

1 Analysis of Information Entropy

1.1 Dependence of information entropy on number of sensors

Useful results are derived next that show how the information entropy and its lower and upper bounds depend on the number of sensors. Let L_M denote the sensor configuration involving M sensors. Define also the expression $L_{M+N} = (L_M^T \ L_N^T)^T$ to represent the sensor configuration that is formed from the configuration L_M and N additional sensors placed on the structure as specified by the configuration L_N . Then, the following proposition holds:

Proposition 1: The information entropy for a sensor configuration L_M involving M sensors is higher than the information entropy for a sensor configuration L_{M+N} involving N additional sensors. That is,

$$H(L_{M+N}; \underline{\theta}_0, \Sigma) \leq H(L_M; \underline{\theta}_0, \Sigma) \quad (1)$$

The proof of Proposition 1 was presented for the special case of uncorrelated prediction errors in [2]. The proof of the Proposition 1 for the general case of correlated prediction errors is more involved and is presented next.

Proof: Using (12), it suffices to show that the following inequality holds for two sensor configurations L_{M+N} and L_M :

$$\det[Q(L_{M+N})] \geq \det[Q(L_M)] \quad (2)$$

where the dependence of $Q(L; \underline{\theta}_0, \Sigma)$ on $\underline{\theta}_0$ and Σ has been dropped for notational convenience. The following statements are next shown to be valid: (i) the matrix $Q(L)$ is symmetric semi-positive definite, and (ii) the matrix $Q(L_{M+N})$ with $N \geq 1$ admits the representation

$$Q(L_{M+N}) = Q(L_M) + \delta Q_{MN}, \quad \delta Q_{MN} \geq 0 \quad (3)$$

where the notation $\delta Q_{MN} \geq 0$ denotes that the matrix δQ_{MN} is a symmetric semi-positive definite matrix. The proof of statement (i) follows by exploiting the special form (13) of the matrix $Q(L)$. It can be readily shown that for every non-zero vector $\underline{y} \in R^{N_\theta}$, the quadratic form

$$\underline{y}^T Q(L) \underline{y} = \sum_{k=1}^N (L \nabla_{\theta} x_k \underline{y})^T (L \Sigma L^T)^{-1} (L \nabla_{\theta} x_k \underline{y}) = \sum_{k=1}^N \underline{z}^T (L \Sigma L^T)^{-1} \underline{z} \geq 0 \quad (4)$$

where $\underline{z} = L \nabla_{\theta} x_k \underline{y}$, is always non-negative since the covariance matrix $L \Sigma L^T$ is by construction symmetric positive definite. Thus, the matrix $Q(L)$ is symmetric semi-positive definite. The proof of statement (ii) is given in the Appendix A. Given the

representation (3) and the fact that δQ_{MN} is semi-positive definite, the proof of Proposition 1 follows the same procedure as presented in [2]. For completeness, the main steps of the proof are given next. Substituting (3) into (2), it remains to show the validity of the inequality

$$\det[Q(L_M) + \delta Q_{MN}] \geq \det[Q(L_M)], \quad N \geq 0 \quad (5)$$

This statement can be shown using the fact that for two symmetric semi-positive definite matrices $A_0 \in R^{N_\theta \times N_\theta}$ and $B_0 \in R^{N_\theta \times N_\theta}$ the following is true:

$$\lambda_r[A_0 + B_0] \geq \lambda_r[A_0] \geq 0, \quad r = 1, \dots, N_\theta \quad (6)$$

where the symbol $\lambda_r[A_0]$ denotes the r -th eigenvalue of the matrix A_0 . The last inequality can be derived from the application of the minimax theorem for eigenvalues of symmetric matrices. Applying the inequality (6) for $A_0 = Q(L_M)$, $B_0 = \delta Q_{MN}$, and using

the fact that $\det A_0 = \prod_{r=1}^{N_\theta} \lambda_r[A_0]$ for any matrix A_0 , the inequality (5) is readily derived.

Proposition 1 implies that the information entropy reduces as additional sensors are placed in a structure. Given the interpretation of the information entropy as a measure of the uncertainty in the parameter estimates, this should be intuitively expected since adding one or more sensors in the structure will have the effect of providing more information about the system parameters and thus reducing the uncertainty in the parameter estimates.

As a direct consequence of the Proposition 1, the following proposition is also true.

Proposition 2: The minimum and maximum information entropies for M sensors are decreasing functions of the number of sensors, M .

This reduction of the information entropy as a function of the number of sensors is expected since increasing the number of sensors has the effect of extracting more information from the data. The Proposition 2 follows directly from the Proposition 1, independent of the correlation model assumed for the prediction error. Thus, the reader is referred to reference [2] for a proof.

Propositions 1 and 2, shown in this work to be valid for spatially correlated prediction error, were employed in [2] to justify the use of the heuristic algorithms FSSP and BSSP for efficiently constructing sub-optimal solutions to the optimal sensor location problem.

1.2 Effect of correlation length on the distance between sensors

Let L_M be a sensor configuration involving M sensors that have already been placed on the structure. Let also Δ be the minimum distance between any two sensors in the sensor configuration L_M . It is assumed that the correlation length λ of the prediction error is small enough compared to the minimum distance Δ between any two sensors in L_M . This is sufficient to guarantee that the prediction errors between the responses at any two sensors in L_M are uncorrelated. Consider a new sensor to be placed on the structure and

let δ be the distance of the new sensor from one of the existing sensors in L_M . Without loss of generality it can be assumed that the new sensor will be placed closer to the M -th sensor in the configuration. Otherwise, the sensor numbering can be re-arranged such that the sensor the closest to the new sensor is the M -th sensor. The sensor location for the new sensor is defined by $L_1 \in R^{1 \times N_d}$. Consider that δ varies from values of zero up to the order of the correlation length λ so that, using the fact that λ is small compared to Δ , the prediction error at the position of the additional sensor is correlated to the closest M -th sensor and uncorrelated to the prediction errors from all other $M-1$ sensors in the sensor configuration L_M . Let $\hat{H}(\delta) \equiv H_\delta(L_{M+1}; \underline{\theta}_0, \Sigma)$ denote the information entropy as a function of the distance δ for $M+1$ sensors located according to the sensor configuration $L_{M+1} = (L_M^T \ L_1^T)^T$. Similarly let also $\hat{Q}(\delta) \equiv Q_\delta(L_{M+1}; \underline{\theta}_0, \Sigma)$ denote the corresponding information matrix.

Proposition 3: Consider the problem of placing an additional sensor on a structure instrumented with M sensors. The information entropy for the sensor configuration L_{M+1} involving $M+1$ sensors is a decreasing function of the shortest distance δ of the new sensor from the other M sensors, provided that δ is sufficiently small. Mathematically, this proposition reads:

$$\hat{H}(\delta_1) < \hat{H}(\delta_2) \quad \text{for any} \quad \delta_1 > \delta_2 \quad (7)$$

or, equivalently, using (12)

$$\det \hat{Q}(\delta_1) > \det \hat{Q}(\delta_2) \quad \text{for any} \quad \delta_1 > \delta_2 \quad (8)$$

where δ_1 and δ_2 are sufficiently.

The proof of Proposition 3 is presented in Appendix B and shows that the proposition holds for distances δ smaller than the characteristic length of the structural dynamic problem under consideration. Considering that the response is a superposition of structural modes, this length is controlled by the characteristic length of the highest contributing mode which defines the length scale over which the response sensitivities with respect to the parameters fluctuate in space.

Expression (7) or (8) implies that sensors locations further away from an existing sensor have a higher information content. Consequently, the spatial correlation of the prediction error tends to shift a sensor away from existing sensor locations. Over a distance larger than the characteristic length, the spatial change of the response sensitivity will eventually control how far away the new sensor is placed from the existing ones.

2 Optimal Sensor Placement in Structural Dynamics

2.1 Bayesian inference for parameter estimation

Consider a parameterized class of structural models (e.g. a class of finite element models or a class of modal models) chosen to describe the input-output behavior of a structure. Let $\underline{\theta} \in R^{N_\theta}$ be the vector of free parameters (physical or modal parameters) in the model class that need to be estimated using measured data D collected from a sensor network.

