Université de Liège
Vibrations et Identification des Structures
Département d'Aérospatiale, Mécanique et Matériaux

On the Model Validation in Non-linear Structural Dynamics

Thèse de doctorat présentée en vue
de l'obtention du grade de
Docteur en Sciences Appliquées

par

Gaëtan KERSCHEN
Ingénieur Civil Electro-Mécanicien (Aérospatiale)
Aspirant F.N.R.S

Décembre 2002

Acknowledgments

I wish to acknowledge the Belgian National Fund for Scientific Research (FNRS) for its financial support to this research.

I am deeply indebted to a number of people who helped me pursuing and finishing this doctoral dissertation.

First of all, I would like to express my gratitude to my advisor Professor Jean-Claude Golinval for the constant help, guidance and encouragement provided during the course of my research.

I am pleased to acknowledge my colleagues at the *Laboratoire de Vibrations & Identification des Structures* who created a pleasant working atmosphere. Particular acknowledgment is given to Vincent Lenaerts, with whom I had a lot of helpful discussions.

I have also enjoyed my stays with Professor Brian Feeny at Michigan State University, Dr. François Hemez at Los Alamos National Laboratory, Professor Jens Timmer at the University of Freiburg and Professor Keith Worden at the University of Sheffield. I am especially grateful to François Hemez and Keith Worden for their warm hospitality and for graciously opening their home to me.

I would like to thank Professors Michel Hogge, Rodolphe Sepulchre, Alexander Vakakis, Hendrik Van Brussel, Louis Wehenkel and Keith Worden for serving on my dissertation examination committee. I am grateful to all of them for accepting the heavy task of going through my dissertation. Special thanks go to Professor Louis Wehenkel for his meticulous reading of the thesis draft.

I am in debt to my parents-in-law for their support and for taking care of my wife during my study stays abroad.

Finally and foremost, I would like to express my love and gratitude to Carine for her infinite patience and understanding and for always being there for me.

Contents

Introduction **1**
Motivation . . . 1
Identification techniques . . . 2
State-of-the-art . . . 3
Finite element model updating techniques . . . 6
Linear updating strategies . . . 6
Non-linear updating strategies . . . 9
Outline of the thesis . . . 12

1 Restoring Force Surface Method **14**
1.1 Introduction . . . 14
1.2 Theoretical background . . . 15
1.3 Data processing . . . 16
1.4 Model selection . . . 17
1.4.1 Significance factor . . . 17
1.4.2 Bayesian model screening . . . 17
1.5 Application to an impacting beam . . . 21
1.5.1 Experimental set-up . . . 21
1.5.2 Piece-wise linear beam . . . 21
1.5.3 Bi-linear beam . . . 25
1.6 Application to wire rope isolators . . . 26
1.6.1 Experimental set-up . . . 27
1.6.2 Identification results . . . 27
1.6.3 Validation of the identified model . . . 31
1.6.4 Comparison with the conditioned reverse path method . . . 33
1.7 Concluding remarks . . . 34

2 Frequency Response Function-Based Identification Methods **35**
2.1 Introduction . . . 35
2.2 Conditioned reverse path method . . . 35
2.2.1 Estimation of the underlying linear system properties . . . 37
2.2.2 Estimation of the non-linear coefficients . . . 40

2.2.3 Coherence functions . . . 41
2.3 Non-linear identification through feedback of the outputs method . . . 42
2.4 Application to a beam with a geometric non-linearity . . . 43
2.4.1 Experimental set-up . . . 43
2.4.2 Identification results with the conditioned reverse path method: horizontal set-up . . . 44
2.4.3 Identification results with the conditioned reverse path method: vertical set-up . . . 49
2.4.4 Identification results with the non-linear through feedback of the outputs method method: vertical set-up . . . 53
2.4.5 Comparison of the results . . . 53
2.5 Concluding remarks . . . 56

3 Proper Orthogonal Decomposition 58
3.1 Introduction . . . 58
3.2 Mathematical formulation . . . 60
3.3 Practical computation of the proper orthogonal decomposition . . . 62
3.3.1 Eigensolutions of the sample covariance matrix . . . 62
3.3.2 Singular value decomposition of the response matrix . . . 63
3.4 Physical interpretation of the proper orthogonal modes . . . 64
3.4.1 Undamped and unforced linear systems . . . 65
3.4.2 Damped and unforced linear systems . . . 68
3.4.3 Response of a linear system to harmonic excitation . . . 69
3.4.4 Random response of a linear system to a white noise excitation . . . 74
3.4.5 Linear normal modes, non-linear normal modes and proper orthogonal modes: a geometric approach . . . 75
3.5 Model reduction of non-linear systems . . . 78
3.5.1 Motivation . . . 78
3.5.2 Reduced-order models . . . 79
3.5.3 Numerical application . . . 80
3.6 Model updating of non-linear systems . . . 87
3.7 Concluding remarks . . . 92

4 Non-linear Proper Orthogonal Decomposition 94
4.1 Introduction . . . 94
4.2 Non-linear proper orthogonal decomposition: a global approach . . . 95
4.2.1 Proper orthogonal decomposition and neural networks . . . 95
4.2.2 Five-layer auto-associative neural networks . . . 99
4.2.3 Bias/variance trade-off . . . 103
4.2.4 Training the network . . . 104
4.2.5 Selection of the number of neurons . . . 106

