The effect of prediction error correlation on optimal sensor placement in structural dynamics

Costas Papadimitriou*, Geert Lombaert

K.U. Leuven, Department of Civil Engineering, Kasteelpark Arenberg 40, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:
Received 29 December 2010
Received in revised form
23 May 2011
Accepted 28 May 2011

Keywords:
Optimal sensor placement
Vibration testing
Information entropy
Parameter estimation
Modal identification
Model updating

ABSTRACT

The problem of estimating the optimal sensor locations for parameter estimation in structural dynamics is re-visited. The effect of spatially correlated prediction errors on the optimal sensor placement is investigated. The information entropy is used as a performance measure of the sensor configuration. The optimal sensor location is formulated as an optimization problem involving discrete-valued variables, which is solved using computationally efficient sequential sensor placement algorithms. Asymptotic estimates for the information entropy are used to develop useful properties that provide insight into the dependence of the information entropy on the number and location of sensors. A theoretical analysis shows that the spatial correlation length of the prediction errors controls the minimum distance between the sensors and should be taken into account when designing optimal sensor locations with potential sensor distances up to the order of the characteristic length of the dynamic problem considered. Implementation issues for modal identification and structural-related model parameter estimation are addressed. Theoretical and computational developments are illustrated by designing the optimal sensor configurations for a continuous beam model, a discrete chain-like stiffness–mass model and a finite element model of a footbridge in Wetteren (Belgium). Results point out the crucial effect the spatial correlation of the prediction errors have on the design of optimal sensor locations for structural dynamics applications, revealing simultaneously potential inadequacies of spatially uncorrelated prediction errors models.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The problem of parameter estimation of structural models using measured dynamic data is important in modal identification, structural model updating, structural health monitoring and structural control. The estimate of the parameter values involves uncertainties that are due to limitations of the mathematical models used to represent the behavior of the real structure, the presence of measurement error in the data and insufficient excitation and response bandwidth. In particular, the quality of information that can be extracted from the data for estimating the model parameters depends on the number and location of sensors in the structure as well as on the type and size of model and measurement error. The objective in an experimental design is to make a cost-effective selection of the optimal number

* Corresponding author. Permanent address: University of Thessaly, Department of Mechanical Engineering, Volos 38334, Greece.
Tel.: +30 2421074006; fax: +30 24210 74012.
E-mail address: costasp@uth.gr (C. Papadimitriou).

0888-3270/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ymssp.2011.05.019

and location of sensors such that the resulting measured data are most informative for estimating the parameters of a mathematical model of the structure.

Information theory based approaches [1–8] have been developed to provide rational solutions to several issues encountered in the problem of selecting the optimal sensor configuration for modal identification and structural parameter estimation. The optimal sensor configuration is selected as the one that maximizes some norm (determinant or trace) of the Fisher information matrix (FIM) [9]. In other studies [10,11], the optimal sensor configuration has been chosen as the one that minimizes the expected Bayesian loss function involving the trace of the inverse of the FIM. A Bayesian framework to optimal sensor location for structural health monitoring (SHM) has also been introduced in [12]. The optimal configuration is chosen to optimize (maximize or minimize) a Bayesian risk-based performance metric related to the probability of damage detection or false alarm of all regions of the structure. A probabilistic approach to optimal sensor location for SHM was proposed in [13], utilizing a priori knowledge of probable damage locations and severities in the structure. The weights of an artificial neural network, trained to detect damage, were used to generate a probability distribution function that is sampled to determine the optimal sensor locations.

The information entropy, measuring the uncertainty in the model parameter estimates, was also introduced [14] for designing optimal sensor configurations. It was shown [15] that, asymptotically for very large number of data, the information entropy depends on the determinant of the FIM, justifying the use of the determinant instead of the trace or other scalar measures of FIM in previous approaches. The information entropy has been applied to design optimal sensor locations for parameter estimation using ambient vibrations [16], model class selection [17] for damage detection, as well as to design the optimal excitation characteristics (e.g. amplitude and frequency content) for the identification of linear and strongly nonlinear models [18].

The optimal sensor location problem is formulated as a single-objective optimization problem involving discrete-valued variables. Computational efficient algorithms for solving the discrete-valued minimization problem have been proposed. Udwadia [6] demonstrated that using the trace of the FIM is computationally very attractive since the solution of the underlying discrete optimization problem is straightforward. However, for other more popular scalar measures of uncertainties such as the determinant of the FIM or the information entropy, an exhaustive search over all possible sensor configurations is required to obtain the exact optimal sensor configuration. This approach is computationally prohibitive even for structures with a relatively small number of degrees of freedom (DOF). Heuristic optimization tools have also been developed as efficient alternatives for efficiently solving these discrete optimization problems involving discrete-valued variables. In the modal identification case, efficient iterative algorithms [19,20] were proposed for sensor placement using the effective independence method. Exploiting theoretical asymptotic results on information entropy and FIM, two computationally efficient heuristic algorithms, the forward sequential sensor placement (FSSP) and the backward sequential sensor placement (BSSP) [15,17] were proposed. These algorithms construct sensor configurations for physical model parameter estimation, corresponding to information entropy values very close to lower or upper bounds of the information entropy. The numerical results indicated that the proposed heuristic algorithms provide sub-optimal sensor configurations that can be extremely good approximations of the optimal sensor configuration [15,17]. Moreover, these heuristic algorithms are very simple to implement in software and computationally very efficient. Alternatively, Genetic Algorithms [21–24] are intelligent techniques that will yield optimal solutions and in this sense can be used whenever it is deemed necessary to complement the heuristic algorithms in an effort to improve estimates.

In most information theory-based methods the effect of spatially correlated prediction errors and its importance was not adequately explored. The present study provides insight into the effect of spatially correlated prediction errors on the design of the optimal sensor locations. The information entropy is used as the performance measure of a sensor configuration. The information entropy is built from the parameter uncertainty identified by applying a Bayesian identification framework. The optimal sensor location problem is formulated as a single-objective optimization problem involving discrete-valued variables. The effectiveness of available heuristic algorithms BSSP and FSSP, known to be computationally very efficient and accurate for uncorrelated prediction errors, is explored.