Let $D = \{\underline{y}_k, k = 1, \dots, N\}$ be the measured sampled response time histories data, where $\underline{y}_k \in R^{N_0}$ refer to output data, N_0 is the number of observed degrees of freedom (DOF) of the structural model, k denotes the time index at time $k\Delta t$, Δt is the sampling interval, and N is the number of sampled data. Let $\underline{x}_k(\underline{\theta}) \in R^{N_d}$, $k = 1, \dots, N$ be the sampled response time histories computed at all N_d model DOFs from a structural model that corresponds to a specific value $\underline{\theta}$ of the model parameters. The measured response and model response predictions at time instant $k\Delta t$ satisfy the prediction error equation

$$\underline{y}_k = L \underline{x}_k(\underline{\theta}) + L \underline{e}_k(\underline{\theta}) \quad (9)$$

where $\underline{e}_k(\underline{\theta})$ is the prediction error due to modelling error and measurement noise. The matrix $L \in R^{N_0 \times N_d}$ is the observation matrix comprised of zeros and ones and maps the model DOFs to the measured DOFs. The matrix L therefore defines the location of the sensors in the structure.

Using a Bayesian identification methodology, the uncertainties in the values of the parameters $\underline{\theta}$ are quantified by probability density functions (PDF) that are obtained using the dynamic test data D and the probability model for the prediction error $\underline{e}_k(\underline{\theta})$. In what follows, the prediction error vector $\underline{e}_k(\underline{\theta})$ at time $k\Delta t$ is modeled as a Gaussian random vector with zero mean and covariance $\Sigma \in R^{N_d \times N_d}$. Also, it is assumed that the prediction errors between different time instances are independent. Applying the Bayes' theorem, the updating PDF $p(\underline{\theta} | \Sigma, D)$ of the set of structural model parameters $\underline{\theta}$ given the measured data D and the prediction error parameters Σ , takes the form:

$$p(\underline{\theta} | \Sigma, D) = c \frac{1}{(\sqrt{2\pi \det \Sigma})^N} \exp\left[-\frac{NN_0}{2} J(\underline{\theta}; \Sigma, D)\right] \pi(\underline{\theta}) \quad (10)$$

where

$$J(\underline{\theta}; \Sigma, D) = \frac{1}{NN_0} \sum_{k=1}^N [\underline{y}_k - L\underline{x}_k(\underline{\theta})]^T \Sigma^{-1} [\underline{y}_k - L\underline{x}_k(\underline{\theta})] \quad (11)$$

represents the measure of fit between the measured and the model response time histories, $\pi(\underline{\theta})$ is the prior distribution for the parameter set $\underline{\theta}$, and c is a normalizing constant chosen such that the PDF in (2) integrates to one.

2.2 Asymptotic estimate of the information entropy

An asymptotic approximation of the information entropy, valid for large number of data ($NN_0 \rightarrow \infty$), is available which is useful in the experimental stage of designing an optimal sensor configuration. The asymptotic approximation is obtained by substituting $p(\underline{\theta} | \underline{\sigma}, D)$ from (10) into the information entropy and observing that the resulting integral can be re-written as a Laplace-type integral which can be approximated by applying Laplace method of asymptotic approximation [1]. Specifically, it can be shown

that for a large number of measured data, i.e. as $NN_0 \rightarrow \infty$, the following asymptotic result holds for the information entropy [2]

$$I(L; D) \square H(L; \underline{\theta}_0, \underline{\sigma}) = \frac{1}{2} N_\theta \ln(2\pi) - \frac{1}{2} \ln \det[Q(L; \underline{\theta}_0, \underline{\sigma}) + Q_\pi(\underline{\theta}_0)] \quad (12)$$

where $\underline{\theta}_0 \equiv \hat{\underline{\theta}}(L, \underline{\sigma}, D)$ is the optimal value of the parameter set $\underline{\theta}$ that minimizes the measure of fit $J(\underline{\theta}; L, D)$ given in (11), and $Q(L; \underline{\theta}, \Sigma)$ is an $N_\theta \times N_\theta$ semi-positive definite matrix defined as $N_D N \nabla_{\underline{\theta}} \nabla_{\underline{\theta}}^T J(\underline{\theta}; \underline{\sigma}, D)$ and asymptotically approximated by

$$Q(L; \underline{\theta}, \Sigma) = \sum_{k=1}^N [L \nabla_{\underline{\theta}} x_k(\underline{\theta})]^T (L \Sigma(\underline{\sigma}) L^T)^{-1} [L \nabla_{\underline{\theta}} x_k(\underline{\theta})] \quad (13)$$

in which $\nabla_{\underline{\theta}} = [\partial/\partial\theta_1, \dots, \partial/\partial\theta_{N_\theta}]^T$ is the usual gradient vector with respect to the parameter set $\underline{\theta}$. The matrix $Q(L, \underline{\theta}, \underline{\sigma})$ is a semi-positive definite matrix, known as the Fisher information matrix (FIM), containing the information about the uncertainty in the values of the parameters $\underline{\theta}$ based on the data from all measured positions specified in L . Details for the derivation in the special case of diagonal covariance matrix $\Sigma = \sigma^2 I$, where I is the identity matrix can be found in [2]. Note that the asymptotic estimate of the matrix in (13) can readily be obtained by following the same steps as the ones presented in [2] for the uncorrelated case.

In the initial stage of designing the experiment, the data and consequently the values of the optimal model parameters $\hat{\underline{\theta}}$ and the form of the prediction error covariance matrix Σ are not available. In practice, useful designs can be obtained by taking the optimal model parameters $\hat{\underline{\theta}}$ and prediction error covariance Σ to have some nominal values $\underline{\theta}_0$ and Σ to arise from a correlation function chosen by the designer to be representative of the system and the expected model and measurement errors.

3 Formulation as a Discrete-Valued Optimization Problem

In experimental design, the sensors are placed in the structure such that the resulting measured data are most informative about the parameters of the model class used to represent the structure behavior. Since the information entropy gives the amount of useful information contained in the measured data, the optimal sensor configuration is selected as the one that minimizes the information entropy [3]. That is,

$$L_{best} = \arg \min_L H(L; \underline{\theta}_0, \underline{\sigma}) \quad (14)$$

where $H(L; \underline{\theta}_0, \underline{\sigma})$ is given by (12) and the minimization is constrained over the set of N_p measurable DOFs. The lower bound of the information entropy is then given by $H_{min} = H(L_{best}; \underline{\theta}_0, \underline{\sigma})$.

It should be noted that the upper bound of the information entropy corresponding to the worst sensor configuration is also useful since, when it is compared with the minimum information entropy for the same number of sensors, it gives a measure of the reduction that can be achieved by optimising the sensor configuration. The maximum information

entropy and the corresponding worst sensor configuration can be obtained by maximizing instead of minimizing the information entropy. The worst sensor configuration is obtained as

$$L_{worst} = \arg \max_L H(L; \underline{\theta}_0, \underline{\sigma}) \quad (15)$$

while the upper bound of the information entropy is given by $H_{max} = H(L_{worst}; \underline{\theta}_0, \underline{\sigma})$

3.1 Computational algorithms

Two heuristic sequential sensor placement (SSP) algorithms, the forward (FSSP) and the backward (BSSP), were proposed [15,17] for constructing predictions of the optimal and worst sensor configurations.

According to FSSP (algorithm 1), the positions of N_0 sensors are computed sequentially by placing one sensor at a time in the structure at a position that results in the highest reduction in information entropy. BSSP (algorithm 2) is used in an inverse order, starting with N_d sensors placed at all DOFs of the structure and removing successively one sensor at a time from the position that results in the smallest increase in the information entropy.

Algorithm 1: Forward Sequential Sensor Placement (FSSP)

1. **Initialize:** no sensors selected, *number of sensors* $N = 0$ and *sensor configuration* $L_N = \{ \}$
2. **While** *number of sensors* $N <$ *maximum number of sensors* N_0 **do**
 - a. Consider combinations with one additional sensor, $N = N + 1$.
 - b. **For** *counter* $i = 1$ **to** *number of possible sensor positions* $N_d - N + 1$
 - i. Obtain configuration L_N by adding sensor i to configuration L_{N-1}
 - ii. Evaluate information entropy of new sensor configuration L_N .
 - c. **End**
 - d. Select the sensor configuration L_N that minimizes the information entropy.
3. **End**

Algorithm 2: Backward Sequential Sensor Placement (BSSP)

1. **Initialize:** all sensors selected, *number of sensors* $N = N_d$ and *sensor configuration* L_N
2. **While** *number of sensors* $N > 1$ **do**
 - a. Consider combinations with one sensor less, $N = N - 1$.
 - b. **For** *counter* $i = 1$ **to** *number of possible sensors to be removed* $N + 1$
 - i. Obtain configuration L_N by removing sensor i from configuration L_{N+1} ,
 - ii. Evaluate information entropy of new sensor configuration L_N .
 - c. **End**

- d. Select the sensor configuration L_N that minimizes the information entropy.

