

Article

Cost–Benefit Optimization of Structural Health Monitoring Sensor Networks

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, Italy; stefano.mariani@polimi.it

² ETH Zürich, Institut für Baustatik und Konstruktion Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland; chatzi@ibk.baug.ethz.ch

* Correspondence: giovanni.capellari@polimi.it; Tel.: +39-02-2399-4267

† These authors contributed equally to this work.

Received: 11 June 2018; Accepted: 4 July 2018; Published: 6 July 2018



Abstract: Structural health monitoring (SHM) allows the acquisition of information on the structural integrity of any mechanical system by processing data, measured through a set of sensors, in order to estimate relevant mechanical parameters and indicators of performance. Herein we present a method to perform the cost–benefit optimization of a sensor network by defining the density, type, and positioning of the sensors to be deployed. The effectiveness (benefit) of an SHM system may be quantified by means of information theory, namely through the expected Shannon information gain provided by the measured data, which allows the inherent uncertainties of the experimental process (i.e., those associated with the prediction error and the parameters to be estimated) to be accounted for. In order to evaluate the computationally expensive Monte Carlo estimator of the objective function, a framework comprising surrogate models (polynomial chaos expansion), model order reduction methods (principal component analysis), and stochastic optimization methods is introduced. Two optimization strategies are proposed: the maximization of the information provided by the measured data, given the technological, identifiability, and budgetary constraints; and the maximization of the information–cost ratio. The application of the framework to a large-scale structural problem, the Pirelli tower in Milan, is presented, and the two comprehensive optimization methods are compared.

Keywords: structural health monitoring; Bayesian inference; cost–benefit analysis; stochastic optimization; information theory; Bayesian experimental design; surrogate modeling; model order reduction

1. Introduction

Structural health monitoring (SHM) allows the detection and estimation of variations in the behavior and thereby condition of engineered systems [1], and therefore helps in making decisions about the actions needed to maintain or recover the overall structural safety [2]. Amongst the available methods for SHM in the literature (see [1]), the Bayesian framework allows both unknown system properties and their associated uncertainties to be estimated, as introduced in [3]. The effectiveness of any SHM system to estimate and detect damage, here assumed as a variation of mechanical properties, depends on both the estimation method exploited to crunch the data and the SHM system itself. From a theoretical point of view, we can interpret SHM as an experimental procedure, where the measurements obtained through the SHM system are exploited to reach the goal of the experiment (i.e., the estimation of the parameters and their associated uncertainties). In this view, the experimental setup in SHM includes all the settings which can affect the measurements—namely, the position of the

sensors on the structure, the physical quantities to be measured, the number of sensors, and the type of sensors.

In the present work, a SHM system cost–benefit optimization method is presented, which allows the most suitable experimental settings to be chosen in order to maximize the estimation potential and simultaneously minimize the cost of the SHM system. Two alternative approaches are proposed and discussed: (i) the system is optimized by maximizing its effectiveness and simultaneously fulfilling the budgetary constraint; (ii) the system is optimized by maximizing the ratio between its effectiveness and its cost.

Two main advantages can be derived from optimizing the sensor network. Let us first assume that by optimizing the sensor network in terms of position and types of sensors, their number can be decreased. This results in a cost reduction of the overall SHM system and a simplification of the data acquisition system and the system assembly phase. Moreover, an additional side benefit lies in the reduction of the amount of data to be processed. In other words, the resulting optimized SHM system is more “informative”, and thus fewer sensors are required to guarantee the same accuracy of the estimated quantities. Consequently, both the cost and the complexity of the data storage system and the required computational resources can be significantly reduced. Moreover, since the amount of data to be processed is lower, the applicability of real-time estimation methods is enhanced, and the required data storage is reduced for non-real-time applications. On the contrary, if the number and type of sensors are assumed to be constant, the optimization of the sensor network guarantees an increase of information provided by the monitoring system, and thus a consequent reduction of the estimate uncertainties.

Several methods have been presented in the literature to optimally design SHM systems: in all of them, the type and number of sensors are assumed to be constant. The vast majority of these prescribes the type and number of sensors as constant parameters in the optimization problem, allowing for the optimization of the spatial configuration alone. Among them, a stochastic approach for the optimal sensor placement, based on the minimization of a Bayesian loss function, was introduced in [4]: the sensor locations were chosen such that the expectation of the squared error loss function between the estimated and the target values of the quantities to be estimated was minimized. In [5], the evaluation of entropy and mutual information was proposed in order to quantify the amount of information, which can be inferred experimentally (and therefore from an SHM system).

In [6,7], Papadimitriou et al. proposed the minimization of the information entropy as a rationale to optimize the spatial configuration of the sensor network. In order to numerically evaluate the associated objective function, the integral terms arising from an analytical manipulation of the information entropy were approximated through the Laplace method of asymptotic expansions, which allows an associated algebraic formulation to be obtained. Unlike in the present study, the objective function was locally approximated through a smooth replica centered on nominal parameter values, which have to be chosen a priori.

The optimization algorithm to obtain the optimal solution (i.e., maximization of the objective function) has mostly been treated as a discrete optimization problem in the existing literature, with genetic algorithms exploited in [8,9]. According to an alternative greedy approach proposed in [7,10], the optimal configuration can be obtained by splitting the optimization problem into a number of sub-problems, where only one sensor is added at each step, so that the increase in the objective function value is maximized.

In contrast to the existing methods, in the present work a method is introduced to comprehensively optimize the sensor network not only in terms of sensor placement, but also in terms of the type and number of sensors. The effectiveness of a sensor network is quantified through an index based on information theory, originally developed within the computer science research community for the quantification of uncertainty relating to random variables. The SHM sensor network is therefore optimized in terms of number, position, and type of sensors by maximizing the relevant expected Shannon information gain, which is a measure of the utility of the measurements with respect to the

quantities to be estimated. It should be underlined that the proposed method was developed within a Bayesian framework and is therefore valid for SHM procedures aiming at Bayesian inference, such as Bayesian model updating or parameter characterization. It can be applied to processes that run both offline and in real-time (as in the reasoning of Kalman filtering-based estimation [11–14]). As in most stochastic approaches, large computational resources would be needed to take the uncertainties into account, preventing the applicability of the method to large structural models. The coupling with surrogate models (polynomial chaos expansion, [15]), which aims at replacing the original computationally expensive numerical model by reproducing the relation between inputs and outputs, and model order reduction strategies (principal component analysis) allow the latter problem to be overcome. The use of stochastic optimization methods (covariance matrix adaptation-evolutionary strategy, [16]) allows high-dimensional problems to be solved, which was not possible with the previously exploited methods.

The method developed herein is applicable to both static and dynamic monitoring applications of diverse sensing capabilities, where a Bayesian framework is intended to be implemented. The method is demonstrated herein for static measurements only. However, the same framework can be implemented for dynamic monitoring as well if the objective function is expressed in terms of frequency and mode-shape matching.

This paper is organized as follows: first, the theoretical framework is presented in Section 2 by introducing the Bayesian experimental design in Section 2.1 and defining the optimization statement in Section 2.2. Then, the approach for numerically evaluating the objective function is discussed in Sections 3 and 4. In Section 5, the whole optimization procedure is summarized. Then, the application of the method to a tall building, namely the Pirelli tower in Milan, is presented in Section 6. Finally, some concluding remarks are gathered in Section 7.

2. Theoretical Basis

2.1. Bayesian Experimental Design

Let the goal of the SHM system be the estimation of a set of parameters (e.g., mechanical properties, geometrical properties, or damage indices) defined within an appropriate numerical model of a structure, used to predict its response to given loads. The following random vectors are defined: the parameter vector $\theta = [\theta_1 \ \theta_2 \ \dots \ \theta_{n_\theta}] \in \mathbb{R}^{n_\theta}$ to be estimated; the data vector $\mathbf{y} \in \mathbb{R}^{n_y}$, with measurements assumed to be collected through a set of sensors. Here, n_θ is the number of parameters and n_y is the number of measurements.

The prior probability density function (pdf) $p(\theta)$ represents the prior knowledge on θ , and it can be suitably chosen in order to take initial information into account, such as that provided by previous experiments or by the subjective belief of an expert. If no previous information is available on the values of parameters in θ , an uninformative distribution can be considered. The pdf may then be updated considering the data \mathbf{y} , through Bayes' theorem:

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)p(\theta)}{p(\mathbf{y})}, \quad (1)$$

where the expression $p(\cdot|\cdot)$ represents the conditional pdf of the first term with respect to the second one. Thus, $p(\theta|\mathbf{y})$ is the posterior pdf (i.e., the probability density function of θ), given \mathbf{y} ; $p(\mathbf{y}|\theta)$ is the likelihood; $p(\mathbf{y})$ is the evidence, that is, the distribution of the observed data, marginalized over θ .

Bayes' theorem is particularly suited within such a context in SHM problems: given a structure, a class of models can be a-priori defined to describe the behaviour of the system, and updated as soon as the structure response is measured. Bayesian model updating was first introduced for structural applications in [3,17], and it allows the posterior probability density function $p(\theta|\mathbf{y})$ to be

obtained based on prior knowledge of θ and the data, that is, the maximum a posteriori estimate $\theta^* = \arg \max_{\theta} [p(\theta|\mathbf{y})]$, along with the related uncertainty level.

Within this framework, the effectiveness of the sensor network can be quantified by following the decision-theoretic approach introduced in [18,19]: in the case at hand, prior to performing the measurements, the choice of the experimental settings—in terms of spatial sensor configuration, number of sensors, and type of sensors—must be made. An additional term—the design variable $\mathbf{d} \in \mathbb{R}^{n_d}$, with n_d designating the dimension of the design variable vector—must be accordingly introduced in the formulation in order to parametrize the network topology and the sensor features. Bayes' theorem in Equation (1) is then modified as:

$$p(\theta|\mathbf{y}, \mathbf{d}) = \frac{p(\mathbf{y}|\theta, \mathbf{d})p(\theta|\mathbf{d})}{p(\mathbf{y}|\mathbf{d})}, \quad (2)$$

where all the previously introduced pdfs are conditioned with respect to the design variable \mathbf{d} , as both the measurements and the parameters to be estimated depend on the experimental settings. Here, the unknown variable \mathbf{d} is supposed to define the spatial configuration of the sensors (e.g., in terms of spatial coordinates), for a constant number of measurements n_y and type of sensors.