Asymptotic approximations, valid for large number of data, available for the information entropy [15] for the case of uncorrelated prediction errors are extended to account for the case of spatially correlated prediction errors. Useful theoretical results are derived that show that the lower and upper bounds of the asymptotic estimate of the information entropy, corresponding, respectively, to the optimal and worst sensor configuration, are a decreasing function of the number of sensors. In addition, it is shown that for up to the characteristic length of the highest contributing mode of the structure, the spatial correlation between prediction errors forces the minimum distance between sensors to be of the order of the prediction error correlation length. Consequently, sensor placement becomes independent of the mesh size of the finite element models used for structural dynamics simulations. For distances between two sensors higher than the characteristic length, the sensor locations are affected also by the spatial variability of the response sensitivities computed by a nominal structural model. Implementation in structural dynamics is concentrated on the design of optimal sensor location for (a) modal identification and (b) estimation of structural model (e.g. finite element) parameters. Theoretical and computational developments are demonstrated by designing the optimal sensor configurations for a simply supported continuous beam model, a discrete chain-like stiffness–mass model and a finite element model of a footbridge in Wetteren (Belgium). It is illustrated that the extent of the spatial correlation of the prediction errors has an important effect on the optimal sensor locations. In addition, inadequacies of the spatially uncorrelated prediction error models are emphasized.
2. Estimation of model parameter uncertainty

2.1. Bayesian statistical framework

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 $\vec{\vartheta} \in \mathbb{R}^N$ 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 $D = \{y_k, k = 1, \ldots, N\}$ be the measured sampled response time histories data, where $y_k \in \mathbb{R}^N$ 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 $k\Delta t$, $\Delta t$ is the sampling interval and $N$ is the number of sampled data. Let $x_k(\vec{\vartheta}) \in \mathbb{R}^{N_L}$, $k = 1, \ldots, N$ be the sampled response time histories computed at all $N_L$ model DOFs from a structural model that corresponds to a specific value $\vec{\vartheta}$ of the model parameters. The measured response and model response predictions at time instant $k\Delta t$ satisfy the prediction error equation:

$$ y_k = Lx_k(\vec{\vartheta}) + L\epsilon_k(\vec{\vartheta}) $$

where $\epsilon_k(\vec{\vartheta})$ is the prediction error due to modelling error and measurement noise. The matrix $L \in \mathbb{R}^{N_0 \times N_L}$ 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 [25,26], the uncertainties in the values of the parameters $\vartheta$ are quantified by probability density functions (PDF) that are obtained using the dynamic test data $D$ and the probability model for the prediction error $\epsilon_k(\vec{\vartheta})$. In what follows, the prediction error vector $\epsilon_k(\vec{\vartheta})$ at time $k\Delta t$ is modeled as a Gaussian random vector with zero mean and covariance $\Sigma_i \in \mathbb{R}^{N_L \times N_L}$. Also, it is assumed that the prediction errors between different time instances are independent. Applying Bayes’ theorem, the updating PDF $p(\vec{\vartheta} | \Sigma_i, D)$ of the set of structural model parameters $\vec{\vartheta}$ given the measured data $D$ and the prediction error parameters $\Sigma_i$ takes the form:

$$ p(\vec{\vartheta} | \Sigma_i, D) = \frac{1}{\sqrt{(2\pi)^N \sqrt{\det \Sigma_i}} \exp \left[ -\frac{NN_0}{2} J(\vec{\vartheta}; \Sigma_i, D) \right] \pi(\vec{\vartheta}) } $$

where

$$ J(\vec{\vartheta}; \Sigma_i, D) = \sum_{k=1}^{N} \frac{1}{NN_0} [y_k - Lx_k(\vec{\vartheta})]^T \Sigma_i^{-1} [y_k - Lx_k(\vec{\vartheta})] $$

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

2.2. Prediction error correlation model

An analysis of the prediction error correlation models is presented next. The prediction error $\epsilon_k = \epsilon_{k,\text{meas}} + \epsilon_{k,\text{mod}}$ in (1) is due to a term, $\epsilon_{k,\text{meas}}$, accounting for the measurement error and a term, $\epsilon_{k,\text{mod}}$, accounting for the model error. Assuming independence between the measurement error and model error, the covariance $\Sigma_i$ of the total prediction error is given in the form:

$$ \Sigma_i = \Sigma + \Sigma $$

where $\Sigma$ and $\Sigma$ are the covariance matrices of the measurement and model errors, respectively.

The designer has to assign values for the individual covariance matrices in (4). Such assumptions may depend on the nature of the problem analyzed. One reasonable choice is to assume that the measurement error is independent of the location of sensors so that the covariance $\Sigma$ takes the diagonal form $\Sigma = \sigma^2 I$, where $I$ is the identity matrix. However, a certain degree of correlation should be expected for the model errors between two neighborhood locations arising from the underlining model dynamics. This correlation can be taken into account by selecting a non-diagonal covariance matrix $\Sigma$.

The correlation $\Sigma_{ij}$ between the predictions $\epsilon_{i,\text{mod}}$ and $\epsilon_{j,\text{mod}}$ at DOFs $i$ and $j$, respectively, is given by

$$ \Sigma_{ij} = E[\epsilon_{i,\text{mod}} \epsilon_{j,\text{mod}}] = \sqrt{\Sigma_{ii} \Sigma_{jj}} R(\delta_{ij}) $$

that accounts for the spatial distance $\delta_{ij}$ between the DOFs $i$ and $j$, where $R(\delta_{ij})$ is a correlation function satisfying $R(0) = 1$. In general, the covariance matrix should be consistent with the actual errors and correlations as observed from measurements. However, in an experimental design stage such measurements are not available to guide the selection of the correlation between prediction errors. Instead a correlation function should be postulated to proceed with the design of the optimal sensor locations. Several correlation functions can be explored. For demonstration purposes in this study, the following exponential correlation function is assumed:

$$ R(\delta) = \exp[-\delta/\lambda] $$

where $\lambda$ is a measure of the spatial correlation length. However, the formulation in this work is general and does not depend on the choice of the correlation model.

### 3. Optimal sensor location methodology

#### 3.1. Asymptotic estimate of the information entropy

The updated PDF $p(\tilde{\theta} | \Sigma; D)$ given in (2) represents the uncertainty in the structural model parameter values based on the information contained in the measured data. The information entropy, defined by

$$h(L; \Sigma; D) = E_0[-\ln p(\tilde{\theta} | \Sigma; D)] = -\int \ln p(\tilde{\theta} | \Sigma; D)p(\tilde{\theta} | \Sigma; D)d\tilde{\theta}$$

(7)

has been introduced [14] to provide a unique scalar measure of the uncertainty in the estimate of the structural parameters $\tilde{\theta}$. The information entropy depends on the available data $D=D(L)$ and the sensor configuration vector $L$.

An asymptotic approximation of the information entropy, valid for large number of data ($NN_0 \to \infty$), is available [15] which is useful in the experimental stage of designing an optimal sensor configuration. The asymptotic approximation is obtained by substituting $p(\tilde{\theta} | \Sigma; D)$ from (2) into (7) 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 [27]. Specifically, it can be shown that for a large number of measured data, i.e. as $NN_0 \to \infty$, the following asymptotic result holds for the information entropy [15]:

$$h(L; \Sigma; D) \approx H(L; \bar{\theta}_0, \Sigma) = \frac{1}{2} NN_0 \ln(2\pi) - \frac{1}{2} \ln |\det Q(L; \bar{\theta}_0, \Sigma)|$$

(8)

where $\bar{\theta}_0 = \bar{\theta}(L; \Sigma; D)$ is the optimal value of the parameter set $\bar{\theta}$ that minimizes the measure of fit $J(\bar{\theta}; L; D)$ given in (3), and $Q(L; \bar{\theta}; \Sigma)$ is an $N_0 \times N_0$ semi-positive definite matrix defined as $NN_0 \sum_{k=1}^{N} J(\bar{\theta}_k; \Sigma; D)$ and asymptotically approximated by

$$Q(L; \bar{\theta}; \Sigma) = \sum_{k=1}^{N} (L\nabla \bar{\theta}_k)^T (L\Sigma L^T)^{-1} (L\nabla \bar{\theta}_k)$$

(9)

in which $\nabla \bar{\theta} = [\partial J/\partial \theta_1, \ldots, \partial J/\partial \theta_{N_0}]^T$ is the usual gradient vector with respect to the parameter set $\bar{\theta}$. For convenience the following notation was introduced $\nabla \bar{\theta}_k = \nabla \bar{\theta}(\bar{\theta}_k)$. The matrix $Q(L; \bar{\theta}; \Sigma)$ is a semi-positive definite matrix, known as the Fisher information matrix (FIM), containing the information about the uncertainty in the parameters $\bar{\theta}$ 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 [15]. Note that the asymptotic estimate of the matrix in (9) can readily be obtained by following the same steps as the ones presented in [15] for the uncorrelated case.