3. End

Using the FSSP algorithm an approximation to the worst sensor configuration can also be obtained efficiently by placing successively one sensor at a time in the position that results in the smallest decrease in information entropy. Similarly, using the BSSP algorithm, an approximation to the worst sensor configuration is obtained by removing successively one sensor at a time from the position that results in the highest increase in the information entropy value.

The computations involved in the SSP algorithms are an infinitesimal fraction of the ones involved in the exhaustive search method and can be done in realistic time, independently of the number of sensors and the number of model DOFs. Although the SSP algorithms are not guaranteed to give the optimal solution, they were found to be effective and computationally attractive alternatives to the GAs. However, when necessary, GAs can improve the SSP estimates, converging to the optimal solutions.

4 Applications in Structural Dynamics

The optimal sensor location methodology is implemented for applications in structural dynamics. For a linear structural model, arising from the discretization of continuous domain using the finite element method, the governing equations of motion are

$$M\ddot{\underline{u}}(t) + C\dot{\underline{u}}(t) + K\underline{u}(t) = \Gamma\underline{z}(t) \quad (16)$$

where $M \in R^{N_d \times N_d}$, $C \in R^{N_d \times N_d}$ and $K \in R^{N_d \times N_d}$ are the mass, damping and stiffness matrices, respectively, $\underline{u}(t) \in R^{N_d}$ is the displacement response vector, $\underline{z}(t) \in R^{N_t}$ is the vector of independent input forces, and $\Gamma \in R^{N_d \times N_t}$ is the input selection matrix. The measured response quantities are assumed to be either displacements, or velocities or absolute accelerations with the sampled measured response $\underline{x}_k \in R^{N_d}$ given respectively by either \underline{u}_k , or $\dot{\underline{u}}_k$ or $\ddot{\underline{u}}_k$. Strain measurements can readily be accommodated in the formulation.

The optimal sensor location design depends on the type of parameters considered for estimation. Two different categories of problems are treated next based on the selection of the model and the parameter set.

4.1 Design of optimal sensor locations for modal identification

The first category deals with the estimation of the modal coordinate vector $\underline{\xi} \in R^m$ ($m \leq N_d$) encountered in modal identification. The objective is to design the sensor configuration that provides the most information in order to estimate the modal coordinate vector $\underline{\xi}$. In this case, the parameter set $\underline{\theta} \equiv \underline{\xi}$. Following the conventional modal analysis, the response vector $\underline{x} \in R^{N_d}$ is given with respect to the parameter set $\underline{\theta}$ in the form $\underline{x} = \Phi\underline{\theta}$, where $\Phi \in R^{N_d \times m}$ is the mode shape matrix for m contributing

modes. Noting that $\nabla_{\underline{\theta}} \underline{x} = \Phi$ and substituting into (13), the information matrix takes the form

$$Q(L, \underline{\theta}_0, \underline{\sigma}) \equiv Q(L, \underline{\sigma}) = (L\Phi)^T (L\Sigma(\underline{\sigma})L^T)^{-1} (L\Phi) \quad (17)$$

which is independent of the nominal parameter values $\underline{\theta}_0$. In addition, the optimal sensor locations are independent of the excitation used.

Based on the form of (17), a non-singular FIM matrix $Q(L, \underline{\sigma})$ is obtained only if the number of sensors, N_0 , is at least equal to the number of contributing modes, m , or the number of parameters, N_θ ($N_\theta = m$). Otherwise, for $N_0 < m$, the matrix $Q(L, \underline{\sigma})$ in (17) is singular and the determinant of the FIM will be zero for any sensor configuration. Thus, for $N_0 < m$ the optimal sensor location problem cannot be performed when a uniform prior PDF is assumed. This means that the information content in the measured data is not sufficient to estimate all the parameters simultaneously. The problem is critical for the FSSP algorithm where one starts with no sensors placed on the structure and sequentially adds one sensor at a time on the structure. The estimation of the sensor locations will be a problem for a small number of sensors, $N_0 < m$, and will considerably affect the optimal sensor location for $N_0 \geq m$. The remedy is to use a non-uniform prior so that the matrix $Q(L, \underline{\sigma}) + Q_\pi(\underline{\theta}_0)$ is non-singular. In this case, the optimal sensor locations will also depend on the prior information along certain directions in the parameter space for which the model itself fails to provide information due to insensitivity of output quantities with respect to the parameters in these directions.

For non-informative uniform prior PDF for which $Q_\pi(\underline{\theta}_0) = 0$, one way to optimally place sensors in the structure for $N_0 < m$ is to maximize the product of the N_0 non-zero eigenvalues in the FIM, instead of maximizing the product of all eigenvalues. This procedure allows to systematically and optimally place sensors in the structure even for the unidentifiable case that arises for a small number of sensors. This procedure considerably improves the FSSP estimates for $N_0 \geq m$.

4.2 Design of optimal sensor locations for model parameter estimation

The second category deals with the estimation of structural-related properties. In this case, the parameter set $\underline{\theta}$ includes variables related to stiffness, mass, and damping characteristics. The system mass $M(\underline{\theta})$, stiffness $K(\underline{\theta})$ and damping $C(\underline{\theta})$ matrices depend on the parameter set $\underline{\theta}$. To compute the response sensitivity matrix $\nabla_{\underline{\theta}} \underline{x}_k$ needed in (13), a sensitivity analysis needs to be performed. Following a conventional modal analysis, writing the response vector in the form $\underline{u}(t) = \Phi \underline{\xi}(t)$, where the modal coordinates $\underline{\xi}(t)$ satisfy the uncoupled system of modal equations

$$\ddot{\underline{\xi}} + C^* \dot{\underline{\xi}} + \Lambda \underline{\xi} = \Phi^T \Gamma \underline{z} \quad (18)$$

and differentiating the modal equations with respect to the i -th component θ_i of the parameter set $\underline{\theta}$, the following analytical differential equations for the components

$\underline{\mu}_i(t) \equiv \frac{\partial \underline{x}(t)}{\partial \theta_i}$, $i = 1, \dots, N_\theta$ of the sensitivity matrix $\nabla_{\theta} \underline{x}(t) = [\underline{\mu}_1(t), \underline{\mu}_2(t), \dots, \underline{\mu}_{N_\theta}(t)]$ are readily obtained in the form:

$$\underline{\mu}_i(t) = \Phi \underline{\eta}_i(t) + \frac{\partial \Phi}{\partial \theta_i} \underline{\xi}(t) \quad (19)$$

where $\underline{\eta}_i(t)$ is computed from the modal equations (18) to satisfy the following system of equations

$$\ddot{\underline{\eta}}_i + C^* \dot{\underline{\eta}}_i + \Lambda \underline{\eta}_i = -\ddot{\underline{\xi}}(t) - \frac{\partial C^*}{\partial \theta_i} \dot{\underline{\xi}} - \frac{\partial \Lambda}{\partial \theta_i} \underline{\xi} + \frac{\partial(\Phi^T \Gamma)}{\partial \theta_i} \underline{z}(t) \quad (20)$$

The matrix Λ in (18) and (20) is diagonal with elements ω_i^2 , while assuming classically damped modes the matrix C^* is also diagonal with elements $2\zeta_i \omega_i$, where ω_i and ζ_i are the i -th modal frequency and modal damping ratio of the structure, respectively. The formulation in the modal space allows one to perform computationally efficient analyses for models with a large number of DOFs by selecting only the contributing modes. The sensitivities of the eigenvalues in Λ and the eigenvectors in Φ are readily obtained from the sensitivities of the mass and stiffness matrices using well established techniques [4,5]. For more information on sensitivity of eigenvalues and eigenvectors for complex eigensystems, the reader is referred to the overview paper in [6].

The optimal sensor locations depend on the location and type of excitations used. Also, in contrast to the modal identification case, the matrix $Q(L, \underline{\theta}_0, \underline{\sigma})$ may be non-singular even for one sensor since the time history response obtained from the model for a given nominal input excitation may contain enough information from all contributing modes of the structure in order to estimate the parameter set $\underline{\theta}$. As before, in the case that FIM is singular, a systematic way to optimally place the first few sensors in FSSP method is to maximize the product of the non-zero eigenvalues of the FIM.

5 Numerical Demonstration

5.1 Simply-supported continuous beam model

The objective is to demonstrate the problem that arises in the design of the optimal sensor location for spatially continuous structures in which two sensors can potentially be placed very close to each other. The design of the optimal location of two sensors for a simply supported uniform beam of length h is considered. The mode shapes needed in (17) are given by $\sin(n\pi x/h)$ for the n mode and they are independent of the material and cross-sectional properties. The sensors can be placed at any point along its axis. To illustrate the effect of spatial correlation, it is assumed that only the third mode contributes to the dynamics of the beam. The correlation function for the prediction error between two points located at distance η is given by $R(\eta) = \exp(-\eta/\lambda)$, with $\bar{\sigma}^2 = 0$ and equal variance values $\Sigma_{ii} = \bar{\sigma}^2$ for all i . A uniform prior PDF is assumed so that $Q_\pi(\underline{\theta}_0) = 0$.

