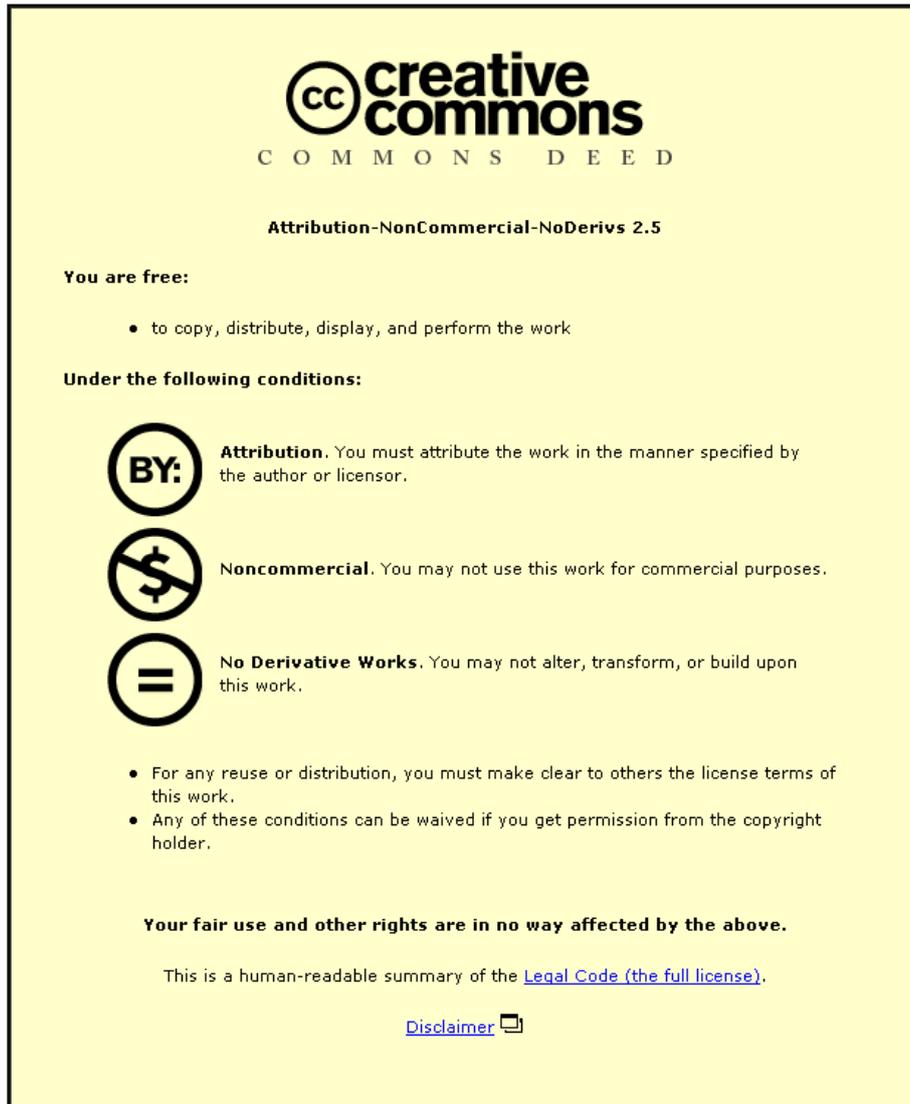


This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



CC creative commons
COMMONS DEED

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

BY: **Attribution.** You must attribute the work in the manner specified by the author or licensor.

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Optimization Algorithms for System Integration

Costas Papadimitriou^{1, a} and Evaggelos Ntotsios^{1, b}

¹University of Thessaly, Department of Mechanical and Industrial Engineering, Volos 38334, Greece

^acostasp@uth.gr, ^bentotsio@uth.gr

Keywords: Structural Dynamics, Structural Identification, Damage Detection, Optimal Sensor Location, Bayesian inference, Information Entropy.

Abstract. This work outlines the optimization algorithms involved in integrating system analysis and measured data collected from a network of sensors. The integration is required for structural health monitoring problems arising in structural dynamics and related to (1) model parameter estimation used for finite element model updating, (2) model-based damage detection in structures and (3) optimal sensor location for parameter estimation and damage detection. These problems are formulated as single- and multi-objective optimization problems of continuous or discrete-valued variables. Gradient-based, evolutionary, hybrid and heuristic algorithms are presented that effectively address issues related to the estimation of multiple local/global solutions and computational complexity arising in single and multi-objective optimization involving continuous and discrete variables.

Introduction

Successful health monitoring and diagnosis of structural systems depends to a large extent on the integration of cost-effective intelligent sensing techniques, accurate physics-based computational models simulating structural behavior, effective system identification methods, sophisticated health diagnosis algorithms, as well as decision-making expert systems to guide management in planning optimal cost-effective strategies for system maintenance, inspection and repair/replacement. A structural health monitoring system incorporates algorithms related to (1) finite element model parameter estimation (updating), (2) structural damage detection based on finite element model updating, and (3) optimal sensor location.

