidence interval centered around the selected natural frequency estimate. This applies for the chosen estimate in Figure 3 (red square), however it would also be true for most of the values of the histogram and their variances. In the second histogram, Model B shows a closer correspondence, hence its first frequency will belong to more of the confidence intervals computed for each frequency estimate. The first frequency of Model C will belong to the confidence interval for most of the values of the third histogram. Based on the displayed estimates, Models B and C yield acceptable solutions for  $\rho_8$  and  $\rho_9$ , and Model A would be rejected most of the time.

### 3.2. Design of an objective function considering the confidence intervals of natural frequency estimates

The normalized difference (6) between the estimated frequency  $\hat{f}_i$  and the model frequency  $f_i^\theta$  is now penalized when  $f_i^\theta$  is outside the 95% confidence interval of  $\hat{f}_i$ . The corresponding difference is defined as

$$\tilde{\Delta}_{f_i}^\theta = \begin{cases} 1 - \frac{f_i^\theta}{\hat{f}_i}, & \text{if } \hat{f}_i^{\min} \leq f_i^\theta \leq \hat{f}_i^{\max} \\ 1, & \text{otherwise} \end{cases} \quad (10)$$

and the corresponding objective function is defined as

$$\tilde{F}(\theta) = \sum_{i=1}^m \left| \tilde{\Delta}_{f_i}^\theta \right| + \sum_{i=1}^m \Delta_{\text{MAC}_i}^\theta. \quad (11)$$

Based on the confidence intervals (9), the acceptance region

$$\tilde{\Theta} = \{\theta : f_i^\theta \in c_{\hat{f}_i} \text{ for all } i = 1, \dots, m\} \quad (12)$$

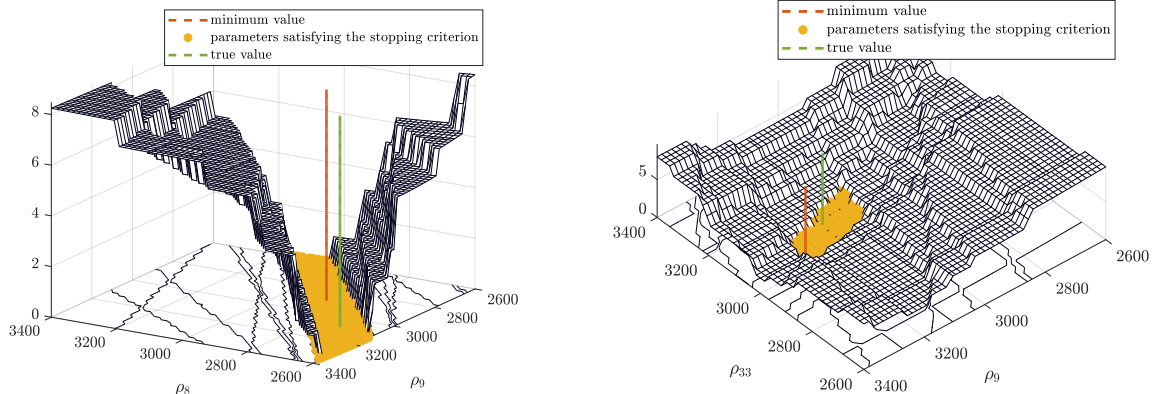


Figure 4: Objective function  $\tilde{F}(\theta)$  for single fault (left) and double fault (right).

can be defined, which comprises all parameter vectors  $\theta$  that yield model frequencies within the confidence intervals of their estimated counterparts. Thus,  $\tilde{\Theta}$  comprises the statistically acceptable solutions for  $\theta_*$  with regards to the considered frequencies, and a stopping criterion of the optimization search can be formulated as  $\theta \in \tilde{\Theta}$ .

The shape of  $\tilde{F}(\theta)$  is illustrated in Figure 4 for the same data set and the same parameter space as for the previously analyzed objective function  $F(\theta)$ . For both damage cases, it can be viewed that  $\tilde{F}(\theta)$  is rugged and steeper than the classic  $F(\theta)$  in Figure 2. Also, the bottom region corresponds by design to all the models where all the model frequencies and the estimated frequencies satisfy (12). The resulting acceptance region is still quite wide, with the minimum and the true value well apart, and does not take into account any statistical information about the mode shapes, which is analyzed next.

#### 4. Objective function considering the uncertainty of both the frequency and MAC estimates

In this section, the uncertainty related to the MAC estimate is investigated for the objective function and the stopping criterion. This uncertainty is particularly relevant when evaluating if the estimated mode shapes  $\hat{\varphi}_i$  and numerical mode shapes  $\psi_i^\theta$  match, which is an objective for model optimization. Due to the statistical nature of the estimated mode shapes, it is not meaningful to evaluate if the MAC is exactly at 1 for a match, but if it is close enough to 1 up to the statistical uncertainty.

While classical Gaussian confidence intervals can be easily defined for the frequencies, as in the previous section, the uncertainty quantification of the MAC is more intricate. The MAC distribution can only be approximated as Gaussian when the two mode shapes are significantly different, i.e., the MAC is well separated from 1. Then, the mode shape covariance can be propagated to the MAC by a first-order sensitivity approach similar to [38, 62]. However, for model optimization it is the goal to find a parameter  $\theta$  such that the estimated and numerical mode shapes match. Then,  $\text{MAC}(\hat{\varphi}_i, \psi_i^\theta)$  should be close to 1 and its limit value

is expected to be 1 for  $N \rightarrow \infty$ . The distribution of  $\text{MAC}(\hat{\varphi}_i, \psi_i^\theta)$  is bounded at 1 where it will accumulate, so it cannot be Gaussian or symmetric anymore. Thus, the MAC distribution for this case needs to be characterized before the uncertainty can be evaluated, which is carried out in detail in Appendix A. Based on the framework of [58], it is shown successively:

1. The statistical distribution of the MAC between a mode shape estimate and a numerical mode shape is characterized by a quadratic form if the mode shapes match. The analysis is based on a second-order Taylor expansion, as shown in Appendix A.1.
2. The distribution of the quadratic form can be approximated by a scaled and shifted  $\chi^2$  distribution, whose parameters are detailed in Appendix A.2.

Finally, this statistical analysis is used to determine whether the underlying mode shapes are matching. The evaluation of the MAC distribution yields a quantile at the desired confidence level and subsequently an interval of plausible values of MAC estimates under the assumption that the true MAC value is 1. Then, this assumption can be accepted or rejected based on the value of the MAC estimate and the associated interval, to evaluate if the mode shape estimate matches the mode shape from the model in the optimization process. The computation of the quantile for the MAC uncertainty quantification is detailed in Section 4.1 and summarized in Algorithm 1, before incorporating it into the objective function in Section 4.2.

#### 4.1. Computing the quantile of MAC as a criterion for the collinearity of $\hat{\varphi}_i$ and $\psi_i^\theta$

The assumption  $\text{MAC}(\varphi_{i,*}, \psi_i^\theta) = 1$  is checked based on the MAC estimate  $\text{MAC}(\hat{\varphi}_i, \psi_i^\theta)$  and its statistical properties to verify if the parameter  $\theta$  corresponds to the true parameter  $\theta_*$  in the optimization process. Under this assumption, it is not useful to define a confidence interval around the MAC estimate in which the *true MAC value* should lie as in the classical Gaussian case, since the true MAC value is known and assumed to be 1. Instead, a quantile  $t_{\text{MAC}_i}$  of the MAC distribution approximation can be obtained for a chosen confidence level  $\gamma$ , e.g.,  $\gamma = 0.95$ , yielding

$$\int_{t_{\text{MAC}_i}}^1 f_{\text{MAC}}(x) dx = \gamma, \quad (13)$$

where  $f_{\text{MAC}}$  is the probability density function (pdf) of the MAC distribution in (A.11). Based on (A.11), the quantile  $t_{\text{MAC}_i}$  is computed as

$$t_{\text{MAC}_i} = 1 - \frac{\beta}{N} - \frac{\alpha}{N} t_{\chi^2}, \quad (14)$$

where  $t_{\chi^2}$  is the  $\gamma$  quantile of the underlying  $\chi^2$  distribution at  $l$  DOF, satisfying  $\int_0^{t_{\chi^2}} f_{\chi^2}(x) dx = \gamma$ . Parameters  $l$ ,  $\alpha$  and  $\beta$  of the MAC distribution are computed in Appendix A.2.

