AN OPTIMAL SENSOR PLACEMENT METHOD FOR SHM BASED ON BAYESIAN EXPERIMENTAL DESIGN AND POLYNOMIAL CHAOS EXPANSION

Giovanni Capellari[1], Eleni Chatzi[2], and Stefano Mariani[1]

[1]Politecnico di Milano, Dipartimento di Ingegneria Civile e Ambientale
Piazza Leonardo da Vinci, 32, 20133 Milano (IT)
e-mail: giovanni.capellari, stefano.mariani@polimi.it

[2] ETH Zürich, Institut für Baustatik und Konstruktion
Stefano-Franscini-Platz, 5, 8093 Zürich (CH)
e-mail: chatzi@ibk.baug.ethz.ch

Keywords: Structural Health Monitoring, Damage Detection, Optimal Sensor Placement, Bayesian Inference, Uncertainty Quantification.

Abstract. *We present an optimal sensor placement methodology for structural health monitoring (SHM) purposes, relying on a Bayesian experimental design approach. The unknown structural properties, e.g. the residual strength and stiffness, are inferred from data collected through a network of sensors, whose architecture, i.e., type and position may largely affect the accuracy of the monitoring system. In tackling this issue, an optimal network configuration is herein sought by maximizing the expected information gain between prior and posterior probability distributions of the parameters to be estimated. Since the objective function linked to the network topology cannot be analytically computed, a numerical approximation is provided by means of a Monte Carlo analysis, wherein each realization is obtained via finite element modeling. Since the computational burden linked to this procedure often grows infeasible, a Polynomial Chaos Expansion (PCE) approach is adopted for accelerating the computation of the forward problem. The analysis expands over joint samples covering both structural state and design variables, i.e., sensor locations. Via increase of the number of deployed sensors in the network, the optimization procedure soon turns computationally costly due to the curse of dimensionality. To this end, a stochastic optimization method is adopted for accelerating the convergence of the optimization process and thereby the damage detection capability of the SHM system. The proposed method is applied to thin flexible structures, and the resulting optimal sensor configuration is shown. The effects of the number of training samples, the polynomial degree of the approximation expansion and the optimization settings are also discussed.*

1 INTRODUCTION

Let us consider a structural system, monitored through a network of sensors. The measured response of the system subjected to a set of given actions is collected in the vector $\mathbf{y} \in \mathbb{R}^{n_\mathbf{y}}$, where entries can be either displacements or rotations. Here $n_\mathbf{y}$ is the number of measurements, i.e., the number of sensors to be deployed multiplied by the number of physical quantities that can be measured by each sensor.

Let us define a set of mechanical properties like Young modulus, stiffness, modal parameters, collected in the vector $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$, to be identified using the measured data $\mathbf{y}$. The capability of a SHM system to detect and identify the said mechanical properties is basically affected by two aspects [1]: the mathematical tools exploited to estimate the parameters; and the experimental setting, i.e., the physical quantities to be measured and the procedure to measure them. Considering the experimental design alone, the estimation capability is affected by several aspect, including:

- sensor characteristics such as accuracy, resolution, sensitivity, noise, etc.;

- number of sensors;

- orientation and deployment of sensors;

which should be appropriately set to ensure high accuracy at a reduced cost.

Optimal sensor placement (OSP) forms a fundamental task for real SHM applications, as it can allow decreasing the cost and complexity of the SHM system. As summarized in [2, 3, 4], among the most commonly adopted OSP methods we find: the effective independence method (EFI); the driving-point residue method (EFI-DPR); the kinetic energy method (KEM); the modified variance method (MVM); methods based on topology optimization [5]. All aforementioned methods do not take into account that measurements are inevitably corrupted with errors, and hence they should be treated as stochastic quantities. The main goal of the present work is to determine the optimal configuration of sensors, that guarantees the best estimation of mechanical parameters or, in other words, the maximum amount of inferred information, under noise corrupted variables. The general purpose framework proposed in [6] is applied to the OSP problem, in order to find the sensor configuration guaranteeing a maximal amount of information, as already proposed in [7] and [8].

The remainders of the paper is organized as follows: in Section 2, the main theoretical aspects of the adopted methodology, such as the problem settings, the numerical solution of the problem and the optimization scheme, are described. In section 3, the framework is implemented on a simple benchmark problem consisting in a flexible simply supported plate. Finally, some concluding remarks and possible future developments are friefly summarized in Section 4.

