RESEARCH ARTICLE

Bayesian optimal sensor placement for modal identification of civil infrastructures

Costas Argyris[1], Costas Papadimitriou[*1] and Panagiotis Panetsos[2]

1. Department of Mechanical Engineering, University of Thessaly, Volos 38221, Greece
2. Capital Maintenance Department, Egnatia Odos S.A., Thermi 57001, Greece

Abstract: A Bayesian optimal experimental design (OED) method is proposed in this work for estimating the best locations of sensors in structures so that the measured data are most informative for estimating reliably the structural modes. The information contained in the data is measured by the Kullback-Leibler (K-L) divergence between the prior and posterior distribution of the model parameters taken in modal identification to be the modal coordinates. The optimal sensor placement that maximizes the expected K-L divergence is shown also to minimize the information entropy of the posterior distribution. Unidentifiability issues observed in existing formulations when the number of sensors is less than the number of identified modes, are resolved using a non-uniform prior in the Bayesian OED. An insightful analysis is presented that demonstrates the effect of the variances of Bayesian priors on the optimal design. For dense mesh finite element models, sensor clustering phenomena are avoided by integrating in the methodology spatially correlated prediction error models. A heuristic forward sequential sensor placement algorithm and a stochastic optimization algorithm are used to solve the optimization problem in the continuous physical domain of variation of the sensor locations. The theoretical developments and algorithms are applied for the optimal sensor placement design along the deck of a 537 m concrete bridge.

Keywords: Bayesian inference, Kullback-Leibler divergence, information entropy, structural dynamics

*Correspondence to: Costas Papadimitriou, Department of Mechanical Engineering, University of Thessaly, Volos 38221, Greece; Email: costasp@uth.gr

Received: October 10, 2016; **Accepted:** October 28, 2016; **Published Online:** December 30, 2016

Citation: Argyris C, Papadimitriou C, and Panetsos P, 2016, Bayesian optimal sensor placement for modal identification of civil infrastructures. *Journal of Smart Cities*, vol.2(2): 69–86. http://dx.doi.org/10.26789/JSC.2016.02.001.

1. Introduction

Experimental measurements from civil infrastructures, including buildings, bridges, offshore structures, wind turbines, and industrial facilities are often used to obtain the modes of these structures. The modes are useful in structural performance evaluation[1], finite element model updating[2-4], and model-based structural health monitoring[5,6]. A number of methodologies have been developed in the past to optimize the location of sensors in order to maximize the information contained in the measurements for identifying the structural modes. Among them, methods based on information theory are used to make rational decisions consistent with the information provided by the measurements. Non-information based methods have also been developed. A review of a number of non-information based methods can be found in a thesis by Li[7].

This work concentrates on optimal sensor placement design methods for modal identification based on information theory. In the past, notable contributions to the sensor placement problem for modal identification have been provided by the effective indepen-

dence (EFI) method[8] for uniaxial and triaxial sensors[9,10]. Information theory measures, based on scalar measures of the Fisher information matrix (FIM)[11,12] and on information entropy[13–15], proposed in the past for structural parameter estimation problems, have been extended to be used for modal identification[16] as well. An optimal sensor placement design for modal identification based on FIM was proposed by Kammer[17]. The information entropy measure for parameter estimation introduced by Papadimitriou *et al.*[13], was extended by Papadimitriou [18] to obtain useful expressions of the information entropy as a function of the number of sensors, and was applied to modal identification problems[16]. In particular, the information entropy measures the uncertainty in the posterior distribution of the model parameters to be identified. The posterior distribution is obtained from a Bayesian analysis that makes use of the prior distribution of the model parameters. Up to now, the effect of the prior distribution in the optimal design has not been adequately explored.

When applying the EFI and information entropy techniques to dense finite element models, the problem of sensor clustering is manifested. The source of the problem is the failure to take into account the redundant information provided from neighborhood sensor locations. The problem was adequately resolved using spatially correlated prediction error models in the Bayesian formulation to exclude redundant information from neighboring sensors[16]. Stephan[19] has also effectively tackled the issue of information redundancy between sensors by introducing a measure of information redundancy.

Optimization algorithms have also been proposed to find the optimal sensor locations. For structural parameter estimation problems multiple local/global optima make the solution of the optimization problem very challenging[20]. Heuristic algorithms such as the backward and forward sequential sensor placement (BSSP and FSSP) algorithms have been proposed to drastically reduce the computational effort[17,18]. The optimal sensor placement strategies based on the information entropy and FSSP was applied to bridge, towers, and timber structures for optimizing the location of uniaxial and triaxial sensors[21-23]. These algorithms are shown to be quite accurate. However, the use of the FSSP algorithm requires that identifiability issues are resolved for small number of sensors. Unidentifiability issues arise from the aforementioned methods for modal identification when the number of sensors placed in the structure is less than the number of identified modes. This is due to the fact that the FIM becomes singular. Yuen and Kuok[24] noted that introducing non-uniform priors in the Bayesian posterior used in the information entropy measures resolves the problem. For uniform priors, Papadimitriou and Lombaert[16] obtained reasonable sensor placement design by excluding the zero eigenvalues from the product of the eigenvalues of the FIM used for computing the determinant of the FIM that is required in information entropy formulations.

In this work we revisit the problem of sensor placement for modal identification. We formulate the optimal experimental design based on expected utility functions[25] and apply it to the case of optimizing the location of sensors for a real bridge. We use as utility function the relative entropy or, equivalently, the Kullback-Leibler divergence between the prior and the posterior distribution of model parameters[26]. For uniform or Gaussian priors and for models for which the output quantities of interest (QoI) depend linearly on the model parameters to be identified, as it is the case of modal identification, we demonstrate that the expected utility function is directly related to the minus of the information entropy of the parameters to be estimated plus a constant that does not depend on the sensor locations. Thus, the optimal design results in maximizing the expected log determinant of the sum of the FIM and the prior Hessian. As a result, with the aid of the information contained in the prior PDF of the model parameters, the combined matrix (FIM and prior Hessian) is non-singular for non-uniform priors and the optimal sensor placement problem can be carried out also for the case where the number of sensors is less than the number of modes.

A novelty in this work is to study the effect of Gaussian prior uncertainties on the optimal sensor design. For this, an insightful analytical expression is developed that shows the effect the prior uncertainties have on the optimal design. The importance of spatial correlation in the prediction error is also pointed out as the means of avoiding sensor clustering phenomena for finite element models used to simulate civil infrastructures.

Theoretical developments are demonstrated by designing the optimal locations of a number of sensors for a 537 m long concrete bridge using a dense finite element mesh of approximately 830,000 degrees of freedom (DOF). The optimization is formulated in the physical continuous space of the design variables.

This avoids the large discrete design space that arises from the extremely large number of possible nodal positions due to dense FE meshes. Multiple local optima are revealed that make the optimization problem very challenging. It is demonstrated that the computationally efficient heuristic FSSP algorithm provides accurate solutions when compared to stochastic algorithms such as the covariance matrix adaptation (CMA-ES)[27] able to estimate the global optimum with substantially higher computational cost. Useful results are obtained which help guide experimentalists as to how the proposed method can be used for designing the optimal sensor locations in practical applications. In particular, we demonstrate that designs are also obtained for the very important case of number of sensors which are less than the number of modes by using the information contained in the prior distribution. In particular, this is important in designing the location of the reference sensors in a multiple sensor configuration set up experiment conducted with limited number of reference and roving sensors in order to obtain the modal frequencies and reliably assemble the mode shapes from multiple setups. The effectiveness of the methodology is illustrated by designing the optimal location of two reference sensors (one transverse and one vertical) for the bridge. Finally, we draw attention to the sensor clustering issue and propose methods to avoid it and finally we show the effect of uncertainty in the prior distribution on the sensor placement designs.

2. Bayesian Parameter Estimation

The Bayesian framework for the estimation of the parameters of finite element models of structures based on experimental data is first outlined and the results are used in the optimal experimental design formulation presented in Section 3. Consider a model of a structural system and let $\underline{\theta} \in R^{N_\theta}$ be the vector of model parameters to be estimated using a set of measured data $\underline{y} \equiv \underline{y}(\underline{\delta}) \in R^{N_0}$ of output quantities that depend on experimental design variables $\underline{\delta}$, where N_0 is the number of sensors. Let $\underline{g}(\underline{\theta};\underline{\delta}) = A(\underline{\delta})\underline{\theta} \in R^{N_0}$ be the vector of the values of the output quantities predicted by a structural model for specific values of the parameter set $\underline{\theta}$, where it is assumed that there is a linear relationship between the output quantity of interest (QoI) and the parameter set $\underline{\theta}$ to be identified from the experiments, and $A(\underline{\delta}) \in R^{N_0 \times N_\theta}$ depends on the structure of the model and the experimental design variables $\underline{\delta}$. The design variables are related to the location of sensors placed in a structure. The location vector $\underline{\delta}$ contains the coordinates of the sensors with respect to a coordinate system.

The following prediction error equation is introduced to model the discrepancy between the measurements and the model predictions

$$\underline{y} = \underline{g}(\underline{\theta};\underline{\delta}) + \underline{e} = A(\underline{\delta})\underline{\theta} + \underline{e} \quad (1)$$

where $\underline{e}$ is the additive prediction error term due to model and measurement error. The prediction error $\underline{e}$ is usually modeled as a Gaussian vector, whose mean value is equal to zero and its covariance is equal to $\Sigma(\underline{\delta};\underline{\sigma}) \in R^{N_0 \times N_0}$, where $\underline{\sigma}$ contains the parameters that define the correlation structure $\Sigma(\underline{\delta};\underline{\sigma})$ of the prediction error. Applying the Bayesian theorem[2], the posterior probability density function (PDF) of the model parameter set $\underline{\theta}$, given the measured data $\underline{y}$, takes the form

$$p(\underline{\theta} \mid \underline{y},\underline{\sigma},\underline{\delta}) = c\frac{1}{(2\pi)^{N_0/2}\left[\det \Sigma(\underline{\delta};\underline{\sigma})\right]^{1/2}} \exp\left[-\frac{1}{2}J(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta})\right] \pi(\underline{\theta}) \quad (2)$$

where

$$J(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta}) = [\underline{y} - A(\underline{\delta})\underline{\theta}]^T \Sigma^{-1}(\underline{\delta};\underline{\sigma})[\underline{y} - A(\underline{\delta})\underline{\theta}] \quad (3)$$

quantifies the discrepancy between the measured and model predicted quantities, $\pi(\underline{\theta})$ is the prior distribution for $\underline{\theta}$, and c is a normalization constant guaranteeing that the posterior PDF integrates to one.