4.2.6 Feature extraction . . . 108
4.2.7 Limitations . . . 110
4.3 Non-linear proper orthogonal decomposition: a local approach . . . 111
4.3.1 Dimension reduction by local linear models . . . 111
4.3.2 Vector quantisation . . . 112
4.3.3 Vector quantisation principal component analysis algorithm . . . 113
4.4 An illustrative example . . . 114
4.4.1 Proper orthogonal decomposition . . . 114
4.4.2 Non-linear principal component analysis . . . 116
4.4.3 Vector quantisation principal component analysis . . . 121
4.5 Concluding remarks . . . 123

5 Parameter Estimation using Optimisation Algorithms 125
5.1 Introduction . . . 125
5.2 Differential evolution . . . 125
5.2.1 Evolutionary computation . . . 126
5.2.2 Description of the algorithm . . . 127
5.2.3 An illustrative example . . . 130
5.2.4 Discussion . . . 132
5.3 Coupling local optimisers with the conditioned reverse path method . . . 132
5.4 A model updating strategy of non-linear vibrating structures . . . 134

6 Model Updating of a Three-dimensional Portal Frame 138
6.1 Introduction . . . 138
6.2 Description of the structure . . . 138
6.3 Experimental measurements and finite element modelling . . . 141
6.4 Feature extraction and correlation study . . . 144
6.5 Parameter selection . . . 151
6.6 Model updating with limited knowledge about the non-linearity . . . 152
6.6.1 Results for the impulse force . . . 152
6.6.2 Results for the random excitation . . . 154
6.6.3 Discussion . . . 155
6.7 Model updating using the conditioned reverse path method . . . 157
6.7.1 Application of the conditioned reverse path method . . . 157
6.7.2 Results for the impulse force . . . 159
6.7.3 Results for the random excitation . . . 161
6.8 Analysis of the results and concluding remarks . . . 162

Conclusions 164
Main contributions . . . 165
Directions for future work . . . 166

A Quantitative Analysis of a Chaotic Signal **168**
A.1 Correlation dimension 169
A.2 Lyapunov exponents 169

List of Acronyms

AIC	Information theoretic criterion
COST	Cooperation in the field of scientific and technical Research
CRP	Conditioned reverse path
DOF	Degree-of-freedom
FDAC	Frequency domain assurance criterion
FEV	Fraction of explained variance
FPE	Final prediction error
FRF	Frequency response function
FT	Fourier transform
LNM	Linear normal mode
LSCE	Least squares complex exponential
MAC	Modal assurance criterion
MCMC	Markov chain Monte Carlo
MDOF	Multi-degree-of-freedom
MSE	Mean square error
NARMAX	Non-linear auto-regressive moving average with exogeneous inputs
NIFO	Non-linear identification through feedback of the outputs
NLPC	Non-linear principal component
NLPCA	Non-linear principal component analysis
NLPOM	Non-linear proper orthogonal mode
NMSD	Normalised mean square distance
NMSE	Normalised mean square error
NNM	Non-linear normal mode
PCA	Principal component analysis
PDF	Probability density function
POD	Proper orthogonal decomposition
POM	Proper orthogonal mode
POV	Proper orthogonal value
PSD	Power spectral density
RFS	Restoring force surface
RMS	Root mean square
RP	Reverse path

SDOF	Single-degree-of-freedom
SIC	Schwarz information criterion
SQP	Sequential quadratic programming
SSE	Sum of squares error
SVD	Singular value decomposition
VQPCA	Vector quantisation principal component analysis
VQPOM	Vector quantisation proper orthogonal mode

Introduction

Motivation

The mathematical modelling of structures is currently one of the most important fields of research in the theory of structural vibration. The purpose is to derive an accurate mathematical model that predicts reliably the dynamic characteristics of the structure. Such a model enables the structural dynamicist to control some important aspects of the design procedure such as the maximum acceptable displacement or the excitation near a natural frequency resulting in an optimised design.

Partial differential equations have been studied over the past 300 years because they provide a natural mathematical description of virtually all phenomena in the physical sciences (e.g., thermodynamics, mechanics). However, their solutions are often extraordinarily complex, and closed analytical expressions can be found only for simple forms of the equations in simple geometric regions.

Thus, scientists and mathematicians have been naturally led to seeking techniques for the approximation of solutions. Due to a confluence of new methods of analysis and the digital computer, numerical algorithms for the approximate solution of partial differential equations have been introduced. Among the numerous methods that have been developed, the finite element method (see e.g., reference [189]) is certainly the most popular and well established numerical technique in structural engineering. It allows for the spatial discretisation of the partial differential equations and yields a system of ordinary differential equations in time. The dynamic response is then obtained through the use of direct time integration techniques.

At present in the industry, most structures are modelled with the assumption that the structural behaviour is linear. Linearity means that the system response to any combination of dynamic loads, simultaneously applied, equals the sum of the individual responses to each of the loads acting separately. In this case, the equation of motion takes the following form:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{f}(t) \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, damping and stiffness matrices respectively;

$\mathbf{x}$, $\dot{\mathbf{x}}$ and $\ddot{\mathbf{x}}$ are the displacement, velocity and acceleration vectors respectively;

$\mathbf{f}$ is the external applied force vector.