2 BAYESIAN EXPERIMENTAL DESIGN FOR OPTIMAL SENSOR PLACEMENT

2.1 Problem settings

Moving from a space discretized system, the position of the sensors is defined through an appropriate design variable vector $\mathbf{d} \in \mathbb{R}^{n_d}$, that can be set alternatively as:

- vector of spatial nodal coordinates

$$\mathbf{d} = \{x_1 \ y_1 \ z_1 \ \boldsymbol{\delta}_1 \ \dots \ x_{n_{sens}} \ y_{n_{sens}} \ z_{n_{sens}} \ \boldsymbol{\delta}_{n_{sens}}\}^T \tag{1}$$

where $1 \leqslant n_{sens} \leqslant n_{nodes}$ is the number of sensors to be deployed; variable $\boldsymbol{\delta}_i = \{u_i, v_i, w_i\}^T$ identifies the orientation of the $i-$th sensor; the spatial coordinates x_i, y_i, z_i are defined in the discrete domain of all the nodes of the Finite Element (FE) model;

- vector of nodal labels

$$\mathbf{d} = \{d_1 \ \boldsymbol{\delta}_1 \ \ldots \ d_{n_{sens}} \ \boldsymbol{\delta}_{n_{sens}}\}^T \tag{2}$$

with d_i designating the FE model node where the $i-$th sensor is placed.

The measured quantities $\mathbf{y}$ are related to the model output $\mathbf{u} \in \mathbb{R}^{n_{nodes}n_{dof}}$, n_{dof} denoting the number of degrees of freedom of each node associated to the FE model, through the relation:

$$\mathbf{y} = \mathbf{Lu} + \boldsymbol{\epsilon} \tag{3}$$

where $\boldsymbol{\epsilon} \in \mathbb{R}^{n_{sens}}$ is the model prediction error, accounting for the modelling and measurement errors, and is therefore related to the type of sensors deployed. $p_{\boldsymbol{\epsilon}} = p(\boldsymbol{\epsilon})$ is defined as the probability density function associated with the model prediction error. $\mathbf{L} \in \mathbb{R}^{n_{sens} \times n_{nodes}n_{dof}}$ is a configuration matrix, linking the measurements $\mathbf{y}$ to the model response $\mathbf{u}$, defined as a boolean matrix, with 1 or 0 entries corresponding respectively to the measured and non-measured quantities. $\mathbf{L}$ depends both on the sensors positions, i.e., the nodes in which the sensors are placed, and the direction of the measured displacement or accelerations. Therefore, $\mathbf{L} = \mathbf{L}(\mathbf{d})$ is function of the experimental variable $\mathbf{d}$.

Since the structure is modeled through a FE model, the model response $\mathbf{u}$ depends on the mechanical properties of the system, according to:

$$\mathbf{u} = \mathbf{K}(\boldsymbol{\theta})^{-1}\mathbf{F} \tag{4}$$

where $\mathbf{K} \in \mathbb{R}^{n_{nodes}n_{dof} \times n_{nodes}n_{dof}}$ is the stiffness matrix associated with the discretized structure and $\mathbf{F} \in \mathbb{R}^{n_{nodes}n_{dof}}$ is the load vector. The load is assumed implemented as a quasi-static way and therefore any dynamic effect on the response is disregarded.
Eqs. (3) and (4) lead to:

$$\mathbf{y} = \mathbf{G}(\mathbf{d}, \boldsymbol{\theta}) + \boldsymbol{\epsilon} \tag{5}$$

where:

$$\mathbf{G}(\mathbf{d}, \boldsymbol{\theta}) = \mathbf{L}(\mathbf{d})\mathbf{K}(\boldsymbol{\theta})^{-1}\mathbf{F} \tag{6}$$

is the forward model operator $\mathbb{R}^{n_{\mathbf{d}}} \times \mathbb{R}^{n_{\boldsymbol{\theta}}} \to \mathbb{R}^{n_{\mathbf{u}}}$ that maps the design variable and the parameters to be estimated onto the response domain.
The main goal of the OSP problem is to determine the optimal sensor configuration $\mathbf{d}^*$, for the Bayesian inference of the mechanical parameters $\boldsymbol{\theta}$. Here we use the general framework proposed in [6, 9], which is applicable to any Bayesian inference problem; $\boldsymbol{\theta}$ and $\mathbf{y}$ are therefore treated as random variables, with their appropriate probability densities, accounting for their respective uncertainties.
Bayes' theorem allows to compute the conditional probability density of the parameters to be inferred, given the acquired data, i.e., the measurements. In the case of experimental design, the Bayes' rule can be specialized, taking also into account the design variable, as:

$$p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d}) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{d})p(\boldsymbol{\theta}|\mathbf{d})}{p(\mathbf{y}|\mathbf{d})} \tag{7}$$

where $p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})$ is the posterior distribution (probability density of $\boldsymbol{\theta}$, given $\mathbf{y}$ and $\mathbf{d}$), $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{d})$ is the likelihood, $p(\boldsymbol{\theta}|\mathbf{d})$ is the prior distribution and $p(\mathbf{y}|\mathbf{d})$ is the evidence. In the following, the

evidence is assumed to be independent of the design variable, therefore $p(\boldsymbol{\theta}|\mathbf{d}) = p(\boldsymbol{\theta})$. Hence, the prior knowledge of $\boldsymbol{\theta}$ is not affected by the experimental procedure.