According to [18], the expected utility of one experiment can be quantified through:

$$U(\mathbf{d}) = \int_{\mathcal{Y}} \int_{\Theta} u(\mathbf{d}, \mathbf{y}, \theta) p(\theta|\mathbf{y}, \mathbf{d}) p(\mathbf{y}|\mathbf{d}) d\theta d\mathbf{y}, \quad (3)$$

where \mathcal{Y} and Θ respectively represent the domains of the measurements \mathbf{y} and of the parameters θ .

The function $u(\mathbf{d}, \mathbf{y}, \theta)$ is called utility function and defines a scalar measure of the usefulness of the experiment. That is, it quantifies the extent to which certain measurement values are preferable to attain the goal of SHM.

Within a stochastic environment and from a decision-theory perspective, the expected utility allows to choose which action should be performed in order to achieve a certain goal. Therefore, it can be defined as the weighted average of the utilities of each possible consequence of a certain action, wherein the weights describe the probabilities that an action would lead to a certain outcome. For SHM applications, the action is represented by the design of the monitoring system, and the goal is the estimation of the unknown structural parameters.

The choice of $u(\mathbf{d}, \mathbf{y}, \theta)$ depends on the goal of the experiment. A thorough review of utility functions is presented in [20]. In the present case, the aim of the experiment is the inference (estimation) of the parameters θ . Therefore, following [21], a suitable utility function is the Kullback-Leibler divergence (KLD) [22,23] (also called relative entropy) between the prior and the posterior pdfs. Supposing that the structural response to loading is linear and the posterior pdfs are Gaussian, the optimization problem results in the so-called Bayesian D-optimality [24], which corresponds to the maximization of the determinant of the Fisher information matrix [25] of the measurements.

The KLD from P to Q , which are two generic probability distributions of a random variable \mathbf{x} , is defined as:

$$D_{KL}(P||Q) = \int_{\mathcal{X}} p(\mathbf{x}) \ln \frac{p(\mathbf{x})}{q(\mathbf{x})} d\mathbf{x}, \quad (4)$$

where \mathcal{X} is the domain of \mathbf{x} ; $p(\mathbf{x})$ and $q(\mathbf{x})$ are the pdfs related to P and Q . The KLD thus measures the increase of information from Q to P . If the two distributions are identical (i.e., $P = Q$ almost everywhere), then $D_{KL}(P||Q) = 0$.

If the goal of the Bayesian inference problem is the design of the sensor network such that the most possible information is provided by the measurements \mathbf{y} on the parameters θ to be estimated, the design variable \mathbf{d} has to be optimized by maximizing the gain between the prior pdf $p(\theta|\mathbf{d})$ and

the posterior pdf $p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})$. By specializing Equation (4) for the problem considered here, the resulting utility function is given as:

$$u(\mathbf{d}, \mathbf{y}, \boldsymbol{\theta}) = D_{KL} [p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d}) || p(\boldsymbol{\theta}|\mathbf{d})] = \int_{\Theta} p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d}) \ln \frac{p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})}{p(\boldsymbol{\theta}|\mathbf{d})} d\boldsymbol{\theta}. \quad (5)$$

It should be noted that in the integral of Equation (5) the parameter vector $\boldsymbol{\theta}$ serves as a dummy variable. Therefore, $u(\mathbf{d}, \mathbf{y}, \boldsymbol{\theta})$ is not a function of $\boldsymbol{\theta}$. Thus, the expected utility function in Equation (3) can be written as:

$$U(\mathbf{d}) = \int_{\mathbf{y}} \int_{\Theta} p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d}) \ln \frac{p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{d})}{p(\boldsymbol{\theta}|\mathbf{d})} p(\mathbf{y}|\mathbf{d}) d\boldsymbol{\theta} d\mathbf{y}, \quad (6)$$

where $U(\mathbf{d})$ is called the expected Shannon information gain [26] or the Lindley information measure [18].

2.2. Optimal Design of the SHM System

In order to provide a comprehensive strategy to optimize the sensor network, the number of measurements n_y and the pdf p_ϵ of the so-called prediction error $\epsilon \in \mathbb{R}^{n_y}$ are taken into account as unknown variables of the relevant optimization problem.

Let the prediction error ϵ be sampled from a zero mean Gaussian noise $p_\epsilon = \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the covariance matrix (however, in principle the proposed method can be applied to any kind of p_ϵ). The expected Shannon information gain then generally depends on the sensor configuration, number of measurements, and prediction error (i.e., $U = U(\mathbf{d}, n_y, \boldsymbol{\Sigma})$).

Within this framework, the prediction error accounts for the measurement errors related to the sensor characteristics and the model error associated with the intrinsic numerical approximations. Assuming independence between these two error sources, the covariance matrix can be written as:

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_m + \boldsymbol{\Sigma}_n, \quad (7)$$

where $\boldsymbol{\Sigma}_m$ and $\boldsymbol{\Sigma}_n$ respectively account for the model and the measurement error. In [27], it has been shown that the optimal sensor configuration can also be affected by the spatial correlation among different measurements, which can be taken into account in $\boldsymbol{\Sigma}_m$. In practice, the correlation between any couple of measurements decays exponentially with the distance between their locations. A spatial correlation length can be then introduced to constrain the optimal spatial configuration. $\boldsymbol{\Sigma}_n$ can be instead related to the type of sensors to be employed in the SHM system—that is, to the instrumental noise which depends on their characteristics (e.g., signal-to-noise ratio).

For the sake of simplicity, it is assumed that: the sensor type is unique, and so the measurement noise can be accounted for through $\boldsymbol{\Sigma}_n = \sigma^2 \mathbf{I}$; and the model error $\boldsymbol{\Sigma}_m$ is treated as a constant that can affect the optimal configuration, but it is not handled as a further object of the optimization procedure. The resulting optimization statement thus reads:

$$(\mathbf{d}^*, n_y^*, \sigma^*) = \arg \max [U(\mathbf{d}, n_y, \sigma)]. \quad (8)$$

Three types of constraints must be taken into account in the problem:

- identifiability constraint: $n_y > n_{iden}$, where n_{iden} is the minimum number of measurements which are required in order to guarantee the identifiability of the parameters $\boldsymbol{\theta}$ (see [17,28,29]);
- technological constraint: $\sigma > \sigma_{sens}$, where σ_{sens} is the lowest standard deviation of the measurement noise, associated with the sensors available on the market and chosen to measure the structural output;
- cost constraint: $C(n_y, \sigma) \leq B$, where $C(n_y, \sigma)$ is the cost model of the SHM system and B is the maximum budget available for SHM.

The whole optimization problem can therefore be stated as:

$$\begin{aligned}
 (\mathbf{d}^*, n_{\mathbf{y}}^*, \sigma^*) &= \arg \max [U(\mathbf{d}, n_{\mathbf{y}}, \sigma)], \\
 \text{subject to } &\begin{cases} n_{\mathbf{y}} > n_{iden}, \\ \sigma > \sigma_{sens}, \\ C(n_{\mathbf{y}}, \sigma) \leq B. \end{cases} \quad (9)
 \end{aligned}$$

Regarding the cost model $C(n_{\mathbf{y}}, \sigma)$, the simplest possible consists of a combination of a sensor network cost C_0 , which for example includes the data acquisition hardware, database, assemblage, etc., plus a variable cost (i.e., the cost of all the sensors to be deployed over the structure). Accordingly:

$$C(n_{\mathbf{y}}, \sigma) = C_0 + c(\sigma) n_{\mathbf{y}}, \quad (10)$$

where $c(\sigma)$ is the unitary cost per sensor.

In order to solve the optimization problem, a possible approach would be to embed the unknown variables $n_{\mathbf{y}}$ and σ into the design variable vector \mathbf{d} . An alternative approach, which is particularly suitable for real applications if only a limited set of sensor types is available, is to explore the function $\bar{U} = U(\mathbf{d}^*, n_{\mathbf{y}}, \sigma)$ [30], which represents the maximum of the expected Shannon information gain over a search grid of points $\{n_{\mathbf{y}}, \sigma\}$. \mathbf{d}^* is then the optimal configuration obtained by solving the relevant optimization statement, with fixed values of $n_{\mathbf{y}}$ and σ . Since \mathbf{d}^* depends on the choice of $(n_{\mathbf{y}}, \sigma)$, it is possible to conclude that the function $\bar{U} = \bar{U}(n_{\mathbf{y}}, \sigma)$ depends exclusively on $(n_{\mathbf{y}}, \sigma)$.

In place of the preceding formulation, based on budget constraint, a procedure based on a cost–benefit analysis can be followed (see [31,32]). In the problem at hand, the benefit is represented by the expected Shannon information gain. Although $U(\mathbf{d}, n_{\mathbf{y}}, \sigma)$ cannot be directly converted into an expected monetary gain (benefit), it is possible to define a utility–cost index (UCI) through [33]:

$$UCI(\mathbf{d}, n_{\mathbf{y}}, \sigma) = \frac{U(\mathbf{d}, n_{\mathbf{y}}, \sigma)}{C(n_{\mathbf{y}}, \sigma)}, \quad (11)$$

whose measurement unit is [nat/€], [nat] standing for the natural unit of information. The associated optimization problem would then be:

$$\begin{aligned}
 (\mathbf{d}^*, n_{\mathbf{y}}^*, \sigma^*) &= \arg \max \left[\frac{U(\mathbf{d}, n_{\mathbf{y}}, \sigma)}{C(n_{\mathbf{y}}, \sigma)} \right], \\
 \text{subject to } &\begin{cases} n_{\mathbf{y}} > n_{iden}, \\ \sigma > \sigma_{sens}, \\ C(n_{\mathbf{y}}, \sigma) \leq B. \end{cases} \quad (12)
 \end{aligned}$$

This optimization formulation allows the most efficient SHM design to be obtained (i.e., to maximize the information per unitary cost).

The same considerations reported previously hold for the optimization problem in Equation (12). The optimal solutions can be obtained by maximizing the associated objective function $\overline{UCI} = UCI(\mathbf{d}^*, n_{\mathbf{y}}, \sigma)$, where \mathbf{d}^* is the optimal spatial configuration for each $\{n_{\mathbf{y}}, \sigma\}$ set in the search grid.

A comparison between the results of the two strategies defined in Equations (9) and (12) is discussed in Section 6. It is important to underline now that, because the measurements \mathbf{y} depend on the loading conditions, the optimal sensor placement depends on them as well. Therefore, in order to obtain a sensor network design which is robust with respect to the input loading, several optimizations under different loads should be performed, and the final sensor network configuration should be chosen as the one providing the maximum value of the objective function or, alternatively, as a kind of envelope of all the available solutions.