As it is of continual interest to expand the performance envelope of structures, they are becoming lighter, more flexible, and consequently, more non-linear. The non-linear behaviour is present in a wide class of engineering problems and can be caused by a variety of different reasons. It can be observed even in rather simple structures like plates and beams, as a result of buckling or large deformation related effects. In the case of an aircraft, typical non-linearities are backlash and friction in control surfaces, joints, stiffening non-linearities in the engine-to-pylon connection plus any underlying distributed non-linearity in the structure. The equation of motion of a non-linear vibrating structure is given by the following expression:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \mathbf{f}_{NL}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \mathbf{f}(t) \tag{2}$$

where vector $\mathbf{f}_{NL}$ is a non-linear function of the displacement and velocity.

If the effect of the non-linearity is not taken into account, the numerical model gives erroneous predictions of the structural behaviour which may in some cases introduce a threat to human life.

In addition, even in the case of simple linear structures and despite the high sophistication of mathematical modelling, practical applications often reveal considerable discrepancies between analytical and experimental results, mainly due to three sources of errors:

1. Modelling errors. They include the simplifications made by the analyst during the modelling process (e.g., approximation of boundary conditions or assumption of proportional damping) but also the discretisation errors that are related to the number of elements, their type and their spatial distribution.

2. Parameter errors. They refer to the inaccuracy of model parameters.

3. Testing errors. Noise or the influence of the test equipment on the structure affect the quality of the measurements. By adopting rigorous testing procedures, it can confidently be assumed that the measured data provides a true representation of the dynamics of the structure.

It is thus of the utmost importance that mathematical models be verified and validated against data from the real structure before using them for tasks such as structural modification analysis or lifetime estimation.

Identification techniques

The goal of system identification is to exploit input and output measurements performed on the real structure using vibration sensing devices to estimate the parameters of the equation of motion of the structure. The identification process for linear systems is now

mature, and the methods almost all yield a final model based upon modal parameters, namely natural frequencies, mode shapes and damping ratios.

As mentioned previously, non-linear behaviour is often encountered in real-life applications, and there is thus considerable interest in the identification of the dynamic characteristics of non-linear structures. This is widely recognised to be a challenging task for the following reasons:

- The basic principles that apply to a linear system and that form the basis of modal analysis are no longer valid. New analysis tools have been developed but involve more advanced mathematics.
- The dynamics of a non-linear system is much richer and much more complex than the dynamics of a linear system. There are phenomena that cannot be encountered in the case of a linear system, e.g., limit cycles, harmonics or chaos.
- There are many relevant sources of non-linearity, e.g., non-linear material behaviour, contact between components or Coulomb friction and energy loss mechanisms.

The identification process involves a progression through three steps, namely characterisation, model selection and parameter estimation and is outlined in Figure 1. A non-linear system is said to be characterised after the type and location of all the non-linearities throughout the system are determined. The purpose of the second step, i.e., model selection, is to describe the non-linear effect using a mathematical model. Polynomial expansion is extensively used in the literature. Finally, the non-linear model parameters need to be determined. This is usually carried out using linear least squares fitting.

State-of-the-art

Non-linear system identification began in 1979 with the introduction of the restoring force surface (RFS) method by Masri and Caughey [126]. Since then, numerous methods have been proposed mainly because it is difficult to find a single method that performs well on a wide variety of non-linearities. It is not our intention to review all the methods available but rather to cite the most popular techniques that have been considered during the last twenty years:

- The first application of the Hilbert transform in the field of vibrating systems was made in the frequency domain [158]. The time-domain Hilbert transform was also utilised to solve an inverse problem [48, 49]. This method presents interesting properties for causal functions and leaves the frequency response functions (FRFs) of a linear system unchanged. This property can be exploited to detect any departure from linearity.