Following [10], the optimal experimental design in a Bayesian sense can be obtained by maximizing the following expression:

$$\max_{\mathbf{d} \in \mathcal{D}} \int_{\mathcal{Y}} \int_{\Theta} U(\mathbf{d}, \mathbf{y}, \boldsymbol{\theta}) p(\boldsymbol{\theta}, \mathbf{y}|\mathbf{d}) d\boldsymbol{\theta} d\mathbf{y} \tag{8}$$

where $\mathcal{Y}$ and Θ respectively designate the supports of $p(\mathbf{y}|\mathbf{d})$ and $p(\boldsymbol{\theta})$. Within a decision theory formulation, $U(\mathbf{d}, \mathbf{y}, \boldsymbol{\theta})$ is a utility function defined in $(\mathcal{D}, \mathcal{Y}, \Theta) \rightarrow \mathbb{R}$ that takes into account the usefulness of the experimental decision $\mathbf{d}$, given a particular value of $\boldsymbol{\theta}$ and $\mathbf{y}$. The optimal experimental design aims to maximize the expected utility function specified by Eq. (8). As summarized in [11], the choice of the utility function is strictly related to the type of problem, whether it is a prediction or an estimation one. As suggested in [7], the most common utility function for the inference problems is the expected Kullbach-Leibler divergence (also called expected gain in Shannon information [12] or relative entropy) between the posterior and the prior probability distributions [13]:

$$D_{KL}(p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})||p(\boldsymbol{\theta})) = \int_{\mathcal{Y}} \int_{\Theta} \log \frac{p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})}{p(\boldsymbol{\theta})} p(\mathbf{y}, \boldsymbol{\theta}|\mathbf{d}) d\mathbf{y} d\boldsymbol{\theta} \tag{9}$$

The rationale behind this choice relies on the concept that the optimal configuration should yield the largest increase in information between the prior and the posterior.

Recasting Eq. (9) into the experimental design problem, the utility function becomes:

$$U(\mathbf{d}, \mathbf{y}, \boldsymbol{\theta}) = U(\mathbf{d}, \mathbf{y}) = \int_{\Theta} \log \frac{p(\bar{\boldsymbol{\theta}}|\mathbf{y}, \mathbf{d})}{p(\bar{\boldsymbol{\theta}})} p(\bar{\boldsymbol{\theta}}|\mathbf{y}, \mathbf{d}) d\bar{\boldsymbol{\theta}} \tag{10}$$

Remebering that $\bar{\boldsymbol{\theta}}$ is a dummy variable and using Eq. (10), the optimization statement of Eq. (8) becomes:

$$\max_{\mathbf{d} \in \mathcal{D}} \int_{\mathcal{Y}} \int_{\Theta} p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d}) \ln \frac{p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})}{p(\boldsymbol{\theta})} d\boldsymbol{\theta} p(\mathbf{y}|\mathbf{d}) d\mathbf{y} \tag{11}$$

Rearranging Eq. (11) using Eq. (7), we can rewrite the optimization problem as:

$$\text{find} \quad \mathbf{d}^* = \arg \max_{\mathbf{d} \in \mathcal{D}} \bar{U}(\mathbf{d})$$

$$\text{such that} \quad \bar{U}(\mathbf{d}) = \int_{\mathcal{Y}} \int_{\Theta} \{\ln[p(\mathbf{y}, \boldsymbol{\theta}|\mathbf{d})] - \ln[p(\mathbf{y}|\mathbf{d})]\} p(\mathbf{y}, \boldsymbol{\theta}|\mathbf{d}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} d\mathbf{y} \tag{12}$$

where $\bar{U}(\mathbf{d})$ is the so-called expected information gain in $\boldsymbol{\theta}$.