In the initial stage of designing the experiment, the data and consequently the values of the optimal model parameters $\bar{\theta}$ and the form of the prediction error covariance matrix $\Sigma$ are not available. In practice, useful designs can be obtained by taking the optimal model parameters $\bar{\theta}$ and prediction error covariance $\Sigma$ to have some nominal values $\bar{\theta}_0$ and $\Sigma$ to arise from a correlation function such as (6), chosen by the designer to be representative of the system and the expected model and measurement errors.

#### 3.2. Formulation as a discrete-valued optimization problem

In experimental design, the sensor 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 [14]. That is

$$L_{\text{best}} = \arg\min_L H(L; \bar{\theta}_0, \Sigma)$$

(10)

where $H(L; \bar{\theta}_0, \Sigma)$ is given by (8) and the minimization is constrained over the set of $N_p$ measurable DOFs. The lower bound of the information entropy is then given by $H_{\text{min}} = H(L_{\text{best}}; \bar{\theta}_0, \Sigma)$.

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

$$L_{\text{worst}} = \arg\max_L H(L; \bar{\theta}_0, \Sigma)$$

(11)

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

---

3.3. Computational algorithms

Two heuristic sequential sensor placement (SSP) algorithms, the forward (FSSP) and the backward (BSSP), were proposed \cite{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 \( N_p=N+1 \)
      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**

The aforementioned algorithms closely resemble the algorithms used to address a qualitatively different problem related to the selection of a suitable subset of parameters for estimation in structural dynamics and structural health monitoring \cite{28–30} based on subset selection techniques introduced in \cite{31}.

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 \cite{15,17}. However, when necessary, GAs can improve the SSP estimates, converging to the optimal solutions.

4. Analysis of information entropy

4.1. 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 \( L_{M+N}=(L_M^T L_N^T) \) 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

\[
H(L_{M+N};\bar{\Theta}_0;\Sigma) \leq H(L_M;\bar{\Theta}_0;\Sigma)
\]

(12)

The proof of Proposition 1 was presented for the special case of uncorrelated prediction errors in [15]. The proof of Proposition 1 for the general case of correlated prediction errors is more involved and is presented next.

**Proof.** Using (8), it suffices to show that the following inequality holds for two sensor configurations \( L_{M+N} \) and \( L_M \):

\[
\det(Q(L_{M+N})) = \det(Q(L_M)) + \delta Q_{MN} \geq 0
\]

where the dependence of \( Q(L; \theta, \Sigma) \) on \( \theta \) and \( \Sigma \) 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

\[
Q(L_{M+N}) = Q(L_M) + Q_{MN}, \quad \delta Q_{MN} = 0
\]

where the notation \( \delta Q_{MN} \geq 0 \) denotes that the matrix \( \delta Q_{MN} \) is a symmetric semi-positive definite matrix. The proof of statement (i) follows by exploiting the special form (9) of the matrix \( Q(L) \). It can be readily shown that for every non-zero vector \( y \in \mathbb{R}^N \), the quadratic form:

\[
y^T Q(L) y = \sum_{k=1}^N (L y_k y_k^T) (L y_k y_k^T)^{-1} (L y_k y_k^T) = \sum_{k=1}^N z^T y_k y_k^T z \geq 0
\]

where \( z = \sum_{k=1}^N \mathbf{y}_k \), is always non-negative since the covariance matrix \( L L^T \) is by construction symmetric positive definite. Thus, the matrix \( Q(L) \) is symmetric semi-positive definite. The proof of statement (ii) is given in Appendix A. Given the representation (14) and the fact that \( \delta Q_{MN} \) is semi-positive definite, the proof of Proposition 1 follows the same procedure as presented in [15]. For completeness, the main steps of the proof are given next. Substituting (14) into (13), it remains to show the validity of the inequality

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

This statement can be shown using the fact that for two symmetric semi-positive definite matrices \( A_0 \in \mathbb{R}^{N_0 \times N_0} \) and \( B_0 \in \mathbb{R}^{N_0 \times N_0} \) the following is true:

\[
\lambda_r[A_0 + B_0] \geq \lambda_r[A_0] \geq 0, \quad r = 1, \ldots, N_0
\]

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 (17) for \( A_0 = Q(L_M), B_0 = \delta Q_{MN} \) and using the fact that \( \det A_0 = \Pi_{r=1}^{N_0} \lambda_r[A_0] \) for any matrix \( A_0 \), the inequality (16) is readily derived. \( \square \)

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 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. Proposition 2 follows directly from Proposition 1, independent of the correlation model assumed for the prediction error. Thus, the reader is referred to Ref. [15] for a proof.

Propositions 1 and 2, shown in this work to be valid for spatially correlated prediction error, were employed in [15] to justify the use of the heuristic algorithms FSSP and BSSP for efficiently constructing sub-optimal solutions to the optimal sensor location problem.

### 4.2 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 \( D \) be the minimum distance between any two sensors in the sensor configuration \( L_M \). It is assumed that the correlation length \( \lambda \) of the prediction error is small enough compared to the minimum distance \( D \) 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 \( \delta \) 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 \( L_1 \in \mathbb{R}^{1 \times N_1} \). Consider that \( \delta \) varies from values of zero up to the order of the correlation length \( \lambda \) so that, using the fact that \( \lambda \) is small compared to \( D \), 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 \( H_\delta(L_M+1;\bar{\theta}_0,\Sigma) \) denote the information entropy as a function of the distance \( \delta \) for \( M+1 \) sensors located according to the sensor configuration \( L_{M+1} = (L_M, 1)^T \). Similarly let also \( Q_\delta(L_M+1;\bar{\theta}_0,\Sigma) \) 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 \( \delta \) of the new sensor from the other \( M \) sensors, provided that \( \delta \) is sufficiently small. Mathematically, this proposition reads:

\[
H(\delta_1) < H(\delta_2) \quad \text{for any} \quad \delta_1 > \delta_2
\]

or, equivalently, using (8)

\[
\det Q(\delta_1) > \det Q(\delta_2) \quad \text{for any} \quad \delta_1 > \delta_2
\]

where \( \delta_1 \) and \( \delta_2 \) are sufficiently small.