3. Numerical Approach

As explained in Section 2.1, the optimal design of the SHM system is obtained by maximizing the expected Shannon information gain $U(\mathbf{d}, n_y, \sigma)$ (see Equation (9)), or a function related to it (see Equation (12)). At given values of n_y and σ , the optimal experimental design \mathbf{d}^* defines the spatial configuration of the network for which the utility is maximized:

$$\mathbf{d}^* = \arg \max_{\mathbf{d} \in \mathcal{D}} \left[\int_{\mathcal{Y}} \int_{\Theta} p(\theta|\mathbf{y}, \mathbf{d}) \ln \frac{p(\theta|\mathbf{y}, \mathbf{d})}{p(\theta|\mathbf{d})} p(\mathbf{y}|\mathbf{d}) d\theta d\mathbf{y} \right] = \arg \max_{\mathbf{d} \in \mathcal{D}} [U(\mathbf{d})] \quad (13)$$

\mathcal{D} being the design space, which is the domain of all the possible experimental settings (e.g., the locations where the sensors can be placed).

Because the experimental design has to be put in place before performing the measurements, the optimal solution \mathbf{d}^* cannot be found by simply maximizing $U(\mathbf{d}, \mathbf{y}, \theta)$ with respect to \mathbf{y} and θ , which are random variables. The optimal point is instead looked for in the design space \mathcal{D} , by exploring the probability distributions $p(\theta|\mathbf{y}, \mathbf{d})$ and $p(\mathbf{y}|\mathbf{d})$ in the domains \mathcal{Y} and Θ .

In order to solve the optimization problem, a strategy to compute $U(\mathbf{d})$ is needed. Since the double integration in Equation (6) generally cannot be performed analytically, a numerical procedure has to be adopted. Following [34,35] and assuming that $p(\theta|\mathbf{d}) = p(\theta)$ (i.e., that the prior distribution is independent of the design variable), Equation (6) can be approximated through the associated Monte Carlo (MC) estimator:

$$\hat{U}(\mathbf{d}) = \frac{1}{N_{out}} \sum_{i=1}^{N_{out}} \left\{ \ln [p(\mathbf{y}^i|\theta^i, \mathbf{d})] - \ln [p(\mathbf{y}^i|\mathbf{d})] \right\}, \quad (14)$$

where N_{out} is the number of samples θ^i and \mathbf{y}^i to be respectively drawn from $p(\theta)$ and $p(\mathbf{y}|\theta = \theta^i, \mathbf{d})$.

The term $p(\mathbf{y}^i|\mathbf{d})$ can be computed through an analogous MC estimator as:

$$p(\mathbf{y}^i|\mathbf{d}) \simeq \frac{1}{N_{in}} \sum_{j=1}^{N_{in}} p(\mathbf{y}^i|\theta^j, \mathbf{d}), \quad (15)$$

where N_{in} is the number of samples θ^j to be drawn from $p(\theta)$.

The computational cost of such an MC approach can be reduced by using the same batch of samples $\theta^i = \theta^j$ in Equations (14) and (15). The resulting number of likelihood function evaluations then decreases from $N_{in} \times N_{out}$ to $N = N_{in} = N_{out}$ (see [34]). The MC estimator of $\hat{U}(\mathbf{d})$ is then obtained as:

$$\hat{U}(\mathbf{d}) = \frac{1}{N} \sum_{i=1}^N \left\{ \ln [p(\mathbf{y}^i|\theta^i, \mathbf{d})] - \ln \left[\frac{1}{N} \sum_{j=1}^N p(\mathbf{y}^i|\theta^j, \mathbf{d}) \right] \right\}. \quad (16)$$

Model Response

In Equation (16), a major issue is represented by the evaluation of the likelihood function $p(\mathbf{y}^i|\theta^j, \mathbf{d})$. Let the structural system, whose SHM network has to be designed, be subjected to a set of forces and constraints. Since its response to the loads depends on the unknown parameters θ , the measurements can be linked to the design variables in accordance with:

$$\mathbf{y} = \mathbf{L}(\mathbf{d})\mathbf{v}(\theta) + \boldsymbol{\epsilon} = \mathcal{M}(\mathbf{d}, \theta) + \boldsymbol{\epsilon}, \quad (17)$$

where $\mathcal{M}(\mathbf{d}, \theta) : \mathbb{R}^{n_d} \times \mathbb{R}^{n_\theta} \rightarrow \mathbb{R}^{n_y}$ is the forward model operator which relates the model inputs (i.e., the design variables \mathbf{d} and θ) with the measurements \mathbf{y} under the considered loading, $\mathbf{L} \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_{dof}}$ is a Boolean operator which aims at selecting from \mathbf{v} the actually measured response components, n_{dof} is the number of degrees of freedom (DOFs) of the numerical model, and $\mathbf{v} \in \mathbb{R}^{n_{dof}}$

is the structural response (e.g., displacements, rotations, etc.) of the numerical model for all the n_{dof} degrees of freedom.

Following [28], the likelihood function can then be expressed as:

$$p(\mathbf{y}^i | \boldsymbol{\theta}^i, \mathbf{d}) = p_{\epsilon}(\mathbf{y}^i - \mathcal{M}(\mathbf{d}, \boldsymbol{\theta}^i)). \quad (18)$$

It can be underlined that the same approach can be applied to dynamic cases as well. To this end, only Equation (17) has to be modified as follows:

$$\mathbf{y} = \mathbf{L}(\mathbf{d})\boldsymbol{\Phi}(\boldsymbol{\theta}) + \boldsymbol{\epsilon}, \quad (19)$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{n_{dof}} \times \mathbb{R}^{N_m}$ is a matrix containing N_m modal shapes of the structure and \mathbf{y} are the relevant measurements. Equation (18) is accordingly changed to [28]:

$$p(\mathbf{y}^i | \boldsymbol{\theta}^i, \mathbf{d}) = \prod_{m=1}^{N_m} p_{\epsilon}(\mathbf{y}^i - \mathbf{L}(\mathbf{d})\boldsymbol{\Phi}_m(\boldsymbol{\theta}^i)). \quad (20)$$

Apart from the definition of the likelihood function in Equation (20), the rest of the framework also remains valid for the dynamic case. Since this application goes beyond the scope of the present paper, future work will be devoted to the implementation of such an approach to dynamic testing.

Returning to the static case and considering Equations (16) and (18) and knowing that $\mathbf{y}^i = \mathcal{M}(\mathbf{d}, \boldsymbol{\theta}^i) + \boldsymbol{\epsilon}^i$ (see Equation (17)), the MC estimator of the expected Shannon information gain is obtained as:

$$\hat{U}(\mathbf{d}) = \frac{1}{N} \sum_{i=1}^N \left\{ \ln [p_{\epsilon}(\boldsymbol{\epsilon}^i)] \right\} - \frac{1}{N} \sum_{i=1}^N \left\{ \ln \left[\frac{1}{N} \sum_{j=1}^N p_{\epsilon}(\mathcal{M}(\mathbf{d}, \boldsymbol{\theta}^j) + \boldsymbol{\epsilon}^i - \mathcal{M}(\mathbf{d}, \boldsymbol{\theta}^j)) \right] \right\}. \quad (21)$$

The MC estimator is thus offered as the sum of two terms: the first one depends on the prediction error only; the second one depends on the design variables as well as on the parameters.

If the prediction error $\boldsymbol{\epsilon}$ can be assumed independent of the design variable \mathbf{d} (i.e., if $p(\boldsymbol{\theta} | \mathbf{d}) = p(\boldsymbol{\theta})$), the first term in Equation (21) turns out to be independent of \mathbf{d} . Therefore, it can be dropped from the computation, as we are interested only in the design vector \mathbf{d}^* providing the optimum. In this way, the computing time required for the evaluation of the objective function would be significantly reduced. This case occurs, for example, if a unique type of sensor is planned to be installed on the structure. If the standard deviation of measurements depends on the design variable (i.e., if $\sigma = \sigma(\mathbf{d})$), the first term must be kept in the objective function because it affects the optimal solution.

4. Surrogate Modeling

The computational cost of the MC estimator in Equation (21), as is true of any other MC analysis, may be attributed to the repeated evaluation of the model response $\mathcal{M}(\mathbf{d}, \boldsymbol{\theta}^i)$, for each of the N samples $\boldsymbol{\theta}^i = \boldsymbol{\theta}^j$ drawn from $p(\boldsymbol{\theta})$. From a practical point of view, the computation of $\hat{U}(\mathbf{d})$ can become infeasible due to the high number of degrees of freedom (DOFs) in the numerical models (e.g., of real-life structures).

To reduce the overall computational costs of the evaluation of the model response in the optimization procedure, the exploitation of surrogate models has been proposed in [35] and applied to SHM sensor network optimization in [36]. A surrogate model (or metamodel) is aimed at providing the relationship between input and output through a more computationally efficient formulation. These approaches can therefore be classified as data-driven ones: the underlying physics of the problem is lost. Accordingly, if the physical behavior of the problem changes, a new surrogate model should be built by using the new relevant input–output data. Alternative methods to reduce the computational cost of model evaluation are model-based (e.g., [37,38]).

One of the most widely exploited types of surrogate models is based on polynomial chaos expansion (PCE). PCE was first introduced in [15,39] for standard Gaussian random variables, and was then generalized to other probability distributions in [40–42].

Following the investigation in [43], it is assumed that the input random vector is constituted by the unknown parameters only, featuring a joint pdf $p(\boldsymbol{\theta})$, whereas the design variable \mathbf{d} is not considered in the surrogate model.