The contour plots of $\det Q(L, \theta_0, \Sigma)$ as a function of the locations x/h of the two sensors along the beam axis are presented in Figure 1. Representative values of the correlation lengths $\lambda = 0, 0.1h, 0.2h$ and $0.4h$ have been selected to demonstrate the effect of the correlation length on the sensor locations. The optimal locations of the two sensors correspond to the combination of approximately the three locations $x/h = 0.167, 0.5$ and 0.833 on which the third mode peaks (negative or positive peaks). It is clearly seen that the spatial correlation of the prediction errors affects the optimal location of sensors. The uncorrelated case ($\lambda = 0$) results in six distinct global optimal sensor configurations arising from all possible combinations of the locations of the three peaks of the third mode shape. However, the three optimal sensor configurations with coordinates $(0.167, 0.167)$, $(0.5, 0.5)$ and $(0.833, 0.833)$ shown in Figure 1(a) correspond to both sensors placed at exactly the same locations. Such sensors locations are contrary to expectations since in practical applications sensors are never placed at the same position or neighborhood positions. These locations have exactly the same information content and such configurations should be avoided.

The selection of these three sensor configurations arises from the uncorrelated assumption used for the prediction errors. In reality, prediction errors between neighborhood locations are correlated due to model error. Such correlation when it is included in the formulation results in optimal sensor locations that are consistent with designer's expectations. This is seen in the results for non-zero correlation length. For $\lambda = 0.1h$, only three global solutions, with coordinates approximately $(0.167, 0.5)$, $(0.167, 0.833)$ and $(0.5, 0.833)$ shown in Figure 1(b), are left as global solutions. As before, such solutions arise from all possible combinations of the locations of the three peaks of the third mode shape, excluding the sensor configurations that involve both sensors at exactly the same position. As the correlation length increases to $\lambda = 0.2$ and $\lambda = 0.4$, only two global solutions remain (positions with coordinates $(0.167, 0.5)$ and $(0.5, 0.833)$) as it is seen in Figures 1(c) and 1(d) which, due to symmetry of the third mode shape, correspond to sensors placed approximately at the peaks of the third mode with opposite sign. The global optimal sensor configuration involving one sensor at the location $x/h = 0.167$ and the second sensor at location $x/h = 0.833$ become suboptimal as the correlation length increases, since the correlation of the prediction errors for these two locations becomes stronger. As a result, one of these two positions is excluded from the optimal positions as the two sensor locations tend to provide similar information content to the total FIM. Finally, it should be noted that higher correlation lengths ($\lambda = 0.4$) have the effect of moving the optimal sensor locations slightly away from the positions where the contributing third mode peaks.

It is clearly demonstrated that the minimum distance between the two sensors depends, among other factors, on the spatial correlation length assumed between the prediction errors. This has implications in designing sensor locations for models with potentially very close sensor locations, as arising from fine mesh discretization of continua using numerical finite element methods. Introducing spatial correlation between prediction errors, the optimal sensor locations become independent of the mesh refinement in finite element models, a property that is desirable in experimental design to avoid redundant information provided from closely spaced sensors.

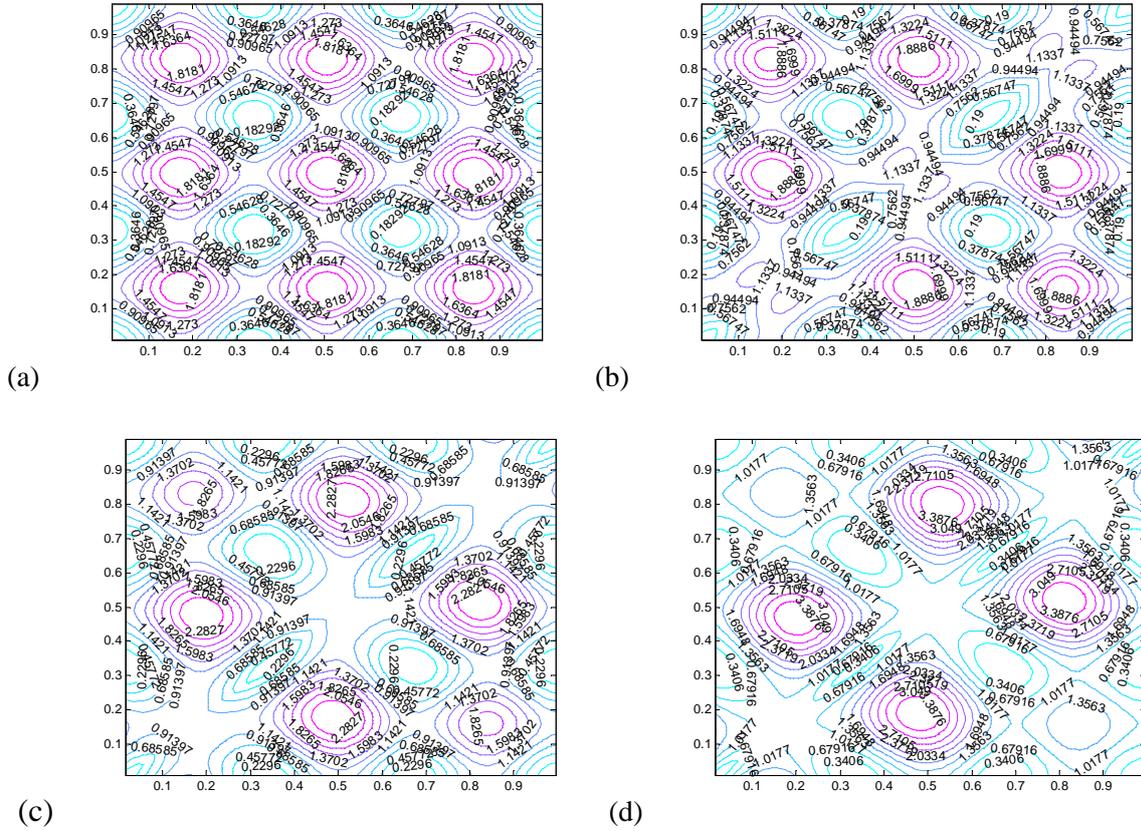


Figure 1: Contour plots of $\det Q(L, \theta_0, \Sigma)$ as a function of the locations x/h of the two sensors along the beam axis (a) $\lambda = 0h$ (uncorrelated case), (b) $\lambda = 0.1h$, (c) $\lambda = 0.2h$, and (d) $\lambda = 0.4h$.

5.2 20-DOF Spring-mass chain-like model

The methodology is applied next to a 20-DOF chain-like spring-mass model, fixed at the bottom spring end and free at the top twentieth mass. The DOF are numbered consecutively starting from the bottom of the chain. A model with a small number of DOFs is purposely chosen to facilitate comparisons between the two SSP algorithms and the exact exhaustive search method that considers all possible sensor configurations. The structure is subdivided into five substructures that each consist of four consecutive masses and springs. The structure is parameterised using five parameters, with the i -th parameter modelling the spring stiffness k_i of the i -th substructure. The masses are considered to be same for all links in the chain. The distance between any two consecutive masses in the chain is chosen to be h . The nominal structure corresponds to a uniform stiffness distribution along the chain. The ratio of the spring stiffness k_i to the mass m_i of a link is chosen to be $k_i/m_i = 1 [l/s^2]$. Classical normal modes are assumed with the modal damping fixed at 5% for all modes. The structure is subjected to an impulse excitation of unit magnitude at the top mass of the model. This impulse excitation can be viewed as simulating the excitation in impact hammer tests.

Optimal sensor placement is applied to identify the optimal sensor locations for the estimation of the stiffness k_i , $i = 1, \dots, 5$ of the five substructures in the system. The time history of the response is sampled with a time step Δt equal to $T_{\min}/6$, with T_{\min} the period of the largest natural frequency of the system. A total of $N = 2048$ points is considered in the response. The total measurement time $N\Delta t$ is slightly larger than 13 times the natural period of the system.