The proof of Proposition 3 is presented in Appendix B and shows that the proposition holds for distances \( \delta \) 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 (18) or (19) 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.

### 5. Implementation 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

\[
\ddot{u}(t) + C\ddot{u}(t) + Ku(t) = \Gamma z(t)
\]

where \( M \in \mathbb{R}^{N_x \times N_x} \), \( C \in \mathbb{R}^{N_x \times N_x} \) and \( K \in \mathbb{R}^{N_x \times N_x} \) are the mass, damping and stiffness matrices, respectively, \( \bar{u}(t) \in \mathbb{R}^{N_u} \) is the displacement response vector, \( z(t) \in \mathbb{R}^{N_z} \) is the vector of independent input forces and \( \Gamma \in \mathbb{R}^{N_z \times N_u} \) 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 \( \bar{X}_k \in \mathbb{R}^{N_z} \) given, respectively, by either \( \bar{u}_k \), or \( \dot{u}_k \) or \( \ddot{u}_k \). 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.

#### 5.1. Design of optimal sensor locations for modal identification

The first category deals with the estimation of the modal coordinate vector \( \bar{\xi} \in \mathbb{R}^m \) 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 \( \bar{\xi} \). In this case, the parameter set \( \bar{\theta} = \bar{\xi} \). Following the conventional modal analysis, the response vector \( \bar{X} \in \mathbb{R}^{N_z} \) is given with respect to the parameter set \( \bar{\theta} = \bar{\xi} \) in the form \( \bar{X} = \Phi \bar{\xi} \), where \( \Phi \in \mathbb{R}^{N_z \times m} \) is the mode shape matrix for \( m \) contributing modes. Noting that \( \bar{X} = \Phi \bar{\xi} \) and substituting into (9), the information matrix takes the form:

\[
Q(L, \bar{\xi}, \Sigma) = (L\Phi)^T(L\Sigma L^T)^{-1}(L\Phi)
\]

which is independent of the nominal parameter values \( \bar{\theta}_0 \). In addition, the optimal sensor locations are independent of the excitation used. The information matrix \( Q(L, \Sigma) \) in (21) has exactly the same form as the one proposed [3,19] for designing the optimal sensor location. So, with the same assumed prediction error correlation matrix \( \Sigma \), the optimal sensor location from the proposed methodology and the effective independence are based on exactly the same information matrix \( Q \) given in (21).

Based on the form of (21), a non-singular FIM matrix \( Q(L, \Sigma) \) 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, \( N_0 (N_0 = m) \). Otherwise, for \( N_0 < m \), the matrix \( Q(L, \Sigma) \) in (21) 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. 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 for a small number of sensors, \( N_0 < m \), and will considerably affect the optimal sensor location for \( N_0 \geq m \). One way to optimally place sensors in the structure for \( N_0 < m \) is to maximize the product of the
The second category deals with the estimation of structural-related properties. In this case, the parameter set \( \theta \) includes variables related to stiffness, mass and damping characteristics. The system mass \( M(\theta) \), stiffness \( K(\theta) \) and damping \( C(\theta) \) matrices depend on the parameter set \( \theta \). To compute the response sensitivity matrix \( \nabla \Sigma_k \) needed in (9), a sensitivity analysis needs to be performed. Following a conventional modal analysis, writing the response vector in the form \( \dot{y}(t) = \Phi \ddot{z}(t) \), where the modal coordinates \( \ddot{z}(t) \) satisfy the uncoupled system of modal equations:

\[
\ddot{z}_i + C \ddot{z}_i + A \ddot{z}_i = \Phi^T \Gamma \ddot{z}
\]

and differentiating the equations with respect to the \( i \)th component \( \theta_i \) of the parameter set \( \theta \), the following analytical differential equations for the components \( \mu_i(t) = \partial \ddot{z}(t)/\partial \theta_i \), \( i = 1, \ldots, N_\theta \) of the sensitivity matrix \( \Sigma_{\theta \theta} = [\mu_1(t), \mu_2(t), \ldots, \mu_{N_\theta}(t)] \) are readily obtained in the form:

\[
\ddot{\mu}_i(t) = \Phi_{i} \mu_i(t) + \frac{\partial \Phi_{i}}{\partial \theta_i} \ddot{z} + \frac{\partial \Phi^T \Gamma}{\partial \theta_i} \dot{z}(t)
\]

where \( \mu_i(t) \) is computed from the modal equations (22) to satisfy the following system of equations:

\[
\ddot{\mu}_i + C \ddot{\mu}_i + A \ddot{\mu}_i = -\Gamma \ddot{z} + \frac{\partial C \ddot{z}}{\partial \theta_i} - \frac{\partial A \ddot{z}}{\partial \theta_i} + \frac{\partial \Phi_j \phi}{\partial \theta_i} \dot{z}(t)
\]

The matrix \( A \) in (22) and (24) is diagonal with elements \( \omega_i^2 \), while assuming classically damped modes the matrix \( C \) is also diagonal with elements \( 2 \omega_i \xi_i \), where \( \omega_i \) and \( \xi_i \) 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 \( A \) and the eigenvectors in \( \Phi \) are readily obtained from the sensitivities of the mass and stiffness matrices using well established techniques [32,33]. For more information on sensitivity of eigenvalues and eigenvectors for complex eigen systems, the reader is referred to the overview paper in [34].

The optimal sensor locations depend on the location and type of excitations used. Also, in contrast to the modal identification case in Section 5.1, the matrix \( Q(L, \mu, \Sigma) \) 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 \( \theta \). 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.

### 6. Applications

#### 6.1. 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 (21) are given by \( \sin(m \pi x/h) \) 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 \( d \) is given by (4)-(6) with \( \sigma^2 = 0 \) and equal variance values \( \Sigma_{ii} \) for all \( i \).

The contour plots of \( \det Q(L, \mu_0, \Sigma) \) as a function of the locations \( x/h \) of the two sensors along the beam axis are presented in Fig. 1. Representative values of the correlation lengths \( \lambda = 0, 0.1, 0.2 \) and 0.4 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 Fig. 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 Fig. 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 $0.4$, only two global solutions remain (positions with coordinates $(0.167, 0.5)$ and $(0.5, 0.833)$) as it is seen in Fig. 1(c) and (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 $x/h = 0.833$ becomes 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.

### 6.2. 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...
$k_i/m_i=1/|s^2|$. 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 $k_i$, $i=1,\ldots,5$ of the five substructures in the system. The time history of the response is sampled with a time step $\Delta t$ equal to $T_{\text{min}}/6$, with $T_{\text{min}}$ 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

$$\text{IEI}(L) = \exp[H(L; \hat{\Omega},\Sigma) - H(L_{\text{ref}}; \hat{\Omega}_{\text{ref}},\Sigma)] = \frac{\det Q(L_{\text{ref}},\hat{\Omega}_{\text{ref}},\Sigma)}{\det Q(L,\hat{\Omega},\Sigma)}^{1/2}$$

where $H(L; \hat{\Omega},\Sigma)$ is the reference information entropy computed for a reference sensor configuration $L_{\text{ref}}$. 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 uncertainty in the parameter estimates relative to the uncertainty obtained for the reference sensor configuration 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 Fig. 2 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 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 Fig. 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 five 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.