Since the assumed true MAC value is at its maximum at 1,  $t_{\text{MAC}_i}$  has been chosen as the lower bound of the range of the plausible MAC estimates in (13), which are contained in the interval  $[t_{\text{MAC}_i}, 1]$ . This interval

---

**Algorithm 1:** Computation of the q-interval related to MAC uncertainty
 

---

**Input** : mode shape estimate  $\hat{\varphi}$ , model mode shape  $\psi^\theta$  and the mode shape covariance estimate  $\hat{\Sigma}_\varphi$ , e.g., from subspace identification [38, 56, 62]

**Output:** quantile  $t_{\text{MAC}}$  of the MAC distribution approximation

- 1 compute Hessian  $H_{\hat{\varphi}}^{\text{MAC}}$  based on (A.8) using  $\hat{\varphi}$ ,  $\psi^\theta$  and  $\hat{k} = \psi^H \hat{\varphi} / \psi^{\theta H} \psi^\theta$ , and  $\bar{H}_{\hat{\varphi}}^{\text{MAC}} = -\frac{1}{2} H_{\hat{\varphi}}^{\text{MAC}}$ ;
  - 2 compute  $c_1$ ,  $c_2$  and  $c_3$  in (A.9) ;
  - 3 compute number of DOF  $l = c_2^3 / c_3^2$ , scaling parameter  $\alpha = c_3 / c_2$  and shift parameter  $\beta = c_1 - c_2^2 / c_3$  (see Appendix A.2);
  - 4 compute quantile  $t_{\text{MAC}}$  for the desired confidence level in (14)
- 

is denoted as the q-interval and contains the MAC estimate with probability  $\gamma$ , if the true MAC value is actually 1. Thus, it can be checked if the MAC estimate lies in the q-interval to verify the assumption that the true MAC is 1. Note that this is different to the Gaussian case, where the true value should lie in the confidence interval around the estimate. Based on  $t_{\text{MAC}_i}$  it can be inferred that the MAC estimates with  $\text{MAC}(\hat{\varphi}_i, \psi_i^\theta) \geq t_{\text{MAC}_i}$  are plausibly close to 1, whereas MAC estimates satisfying  $\text{MAC}(\hat{\varphi}_i, \psi_i^\theta) < t_{\text{MAC}_i}$  are not. In the latter case it can be inferred that  $\text{MAC}(\varphi_{i,*}, \psi_i^\theta) \neq 1$  for the considered confidence level, thus the true mode shape  $\varphi_{i,*}$  and the model mode shape  $\psi_i^\theta$  do not match, and consequently  $\psi_i^\theta$  does not correspond to  $\theta_*$ .

The proposed framework is applied to match the first mode shape computed from data of the beam example (see Section 2.1) with the first mode shape of the numerical models B and C as stated in Table 2. Monte Carlo histograms of the MAC estimates, along with the corresponding distribution quantiles, are illustrated in Figure 5. For Model B, the MAC histogram is clearly Gaussian shaped in Figure 5 (left). A Gaussian confidence interval for a selected MAC estimate is obtained based on (A.2)–(A.4) for a 95%

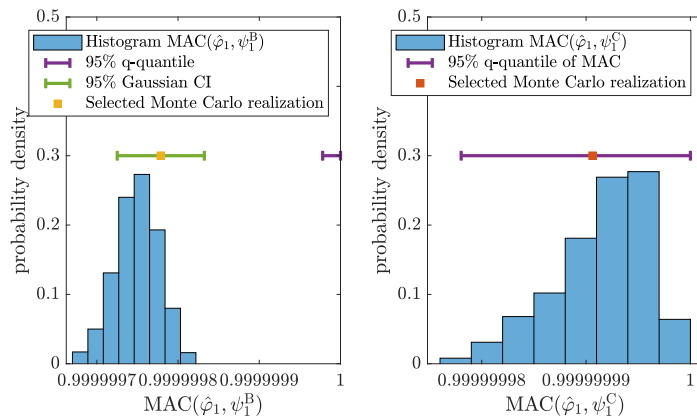


Figure 5: Examples of Gaussian confidence interval and quadratic quantile for the MAC histograms.

confidence level (green error bar), and it can be seen that it covers most of the histogram including the mean value. The entire histogram lies outside the q-interval of the selected estimate for a 95% confidence level (purple error bar), which indeed indicates that the mode shape estimate is not corresponding to the model with statistical significance. Conversely, for the MAC histogram for Model C in Figure 5 (right), the q-interval contains around 95% of the MAC histogram values, indicating matching mode shapes. Thus testing if the estimate is within  $[t_{\text{MAC}_i}, 1]$  leads to the acceptance or the rejection of the hypothesis that the numerical mode shape and the estimated mode shape correspond to the same model.

#### 4.2. Design of an objective function considering uncertainty information for frequency and MAC

If the MAC between the tested numerical mode shape and the reference estimate does not belong to the q-interval for the chosen confidence level, the numerical mode shape is considered to be different from the estimate. In this case, the corresponding value (7) is penalized in the objective function, and the difference between the MAC and 1 is defined as

$$\tilde{\Delta}_{\text{MAC}_i}^\theta = \begin{cases} 1 - \text{MAC}(\hat{\varphi}_i, \psi_i^\theta), & \text{if } t_{\text{MAC}_i} \leq \text{MAC}(\hat{\varphi}_i, \psi_i^\theta) \\ 1, & \text{otherwise} \end{cases} \quad (15)$$

The resulting objective function writes

$$\bar{F}(\theta) = \sum_{i=1}^m \left| \tilde{\Delta}_{f_i}^\theta \right| + \sum_{i=1}^m \tilde{\Delta}_{\text{MAC}_i}^\theta, \quad (16)$$

and the corresponding acceptance region is defined as

$$\bar{\Theta} = \{\theta : f_i^\theta \in c_{\hat{f}_i} \text{ and } \text{MAC}(\hat{\varphi}_i, \psi_i^\theta) > t_{\text{MAC}_i} \text{ for all } i = 1, \dots, m\}. \quad (17)$$

Thus,  $\bar{\Theta}$  comprises the statistically acceptable solutions for  $\theta_*$  with regards to the considered frequencies and MAC values, and a stopping criterion of the optimization search can be formulated as  $\theta \in \bar{\Theta}$ .

The resulting plot of the objective function for the considered numerical example from Section 2.1 is displayed in Figure 6, where it can be seen that the stopping region is narrower than in Figure 4. Taking into account the MAC uncertainty provides more information and reduces the pool of accepted model candidates.

#### 4.3. Model identifiability discussion

Due to limited sensor instrumentation, the acceptance region for  $\bar{F}(\theta)$  in (17) does not only contain a single value of parameter  $\theta$ , but there are many physical models that yield the same frequencies and mode shapes. For example, the region highlighted in Figure 7 is the acceptance region for damage case 1, corresponding to a slice of the discretized parameter space for three parameters. As it can be seen, this region is a plane encompassing dependent parameter components. All the models corresponding to these values will generate statistically equivalent natural frequencies and mode shapes. As such, no separation between

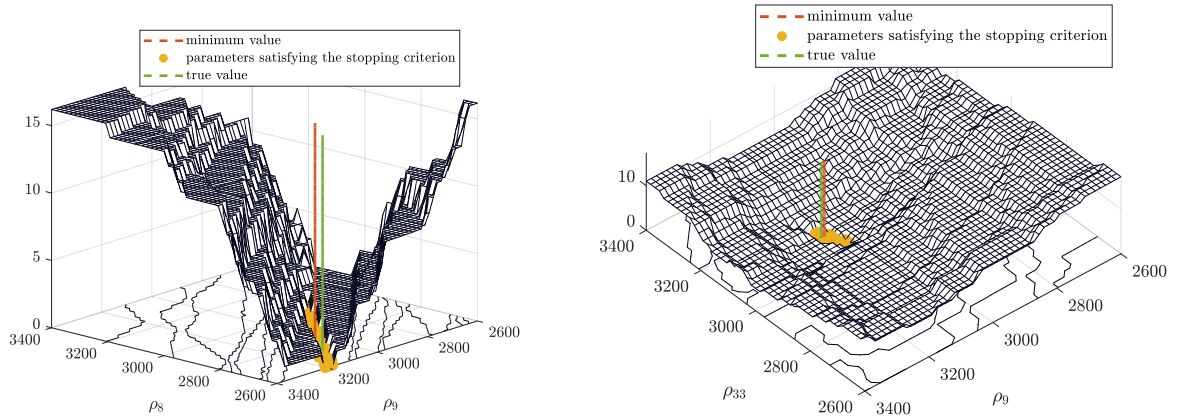


Figure 6: Objective function  $\bar{F}(\theta)$  for single fault (left) and double fault (right).