2.2 Surrogate modeling

The optimization problem defined in Eq. (12) is handled through Monte Carlo sampling as:

$$\bar{U}(\mathbf{d}) \approx \frac{1}{N_{out}} \sum_{i=1}^{N_{out}} \left\{ \ln\left[p(\mathbf{y}^i|\boldsymbol{\theta}^i, \mathbf{d})\right] - \ln\left[p(\mathbf{y}^i|\mathbf{d})\right] \right\} \tag{13}$$

where N_{out} represents the number of samples $\boldsymbol{\theta}^i$ to be drawn in the outer Monte Carlo sum. The evidence computed at $\mathbf{y}^{(i)}$ is approximated as:

$$p(\mathbf{y}^i|\mathbf{d}) \approx \frac{1}{N_{in}} \sum_{j=1}^{N_{in}} p(\mathbf{y}^i|\boldsymbol{\theta}^j, \mathbf{d}) \tag{14}$$

The samples $\boldsymbol{\theta}^i$ and $\boldsymbol{\theta}^j$ are drawn from the prior distribution $p(\boldsymbol{\theta})$, which is chosen a priori. In SHM applications, the main goal is to identify the mechanical properties of a damaged structure, and it is common to assume parameters $\boldsymbol{\theta}$ as drawn from a uniform distribution $\mathcal{U}(\boldsymbol{\Theta})$. In reducing the associated computational cost, the same set of samples $\boldsymbol{\theta}$ is used for both the inner and the outer sum, resulting in $N_{in} = N_{out}$; the computational complexity is thus reduced from $O(N_{out}N_{in})$ to $O(N_{out})$. For further details on the bias introduced by this assumption and on the numerical approximation, the interested reader is referred to [6].

The term $p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{d})$ corresponds to the probability density of the measured data $\mathbf{y}$, given a specific set of mechanical parameters $\boldsymbol{\theta}$ and a specific sensor configuration $\mathbf{d}$. Using Eq. (5), we can conclude that:

$$p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{d}) = p_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}) = p_{\boldsymbol{\epsilon}}(\mathbf{y} - \mathbf{G}(\mathbf{d}, \boldsymbol{\theta})) \tag{15}$$

Hence, via use of Eqs. (6) and (15), Eq. (13) becomes:

$$\begin{aligned}
\bar{U}(\mathbf{d}) &\approx \frac{1}{N_{out}} \sum_{i=1}^{N_{out}} \left\{ \ln\left[p_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}^i)\right] - \ln\left[\frac{1}{N_{in}} \sum_{j=1}^{N_{in}} p_{\boldsymbol{\epsilon}}(\mathbf{G}(\mathbf{d}, \boldsymbol{\theta}^i) + \boldsymbol{\epsilon}^i - \mathbf{G}(\mathbf{d}, \boldsymbol{\theta}^j))\right]\right\} = \\
&= \frac{1}{N_{out}} \sum_{i=1}^{N_{out}} \left\{ \ln\left[p_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}^i)\right] - \ln\left[\frac{1}{N_{in}} \sum_{j=1}^{N_{in}} p_{\boldsymbol{\epsilon}}(\mathbf{L}(\mathbf{d})[\mathbf{K}^{i^{-1}} - \mathbf{K}^{j^{-1}}]\mathbf{F} + \boldsymbol{\epsilon}^i)\right]\right\}
\end{aligned} \tag{16}$$

where $\mathbf{K}^{i,j} = \mathbf{K}(\boldsymbol{\theta}^{i,j})$. The samples $\boldsymbol{\epsilon}^i$ are drawn from the density distribution $p_{\boldsymbol{\epsilon}}$, that is chosen once again a priori.

From a computational point of view, the most time-consuming part of the procedure is the evaluation of the system response through the model mapping $\mathbf{G}(\mathbf{d}, \boldsymbol{\theta})$ of Eq. (6). Assuming that we use the same batch of samples $\boldsymbol{\theta}$ for both sums, the Monte Carlo approximation would require N_{out} evaluations of the system output, each one associated to a stiffness matrix inversions, rendering the computation practically infeasable. In order to overcome this problem, in [6] the adoption of a Polynomial Chaos Expansion (PCE) scheme has been proposed. Given an input random vector with independent components $\mathbf{x} \in \mathbb{R}^{n_{\mathbf{x}}}$, described by a joint probability density function $f_{\mathbf{x}}$, and an output random vector $\mathbf{y}$ calculated through a certain computational model $\mathcal{M}$, the PCE allows to build a map, or metamodel, relating the two spaces. Assuming that:

$$\mathbb{E}[\mathbf{y}^2] = \int_{\mathcal{D}_{\mathbf{x}}} \mathcal{M}^2(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} < \infty \tag{17}$$

it is possible to define (see [14]):

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\mathbf{x}) \tag{18}$$

where $\boldsymbol{\psi}_{\boldsymbol{\alpha}}$ are multivariate polynomials of order p orthonormal to $f_{\mathbf{x}}$, $\boldsymbol{\alpha} \in \mathbb{N}^M$ is a multi-index associated with $\boldsymbol{\psi}_{\boldsymbol{\alpha}}$, and $y_{\boldsymbol{\alpha}} \in \mathbb{R}$ are the corresponding coefficients. For practical purposes, the sum is truncated to a sub-set of the multi-indeces $\boldsymbol{\alpha} \in \mathcal{A} \subset \mathbb{N}^M$, according to:

$$\mathbf{y} \cong \mathbf{y}^{PCE} = \mathcal{M}^{PCE}(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\mathbf{x}) \tag{19}$$

where $\mathcal{M}^{PCE}$ stands for the surrogate model, and $\mathbf{y}^{PCE}$ is the approximated response computed through it.