Fig. 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) $\lambda = h$. Please cite this article as: C. Papadimitriou, G. Lombaert, The effect of prediction error correlation on optimal sensor placement in structural dynamics, Mechanical Systems and Signal Processing (2011), doi:10.1016/j.ymssp.2011.05.019
The corresponding condition numbers for the information matrix \( Q(L, \theta_0, \Sigma) \) are also computed using the exhaustive search method and the two SSP methods and shown in Fig. 3. Reasonable values of the condition numbers are observed in Fig. 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 Fig. 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 FSSP 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 Fig. 4(a) and (b) with (c) and (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.

6.3. Finite element model of the footbridge in Wetteren (Belgium)

Optimal sensor locations are next designed for the footbridge in Wetteren (Belgium) shown in Fig. 5. The footbridge is a recently built (2003) steel bridge that crosses the E40 highway between Brussels and Ghent in Wetteren (Belgium). The bridge has a short span of 30.33 m and a large bow-string span of 75.23 m. The arches of the bow-string have an inclination of $13.78^\circ$ with respect to the vertical plane.

The bridge has been used as an in situ test case for recently developed Operational Modal Analysis (OMA) [35] and Operational Modal Analysis with eXogenous inputs (OMAX) [36] techniques. Furthermore, several tests have been carried out to validate methods for the prediction of the response of the structure to footfall excitation. For the OMA tests, measurements have been performed in a total of 72 measurement channels shown in Fig. 6. The measurements were made in 44 points (points 1–44 in Fig. 6) on the bridge deck that are located on top of the transverse beams in order to avoid local bending modes of the bridge deck. The distance between two consecutive sensors along the deck is approximately $h=5$ m. The width of the deck is 3 m. In these 44 points, 66 accelerations are measured. The first 22 measurement channels correspond to the lateral accelerations on the west side of the bridge (points 1–22 in Fig. 6), measurements channels 23–66 correspond to the vertical accelerations in all points 1–44 in Fig. 6, and measurement channel 67 corresponds to the lateral acceleration of a single point 34 on the east side of the bridge. Additionally, 5 points are located on the two arches near the connection to the bridge deck (points 45–49 in Fig. 6). Other points on the arches could not be reached from the bridge deck. The lateral accelerations in these 5 points are measurement channels 68–72.

The problem considered is the design of optimal sensor locations for modal identification discussed in Section 5.1. A finite element (FE) model of the bridge developed using the FE program ANSYS is used to demonstrate the methodology. In the FE model, the bridge deck was modeled using the ANSYS shell element SHELL63, the longitudinal and transversal beams of the bridge deck, as well as the bows, connections of the bows and the supports were modeled using the beam element BEAM188. The cables were modeled using a 3D truss element LINK8, taking into account the effective stiffness $E_{eff}$ of the cable based on the tensile cable force. The FE model has a total of 507 nodes, 1010 elements and 3042 degrees of freedom (DOFs). More details on the model can be found in [37].

The selection of the set of reference sensors will be made based on the total number of 72 sensors considered in the OMA test. The effect of spatially correlated prediction errors on the optimal location of sensors is investigated. Since the measurement points on the bridge (Fig. 6) have been chosen according to a relatively dense grid, it can already be anticipated that taking into account correlation between the sensors is important to maximize the information obtained from the selected subset. In the calculations, it is assumed that the prediction errors in the lateral accelerations are mutually correlated depending on the relative distance. The correlation function is given by (6). To study the effect of

![Fig. 5. Footbridge in Wetteren (Belgium).](image-url)
Fig. 6. Layout of potential measurements points on the bridge.
corresponding to the UNC-PE case, and \( \lambda = 0.002h \), corresponding to the UNC-PE case, and \( \lambda = 2h = 10 \text{ m} \), chosen to be of the order of twice the average distance \( h = 5 \text{ m} \) between the potential sensor points, strongly correlating the prediction errors between measurements at any two to three consecutive points shown in Fig. 6 along the deck. The \( \lambda = 2h = 10 \text{ m} \) correlation length is chosen for demonstrating the effect of correlation length in relation to the characteristic length of the contributing modes. The actual correlation length in the phase of optimal sensor location is unknown and depends on the measurement as well as model error which is expected to be distributed along the structure in the present case. A similar correlation is taken into account for the vertical accelerations on the bridge deck. The prediction errors between the vertical and the lateral accelerations are assumed to be uncorrelated. This choice is partially supported by the fact that the modes that contribute significantly to the vertical response of the deck differ from the modes that contribute significantly to the response in the transverse direction. Also the prediction errors for the lateral accelerations at the measurement channels from 68 to 72 on the arches are also taken to be mutually uncorrelated and also uncorrelated with the prediction errors of the deck channels.

To emphasize the importance of correlated prediction errors in the identification of the optimal sensor location, three modal Cases A, B and C are considered that each contain a limited number of modes. In modal Case A, only modes 1 and 2 (Fig. 7(a) and (b)) are considered that mainly involve lateral movements of the arches. In modal Case B, only modes 3 and 4 (Fig. 7(c) and (d)) are considered with coupled bending-torsional deformation of the long span and bending deformations of the small span, respectively. Finally, in modal Case C, all four modes 1–4 are considered.

6.3.1. Information entropy index

The minimum (best) and maximum (worst) IEI values for Case A as a function of the number of sensors computed by the FSSP and BSSP algorithms are shown in Fig. 8 for both the UNC-PE and COR-PE cases. Similar results for the Cases B and C are given in Figs. 9 and 10, respectively.

For all modal Cases A, B and C considered, the predictions of the IEI provided by the FSSP and the BSSP algorithms almost coincide for the UNC-PE case (see Figs. 8(a), 9(a) and 10(a)), while they differ slightly for the COR-PE cases (see Figs. 8(c), 9(c) and 10(c)) depending on the number of sensors in a configuration. The differences are more pronounced for the maximum IEI predictions for both the UNC-PE and COR-PE cases (see Fig. 8(b,d), 9(b,d) and 10(b,d)).

For comparison purposes, results from the exact exhaustive search method are also added in Figs. 8 and 9 for up to four sensors. For more sensors, the exact results from the exhaustive search method are not available due to the prohibitively large number of computations required. For up to 4 sensors, the SSP algorithms are also shown in Figs. 8 and 9 to be very accurate in predicting the minimum IEI.

The two SSP algorithms are computationally very efficient algorithms for designing approximations of the optimal and worst sensor configurations for even large number of sensors and large number of DOFs. In addition, the FSSP algorithm requires one to two orders of magnitude less computational effort than the BSSP algorithm since in most practical cases the sensor configurations are designed for a relatively small number of sensors in relation to the model DOFs. However, the FSSP results in Figs. 8–10 are slightly less accurate than the BSSP results. Given the manageable computational effort involved in both heuristic algorithms [15], an effective use of the FSSP and BSSP algorithms is to combine their predictions by selecting the optimal (respectively worst) sensor configuration as the one that corresponds to the minimum (respectively maximum) information entropy value.