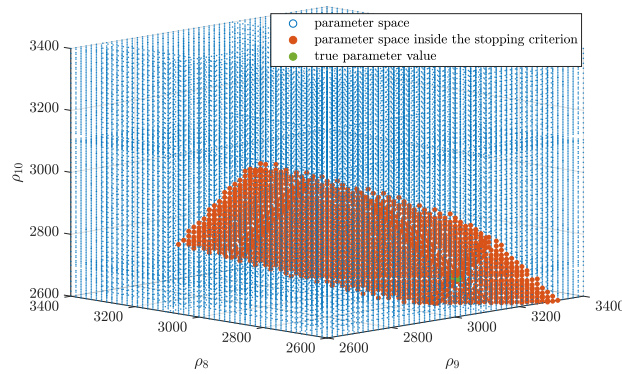


Figure 7: Discretized parameter space of the beam for parameter triplet  $(\rho_8, \rho_9, \rho_{10})$ .

these models is possible, which relates to the identifiability of the model through the available limited set of sensors and modes. All these models are equally valid as a solution of the optimization problem. Since the ultimate goal is the quantification of damage, the respective parameter changes should not be evaluated for each parameter component individually, but should also include the dependent parameter components, as described in the following section.

## 5. Damage quantification strategy

In the previous section, an objective function has been elaborated whose set of minima corresponds to models that are statistically compliant with the modal parameter estimates. As such, looking for the minima of this function will yield the sought models. This, however, is not trivial. For example, some optimization strategies might be stuck in local minima due to the rugged nature of the proposed objective function, and an exhaustive search can be time consuming. Moreover, the lack of identifiability for parameters of large FE models requires clustering in order to estimate the change over subsets of similar parameters, and



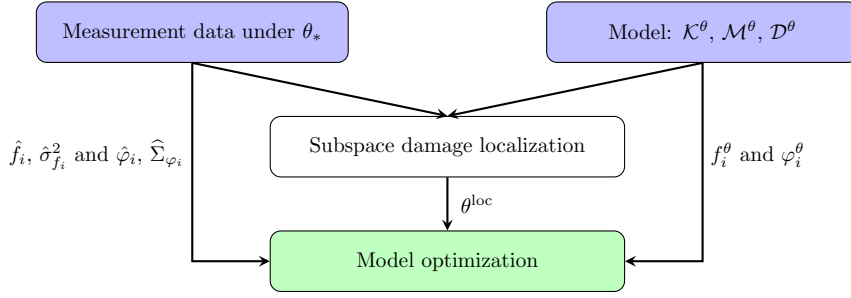


Figure 8: Damage quantification flow chart.

consequently to quantify the possible damage.

In this paper, it is proposed to chain a damage localization approach with a model optimization procedure, to minimize the designed objective function (16) over the subset of parameters that have been recognized as damaged by the localization strategy. In this way, the search domain for the optimization approach can be reduced efficiently. The resultant damage quantification framework is summarized in Figure 8. For damage localization, the subspace-based damage localization approach [7] is chosen, and the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [30] is adapted as the optimization approach. Both methods are outlined in the following subsections. The damage quantification procedure comprising the aforementioned strategies is summarized in Algorithm 2.

### 5.1. Damage localization and clustering by the subspace-based damage localization method

The subspace-based damage localization strategy has been proposed in [65] and refined in [7]. This method is based on a data-driven feature vector that compares the subspaces of output covariance Hankel matrices between the reference and damaged states. The feature vector is statistically evaluated, taking into account its covariance and its sensitivity with respect to FE model parameter vector  $\theta$ , in order to determine which components of  $\theta$  have changed for damage localization. Due to a possibly large FE model-based parametrization  $\theta$  on one side and limited measurements regarding the number of sensors and modes on the other side, the sensitivity of the feature vector with respect to some components of  $\theta$  may be equal or be very close. Thus, such parameter components are indistinguishable for damage localization, and a clustering of the parameters is performed to ensure distinguishable clusters for localization. Hereby, parameter components of  $\theta$  for which the sensitivities are equal, or are very close, form a cluster. The localization resolution is defined by the clusters, and no separation is possible between the elements of the same cluster. The elements of one cluster are all either detected as damaged, even if only one of them is actually damaged, or none of the cluster elements is recognized as damaged. In [7], it has been shown how the clusters can be defined in the reference state and a procedure for optimal clustering has been developed in [66]. The parameter clusters established for the beam example from Section 2.1 are shown in Table 3.

The output of the localization procedure is a subset of the parameter vector  $\theta$ , denoted by  $\theta^{\text{loc}}$ , which contains all parameters of the clusters that are recognized as damaged. Hence, it can be considered that the remaining parameters in  $\theta$  are unchanged. Subsequently, the damage extent can be quantified by updating only the parameter subset  $\theta^{\text{loc}}$ . Thus, linking the damage localization procedure with an adequate model optimization method should lead to a complete damage quantification strategy, where the dimension of the search domain is reduced by the localization approach.

4-13	14-19	20-23	27-30	31-42	43-69	73-76	77-83	84-89	90-92
------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Table 3: Parameter clusters for the numerical example from subspace-based damage localization.

### 5.2. Model optimization with modified CMA-ES

Starting with the initial value  $\theta = \theta_{\text{init}}$  corresponding to the model in the reference state, the CMA algorithm consists in generating  $\lambda$  model candidates  $\theta_j^g$ ,  $j = 1, \dots, \lambda$ , in each population  $g$ , by sampling a multivariate Gaussian distribution. The model candidates are only modified on the parameter subset  $\theta^{\text{loc}}$  that is given by the localization approach. The sampling is carried out on the considered parameter subset for the subsequent population  $g + 1$  as

$$(\theta_j^{\text{loc}})^{g+1} = m^g + \varepsilon_j \text{ where } \varepsilon_j \sim \sigma^g \mathcal{N}(0, C^g), \quad \text{for } j = 1, \dots, \lambda \quad (18)$$

where  $m^g$  is a weighted mean of the model candidates  $(\theta_j^{\text{loc}})^g$  in the parent generation. Then, the parameter subset of the full parameter vector  $\theta_j^{g+1}$  is updated with  $(\theta_j^{\text{loc}})^{g+1}$ . Equation (18) represents a mutation and recombination into offsprings, for which the CMA-ES algorithm adapts the parameters  $C^g$  and  $\sigma^g$  in each generation. The covariance matrix  $C^g$  of the added Gaussian noise represents the amplitude for the sampling to occur, and the scaling factor  $\sigma^g$  determines the range of the considered mutation. Consequently, the optimization continues and the best parent solutions replace the offspring until it converges to a solution. For CMA-ES, the covariance matrix  $C^g$  is incrementally updated with rank-one matrices representing the direction between the best parent solutions at two consecutive generations, such that the likelihood of previously successful search steps is increased [30].

For the convergence to a solution, a stopping criterion is defined based on the acceptance region (17). Since all models therein are equally optimal in the statistical sense, the CMA-ES algorithm can stop once a number  $t_{\text{opt}}$  of population model candidates are inside the region, where  $t_{\text{opt}} \geq 1$ . A higher value leads to more confidence in the set of retained solutions. Once inside the acceptance region, there is no need to further minimize the objective function, avoiding unnecessary additional computations. Since the acceptance region yields a hyperplane, the mean of the selected model candidates  $\theta_{j_k}^g \in \bar{\Theta}$  in the last population  $g$  is also an

acceptable solution for the final parameter value, with

$$\theta_{\text{sol}} = \frac{1}{t_{\text{opt}}} \sum_{k=1}^{t_{\text{opt}}} \theta_{j_k}^g. \quad (19)$$

Using (19) centers the solution among the model candidates satisfying the stopping criterion. It allows to ponder some possibly erroneously selected model candidates, since the considered stopping criterion is defined for a given confidence level.

Finally, the change in the parameter vector is evaluated for damage quantification. The changes in the parameter components of each cluster are indistinguishable and cannot be evaluated separately, as described in the previous section. However, the global change for each cluster  $c$  can be evaluated as

$$\hat{\delta}_c = \sum_{k=1}^{p_c} \theta_{\text{sol}}^{i_c(k)} - \theta_{\text{init}}^{i_c(k)} \quad (20)$$

where  $i_c(1), \dots, i_c(p_c)$  are the indices of the components of parameter vector  $\theta$  that correspond to cluster  $c$ , and  $p_c$  is the number of elements in  $c$ .