Finite element model updating methods based on vibration data are often used to develop high fidelity models so that model predictions are consistent with measured data. The need for model updating arises because there are always assumptions and numerical errors associated with the process of constructing a theoretical model of a structure and predicting its response using the underlined model. Reviews of model updating methodologies based on modal data can be found in [1]. Moreover, model updating methodologies are useful in predicting the structural damage by continually updating the structural model using vibration data [2-8]. Such updated models obtained periodically throughout the lifetime of the structure can be further used to update the response predictions and lifetime structural reliability based on available data [9]. Optimal sensor location methods refer to algorithms for optimizing the location and number of sensors in the structure such that the measure data contain the most important information for structural identification purposes. Algorithms based on information theory and using a nominal finite element model of the structure, have been proposed to address this problem [10-13]. Effective heuristic optimization tools [14,15] have also been developed for efficiently solving the resulting nonlinear single- and multi-objective optimization problems involving discrete-valued variables.

This work first presents a Bayesian methodology for the parameter estimation and damage detection used in the structural health monitoring system. It is shown that the formulations result in optimization problems with respect to continuous variables. Computational algorithms for solving these optimization problems are proposed to overcome convergence problems and premature convergence to local optima. In particular, hybrid algorithms based on evolution strategies and

gradient methods are necessary and well-suited optimization tools for solving the resulting non-convex single-objective optimization problem and identifying the global optimum from multiple local ones. For gradient-based algorithms, computationally efficient schemes for estimating the gradients and Hessians of the objective functions are proposed and shown to significantly reduce the computational effort and the number of iterations required for convergence. Next, theoretical and computational issues arising in the selection of the optimal sensor location for parameter estimation and damage detection are addressed. The information entropy is used as the performance measure of a sensor configuration. The optimal sensor location problem is formulated as single- and multi-objective optimization problems involving discrete-valued variables. Accurate and computationally efficient heuristic algorithms for solving these problems are outlined.

Bayesian Methodology for Model Parameter Estimation

Consider a parameterized class \mathcal{M} 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 $\boldsymbol{\theta} \in R^{N_\theta}$ be the vector of free parameters (physical or modal parameters) in the model class. A Bayesian statistical system identification methodology [16,17] is used to estimate the values of the parameter set $\boldsymbol{\theta}$ and their associated uncertainties using the information provided from dynamic test data. For this, the uncertainties in the values of the structural model parameters $\boldsymbol{\theta}$ are quantified by probability density functions (PDF) that are updated using the dynamic test data.

According to the Bayesian structural identification methodology, assuming independent and zero-mean Gaussian prediction errors $e_j(k)$ with variance σ_j^2 , the updating PDF $p(\boldsymbol{\theta}, \boldsymbol{\sigma} | D)$ of the parameter sets $\boldsymbol{\theta}$ and $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{N_0})$, given the measured data D and the class of models \mathcal{M} , takes the form [18]:

$$p(\boldsymbol{\theta}, \boldsymbol{\sigma} | D) = \frac{\tilde{c}}{(\sqrt{2\pi})^{N_D N_0} \rho(\boldsymbol{\sigma})} \exp\left\{-\frac{N_D N}{2} J(\boldsymbol{\theta}; \boldsymbol{\sigma})\right\} \pi_\theta(\boldsymbol{\theta}) \pi_\sigma(\boldsymbol{\sigma}) \quad (1)$$

where $J(\boldsymbol{\theta}; \boldsymbol{\sigma})$ is a measure of fit of the measured response characteristics and the corresponding response characteristics predicted by a particular model in the model class \mathcal{M} , and $\rho(\boldsymbol{\sigma})$ is a function of the prediction error parameters $\boldsymbol{\sigma}$, $\pi_\theta(\boldsymbol{\theta})$ and $\pi_\sigma(\boldsymbol{\sigma})$ are the prior distribution for the parameter sets $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$, respectively, $N = N_0$, N_0 is the number of response characteristics, N_D is the number of measured data sets, and \tilde{c} is a normalizing constant chosen such that the PDF in (1) integrates to one.

Time History Data. For the case for which the response characteristics consist of the response time histories data $D = \{\hat{x}_j(k\Delta t) \in R^{N_0}, j = 1, \dots, N_0, k = 1, \dots, N_D\}$ at N_0 measured DOFs, where N_D is the number of the sampled data using a sampling rate Δt , then

$$J(\boldsymbol{\theta}; \boldsymbol{\sigma}) = \frac{1}{N_0} \sum_{j=1}^{N_0} \frac{1}{\sigma_j^2} J_j(\boldsymbol{\theta}), \quad J_j(\boldsymbol{\theta}) = \frac{1}{N_D} \sum_{k=1}^{N_D} [\hat{x}_j(k) - x_j(k; \boldsymbol{\theta})]^2 \quad (2)$$

$\rho(\boldsymbol{\sigma}) = \prod_{j=1}^{N_0} \sigma_j^{N_D}$, and $x_j(k; \boldsymbol{\theta})$, $j = 1, \dots, N_d$, are the predictions of the sampled response time histories obtained from a particular model in the model class \mathcal{M} corresponding to a specific value of the parameter set $\boldsymbol{\theta}$.