6.3.2. Design of optimal sensor locations for modal Case A

For the modal Case A, the optimal sensor locations as a function of the number of sensors are shown in Fig. 11 for the two SSP algorithms and for both the UNC-PE and COR-PE cases. These results are also compared to the ones obtained from the exact exhaustive search method for up to 4 sensors. Similar results for the Cases B and C are given in Figs. 12 and 13 for the optimal sensor locations. Comparing Figs. 11(a,b), 12(a,b) and 13(a,b) with Figs. 11(c,d), 12(c,d) and 13(c,d), respectively, it is clear that the spatial correlation of the prediction error affects considerably the optimal location of sensors.

Fig. 7. Finite element modal analysis results: modes 1–4. Top: y-displacements, middle: z displacements, bottom: displacement vector sum.
Fig. 8. Information entropy index for Case A as a function of the number of sensors for (a,c) the optimal and (b,d) the worst sensor configuration; (a,b) $\lambda = 0.002h$ and (c,d) $\lambda = 2h$.

Fig. 9. Information entropy index for Case B as a function of the number of sensors for (a,c) the optimal and (b,d) the worst sensor configuration; (a,b) $\lambda = 0.002h$ and (c,d) $\lambda = 2h$. 

Specifically, for the UNC-PE model, the optimal sensor locations for the first three sensors shown in Fig. 11(b) are placed by the BSSP algorithm at the measurement channels 69, 72 and 68 (points 46, 49 and 45 in Fig. 6) on the left and right arches. The preference of the sensor locations on the arches is justified from the fact that the arch motion is dominant for the two contributing modes included in sensor design methodology. The remaining two measurement points on the arches are selected as the fifth and the seventh optimal sensor locations by the BSSP algorithm, while the fourth and the sixth optimal sensor locations are selected to be the lateral deck measurement channels 13 (point 13 in Fig. 6) and 67 (point 34 in Fig. 6) of the west and east side of the bridge, respectively. The optimal locations for the next 17 sensors (from 8th to 24th) are the lateral acceleration channels along the bridge deck, which are clustered close to the 13 measurement channel. Sensor measurements along the vertical direction are not predicted by the algorithm for the first 24 optimal sensor locations. The optimal positions of the vertical sensors are measurement channels 35 and 57 (see Fig. 11(b)) which correspond to vertical points 13 and 35 in Fig. 6. As the number of sensors increases, the optimal sensor locations for the vertical sensors tend to cluster close to channel points 35 and 57.

For the COR-PE model, the optimal locations for the first three sensors shown in Fig. 11(d) are predicted by the BSSP algorithm at the arch measurement points exactly at the same positions as the ones predicted by the UNC-PE case. This is expected since the arch motion dominates and no spatial correlation was assumed for the measurement points at the arches. In addition, the optimal sensor configuration for the first six sensors includes all five measurement points at the arches as well as the single lateral measurement channel 67 (point 34 in Fig. 6) available at the east side of the bridge. In contrast to the UNC-PE case in Fig. 11(b) where the next group of seventeen sensors is clustered close to lateral channel point 13, for the COR-PE case in Fig. 11(d) the next group of seventeen sensors are uniformly distributed along the bridge deck in both lateral and vertical measurement channels. For the COR-PE case, clustering effects with sensor positions in neighboring locations are avoided due to the correlation assumed between prediction errors. As a result, sensors are more rationally distributed, allowing for both vertical and lateral measurements to be selected earlier in the optimal sensor location design process which is more reasonable, given that the first two modes do have a non-negligible contribution to both the lateral and vertical directions of motion. It is worth observing from Figs. 11(b) and (d) that the vertical measurement channels 35 and 57 are favored as most optimal locations for both the UNC-PE and COR-PE case. Despite the relatively large correlation length compared to the width of the deck, these two channels correspond to opposite neighborhood points 13 and 37 of the west and east side of the bridge, respectively. Their selection is due to the deck torsional effects manifested in the mode.

Fig. 10. Information entropy index for Case C as a function of the number of sensors for (a,c) the optimal and (b,d) the worst sensor configuration; (a,b) $\lambda = 0.002h$ and (c,d) $\lambda = 2h$. 

It should be noted that the FSSP algorithm in Fig. 11(a) provides very similar results as the BSSP algorithm in Fig. 11(b) for the UNC-PE case. Comparing the COR-PE cases in Figs. 11(c) and (d), the optimal sensor location results for the lateral and vertical measurement channels along the deck have some differences for sensor configurations involving from 5 up to 20 sensors. These differences are due to the fact that the FSSP algorithm is less accurate than the BSSP algorithm as it is seen by the higher IEI values predicted by the FSSP algorithm in Fig. 8(a).

6.3.3. Design of optimal sensor locations for modal Case B

In modal Case B the contributing modes of the bridge are the 3rd and 4th mode shown in Fig. 7(c) and (d). These modes are dominated by coupled lateral bending-torsional deformation of the long span and vertical bending deformation of the short span of the bridge. For the sensor configuration involving one sensor, the exact method predicts the optimal location of one sensor in Fig. 12 to be the vertical measurement channel 46 (point 24 in Fig. 6), while the FSSP or BSSP algorithms predict the vertical measurement channel 24 (point 2 in Fig. 6). It should be noted that these two different predictions correspond to exactly the same information entropy which is also justified from the fact that these two measurement channels 24 and 46 are located at the same position on the bridge at the opposite west and east sides of the deck.

According to the exact predictions for a sensor configuration involving two sensors as shown in Fig. 12(a,b) and (c,d) for the UNC-PE and COR-PE cases, respectively, the optimal location of the first sensor is the vertical measurement channel 24 (point 2 in Fig. 6) on the short span of the bridge, while the optimal location of the second sensor is the lateral measurement channel 15 (point 15 in Fig. 6) of the long span of the bridge. The BSSP and FSSP algorithms correctly predict these optimal sensor locations for the UNC-PE and COR-PE cases, respectively. The results are consistent with the dominant deflections appearing in the contributing modes. The optimal location of the rest of the sensor depends on the correlation model assumed.

For the UNC-PE case, the optimal location of the third sensor predicted by the BSSP algorithm is the vertical measurement channel 46 (point 24 in Fig. 6) of the long span of the bridge. The optimal locations of the next group of sixteen sensors tend to cluster in the neighborhood of the first three optimal sensor locations. Sensors positioned on the lateral measurement channels along the deck are strongly preferred. Lateral measurement channels on the two arches are not contained in the optimal sensor configuration involving up to 19 sensors. The optimal locations of the 20th, 22nd and 27th sensors are measurement points on the two arches. The 28th and the 29th sensor are optimally located at vertical deck measurement channels 37 and 59 (points 15 and 37 in Fig. 6), while the following group of sensors is clustered in the neighborhood of these vertical measurement channels.