Due to the linear relationship between the output QoI and the model parameters, the function $J(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta})$ is quadratic in $\underline{\theta}$. Assuming a uniform prior with wide enough bounds or a Gaussian prior, the posterior PDF for the model parameters $\underline{\theta}$ is Gaussian, denoted by $N(\underline{\theta};\underline{\hat{\theta}},C)$, where $\underline{\hat{\theta}} \equiv \underline{\theta}(\underline{y};\underline{\sigma},\underline{\delta})$ is the most probable value obtained by minimizing the function $-\ln p(\underline{\theta} \mid \underline{y},\underline{\sigma},\underline{\delta})$, *i.e.*,

$$\underline{\hat{\theta}}(\underline{y};\underline{\sigma},\underline{\delta}) = \arg\min_{\underline{\theta}}[J(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta}) - \ln \pi(\underline{\theta})] \quad (4)$$

and the covariance matrix $C = C(\underline{\hat{\theta}};\underline{y},\underline{\sigma},\underline{\delta})$ equals to the inverse of the Hessian of $-\ln p(\underline{\theta} \mid \underline{y},\underline{\sigma},\underline{\delta})$, *i.e.*,

$$C^{-1}(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta}) = \underline{\nabla}_{\theta}\underline{\nabla}_{\theta}^{T}[-\ln p(\underline{\theta}\,|\,\underline{y},\underline{\sigma},\underline{\delta})] = Q_L(\underline{\delta},\underline{\sigma}) + Q_{\pi} \tag{5}$$

evaluated at the most probable value $\underline{\hat{\theta}}(\underline{y};\underline{\sigma},\underline{\delta})$, where $Q_L(\underline{\delta},\underline{\sigma})$ is the FIM obtained from

$$Q_L(\underline{\delta},\underline{\sigma}) = \frac{1}{2}\underline{\nabla}_{\theta}\underline{\nabla}_{\theta}^{T} J(\underline{\theta};\underline{y},\underline{\sigma},\underline{\delta}) = A^{T}(\underline{\delta})\Sigma^{-1}(\underline{\delta};\underline{\sigma})A(\underline{\delta}) \tag{6}$$

to be independent of $\underline{\theta}$ and the experimental values $\underline{y}$, and $Q_{\pi}(\underline{\theta}) = \underline{\nabla}_{\theta}\underline{\nabla}_{\theta}^{T}[-\ln \pi(\underline{\theta})]$ is the zero matrix $Q_{\pi} = 0$ for uniform prior and the constant matrix $Q_{\pi} = S^{-1}$ for a Gaussian prior PDF with covariance matrix S. The covariance matrix C of the Gaussian posterior PDF does not depend on the values of the parameters $\underline{\theta}$ and the experimental data $\underline{y}$. The dependence on the experimental design variables $\underline{\delta}$ comes from the information matrix $Q_L(\underline{\delta},\underline{\sigma})$.

3. Bayesian Optimal Experimental Design

3.1 Expected Utility Function

The expected utility function has been introduced by Lindley[25] to measure the information contained in the experimental data for estimating the parameters of the model. The expected utility function has the form

$$U(\underline{\delta}) = \int_{\Sigma}\int_{\Upsilon}\int_{\Theta} u(\underline{\delta};\underline{\theta},\underline{y},\underline{\sigma})\; p(\underline{\theta},\underline{y},\underline{\sigma}\,|\,\underline{\delta})\; d\underline{\theta}\; d\underline{y}\; d\underline{\sigma} \tag{7}$$

where $u(\underline{\delta};\underline{\theta},\underline{y},\underline{\sigma})$ is the utility function given a particular value of the model parameter set $\underline{\theta}$, the outcome $\underline{y}$ from the experiment, and the parameter set $\underline{\sigma}$, $p(\underline{\theta},\underline{y},\underline{\sigma}\,|\,\underline{\delta}) = p(\underline{\theta}\,|\,\underline{y},\underline{\sigma},\underline{\delta})\; p(\underline{y}\,|\,\underline{\sigma},\underline{\delta})\; p(\underline{\sigma}\,|\,\underline{\delta})$, $p(\underline{\theta}\,|\,\underline{y},\underline{\sigma},\underline{\delta})$ is the posterior uncertainty in the model parameters given the outcome $\underline{y}$, $p(\underline{y}\,|\,\underline{\sigma},\underline{\delta})$ is the uncertainty in the data, and $p(\underline{\sigma}\,|\,\underline{\delta}) = \pi(\underline{\sigma})$ quantifies the prior uncertainty in the parameter set $\underline{\sigma}$. Herein, the expected utility function in (7) has been extended to include the uncertainties in the model prediction error parameters $\underline{\sigma}$ by taking the expectation over the parameter space $\underline{\sigma}$ as well.

Based on information theory, the utility function can be chosen to be the relative entropy or the Kullback-Leibler divergence[25,26]

$$u(\underline{\delta};\underline{\theta},\underline{y},\underline{\sigma}) = p(\underline{\theta}\,|\,\underline{y},\underline{\sigma},\underline{\delta})\; \ln \frac{p(\underline{\theta}\,|\,\underline{y},\underline{\sigma},\underline{\delta})}{\pi(\underline{\theta})} \tag{8}$$

between the prior and posterior PDF of the model parameters $\underline{\theta}$ given an outcome $\underline{y}$ obtained from an experimental design $\underline{\delta}$. Substituting (8) into (7), one has that the expected utility function is an average of the K-L divergence over all possible values of the model parameters $\underline{\theta}$ as they are inferred from the data, all the possible outcomes $\underline{y}$ of the experiment, and all possible values of the model prediction error parameters $\underline{\sigma}$. For Gaussian posterior PDF in $\underline{\theta}$ that does not depend on the data $\underline{y}$, the inner double integral in the expected utility function simplifies considerably[28] and the expected utility function takes the form

$$U(\underline{\delta}) = -\int H(\underline{\delta},\underline{\sigma})\; \pi(\underline{\sigma})\; d\underline{\sigma} - c \tag{9}$$

where

$$H(\underline{\delta},\underline{\sigma}) = \frac{1}{2}N_{\theta}[\ln(2\pi)+1] - \frac{1}{2}\ln\det[Q_L(\underline{\delta};\underline{\sigma}) + S^{-1}] \tag{10}$$

is the information entropy of the model parameters $\underline{\theta}$ given $\underline{\sigma}$, while c is a quantity that does not depend on the design variables $\underline{\delta}$ so that it can be treated as constant. This integral in the right-hand-side of (9) represents the robust measure of the information entropy over all possible values of the prediction error parameters quantified by the prior PDF $\pi(\underline{\sigma})$. For a small number (one or two) of parameters in $\underline{\sigma}$, the integral can be carried out using a numerical integration algorithm. For the sake of simplicity, in this work the values of $\underline{\sigma}$ are assumed known (deterministic). The utility function then takes the form

$$U(\underline{\delta}) = -H(\underline{\delta},\underline{\sigma}) - c \tag{11}$$

It can be seen that the expected relative information or expected K-L divergence has a direct connection to the robust information entropy proposed by Papadimitriou *et al.*[13] and extended by Yuen and Kuok[24] for non-uniform distributions.

3.2 Optimal Designs

The optimal experimental design problem is formulated as finding the values $\underline{\delta}_{opt}$ of the design variables

$\underline{\delta}$ that maximize the utility function $U(\underline{\delta})$ or, equivalently, minimize the information entropy $H(\underline{\delta},\underline{\sigma})$, *i.e.*,

$$\underline{\delta}_{opt} = \arg\max_{\underline{\delta}} U(\underline{\delta}) = \arg\min_{\underline{\delta}} H(\underline{\delta},\underline{\sigma}) \qquad (12)$$

Optimal experimental design problems involving the design of the position of sensors often result in multiple local and global solutions. This will be also evident in the results presented in the application section. Also the gradient of the objective function with respect to the design variables in most cases of practical interest cannot be evaluated analytically. To avoid premature convergence to a local optimum and the evaluation of sensitivities of the utility function with respect to the design variables, stochastic optimization algorithms can be used to find the optimum. Herein the CMA-ES algorithm[27] is used for solving the optimization problem, requiring only evaluation of the objective function at different values of the design variables. For this, the problem is formulated as a continuous optimization problem where the design variables are related to the coordinates of the sensors along the physical domain of the structure. To account, however, of curved and disconnected one-dimensional or two-dimensional domains, as well as to take into account the different types of sensors (vertical or transverse) that can be placed along the curved domain, a mapping technique can be used to map the physical design space into a regular one-dimensional parent domain. The optimization is then conveniently carried out in the parent domain. The method will be illustrated for one-dimensional curved domains in the application section.

Heuristic algorithms have also been proposed to provide sub-optimal solutions. Notable are the forward and backward sequential sensor placement algorithms (FSSP and BSSP) in which the optimization is carried out sequentially. In the FSSP algorithm[18] the optimization is carried out sequentially for a single sensor, say the $i+1$ sensor, given that i sensors have already been placed in their optimal locations. This procedure is repeated for $i=1,\cdots,N_0$, where N_0 is the total number of sensors to be placed in the structure. The heuristic algorithms have been shown to be effective and provided for several problems near optimal solutions[16]. The FSSP algorithm is used here for investigating its effectiveness for the optimal sensor placement problem for modal identification.

3.3 Properties of the Expected Relative Entropy or K-L Divergence

The robust information entropy involved in the expected utility function in (9) has exactly the same structure as the information entropy derived asymptotically by Papadimitriou[18] for nonlinear models. Based on the structure of the information entropy measure in (10) it was shown analytically that the information entropy follows certain properties[16,18]. Using the relationship (9) or (11) between the expected K-L divergence and the information entropy, one can easily extend the validity of these properties to the expected utility function. Specifically, following Papadimitriou's paper[18], one can confirm that the expected utility function for a given number of sensors placed in the structure increases when one or more additional sensors are placed in the structure. As a result, the maximum value of the expected utility function attained for a certain number of sensors is a non-decreasing function of the number of sensors. Thus the selection of the optimal number of sensors is based only on the level of the information that one can afford loosing when stops placing additional sensors in the structure. In addition, spatially correlated prediction error models avoid the problem of the redundant information that is usually provided by neighborhood sensors[16]. This is true when the distance of the sensors is less than a characteristic length of the contributing mode. In this case the spatial correlation length in spatially correlated prediction error models should be selected based on the characteristic length of the highest contributing mode. Spatially uncorrelated prediction error models fail to properly consider the redundant information and should be avoided since otherwise they result in sensor clustering.