Assuming a finite variance model, the PCE of the response \mathbf{v} in Equation (17) reads:

$$v_k = \mathcal{M}_k(\boldsymbol{\theta}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} \phi_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \quad k = 1, \dots, n_{dof}, \quad (22)$$

where $\Psi_{\boldsymbol{\alpha}}$ are multivariate polynomials which are orthonormal with respect to $p(\boldsymbol{\theta})$; $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_M\} \in \mathbb{N}^M$ is a multi-index associated with the components of Ψ , and $\phi_{\boldsymbol{\alpha}} \in \mathbb{R}$ are the related coefficients. For real-life applications, the sum in Equation (22) is truncated by retaining only those polynomials whose total degree $|\boldsymbol{\alpha}|$ is less than a certain value p :

$$v_k \simeq \mathcal{M}_k^{PCE}(\boldsymbol{\chi}) = \sum_{|\boldsymbol{\alpha}| \leq p} \phi_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\theta}) \quad k = 1, \dots, n_{dof}, \quad (23)$$

where $|\boldsymbol{\alpha}| = \sum_{i=1}^M \alpha_i$ and \mathcal{M}_k^{PCE} is the surrogate model. The response \mathbf{v} can be approximated in a component-wise fashion by building a set of n_y PCE surrogate models according to:

$$\mathbf{v} \simeq \mathcal{M}^{PCE}(\boldsymbol{\theta}) = \begin{Bmatrix} \mathcal{M}_1^{PCE}(\boldsymbol{\theta}) \\ \vdots \\ \mathcal{M}_{n_{dof}}^{PCE}(\boldsymbol{\theta}) \end{Bmatrix}. \quad (24)$$

In order to compute the unknown polynomial coefficients $\phi_{\boldsymbol{\alpha}}$ for each surrogate, both intrusive and non-intrusive methods can be adopted [44]. Intrusive approaches rely on the projection of the original computational model onto the subspace spanned by the PCE, through a Galerkin projection [45]. In such methods, the variables in the governing equations are replaced by their polynomial chaos expansions. For instance, for linear structural problems, the stiffness matrix and the response vector are approximated through a truncated expansion, leading to a linear system of equations to be solved [46]. While these methods demonstrate an increase of the computational cost which is linear with the number of basis polynomials, they are not suitable for the current purposes, since they require the custom modification of the computational solver.

Non-intrusive methods instead allow the bases to be computed by simply processing a batch of sampled input variables $\boldsymbol{\theta}$ and the corresponding model evaluations \mathbf{v} , which form the so-called experimental design. No manipulations of the solver are needed, rendering this approach particularly suitable for general-purpose problems. Two methods can be used for non-intrusively computing the coefficients: the projection approach [47,48], where the computation of each coefficient is formulated as a multi-dimensional integral; and least-squares minimization [49].

The latter method is employed here, since an arbitrary number of samples can be used in order to estimate the coefficients of the expansion in Equation (23). The corresponding formulation of the least-square minimization problem is:

$$\hat{\boldsymbol{\phi}} = \arg \min_{\boldsymbol{\phi}} \mathbb{E} \left\{ \left[\boldsymbol{\phi}^T \boldsymbol{\Psi}(\boldsymbol{\chi}) - \mathcal{M}(\boldsymbol{\chi}) \right]^2 \right\}, \quad (25)$$

where $\hat{\boldsymbol{\phi}}$ is the whole set of coefficients to be estimated. In order to reduce the computational cost of the least-squares approach, a method based on least angle regression introduced in [50,51] is adopted. The method relies on the selection of the most significant coefficients of the PC expansion, allowing a

reduction in the number of model evaluations, which are required to build the experimental design for the coefficient estimation.

Following the previously described non-intrusive method, a set of $N^{PCE} \ll N$ samples of the input variable θ must be drawn from $p(\theta)$, and the corresponding model responses \mathbf{v} are numerically computed through the model $\mathcal{M}(\theta, \mathbf{d})$. Once the metamodel \mathcal{M}^{PCE} is built, the N samples required for the estimation of $\hat{U}(\mathbf{d})$ can be computed through the surrogate. The number N^{PCE} of input–output samples needed to build such a surrogate model should be chosen by considering the required accuracy of the metamodel in predicting the response of the original model.

According to the adopted formulation for surrogate modeling, n_{dof} PCE surrogates would be required (see Equation (24)), thus making the computation unbearable. Dimensionality reduction strategies are further required to overcome this computational issue. Principal component analysis (PCA) offers a statistical tool for handling large datasets, first introduced by Pearson [52] and Hotelling [53], and later developed in [54,55] for different fields of application: Karhunen–Loeve decomposition (KLD) [56,57] in signal processing, proper orthogonal decomposition (POD) [58] in mathematics, and singular value decomposition (SVD) in mechanical engineering [59]. Some examples of the application of POD in structural health monitoring can be found in [60–63], where the method has been employed for the order reduction of dynamical models feeding Bayesian updating schemes.

PCA allows the computational burden to be reduced according to the following procedure. Let the model parameters (i.e., the input variables of the surrogate) be sampled from the prior pdf: $\theta_i \sim p(\theta)$, with $i = 1, \dots, N^{PCE}$. Compute the response vectors, which are instead the output variable of the surrogate, as $\mathbf{v}_i = \mathbf{v}(\theta_i)$ through the full-order numerical model, building the so-called experimental design of the surrogate model. The model response data are gathered in the matrix $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_{N^{PCE}}] \in \mathbb{R}^{n_{dof} \times N^{PCE}}$. \mathbf{V} is projected onto a new space of n_{dof} of uncorrelated variables:

$$\mathbf{T} = \mathbf{W}\mathbf{V}, \quad (26)$$

where $\mathbf{W} \in \mathbb{R}^{n_{dof} \times n_{dof}}$ is a square orthogonal matrix, whose rows are the eigenvectors of the matrix $\mathbf{V}^T\mathbf{V}$ and form an orthogonal basis; $\mathbf{T} \in \mathbb{R}^{n_{dof} \times N^{PCE}}$ is the matrix of the principal component scores (i.e., the representation of \mathbf{V} in the principal component space).

The dimension of the response matrix is reduced through the PCA by retaining only the first $l \ll n_{dof}$ components in the solution:

$$\mathbf{T}_l = \mathbf{W}_l\mathbf{V}, \quad (27)$$

where $\mathbf{T}_l \in \mathbb{R}^{l \times N^{PCE}}$ is the reduced-order response matrix and $\mathbf{W}_l \in \mathbb{R}^{l \times n_{dof}}$ is orthonormal.

The formulation defined in Equation (24) is then modified by setting the output of the PCE surrogate model as the reduced-dimension response vector, thus establishing a relation between the parameters θ and the first principal components of \mathbf{v} .

In conclusion, by combining the PCE surrogate and the PCA dimensionality reduction technique, the original model response $\mathcal{M}(\theta, \mathbf{d})$ can be approximated through:

$$\mathcal{M}(\theta, \mathbf{d}) \cong \mathcal{M}^{META}(\theta, \mathbf{d}) = \mathbf{L}(\mathbf{d})\mathbf{W}_l^T \mathcal{M}^{PCE}(\theta), \quad (28)$$

where $\mathcal{M}^{META}(\theta, \mathbf{d})$ is the PCA-PCE-based metamodel.

The resulting MC estimator then reads:

$$\hat{U}(\mathbf{d}) = \frac{1}{N} \sum_{i=1}^N \left\{ \ln \left[p_\epsilon(\epsilon^i) \right] \right\} - \frac{1}{N} \sum_{i=1}^N \left\{ \ln \left[\frac{1}{N} \sum_{j=1}^N p_\epsilon \left(\mathbf{L}(\mathbf{d})\mathbf{W}_l^T \mathcal{M}^{PCE}(\theta^j) + \epsilon^i - \mathbf{L}(\mathbf{d})\mathbf{W}_l^T \mathcal{M}^{PCE}(\theta^j) \right) \right] \right\}, \quad (29)$$

where, according to the formulation presented, the design variable \mathbf{d} that defines the spatial configuration of the sensor network is defined as follows:

$$\mathbf{d} = \left\{ \begin{array}{c} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_s \\ \vdots \\ \mathbf{d}_{n_y} \end{array} \right\}, \quad (30)$$

where:

$$\mathbf{d}_s = \{x_1^s \ x_2^s \ x_3^s \ \delta_s\}^T, \quad (31)$$

and x_1^s , x_2^s , x_3^s are the coordinates of the location where the s -th measurement is supposed to be taken. δ_s is a scalar integer value which defines the measured DOF—either a displacement or a rotation.

A possible alternative formulation relies on the nodal labeling of the numerical model DOFs. Despite the beneficial dimension reduction that can be reached, the adoption of this formulation would be detrimental in the solution of the optimization problem, as it would lead to consistent discontinuities of the objective function in the associated search space.

5. Optimization Procedure

The proposed procedure for optimal sensor placement is based on the estimation of the expected Shannon information gain $U(\mathbf{d})$ through the MC estimator $\hat{U}(\mathbf{d})$.

According to [7], a sequential strategy can be adopted to solve the problem. At each iteration, only the position of one sensor is optimized, while all other sensors deployed in the previous algorithm steps are held fixed. This strategy was termed forward sequential sensor placement. On the contrary, with the backward sequential sensor placement strategy, the initial configuration is populated with sensors at all of the nodes, and they are later dropped from the optimal configuration one-by-one. In this regard, such sequential strategies are expected to yield sub-optimal solutions, since they cannot guarantee that the optimal solution (i.e., the global maximum of the objective function) is attained. Independently of the method adopted, the iterations are stopped when the desired number n_y of sensors are placed over the structure.

Since the estimator $\hat{U}(\mathbf{d})$ is based on Monte Carlo sampling of measurement error $\epsilon \sim p_\epsilon$ and parameters vector $\theta \sim p(\theta|\mathbf{d})$, the resulting objective function becomes noisy. As discussed in Section 2.1, the prior pdf can be assumed to be independent of the position of the sensors (i.e., $\theta \sim p(\theta|\mathbf{d}) = p(\theta)$). The same batch of samples θ can be used for each realization of \mathbf{d} , resulting in a less-noisy objective function. It should be underlined that, following this assumption, the objective function will be affected by a constant bias, which therefore will not influence the resulting optimal solutions in terms of sensor configuration. Moreover, since there is no need to re-sample θ and compute the corresponding structural response for each different sample of \mathbf{d} at each iteration of the optimization procedure, a consistent reduction in the overall computational cost is achieved.

Due to the noisy objective function, in order to avoid the attainment of a false local optimum, the covariance matrix adaptation evolution strategy (CMA-ES) [64] is adopted here. It is an iterative evolutionary derivative-free algorithm that is suitable for stochastic optimization problems, introduced in [65,66].