---

**Algorithm 2:** Model optimization

---

**Input** : frequency and mode shape estimates  $\hat{f}_i, \hat{\varphi}_i, i = 1, \dots, m$ , and (co-)variances  $\hat{\sigma}_{f_i}^2, \hat{\Sigma}_{\varphi_i}$  ;  
 physical parameter  $\theta_{\text{init}}$  and its subset  $\theta_{\text{init}}^{\text{loc}}$  corresponding to located damaged clusters ;  
 parametrized structural model  $\mathcal{K}^\theta, \mathcal{M}^\theta, \mathcal{D}^\theta$  ;  
 set of initial optimization parameters from Table 4 ;

**Output:** estimated parameter value  $\theta_{\text{sol}}^{\text{loc}}$  and estimated total change  $\hat{\delta}$  for each damaged cluster ;

- 1 compute  $c_{\hat{f}_1}, \dots, c_{\hat{f}_m}$  in (9) ;
  - 2 create  $\mathcal{K}^{\theta_{\text{init}}}, \mathcal{M}^{\theta_{\text{init}}}, \mathcal{C}^{\theta_{\text{init}}}$  and evaluate  $f_1^{\theta_{\text{init}}}, \dots, f_m^{\theta_{\text{init}}}, \psi_1^{\theta_{\text{init}}}, \dots, \psi_m^{\theta_{\text{init}}}$  ;
  - 3 compute  $t_{\text{MAC}_1}, \dots, t_{\text{MAC}_m}$  in Algorithm 1 ;
  - 4 evaluate  $\bar{F}(\theta_{\text{init}})$  in (16) ;
- }

**Model**

**evaluation**
- 5 **repeat**
  - 6 | get  $\lambda$  model candidates  $\theta_1, \dots, \theta_\lambda$  in (18) ;
  - 7 | repeat **Model evaluation** steps 2, 3 and 4 with  $\theta_1, \dots, \theta_\lambda$  ;
  - 8 | update CMA-ES optimization parameters after [30] ;
  - 9 | count the number  $t_{\bar{\Theta}}$  of model candidates  $\theta_j$  with  $\theta_j \in \bar{\Theta}$  in the acceptance region, see (17) ;
  - 10 **until**  $t_{\bar{\Theta}} \geq t_{\text{opt}}$ ;
  - 11 compute  $\theta_{\text{sol}}$  in (19) as the mean of all model candidates in  $\bar{\Theta}$  ;
  - 12 compute the change in the parameter value  $\hat{\delta}_c$  for each damaged cluster  $c$  in (20)
-

## 6. Numerical application

This section is devoted to the application of the proposed scheme to quantify damage exerted in the numerical beam example introduced in Section 2.1, for which simultaneous faults in parameters  $\rho_9$  and  $\rho_{33}$  are investigated. Note that these parameters belong to different clusters, as shown in Table 3.

First, the damage case 2 comprising 15% reduction in both  $\rho_9$  and  $\rho_{33}$  is analyzed. Second, five cases consisting of different fault combinations of 2.5%, 5%, 10% and 15% damage in  $\rho_9$  and  $\rho_{33}$  are examined. The beam data are simulated after the procedure in Section 2.1, and the natural frequencies  $\hat{f}_1, \dots, \hat{f}_{12}$  and mode shapes  $\hat{\varphi}_1, \dots, \hat{\varphi}_{12}$  together with their corresponding covariances are computed with the SSI-UPC algorithm [62]. The fully parametrized beam model comprises 80 parameters (excluding the junction elements in Figure 1), whose updating can involve a considerable number of iterations and may not be efficient for damage quantification in practice. To reduce the search parameter space and focus the optimization algorithm on the potentially damaged parameters, the subspace-based damage localization is applied.

The Fisher information related to the subspace damage localization residual is used for complete linkage parameter clustering [7], and the resultant parameter clusters are shown in Table 3. After the subspace-based localization is applied, both clusters containing the faulty parameter components, namely the first and the fifth cluster of Table 3, are detected. Both clusters contain 22 parameters, which are used instead of the initial 80 parameters in the optimization procedure. The parameters used to initialize the optimization algorithm are depicted in Table 4. The first model candidate arrives in the acceptance region after 38 iterations in 550 seconds in Matlab 2018b on a i7 2.9GHz CPU with 16GB of RAM. The whole model optimization takes 56 iterations to converge to a solution with 60% of the population in the acceptance region, lasting 880 seconds. The model candidates retained in the acceptance region are visualized in the modal parameter space for the first and the twelfth mode in Figure 9. It can be viewed that in both cases the difference between the estimated and the model natural frequencies are very small, and the corresponding MAC values are approximately 1 for the first mode and higher than 0.99 for the twelfth mode. The parameter value that corresponds to the exact model is within both regions. Furthermore, it can be observed that the accepted MAC and frequency pairs create an elliptic surface when projected into a two-dimensional modal parameter space, where the ellipse is the 95% confidence region of the accepted samples, as illustrated in Figure 9. The shape of the ellipse signifies the boundary of the acceptance region and differs for different modes. Notice that those plots are just a slice of the multidimensional parameter search. As the clusters of

Table 4: CMA-ES optimization parameters.

$\sigma_{\text{init}}$	$\theta_{\text{init}}^i$	$\theta_{\text{min-max}}^i$	$\lambda$	$t_{\text{opt}}$
810	2700	2295-3510	1000	60% $\lambda$

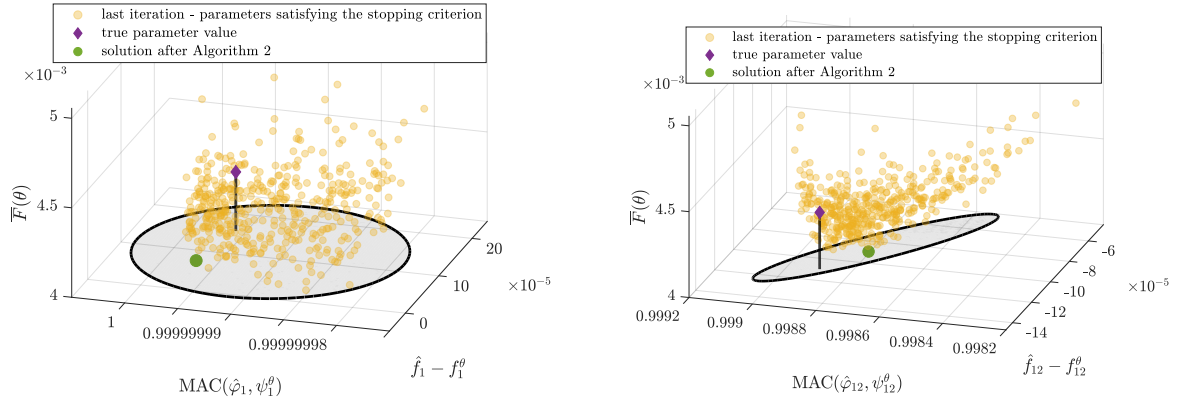


Figure 9: Natural frequency and MAC differences at the last iteration: first mode (left), twelfth mode (right).

interest have altogether 22 parameter components, the search is over 22 dimensions, which is significantly less than the original parameter size.

Prior to damage quantification, the statistical uncertainty of the MAC estimates is examined in more detail. The uncertainty of the MAC has been derived and incorporated into the objective function in this paper, and it takes an important role in the definition of the acceptance region. To validate the derived statistical properties of the MAC, a Monte Carlo experiment with 1000 simulations of the measurement data and the optimization procedure from Figure 8 is considered for each simulation. Two scenarios for the MAC computation are distinguished. First, the MAC is computed between the mode shape estimate  $\hat{\varphi}_i$  of each simulation and the true numerical mode shape  $\psi_{i,*}$ . Second, the MAC is computed between the mode shape estimate  $\hat{\varphi}_i$  and the numerical mode shape  $\psi_i^{\text{sol}}$  corresponding to the mean solution of the optimization algorithm in each simulation. The respective histograms for the first and the twelfth mode shape are shown in Figure 10 together with a MAC distribution fit, whose parameters are computed for one realization of the Monte Carlo simulation with Algorithm 1. These results illustrate that the distribution of the MAC estimates is well approximated by the scaled and shifted  $\chi^2$  distribution, as developed in Appendix A.