For the COR-PE case, the optimal location of the third sensor predicted by the more accurate FSSP algorithm is lateral measurement channel 67 (point 34 in Fig. 6) on the east side of the bridge deck. According to BSSP algorithm, shown to be more accurate in Fig. 9 for more than 7 sensors, the first 20 sensors are uniformly distributed in positions close to lateral deck measurement channel 15 (point 15 in Fig. 6) and vertical deck measurement channels 24 and 36, 46 and 58 (points 2, 14, 24 and 36 in Fig. 6), avoiding sensor placement at neighboring measurement channels. More specifically, the group of 12 sensors (Fig. 12(b)) within the optimal sensor configuration involving 20 sensors, promoted by the UNC-PE case to be mainly clustered at lateral measurement channels around measurement channel 15, is replaced by only 4 sensors in the COR-PE case, while the remaining 8 sensors are distributed for the COR-PE case along the deck vertical measurement channels and the lateral measurement channels on the two arches. This saves sensors in order to promote earlier (compared to the UNC-PE case) in the design the measurement channels on the arches as optimal locations of sensors 15 and 16, as well as the vertical measurement channels on the deck close to channels 36 and 58 shown in Fig. 12(d).

6.3.4. Design of optimal sensor locations for modal Case C
In modal Case C, all four modes shown in Fig. 7 are contributing. Based on the results in Fig. 13, for both the UNC-PE and COR-PE cases the optimal configuration involving four sensors promote a combination of lateral and vertical measurement channels on the deck and lateral measurement channels on the arches. Specifically, the first optimal sensor location is on the deck along the vertical direction (point 2 on the west side or point 24 on the opposite east side in Fig. 6), the second is on the deck along the lateral direction (point 16 in Fig. 6), while the third and fourth optimal sensors are located on the two arches (points 49 and 46 on the east side or 45 on the opposite west side in Fig. 6). It becomes evident that the optimal positions for the four sensors monitor motions in the vertical and lateral directions of the bridge deck as well as the lateral directions of the two arches. Both the UNC-PE and COR-PE models predict very similar optimal positions for the first four sensors since the prediction errors associated with these positions (two lateral on the arches, one deck lateral and one deck vertical) are assumed independent for both cases. Differences between the UNC-PE and COR-PE cases are observed for the position of the rest of the sensors. For each correlation model considered (UNC-PE or COR-PE), the positions of these sensors follow a pattern very similar to the one observed in modal Case B or A. Once again, in the spatially correlated prediction error case, the sensors are more uniformly distributed along the bridge deck, avoiding redundant information from sensor clustering predicted by the UNC-PE case.
Summarizing, it can be concluded that similar results for the effect of spatial correlation of the prediction errors are observed for all Cases A, B and C. Significant qualitative differences between the UNC-PE and CORR-PE cases are observed. Specifically, in the UNC-PE case, the optimal sensor locations are clustered in the neighborhood of specific lateral and vertical measurement channels. In the COR-PE case, sensors are more uniformly distributed along the deck, avoiding clustering effect and neighboring sensor positions with redundant information. In addition, less important sensor locations predicted in the UNC-PE case very late in the design process are promoted by the COR-PE case much earlier in the design process. The results are consistent with the derived theoretical results showing that the minimum distance between sensor locations is controlled by the correlation length assumed for the prediction errors.

7. Conclusions

The theoretical developments provide valuable insight into the effect of spatial correlation of the prediction error on the optimal placement of sensors for modal identification or parameter estimation in finite element model updating problems encountered in structural dynamics. Spatially uncorrelated prediction errors between two observations should be selected when these contain qualitatively different information for the dynamics of the structure. Examples of such observations include the transverse and vertical deck accelerations of the footbridge. The spatial correlation is important to consider in order to avoid redundant information provided by neighboring sensors with distance less than the characteristic length of the highest contributing mode. Selected applications on modal identification and parameter estimation problems on simple models and a finite element model of a bridge, clearly demonstrated this effect of prediction error correlation. It was also demonstrated that heuristic sequential sensor placement algorithms maintain their advantages, in terms of both computational efficiency and accuracy, also for the case of spatially correlated prediction errors.

The developments have obvious implications on the optimal distance between two sensors in dense meshes often encountered in finite element models of structures. In such models, the potential distance between two sensors is of the order of the element size. Spatially uncorrelated prediction errors allow two sensors to be placed at the same or neighboring nodes of the finite element mesh. This placement may be desirable when sensors measure in different directions at the nodes of the finite element mesh, thus providing, depending on the structural model, qualitatively different information. For sensors measuring along the same direction, a sensor design with closely spaced sensors may not be acceptable since sensors at an intermediate distance smaller than the characteristic length determined by the highest contributing mode, tend to have very similar information content. Spatially correlated prediction errors between
two neighboring nodes prohibit the placement of sensors close to each other. Within the characteristic length, the minimum distance between any two sensors is determined by the correlation length, while for distances larger than the characteristic length it is controlled by the detailed characteristics of the response sensitivities. Unacceptable designs with very close sensors are avoided by selecting the correlation length of the prediction errors to be at least of the order of the characteristic length of the problem.

The correct correlation structure and correlation length over different regions of a structure, however, remain an issue since in the initial experimental design phase no measurements are available to support the selection of a correlation structure consistent with the data and the structural model on which the optimal sensor location is based.

Acknowledgements

The second author would like to acknowledge the support of K.U. Leuven PFV/10/002 for the Optimization in Engineering Center (OPTEC).

Appendix A. Proof of representation (14)

The representation (14) and the semi-positive definiteness of the matrix $\delta Q_{MN}$ is shown as follows. First note that using (9) the matrix $Q(L_{M+N})$ for the sensor configuration $L_{M+N}$ involving $M+N$ sensors admits the representation

$$Q(L_{M+N}) = \sum_{k=1}^{N} (L_{M+N} \Sigma \phi_k \Sigma) A^{-1} (L_{M+N} \Sigma \phi_k \Sigma)^T$$

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

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

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:

$$A_M B_M + A_{NM} B_{NM} = I_{MM}$$
$$A_M B_{NM} + A_{NM} B_N = 0_{MN}$$
$$A_N B_{NM} + A_{NM} B_{NN} = I_{NN}$$
$$A_N B_N + A_{NN} B_{NM} = 0_{NM}$$

(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

$$B_M = A_M^{-1} - A_M^{-1} A_{NM} B_{NM}$$

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

(A.5)
$$B_N = [A_N - A_{NM} A_{NM}^{-1} A_N^{-1}]^{-1}$$

(A.6)

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

$$Q(L_{M+N}) = \sum_{k=1}^{N} (L_{M+N} \Sigma \phi_k \Sigma)^T B_M (L_{M+N} \Sigma \phi_k \Sigma) + \delta Q_{MN}^*$$

(A.7)

where $\delta Q_{MN}^*$ is given by

$$\delta Q_{MN}^* = \sum_{k=1}^{N} [(L_N \Sigma \phi_k \Sigma)^T B_{MN} (L_{M+N} \Sigma \phi_k \Sigma) + (L_N \Sigma \phi_k \Sigma)^T B_{NM} (L_{M+N} \Sigma \phi_k \Sigma) + (L_N \Sigma \phi_k \Sigma)^T B_N (L_N \Sigma \phi_k \Sigma)$$