The pseudo-code of the CMA-ES is listed in Algorithm 1. The algorithm is based on an evolutionary strategy, where at each iteration i , a total number N^{opt} of samples \mathbf{d} are drawn from a multivariate normal distribution $\mathbf{d}_j \sim \mathbf{m} + \sigma_c \mathcal{N}(\mathbf{0}, \mathbf{C})$, where $\mathbf{C} \in \mathbb{R}^{n_d \times n_d}$ is the covariance matrix, $\mathbf{m} \in \mathbb{R}^{n_d}$ is the mean of the design points distribution, and σ_c is the step size. Then, the values of \mathbf{m} , \mathbf{C} , and σ_c are updated in order for the population of new points $\mathbf{d}_{1:N^{opt}}$ to move towards the maximum of the objective function $\hat{U}(\mathbf{d})$. The evolution of the design variable (i.e., the sequence of consecutive steps

of the mean \mathbf{m}) is performed through the so-called cumulation technique, detailed in Algorithm 1, moving from the initial condition \mathbf{d}_0 , and μ and μ_w are parameters needed to control the update phase. In Algorithm 1, $c_c, c_{\sigma_c}, c_1, d_{\sigma_c}$ are parameters which control the optimization procedure and have to be set empirically for each numerical application. The iterations are stopped whenever at least one of the following criteria are fulfilled:

$$\begin{cases} |\hat{U}(\mathbf{d}_k) - \hat{U}(\mathbf{d}_{k-1})| \leq \rho_U, \\ \|\mathbf{d}_k - \mathbf{d}_{k-1}\| \leq \rho_d, \end{cases} \quad (32)$$

where the symbol $|\cdot|$ stands for the absolute value of the argument; $\|\cdot\|$ represents an appropriately chosen norm of vectors (e.g., the L^2 norm used here); and ρ_U and ρ_d are parameters that tolerances the accuracy of the solution in terms of objective function and design variable, respectively. These parameters cannot be chosen a priori, as they are dependent on the specific application, on the model discretization, and on the desired accuracy. For further details on the algorithm, the interested reader may refer to [16,67].

Algorithm 1 Covariance matrix adaptation evolution strategy

Input : $\mathbf{m} \in \mathbb{R}^n, \sigma_c \in \mathbb{R}_+, N^{opt}, \mathbf{d}_0$

Output: Optimal configuration \mathbf{d}^*

Initialize parameters $c_c, c_{\sigma_c}, c_1, d_{\sigma_c}, i$

while Stopping criteria in Equation (32) are not met **do**

for $j = 1, \dots, N^{opt}$ **do**

 Sample the design variable: $\mathbf{d}_j = \mathbf{m} + \sigma_c \gamma_j, \gamma_j \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$

 Evaluate the objective function through the estimator in Equation (16): $\hat{U}_j = \hat{U}(\mathbf{d}_j)$

end

Update parameters

begin

 Sort samples: $\hat{U}(\mathbf{d}_{1:N^{opt}}) \leq \dots \leq \hat{U}(\mathbf{d}_{i:N^{opt}}) \dots \leq \hat{U}(\mathbf{d}_{N^{opt}:N^{opt}})$

 Update mean: $\mathbf{m} = \sum_{i=1}^{\mu} w_i \mathbf{d}_{i:N^{opt}} = \mathbf{m} + \sigma_c \gamma_w$, with $\gamma_w = \sum_{i=1}^{\mu} w_i \gamma_{i:N^{opt}}$

 Cumulation for \mathbf{C} : $\mathbf{p}_c^i = (1 - c_c) \mathbf{p}_c^{i-1} + \sqrt{1 - (1 - c_c)^2} \sqrt{\mu_w} \gamma_w$

 Cumulation for σ_c : $\mathbf{p}_{\sigma_c}^i = (1 - c_{\sigma_c}) \mathbf{p}_{\sigma_c}^{i-1} + \sqrt{1 - (1 - c_{\sigma_c})^2} \sqrt{\mu_w} (\mathbf{C}^{i-1})^{-1/2}$

 Update \mathbf{C} : $\mathbf{C}^i = (1 - c_1 - c_{\mu}) \mathbf{C}^{i-1} + c_1 + \mathbf{p}_c^{i-1} \mathbf{p}_c^{i-1T} + c_{\mu} \sum_{i=1}^{\mu} w_i \gamma_{i:N^{opt}} \gamma_{i:N^{opt}}^T$

 Update σ_c : $\sigma_c^i = \sigma_c^{i-1} e^{\frac{c_{\sigma_c}}{d_{\sigma_c}} \left(\frac{\|\mathbf{p}_{\sigma_c}^i\|}{\|\mathbf{p}_{\sigma_c}^{i-1}\|} - 1 \right)}$

 Update $i = i + 1$

end

end

The overall procedure for computing the optimal sensor configuration is listed in Algorithm 2, and the corresponding flowchart is shown in Figure 1. First, the parameter vector θ is sampled from the prior pdf $p(\theta)$, which is chosen a priori. For each sample θ_i (with $i = 1, \dots, N^{PCE}$), the corresponding response $\mathbf{v}_i = \mathbf{v}(\theta_i)$ is computed through the numerical model. Then, the dimension of the response vector is reduced from n_{dof} to l by performing the PCA of $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_{N^{PCE}}]$. A total number l of model surrogates is built by considering θ as the input variable and the components of the reduced-space vector \mathbf{T}_l as the output variables. A fresh batch of $N \gg N^{PCE}$ samples θ is drawn from prior $p(\theta)$, and the corresponding system response is computed through the PCE surrogates. In the end, the optimal configuration is obtained through the CMA-ES optimization method listed in Algorithm 1, where the evaluation of the objective function is performed through the MC estimator defined in Equation (16).

Algorithm 2 Algorithm for the optimization of SHM sensor networks through Bayesian experimental design

Input : Prior pdf $p(\boldsymbol{\theta})$
Output: Optimal configuration \mathbf{d}^*
for $i = 1 : N^{PCE}$ **do**
 Sample parameters values: $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$
 Compute system response: $\mathbf{v}_i = \mathbf{v}(\boldsymbol{\theta}_i)$
end
Perform PCA: find $\mathbf{W}_l, \mathbf{T}_l$ s.t. $[\mathbf{v}_1 \dots \mathbf{v}_{N^{PCE}}] = \mathbf{V} \cong \mathbf{W}_l^T \mathbf{T}_l$
Compute surrogate model $\mathcal{M}^{PCE}(\boldsymbol{\theta})$, which links the inputs $\boldsymbol{\theta}$ with the output \mathbf{v}_l
for $i = 1 : N$ **do**
 Sample parameters values: $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$
 Compute system response: $\mathbf{v}_i = \mathbf{W}_l^T \mathcal{M}^{PCE}(\boldsymbol{\theta}_i)$
end
Optimization: run Algorithm 1

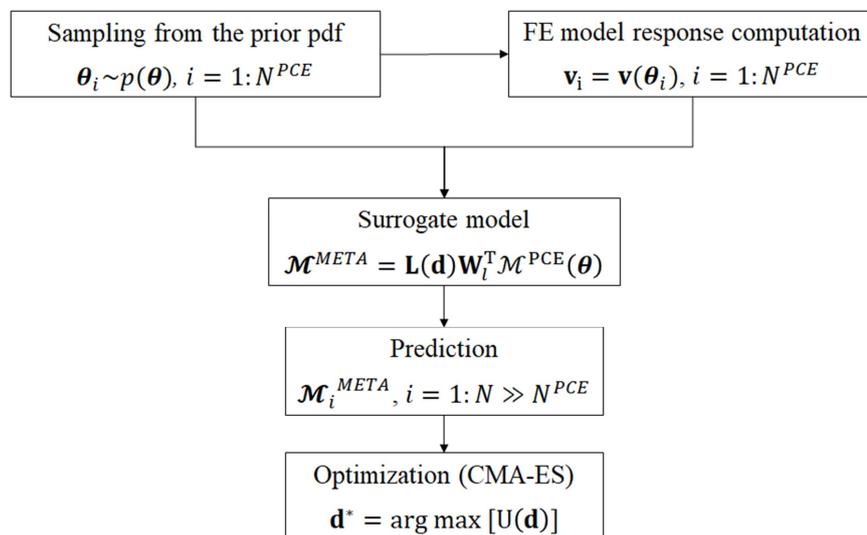


Figure 1. Flowchart of the proposed procedure. CMA-ES: covariance matrix adaptation evolution strategy; FE: finite element.

6. Results: Application to the Monitoring of a Tall Building

The framework described in the present work was applied to a real large-scale structure: the Pirelli Tower, a 130-m-tall building in Milan (Italy). The building consists of 35 stories out of ground, which are approximately 70 m long and 30 m wide (see Figure 2). The structural system is entirely made of reinforced concrete. Four symmetric triangular cores at the two extremities of each storey are connected by T-shaped beams (see Figure 2b). The structure was modeled using the commercial software SAP2000 v19 (Computer and Structures, Inc., Berkeley, CA, USA) and the associated finite element (FE) model, shown in Figure 2a, consists of 4106 nodes with 6 DOFs each (the 3 displacements $u_{x_1}, u_{x_2}, u_{x_3}$, and the 3 rotations $\varphi_{x_1}, \varphi_{x_2}, \varphi_{x_3}$ about the axes of the reported orthonormal reference frame), resulting in a total number of DOFs $n_{dof} = 24,500$. The structure is supposed to be subjected to a horizontal force acting in the x_2 direction on the top floor (Figure 2b), see also [14,38,68,69]. The force was assumed to be eccentric, as shown in Figure 2b, in order to induce a complex bending-torsional mechanical response of the tower. For further details on the structural characteristics and on the FE model, the readers may refer to [68,70,71].

The sensor network was assumed to be optimized in terms of the $n_\theta = 6$ parameters, which are listed in Table 1. The parameters were chosen to render the example as general as possible.

Both mechanical and geometrical properties are handled, associated to both vertical and horizontal structural members. The chosen parameters were the Young's moduli of column groups LC and RC, the Young's moduli of beam groups LB, CB, and RB, and the beam thickness of group CB. The prior pdfs of each parameter are also shown in Table 1. The prior pdfs of the concrete Young's modulus were assumed to be uniform, with lower and upper bounds respectively equal to 24 GPa and 36 GPa. The prior pdf of the beam thickness was considered to be uniform as well, with lower and upper bounds respectively equal to 0.7 m and 0.9 m.

Since the structural model features both displacement and rotation DOFs at each node, the design variable \mathbf{d} must be defined such that both the spatial position of the sensors and the physical quantity to be measured are taken into account, in accordance with Equations (30) and (31).