4. Optimal Sensor Placement for Modal Identification

4.1 Formulation for Modal Identification

The optimal experimental design methodology is next implemented in structural dynamics for optimally placing the sensors in the structure for modal identification. Considering a linear finite element model of a structure, the equations of motion are given by

$$M\ddot{\underline{u}} + \tilde{C}\dot{\underline{u}} + K\underline{u} = \underline{f} \qquad (13)$$

where M, $\tilde{C}$ and $M \in \mathrm{R}^{n\times n}$ are the mass, damping and stiffness matrices, respectively, $\underline{u} \in R^n$ is the di-

splacement vector and $\underline{f}$ is the forcing vector. Using modal analysis and assuming classically damped modes, the response displacement and acceleration vectors are given by $\underline{u} = \Phi\underline{\xi}$ and $\ddot{\underline{u}} = \Phi\ddot{\underline{\xi}}$, respectively, where $\Phi = [\underline{\phi}_1, \cdots, \underline{\phi}_m] \in \mathrm{R}^{n\times m}$ is the matrix of mode shapes involving m contributing modes ($m \leqslant n$) that can be obtained by solving the eigenvalue problem $K\Phi = \Lambda M\Phi$, Λ is the diagonal matrix of eigenvalues, $\underline{\xi} = [\xi_1, \cdots, \xi_m]^T \in R^m$ is the vector of modal coordinates satisfying

$$\ddot{\xi}_r + 2\zeta_r\omega_r\dot{\xi}_r + \omega_r^2\xi_r = \underline{\phi}_r^T\,\underline{f} \tag{14}$$

$r = 1, \cdots, m$, and ζ_r is the modal damping ratio. The strain vector is given by a similar expression $\underline{\varepsilon} = E\underline{\xi}$, where E depends through the finite element modeling on the elements in the mode shape matrix Φ.

The problem of estimating the modal coordinate vector $\underline{\xi}$ or $\ddot{\underline{\xi}}$ using displacement/strain or acceleration measurements is investigated in this work. The modal coordinates $\underline{\xi}$ or $\ddot{\underline{\xi}}$ contain the modal properties (modal frequencies, modal damping ratios, participation factors). The objective in modal identification is to place sensors (displacement, acceleration and/or strain sensors) so that the information contained in the measured data is sufficient to estimate the modal coordinate vectors $\underline{\xi}$ or $\ddot{\underline{\xi}}$, depending on the sensor type used. Introducing the parameter set $\underline{\theta}$ to be either $\underline{\xi}$ or $\ddot{\underline{\xi}}$ and denoting by $\underline{g}(\underline{\theta};\underline{\delta}) \in R^{N_0}$ the response quantity that is measured by the sensors, one has the following equation between the modal model predictions and the parameter set $\underline{\theta}$

$$\underline{g}(\underline{\theta};\underline{\delta}) = \Phi(\underline{\delta})\underline{\theta} = L(\underline{\delta})\Phi\,\underline{\theta} \tag{15}$$

where the matrix $L(\underline{\delta}) \in R^{N_0\times n}$ is the observation matrix and maps the n model DOF to the N_0 measured positions. The matrix $L(\underline{\delta})$ depends on the location vector $\underline{\delta}$ defining the locations of the sensors in the structure. If the measured positions coincide with the DOF of the model then the matrix $L(\underline{\delta})$ is comprised of zeros and ones. For the general case for which the measured locations do not coincide with the DOF of the FE model, the matrix $L(\underline{\delta})$ depends on the interpolation scheme used to obtain the response within a finite element in terms of the finite element nodal responses. The above formulation allows sensors placed in any point in the structure, not only at nodal points. Also it gives the flexibility to covert the optimization problem for estimating the design variables to a continuous optimization problem over the physical domain of the structure. For strain measurements the aforementioned analysis is the same provided that the mode shape matrix Φ is replaced by the matrix E.

The model equation in (15) is the same as the one used in (1) so that the expected utility function (9) or (11) applies with the FIM given by

$$Q_L(\underline{\delta};\underline{\sigma}) = [L(\underline{\delta})\Phi]^T\,\Sigma^{-1}(\underline{\delta};\underline{\sigma})[L(\underline{\delta})\Phi] \tag{16}$$

in terms of the mode shape components at the measured locations.

Based on the form (16), and using the dimensions of the matrices $[L(\underline{\delta})\Phi]^T \in R^{m\times N_0}$, $\Sigma^{-1}(\underline{\delta};\underline{\sigma}) \in R^{N_0\times N_0}$ and $L(\underline{\delta})\Phi \in R^{N_0\times m}$ in (16), a non-singular FIM matrix $Q_L(\underline{\delta};\underline{\sigma})$ is obtained only if the number of sensors N_0 is at least equal to the number of contributing modes m ($N_0 \geqslant m$). For $N_0 < m$, the matrix $Q_L(\underline{\delta};\underline{\sigma})$ in (15) is by construction singular and for uniform prior PDF the determinant of the combined matrix $Q_L(\underline{\delta};\underline{\sigma}) + S^{-1}$ in (10) will be zero for any sensor configuration. Thus, for $N_0 < m$ the optimal sensor location problem cannot be performed for uniform prior PDF. This means that the information content in the measured data and the prior is not sufficient to estimate all the parameters simultaneously.

The Bayesian optimal experimental design formulation yields a nonsingular matrix for $N_0 < m$ only if the prior is non-uniform distribution. The non-uniform prior, say Gaussian, yields a Hessian matrix S^{-1} that is added to the Fisher information matrix $Q_L(\underline{\delta};\underline{\sigma})$ and makes the combined matrix $Q_L(\underline{\delta};\underline{\sigma}) + S^{-1}$ non-singular. The information matrix $Q_L(\underline{\delta};\underline{\sigma})$ in (16) has exactly the same form as the one proposed by Yuen[4] for designing the optimal sensor locations using the EFI algorithm. The difference of the present Bayesian formulation to the EFI algorithm is in the use of the prior information for the model parameters which permits the design of optimal sensor locations for the case of number of sensors which is less than the number of modes. The contribution from the prior is the result of the application of the Bayesian optimal

experimental design proposed herein based on optimizing the expected K-L divergence. Alternatively, Yuen and Kuok[24] have also proposed a non-uniform prior on the information entropy measure in order to solve this unidentifiability problem. For uniform prior and unidentifiable case, Papadimitriou and Lombaert[16] proposed the sum of the log of the non-zero eigenvalues in the FIM to be maximized instead of the sum of the log of all eigenvalues. Herein these results are generalized to incorporate prior uncertainty in the parameter estimates. This procedure allows to systematically place the sensors optimally in the structure even for the unidentifiable case that arises for a small number of sensors.

4.2 Effect of Prior Uncertainty

Next a diagonal covariance $S = diag(s_1^2, \cdots, s_m^2)$ for the Gaussian prior is assumed and the effect of the values of the variances is examined. These variances control the prior uncertainty in the values of the model parameters. A theoretical result is provided that shows the effect of the assigned prior uncertainties on the optimal design.

For simplicity, the case of optimally placing one sensor is considered first and then generalized for the multiple sensor case. In the case of a single sensor, the FIM reduces to the form $Q_L(\underline{\delta};\underline{\sigma}) = \sigma^{-2}\underline{\varphi}_i\underline{\varphi}_i^T$, where $\underline{\varphi}_i = [\Phi_{i1}, \Phi_{i2}, \cdots, \Phi_{im}]^T$ is a vector of dimension m that consists of the values of each mode shape at the sensor location denoted here as i. Note that for one sensor the prediction error covariance $\Sigma(\underline{\delta};\underline{\sigma})$ is scalar with $\Sigma(\underline{\delta};\underline{\sigma}) = \sigma^2$. Using the following known result for a square matrix B and two vectors $\underline{u}$ and $\underline{v}$

$$det(\underline{u}\underline{v}^T + B) = (1 + \underline{v}^T B^{-1}\underline{u})\, det(\mathrm{B}) \tag{17}$$

the $\det Q \equiv \det[Q_L(\underline{\delta};\underline{\sigma}) + S^{-1}]$ in (9) takes the form

$$\begin{aligned}\det[Q] &= det[\sigma^{-2}\underline{\varphi}_i\underline{\varphi}_i^T + S^{-1}] = \\ &[1 + \sigma^{-2}\underline{\varphi}_i^T S\underline{\varphi}_i]\, det(S^{-1}) = \\ &[1 + \sigma^{-2}\sum_{k=1}^{m} s_k^2\Phi_{ik}^2]\, det(S^{-1})\end{aligned} \tag{18}$$

Note that the variance s_k^2 of the k-th parameter (modal coordinate) weights the contribution in the sum of the value of the k-th mode shape at the sensor location. The higher the value of the variance s_k^2, the higher the contribution of the k-th mode shape on the $\det[Q]$. So it is evident that the optimal design will give preference to the modes that have higher prior uncertainty, *i.e.*, higher s_k^2 values.

Next we complete the proof for the general case of arbitrary number of N_0 sensors. For this, we use the following known result

$$\begin{aligned}&det(UWV^T + B) = \\ &\quad det(W^{-1} + V^T B^{-1}U)\, det(W)\, det(B)\end{aligned} \tag{19}$$

Setting for simplicity $\Psi = L(\underline{\delta})\Phi$ and using (16) and (10), the $\det Q = \det[Q_L(\underline{\delta};\underline{\sigma}) + S^{-1}]$ takes the form

$$\begin{aligned}\det[Q] &= \det[\Psi^T\Sigma^{-1}(\underline{\delta};\underline{\sigma})\Psi + S^{-1}] \\ &= \det[\Sigma(\underline{\delta};\underline{\sigma}) + \Psi S\Psi^T]\det[S^{-1}]\det[\Sigma^{-1}(\underline{\delta};\underline{\sigma})] \\ &= \det\left[\Sigma(\underline{\delta};\underline{\sigma}) + \sum_{k=1}^{m} s_k^2\psi_k(\underline{\delta})\psi_k^T(\underline{\delta})\right] \\ &\quad \det[S^{-1}]\det[\Sigma^{-1}(\underline{\delta};\underline{\sigma})] \\ &= \det\left[\Sigma(\underline{\delta};\underline{\sigma}) + \sum_{k=1}^{m} s_k^2\,\|\psi_k(\underline{\delta})\|^2\,\tilde{\psi}_k(\underline{\delta})\tilde{\psi}_k^T(\underline{\delta})\right] \\ &\quad \det[S^{-1}]\det[\Sigma^{-1}(\underline{\delta};\underline{\sigma})]\end{aligned} \tag{20}$$

where $\psi_k(\underline{\delta})$ is the k-th mode shape evaluated at the sensor locations, $\tilde{\psi}_k(\underline{\delta})$ is the unit-normalized mode shape, $\|\psi_k(\underline{\delta})\|^2$ is the Euclidean norm square of the mode shape.