The practical application of the statistical analysis is the estimation of the corresponding q-intervals based on the quantile  $t_{\text{MAC}_i}$  in (14), to infer if the MAC between a mode shape estimate and a numerical mode shape is significantly close to 1 in the optimization with the chosen confidence level of  $\gamma = 0.95$ . The computation of these intervals can be validated in the Monte Carlo experiment, where 95% of the intervals  $[t_{\text{MAC}_i}, 1]$  should contain the respective MAC estimate when the mode shapes match. The ratios of occurrences where the estimates of MAC are indeed in the interval are shown in Table 5. It can be viewed that approximately 95% of all MAC estimates are contained in the interval  $[t_{\text{MAC}_i}, 1]$ , showing that the proposed uncertainty quantification scheme is accurate. Even if the displayed percentages are slightly higher than expected, they are within 1% of each other when the estimated mode shapes are compared to the mode shape of either the true model or the result of the optimization. With these results, the statistical

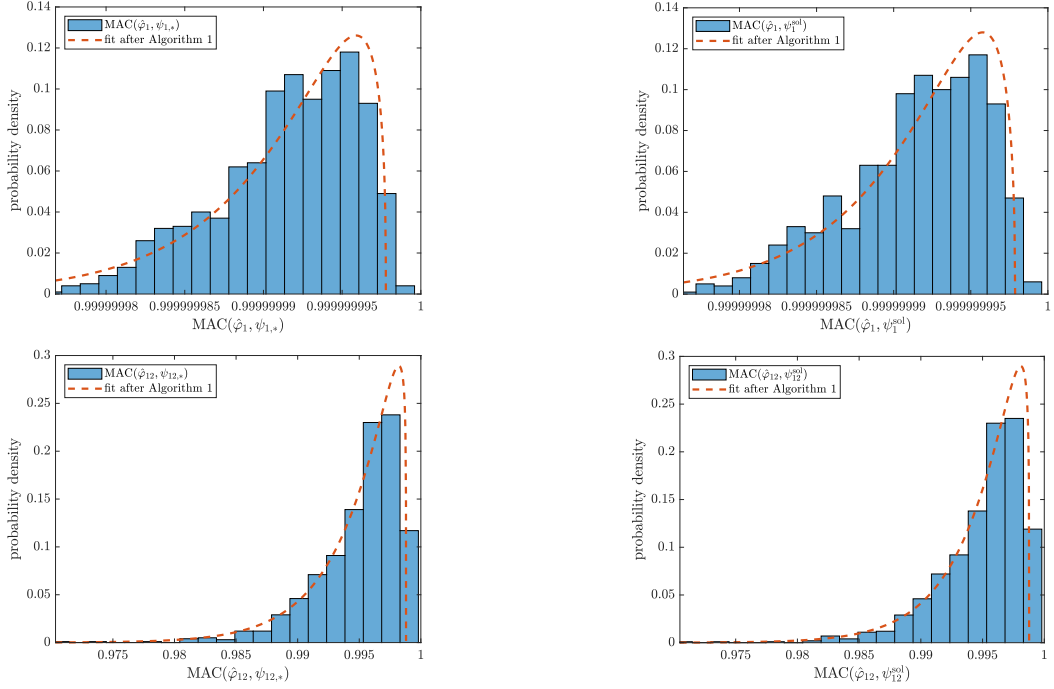


Figure 10: MAC distribution fits to histograms of  $\text{MAC}(\hat{\varphi}_1, \psi_{*,1})$  (top left),  $\text{MAC}(\hat{\varphi}_1, \psi_1^{\text{sol}})$  (top right),  $\text{MAC}(\hat{\varphi}_{12}, \psi_{*,12})$  (bottom left) and  $\text{MAC}(\hat{\varphi}_{12}, \psi_{12}^{\text{sol}})$  (bottom right).

Table 5: Percentage of MAC estimates contained in the 95% q-intervals of the MAC distribution approximation.

Mode $i$	1	2	3	4	5	6	7	8	9	10	11	12
$\text{MAC}(\hat{\varphi}_i, \psi_{i,*})$	96.4%	97.8%	97.4%	96.2%	97.4%	96.8%	97.4%	97.3%	97.0%	97.8%	97.4%	96.7%
$\text{MAC}(\hat{\varphi}_i, \psi_i^{\text{sol}})$	97.5%	98.7%	98.5%	97.0%	98.1%	97.9%	98.3%	98.0%	97.9%	98.5%	98.1%	97.4%

properties of the MAC estimates are validated for use in the proposed optimization procedure.

The goal of the damage quantification approach is to estimate the change in the parameters of each damaged cluster. This change is expressed as the sum of the differences between the considered parameter components of  $\theta_{\text{sol}}$  and their initial values  $\theta_{\text{init}}$  in the cluster in (20). In the iterations of the optimization procedure, each population contains  $\lambda$  model candidates, from which  $t_{\text{opt}}$  are required to be within the acceptance region for the algorithm to stop. The mean of these candidates is defined as  $\theta_{\text{sol}}$ . The estimated change can be examined for every population of model candidates at each iteration of the algorithm, illustrating its performance, as depicted in Figure 11. It can be seen that the estimated change reaches the lowest values of the objective function for the parameters inside the acceptance region. The estimated cluster changes corresponding to the mean solution  $\theta_{\text{sol}}$  and the true value  $\theta_*$  are within the envelope of the cluster changes corresponding to the accepted models of the final population. The final estimates based on  $\theta_{\text{sol}}$  as shown in Figure 11 (green circle) are  $\hat{\delta}_{4-13} = 423.8$  and  $\hat{\delta}_{31-42} = 390.2$ . Their standard deviations

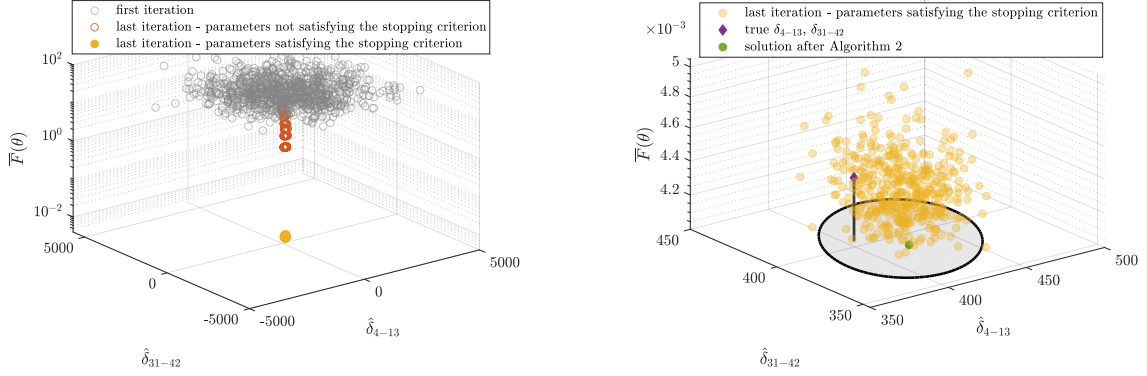


Figure 11: CMA-ES optimization of the change in parameter clusters (left) with zoom on the cluster changes corresponding to the acceptance region in the last iteration (right).

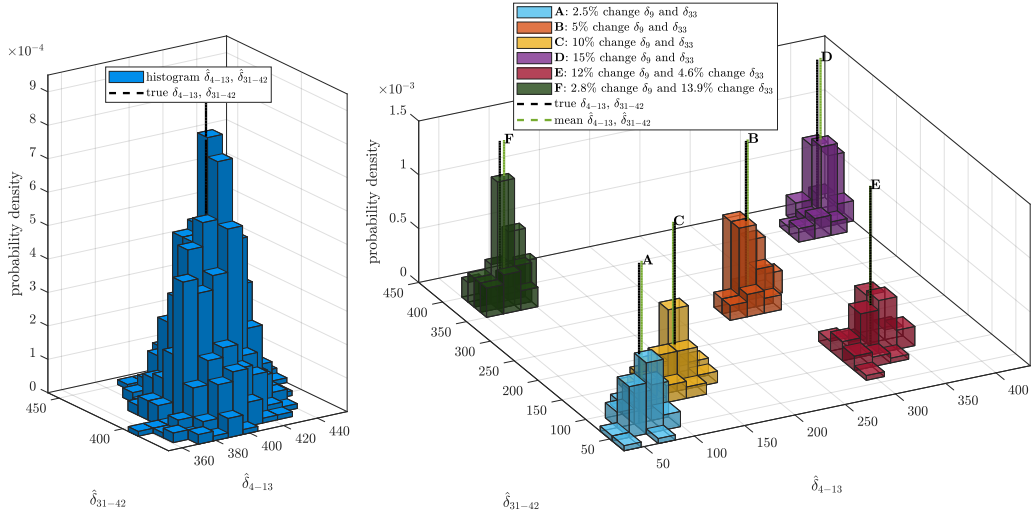


Figure 12: Damage extent estimation for different scenarios.

originating from the final population of the optimization algorithm are 19.3 and 15.7, respectively, where the true value of the change is 405 in both cases.