Normalized information entropy results are presented by defining the information entropy index (IEI) as

$$\text{IEI}(L) = \exp[H(L; \underline{\theta}_0, \Sigma) - H(L_{ref}; \underline{\theta}_0, \Sigma)] = \left[\frac{\det Q(L_{ref}, \underline{\theta}_0, \Sigma)}{\det Q(L, \underline{\theta}_0, \Sigma)} \right]^{1/2} \quad (21)$$

where $H(L; \underline{\theta}_0, \Sigma)$ is the reference information entropy computed for a reference sensor configuration L_{ref} . The IEI is a measure of the uncertainty in the parameter estimates relative to the uncertainty obtained for the reference sensor configuration. The reference sensor configuration is selected as the one involving sensors at all model DOFs so that the $\text{IEI}(L)$ values, when compared to one, give the effectiveness of the sensor configuration and the maximum improvement that can be achieved by sensor configuration strategies.

The prediction error correlation model $R(\eta) = \exp(-\eta/\lambda)$ is used. To study the effect of spatial correlation on the optimal sensor locations, results are presented for two values $\lambda = 0.002h$ and $\lambda = h$ of the correlation length. The first value $\lambda = 0.002h$, chosen to be very small compared to the distance between the masses in the chain, corresponds to spatially uncorrelated prediction errors. The second value $\lambda = h$, chosen to be of the order of the distance between the masses in the chain, strongly correlates the prediction

errors between measurements at any two consecutive masses. To facilitate the interpretation of the optimal sensor configuration, and to clearly demonstrate the effect of spatial correlation, the number of contributing modes is kept small and equal to three so that the characteristic length of the problem, defined by the characteristic length of the highest mode (for the third mode is approximately $20h/3 \approx 7h$), is significantly larger than the correlation length $\lambda = h$.

The minimum (best) and maximum (worst) information entropy index values $IEI(L)$ as a function of the sensors, computed by the exhaustive search method (exact method) and the FSSP and BSSP algorithms, are shown in Figure 3 for both the spatially uncorrelated prediction error (UNC-PE) and spatially correlated prediction error (COR-PE) cases. The IEI values predicted by the heuristic FSSP and BSSP methods are extremely good estimates of the minimum information entropy predicted by the exhaustive search method. Comparing the optimal predictions from the FSSP and BSSP methods, the results are indistinguishable for all sensor numbers considered and for both the UNC-PE and COR-PE cases. Differences between the two heuristic methods exist for the maximum (worst) IEI values for the UNC-PE case for a small number of sensors where the BSSP method fails to give the exact estimates. If necessary, these estimates for the worst IEI can be improved to match the exact estimates by using GAs.

The minimum and maximum IEI are decreasing functions of the number of sensors placed in the structure at the optimal and worst positions, respectively. This is consistent with the theoretical result stated in Proposition 2. Comparing the variation of the minimum IEI values as a function of the number of sensors, it can be observed that a significantly higher reduction rate is observed for the COR-PE case in Figure 3(b). Specifically, a drastic reduction in the minimum IEI is observed for the first 5 sensors, accounting almost for the most information provided by the data, while the remaining fifteen sensors cause only a relatively small reduction. It is evident that a small number of sensors placed at their optimal locations may contain more valuable information than a higher number of sensors arbitrarily placed in the structure. For example, for the COR-PE case, one, four and five sensors placed at their optimal locations yield better information than seven, seventeen and eighteen sensors, respectively, placed at their worst sensor locations.

The corresponding condition numbers for the information matrix $Q(L, \underline{\theta}, \Sigma)$ are also computed using the exhaustive search method and the two SSP methods and shown in Figure 3. Reasonable values of the condition numbers are observed in Figure 3 which indicates that in all cases considered the five parameter values are identifiable, independent of the number of sensors.

The optimal sensor locations as a function of the number of sensors are shown in Figure 4 for the FSSP and BSSP algorithms and for both the UNC-PE and COR-PE cases. Given a prediction error model (UNC-PE or COR-PE), it is seen that both the FSSO and BSSP algorithms give exactly the same estimates. These results are also compared to the ones obtained from the exact method for up to 7 sensors. The SSP predictions coincide in all cases with the exact one provided by the exhaustive search method.

Comparing Figures 4(a,b) with Figures 4(c,d), it is clear that the prediction error correlation affects considerably the optimal location of sensors. Specifically, for the

UNC-PE case, the optimal locations of the first three sensors are concentrated at the top masses (DOFs 20, 19 and 18) of the chain model, while for the COR-PE case the sensors are uniformly distributed along the chain at the 20th, 12th and 4th DOFs. Neighboring sensor locations are not promoted in the case of spatial correlation due to the fact that these locations provide similar information. It should be noted that as the number of sensors increases, the locations are uniformly distributed along the chain for the COR-PE case, while for the UNC-PE case the location of the sensors tend to cluster around the 20th, 12th and 4th DOFs which are the first three optimal locations predicted by the COR-PE model. The results in this application are consistent with the derived theoretical results showing that introducing spatial correlation between prediction errors tends to keep the sensors apart at a distance controlled by the magnitude of the correlation length. In contrast, UNC-PE models tend to cluster sensors at locations that provide redundant information. In this case, a number of sensors are wasted before new sensors are installed at locations that provide more useful information.

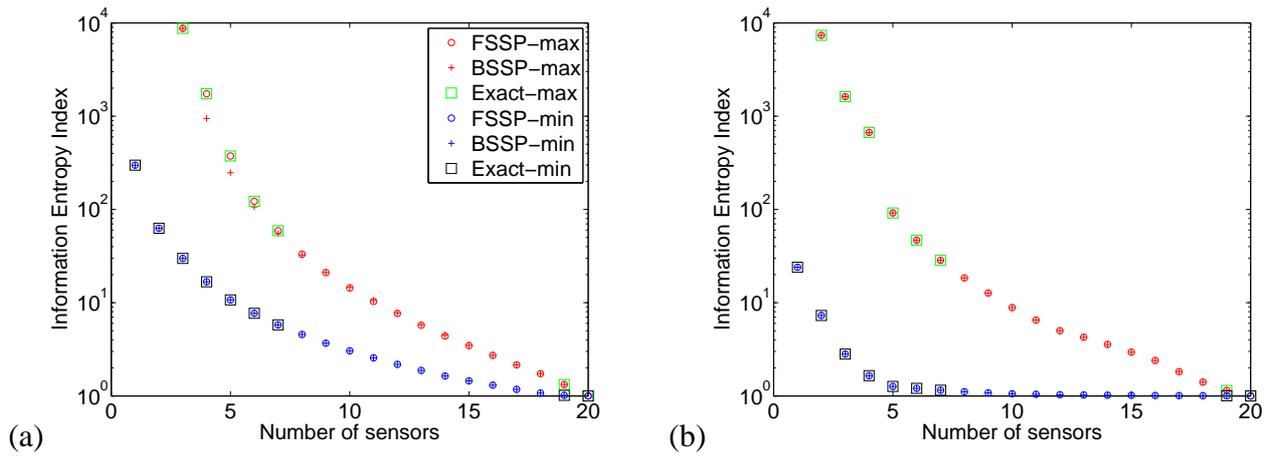


Figure 2: Information entropy index as a function of the number of sensors for the optimal and worst sensor configuration; (a) $\lambda = 0.002h$ and (b) $\lambda = h$.

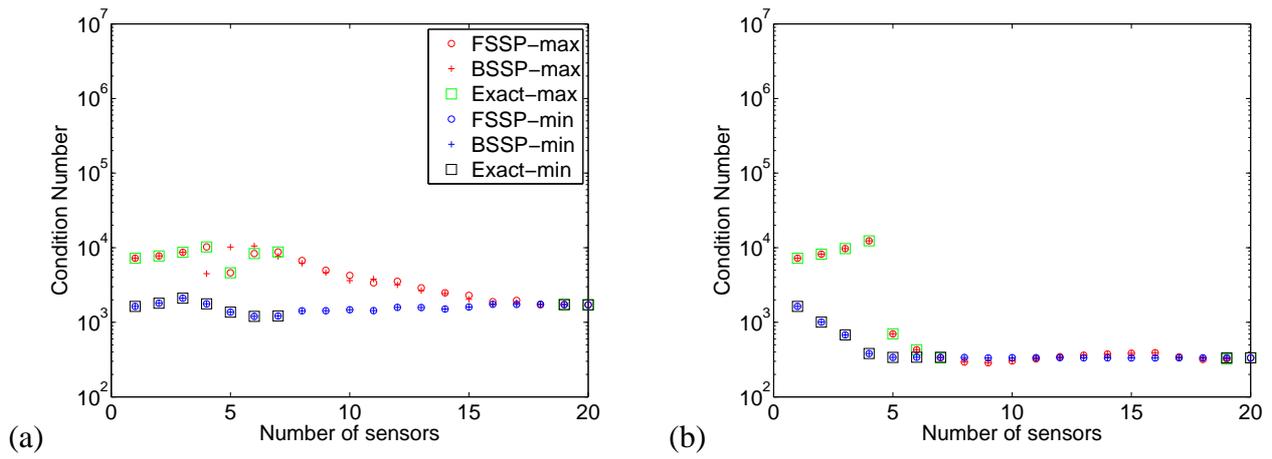


Figure 3: Condition number of matrix $Q(L, \underline{\theta}_0, \Sigma)$ as a function of the number of sensors for the optimal and worst sensor configuration (a) $\lambda = 0.002h$ and (b) $\lambda = h$.