Modal Data. For the case where the response characteristics consist of modal data $D = \{\hat{\omega}_r^{(k)}, \hat{\phi}_r^{(k)} \in R^{N_0}, r = 1, \dots, m\}$, where $\hat{\omega}_r^{(k)}$ are the modal frequencies and $\hat{\phi}_r^{(k)}$ are modeshape

components at N_0 measured DOFs, m is the number of observed modes and N_D is the number of modal data sets available, then

$$J(\boldsymbol{\theta}; \boldsymbol{\sigma}) = \frac{1}{N_D N} \sum_{r=1}^m \left[\frac{1}{\sigma_{\omega_r}^2} \varepsilon^2(\omega_r, \hat{\omega}_r) + \frac{N_0}{\sigma_{\phi_r}^2} \varepsilon^2(L\phi_r, \hat{\phi}_r) \right] \quad (3)$$

where

$$\varepsilon(\omega_r, \hat{\omega}_r) = \varepsilon_{\omega_r} = \frac{\omega_r^2 - \hat{\omega}_r^2}{\hat{\omega}_r^2} \quad \text{and} \quad \varepsilon(L\phi_r, \hat{\phi}_r) = \varepsilon_{\phi_r} = \frac{\|\beta_r L\phi_r - \hat{\phi}_r\|}{\|\hat{\phi}_r\|} \quad (4)$$

$\omega_r(\boldsymbol{\theta})$ and $\phi_r(\boldsymbol{\theta}) \in R^{N_d}$, $r = 1, \dots, m$, are the predictions of the modal frequencies and modeshapes obtained for a particular value of the model parameter set $\boldsymbol{\theta}$, $N = m(N_0 + 1)$ is the number of measured data per modal set, $N_D = 1$ and $\beta_r = \hat{\phi}_r^T(L\phi_r)/(L\phi_r)^T(L\phi_r)$ is a normalization constant that accounts for the different scaling between the measured and the predicted modeshape.

Most Probable Model. Note that the probability density function in (1) quantifies the uncertainty in the values of the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\sigma}$ given the data D . For a non-informative prior distributions $\pi_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $\pi_{\boldsymbol{\sigma}}(\boldsymbol{\sigma})$, the optimal value $\hat{\boldsymbol{\theta}}$ of the parameter set $\boldsymbol{\theta}$ given the data is obtained by [19] minimizing the measure of fit $J(\boldsymbol{\theta}; \boldsymbol{\sigma})$.

Bayesian Methodology for Damage Detection

Damage detection is accomplished by introducing a family of μ model classes $M_1 \dots M_{\mu}$ and associating each model class to a damage pattern in the structure, indicative of the location of damage. Each model class M_i is assumed to be parameterized by a number of structural model parameters $\boldsymbol{\theta}_i$ scaling the stiffness contributions of a ‘‘possibly damaged’’ substructure, while all other substructures are assumed to have fixed stiffness contributions equal to those corresponding to the undamaged structure. Using a Bayesian model selection framework, the probable damage locations are ranked according to the posterior probabilities of the corresponding model classes. The most probable model class will be indicative of the substructure that is damaged, while the probability distribution of the model parameters of the corresponding most probable model class will be indicative of the severity of damage in the identified damaged substructure.

Using Bayes’ theorem, the posterior probabilities of the various model classes given the data D is

$$P(M_i | D) = \frac{p(D | M_i) P(M_i)}{d} \quad (5)$$

where $p(D | M_i)$ is the probability of observing the data from the model class M_i , $P(M_i)$ is the prior probability of the model class M_i , while d is selected so that the sum of all model probabilities equals to one. Assuming there is no prior preference as to what class of models we choose, we may set that $P(M_i) = 1/\mu$ in (5).

The following asymptotic approximation has been introduced to give a useful and insightful estimate of the integral involved in $p(D | M)$ in (5) [19-20].

$$p(D | M_i) \sim (2\pi)^{N_\theta/2} \frac{\pi_\theta(\hat{\theta}_i | M_i) [J_M(\hat{\theta}_i, \hat{\sigma}_i)]^{-N_D N}}{\sqrt{\det \mathbf{h}_{M_i}(\hat{\theta}_i, \hat{\sigma}_i)}} \quad (6)$$

where, for uniform prior distribution of the parameters θ_i in a model class M_i , $\hat{\theta}_i$ is the value that minimizes $J_{M_i}(\hat{\theta}_i, \hat{\sigma}_i)$, and $\mathbf{h}_{M_i}(\hat{\theta}_i, \hat{\sigma}_i)$ is defined by

$$\mathbf{h}_{M_i}(\hat{\theta}_i, \hat{\sigma}_i) = -\nabla_{\theta_i} \nabla_{\theta_i}^T \ln[J(\theta_i, \sigma_i)]^{-N_D N} \Big|_{\theta=\hat{\theta}} \quad (7)$$

in which $\nabla_{\theta} = [\partial/\partial\theta_1, \dots, \partial/\partial\theta_{N_\theta}]^T$ is the usual gradient vector with respect to the parameter set θ_i , and $\hat{\sigma}_i^2$ is the optimal prediction error variance for the model class M_i . The approximate estimate is unreliable when the optimal $\hat{\theta}$ is outside the region $\Theta_i = \{\theta_i : \theta_i \leq \theta_i^u\}$ of variation of θ_i , where θ_i^u are the values of θ_i at the undamaged condition. Alternatively, one can use importance sampling method to compute the integral involved in estimating $p(D | M)$ in (5) [19].