Recalling Eq. (5), the forward model maps the design variable and the parameters into the measurement domain $(\mathcal{D}, \boldsymbol{\Theta}) \rightarrow \mathcal{Y}$ as:

$$\mathbf{y} = \mathcal{M}(\mathbf{x}) = \mathbf{G}(\mathbf{d}, \boldsymbol{\theta}) \tag{20}$$

As suggested in [6], the PCE is applied jointly as:

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{d} \end{bmatrix} \tag{21}$$

with $\mathbf{x} \in \mathbb{R}^{n_{\boldsymbol{\theta}}+n_{\mathbf{d}}}$. In this way, it is not necessary to build a different PCE for each possible design variable $\mathbf{d}$, which would lead to a larger computational cost. In order to compute the surrogate model, two sets of samples $\boldsymbol{\theta}_k$ and $\mathbf{d}_k$ with $k = 1, \cdots, N_{PCE}$ are drawn from the respective aforementioned probability densities; the output samples are then calculated through the complete model $\mathbf{y}_k = \mathcal{M}(\mathbf{x}_k)$. Once the input and output populations have been built, the truncated bases and the relative coefficients are computed through the Least Angle Regression (LAR) algorithm, as specified in [14, 15]. The accuracy of the surrogate model with respect to the complete one is affected by both the population size N_{PCE} and the poynomial degree p; these settings should be chosen so as to minimize the leave-one-out (LOO) cross-validation error, as defined in [16].

2.3 Optimization algorithm

The optimization statement summarized in Eq. (12) is herein solved by applying the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [17], which comprises a stochastic search algorithm for numerical optimization of non-linear non-convex functions. The basic idea is to sample the design variable $\mathbf{d}$ from a multivariate normal distribution, and produce a new mutation of the generated population that is closer to the maximum of the function $\bar{U}(\mathbf{d})$. The algorithm can be summarized as follows:

1. for each iteration, a population of design points $\mathbf{d}$ is sampled from the Normal multivariate distribution $\mathbf{d}_i \sim \mathbf{m} + \sigma \mathcal{N}_i(\mathbf{0}, \mathbf{C})$ for $i = 1, \cdots, \lambda$, where $\mathbf{m} \in \mathbb{R}^{n_{\mathbf{d}}}$ is the mean vector, $\sigma \in \mathbb{R}$ is the step size and $\mathbf{C} \in \mathbb{R}^{n_{\mathbf{d}} \times n_{\mathbf{d}}}$ is the covariance matrix;

2. $\mathbf{m}$ and $\mathbf{C}$ are updated through cumulation, in order the likelihood of the successive iterations to be increased;

3. the previous steps are repeated until a certain tolerance on the function evaluation is attained.

Further details on the algorithm can be found in [18].

3 RESULTS

The method discussed in Section 2 is now applied to a simple SHM benchmark case; a square plate, simply supported at four edges and subjected to a static force applied at its centre point (see node 1 in Figure 1). The structure is ideally subdivided into four zones and the aim of the SHM system is the estimation of the Young modulus E_j, $j = 1, ..., 4$, of each zone. The sensors can be positioned at each node of the adopted 10x10 FE model and, for the sake of simplicity, only the out-of-plane deflections are assumed to be measured. The structure is discretized through conventional shell elements, with 6 degrees of freedom per node. The mechanical behaviour of the material is supposed to be linear elastic.
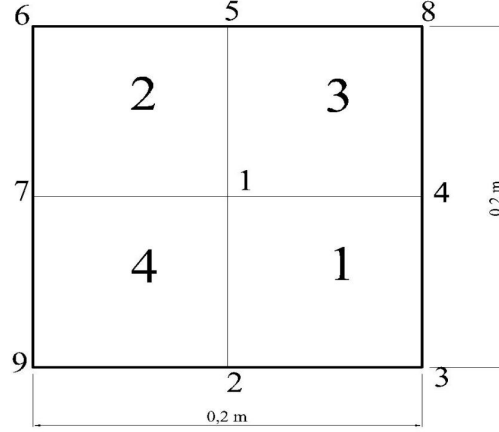Two cases are considered herein:

Figure 1: Benchmark structural case.