Finally, the variability of  $\hat{\delta}_{4-13}$  and  $\hat{\delta}_{31-42}$  is illustrated when estimated on different data sets in the Monte Carlo experiment. This variability stems from the statistical uncertainty of the modal parameter estimates and is shown in a bivariate histogram of  $\hat{\delta}_{4-13}$  and  $\hat{\delta}_{31-42}$  for 1000 simulations of damage case 2 in Figure 12 (left). Furthermore, histograms of progressive damage built from 50 simulations are shown in Figure 12 (right). These histograms indicate that the proposed damage quantification approach yields accurate estimates of the change in the considered parameter clusters, regardless of the varying fault extent. It also can be seen that the true change is always contained inside the support of each histogram.

## 7. Conclusions

In this paper, a model optimization framework has been proposed to quantify the extent of structural damage. The framework takes into account the uncertainty of the estimated modal parameters by incorporating it in both the design of the objective function and the stopping criterion of the optimization procedure. This has been made possible by the evaluation of the statistical properties of the MAC between the mode shapes from the numerical model and the mode shape estimates from measurements. An improved optimization procedure has then been designed whose outcome is coherent with the data-based uncertainties. Finally, the potentially large dimension of the model parameter space has been addressed by linking statistical clustering and localization to focus the model optimization on a subset of potentially damaged elements for damage quantification. The proposed framework has been validated on a numerical simulation of a beam, where the damage extent of the damaged clusters has been correctly quantified for multiple damage scenarios. The framework has been used with the subspace-based damage localization method and CMA-ES for the optimization, but other methods can be used similarly. While in this work data-based uncertainties have been addressed, future work should include the handling of uncertainty due to modeling errors for the proposed damage quantification framework.

## Appendix A. Approximation of $\text{MAC}(\hat{\varphi}, \psi_*)$ distribution

### Appendix A.1. Expression of the quadratic form

A classic approach to quantify the statistical uncertainty of estimated parameters is to propagate the sample covariance of measurement-related quantities (like output covariances) onto computed parameters based on their sensitivities. Their distribution function is often inferred with the first-order delta method, which allows to characterize the probability distribution of a function of a Gaussian variable also as Gaussian. Consider two complex-valued mode shapes  $\varphi_*$  and  $\psi_*$  that are possibly equal. A consistent estimate  $\hat{\varphi}$  of  $\varphi_*$  is obtained from measurements, and it is assumed that its uncertainty is already quantified, e.g., by subspace identification methods, with

$$(\hat{\varphi} - \varphi_*)_{\text{Re}} = \begin{bmatrix} \Re(\hat{\varphi} - \varphi_*) \\ \Im(\hat{\varphi} - \varphi_*) \end{bmatrix} \approx \mathcal{N}\left(0, \frac{1}{N}\Sigma_{\varphi_*}\right), \quad (\text{A.1})$$

cf. (4). Consider  $\text{MAC}(\hat{\varphi}, \psi_*)$  and its derivative w.r.t. the first variable while the second variable is constant

$$\mathcal{J}_{\varphi_*}^{\text{MAC}} = \begin{bmatrix} \frac{\partial \text{MAC}}{\partial \Re(\varphi)} \Big|_{\varphi=\varphi_*} \\ \frac{\partial \text{MAC}}{\partial \Im(\varphi)} \Big|_{\varphi=\varphi_*} \end{bmatrix} = \begin{bmatrix} \frac{2\Re(\psi_* \psi_*^H \varphi_*)^T}{\varphi_*^H \varphi_* \psi_*^H \psi_*} - \frac{2\Re(\varphi_*)^T \text{MAC}(\varphi_*, \psi_*)}{\varphi_*^H \varphi_*} \\ \frac{2\Im(\psi_* \psi_*^H \varphi_*)^T}{\varphi_*^H \varphi_* \psi_*^H \psi_*} - \frac{2\Im(\varphi_*)^T \text{MAC}(\varphi_*, \psi_*)}{\varphi_*^H \varphi_*} \end{bmatrix} \quad (\text{A.2})$$

following from (5). Then, based on the first-order Taylor expansion

$$\text{MAC}(\hat{\varphi}, \psi_*) = \text{MAC}(\varphi_*, \psi_*) + \mathcal{J}_{\varphi_*}^{\text{MAC}}(\hat{\varphi} - \varphi_*)_{\text{Re}} + o(\|\hat{\varphi} - \varphi_*\|), \quad (\text{A.3})$$



the first-order delta method [55] states that for  $\text{MAC}(\hat{\varphi}, \psi_*)$  it holds

$$(\text{MAC}(\hat{\varphi}, \psi_*) - \text{MAC}(\varphi_*, \psi_*)) \approx \mathcal{N}\left(0, \frac{1}{N} \sigma_{\text{MAC}}^2\right), \quad (\text{A.4})$$

where  $\sigma_{\text{MAC}}^2 = \mathcal{J}_{\varphi_*}^{\text{MAC}} \Sigma_{\varphi_*} (\mathcal{J}_{\varphi_*}^{\text{MAC}})^T$ . This approximation of the asymptotic MAC distribution is only valid if  $\mathcal{J}_{\varphi_*}^{\text{MAC}} \neq 0$ . However, when the limit of  $\hat{\varphi}$  coincides with  $\psi_*$  then  $\text{MAC}(\varphi_*, \psi_*) = 1$ , and it can easily be shown that the MAC derivative  $\mathcal{J}_{\varphi_*}^{\text{MAC}}$  in (A.2) is zero if and only if  $\text{MAC}(\varphi_*, \psi_*) = 1$  (or 0, which is not of interest here). In this case, the first-order delta method for approximating the distribution from the covariance of  $\text{MAC}(\hat{\varphi}, \psi_*)$  cannot be used, since the linearization by the first-order Taylor expansion from (A.3) is insufficient. Then, a second-order Taylor expansion yields

$$\text{MAC}(\hat{\varphi}, \psi_*) \approx \underbrace{\text{MAC}(\varphi_*, \psi_*)}_{=1} + \underbrace{\mathcal{J}_{\varphi_*}^{\text{MAC}}}_{=0} (\hat{\varphi} - \varphi_*)_{\text{Re}} + \frac{1}{2} (\hat{\varphi} - \varphi_*)_{\text{Re}}^T \mathbf{H}_{\varphi_*}^{\text{MAC}} (\hat{\varphi} - \varphi_*)_{\text{Re}}, \quad (\text{A.5})$$

where  $\mathbf{H}_{\varphi_*}^{\text{MAC}} \in \mathbb{R}^{2r \times 2r}$  is the Hessian, i.e., the second derivative of  $\text{MAC}(\varphi, \psi)$  as detailed in the next section. Let  $\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}} = -\frac{1}{2} \mathbf{H}_{\varphi_*}^{\text{MAC}}$ , then the asymptotic properties of  $\text{MAC}(\hat{\varphi}, \psi_*)$  follow from (A.5) with

$$N(1 - \text{MAC}(\hat{\varphi}, \psi_*)) \approx \sqrt{N} (\hat{\varphi} - \varphi_*)_{\text{Re}}^T \bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}} \sqrt{N} (\hat{\varphi} - \varphi_*)_{\text{Re}}, \quad (\text{A.6})$$

which is a quadratic form of the Gaussian variable  $\sqrt{N} (\hat{\varphi} - \varphi_*)_{\text{Re}}$ . 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 [67] can be applied when the matrix  $\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}}$  is positive semi-definite. With this approximation, the statistical properties of the quadratic form can be related to the asymptotic covariance  $\Sigma_{\varphi_*}$  in (A.1), and subsequently the pdf of the MAC estimate can be retrieved through the relation (A.6) as detailed in the next section.

#### Appendix A.2. Approximation of the quadratic form

The quadratic form in (A.6) is characterized by its inner matrix  $\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}}$ , which is related to the Hessian of the MAC. Since the MAC distribution is developed under the assumption that  $\varphi_*$  and  $\psi_*$  are corresponding mode shapes, they are collinear with

$$\varphi_* = k \cdot \psi_* \quad (\text{A.7})$$

for some  $k \in \mathbb{C}$ . Taking the derivatives of (A.2) and plugging in (A.7) leads to the Hessian

$$\mathbf{H}_{\varphi_*}^{\text{MAC}} = \frac{2}{|k|^2 d_*^2} \begin{bmatrix} M_{xx}^* + M_{yy}^* - d_* I_r & M_{xy}^* - M_{yx}^* \\ M_{yx}^* - M_{xy}^* & M_{xx}^* + M_{yy}^* - d_* I_r \end{bmatrix}, \quad (\text{A.8})$$

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  $d_* = \psi_*^H \psi_*$ . At the end of this section it is shown that  $\mathbf{H}_{\varphi_*}^{\text{MAC}}$  is negative semi-definite, and thus  $\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}} = -\frac{1}{2} \mathbf{H}_{\varphi_*}^{\text{MAC}}$  is positive semi-definite, so the distribution approximation of the quadratic form from [67] can be applied.