Under the assumption that the prior distributions $\pi_\theta(\theta_i)$ are non-informative uniform distributions over the range of variation of θ_i , the probability of the model class M_i is given by [20]

$$\log P(M_i | D) = -N_J \log [J_M(\hat{\theta}_i, \hat{\sigma}_i; M_i)] + \beta(\hat{\theta}_i; M_i) + P(M_i) + d \quad (8)$$

where d is constant independent of the model class M_i , $N_J = N_D N / 2$, and the factor $\beta(\hat{\theta}_i; M_i)$ in (8), known as the Ockham factor, simplifies for large number of data N_J to [21, 22]

$$\beta(\hat{\theta}_i; M_i) = \beta_i = -\frac{N_{\theta_i}}{2} \log N_J \quad (9)$$

where it is evident that it depends from the number N_{θ_i} of the model parameters involved in the model class M_i .

The optimal model class M_{best} is selected as the model class that maximizes the probability $P(M_i | D)$ given by (8). It is evident that the selection of the optimal model class depends on the measure of fit $J_M(\hat{\theta}_i, \hat{\sigma}_i; M_i)$ between the measured characteristics and the characteristics predicted by the optimal model of a model class M_i . Thus, the first term in (8) gives the dependence of the probability of a model class M_i from how well the model class predicts the measurements. The smaller the value of $J_M(\hat{\theta}_i, \hat{\sigma}_i; M_i)$, the higher the probability $P(M_i | D)$ of the model class M_i . Based on the Ockham factor β_i simplified in (9), the ordering of the model classes in (8) also depends on the number N_{θ_i} of the structural model parameters that are involved in each model class. Specifically, model classes with large number of parameters are penalized in the selection of the optimal model class.

It should be pointed out that damage detection using the methodology requires the solution of μ optimization problems with the objective function for each model class M_i to be the measure of fit function $J_{M_i}(\theta_i, \sigma_i)$.

Optimization Issues Related to Model Updating and Damage Detection

The resulting optimization problems are solved using available numerical algorithms. The optimization of $J_{M_i}(\boldsymbol{\theta}_i, \boldsymbol{\sigma}_i)$ with respect to $\boldsymbol{\theta}$ for given $\boldsymbol{\sigma}$ can readily be carried out numerically using any available algorithm for optimizing a nonlinear function of several variables. These single objective optimization problems may involve multiple local/global optima. Conventional gradient-based local optimization algorithms lack reliability in dealing with the estimation of multiple local/global optima observed in structural identification problems [18, 23], since convergence to the global optimum is not guaranteed. Evolution strategies [24] are more appropriate and effective to use in such cases. Evolution strategies are random search algorithms that explore better the parameter space for detecting the neighborhood of the global optimum, avoiding premature convergence to a local optimum. A disadvantage of evolution strategies is their slow convergence at the neighborhood of an optimum since they do not exploit the gradient information. A hybrid optimization algorithm [19] should be used that exploits the advantages of evolution strategies and gradient-based methods. Specifically, an evolution strategy is used to explore the parameter space and detect the neighborhood of the global optimum. Then the method switches to a gradient-based algorithm starting with the best estimate obtained from the evolution strategy and using gradient information to accelerate convergence to the global optimum.

Gradient and Hessian Computations. In order to guarantee the convergence of the gradient-based optimization methods for structural models involving a large number of DOFs with several contributing modes, the gradient of the objective function with respect to the parameter set $\boldsymbol{\theta}$ has to be estimated accurately. It has been observed that numerical algorithms such as finite difference methods for gradient evaluation does not guarantee convergence. Moreover, gradient computations with respect to the parameter set using the finite difference method requires the solution of as many eigenvalue problems as the number of parameters.

Analytical expressions for the gradient of the objective functions can be used to overcome the convergence problems. These analytical expressions are next given for the case of structural identification based on modal data. In particular, Nelson's method [25] is used for computing analytically the first derivatives of the eigenvalues and the eigenvectors. The advantage of the Nelson's method compared to other methods is that the gradient of eigenvalue and the eigenvector of one mode are computed from the eigenvalue and the eigenvector of the same mode and there is no need to know the eigenvalues and the eigenvectors from other modes. For each parameter in the set $\boldsymbol{\theta}$ this computation is performed by solving a linear system of the same size as the original system mass and stiffness matrices. Nelson's method has also been extended in this work to compute the second derivatives of the eigenvalues and the eigenvectors.