(A.8)

The representation (14) follows from (A.7) by substituting $B_M$ from (A.4), expanding the first term in the right hand side and noting that $Q(L_M) = \sum_{k=1}^{N} (L_M \Sigma \phi_k \Sigma)^T A_M^{-1} (L_M \Sigma \phi_k \Sigma)$, while $\delta Q_{MN}$ in (14) is given by the remaining terms as

$$\delta Q_{MN} = - \sum_{k=1}^{N} (L_M \Sigma \phi_k \Sigma)^T A_{MN}^T A_{NM} B_{MN} (L_M \Sigma \phi_k \Sigma) + \delta Q_{MN}^*$$

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

$$\delta Q_{NM} = \sum_{k=1}^{N} (l_M \sum_{i} x_i \bar{x}_k)^T B_{NM}^T B_N^{-1} B_NM (l_M \sum_{i} \bar{x}_i) + \delta Q_{NM}^*$$

(A.10)

Introducing the auxiliary matrices $E_{M,k} = (l_M \sum_{i} x_i \bar{x}_k)$ and $E_{N,k} = (l_M \sum_{i} \bar{x}_i)$ in (A.8) and (A.10), one finally simplifies the expression for $\delta Q_{NM}$ in the form:

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

(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

$$\delta Q_{NM} = \sum_{k=1}^{N} [(E_{M,k} B_{NM} + E_{N,k} B_{NM}) B_N^{-1} B_NM + (E_{N,k} B_{NM} + E_{N,k} B_{NM}) E_{N,k}] = \sum_{k=1}^{N} [(B_{NM} E_{M,k} + B_N E_{N,k}) B_N^{-1} (B_{NM} E_{M,k} + B_N E_{N,k})]$$

(A.12)

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

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

(A.13)

where $\bar{z} = (B_{NM} E_{M,k} + B_N E_{N,k}) \bar{y}$, 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 (14) 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

$$A_1 = L_1 \Sigma L_1 = \sigma^2 + s^2$$

(B.1)

and

$$A_{1M} = L_1 \Sigma L_{M} = [0, \ldots, 0, \sigma^2 R(\delta)]_{M-1}$$

(B.2)

where $\sigma^2$ and $\delta^2$ 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:

$$\hat{Q}(\delta) = Q(L_{M}) + \sum_{k=1}^{N} [(B_{1M} E_{M,k} + B_1 E_{1,k}) B_1^{-1} (B_{1M} E_{M,k} + B_1 E_{1,k})]$$

(B.3)

Note that the right-hand-side of (B.2) is valid based on the assumption that the correlation length $\lambda$ is small compared to the minimum distance $\lambda$ 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 $B_1^T = -A_{1M}^{-1} A_{NM} B_1$, where $B_1 = B_1(\delta)$ is a scalar given by (A.6) in the form:

$$B_1 = [A_1 - A_{1M} A_{NM}^{-1} A_{1M}^T]^{-1}$$

(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 $1/[\sigma^2 + s^2]$, the scalar $B_1(\delta)$ admits the representation

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

(B.5)

Substituting $B_1^T = -A_{1M}^{-1} A_{1M}^T B_1$ from (A.5), noting that $A_{1M}$ is given by (B.2), also that $A_{1M} E_{M,k} = \sum_{i} x_{M,k}$ and $E_{1,k} = \sum_{i} x_{N,k}$, 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

$$\hat{Q}(\delta) = Q(L_{M}) + B_1(\delta) G_{1M}(\delta)$$

(B.6)
where \( G_{IM} (\delta) \) a semi-positive definite matrix given by

\[
G_{IM} (\delta) = \sum_{k = 1}^{N} \left( -\frac{\sigma R(\delta)}{\delta + s^2} \sum_{j} x_{M,k} + \sum_{j} x_{N,k} \right) \left( -\frac{\sigma R(\delta)}{\delta + s^2} \sum_{j} x_{M,k} + \sum_{j} x_{N,k} \right)
\]  

(B.7)

Introducing the vector \( \hat{u}_k (\delta) = \sqrt{F_{N,k} - \sqrt{F_{M,k}}}, \) the matrix \( G_{IM} (\delta) \) takes the form:

\[
G_{IM} (\delta) = \sum_{k = 1}^{N} \left[ I(\delta) \sum_{j} x_{M,k} + \hat{u}_k (\delta) \right] \left[ I(\delta) \sum_{j} x_{M,k} + \hat{u}_k (\delta) \right]
\]

(B.8)

where

\[
I(\delta) = 1 - \frac{\sigma^2}{\delta^2 + s^2} \frac{R(\delta)}{R(\delta)} > 0
\]

(B.9)

Note that the sum in the first term is independent of \( \delta \). Also the elements of the response sensitivity vector \( \sqrt{F_{M,k}} \) 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 \( \delta \) compared to the characteristic length of the dynamic problem, the vector \( \sqrt{F_{N,k}} \) at the new sensor location does not vary significantly from the vector \( \sqrt{F_{M,k}} \) in the neighborhood sensor location. In this case, the vector \( \hat{u}_k (\delta) \) 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 \( \delta_1 \) and \( \delta_2 \) with \( \delta_1 > \delta_2 \) the following inequality \( R(\delta_1) > R(\delta_2) \) holds. Using (B.5) and (B.9), it is straightforward to show that for \( \delta_1 > \delta_2 \), the following inequality holds \( I^2(\delta_1) > I^2(\delta_2) \). The last expression along with the fact that the first term dominates the right hand side of (B.8) for sufficiently small distances \( \delta_1 \) and \( \delta_2 \) determined by the characteristic length of the problem, results in \( B_1 (\delta_1) G_{IM} (\delta_1) - B_2 (\delta_2) G_{IM} (\delta_2) \) to be a positive definite matrix. Applying (B.6) for \( \delta_1 \), noting that

\[
\hat{Q}(\delta_1) = Q(L_2) + B_1 (\delta_1) G_{IM} (\delta_1) + [B_1 (\delta_1) G_{IM} (\delta_1) - B_2 (\delta_2) G_{IM} (\delta_2)] = \hat{Q}(\delta_2) + [B_1 (\delta_1) G_{IM} (\delta_1) - B_2 (\delta_2) G_{IM} (\delta_2)]
\]

(B.10)

the validity of (19) follows immediately from the relation (17) with \( A_0 = \hat{Q}(\delta_2) \) and \( B_0 = [B_1 (\delta_1) G_{IM} (\delta_1) - B_2 (\delta_2) G_{IM} (\delta_2)] \).

It should be noted that for the nearly uncorrelated prediction error case, the correlation length \( \lambda \) is significantly smaller than the characteristic length of the problem, and the factor \( I^2(\delta)/B_1 (\delta) \) tends to \( (\sigma^2 + s^2)^{-1} \) independent of \( \delta \), provided that \( \delta \) is large compared to the correlation length. In this case the dominant term in (B.8) is independent of \( \delta \) 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 by the second and third terms.

References