Denote the first three cumulants of the quadratic form (A.6) as  $c_1$ ,  $c_2$  and  $c_3$ . They satisfy [67]

$$c_j = \text{tr} \left( (\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}} \Sigma_{\varphi_*})^j \right), \quad j = 1, 2, 3. \quad (\text{A.9})$$

The approximate pdf  $f_Q$  of the quadratic form is then linked to the  $\chi^2$  distribution and writes [67]

$$f_Q(x) = \frac{1}{\alpha} f_{\chi_l^2} \left( \frac{x - \beta}{\alpha} \right), \quad x \in [\beta, +\infty), \quad (\text{A.10})$$

and  $f_Q(x) = 0$  for  $x < \beta$ , where  $\alpha = c_3/c_2$  and  $\beta = c_1 - c_2^2/c_3$  are scaling and shift parameters, and  $l = c_2^3/c_3^2$  is the number of DOF. Note that  $l$  is in general not an integer. Strictly speaking,  $f_{\chi_l^2}$  is always the pdf of a Gamma distribution, generalizing the  $\chi^2$  distribution family for non-integer values of  $l$ . For simplicity, this distribution is still denoted as  $\chi^2$ . By (A.10), the distribution of  $\text{MAC}(\hat{\varphi}, \psi_*)$  follows immediately based on relation (A.6), it yields the pdf

$$f_{\text{MAC}}(x) = \frac{N}{\alpha} f_{\chi_l^2} \left( \frac{N(1-x) - \beta}{\alpha} \right), \quad x \in (-\infty, 1 - \frac{\beta}{N}]. \quad (\text{A.11})$$

The MAC distribution is hence approximated by a scaled and shifted  $\chi^2$  distribution under the assumption that  $\text{MAC}(\varphi_*, \psi_*) = 1$ . For the computation of the distribution parameters  $l$ ,  $\alpha$  and  $\beta$  of the  $\chi^2$  approximation, estimates of the mode shape covariance  $\Sigma_{\varphi_*}$  and of the Hessian  $\mathbf{H}_{\varphi_*}^{\text{MAC}}$  are required for (A.9). The mode shape covariance estimate  $\hat{\Sigma}_{\varphi_*}$  is obtained from the system identification, and an estimate  $\mathbf{H}_{\hat{\varphi}}^{\text{MAC}}$  of the Hessian is obtained by evaluating (A.8) with  $\hat{\varphi}$  and  $\hat{k} = \psi_*^H \hat{\varphi} / \psi_*^H \psi_*$  based on (A.7).

To finalize this section, it is proved that  $\mathbf{H}_{\varphi_*}^{\text{MAC}}$  is negative semi-definite and thus  $\bar{\mathbf{H}}_{\varphi_*}^{\text{MAC}} = -\frac{1}{2} \mathbf{H}_{\varphi_*}^{\text{MAC}}$  is positive semi-definite, which is a requirement for the applicability of the approximation from [67]. From (A.8) follows

$$\begin{aligned} \frac{|k|^2 d_*}{2} \mathbf{H}_{\varphi_*}^{\text{MAC}} &= \frac{1}{d_*} \begin{bmatrix} M_{xx}^* + M_{yy}^* & M_{xy}^* - M_{yx}^* \\ M_{yx}^* - M_{xy}^* & M_{xx}^* + M_{yy}^* \end{bmatrix} - I_{2r} \\ &= LL^T - I_{2r}, \quad \text{where } L = \frac{1}{\sqrt{d_*}} \begin{bmatrix} \Re(\psi_*) & -\Im(\psi_*) \\ \Im(\psi_*) & \Re(\psi_*) \end{bmatrix} \in \mathbb{R}^{2r \times 2}. \end{aligned} \quad (\text{A.12})$$

Notice that the norm of both columns of  $L = \begin{bmatrix} L_1 & L_2 \end{bmatrix}$  is 1 since  $\sqrt{d_*} = \sqrt{\Re(\psi_*)^T \Re(\psi_*) + \Im(\psi_*)^T \Im(\psi_*)}$ , and they are orthogonal. There exist  $2r - 2$  linearly independent vectors  $q_1, \dots, q_{2r-2} \in \mathbb{R}^{2r}$  such that matrix  $\tilde{L} = \begin{bmatrix} L_1 & L_2 & q_1 & \dots & q_{2r-2} \end{bmatrix} \in \mathbb{R}^{2r \times 2r}$  is orthogonal, i.e.,  $\tilde{L}^T \tilde{L} = \tilde{L} \tilde{L}^T = I_{2r}$ . Then, it follows from (A.12)

$$\begin{aligned} \frac{|k|^2 d_*}{2} \mathbf{H}_{\varphi_*}^{\text{MAC}} &= \tilde{L} \text{diag}(1, 1, 0, \dots, 0) \tilde{L}^T - \tilde{L} \tilde{L}^T \\ &= \tilde{L} \text{diag}(0, 0, -1, \dots, -1) \tilde{L}^T \end{aligned}$$

which is an eigenvalue decomposition. Since all eigenvalues are non-positive and  $\frac{|k|^2 d_*}{2} > 0$ , it follows that  $\mathbf{H}_{\varphi_*}^{\text{MAC}}$  is negative semi-definite, which concludes the proof.

## References

- [1] A. Rytter, Vibrational based inspection of civil engineering structures, Ph.D. thesis, Aalborg University, Denmark (1993).
- [2] M. Döhler, L. Mevel, F. Hille, Subspace-based damage detection under changes in the ambient excitation statistics, *Mechanical Systems and Signal Processing* 45 (1) (2014) 207 – 224.
- [3] M. D. Ulriksen, D. Tcherniak, P. H. Kirkegaard, L. Damkilde, Operational modal analysis and wavelet transformation for damage identification in wind turbine blades, *Structural Health Monitoring* 15 (4) (2016) 381–388.
- [4] K. Worden, G. Manson, N. R. J. Fieller, Damage detection using outlier analysis, *Journal of Sound and Vibration* 229 (3) (2000) 647 – 667.
- [5] D. Bernal, Kalman filter damage detection in the presence of changing process and measurement noise, *Mechanical Systems and Signal Processing* 39 (1) (2013) 361 – 371.
- [6] A.-M. Yan, J.-C. Golinval, Null subspace-based damage detection of structures using vibration measurements, *Mechanical Systems and Signal Processing* 20 (3) (2006) 611 – 626.
- [7] S. Allahdadian, M. Döhler, C. Ventura, L. Mevel, Towards robust statistical damage localization via model-based sensitivity clustering, *Mechanical Systems and Signal Processing* 134 (2019) 106341.
- [8] D. Bernal, M. D. Ulriksen, Subspace exclusion zones for damage localization, *Mechanical Systems and Signal Processing* 114 (2019) 120 – 127.
- [9] M. D. Ulriksen, D. Bernal, L. Damkilde, Shaped input distributions for structural damage localization, *Mechanical Systems and Signal Processing* 110 (2018) 499 – 508.
- [10] M. D. Ulriksen, L. Damkilde, Structural damage localization by outlier analysis of signal-processed mode shapes – analytical and experimental validation, *Mechanical Systems and Signal Processing* 68-69 (2016) 1 – 14.
- [11] M. Döhler, L. Mevel, Q. Zhang, Fault detection, isolation and quantification from gaussian residuals with application to structural damage diagnosis, *Annual Reviews in Control* 42 (2016) 244–256.
- [12] N. V. Hà, J.-C. Golinval, Localization and quantification of damage in beam-like structures using sensitivities of principal component analysis results, *Mechanical Systems and Signal Processing* 24 (6) (2010) 1831 – 1843.
- [13] A. Yan, J.-C. Golinval, Structural damage localization by combining flexibility and stiffness methods, *Engineering Structures* 27 (12) (2005) 1752–1761.
- [14] J. Mottershead, M. Friswell, Model updating in structural dynamics: A survey, *Journal of Sound and Vibration* 167 (2) (1993) 347 – 375.
- [15] M. Friswell, J. E. Mottershead, *Finite element model updating in structural dynamics*, Vol. 38, Springer Science & Business Media, 2013.
- [16] I. Behmanesh, B. Moaveni, C. Papadimitriou, Probabilistic damage identification of a designed 9-story building using modal data in the presence of modeling errors, *Engineering Structures* 131 (2017) 542 – 552.
- [17] O. Huth, G. Feltrin, J. Maeck, N. Kilic, M. Motavalli, Damage identification using modal data: Experiences on a prestressed concrete bridge, *Journal of Structural Engineering* 131 (12) (2005) 1898–1910.
- [18] E. Reynders, G. D. Roeck, P. G. Bakir, C. Sauvage, Damage identification on the tilff bridge by vibration monitoring using optical fiber strain sensors, *Journal of Engineering Mechanics* 133 (2) (2007) 185–193.
- [19] A. Teughels, G. D. Roeck, Damage detection and parameter identification by finite element model updating, *Revue Européenne de Génie Civil* 9 (1-2) (2005) 109–158.
- [20] 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.
- [21] G. Steenackers, P. Guillaume, Finite element model updating taking into account the uncertainty on the modal parameters estimates, *Journal of Sound and Vibration* 296 (4) (2006) 919 – 934.