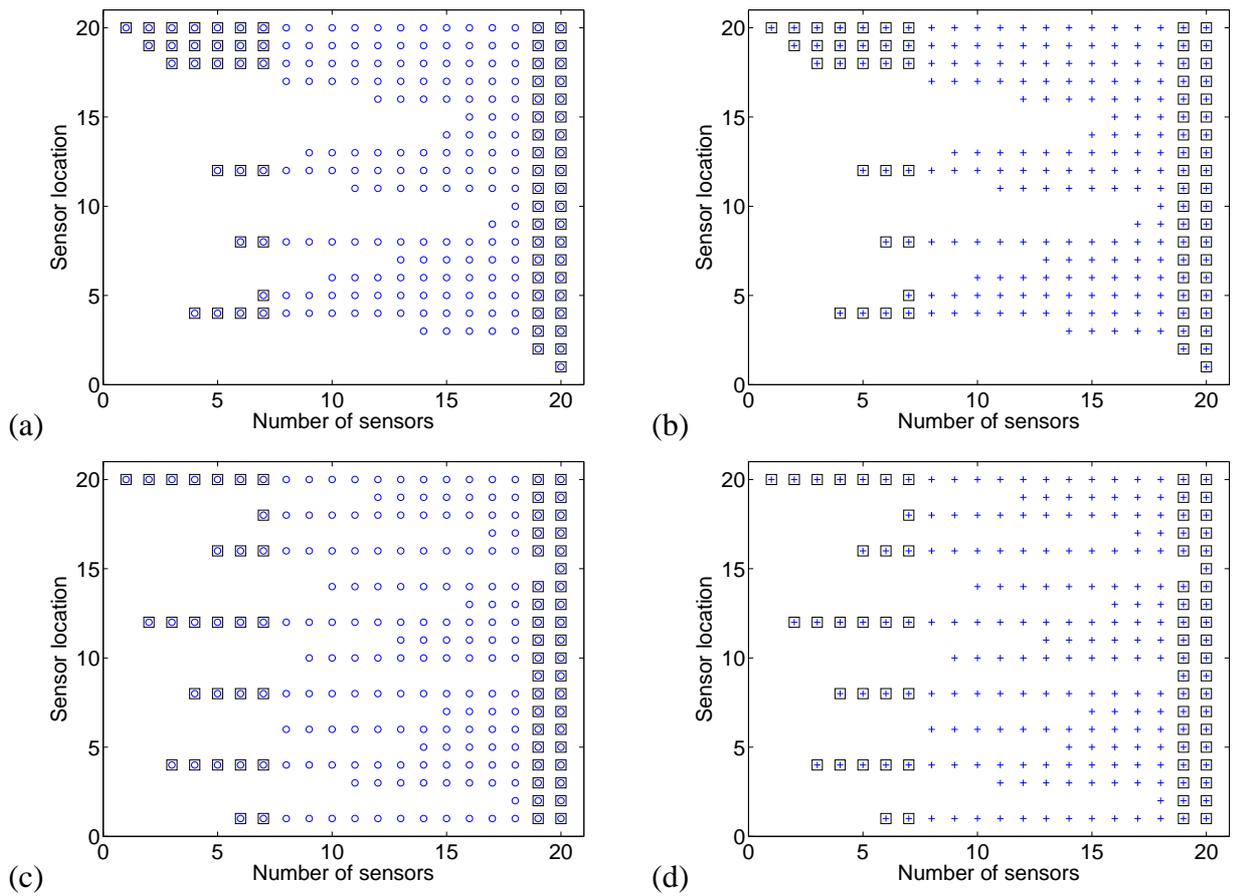


Figure 4: Optimal sensor configuration (a) FSSP for $\lambda = 0.002h$, (b) BSSP for $\lambda = 0.002h$, (c) FSSP for $\lambda = h$, (d) BSSP for $\lambda = h$; Exhaustive search (black squares).

Appendix A: Proof of Representation (3)

The representation (3) and the semi-positive definiteness of the matrix δQ_{MN} is shown as follows. First note that using (13) the matrix $Q(L_{M+N})$ for the sensor configuration L_{M+N} involving $M + N$ sensors admits the representation

$$Q(L_{M+N}) = \sum_{k=1}^N (L_{M+N} \underline{\nabla}_{\theta} \underline{x}_k)^T A^{-1} (L_{M+N} \underline{\nabla}_{\theta} \underline{x}_k) \quad (\text{A.1})$$

where the matrix A is given by $A = L_{M+N} \Sigma L_{M+N}^T$ with L_{M+N} given by $L_{M+N} = (L_M^T \ L_N^T)^T$. Denoting by $B \equiv A^{-1}$ the inverse of A , partitioning the matrices A and B according to the formulas

$$A = \begin{bmatrix} A_M & A_{MN} \\ A_{NM} & A_N \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_M & B_{MN} \\ B_{NM} & B_N \end{bmatrix} \quad (\text{A.2})$$

where $A_M = L_M \Sigma L_M^T$, $A_{MN} = L_M \Sigma L_N^T$, $A_N = L_N \Sigma L_N^T$, and using the fact that $AB = I$, the partitions of B are given with respect to the partitions of A in the form

$$\begin{aligned} A_M B_M + A_{NM}^T B_{NM} &= I_{MM} \\ A_M B_{NM}^T + A_{NM}^T B_N &= O_{MN} \\ A_{NM} B_{NM}^T + A_N B_N &= I_{NN} \\ A_{NM} B_M + A_N B_{NM} &= O_{NM} \end{aligned} \quad (\text{A.3})$$

Note that the covariance matrices A , A_M and A_N are by construction symmetric positive definite matrices which also imply that B and the partitions B_M and B_N are also symmetric positive definite matrices. The symmetry of the covariance matrices A_N and B_N has been used to simplify the expressions in (A.3). Solving the system of (A.3) with respect to the partitions of B , one readily derives

$$B_M = A_M^{-1} - A_M^{-1} A_{NM}^T B_{NM} \quad (\text{A.4})$$

$$B_{MN} = B_{NM}^T = -A_M^{-1} A_{NM}^T B_N \quad (\text{A.5})$$

$$B_N = [A_N - A_{NM} A_M^{-1} A_{NM}^T]^{-1} \quad (\text{A.6})$$

Substituting $A^{-1} = B$ in (A.1), noting that $(L_{M+N} \underline{\nabla}_{\theta} \underline{x}_k)^T = [(L_M \underline{\nabla}_{\theta} \underline{x}_k)^T \ (L_N \underline{\nabla}_{\theta} \underline{x}_k)^T]$ and expanding the right hand side using the partitions of B as obtained in (A.4)-(A.6), one readily derives that

$$Q(L_{M+N}) = \sum_{k=1}^N (L_M \underline{\nabla}_{\theta} \underline{x}_k)^T B_M (L_M \underline{\nabla}_{\theta} \underline{x}_k) + \delta Q_{NM}^* \quad (\text{A.7})$$

where δQ_{NM}^* is given by

$$\delta Q_{NM}^* = \sum_{k=1}^N [(L_N \underline{\nabla}_\theta \underline{x}_k)^T B_{NM} (L_M \underline{\nabla}_\theta \underline{x}_k) + (L_M \underline{\nabla}_\theta \underline{x}_k)^T B_{NM}^T (L_N \underline{\nabla}_\theta \underline{x}_k) + (L_N \underline{\nabla}_\theta \underline{x}_k)^T B_N (L_N \underline{\nabla}_\theta \underline{x}_k)] \quad (\text{A.8})$$

The representation (3) follows from (A.7) by substituting B_M from (A.4), expanding the first term in the right hand side and noting that $Q(L_M) = \sum_{k=1}^N (L_M \underline{\nabla}_\theta \underline{x}_k)^T A_M^{-1} (L_M \underline{\nabla}_\theta \underline{x}_k)$, while δQ_{NM} in (3) is given by the remaining terms as

$$\delta Q_{NM} = - \sum_{k=1}^N (L_M \underline{\nabla}_\theta \underline{x}_k)^T A_M^{-1} A_{NM}^T B_{NM} (L_M \underline{\nabla}_\theta \underline{x}_k) + \delta Q_{NM}^* \quad (\text{A.9})$$