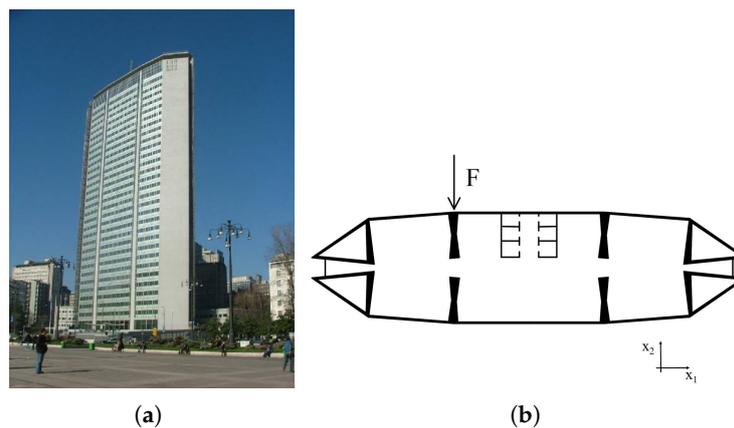


Figure 2. Structural details of the Pirelli Tower: (a) 3D view and (b) plan representation.

Table 1. Definition of parameters θ (see Figure 2) and related prior probability density function (pdf) $p(\theta)$.

Position	Physical Quantity	Prior pdf
20th floor left columns (LC)	Young's modulus E (GPa)	$\mathcal{U}(24, 36)$
20th floor right columns (RC)	Young's modulus E (GPa)	$\mathcal{U}(24, 36)$
20th floor left beams (LB)	Young's modulus E (GPa)	$\mathcal{U}(24, 36)$
20th floor right beams (RB)	Young's modulus E (GPa)	$\mathcal{U}(24, 36)$
20th floor central beams (CB)	Young's modulus E (GPa)	$\mathcal{U}(24, 36)$
20th floor central beams (CB)	Beam thickness t (m)	$\mathcal{U}(0.7, 0.9)$

As discussed in Section 2.2, in the optimization procedure it is assumed that only the standard deviation σ associated with the measurement error can be varied, while the model error is supposed to be a constant term. As σ is assumed to be dependent on the sensor characteristics, we also aim to provide a procedure which allows the optimal type of sensor to be chosen to better estimate the chosen parameters, such as possible variations of the estimated properties from the initial health state of the structure.

The contour plot of the objective function $\bar{U}(n_y, \sigma) = U(\mathbf{d}^*, n_y, \sigma)$ is shown in Figure 3. Here, the objective function is computed at the following discrete points of the grid $(n_y, \sigma^2) = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \times \{10^{-8}, 10^{-7.5}, 10^{-7}, 10^{-6.5}, 10^{-6}\}$. As expected, the maximum value of the expected Shannon information gain increases as the number of sensors increases, as analytically proven in [6] and numerically shown in [72], while the standard deviations decreases, since more information is provided by the SHM system. The associated optimal sensor configuration, which corresponds to the maximum of the objective functions, is shown in Figure 4. A further discussion of the optimal sensor placement with this method can be found in [43].

It can be also observed that the increase in the expected Shannon information gain decreases as more measurements are allowed for. The quantity $\frac{\partial U}{\partial n_y}$ is therefore a decreasing function of n_y .

From a decision-making perspective, it is interesting to underline that this behaviour corresponds to the so-called “law of diminishing marginal utility”, also known as Gossen’s First Law [73], which is used in economics for the optimization of resource allocation. The law states that the marginal utility of each unit decreases as the supply of units increases. In the problem of optimal SHM system design, the utility of the sensor network is quantified by the expected Shannon information gain [18], and the unit is represented by each measurement. Applications of this law to sensor network optimization in different engineering fields can also be found in [74–76].

As a simple linear cost model is assumed (see Equation (10)), the red lines in Figure 3 represent different budget constraints (i.e., the solutions $\{\sigma, n_y\}$ of the equation $B = C_0 + c(\sigma) n_y$, where B is the total available budget). By using this chart, it is possible to optimally design the SHM network: the optimal point $\{\sigma^*, n_y^*\}$, for which $\bar{U}(n_y, \sigma)$ is maximum, is ruled by the available budgetary constraint and it is uniquely associated with the corresponding optimal configuration \mathbf{d}^* .

A different approach for decision-making is to define a Pareto-like graph, as shown in Figure 5. Each line corresponds to the optimal design for a certain standard deviation (i.e., a certain type of sensors). The cost saving is defined in order to normalize the cost function with respect to the chosen budget. Accordingly, any solution point located at the left of the vertical line represents a non-optimal design solution, since the associated cost does not correspond to the best choice of $\{\mathbf{d}, n_y, \sigma\}$. The oscillations of the objective functions in Figure 5 are due to the possible presence of local maxima. Although this problem cannot be solved a priori, it can be mitigated by running the optimization algorithm several times with different initial points \mathbf{d}_0 , and choosing the optimal configuration which corresponds to the maximum value of the objective function among the correspondent solutions.

This graph can be particularly useful in appropriately allocating economic resources. For a chosen budget, it is possible to select the type of sensors which results in the highest accuracy as indicated in Figure 5, the number of sensors, and their location, associated with the maximum possible expected information gain. The trend of each Pareto front provides an indication about the change of maximum utility due to a variation of budget, and thus it helps to decide if additional spending is justified. Moreover, given the value of U , it is possible to compare, from an economic point of view, different solutions in terms of the number and type of sensors.

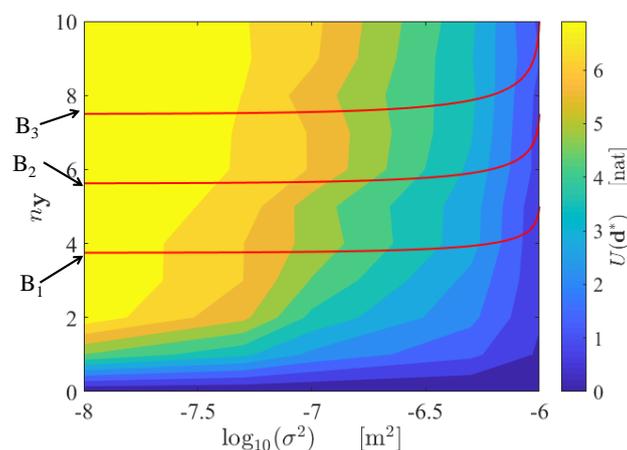


Figure 3. Contour plot of $\bar{U}(n_y, \sigma)$, and curves representing the budget constraints $B = C(\sigma, n_y)$, with $B_1 = 2000$ €, $B_2 = 2500$ €, $B_3 = 3000$ €.

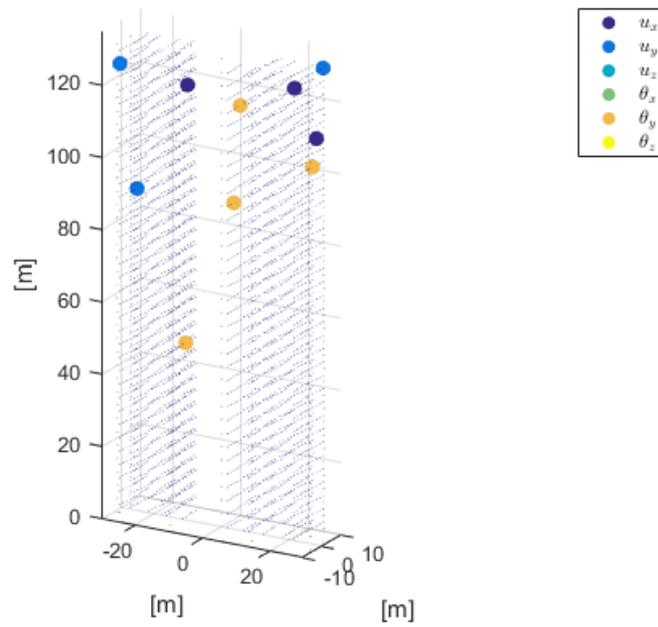


Figure 4. Optimal sensor placement and physical quantity to be measured, with $n_y = 10$ and $\sigma = 10^{-8}$.

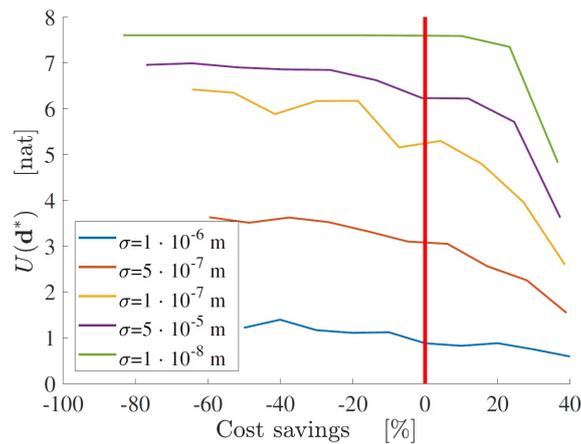


Figure 5. Pareto fronts of the structural health monitoring (SHM) sensor network optimization problem, for different values of standard deviation σ .

An alternative design approach discussed in Section 2.2 with Equation (11) is based on the maximization of the ratio $UCI(n_y, \sigma) = \frac{U(\mathbf{d}^*, n_y, \sigma)}{C(n_y, \sigma)}$. The resulting optimal solution depends on the cost model. In Figure 6a the SHM system is supposed to have a low initial cost (i.e., $C_0 = 500$ €). In Figure 6b the SHM system is supposed to have a high initial cost (i.e., $C_0 = 1000$ €). In both cases, the most efficient employment of resources is reached if the best sensor in terms of measurement noise is chosen, while the optimal number of sensors depends on the cost model.