Equation (18) is a special case of equation (20) when only one sensor is used and Ψ becomes a row vector and its columns become scalars. In the multiple sensor case we are dealing with a sum of rank-one matrices over the model parameters, where each rank-one matrix $\psi_k(\underline{\delta})\psi_k^T(\underline{\delta})$ is formed from the k-th mode shape $\psi_k(\underline{\delta})$, weighted by the correspondding prior variance s_k^2 of that mode. So we see again that, similarly to the one sensor case, the Gaussian prior variance s_k^2 acts as weighting factor, this time by giving greater weight to the matrix $\psi_k(\underline{\delta})\psi_k^T(\underline{\delta})$ related to the k-th mode. The variance s_k^2 of the k-th mode weights the contribution in the sum of the value of the k-th rank-one matrix formed from the k-th mode shape that is evaluated at the sensor locations. The higher the value of the variance s_k^2, the higher the contribution of the rank-one matrix of the k-th mode shape on the $\det[Q]$. So the optimal de-

sign of sensor locations again gives preference to the modes that are assigned by a user to have higher prior uncertainty.

It should be noted that the idea of favoring one or more modes in the design of the optimal sensor configuration by appropriately selecting the prior variances can be useful in model updating and damage detection applications. Specifically, in damage detection and localization the proposed optimal sensor placement design can be used to increase the information about the damage location and size contained in the measured data by favoring, for example, local modes known to be more sensitive to local damage. To avoid reducing the robustness of the proposed optimal sensor placement algorithm for identifying the least favored modes, one may choose to optimally allocate a fraction of the available sensors in an effort to favor a small number of the contributing modes, ensuring that the optimal placement of the rest of the sensors maintain adequate levels of robustness.

5. Application

The proposed methodology is used to optimize the location of acceleration sensors placed at the deck level of the Metsovo Bridge. Numerical results are shown for up to 30 sensors for the spatially uncorrelated and spatially correlated prediction error models. The effect of the spatial correlation as well as the effect of the selection of the uncertainty in the prior distribution on the optimal designs is also investigated. Finally, the optimal design of a small number of reference sensors to be used as common sensors in multiple configuration setups is illustrated.

5.1 Bridge Description and FE Model

The Metsovo bridge of Egnatia Motorway, schematically shown in Figure 1, is crossing the deep ravine of Metsovitikos river, 150 m over the riverbed. This is the highest bridge of the Egnatia Motorway, with the height of the tallest pier equal to 110 m. The total length of the bridge is 537 m. The bridge has 4 spans of length 44.78 m, 117.87 m, 235 m, 140 m and 3 piers of which M1 (45 m) supports the boxbeam superstructure through pot bearings (movable in both horizontal directions), while M2 (110 m) and M3 (35 m) piers connect monolithically to the structure.

A detailed FE model of the bridge is created using 3-dimensional tetrahedron quadratic Lagrange finite elements. The mesh is chosen to accurately predict the lowest twenty modal frequencies and mode shapes of the bridge. The model also considers the interaction with soil by modeling the soil stiffness with solid blocks surrounding the piers and abutments. The FE model has 830,000 DOF. The size of the elements in the mesh is controlled by the thickness of the deck and piers box-like cross-sections. The typical element length is of the order of the thickness of the deck cross-section. The complex 3D geometry of the bridge was designed in SolidWorks and the FE model was created and solved in COMSOL Multiphysics by importing the 3D geometry from SolidWorks. The modulus of elasticity for the concrete was taken from design considerations to be 37×10^9 Pa and 34×10^9 Pa for the deck and the piers, respectively. The modulus of elasticity for the solid blocks was taken to be 9×10^9 Pa. The stiffness and mass matrices of the finite element model are used to obtain the mass normalized mode shapes to be used in the optimal sensor place ment methodology.

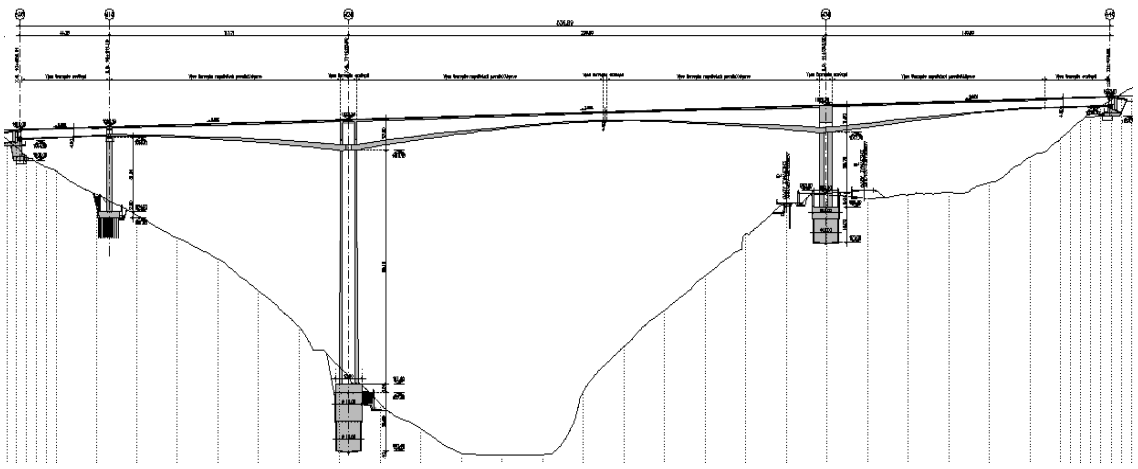


Figure 1. Longitudinal view of the Metsovo Bridge

5.2 Optimal Sensor Placement

The purpose of the OED is to optimally locate a number of acceleration sensors such that one gets the most informative data for identifying the lowest 10 modes. Using the FE model, the lowest 10 mode shapes needed in the design of the optimal sensor locations were computed and stored. These modes consist of 5 transverse modes, with mode shapes that deform the bridge and deck in the transverse direction, and 5 vertical modes mode, with mode shapes that bend the deck in the vertical direction. The transverse mode numbers are $\{1,2,4,6,8\}$, while the bending mode number are $\{3,5,7,9,10\}$. The optimal locations of sensors are obtained by maximizing the expected utility function or minimizing the information entropy. Herein, results will be presented in terms of the information entropy measure. It should be noted that in our case the number of parameters (modal coordinates) is $N_\theta = m = 10$.

5.2.1 Continuous Optimization in Parent Domain

Due to the fact that measurements on the bridge deck have to be taken without traffic interruptions, the sensors in the structure are only allowed to be placed along the pedestrian sidewalks (left and right). For demonstration purposes, in this study it will be assumed that sensors are placed along the one sidewalk marked in Figure 2 with red line. The line along the sidewalk over which sensors can be placed is a curved one and in order to perform the optimization problem along the curved line we develop a mapping of the curved line in a much simpler parent domain of a straight line. Each point in the physical curved line is mapped to a point in the straight parent line. To introduce such a mapping one can follow concepts developed in finite element analysis to map an arbitrary one-dimensional element in space in a parent element of specified length. It should be noted that the curved one-dimensional domains (red line) in Figure 2 consists of the straight elements that are the vertices of the neighbor tetrahedral finite elements used to model the bridge. So these straight vertices are mapped in the sequence that are encountered from left to the right edge of the bridge to a parent element so that the left corner of the red curve is mapped to the parent location 0, while the right corner of the red curve is mapped to the parent location 1. Specifically, the 0 m to 537 m curved line along the deck is mapped to the parent domain from 0 to 1.

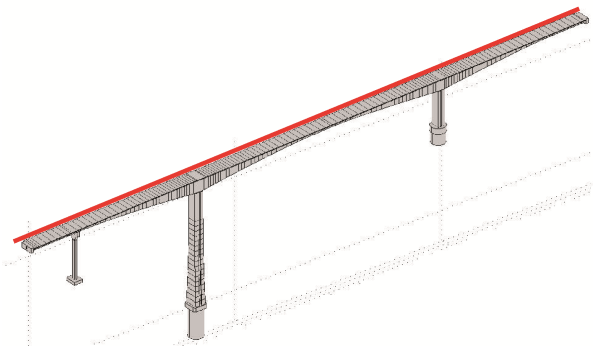


Figure 2. Optimization along pedestrian walkway of the deck

To account for the different type of sensors that can be placed on the physical structures, in our case sensors measuring in the vertical and the transverse directions, the physical curved line with transverse sensors placed on it is mapped to a parent element from 0 to 1, while the physical curved line with vertical sensors placed on it is mapped to a parent element from 1 to 2. In this way one can efficiently handle the two types of transverse and vertical sensors without the need to distinguish them during the optimization. The optimization is performed in the parent domain with the design variables, indicating the location of the sensors in the parent domain, to vary from 0 to 2. A design value inside the parent domain [0,1] corresponds to a transverse sensor which is mapped through the aforementioned mapping to a point on the physical domain (red curve) in Figure 2. Similarly, a design value inside the parent domain [1,2] corresponds to a vertical sensor which is mapped to a point on the physical domain (red curve) in Figure 2.

The mode shape value at a point in the parent domain is obtained by the mode shape value at the corresponding point in the physical domain. Consider the mapping of points from 0 to 1 in the parent domain to the physical domain. Since the mapped point in the physical domain might not correspond to a node of the finite element model, the mode shape value is obtained by interpolation using the mode shape values at the two neighboring nodes of that point in the transverse direction. A similar procedure is used for points between 1 and 2 in the parent domain to find the mode shape component in the vertical direction.