Noting from (A.5) that $-A_M^{-1} A_{NM}^T = B_{NM}^T B_N^{-1}$ and substituting in (A.9) results in

$$\delta Q_{NM} = \sum_{k=1}^N (L_M \underline{\nabla}_\theta \underline{x}_k)^T B_{NM}^T B_N^{-1} B_{NM} (L_M \underline{\nabla}_\theta \underline{x}_k) + \delta Q_{NM}^* \quad (\text{A.10})$$

Introducing the auxiliary matrices $E_{M,k} = (L_M \underline{\nabla}_\theta \underline{x}_k)$ and $E_{N,k} = (L_N \underline{\nabla}_\theta \underline{x}_k)$ in (A.8) and (A.10), one finally simplifies the expression for δQ_{NM} in the form

$$\delta Q_{NM} = \sum_{k=1}^N [(B_{NM} E_{M,k})^T B_N^{-1} B_{NM} E_{M,k} + E_{N,k}^T B_{NM} E_{M,k} + (B_{NM} E_{M,k})^T E_{N,k} + E_{N,k}^T B_N E_{N,k}] \quad (\text{A.11})$$

Factoring out $B_N^{-1} B_{NM} E_{M,k}$ from the first two terms and $E_{N,k}$ from the last two terms on derives

$$\begin{aligned} \delta Q_{NM} &= \sum_{k=1}^N [(E_{M,k}^T B_{NM}^T + E_{N,k}^T B_N) B_N^{-1} B_{NM} E_{M,k} + (E_{M,k}^T B_{NM}^T + E_{N,k}^T B_N) E_{N,k}] \\ &= \sum_{k=1}^N [(B_{NM} E_{M,k} + B_N E_{N,k})^T B_N^{-1} (B_{NM} E_{M,k} + B_N E_{N,k})] \end{aligned} \quad (\text{A.12})$$

Note that the last expression is symmetric and semi-positive definite since for every non-zero vector $\underline{y} \in R^{N_\theta}$ ($\underline{y} \neq \underline{0}$) the quadratic form

$$\underline{y}^T \delta Q_{NM} \underline{y} = \sum_{k=1}^N [\underline{y}^T (B_{NM} E_{M,k} + B_N E_{N,k})^T B_N^{-1} (B_{NM} E_{M,k} + B_N E_{N,k}) \underline{y}] = \sum_{k=1}^N \underline{z}^T B_N^{-1} \underline{z} \geq 0 \quad (\text{A.13})$$

where $\underline{z} = (B_{NM} E_{M,k} + B_N E_{N,k}) \underline{y}$, is non-negative since the matrix B_N (and thus its inverse B_N^{-1}) is symmetric positive definite.

Appendix B: Proof of Proposition 3

The proof uses relationships and notations introduced in Appendix A. Using (3) and (A.12), replacing L_N by the sensor configuration L_1 of the new sensor, using the sensor configuration $L_{M+1} = (L_M^T \ L_1^T)^T$, and noting that A_1 and A_{1M} in (A.2) are given respectively by

$$A_1 = L_1 \Sigma L_1 = \sigma^2 + s^2 \quad (\text{B.1})$$

and

$$A_{1M} = L_1 \Sigma L_M = [0, \dots, 0, \sigma \hat{\sigma} R(\delta)] \quad (\text{B.2})$$

$\underbrace{\hspace{10em}}_{M-1}$

where σ^2 and $\hat{\sigma}^2$ are the variances of the model prediction errors at the new sensor location and the location of the M sensor in the configuration L_M , respectively, it can be readily shown that the information matrix $\hat{Q}(\delta)$ for a sensor configuration $L_{M+1} = (L_M^T \ L_1^T)^T$ takes the form:

$$\hat{Q}(\delta) = Q(L_M) + \sum_{k=1}^N [(B_{1M} E_{M,k} + B_1 E_{1,k})^T B_1^{-1} (B_{1M} E_{M,k} + B_1 E_{1,k})] \quad (\text{B.3})$$

Note that the right-hand-side of (B.2) is valid based on the assumption that the correlation length λ is small to the minimum distance Δ between sensors on configuration L_M , guaranteeing that the prediction errors between the new sensor (placed close to the M sensor) and the existing $M - 1$ sensors are uncorrelated. The matrix A_M is also diagonal due to the uncorrelated prediction errors between the sensors in the configuration L_M .

Using (A.5) for $N = 1$, the following relation holds $B_{1M}^T = -A_M^{-1} A_{NM}^T B_1$, where $B_1 \equiv B_1(\delta)$ is a scalar given by (A.6) in the form

$$B_1 = [A_1 - A_{1M} A_M^{-1} A_{1M}^T]^{-1} \quad (\text{B.4})$$

Noting that A_1 is given by (B.1), A_{1M} by (B.2) and that the M -th diagonal element of the inverse of A_M equals to $1/[\hat{\sigma}^2 + s^2]$, the scalar $B_1(\delta)$ admits the representation

$$B_1(\delta) = \left[\sigma^2 + s^2 - \frac{\sigma^2 \hat{\sigma}^2}{\hat{\sigma}^2 + s^2} R^2(\delta) \right]^{-1} > 0 \quad (\text{B.5})$$

Substituting $B_{1M}^T = -A_M^{-1} A_{NM}^T B_1$ from (A.5) into (B.3), noting that A_{1M} is given by (B.2), also that $A_{1M} E_{M,k} = \nabla_{\theta}^T x_{M,k}$ and $E_{1,k} = \nabla_{\theta}^T x_{N,k}$, where $x_{M,k}$ and $x_{N,k}$ are respectively the responses of the M sensor and the new $M + 1$ sensor, one readily derives that

$$\hat{Q}(\delta) = Q(L_M) + B_1(\delta) G_{1M}(\delta) \quad (\text{B.6})$$

where $G_{1M}(\delta)$ a semi-positive definite matrix given by

$$G_{1M}(\delta) = \sum_{k=1}^N \left(-\frac{\sigma \hat{\sigma}}{\hat{\sigma}^2 + s^2} R(\delta) \underline{\nabla}_{\theta} x_{M,k} + \underline{\nabla}_{\theta} x_{N,k} \right) \left(-\frac{\sigma \hat{\sigma}}{\hat{\sigma}^2 + s^2} R(\delta) \underline{\nabla}_{\theta}^T x_{M,k} + \underline{\nabla}_{\theta}^T x_{N,k} \right) \quad (\text{B.7})$$

Introducing the vector $\underline{v}_k(\delta) = \underline{\nabla}_{\theta} x_{N,k} - \underline{\nabla}_{\theta} x_{M,k}$, the matrix $G_{1M}(\delta)$ takes the form:

$$\begin{aligned} G_{1M}(\delta) &= \sum_{k=1}^N \left[\Gamma(\delta) \underline{\nabla}_{\theta} x_{M,k} + \underline{v}_k(\delta) \right] \left[\Gamma(\delta) \underline{\nabla}_{\theta}^T x_{M,k} + \underline{v}_k^T(\delta) \right] \\ &= \Gamma^2(\delta) \sum_{k=1}^N \underline{\nabla}_{\theta} x_{M,k} \underline{\nabla}_{\theta}^T x_{M,k} + \Gamma(\delta) \sum_{k=1}^N \left[\underline{\nabla}_{\theta} x_{M,k} \underline{v}_k^T(\delta) + \underline{v}_k(\delta) \underline{\nabla}_{\theta}^T x_{M,k} \right] + \sum_{k=1}^N \underline{v}_k(\delta) \underline{v}_k^T(\delta) \end{aligned} \quad (\text{B.8})$$

where $\Gamma(\delta)$ is given by

$$\Gamma(\delta) = 1 - \frac{\sigma \hat{\sigma}}{\hat{\sigma}^2 + s^2} R(\delta) > 0 \quad (\text{B.9})$$

Note that the sum in the first term is independent of δ . Also the elements of the response sensitivity vector $\underline{\nabla}_{\theta} x_{M,k}$ are expected to be large since these response sensitivities correspond to a sensor location in the configuration L_M and the methodology selects the locations with the highest response sensitivities. For sufficiently small δ compared to the characteristic length of the dynamic problem, the vector $\underline{\nabla}_{\theta} x_{N,k}$ at the new sensor location does not vary significantly from the vector $\underline{\nabla}_{\theta} x_{M,k}$ in the neighborhood sensor location. In this case, the vector $\underline{v}_k(\delta)$ becomes sufficiently small and the first term in (B.8) dominates the other two terms.