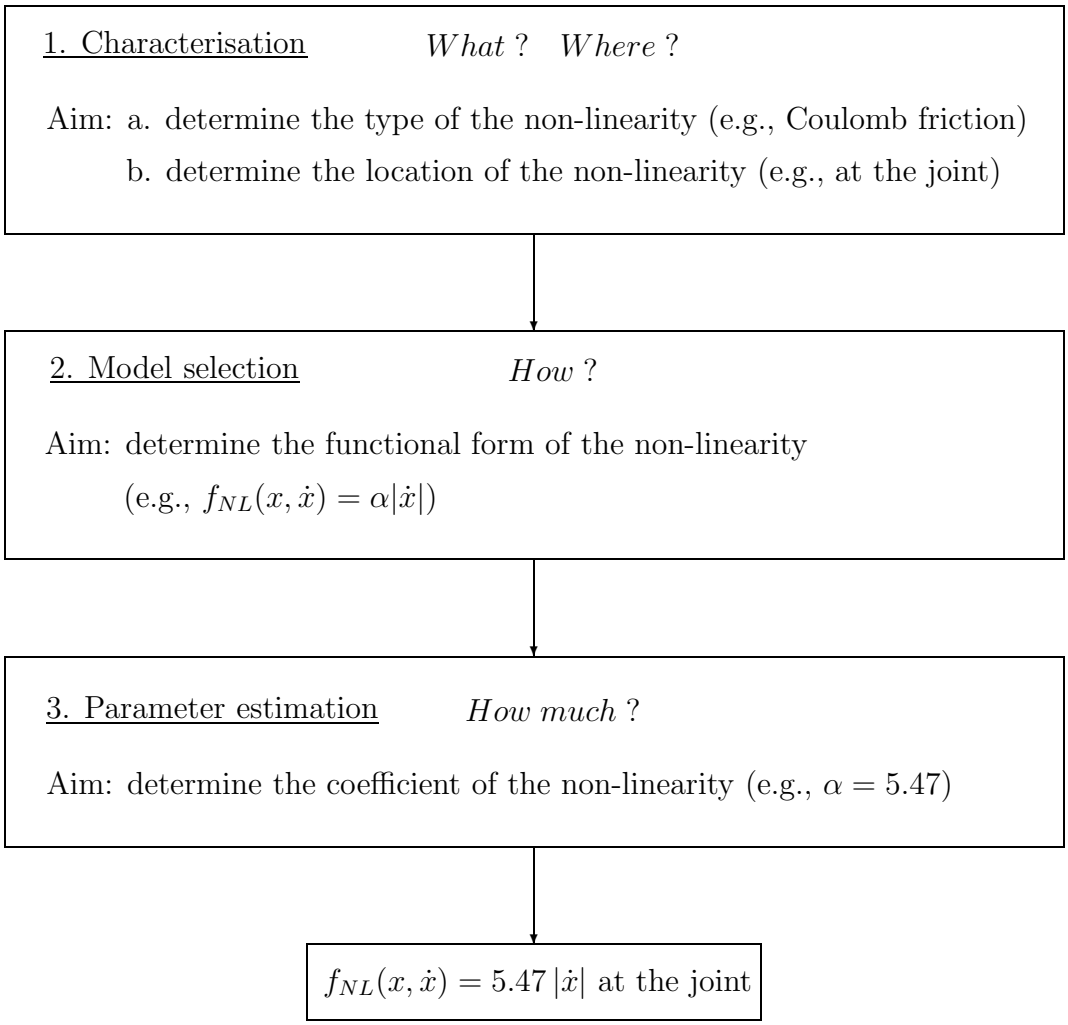


Figure 1: Identification process.

- The Volterra series provide, under some assumptions, an explicit relationship between the input and the output of a non-linear system. The derivation of the series is discussed at length in reference [153]. The use of the Volterra series in the field of structural dynamics began in the late 1980s with the work of Gifford and Tomlinson [61]. Since then, it was also applied by several researchers including Worden and Manson [182, 183], and Kahn and Vyas [82]. However, the method is associated with quite serious drawbacks. For instance, the convergence of the series is not always guaranteed, and it also requires the non-linear functions to be sufficiently differentiable.

- NARMAX models (Non-linear Auto-Regressive Moving Average with eXogeneous inputs) were introduced by Leontaritis and Billings [112, 113] and are based on ARMAX models. An improvement of the original scheme was proposed in reference

[172] where the number of parameters to identify was significantly reduced by using modal co-ordinates to express the NARMAX model. The number of all possible candidate terms may still remain large which renders the identification process complicated and time consuming. It should also be noted that the higher-order FRFs can be computed from the NARMAX model [20, 21].

A considerably more detailed discussion about the foregoing methods is available in reference [185]. However, these methods are mostly limited to systems with a relatively small number of degree-of-freedom (DOF), typically two or three, which represents a severe limitation when investigating real applications.

The development of methodologies capable of dealing with more complex non-linear systems is thus a central issue in structural dynamics. In particular, the identification of continuous systems where a lumped parameter representation is inappropriate needs to be investigated.

The last five years have witnessed impressive progress in this direction:

- An approach using a model expressed in modal space has been proposed by Wright *et al.* [186]. The basic philosophy of this methodology aims at performing a multi-stage identification by considering a single mode or a group of modes instead of treating all the modes of the structure in a single step. To this end, it combines the resonant decay, the force appropriation and the RFS methods.

- The development of frequency response function (FRF)-based approaches has received increasing attention in recent years. In particular, the conditioned reverse path (CRP) and non-linear identification through feedback of the outputs (NIFO) formulations have been introduced to eliminate the distortions in FRFs caused by the presence of a non-linearity. Both methods allow for the estimation of the non-linear coefficients together with the FRFs of the underlying linear system.

The work performed in the framework of the European Cooperation in the field of Scientific and Technical Research (COST) Action F3 working group on "Identification of non-linear systems" (1997-2001) has also stimulated the development of new methodologies:

- In the contribution of Bellizzi and Defilippi [12], the identification of the non-linear system is performed by fitting the analytical approximation of the power spectral density (PSD) function response with the one derived from the measurements. The PSD approximation is defined as the PSD response matrix function of an equivalent linear system with random parameters.

- Gibert and Thouverez [60] have introduced a methodology in which a frequency response model based on the concept of non-linear normal modes is built. The analytical expression of the non-linear compliance is expressed as a linear combination of single mode contributions and depends on complex modal amplitudes. The

non-linear modal model of the structure is computed by fitting test data to the non-linear flexibility functions.

Finite element model updating techniques

Linear updating strategies

Recently, there has been a growth in interest in a particular class of identification techniques based on finite element models and referred to as finite element model updating techniques.

Model updating was first introduced in the context of linear structural dynamics. For a deeper discussion of the work in this field, we invite the interested reader to refer to the literature (e.g., [54, 115] and references therein).

Model updating methods can be broadly classified into global and local methods. Global methods update directly the elements of the structural matrices. The resulting matrices reproduce the dynamic characteristics of the real structure accurately but the corrections suggested are not necessarily physically meaningful. In addition, essential properties of the system such as the structural connectivity or the positive definiteness of matrices may be lost. Local methods only consider the estimation of the uncertain physical parameters such as thickness or Young's modulus and preserves the mathematical properties of the initial model.