(a) the position of the damaged zone is known, therefore the input variables of the PCE surrogate model are summarized in vector:

$$\mathbf{x} = \{E_2 \ d_1 \ \cdots \ d_{n_{sens}}\}^T \tag{22}$$

(b) the position of the damaged zone is unknown:

$$\mathbf{x} = \{E_1 \ E_2 \ E_3 \ E_4 \ d_1 \ \cdots \ d_{n_{sens}}\}^T \tag{23}$$

Let us consider the simplest case: only one sensor is deployed. Figures 2 and 3 demonstrate the contour plot of the objective function defined by Eq. (12) computed on each node of the plate, respectively related to cases (a) and (b). The parameters to be estimated, aggregated in vector $\boldsymbol{\theta}$, are sampled from a uniform distribution $\mathcal{U}(0, E)$, where E is the Young modulus of the undamaged material. The probability distribution of the model prediction error ϵ is supposed to be Gaussian $\mathcal{N}(0, \sigma^2)$ with standard deviation $\sigma = 10^{-5}$ m. In order to compute the PCE coefficients and bases, the design variable $\mathbf{d}$ is sampled uniformly with coordinates defined over the interval $[0, 0.2]$, both along the width and the length of the plate.

The optimal configuration, i.e., maximizing the objective function, is determined at position $(x^*, y^*) = (0.08, 0.12)$ for case (a), i.e., as expected near the zone in which the damage will occur, see [5]. In case (b), the problem settings are perfectly symmetric and therefore the objective function as well: the optimal is determined in point $(x^*, y^*) = (0.10, 0.10)$. The figures show the optimization paths obtained through the CMA-ES algorithm: each path corresponds to a different initial condition, i.e., the four corner points, and the red circles identify the results reached after 25 objective function evaluations. It is demonstrated that, for both cases, the results are stable with respect to the initialization settings and they correspond to the maximum of the contour plot.

The accuracy of the optimization process is largely affected by the PCE approximation phase. The error introduced by the PCE surrogate model with respect to the FE model can be measured through the leave-one-out (LOO) error, as defined in [16]:

$$\epsilon_{LOO} = \frac{\sum_{i=1}^{N} \left(\mathcal{M}(\mathbf{x}^i) - \mathcal{M}^{PCE}(\mathbf{x}^i)\right)^2}{1 - h_i} \bigg/ \left(\sum_{i=1}^{N} \left(\mathcal{M}(\mathbf{x}^i) - \hat{\mu}_Y\right)^2\right) \tag{24}$$
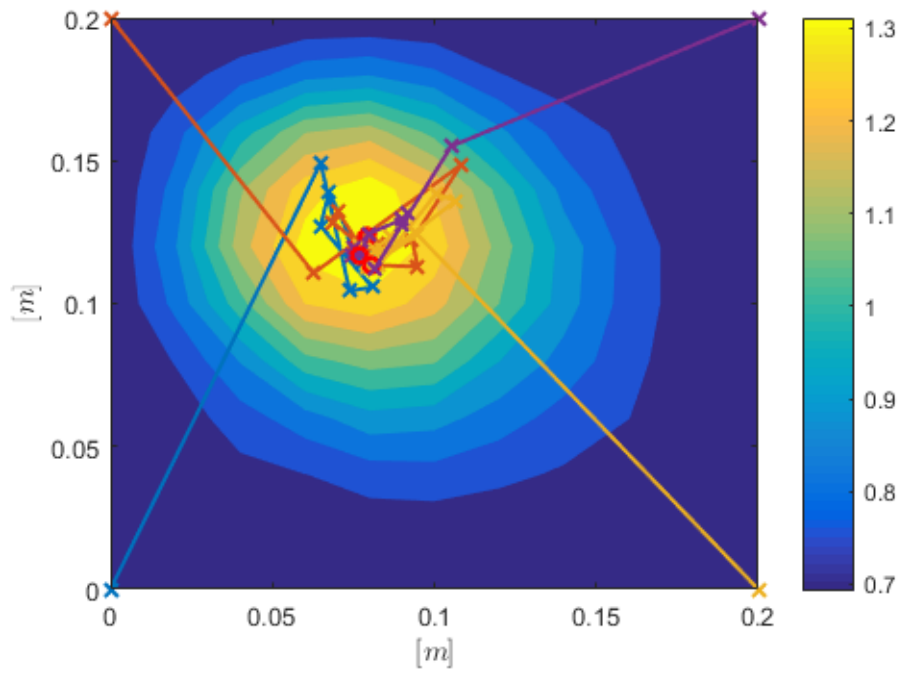
Figure 2: Contour plot of the objective function $\hat{U}(\mathbf{d})$ with one sensor for case (a) and CMA-ES optimization paths ($N^{PCE} = 10^4, p = 10, N^{MC} = 5 \cdot 10^3$).
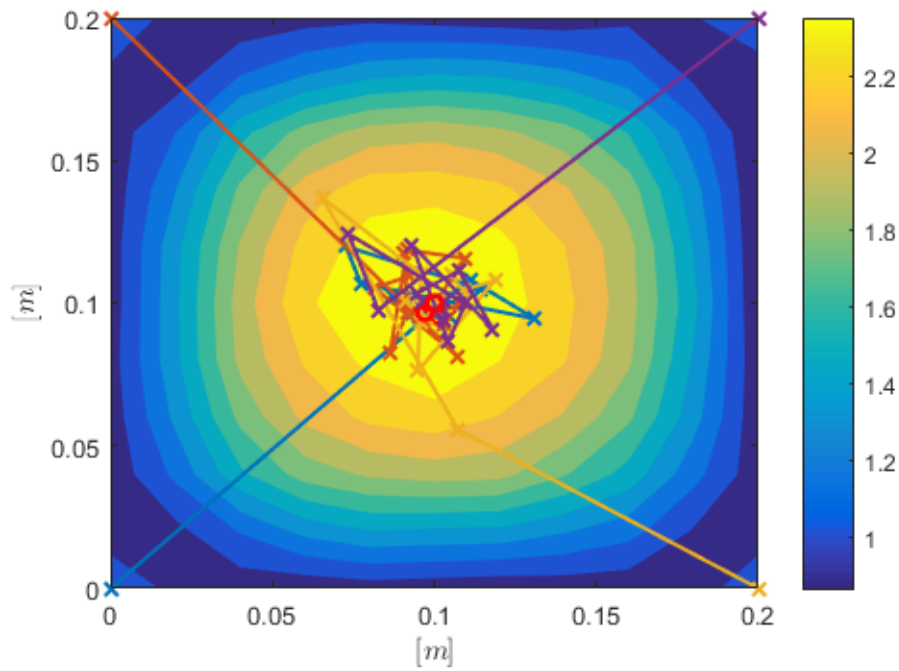