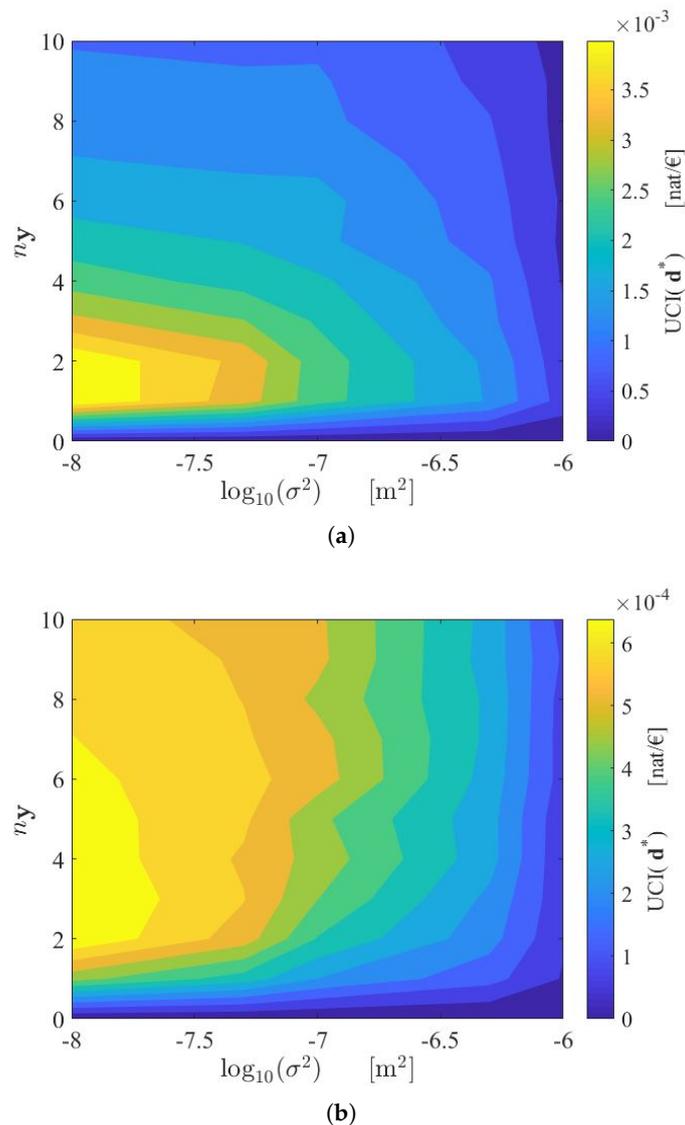


Figure 6. Contour plot of $UCI(n_y, \sigma) = \frac{\bar{U}(\mathbf{d}^*, n_y, \sigma)}{C(n_y, \sigma)}$, with (a) $C_0 = 500$ € and (b) $C_0 = 1000$ €. UCI: utility–cost index.

Note that while the objective function $\bar{U}(n_y, \sigma)$ always increases with n_y and σ , the function $UCI(n_y, \sigma)$ presents a maximum for $n_y < \infty$. This is because, as previously discussed, the increase in information associated with each additional sensor decreases as more sensors are added to the monitoring system. From a cost–benefit point of view, it is therefore worthless to add sensors (i.e., to increase the SHM cost) if the resulting benefit in terms, for example, of the additional expected Shannon information gain is very low.

7. Conclusions

The present paper presents a stochastic cost–benefit methodology to optimally design structural health monitoring systems.

The benefit or usefulness of the SHM system is quantified through the expected Shannon information gain between the prior and the posterior pdfs of the parameters to be estimated. By maximizing this objective function, it is possible to choose the best position, type, and number of sensors, which guarantees the minimization of the uncertainties associated with the quantities to be estimated, or in other words, the maximization of the information obtained through the measurements.

The objective function can be numerically approximated through a Monte Carlo sampling approach. The resulting estimator is expressed as a double sum of terms, which depend on the likelihood function. Since a high number of model response evaluations is required, a procedure based on surrogate models and model order reduction strategies is proposed. The combination of the PCE surrogate model and a model order reduction technique (PCA) allows a computationally efficient meta-model to be built in order to mimic the relation between input and output variables. Since the resulting objective function is affected by noise, leading to possible undesired local maxima, an evolutionary strategy (CMA-ES) suitable for stochastic problems is used.

In order to find the optimal solution, the cost, identifiability of the model parameters, and technological constraints have to be taken into account. A further optimization problem is considered, established by maximizing the information gain per unitary cost by means of a cost–benefit analysis.

Application of the framework to a large-scale numerical model demonstrates that the maximum expected Shannon information gain of the SHM system increases as more sensors are added to the system and lower standard deviations of the prediction error are considered. The optimal solution, in terms of maximum information gain, does not necessarily correspond to the most efficient one (see Figure 3), in terms of the ratio between information and cost. This is because the increase in information gain due to additional sensors is reduced as more measurements are considered. A Pareto-front approach can also be followed in order to choose the best solution, both in terms of maximum information and minimum cost (Figure 5).

An alternative procedure based on the maximization of the utility–cost ratio can be implemented to optimally allocate the available resources. In this case, the optimal solution depends on the variation of the sensor network cost with respect to the number of measurements and the sensor type (see Figure 6a,b). It is worth noting that the same consideration also holds for the case where only a few types of sensors are available, and therefore if it is not possible to establish a cost model. The optimization can be performed in the same way. That is, by computing the maximum values of the objective function (which correspond to the optimal spatial configurations) over the discrete search grid.

The proposed strategy is completely non-intrusive, in that it does not require computation of the gradient of the objective function, but instead exclusively relies on evaluations of model response. Moreover, the method is general and no restrictive assumptions, such as linearity or Gaussianity, are placed.

Further work will be dedicated to the application of this framework to dynamic testing and to more complex structural models.

Author Contributions: Authors contributed equally to this work.

Funding: Eleni Chatzi would like to acknowledge the support of the ERC Starting Grant award WINDMIL (ERC-2015-StG # 679843) on the topic of “Smart Monitoring, Inspection and Life-Cycle Assessment of Wind Turbines”.

Acknowledgments: Giovanni Capellari acknowledges the financial support by IDEA League through a Ph.D. Student Grant. The authors also acknowledge the Chair of Risk, Safety and Uncertainty Quantification and the Computational Science and Engineering Laboratory at ETH Zürich for having provided the MATLAB-based software UQLab and CMA-ES, used in the implementation of the method. Authors are indebted to Gianluca Barbella and Federico Perotti, who provided the numerical model of the Pirelli tower.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SHM	Structural Health Monitoring
PCE	Polynomial Chaos Expansion
PCA	Principal Component Analysis
CMA-ES	Covariance Matrix Adaptation Evolutionary Strategy
KLD	Kullback–Leibler Divergence
UCI	Utility-Cost Index
MC	Monte Carlo
MI	Mutual Information
DOF	Degree of Freedom
FE	Finite Element

References

- Balageas, D.; Fritzen, C.P.; Güemes, A. *Structural Health Monitoring*; Wiley-ISTE: London, UK, 2006.
- Farrar, C.; Worden, K. An introduction to structural health monitoring. *Philos. Trans. R. Soc. A* **2007**, *365*, 303–315. [[CrossRef](#)] [[PubMed](#)]
- Beck, J.; Katafygiotis, L. Updating models and their uncertainties. I: Bayesian statistical framework. *J. Eng. Mech.* **1998**, *124*, 455–461. [[CrossRef](#)]
- Heredia-Zavoni, E.; Esteva, L. Optimal instrumentation of uncertain structural systems subject to earthquake ground motions. *Earthq. Eng. Struct. Dyn.* **1998**, *27*, 343–362. [[CrossRef](#)]
- Sobczyk, K. Theoretic information approach to identification and signal processing. *Reliab. Optim. Struct. Syst.* **1987**, *33*, 373–383.
- Papadimitriou, C.; Beck, J.; Au, S.K. Entropy-based optimal sensor location for structural model updating. *J. Vib. Control* **2000**, *6*, 781–800. [[CrossRef](#)]
- Papadimitriou, C. Optimal sensor placement methodology for parametric identification of structural systems. *J. Sound Vib.* **2004**, *278*, 923–947. [[CrossRef](#)]
- Yao, L.; Sethares, W.; Kammer, D. Sensor placement for on-orbit modal identification via a genetic algorithm. *AIAA J.* **1993**, *31*, 1922–1928. [[CrossRef](#)]
- Chisari, C.; Macorini, L.; Amadio, C.; Izzuddin, B. Optimal sensor placement for structural parameter identification. *Struct. Multidiscip. Optim.* **2017**, *55*, 647–662. [[CrossRef](#)]
- Kammer, D. Sensor placement for on-orbit modal identification and correlation of large space structures. *J. Guid. Control Dyn.* **1991**, *14*, 251–259. [[CrossRef](#)]
- Eftekhari Azam, S.; Mariani, S. Dual estimation of partially observed nonlinear structural systems: A particle filter approach. *Mech. Res. Commun.* **2012**, *46*, 54–61. [[CrossRef](#)]
- Chatzi, E.N.; Fuggini, C. Online correction of drift in structural identification using artificial white noise observations and an unscented Kalman filter. *Smart Struct. Syst.* **2015**, *16*, 295–328. [[CrossRef](#)]
- Eftekhari Azam, S.; Chatzi, E.; Papadimitriou, C. A dual Kalman filter approach for state estimation via output-only acceleration measurements. *Mech. Syst. Signal Process.* **2015**, *60*, 866–886. [[CrossRef](#)]
- Capellari, G.; Eftekhari Azam, S.; Mariani, S. Towards real-time health monitoring of structural systems via recursive Bayesian filtering and reduced order modelling. *Int. J. Sustain. Mater. Struct. Syst.* **2015**, *2*, 27–51. [[CrossRef](#)]
- Ghanem, R.; Spanos, P. Polynomial chaos in stochastic finite elements. *J. Appl. Mech.* **1990**, *57*, 197–202. [[CrossRef](#)]
- Hansen, N. The CMA evolution strategy: A comparing review. In *Towards a New Evolutionary Computation. Advances on Estimation of Distribution Algorithms*; Springer: Berlin, Germany, 2006; pp. 75–102.
- Katafygiotis, L.; Beck, J. Updating models and their uncertainties. II: Model identifiability. *J. Eng. Mech.* **1998**, *124*, 463–467. [[CrossRef](#)]
- Lindley, D. On a measure of the information provided by an experiment. *Ann. Math. Stat.* **1956**, *27*, 986–1005. [[CrossRef](#)]
- Raiffa, H.; Schlaifer, R. *Applied Statistical Decision Theory*; Wiley-Interscience: Hoboken, NJ, USA, 1961.
- Chaloner, K.; Verdinelli, I. Bayesian experimental design: A review. *Stat. Sci.* **1995**, *10*, 273–304. [[CrossRef](#)]