Finally, the computation of the gradients and the Hessian of the objective functions is shown to involve the solution of a single linear system, instead of N_θ linear systems required in usual computations of the gradient and $N_\theta(N_\theta + 1)$ linear systems required in the computation of the Hessian. This reduces considerably the computational time, especially as the number of parameters in the set $\boldsymbol{\theta}$ increase. The expressions for the first and second derivatives of the objective functions are next presented. Due to space limitations details of the derivations are not shown.

The gradient of square errors $\varepsilon_{\omega_r}^2(\boldsymbol{\theta})$ and $\varepsilon_{\phi_r}^2(\boldsymbol{\theta})$ involved in objectives (see equations (3) and (4)) are given by

$$\frac{\partial \varepsilon_{\omega_r}^2(\boldsymbol{\theta})}{\partial \theta_j} = \left[\frac{2\varepsilon_{\omega_r}(\boldsymbol{\theta})}{\hat{\omega}_r^2} \phi_r^T \right] (K_j - \omega_r^2 M_j) \phi_r^T \quad \text{and} \quad \frac{\partial \varepsilon_{\phi_r}^2(\boldsymbol{\theta})}{\partial \theta_j} = \mathbf{x}_r^{*T} \mathbf{F}_{r,j} \quad (10)$$

where $\mathbf{F}_{r,j} = -(I - M \phi_r \phi_r^T)(K_j - \omega_r^2 M_j) \phi_r$ and \mathbf{x}_r is given by the solution of the linear system

$$A_r^* X_r = D_r \quad (11)$$

with $D_r = L^T 2\beta_r (\beta_r L\phi_r - \hat{\phi}_r) / \|\hat{\phi}_r\|^2$ and X_r replaced by \mathbf{x}_r . For notational convenience, the dependence of several variables on the parameter set $\boldsymbol{\theta}$ has been dropped. For an $n \times n$ matrix $A_r = K - \omega_r^2 M$, A_r^* is used to denote the modified matrix derived from A_r by replacing the elements of the k -th column and the k -th row by zeroes and the (k, k) element of A_r by one, where k denotes the element of the modeshape vector ϕ_r with the highest absolute value. Also, the n vector \mathbf{b}_r^* is used to denote the modified vector derived from \mathbf{b}_r by replacing the k -th element of the vector \mathbf{b}_r by zero. Also, K_j and M_j in the formulation denote the quantities $\partial K / \partial \theta_j$ and $\partial M / \partial \theta_j$ that can be obtained either analytically or numerically using finite element methods.

Similarly, it can be shown that the (i, j) element of the Hessian of $\varepsilon_{\omega_r}^2(\boldsymbol{\theta})$ and $\varepsilon_{\phi_r}^2(\boldsymbol{\theta})$ can be adequately approximated in the form (assuming that $M_j = 0$)

$$\frac{\partial^2 \varepsilon_{\omega_r}^2(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = \frac{2}{\hat{\omega}_r^4} [\phi_r^T (K_i - \omega_r^2 M_i) \phi_r] [\phi_r^T (K_j - \omega_r^2 M_j) \phi_r] \quad (12)$$

and

$$\frac{\partial^2 \varepsilon_{\phi_r}^2(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = - \frac{2}{\|\hat{\phi}_r\|^2 \|L\phi_r\|^2} (\mathbf{z}_r^{*T} \mathbf{F}_{r,i}) (\mathbf{z}_r^{*T} \mathbf{F}_{r,j}) - \beta_r^2 \|L\phi_r\|^2 \mathbf{F}_{r,j}^{*T} X_r X_r^T \mathbf{F}_{r,i}^* \quad (13)$$

where \mathbf{z}_r is given by the solution of the linear system (11) with $D_r = (I - M^T \phi_r \phi_r^T) L^T (2\beta_r L\phi_r - \hat{\phi}_r)$ and X_r is given by (11) with $D_r = (I - M^T \phi_r \phi_r^T) L^T$.

It is clear that the computation of the first and second derivatives of the square errors for the modal properties of the r -th mode with respect to the parameters in $\boldsymbol{\theta}$ requires only the solutions of the linear system (11), independent of the number of parameters in $\boldsymbol{\theta}$. For a large number of parameters in the set $\boldsymbol{\theta}$, the above formulation for the gradients and Hessian of the mean errors in modal frequencies and in the modeshape components in (4) are computationally very efficient and informative.