Figure 3: Contour plot of the objective function $\hat{U}(\mathbf{d})$ with one sensor for case (b) and CMA-ES optimization paths ($N^{PCE} = 10^4, p = 10, N^{MC} = 5 \cdot 10^3$).

where $\hat{\mu}_Y = \frac{1}{N}\sum_{i=1}^{N} \mathcal{M}(\mathbf{x}^i)$ is the sample mean of the set of quadrature points and h_i is the ith component of the vector given by:

$$h = \text{diag}\left(\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\right) \tag{25}$$

where $\mathbf{A}$ is the experimental matrix that contains the values of all the basis polynomials in the experimental design points.

Figures 4 and 5 report the LOO error as a function of the polynomial order p and the number of samples N^{PCE} in a log-log plot. Figures show that, as expected, the error decreases as N^{PCE} and p increase; we can point out that for increase beyond a certain value of N^{PCE}, the error trend remains basically stable.
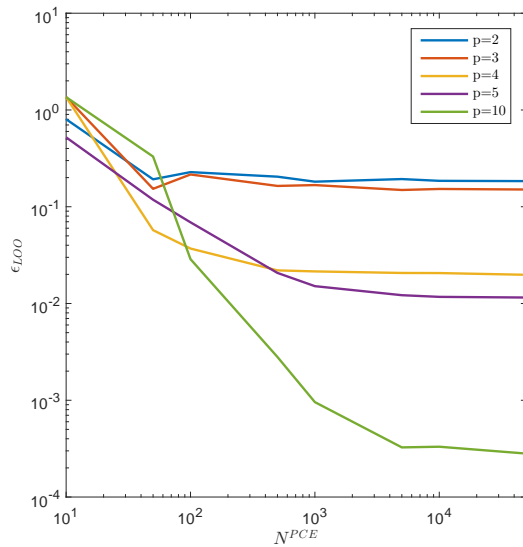

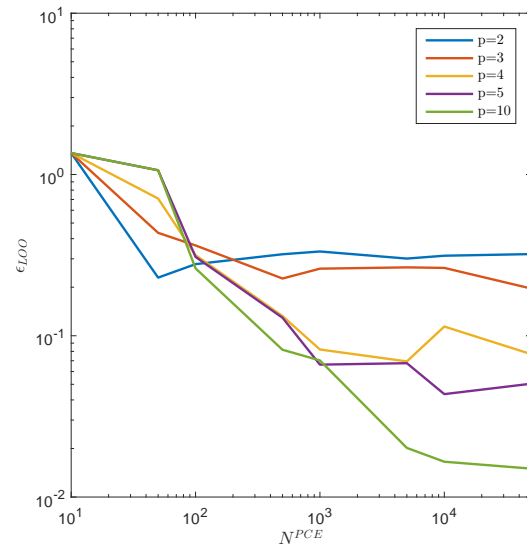Figure 4: LOO error ϵ_{LOO} associated to the PCE surrogate model for case (a).