21. De Groot, M. Uncertainty, information, and sequential experiments. *Ann. Math. Stat.* **1962**, *33*, 404–419. [[CrossRef](#)]
22. Kullback, S.; Leibler, R. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [[CrossRef](#)]
23. Kullback, S. *Statistics and Information Theory*; Courier Corporation: Chelmsford, MA, USA, 1959.
24. Bernardo, J.M. Expected information as expected utility. *Ann. Stat.* **1979**, *7*, 686–690. [[CrossRef](#)]
25. Schervish, M.J. *Theory of Statistics*; Springer: Berlin, Germany, 2012.
26. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
27. Papadimitriou, C.; Lombaert, G. The effect of prediction error correlation on optimal sensor placement in structural dynamics. *Mech. Syst. Signal Process.* **2012**, *28*, 105–127. [[CrossRef](#)]
28. Yuen, K.V. *Bayesian Methods for Structural Dynamics and Civil Engineering*; John Wiley & Sons: Hoboken, NJ, USA, 2010.
29. Capellari, G.; Chatzi, E.; Mariani, S. Parameter identifiability through information theory. In Proceedings of the 2nd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering (UNCECOMP), Rhodes Island, Greece, 15–17 June 2017.
30. Capellari, G.; Chatzi, E.; Mariani, S. Optimal design of sensor networks for damage detection. *Procedia Eng.* **2017**, *199*, 1864–1869. [[CrossRef](#)]
31. Khoshnevisan, M.; Bhattacharya, S.; Smarandache, F. Utility of choice: An information theoretic approach to investment decision-making. *arXiv* **2002**, arXiv:math/0212134. [[CrossRef](#)]
32. Parnell, G.; Driscoll, P.; Henderson, D. *Decision Making in Systems Engineering and Management*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 81.
33. Capellari, G.; Chatzi, E.; Stefano, M. Cost-Benefit Optimization of Sensor Networks for SHM Applications. *Proceedings* **2018**, *2*, 132. [[CrossRef](#)]
34. Ryan, K. Estimating expected information gains for experimental designs with application to the random fatigue-limit model. *J. Comput. Graph. Stat.* **2003**, *12*, 585–603. [[CrossRef](#)]
35. Huan, X.; Marzouk, Y. Simulation-based optimal Bayesian experimental design for nonlinear systems. *J. Comput. Phys.* **2013**, *232*, 288–317. [[CrossRef](#)]
36. Capellari, G.; Chatzi, E.; Mariani, S. An optimal sensor placement method for SHM based on Bayesian experimental design and Polynomial Chaos Expansion. In Proceedings of the VII European Congress on Computational Methods in Applied Sciences and Engineering, Crete, Greece, 5–10 June 2016; Volume 3, pp. 6272–6282.
37. Eftekhari Azam, S.; Mariani, S.; Attari, N. Online damage detection via a synergy of proper orthogonal decomposition and recursive Bayesian filters. *Nonlinear Dyn.* **2017**, *89*, 1489–1511. [[CrossRef](#)]
38. Eftekhari Azam, S.; Mariani, S. Online damage detection in structural systems via dynamic inverse analysis: A recursive Bayesian approach. *Eng. Struct.* **2018**, *159*, 28–45. [[CrossRef](#)]
39. Wiener, N. The homogeneous chaos. *Am. J. Math.* **1938**, *60*, 897–936. [[CrossRef](#)]
40. Xiu, D.; Karniadakis, G. The Wiener—Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **2002**, *24*, 619–644. [[CrossRef](#)]
41. Xiu, D.; Lucor, D.; Su, C.H.; Karniadakis, G. Stochastic modeling of flow-structure interactions using generalized polynomial chaos. *J. Fluids Eng.* **2002**, *124*, 51–59. [[CrossRef](#)]
42. Xiu, D.; Karniadakis, G. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.* **2003**, *187*, 137–167. [[CrossRef](#)]
43. Capellari, G.; Chatzi, E.; Mariani, S. Structural Health Monitoring Sensor Network Optimization through Bayesian Experimental Design. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2018**, *4*, 04018016. [[CrossRef](#)]
44. Le Maître, O.; Knio, O.M. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*; Springer: Berlin, Germany, 2010.
45. Ghanem, R.; Spanos, P. *Stochastic Finite Elements: A Spectral Approach*; Courier Corporation: London, UK, 2003.
46. Herzog, M.; Gilg, A.; Paffrath, M.; Rentrop, P.; Wever, U. Intrusive versus non-intrusive methods for stochastic finite elements. In *From Nano to Space*; Springer: Berlin, Germany, 2008; pp. 161–174.
47. Ghiocel, D.M.; Ghanem, R.G. Stochastic finite-element analysis of seismic soil–structure interaction. *J. Eng. Mech.* **2002**, *128*, 66–77. [[CrossRef](#)]
48. Le Maître, O.; Reagan, M.; Najm, H.; Ghanem, R.; Knio, O. A stochastic projection method for fluid flow: II. Random process. *J. Comput. Phys.* **2002**, *181*, 9–44. [[CrossRef](#)]

49. Berveiller, M.; Sudret, B.; Lemaire, M. Stochastic finite element: A non intrusive approach by regression. *Eur. J. Comput. Mech.* **2006**, *15*, 81–92. [[CrossRef](#)]
50. Blatman, G. Adaptive Sparse Polynomial Chaos Expansions for Uncertainty Propagation and Sensitivity Analysis. Ph.D. Thesis, Blaise Pascal University, Clermont-Ferrand, France, 2009.
51. Blatman, G.; Sudret, B. Adaptive sparse polynomial chaos expansion based on least angle regression. *J. Comput. Phys.* **2011**, *230*, 2345–2367. [[CrossRef](#)]
52. Pearson, K. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [[CrossRef](#)]
53. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417. [[CrossRef](#)]
54. Liang, Y.; Lee, H.; Lim, S.; Lin, W.; Lee, K.; Wu, C. Proper Orthogonal Decomposition and Its applications, Part I: Theory. *J. Sound Vib.* **2002**, *252*, 527–544. [[CrossRef](#)]
55. Wu, C.; Liang, Y.; Lin, W.; Lee, H.; Lim, S. A note on equivalence of proper orthogonal decomposition methods. *J. Sound Vib.* **2003**, *265*, 1103–1110. [[CrossRef](#)]
56. Karhunen, K. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fenn. Math.* **1947**, *37*, 1–79.
57. Loeve, M. Asymptotical Study of Dependent Random Variables. Ph.D. Thesis, Universite de Paris, Paris, France, 1941.
58. Kosambi, D. Statistics in function space. *J. Indian Math. Soc.* **1948**, *7*, 76–88.
59. Mees, A.; Rapp, P.; Jennings, L. Singular-value decomposition and embedding dimension. *Phys. Rev. A* **1987**, *36*, 340–346. [[CrossRef](#)]
60. Eftekhar Azam, S. *Online Damage Detection in Structural Systems*; Springer: Berlin, Germany, 2014.
61. Capellari, G.; Eftekhar Azam, S.; Mariani, S. Hybrid Reduced-Order Modeling and Particle-Kalman Filtering for the Health Monitoring of Flexible Structures. Available online: https://www.researchgate.net/publication/269045554_Hybrid_Reduced-Order_Modeling_and_Particle-Kalman_Filtering_for_the_Health_Monitoring_of_Flexible_Structures (accessed on 11 June 2018).
62. Mirzazadeh, R.; Eftekhar Azam, S.; Jansen, E.; Mariani, S. Uncertainty quantification in polysilicon MEMS through on-chip testing and reduced-order modelling. In Proceedings of the 18th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), Dresden, Germany, 2–5 April 2017; pp. 1–8.
63. Mirzazadeh, R.; Eftekhar Azam, S.; Mariani, S. Mechanical Characterization of Polysilicon MEMS: A Hybrid TCMC/POD-Kriging Approach. *Sensors* **2018**, *18*, 1243. [[CrossRef](#)] [[PubMed](#)]
64. Hansen, N.; Müller, S.D.; Koumoutsakos, P. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.* **2003**, *11*, 1–18. [[CrossRef](#)] [[PubMed](#)]
65. Hansen, N.; Ostermeier, A.; Gawelczyk, A. On the adaptation of arbitrary normal mutation distributions in evolution strategies: The generating set adaptation. In Proceedings of the 6th International Conference on Genetic Algorithms, Pittsburgh, PA, USA, 15–19 July 1995; pp. 57–64.
66. Hansen, N.; Ostermeier, A. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* **2001**, *9*, 159–195. [[CrossRef](#)] [[PubMed](#)]
67. Hansen, N. The CMA evolution strategy: A tutorial. *arXiv* **2016**, arXiv:1604.00772.
68. Eftekhar Azam, S. Dual Estimation and Reduced Order Modelling of Damaging Structures. Ph.D. Thesis, Politecnico di Milano, Milano, Italy, 2012.
69. Capellari, G.; Eftekhar Azam, S.; Mariani, S. Damage detection in flexible plates through reduced-order modeling and hybrid particle-Kalman filtering. *Sensors* **2016**, *16*, 2. [[CrossRef](#)] [[PubMed](#)]
70. Barbella, G. Frequency Domain Analysis of Slender Structural Systems under Turbulent Wind Excitation. Ph.D. Thesis, Politecnico di Milano, Milano, Italy, 2009.
71. Barbella, G.; Perotti, F.; Simoncini, V. Block Krylov subspace methods for the computation of structural response to turbulent wind. *Comput. Methods Appl. Mech. Eng.* **2011**, *200*, 2067–2082. [[CrossRef](#)]
72. Capellari, G.; Chatzi, E.; Mariani, S. Optimal sensor placement through Bayesian experimental design: Effect of measurement noise and number of sensors. *Proceedings* **2017**, *1*, 41. [[CrossRef](#)]
73. Gossen, H. *The Laws of Human Relations and the Rules of Human Action Derived Therefrom*; MIT Press: Cambridge, MA, USA, 1983.
74. Lee, R.W.; Kulesz, J.J. A risk-based sensor placement methodology. *J. Hazard. Mater.* **2008**, *158*, 417–429.

[CrossRef] [PubMed]

75. Marbukh, V.; Sayrafian-Pour, K. Mobile sensor networks self-organization for system utility maximization: Work in progress. In Proceedings of the Fifth International Conference on Wireless and Mobile Communications (ICWMC '09), Nice, France, 23–27 July 2009; pp. 416–419.
76. Tan, L.; Zhang, Y. Optimal resource allocation with principle of equality and diminishing marginal utility in wireless networks. *Wirel. Pers. Commun.* **2015**, *84*, 671–693. [CrossRef]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).