Using this mapping strategy the optimization is performed in a more sophisticated way than what was used in the past. Instead of performing multiple optimizations for one sensor at a time which is predetermined to be either a transverse or vertical sensor, now

it is allowed to optimize for several sensors simultaneously on the bridge pedestrian sideway, without predetermining the types of the sensors. Instead, it is left for the optimization to decide which of the sensors will be transverse and which will be vertical, depending on what region they lie in the parent domain.

So the optimization proceeds as follows. The design vector $\underline{\delta}$ is a vector with as many elements as the number of uniaxial sensors. Each element of this vector describes both the location and the type (transverse or vertical) of a sensor. Each element of the design vector $\underline{\delta}$ is a number on the interval [0 2]. This effectively allows for the optimization of many sensors simultaneously, which includes making the decision about the types of sensors as well along with their location. This is more realistic and gives more freedom to the approach than predetermining how many sensors of each type one should include in $\underline{\delta}$.

Note that the optimal sensor placement methodology can also be readily applied to design a set of triaxial instead of uniaxial sensors. In this case the information from each triaxial sensor will be the combined information obtained from the three sensor components in each direction. The aforementioned concept can also be extended to design a combination of uniaxial (*e.g.*, transverse and vertical) and triaxial sensors by mapping the physical curved line with triaxial sensors placed on it to a parent element from 2 to 3. In this way the three type of sensors can efficiently be handled. However, the triaxial sensors will always be preferred to the transverse or vertical uniaxial sensors since the information gain from a uniaxial sensor placed in an optimal position will always be less than the information gain from a triaxial sensor placed at the same position. To favor a uniaxial sensor in relation to a triaxial sensor one has to introduce additional constrains that penalize the placement of triaxial sensors by taking into account the extra cost of the sensor and the information gain a triaxial sensor will provide in relation to a uniaxial sensor. However, this complicates the optimization, with the subject falling outside the scope of the present work. The previous concepts can be extended to design a combination of different uniaxial and triaxial sensor types such as acceleration, displacement and strain sensors.

5.2.2 Numerical Results for Spatially Uncorrelated Prediction Error Model

An uncorrelated prediction error model is used with diagonal covariance matrix $\Sigma(\underline{\delta};\underline{\sigma}) = \sigma^2 I_{N_0}$, where $I_{N_0} \in R^{N_0 \times N_0}$ is the identity matrix, with the value of the single prediction error parameter chosen to be $\sigma = 0.01$. A Gaussian prior with relatively large uncertainties is used. As a result, the posterior covariance matrix $Q_L(\underline{\delta};\underline{\sigma})$ is non-singular and the design can proceed for any number of sensors. The covariance matrix of the Gaussian prior is set to $S = 10^3 I_m$, with $s_1^2 = \cdots = s_m^2 = 10^3$.

Results for up to $N_0 = 9$ sensors are first obtained. Note that in this case the number of sensors is less than the number of modes ($m = 10$) and so the FIM is singular. The FSSP method is used to obtain the optimal sensor locations. It should be noted, however, that the sensor locations obtained by the heuristic FSSP method are verified that are the global optima, also obtained by the CMA-ES global optimizer[27].

The optimal location of the first sensor corresponds to the minimum of the information entropy. The information entropy as a function of the location of the sensor is drawn in Figure 3(A) at the parent domain. It can be seen that there are 9 local minima. Figure 3(B) shows the contour plots of the information entropy as a function of two sensor locations in the parent domain. A large number of local optima is also observed (blue color). The number of optima are expected to increase as the dimension of the design space increases. Figures 3(A) and 3(B) confirm that there are a large number of local optima and so gradient-based optimization methods will be trapped to a local optimum and will not be able to obtain a global optimum. Stochastic algorithms such as CMA-ES have much higher chances to pinpoint the global optimum.

For the case of optimizing the location of the first sensor, a vertical sensor was obtained rather than a transverse one. The two minima at 1.58 and 1.91 in the parent domain in Figure 3(A) are smaller than all the rest, which implies that for the 1st sensor it is best to be a vertical sensor in one of these two locations which correspond to distances on the bridge of 310 m and 487 m, respectively, from the left end. The minimum at 1.91 (487 m) is slightly smaller, but the difference is negligible for practical purposes.

Another interesting observation in Figure 3(A) is that at the two ends of the bridge (at points 0, 1, and 2 in the parent domain) the value of the posterior entropy is 48.72, which is exactly the information entropy

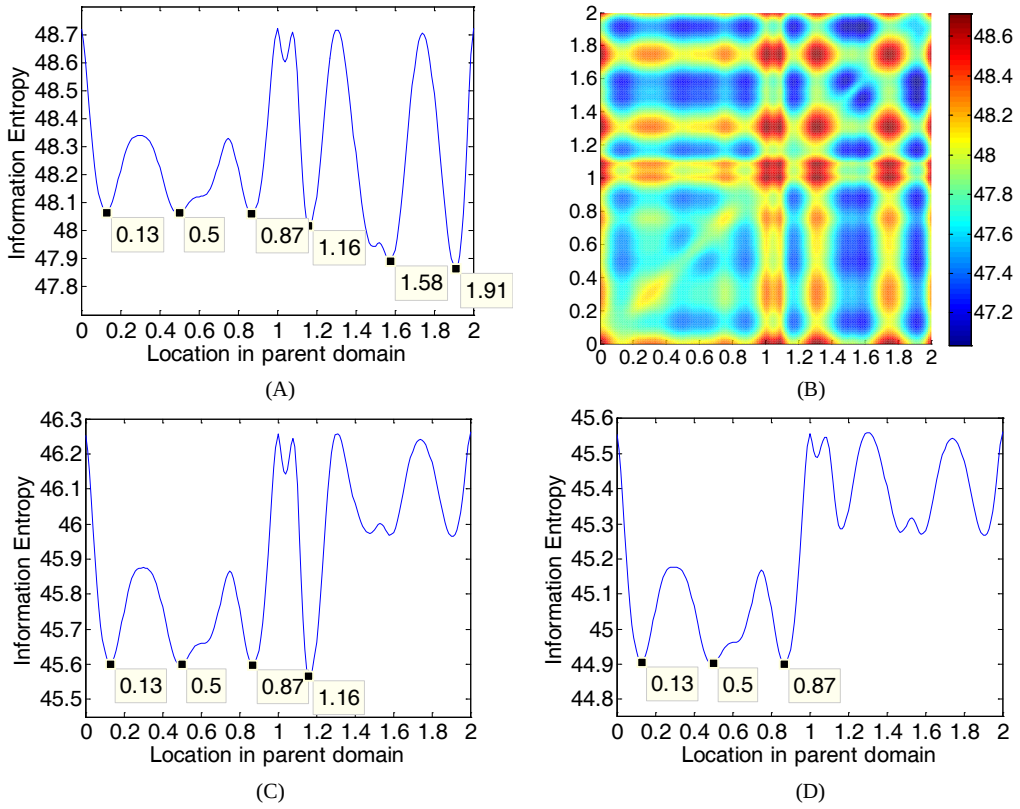


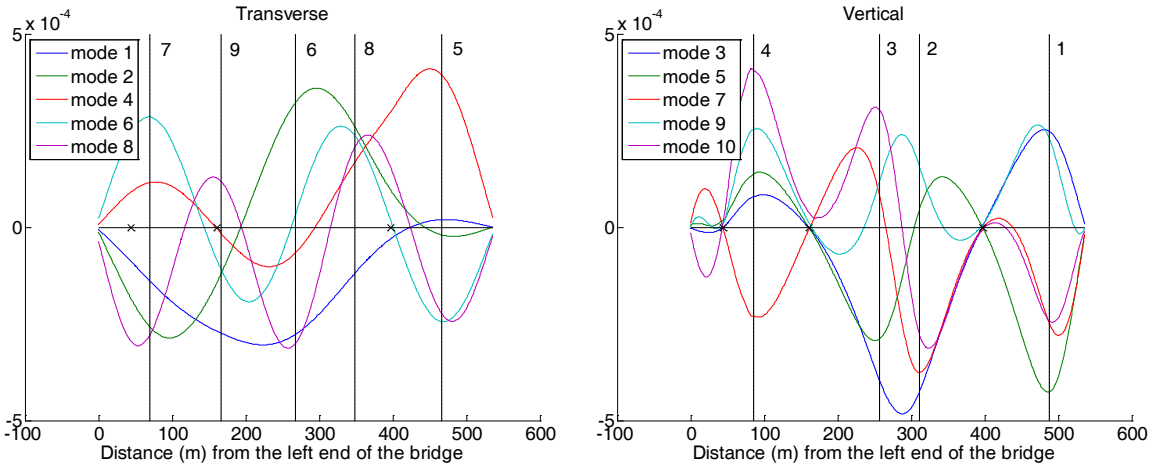
Figure 3. Information entropy vs. the location of a sensor in parent domain (A) 1st sensor, (B) 1st, and 2nd sensor (contour plots), (C) 4th sensor, and (D) 5th sensor.

of the Gaussian prior. This makes sense, because in these points the mode shapes have zero components and the FIM is zero, and therefore the only contribution to the posterior comes from the prior. There are also other points which result in the posterior entropy being almost equal to the prior entropy, which correspond to vertical sensors being placed in points of zero vertical response (locations 1.08, 1.3 and 1.74 in the parent domain) for all vertical mode shapes. These are points in the deck that are above the piers where the vertical motion of the bridge is almost restrained. As one would intuitively expect, the information gain from these designs (which is the difference between the prior and posterior entropies) is zero.

The information entropy as a function of the location of the fourth sensor in the parent domain given that the first three sensors have been placed in their optimal position is shown in Figure 3(C). Similar plot for the information entropy as a function of the position of the 5th sensor given that the first four sensors have been placed in their optimal position, is shown in Figure 3(D). The optimal locations of the first nine sensors are shown in Figures 4(A) and 4(B) for both the transverse and vertical sensors. Also, information of the optimal locations of sensors and the corresponding information entropy values is given in Table 1.

Note that the optimal location of the first two sensors is made up of the location of the global and the next local optimum predicted in Figure 3(A). Also, from the results in Figure 4(B) the first 4 sensors are vertical sensors. This is because, as can be seen from Figures 4(A) and 4(B), vertical modes have slightly higher displacements and therefore contribute more than the transverse modes to the FIM, and therefore to the posterior covariance matrix. Also the vertical modes have more "convenient" points where several modes have high responses compared to the transverse modes. These result in preferring the vertical modes compared to the transverse modes. Note that all modes are treated equally by selecting the prior uncertainty to be the same for all modal coordinates. Finally, results in Figure 4(A) suggest that the last 5 sensors (5th to 9th sensors) are selected to measure in the transverse direction.