The updating procedure, outlined in Figure 2, starts with the model matching step because finite element models require the definition of far more DOF than in the experimental test. In order to solve this mesh incompatibility, either the structural matrices are reduced to the number of measured DOF or the experimental data are expanded to the number of DOF present in the model.

The correlation step aims at deciding whether the model is accurate enough for the purposes for which it is intended. To this end, model updating could involve, in principle, the raw comparison between the time histories of the measured and predicted data. But in many applications, the comparison of features is preferred. Generally speaking, a feature is defined as a variable identified from the physical response that efficiently captures the information contained in the data. The definition of suitable features in a model updating context is motivated by several reasons:

- Estimating the predictive capability of a numerical model based only on its ability to match measured time series may be hazardous. The comparison of features gives valuable information about the system that should be exploited.

- Multi-variate data occupying n dimensions may be highly redundant and are almost never n-dimensional. The elimination of a significant number of dimensions to obtain an efficient representation of the underlying structure should improve the

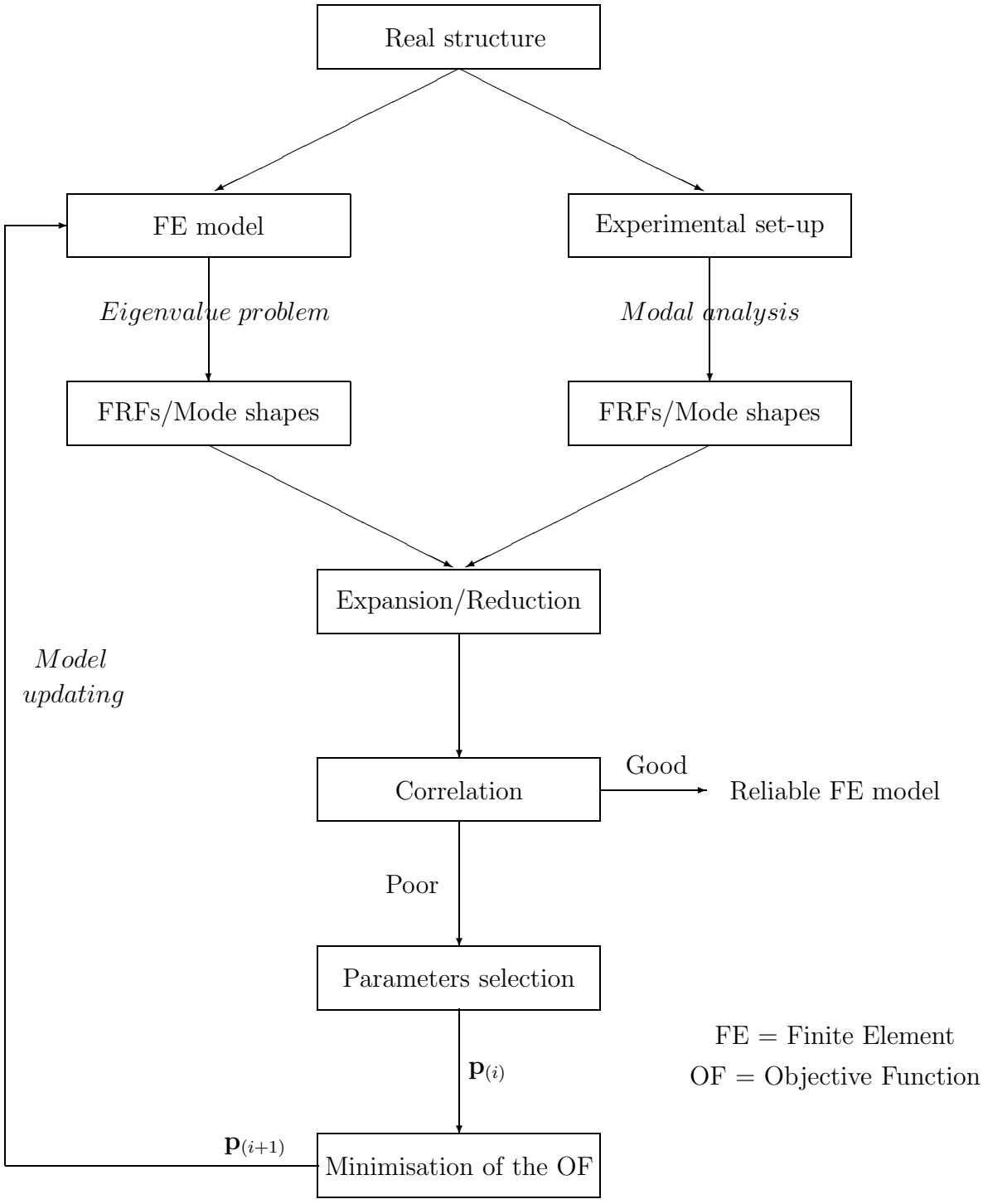


Figure 2: Model updating sequence of linear systems.

performance of the updating algorithm. The fact that such dimensionality reduction can lead to better performance may at first appear as a paradox since in most cases it results in loss of information. Dimension reduction techniques do not create new information, rather they allow relevant information contained in the signal to be enhanced.

- Noise always corrupts the data and features may be less sensitive to it.