Optimal Sensor Location Methodology for Model Parameter Estimation

The information entropy $H(\boldsymbol{\delta}, D)$ [13], introduced as a unique scalar measure of the uncertainty in the estimate of the structural parameters $\boldsymbol{\theta}$, is used for optimizing the sensor configuration in the structure for identifying the parameters in a model class \mathcal{M} . The information entropy depends on the available data D and the sensor configuration vector $\boldsymbol{\delta}$. It has been shown [14] that for a large number of measured data, i.e. as $N_D N \rightarrow \infty$, the following asymptotic result holds for the information entropy for a model class \mathcal{M}

$$H_{\mathcal{M}}(\boldsymbol{\delta}, D) \sim H_{\mathcal{M}}(\boldsymbol{\delta}; \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}) = \frac{1}{2} N_{\boldsymbol{\theta}} \ln(2\pi) - \frac{1}{2} \ln[\det \mathbf{h}_{\mathcal{M}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}; \boldsymbol{\delta})] \quad (14)$$

where $\hat{\boldsymbol{\theta}} \equiv \hat{\boldsymbol{\theta}}(\boldsymbol{\delta}, D) = \arg \min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}; D)$ is the optimal value of the parameter set $\boldsymbol{\theta}$ that minimizes the measure of fit $J_{\mathcal{M}}(\boldsymbol{\theta}; \boldsymbol{\sigma})$ given in (2) or (3) for a model class \mathcal{M} , and $\mathbf{h}_{\mathcal{M}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}) \equiv \mathbf{h}_{\mathcal{M}}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}; \boldsymbol{\delta})$ is an $N_{\boldsymbol{\theta}} \times N_{\boldsymbol{\theta}}$ positive definite matrix defined by (7) and asymptotically approximated by [14]

$$\mathbf{h}_M(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}; \boldsymbol{\delta}) \sim \mathbf{Q}_M(\boldsymbol{\delta}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}) \quad (15)$$

as $N_D N \rightarrow \infty$.

Time History Data. For response time history data, substituting (2) into (7) and considering the limiting case $N_D N \rightarrow \infty$, the resulting matrix $\mathbf{Q}(\boldsymbol{\delta}, \boldsymbol{\theta})$ appearing in (15) simplifies to a positive semi-definite matrix of the form

$$\mathbf{Q}_M(\boldsymbol{\delta}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\sigma}}) = \frac{N_D}{2} \sum_{j=1}^{N_d} \delta_j \frac{1}{\hat{\sigma}_j^2} \mathbf{P}_M^{(j)}(\hat{\boldsymbol{\theta}}) \quad (16)$$

that contains the information about the values of the parameters $\boldsymbol{\theta}$ based on the data from all measured positions specified in $\boldsymbol{\delta}$, while the optimal prediction error variances $\hat{\sigma}^2$ are given by $\hat{\sigma}_j^2 = J_j(\hat{\boldsymbol{\theta}})$. The matrix

$$\mathbf{P}_M^{(j)}(\boldsymbol{\theta}) = \sum_{k=1}^{N_D} \nabla_{\boldsymbol{\theta}} x_j(k; \boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T x_j(k; \boldsymbol{\theta}) \quad (17)$$

is a positive semi-definite matrix containing the information about the values of the parameters $\boldsymbol{\theta}$ based on the data from one sensor placed at the j -th DOF. The prediction error $\hat{\sigma}_j^2 = J_j(\hat{\boldsymbol{\theta}})$ in (16) is computed from $\hat{\sigma}_j^2 = s_1^2 + s_2^2 g_j(\hat{\boldsymbol{\theta}})$, where the first term accounts for constant measurement error and the second term accounts for model error that depends on the strength $g_j(\hat{\boldsymbol{\theta}})$ of the response characteristics with the values of s_1^2 and s_2^2 giving the relative size of measurement and model errors.

Modal Data. For modal data, the resulting matrix $\mathbf{Q}_M(\boldsymbol{\delta}, \boldsymbol{\theta})$ simplifies to a positive semi-definite matrix given by

$$\mathbf{Q}_M(\boldsymbol{\delta}, \boldsymbol{\theta}) = \frac{N_D}{2} \sum_{r=1}^m \left[\frac{\nabla_{\boldsymbol{\theta}} \omega_r(\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T \omega_r(\boldsymbol{\theta})}{s_1^2 + s_2^2 \omega_r^2(\boldsymbol{\theta})} + \sum_{j=1}^{N_d} \delta_j \frac{\nabla_{\boldsymbol{\theta}} L_0 \phi_{jr}(\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}}^T L_0 \phi_{jr}(\boldsymbol{\theta})}{s_1^2 + s_2^2 \|L_0 \phi_r(\boldsymbol{\theta})\|^2 / N_0} \right] \quad (18)$$

containing the information about the values of the model parameters $\boldsymbol{\theta}$ based on the modal data from all sensors placed in the structure.

The asymptotic approximation of the information entropy is useful in the experimental stage of designing an optimal sensor configuration. Specifically, the information entropy for a model class M is completely defined by the optimal value $\hat{\boldsymbol{\theta}}$ of the model parameters and the optimal prediction error $\hat{\sigma}_j^2 = J_j(\hat{\boldsymbol{\theta}})$, $j=1, \dots, N_0$, expected for a set of test data, while the time history details of the measured data do not enter explicitly the formulation. The optimal sensor configuration is selected as the one that minimizes the information entropy [13] with respect to the set of N_0 measurable DOFs. However, in the initial stage of designing the experiment the data are not available, and thus an estimate of the optimal model parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\sigma}}$ cannot be obtained from analysis. In practice, useful designs can be obtained by taking the optimal model parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\sigma}}$ to have some nominal values chosen by the designer to be representative of the system.