From the results in Figure 3(C) of the information entropy as a function of the location of the 4th sensor, given that the first 3 sensors are placed at their optimal location, it is noticed that the vertical sensor at 85 m is



(A) (B)

Figure 4. Optimal locations of (A) transverse sensors, and (B) vertical sensors (Case of 9 sensors with uncorrelated prediction error). The symbol x denotes the location of the piers along the deck.

Table 1. Optimal location in the physical domain and minimum information entropy for the first nine sensors ($N_0 < m$). V = Vertical, T = Transverse.

Sensor	1	2	3	4	5	6	7	8	9
Type	V	V	V	V	T	T	T	T	T
Optimal Location (m)	487	310	257	85	466	267	69	348	166
Information Entropy	47.86	47.03	46.26	45.57	44.89	44.23	43.58	43.02	42.63

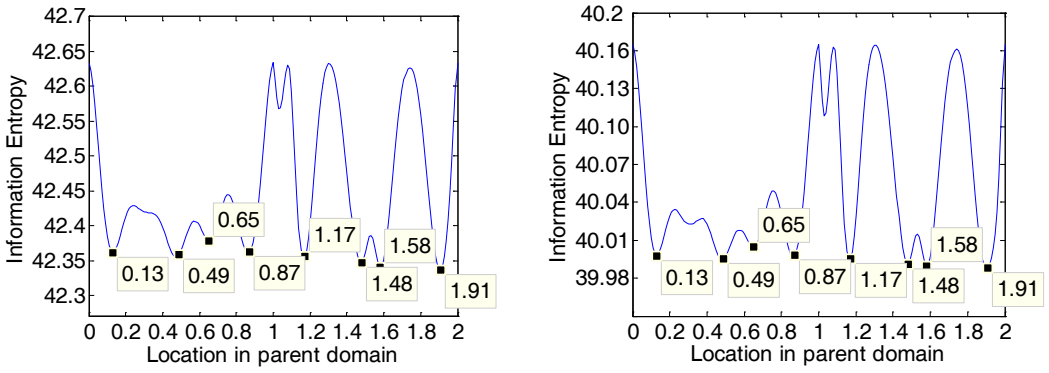
slightly preferred from the three transverse sensors at 69 m, 267 m, and 466 m since the minimum value of 45.57 in the information entropy for the vertical sensor is slightly smaller than the local minimum value of 45.58 for the information entropy for the three transverse sensors. From the results in Figure 3(D) of the information entropy as a function of the location of the 5th sensor, given that the first 4 sensors are placed at their optimal location, it can be observed that any of the locations 69, 267 and 466 m for a transverse sensor are candidates. These sensor locations are the optimal sensor locations for the 5th, 6th and 7th sensor. The optimal sensors for the 8th and 9th sensor are transverse ones, with optimal locations shown in Figure 4(A) and Table 1.

Observing the optimal sensor locations in relation to the mode shapes drawn in Figures 4(A) and 4(B), one should note that the results are reasonable for placing sensors in the suggested vertical or transverse locations since in these locations the mode shape components correspond, in general, to their higher values.

For up to nine sensors it is observed that placing two or more sensors in the same position corresponds to the worst sensor location. This is confirmed also by the contour plots in Figure 3(B) where the placing of the two sensors in the same position is not preferred (red colors). Also it is confirmed by the plots in Figure 3 where it is clear that when a new sensor location coincides with an already placed sensor location it gives large values of the information entropy as compared to the optimal one. So in the singular FIM case with $N_0 < m$, the problem of sensor clustering[16] due to uncorrelated prediction error does not occur.

The optimal location and type (transverse or vertical) of the next 11 sensors (10th to 20th) is also considered. The design is performed using the FSSP algorithm. However, for selected number of sensors, the accuracy of the results obtained from the FSSP algorithm is confirmed by running also the CMA-ES algorithm. Both algorithms provide the same estimates. The information entropy as a function of the location of the 10th sensor given that the first 9 sensors are placed at their optimal positions (shown in Table 1) is presented in Figure 5(A). Figure 5(B) gives similar information but for optimizing the location of the 19th sensor. The optimal sensor locations for the sensors from 10th to 20th, their type (vertical or transverse) and the minimum information entropy are also given in Table 2 and in Figure 6.

It is clear from Figure 5(A) that the optimal location of the 10th sensor coincides with the location of the 1st sensor. Actually the information entropy as a function of the location of the 10th sensor is qualitatively similar to the information entropy for one sensor in Figure 3(A)



(A) (B)

Figure 5. Information entropy vs. the location of a sensor in parent domain (A) 10th sensor, (B) 19th sensor.

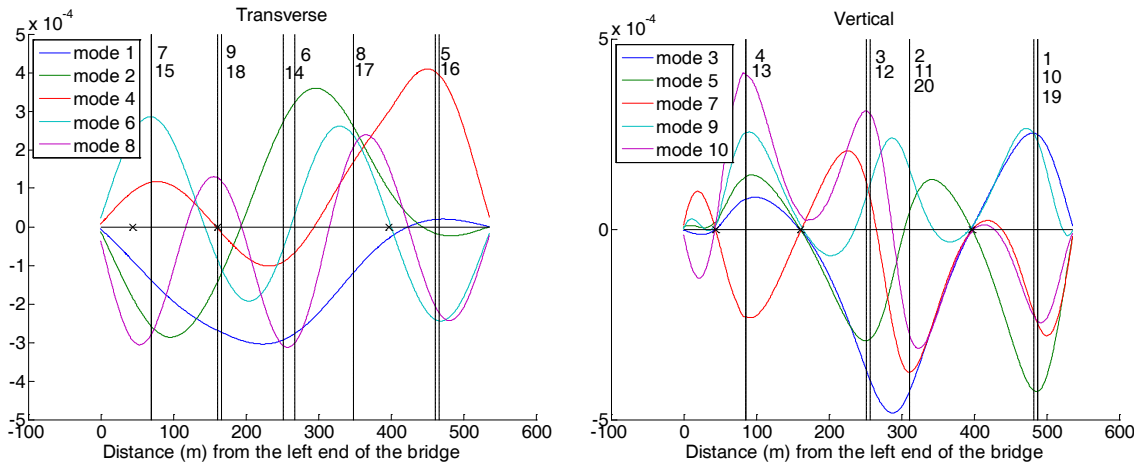
Table 2. Optimal location in the physical domain and minimum information entropy for the 10th up to the 20th sensor ($N_0 \geqslant m$). V = Vertical, T = Transverse.

Sensor	10	11	12	13	14	15	16	17	18	19	20
Type	V	V	V	V	T	T	T	T	T	V	V
Optimal Location (m)	487	310	251	85	251	69	460	348	160	482	310
Information Entropy	42.33	42.04	41.75	41.47	41.19	40.91	40.64	40.39	40.17	39.98	39.80

as a function of its location. Comparing Figures 5(A) and 3(A), the local/global optimal appears at the same locations, suggesting that the optimal locations of the next nine sensors will be close to the optimal locations of the first nine sensors. This is confirmed by comparing also the results in Table 2 with the results in Table 1. Comparisons clearly demonstrate (see also Figure 6) that the optimal locations of the sensors 9th to 20th coincide or they are very close to the optimal locations estimated for the first 9 sensors. This sensor clustering is due to the incorrect assumption of the spatially uncorrelated prediction errors[16]. However, it seems that the sensor design obtained with uncorrelated prediction errors are very reasonable and intuitive for the first nine sensors which correspond to the case of singular FIM with $N_0 < m$. To correct the problem of sensor clustering for more than 9 sensors one has to introduce spatial correlation in the prediction errors.

5.2.3 Numerical Results for Spatially Correlated Prediction Error Model

A spatially correlated prediction error model is assumed next. For this the covariance $\Sigma(\underline{\delta};\underline{\sigma})$ of the prediction error at the sensor locations is selected to be non-diagonal with the (i, j) element of the form



(A) (B)

Figure 6. Optimal locations of (A) transverse sensors, and (B) vertical sensors (Case of 20 sensors with uncorrelated prediction error). The symbol x denotes the location of the piers along the deck.

$\Sigma_{ij}(\underline{\delta};\underline{\sigma}) = \sigma^2 R(\delta_i - \delta_j)$, where $R(\delta_i - \delta_j)$ is the spatial correlation structure of the prediction error, δ_i and δ_j are the locations of the i and j sensors, and σ^2 is the strength of the prediction error. An exponentially decaying correlation structure of the form $R(\delta_i - \delta_j) = \exp(-|\delta_i - \delta_j|/\lambda)$ is selected, where λ is the correlation length.

Since in this formulation we allow for each sensor to be either transverse or vertical, correlation is limited to sensors of the same type. That is, correlation exists between any two transverse sensors or any two vertical sensors, but not between a transverse and a vertical sensor. The described model allows for two different correlation length parameters (or correlation functions in general) to be used for the transverse and vertical sensors, respectively. In the numerical results that follow, the correlation parameters are chosen to be $\sigma = 0.01$ and $\lambda_{trans} = \lambda_{vert} = 10m$ or $\lambda_{trans} = \lambda_{vert} = 20m$.