It should be noted that features can also been employed for other purposes such as damage detection. For instance, natural frequencies have been exploited as damage indicators (see e.g., reference [23]). Sohn *et al.* [160] have applied an outlier analysis to the auto-regressive coefficients in order to detect any abnormal changes of the system. Trendafilova and Van Brussel [173, 174] have investigated non-linear invariants like Lyapunov exponents for condition monitoring of robot joints.

The most popular correlation techniques make use of natural frequencies ω and mode shapes $\boldsymbol{\phi}$. In this case, the correlation between the test $(\omega_t, \boldsymbol{\phi}_t)$ and the simulation $(\omega_s, \boldsymbol{\phi}_s)$ is given by the frequency deviation Δf and the modal assurance criterion (MAC):

$$\Delta f(\omega_t, \omega_s) = \frac{|\omega_t - \omega_s|}{\omega_t} \quad \text{and} \quad \text{MAC}(\boldsymbol{\phi}_t, \boldsymbol{\phi}_s) = \frac{(\boldsymbol{\phi}_t^T \boldsymbol{\phi}_s)^2}{(\boldsymbol{\phi}_t^T \boldsymbol{\phi}_t)(\boldsymbol{\phi}_s^T \boldsymbol{\phi}_s)} \tag{3}$$

MAC values oscillate between 0 and 1. A unitary value means perfect correlation.

The natural frequencies are always estimated with a high level of confidence and do not suffer from the problem of mesh correspondence between the experimental set-up and the finite element model. The utility of mode shapes relies on their capacity to reduce the amount of information that has to be analysed and in the spatial information they provide about the structure.

Another well established technique is to use data in the frequency domain because the effort of experimental modal analysis is avoided, and averaging to reduce noise effects is straightforward. The frequency domain assurance criterion (FDAC) introduced in reference [139] can be regarded as equivalent to MAC in the frequency domain.

If correlation is not satisfactory, the finite element model is to be corrected. The correction of the model begins with the selection of the updating parameters. In this context, error localisation techniques, sensitivity analysis and engineering judgement are helpful. New values for the updating parameters are computed through the minimisation of an objective function J:

$$\min_{\mathbf{p}} J = \|R(\mathbf{p})\|^2 \tag{4}$$

where vector $\mathbf{p}$ contains the updating parameters. Many choices for the residue $R(\mathbf{p})$ are available (e.g., residual energy [39, 140] or residual force [70] methods), the simplest of all being the difference between the test and analytical modal parameters:

$$\|R(\mathbf{p})\|^2 = \frac{1}{2r} \sum_{i=1}^{r} (1 - \text{MAC}_{(i)}) + \Delta f(\omega_{t(i)}, \omega_{s(i)}) \tag{5}$$

where $\text{MAC}_{(i)}$ refers to the paired MAC value corresponding to the ith experimental mode, and r to the number of modes considered in the study.

The objective function is generally non-linear with respect to the updating parameters, and it is necessary to use an iterative procedure in order to update the model.

Non-linear updating strategies

The success of linear model updating mainly relies on two factors:

1. Natural frequencies and mode shapes have the capability to reduce the amount of information that has to be analysed and provide useful insight into the dynamics of interest;

2. The initial finite element model may often be assumed to be powerful enough to avoid being too far from the global optimum of the objective function. Under this assumption, the use of local optimisation techniques like the conjugate gradient method seems to be reasonable although global optimisers have also been investigated, as in reference [114] for instance.

But our intention is to focus on finite element model updating techniques of non-linear structures. Among the earliest approaches that have tackled this problem, the following studies can be cited:

- The minimum model error estimation algorithm is exploited in reference [40] to produce accurate models of non-linear systems. In this algorithm, a two-point boundary value problem is solved in order to obtain estimates of the optimal trajectories together with the model error.

- The hysteretic behaviour of the friction force is frequently encountered in practice, particularly for machine tools [6, 144, 169]. When investigating hysteretic systems, internal variables may be needed for the model description which rules out conventional identification methods. This is the reason why, in reference [103], non-linear system identification is realised through the minimisation of an objective function based on the difference between the measured and predicted responses. The optimisation is achieved using the differential evolution algorithm that belongs to the class of genetic algorithms.

- The formulation proposed by Meyer and Link [129] constitutes an extension of linear model updating methods. It involves a linearisation of the non-linear equilibrium equations of the structure using the harmonic balance method. This procedure results in a linearised model description in the frequency domain. An initial model is then used in combination with techniques to update the parameters describing the linear and non-linear behaviours. This is carried out by minimising the deviations between measured and predicted displacement responses.

- A statistics-based model updating and validation philosophy is proposed in references [41, 154]. The motivation for including statistical analysis is mainly driven by the desire to account for the effects of environmental and experimental variability. The demonstration application consists in analysing the response of a steel/polymer foam assembly during a drop test. The peak response and time of arrival are defined as features in order to study the transient dynamics of this visco-elastic material. The feature comparison is implemented using metrics such as Mahalanobis distance and Kullback-Leibler relative entropy function. It should also be noted that the finite element model is replaced by an equivalent meta-model with a much smaller analytical form. This strategy aims at reducing the number of detailed, costly computer simulations required during optimisation while maintaining the pertinent characteristics of the problem.

For a comprehensive literature review, the reader is referred to references [72, 73].