Optimal Sensor Location Methodology for Damage Detection

The design of optimal sensor configurations for providing informative measurements for multiple model classes M_1, \dots, M_{μ} is next addressed. Let $I_i(\boldsymbol{\delta}) \equiv IEI_i(\boldsymbol{\delta})$ be the effectiveness of a sensor

configuration δ for the i th model class M_i , where $IEI_i(\delta)$ is the information entropy index given by

$$IEI_i(\delta) = \frac{H_i(\delta) - H_i(\delta_{i,best})}{H_i(\delta_{i,worst}) - H_i(\delta_{i,best})} \quad (19)$$

with $H_i(\delta) \equiv H_{M_i}(\delta; \hat{\theta}, \hat{\sigma})$. The optimal sensor configuration for the model class M_i is selected as the one that minimizes the information entropy index $I_i(\delta)$. In (19), $\delta_{i,best}$ is the optimal sensor configuration and $\delta_{i,worst}$ is the worst sensor configuration for the i th model class. The values of $IEI_i(\delta)$ range from zero to one. The most effective configuration corresponds to value of $IEI_i(\delta)$ equal to zero, while the least effective configuration corresponds to value of $IEI_i(\delta)$ equal to one.

The problem of identifying the optimal sensor locations that minimize the information entropy indices for all μ model classes is formulated as a multi-objective optimization problem stated as follows. Find the values of the discrete-valued parameter set δ that simultaneously minimizes the objectives [15].

$$I(\delta) = (I_1(\delta), \dots, I_\mu(\delta)) \quad (20)$$

For conflicting objectives $I_1(\delta), \dots, I_\mu(\delta)$, there is no single optimal solution, but rather a set of alternative solutions, which are optimal in the sense that no other solutions in the search space are superior to them when all objectives are considered. Such alternative solutions, trading-off the information entropy values for different model classes, are known in multi-objective optimization as Pareto optimal solutions.

An advantage of the multi-objective identification methodology is that all admissible solutions are obtained which constitute model trade-offs in reducing the information entropies for each model class. These solutions are considered optimal in the sense that the corresponding information entropy for one model class cannot be improved without deteriorating the information entropy for another model class. The optimal points along the Pareto trade-off front provide detailed information about the effectiveness of the sensor configuration for each model class.

Optimization Issues for Optimal Sensor Location Problems

Model Parameter Estimation. An exhaustive search over all sensor configurations for the computation of the optimal sensor locations is extremely time consuming and in most cases prohibitive, even for models with a small number of degrees of freedom. Alternative techniques must be used to solve the discrete-variable optimization problem. Genetic algorithms [16] are well suited for solving general optimization problems involving discrete-valued variables. A more systematic and computationally very efficient approach for obtaining a good approximation of the optimal sensor configurations for a fixed number of N_0 sensors is to use a sequential sensor placement (SSP) approach [14]. The computations involved in the SSP algorithm are an infinitesimal fraction of the ones involved in exhaustive search method and can be done in realistic time, independent of the number of sensors and the number of model degrees of freedom. For essentially the same accuracy, GA algorithms require significantly more computational time than the heuristic SSP algorithms. In all example cases considered, the SSP algorithms outperform, in terms of accuracy and computational time, the GA algorithms.

Damage Detection. Genetic algorithms are well suited for performing the multi-objective optimization involving discrete variables. In particular, the strength Pareto evolution algorithm [27] based on genetic algorithms is most suitable for solving the resulting discrete multi-objective optimization problem and providing near optimal solutions [15]. A more systematic and computationally very efficient approach for obtaining a good approximation of the Pareto optimal

front and the corresponding Pareto optimal sensor configurations for a fixed number of N_0 sensors is to use a sequential sensor placement (SSP) approach [15], extending the SSP algorithm [14] to handle Pareto optimal solutions. The total number of vector function evaluations using the extended SSP algorithm is infinitesimally small compared to the number of vector function evaluations required in an exhaustive search method. Numerical applications [14, 15] show that the Pareto front constructed by this heuristic algorithm, in most cases examined, coincides with, or is very close to, the exact Pareto front. In all cases, the extended SSP algorithm outperforms, in terms of accuracy and computational time, available discrete multi-objective optimization algorithms such as the strength Pareto evolution algorithm based [27] on genetic algorithms.