Figure 5: LOO error ϵ_{LOO} associated to the PCE surrogate model for case (b).

4 CONCLUSIONS

In the present paper, a new method has been proposed for optimal sensor placement in structural health monitoring applications, having supposed the measurements are noise-corrupted. The method is adapted from the general framework proposed in [6] for the optimal design of experiments; the procedure is coupled with a Covariance Matrix Adaptation Evolution Strategy optimization scheme, which allows for determining the optimal sensor configuration, associated with the maximum amount of information given by the measurements. In order to reduce the computational cost linked to the multiple forward problem simulations required, a Polynomial Chaos Expansion metamodeling scheme is adopted.

The method is herein applied to a simple benchmark problem of a simply supported plate under bending loads. It has been demonstrated that the optimization algorithm is able to find the optimal sensor configuration, further taking also into account measurement error. The method proves successful and yielding a design of experiments corresponding to reduced approximation error.

Further developments will concern the application to more complex structures, with multiple sensors configurations.

REFERENCES

[1] D. Balageas, C. P. Fritzen, A. Gemes, *Structural Health Monitoring*. Wiley-ISTE, 2006.

[2] C. Leyder, V. Ntertimanis, E. Chatzi, A. Frangi, Optimal Sensors Placement for the Modal Identification of an Innovative Timber Structure. *Proceedings of the 1st International Conference on Uncertainty Quantification in Computational Sciences and Engineering*, 467-476, 2015.

[3] M. Meo, G. Zumpano, On the optimal sensor placement techniques for a bridge structure. *Engineering Structures*, 27:14881497, 2005.

[4] M. Chang, S. Pakzad, Optimal Sensor Placement for Modal Identification of Bridge Systems Considering Number of Sensing Nodes. *Journal of Bridge Engineering*, 19(6):04014019, 2014.

[5] M. Bruggi, S. Mariani, Optimization of sensor placement to detect damage in flexible plates. *Engineering Optimization*, 45(6), 659-676, 2013.

[6] X. Huan, Y. M. Marzouk, Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232(1):288-317, 2013.

[7] K. J. Ryan, Estimating Expected Information Gains for Experimental Designs with Application to the Random Fatigue-Limit Model. *Journal of Computational and Graphical Statistics*, 12(3):585-603, 2003.

[8] C. Papadimitriou, Optimal Sensor Placement Methodology for Parametric Identification of Structural Systems. *Journal of Sound and Vibration*, 278:923947, 2004.

[9] X. Huan, Y. M. Marzouk, Gradient-Based Stochastic Optimization Methods in Bayesian Experimental Design. *International Journal for Uncertainty Quantification*, 4(6):479-510, 2014.

[10] D. V. Lindley, *Bayesian Statistics, A Review*, Society for Industrial and Applied Mathematics (SIAM), 1972.

[11] K. Chaloner, I. Verdinelli, Bayesian Experimental Design: A Review. *Statistical Science*, 10(3):273-304, 1995.

[12] C. E. Shannon, A Mathematical Theory of Communication. *Bell System Technical Journal*, 27:379-423, 1948.

[13] D. V. Lindley, On a Measure of the Information Provided by an Experiment. *The Annals of Mathematical Statistics*, 27(4):986-1005, 1956.

[14] S. Marelli, B. Sudret, *UQLab User Manual*. Chair of Risk, Safety & Uncertainty Quantification, ETH Zürich, 2015.

[15] G. Blatman, B. Sudret, Adaptive Sparse Polynomial Chaos Expansion Based on Least Angle Regression. *Journal of Computational Physics* , 230:2345-2367, 2011.

[16] G. Blatman, *Adaptive Sparse Polynomial Chaos Expansion for Uncertainty Propagation and Sensitivity Analysis*. Ph.D. thesis, Blaise Pascal University, Clermont-Ferrand, 2009.

[17] N. Hansen, A. Ostermeier, Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, 312-317, 1996.

[18] N. Hansen, S. D. Müller, P. Koumoutsakos, Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1-18, 2003.