The reason why the literature about this subject matter is rather sparse is that the model updating strategies developed in the linear case are no longer useful. Indeed, two severe problems are introduced when investigating non-linear structures:

1. Feature extraction

The validity of modal analysis is limited to structures showing a linear behaviour. Accordingly, natural frequencies and mode shapes can no longer be exploited.

An extension of the concept of mode shapes to discrete conservative non-linear systems was proposed by Rosenberg in references [150, 151]. He defined a non-linear normal mode (NNM) as a synchronous periodic oscillation where all points pass through the equilibrium position at the same time. Other methodologies [156, 157] have been developed for the study of non-conservative, gyroscopic and continuous systems. These methods define NNMs in terms of motion that occurs on two-dimensional invariant manifolds in the system phase space. While the NNMs provide a valuable theoretical tool for understanding certain dynamic phenomena such as mode bifurcations or non-linear mode localization [175, 176], their computation still remains difficult for practical applications. It is therefore a little too early to tell whether they may be of substantial help in a model updating context.

Progress in the definition of features for non-linear systems has been realised recently with the introduction of the proper orthogonal decomposition (POD) in structural dynamics (see e.g., references [35, 50, 120]). The modes extracted from the POD and referred to as proper orthogonal modes (POMs) have the nice property to be optimal for the reconstruction of the dynamic response of a system. They also provide important insight into the spatial behaviour of the system. This is the reason why the POMs have been exploited as a promising alternative to the mode shapes in a model updating context [66, 71, 109, 111]. In reference [109] and in a companion thesis [111], the proposed model updating procedure relies on the solution of an optimisation problem that consists in

minimising the difference between the POMs of the measured and predicted data. The Wavelet transform of their time modulations is also included in the objective function. However, it should be borne in mind that the POD is still a linear technique. The linear nature of the method is appealing because the theory of linear operators is available but it also represents its major limitation when the data lies on a non-linear manifold.

2. Minimisation of the objective function

A priori knowledge about the non-linearity is often limited, and the initial model cannot generally be assumed to be near the actual model. In that case, a local optimiser is likely to be trapped in a local optimum resulting in a poor solution. Global optimisation algorithms have been introduced but are far less developed than their linear counterpart.

Our major concern is thus to address the two aforementioned problems in order that the structural dynamicist may handle model updating of non-linear structures more efficiently.

Firstly, the definition of features that accurately characterise the behaviour of the system is thoroughly investigated. This study parallels the work realised in references [107, 108, 109, 110, 111] with the POD. But in this thesis, a strong emphasis is also put on the non-linear generalisations of the POD. It is verified that the proposed features enable the formulation of appropriate objective functions showing a smooth topological search space with a small number of local minima around the global minimum. In addition, they should guarantee sufficient sensitivity with respect to the updating parameters.

Secondly, the performance of a global optimisation method called differential evolution is analysed. This method belongs to the class of evolutionary computation. It is also inquired whether the computational load imposed by this algorithm is prohibitive in the particular context of model updating. Another means of tackling the minimisation of the objective function is to exploit the CRP method. It should help to alleviate our lack of a priori knowledge about the non-linearity and should enable the determination of tight constraints. In that case, it is investigated whether the use of a local optimiser is appropriate.

Outline of the thesis

The objective of this doctoral dissertation is to generate accurate and robust mathematical models of non-linear vibrating structures through the development and implementation of identification and finite element model updating techniques. The thesis consists of two parts.

The first part (chapters 1 and 2) focuses on identification techniques. The key advantage of these methods is that they provide an estimation of the non-linearity without building a finite element model of the structure. They require less assumptions in comparison with a model updating procedure.

Chapter 1 makes a brief review of the RFS method. This method has been chosen for its simplicity and reliability for identification of single-degree-of-freedom (SDOF) systems. In collaboration with Dr. François Hemez and Mandy Cundy from Los Alamos National Laboratory, a Bayesian screening method has been developed to determine an appropriate model of the restoring force. The identification of two experimental structures is then realised. The first study concerns an impacting beam and was carried out while visiting Professor Keith Worden at the University of Sheffield (U.K.). This benchmark was considered in order to analyse the contact phenomenon frequently encountered in the industry. The second structure was proposed within the framework of the European COST Action F3. It consists of resilient mounts used for vibration isolation in industrial applications.

Chapter 2 describes two frequency domain identification techniques, namely the CRP and NIFO formulations. They deserve our attention because they have the ability to identify more complex non-linear systems. In this context, these two techniques and the RFS method complement each other quite well. Their overall performances are assessed by means of a continuous multi-degree-of-freedom (MDOF) structure with a geometrical non-linearity. This example is challenging regarding the applications currently studied in the literature. A discussion about the respective merits of each technique is also proposed.

The second part of the thesis (chapters 3 to 6) deals with finite element model updating strategies.

Chapter 3 introduces the POD and its characteristics that are relevant in the context of this study. In particular, a physical interpretation of the modes extracted from the decomposition is proposed. The utility of the method for model reduction purposes has been studied while visiting Professor Brian Feeny at Michigan State University (MI, U.S.A.). The application of the POD to model updating is then discussed.

Chapter 4 is concerned with non-linear generalisations of the POD. It is our intention to take advantage of the work performed within the statistical community to define features that allow for non-linear dependencies between the spatial co-ordinates. The benefits of two approaches, namely non-linear principal component analysis (NLPCA) and vector quantisation principal component analysis (VQPCA), are explored. These methodologies are applied to a synthetic data set sampled from a NNM motion. Synthetic data is used

to develop intuition about their implementation.