Summary

Optimization problems encountered in integrating system analysis and measured data (response time histories or modal data) collected or identified using a sensor network, are reviewed. The present study concentrates on structural identification problems related to (1) model parameter estimation used for finite element model updating, (2) model-based damage detection in structures and (3) optimal sensor location for parameter estimation and damage detection. A Bayesian methodology was presented to address the first two problems and an information entropy method was used to address the third problem. These problems are finally formulated as single and multi-objective optimization problems involving continuous or discrete-valued variables. Gradient-based, evolutionary, hybrid and heuristic algorithms were proposed to efficiently solve these problems as well as effectively address issues related to multiple local/global solutions and computational complexity. For the first two problems, computationally efficient methods for estimating the gradients and Hessians of the objective functions with respect to the model parameters are shown to significantly reduce the computational effort for solving the optimization problems. Heuristic algorithms are effective for solving the single- and multi-objective optimization problems arising in optimal sensor location formulations. These algorithms are based on sequential sensor placement schemes and outperform, in terms of accuracy and computational efficiency, available GA-based algorithms.

Acknowledgements

This research was co-funded 75% from the European Union (European Social Fund), 25% from the Greek Ministry of Development (General Secretariat of Research and Technology) and from the private sector, in the context of measure 8.3 of the Operational Program Competitiveness (3rd Community Support Framework Program) under grant 03-EA-524 (PENED 2003). This support is gratefully acknowledged.

References

- [1] J.E. Mottershead and M.I. Friswell: *Journal of Sound and Vibration* Vol. 167(2) (1993), p. 347
- [2] H. Sohn, K.H. Law: *Earthquake Engineering and Structural Dynamics* Vol. 26 (1997), p.1259
- [3] C.P. Fritzen, D. Jennewein and T. Kiefer: *Mechanical Systems and Signal Processing* Vol. 12(1) (1998), p. 163
- [4] M.W. Vanik, J.L. Beck and S.K. Au: *Journal of Engineering Mechanics (ASCE)* Vol. 126(7) (2000), p. 738
- [5] C. Papadimitriou: 9th ASCE Joint Specialty Conference on Probabilistic Mechanics and Structural Reliability (2004), Albuquerque, New Mexico.
- [6] H.F. Lam, C.T. Ng and M. Veidt: *Journal of Sound and Vibration* Vol. 305(1-2) (2006), p. 34

- [7] H.F. Lam, L.S. Katafygiotis and N.C. Mickleborough: *Journal of Engineering Mechanics (ASCE)* Vol. 130(1) (2004), p. 34
- [8] A. Teughels and G. De Roeck G: *Archives of Computational Methods in Engineering* Vol. 12(2) (2005), p.123
- [9] C. Papadimitriou, J.L. Beck and L.S. Katafygiotis: *Probabilistic Engineering Mechanics* Vol. 16(2) (2001), p. 103
- [10] P.H. Kirkegaard, R. Brincker: *Mechanical Systems and Signal Processing* Vol. 8 (1994), p. 639
- [11] P. Shah and F.E. Udwadia: *Journal of Applied Mechanics* Vol. 45 (1978), p. 188
- [12] D.C. Kammer: *Journal of Guidance, Control and Dynamics*, Vol. 14 (1991), p. 251
- [13] C. Papadimitriou, J. L. Beck and S.K. Au: *Journal of Vibration and Control*, Vol. 6(5) (2000), p. 781
- [14] C. Papadimitriou: *Journal of Sound and Vibration*, Vol. 278(4) (2004), p. 923
- [15] C. Papadimitriou: *Computer Methods in Applied Mechanics and Engineering*, Vol. 194(12-16) (2005), p. 1655
- [16] J.L. Beck and L.S. Katafygiotis: *Journal of Engineering Mechanics (ASCE)* Vol. 124(4) (1998), p. 455
- [17] L.S. Katafygiotis, C. Papadimitriou and H.F. Lam: *International Journal of Soil Dynamics and Earthquake Engineering* Vol. 17 (1998), p. 495
- [18] K. Christodoulou and C. Papadimitriou: *Mechanical Systems and Signal Processing* Vol. 21 (2007), p. 4
- [19] C. Papadimitriou and L.S. Katafygiotis, in: *Engineering Design Reliability Handbook*, edited by N. Nikolaidis, D.M. Ghiocel, S. Singhal, CRC Press. (2004).
- [20] E. Ntotsios, C. Papadimitriou, P. Panetsos, G. Karaiskos, K. Perros and Ph. Perdikaris: *Bulletin of Earthquake Engineering* (2008), in press.
- [21] J.L. Beck, K.V. Yuen: *Journal of Engineering Mechanics (ASCE)* Vol. 130(2) (2004), p. 192
- [22] K.V. Yuen: EERL Report 2002-03, Caltech, Pasadena.
- [23] A. Teughels, G. De Roeck and J.A.K. Suykens: *Computers and Structures* Vol. 81(24-25) (2003), p. 2337
- [24] H. G. Beyer: *The theory of evolution strategies* (Berlin, Springer-Verlag, 2001).
- [25] R.B. Nelson: *AIAA Journal* Vol. 14(9) (1976), p. 1201
- [26] D.E. Goldberg: *Genetic Algorithms in Search, Optimisation and Machine Learning* (Addison-Wesley, Reading, MA, 1989)
- [27] E. Zitzler, L. Thiele: *IEEE Transactions on Evolutionary Computation* Vol. 3 (1999), p. 257