- [22] H. B. Başıağ, T. Türker, A. Bayraktar, A model updating approach based on design points for unknown structural parameters, *Applied Mathematical Modelling* 35 (12) (2011) 5872–5883.
- [23] K. Christodoulou, C. Papadimitriou, Structural identification based on optimally weighted modal residuals, *Mechanical Systems and Signal Processing* 21 (1) (2007) 4 – 23.
- [24] H. Shahverdi, C. Mares, W. Wang, J. Mottershead, Clustering of parameter sensitivities: Examples from a helicopter airframe model updating exercise, *Shock and Vibration* 16 (01 2009).
- [25] J. Jang, A. Smyth, Bayesian model updating of a full-scale finite element model with sensitivity-based clustering, *Structural Control and Health Monitoring* 24 (11) (2017) e2004.
- [26] J. Jang, A. W. Smyth, Model updating of a full-scale fe model with nonlinear constraint equations and sensitivity-based cluster analysis for updating parameters, *Mechanical Systems and Signal Processing* 83 (2017) 337 – 355.
- [27] D. T. Bartilson, J. Jang, A. W. Smyth, Finite element model updating using objective-consistent sensitivity-based parameter clustering and bayesian regularization, *Mechanical Systems and Signal Processing* 114 (2019) 328 – 345.
- [28] G. Gautier, L. Mevel, J.-M. Mencik, R. Serra, M. Döhler, Variance analysis for model updating with a finite element based subspace fitting approach, *Mechanical Systems and Signal Processing* 91 (2017) 142 – 156.
- [29] G. Gautier, J.-M. Mencik, R. Serra, A finite element-based subspace fitting approach for structure identification and damage localization, *Mechanical Systems and Signal Processing* 58 (2015) 143–159.
- [30] N. Hansen, A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, *Evolutionary Computation* 9 (2) (2001) 159–195.
- [31] K. Agathos, E. Chatzi, S. Bordas, Multiple crack detection in 3D using a stable XFEM and global optimization, *Computational Mechanics* 62 (02 2018).
- [32] N. Hansen, A. Ostermeier, Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The  $(\mu/\mu_i, \lambda)$ -CMA-ES, *EUFIT (1997)* 650–654.
- [33] A. Auger, N. Hansen, A restart cma evolution strategy with increasing population size, in: *IEEE Congress on Evolutionary Computation*, Vol. 2, IEEE, 2005, pp. 1769–1776.
- [34] N. Hansen, R. Ros, N. Mauny, M. Schoenauer, A. Auger, Impacts of invariance in search: When cma-es and pso face ill-conditioned and non-separable problems, *Applied Soft Computing* 11 (8) (2011) 5755–5769.
- [35] A. Chaudhuri, R. T. Haftka, A stopping criterion for surrogate based optimization using EGO, in: *10th World Congress on Structural and Multidisciplinary Optimization*, 2013, pp. 20–24.
- [36] K. Zielinski, R. Laur, Stopping criteria for differential evolution in constrained single-objective optimization, in: *Advances in differential evolution*, Springer, 2008, pp. 111–138.
- [37] E. Simoen, G. De Roeck, G. Lombaert, Dealing with uncertainty in model updating for damage assessment: A review, *Mechanical Systems and Signal Processing* 56 (2015) 123–149.
- [38] 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.
- [39] J. D. Collins, G. C. Hart, T. K. Hasselman, B. Kennedy, Statistical identification of structures, *AIAA Journal* 12 (2) (1974) 185–190.
- [40] M. Friswell, The adjustment of structural parameters using a minimum variance estimator, *Mechanical Systems and Signal Processing* 3 (2) (1989) 143 – 155.
- [41] X. G. Hua, Y. Q. Ni, Z. Q. Chen, J. M. Ko, An improved perturbation method for stochastic finite element model updating, *International Journal for Numerical Methods in Engineering* 73 (13) (2008) 1845–1864.
- [42] H. H. Khodaparast, J. E. Mottershead, M. I. Friswell, Perturbation methods for the estimation of parameter variability in stochastic model updating, *Mechanical Systems and Signal Processing* 22 (8) (2008) 1751 – 1773.
- [43] N. A. Husain, H. H. Khodaparast, H. Ouyang, Parameter selection and stochastic model updating using perturbation

- methods with parameter weighting matrix assignment, *Mechanical Systems and Signal Processing* 32 (2012) 135 – 152, uncertainties in Structural Dynamics.
- [44] E. Viefhues, M. Döhler, Q. Zhang, F. Hille, L. Mevel, Subspace-based damage detection with rejection of the temperature effect and uncertainty in the reference, in: *IOMAC 2019 - 8th International Operational Modal Analysis Conference*, Copenhagen, Denmark, 2019, pp. 1–11.
- [45] E. J. Cross, K. Worden, Q. Chen, Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 467 (2133) (2011) 2712–2732.
- [46] B. Qadri, M. Ulriksen, L. Damkilde, D. Tcherniak, Cointegration for detecting structural blade damage in an operating wind turbine: An experimental study, Vol. 2 of *Conference Proceedings of the Society for Experimental Mechanics Series*, Springer, Germany, 2020, pp. 173–180.
- [47] H. Shi, K. Worden, E. J. Cross, A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring, *Mechanical Systems and Signal Processing* 103 (2018) 381–397.
- [48] L. D. Avendaño-Valencia, E. N. Chatzi, D. Tcherniak, Gaussian process models for mitigation of operational variability in the structural health monitoring of wind turbines, *Mechanical Systems and Signal Processing* 142 (2020).
- [49] L. D. Avendaño-Valencia, E. N. Chatzi, Multivariate GP-VAR models for robust structural identification under operational variability, *Probabilistic Engineering Mechanics* 60 (2020).
- [50] C. Argyris, C. Papadimitriou, P. Panetsos, P. Tsopelas, Bayesian model-updating using features of modal data: Application to the metsovo bridge, *Journal of Sensor and Actuator Networks* 9 (2) (2020) 27.
- [51] 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) 455–459.
- [52] 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.
- [53] 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.
- [54] D. Hogben, The distribution of the sample correlation coefficient with one variable fixed, *Journal of Research of the National Bureau of Standards - B. Mathematical Sciences* 72B (1) (1967).
- [55] G. Casella, R. L. Berger, *Statistical inference*, 2nd Edition, Cengage Learning, 2001.
- [56] M. Döhler, L. Mevel, Efficient multi-order uncertainty computation for stochastic subspace identification, *Mechanical Systems and Signal Processing* 38 (2) (2013) 346–366.
- [57] S. Greś, M. Döhler, P. Andersen, L. Mevel, Uncertainty quantification for the Modal Phase Collinearity of complex mode shapes, *Mechanical Systems and Signal Processing* 152 (2021) 107436.
- [58] S. Greś, M. Döhler, L. Mevel, Uncertainty quantification of the Modal Assurance Criterion in operational modal analysis, *Mechanical Systems and Signal Processing* 152 (2021) 107457.
- [59] M. Döhler, L. Mevel, Fast multi-order computation of system matrices in subspace-based system identification, *Control Engineering Practice* 20 (9) (2012) 882 – 894.
- [60] 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.
- [61] 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.
- [62] 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.
- [63] R. J. Allemang, The Modal Assurance Criterion (MAC): Twenty years of use and abuse, *Sound and Vibration* 37 (8)

(2003) 14–21.

- [64] E. Reynders, J. Houbrechts, G. De Roeck, Fully automated (operational) modal analysis, *Mechanical Systems and Signal Processing* 29 (2012) 228 – 250.
- [65] E. Balmès, M. Basseville, L. Mevel, H. Nasser, W. Zhou, Statistical model-based damage localization: A combined subspace-based and substructuring approach, *Structural Control and Health Monitoring* 15 (6) (2008) 857–875.
- [66] A. Mendler, M. Döhler, C. Ventura, L. Mevel, Clustering of redundant parameters for fault isolation with Gaussian residuals, in: 21st IFAC World Congress, 2020.
- [67] 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.