Using the fact that a correlation function $R(\delta)$ attains the maximum at $\delta = 0$, for sufficiently small δ_1 and δ_2 with $\delta_1 > \delta_2$ the following inequality $R(\delta_2) > R(\delta_1)$ holds. Using (B.5) and (B.9), it is straightforward to show that for $\delta_1 > \delta_2$, the following inequality holds $\Gamma^2(\delta_1) / B_1(\delta_1) > \Gamma^2(\delta_2) / B_1(\delta_2)$. The last expression along with the fact that the first term dominates the right hand side of (B.8) for sufficiently small distances δ_1 and δ_2 determined by the characteristic length of the problem, results in $B_1(\delta_1)G_{1M}(\delta_1) - B_2(\delta_2)G_{1M}(\delta_2)$ to be a positive definite matrix. Applying (B.6) for δ_1 , noting that

$$\begin{aligned} \hat{Q}(\delta_1) &= Q(L_M) + B_1(\delta_2)G_{1M}(\delta_2) + [B_1(\delta_1)G_{1M}(\delta_1) - B_1(\delta_2)G_{1M}(\delta_2)] \\ &= \hat{Q}(\delta_2) + [B_1(\delta_1)G_{1M}(\delta_1) - B_1(\delta_2)G_{1M}(\delta_2)] \end{aligned} \quad (\text{B.10})$$

the validity of (8) follows immediately from the relation (6) with $A_0 \equiv \hat{Q}(\delta_2)$ and $B_0 \equiv [B_1(\delta_1)G_{1M}(\delta_1) - B_1(\delta_2)G_{1M}(\delta_2)]$.

It should be noted that for the nearly uncorrelated prediction error case, the correlation length λ is significantly smaller than the characteristic length of the problem, and the factor $\Gamma^2(\delta) / B_1(\delta)$ tends to $(\sigma^2 + s^2)^{-1}$ independent of δ , provided that δ is large

compared to the correlation length. In this case the dominant term in (B.8) is independent of δ within the characteristic length of the problem. Thus the decrease or increase of the information entropy does not depend on the first term and it is controlled on the second and third terms.

References

- [1] N. Bleistein, R. Handelsman, *Asymptotic Expansions for Integrals*, Dover Publications, New York, 1986.
- [2] C. Papadimitriou, Optimal Sensor Placement Methodology for Parametric Identification of Structural Systems. *Journal of Sound and Vibration* 278(4) (2004) 923-947.
- [3] C. Papadimitriou, J.L. Beck, S.K. Au, Entropy-based optimal sensor location for structural model updating. *Journal of Vibration and Control* 6(5) (2000) 781-800.
- [4] R.L. Fox, M.P. Kapoor, Rates of Changes of Eigenvalues and Eigenvectors, *AIAA Journal* 6 (1968) 2426-2429.
- [5] R.B. Nelson, Simplified calculation of eigenvector derivatives. *AIAA Journal* 14(9) (1976) 1201-1205.
- [6] S. Adhikari, M.I. Friswell, Eigenderivative analysis of asymmetric non-conservative systems. *International Journal for Numerical Methods in Engineering* 51(6) (2001) 709-733.

References for Optimal Sensor Placement Methods

- [1] P. Shah, F.E. Udawadia, A methodology for optimal sensor locations for identification of dynamic systems. *Journal of Applied Mechanics* 45 (1978) 188-196.
- [2] K. Sobczyk, Theoretic information approach to identification and signal processing. In: *Proc. of the IFIP Conference on Reliability and Optimisation of Structural Systems*, 1987.
- [3] D.C. Kammer, Sensor placements for on orbit modal identification and correlation of large space structures. *Journal of Guidance, Control and Dynamics* 14 (1991) 251-259.
- [4] D.C. Kammer, Effects of noise on sensor placement for on-orbit modal identification of large space structures, *Journal of dynamic systems, measurements and control - Transactions of the ASCE* (1992) 114 (3) 436-443.
- [5] P.H. Kirkegaard, R. Brincker, On the optimal locations of sensors for parametric identification of linear structural systems. *Mechanical Systems and Signal Processing* 8 (1994) 639-647.
- [6] F.E. Udawadia, Methodology for optimal sensor locations for parameter identification in dynamic systems. *Journal of Engineering Mechanics (ASCE)*, 120(2) (1994) 368-390.
- [7] D.S. Li, H.N. Li, C.P. Fritzen, The connection between effective independence and modal kinetic energy methods for sensor placement, *Journal of Sound and Vibration* 305 (2007) 945-955.
- [8] H.M. Chow, H.F. Lam, T. Yin, S.K. Au, Optimal sensor configuration of typical transmission tower for the purpose of structural model updating, *Structural Control and Health Monitoring* (2010) DOI: 10.1002/stc.372.
- [9] Z.H. Qureshi, T.S. Ng, G.C. Goodwin, Optimum experimental design for identification of distributed parameter systems, *International Journal of Control* 31 (1980) 21-29.
- [10] E. Heredia-Zavoni, R. Montes-Iturrizaga, L. Esteva, Optimal instrumentation of structures on flexible base for system identification. *Earthquake Engineering and Structural Dynamics* 28(12) (1999) 1471-1482.

- [11] E. Heredia-Zavoni, L. Esteva, Optimal instrumentation of uncertain structural systems subject to earthquake motions. *Earthquake Engineering and Structural Dynamics* 27(4) (1998) 343-362.
- [12] E.B. Flynn, M.D. Todd, A Bayesian approach to optimal sensor placement for structural health monitoring with application to active sensing. *Mechanical Systems and Signal Processing* 24 (2010) 891-903.
- [13] M. Azarbajegani, A.I. Eli-Osery, K.K. Choi, M.M. Reda Taha, A probabilistic approach for optimal sensor allocation in structural health monitoring. *Smart Materials and Structures* 17 (2008) doi: 10.1088/0964-1726/17/5/055019.
- [14] C. Papadimitriou, J.L. Beck, S.K. Au, Entropy-based optimal sensor location for structural model updating. *Journal of Vibration and Control* 6(5) (2000) 781-800.
- [15] C. Papadimitriou, Optimal Sensor Placement Methodology for Parametric Identification of Structural Systems. *Journal of Sound and Vibration* 278(4) (2004) 923-947.
- [16] K.V. Yuen, L.S. Katafygiotis, C. Papadimitriou, N.C. Mickleborough, Optimal sensor placement methodology for identification with unmeasured excitation. *Journal of Dynamic Systems, Measurement and Control* 123(4) (2001) 677-686.
- [17] C. Papadimitriou, Pareto Optimal Sensor Locations for Structural Identification. *Computer Methods in Applied Mechanics and Engineering* 194(12-16) (2005) 1655-1673.
- [18] P. Metallidis, G. Verros, S. Natsiavas, C. Papadimitriou, Identification, fault detection and optimal sensor location in vehicle suspensions. *Journal of Vibration and Control* 9 (2003) 9(3-4) 337-359.
- [19] D.C. Kammer and L. Yao, Enhancement of on-orbit modal identification of large space structures through sensor placement. *Journal of Sound and Vibration* 171(1) (1994) 119-139.
- [20] D.S. Li, H.N. Li, C.P. Fritzen, A note on fast computation of effective independence through QR downdating for sensor placement, *Mechanical Systems and Signal Processing* 23 (2009) 1160-1168.
- [21] H. Bedrossian, S.F. Masri, Optimal placement of sensors and shakers for modal identification, in: P.D. Spanos and G. Deodatis (Eds.), *Computational Stochastic Mechanics*, Millpress, Rotterdam, 2003, pp. 53-57.
- [22] N.M. Abdullah, A. Richardson, J. Hanif, Placement of Sensors/Actuators on Civil Structures Using Genetic Algorithms. *Earthquake Engineering and Structural Dynamics* 30(8) (2001) 1167-1184.
- [23] L. Yao, W.A. Sethares, D.C. Kammer, Sensor placement for on orbit modal identification via a genetic algorithm. *AIAA Journal* 31 (1993) 1167-1169.
- [24] C. Papadimitriou, Applications of Genetic Algorithms in Structural Health Monitoring. Proc. 5th World Congress on Computational Mechanics (<http://wccm.tuwien.ac.at>), Vienna, Austria), 2002.
- [25] J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties – I: Bayesian statistical framework. *Journal of Engineering Mechanics (ASCE)* 124(4) (1998) 455-461.
- [26] L.S. Katafygiotis, C. Papadimitriou, H.F. Lam, A probabilistic approach to structural model updating. *International Journal of Soil Dynamics and Earthquake Engineering* 17(7-8) (1998) 495-507.
- [27] M.I. Friswell, J.E.T. Penny, S.D. Garvey, Parameter subset selection in damage location. *Inverse Problems in Engineering* 5 (1997) 189-215.