Optimal sensor placement results for the correlated prediction error models are shown in Figure 7 for two different correlation length of 10 m and 20 m respectively. The Gaussian prior is selected to be the same as in the uncorrelated prediction error case. Results have been derived using the FSSP algorithm and their accuracy has been also confirmed for representative sensor cases using the CMA-ES algorithm. Comparing with the results of the uncorrelated case in Figure 6 it can be observed that the optimal sensor locations for the first 9 sensors are the same as the uncorrelated prediction error case. For more than 9 sensors, the sensor clustering problem is not present in the spatially correlated prediction error case. The optimal locations of sensors 10th to 20th are not close to the

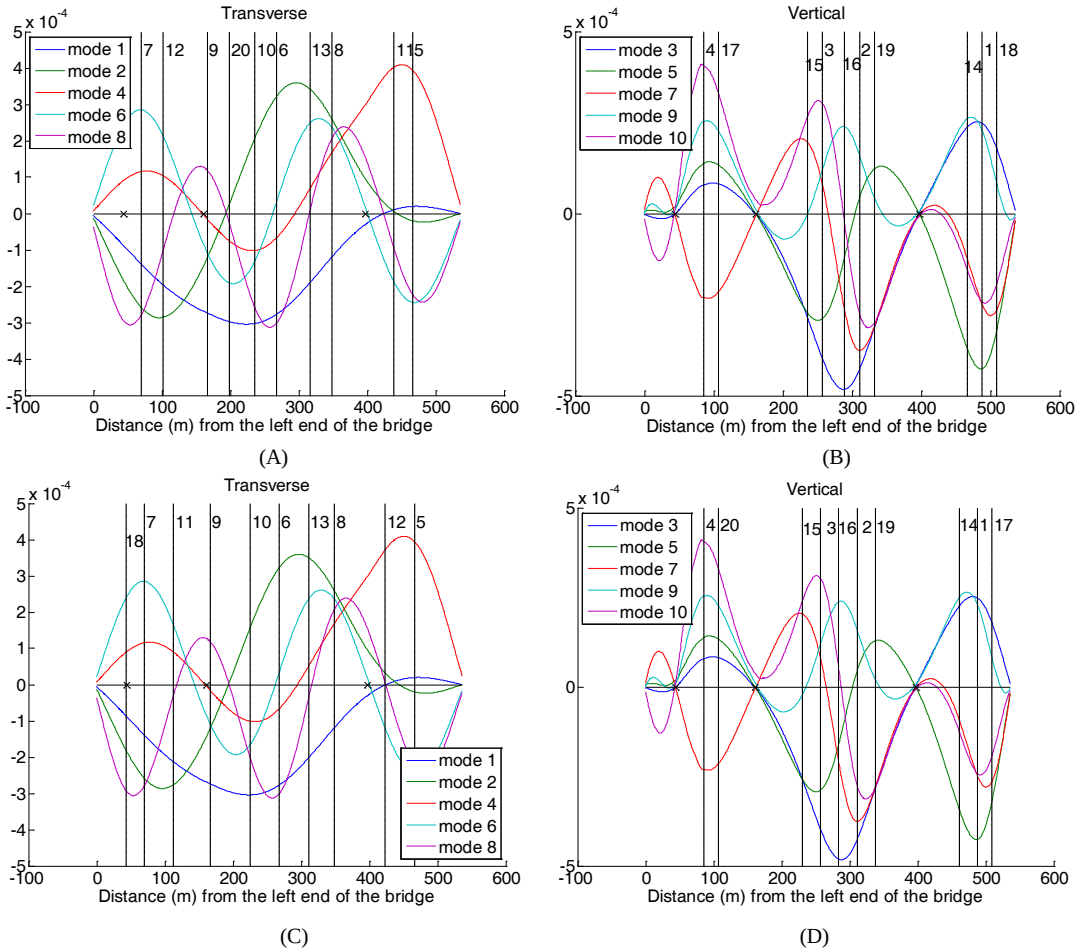


Figure 7. Optimal locations of (A) transverse sensors ($\lambda = 10m$), (B) vertical sensors ($\lambda = 10m$), (C) transverse sensors ($\lambda = 20m$), and (D) vertical sensors ($\lambda = 20m$) (Case of 20 sensors with spatially correlated prediction error). The symbol x denotes the location of the piers along the deck.

locations obtained for the first 9 sensors. In fact, they are more uniformly distributed in some of the areas of the bridge deck. Comparing Figures 7(A,B) with Figures 7(C,D) it can be observed that the spacing of the sensors tends to increase as one increases the correlation length from 10 m to 20 m. This is consistent with the theoretical results obtained by Papadimitriou and Lombaert[16].

In Figure 8 the optimal information entropy as a function of the number of sensors is shown for the uncorrelated and correlated prediction error cases and for 1 up to 30 sensors. The 0 sensor case corresponds to the case where no sensors are placed and so the information entropy is that of the Gaussian prior PDF. As expected, we notice that the optimal value of the information entropy for a given number of sensors decreases as the number of sensors increases. For up to 9 sensors the entropies are identical between the uncorrelated and correlated prediction error models. From the 10th sensor and on, the correlated prediction error models lead to more information entropy (less information) and this entropy increases with the correlation length. This is due to the fact that the information provided by neighbor sensors within the correlation length assumed is not significantly different and so this results in a drop of the total information for the same number of sensors. Also, the curves with high correlation suggest that the information provided by adding sensors in the structure is decreasing and eventually after a number of sensors there is no significant information offered by additional sensors. This plot can be used to decide on the number of sensors to be placed in the structure, given the correlation length. The uncorrelated prediction error models provide misleading results since continuing adding sensors in the structure has the effect of gaining additional information, independent of the number of added sensors, which is counter-intuitive. Finally, the CMA-ES algorithm is also used to design the optimal sensor locations and the resulting information entropy values for representative sensor cases, shown in Figure 8 for correlation length 20 m, match exactly the information entropy values obtained using the FSSP algorithm, confirming in this case the accuracy of FSSP algorithm.

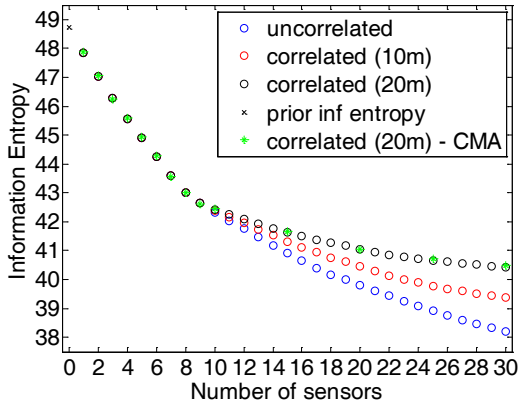
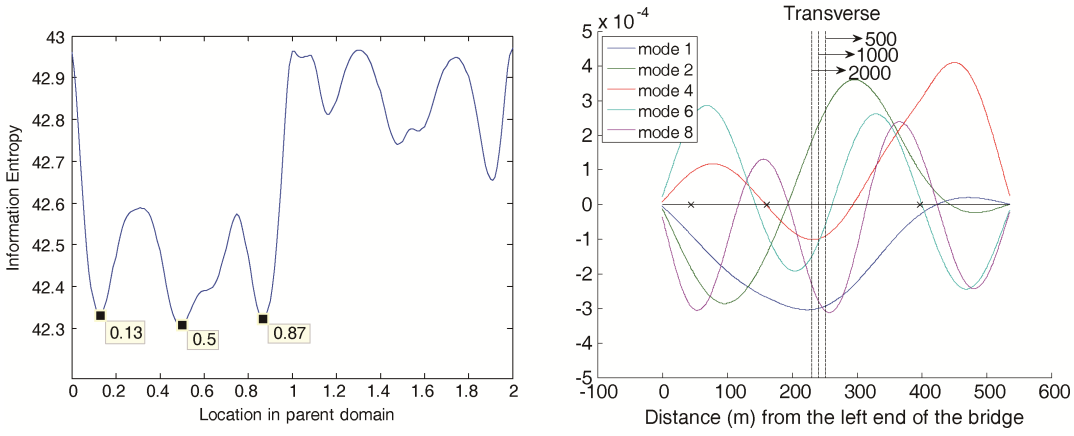


Figure 8. Optimal information entropy vs. number of sensors

5.3 Effect of Prior Uncertainty on Optimal Sensor Placement

The results obtained so far correspond to an isotropic Gaussian prior where all prior uncertainties in the parameters were selected to be the same. This gives equal weight to all parameters as far as the prior is concerned, so there is no preference of a specific mode over another. The Gaussian prior covariance matrix is a modelling choice that depends on user preference. Therefore, it can be fully manipulated according to the needs. Equation (19) suggests that by giving larger prior uncertainties to some specific modes we are essentially giving more weight in these modes in the selection of the optimal design. The insightful result from (19) states that a Gaussian prior can be used as a means to perform more sophisticated OED, where we give preference to some selected modes over others. Different Gaussian prior variances for the different modes get transferred to the posterior and result in different optimal designs, favoring the identification of modes with the largest prior variances. The Bayesian framework for OED provides the means to fully quantify this preference of some modes (or parameters for identification in general) over others through the prior.

In order to illustrate this, the simple case of one sensor is examined. We would like to give more weight to the identification of the modal coordinates of the transverse modes. For this we lower the variances of the bending modes from 1000 to 100, and we keep the variances of the transverse modes at 1000. The information entropy as a function of the position of the 1st sensor is shown in Figure 9(A). We see that the optimal design for the 1st sensor is now a transverse sensor in one of the already found optimal locations for transverse sensors, shown in Figure 3(D). In fact, the plot for one sensor in Figure 9(A) is now qualitatively very similar to the plot of the 5th sensor shown in Figure 3(D), where the first transverse sensor appears as optimal. With this change in the prior



(A) (B)

Figure 9. (A) Information entropy of 1st sensor vs location in parent domain using Gaussian prior variance $s^2 = 1000$ for transverse modes and $s^2 = 100$ for vertical modes. (B) Optimal locations for 1st sensor for prior variance of 1st mode equal to 500, 1000 and 2000. All other prior variances equal to 100.

variances giving weight to the transverse modes, the optimal designs have a preference for the transverse sensors now to show up as the 1st sensor and not the 5th.

Consider next the case of giving preference to the first modal coordinate which is transverse. The design is performed by setting all prior variances, except the first, equal to 100, while we set the prior variance of the first modal coordinate to different values $s_1^2 = 300$, 500, 1000 and 2000. For $\sigma_1^2 = 300$ or for smaller values of s_1^2, the optimal sensor is selected to be a vertical sensor. For values of $s_1^2 = 500$, 100 and 2000, the transverse sensor is selected. The location of the transverse sensor for these higher values is shown in Figure 9(B). It is clear that increasing the value of s_1^2 the sensor is located closer to the location where the first mode has its maximum absolute deflection. This is consistent with the theoretical result in Section 4.2 which states that increasing the prior variance of the first mode gives more and more weight to the identification of the first mode, making it more important in the optimal sensor placement design.

5.4 Optimal Sensor Placement of Reference Sensors

The OED is next used to address the important problem of selecting the optimal location of a few reference sensors in a multiple sensor configuration set up experiment conducted with a limited number of reference and roving sensors in order to obtain the modal frequencies and assemble the mode shapes from the multiple setups. The number of reference sensors is in most cases significantly smaller than the number of modes to be identified. It is important in this case that the reference sensors, common in most setups, contain the maximum possible information for all modes that are planned to be identified. Wrong locations of the reference sensors may degrade the modal information for one or more modes, degrading the accuracy of the corresponding assembled mode shape since such accuracy is based solely on the information contained in the reference sensors.

The effectiveness of the methodology is illustrated by designing the optimal locations of one vertical and one transverse reference sensor for identifying the lowest 10 modes of the bridge. It is clear from the results in Figure 3(A) that the best location of the 1st sensor is at 487 m measuring along the vertical direction. This is also taken as the location of the vertical reference sensor. Figure 4(B) demonstrates that the design of the reference sensor at 487 m is rational since the deflection of all five vertical mode shapes is high. To design the location of the 1st transverse reference sensor one could use the results in Figure 9(A), obtained for the transverse modes after selecting the prior uncertainty to give in the optimal design preferential treatment to these modes over the vertical modes. It is clear that the transverse reference sensor can be selected as the location 268 m (parent location 0.5) that corresponds to the minimum information entropy in Figure 9(A). From Figure 4(A) it is seen that the transverse reference sensor location 268 m corresponds to high deflections of four out of the five transverse modes. One of the transverse modes has relatively small deflection, a problem that arises from the trade-off that has to be made in the design to get

maximum information from all five modes. A solution to this problem is to place a second transverse sensor on the structure at position 70 m (the second best local optimum in Figure 4(A)). In practice, using more reference sensors that the minimum required (two in this case) is a good way to make sure that important information from reference sensors will not be lost. Concluding, the vertical and transverse reference sensors in experiments with multi sensor configuration setups could be 487 m and 268 m (and/or 70 m), respectively.

6. Conclusions

An optimal sensor placement design for modal identification based on the expected K-L divergence as a measure of the information contained in the data is shown to be equivalent to the optimal sensor placement design based on the information entropy proposed by Papadimitriou[18]. Using non-uniform priors, the Bayesian OED allows for optimal sensor placement to be performed even for the case when the number of sensors is less than the number of identified modes. This is important when designing the optimal locations of a very small number of reference sensors for the purpose of assembling the mode shapes using reference and roving sensors in multiple sensor configuration setups. In this study the effect of the Gaussian prior on the optimal design was thoroughly investigated. Insightful analytical expressions were derived to show that larger uncertainty in the prior of a subset of modal coordinates can be used to give preference in this subset in the optimal design of the sensor locations. The prior variances for all modes to be identified can be altered to weight the importance of different modes in the design, favoring a number of modes against the rest of the modes. The prior is the users' choice and it can be used in different ways to achieve different results, which is one of the strengths of Bayesian OED.

The methodology was applied to a 537 m long reinforced concrete bridge in order to design the optimal sensor configuration for identifying the lowest 10 modes. The optimization was performed in the continuous space of the design variables, through appropriate mapping from the physical space to a parent domain. A large number of local optima were observed that result in a challenging optimization problem. The problem is overcome using computationally efficient heuristic FSSP algorithms. The accuracy of the FSSP algorithm was confirmed using the computationally demanding CMA-ES algorithm. A thorough investigation of the effect of correlated and uncorrelated prediction error models was also performed. The design for a smaller number of sensors than the number of modes was shown to be the same for spatially uncorrelated and correlated prediction error models. Rational and intuitive results were obtained. For more sensors than the number of modes the spatially correlated prediction error model gave intuitively reasonable results, avoiding sensor clustering observed for uncorrelated prediction error models.

The proposed method offers a useful decision tool for designing the sensor locations in a structure in order to obtain the maximum information for reliable modal identification of civil infrastructures and industrial facilities using vibration measurements.

Conflict of Interest and Funding

No conflict of interest was reported by the authors. This research has been implemented under the "ARISTEIA" Action of the "Operational Programme Education and Lifelong Learning" and was co-funded by the European Social Fund (ESF) and Greek National Resources.

References

1. Simoen E, Moaveni B, Conte J P, *et al*., 2013, Uncertainty quantification in the assessment of progressive damage in a 7-story full-scale building slice. *Journal of Engineering Mechanics – ASCE*, vol.139(12): 1818–1830. http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0000610
2. Beck J L and Katafygiotis L S, 1998, Updating models and their uncertainties. I: Bayesian statistical framework. *Journal of Engineering Mechanics – ASCE*, vol.124(4): 455–461. http://dx.doi.org/10.1061/(ASCE)0733-9399(1998)124:4(455)
3. Yuen K V, 2012, Updating large models for mechanical systems using incomplete modal measurement. *Mechanical Systems and Signal Processing*, vol.28: 297– 308. http://dx.doi.org/10.1016/j.ymssp.2011.08.005
4. Yuen K V, 2010, *Bayesian Methods for Structural Dynamics and Civil Engineering*, John Wiley and Sons: NJ. http://dx.doi.org/10.1002/9780470824566
5. Lam H, Katafygiotis L and Mickleborough N, 2004, Application of a statistical model updating approach on phase I of the IASC-ASCE structural health monitoring benchmark study. *Journal of Engineering Mechanics – ASCE*, vol.130(1): 34–48. http://dx.doi.org/10.1061/(ASCE)0733-9399(2004)130:1(34)
6. Vanik M, Beck J and Au S, 2000, Bayesian probabilistic approach to structural health monitoring, *Journal of En-*

gineering Mechanics – ASCE, vol.126(7): 738–745. http://dx.doi.org/10.1061/(ASCE)0733-9399(2000)126:7(738)

7. Li D, 2011, *Sensor placement methods and evaluation criteria in structural health monitoring*, PhD. thesis, University of Siegen, Siegen.
8. Kammer D C, 1991, Sensor placements for on-orbit modal identification and correlation of large space structures, *Journal of Guidance, Control and Dynamics*, vol.14(2): 251–259. http://dx.doi.org/10.2514/3.20635
9. Kammer D C, 1992, Effects of noise on sensor placement for on-orbit modal identification of large space structures, *Journal of Dynamic Systems, Measurements and Control*, vol.114(3): 436–443. http://dx.doi.org/10.1115/1.2897366
10. Li D S, Li H N and Fritzen C P, 2009, A note on fast computation of effective independence through QR downdating for sensor placement. *Mechanical Systems and Signal Processing*, vol.23(4): 1160–1168. http://dx.doi.org/10.1016/j.ymssp.2008.09.007
11. Shah P C and Udwadia F E, 1978, A methodology for optimal sensor locations for identification of dynamic systems, *Journal of Applied Mechanics*, vol.45(1): 188–196. http://dx.doi.org/10.1115/1.3424225
12. Udwadia F, 1994, Methodology for optimal sensor locations for parameter identification in dynamic systems, *Journal of Engineering Mechanics* (*ASCE*), vol.120(2): 368–390. http://dx.doi.org/10.1061/(ASCE)0733-9399(1994)120:2(368)
13. Papadimitriou C, Beck J L and Au S K, 2000, Entropy-based optimal sensor location for structural model updating, *Journal of Vibration and Control*, vol.6(5):-781–800. http://dx.doi.org/10.1177/107754630000600508
14. Yuen K V, Katafygiotis L S, Papadimitriou C, *et al*., 2001, Optimal sensor placement methodology for identification with unmeasured excitation. *Journal of Dynamic Systems, Measurement and Control*, vol.123(4): 677–686. http://dx.doi.org/10.1115/1.1410929
15. Ye S Q and Ni Y Q, 2012 Information entropy based algorithm of sensor placement optimization for structural damage detection. *Smart Structures and Systems*, vol. 10(4–5): 443–458. http://dx.doi.org/10.12989/sss.2012.10.4_5.443
16. Papadimitriou C and Lombaert G, 2012, The effect of prediction error correlation on optimal sensor placement in structural dynamics. *Mechanical Systems and Signal Processing*, vol.28: 105–127. http://dx.doi.org/10.1016/j.ymssp.2011.05.019
17. Kammer D C, 2005, Sensor set expansion for modal vibration testing. *Mechanical Systems and Signal Processing*, vol.19(4): 700–713. http://dx.doi.org/10.1016/j.ymssp.2004.06.003
18. Papadimitriou C, 2004, Optimal sensor placement methodology for parametric identification of structural systems. *Journal of Sound and Vibration*, vol.278(4): 923–947. http://dx.doi.org/10.1016/j.jsv.2003.10.063
19. Stephan C, 2012, Sensor placement for modal identification. *Mechanical Systems and Signal Processing*, vol.27: 461–470.http://dx.doi.org/10.1016/j.ymssp.2011.07.022
20. Papadimitriou D and Papadimitriou C, 2015, Optimal sensor placement for the estimation of turbulence model parameters in CFD. *International Journal for Uncertainty Quantification*, vol.5(6): 545–568. http://dx.doi.org/10.1615/Int.J.UncertaintyQuantification.2015015239
21. Leyder C, Ntertimanis V K, Chatzi E, *et al*., 2015, Optimal sensor placement for the modal identification of an innovative timber structure. *Volume 1st ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2015*, pp. 467–476, National Technical University of Athens.
22. Leyder C, Chatzi E, Frangi A, *et al*., 2016, Comparison of optimal sensor placement algorithms via implementation on an innovative timber structure, IALCCE Conference, *Fifth International Symposium on Life-Cycle Civil Engineering (IALCCE 2016)*, 16–19 October, 2016, Delft, The Netherlands.
23. Chow H M, Lam H F, Yin T, *et al*., 2011, Optimal sensor configuration of a typical transmission tower for the purpose of structural model updating. *Structural Control and Health Monitoring*, vol.18(3): 305–320. http://dx.doi.org/10.1002/stc.372
24. Yuen K V and Kuok S C, 2015, Efficient Bayesian sensor placement algorithm for structural identification: a general approach for multi-type sensory systems. *Earthquake Engineering and Structural Dynamics*, vol.44(5):757–774. http://dx.doi.org/10.1002/eqe.2486
25. Lindley D V, 1956, On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, vol.27: 986–1005. http://dx.doi.org/10.1214/aoms/1177728069
26. Huan X and Marzouk Y M, 2013, Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, vol.232(1): 288– 317. http://dx.doi.org/10.1016/j.jcp.2012.08.013
27. Hansen N, Muller S D, and Koumoutsakos P, 2003, Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, vol.11(1): 1–18. http://dx.doi.org/10.1162/106365603321828970
28. Chaloner K and Verdinelli I, 1995, Bayesian experimental design: a review. *Statistical Science*, vol.10(3): 273– 304. http://dx.doi.org/10.1214/ss/1177009939