Chapter 5 discusses the use of differential evolution and of local optimisers coupled with the CRP method to accurately locate the global minimum of the objective function. A model updating strategy of non-linear vibrating structures based upon the developments made in this chapter and in chapters 3 and 4 is then presented.

Chapter 6 aims at assessing the performance of the proposed model updating strategy using simulated data from a three-dimensional portal frame. The example is also exploited to determine the respective merits of the features extracted from the POD and its non-linear generalisations.

Finally, conclusions regarding the completed research and the associated contributions to the field are drawn. A discussion of the ways in which this research may be extended is also given.

Chapter 1

Restoring Force Surface Method

1.1 Introduction

An efficient and reliable identification technique of non-linear mechanical structures that is not confined to pure detection was introduced with the restoring force surface (RFS) method. An equivalent method, referred to as force-state mapping, was proposed independently by Crawley, Aubert and O'Donnell [33, 34]. Masri *et al.* laid down the foundations of the method [126] for single-degree-of-freedom (SDOF) systems, and significant improvements were brought about since the original papers. Al-Hadid and Wright [3] proposed a sensitivity approach for the estimation of the mass or modal mass prior to identification. The Chebychev polynomial representation was replaced by the use of ordinary polynomials in reference [2]. Elaborated interpolation procedures used to overcome the problem of an inadequately covered state plane were presented by Worden and Tomlinson [179]. Duym and Schoukens [42] designed optimised excitation signals in order to guarantee the quality of the fit by uniformly covering the phase plane. They also used a local non-parametric identification of the restoring force [43].

Since the technique is well documented in the literature (see e.g., references [44, 111, 185]), the next section only aims at recalling the key aspects of the RFS method. In particular, only the formulation for SDOF systems is described although the method can be extended to identify multi-degree-of-freedom (MDOF) non-linear systems [127]. This is realised by assuming that the modal matrix is known and by transforming the measured data into modal space. The main drawback of this approach is that the non-linearity is responsible for a modal coupling, and the key advantage of the method that lies in the two-dimensional representation of the restoring force for SDOF systems is lost.

The identification of two experimental structures is realised. The first structure is an impacting beam. This benchmark was considered in order to investigate the contact phenomenon that is frequently encountered in the industry. To the best of our knowledge, only one previous paper [181] tried to identify such structure using the RFS method but poor instrumentation forced bad results.

The second structure consists of resilient mounts used for vibration isolation in indus-

trial applications. It was proposed within the framework of the European COST Action F3 working group on "Identification of non-linear systems" which allows to compare our results with those obtained by other research groups.

1.2 Theoretical background

The RFS method is based on Newton's second law:

$$m\ddot{x}(t) + f_{NL}(x(t), \dot{x}(t)) = f(t) \tag{1.1}$$

where $f(t)$ is the applied force and $f_{NL}(x(t), \dot{x}(t))$ is the restoring force, i.e., a non-linear function of the displacement and velocity. The time histories of the displacement and its derivatives, and of the applied force are assumed to be measured. In practice, the data should be sampled simultaneously at regular intervals. From equation (1.1), it is possible to compute the restoring force:

$$f_{NL(i)} = f_{(i)} - m\,\ddot{x}_{(i)} \tag{1.2}$$

where subscript i refers to the ith sampled value. At each sampling instant, a triplet $(x_{(i)}, \dot{x}_{(i)}, f_{NL(i)})$ is found, i.e., the value of the restoring force is known for each sampled point $(x_{(i)}, \dot{x}_{(i)})$ in the phase plane.

The next step consists in describing the system by means of a mathematical model. The usual way is to fit to the data a polynomial model of the form:

$$f_{NL}(x, \dot{x}) = \sum_{i=0}^{m} \sum_{j=0}^{n} \alpha_{ij}\, x^i\, \dot{x}^j \tag{1.3}$$

Let us assume that N samples of the restoring force are collected in a vector $\mathbf{f}_{NL}$. Equation (1.3) can be written in the following matricial form:

$$\mathbf{f}_{NL} = \mathbf{X}\,\boldsymbol{\alpha} \tag{1.4}$$

where matrix $\mathbf{X}$ evaluates the $(m+1)\times(n+1)$ effects $x^i\,\dot{x}^j$ for each N observations, and the unknown coefficients α_{ij} are contained in vector $\boldsymbol{\alpha}$. The commonly encountered method of fitting the coefficients $\boldsymbol{\alpha}$ is to define an objective function that represents the prediction error and minimise it. The most straightforward choice is to adopt the Euclidean norm of the prediction error, in which case the best, linear, unbiased estimator of the coefficients $\boldsymbol{\alpha}$ is provided by:

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\,\mathbf{f}_{NL} \tag{1.5}$$

To have some means of assessing the error between the measured value $f_{NL(i)}$ and the predicted value $\hat{f}_{NL(i)}$ of the restoring force, the normalised mean square error (NMSE) indicator is defined as:

$$NMSE(\hat{f}_{NL}) = \frac{100}{N\,\sigma^2_{f_{NL}}} \sum_{i=1}^{N} (f_{NL(i)} - \hat{f}_{NL(i)})^2 \tag{1.6}$$

where $\sigma^2_{f_{NL}}$ is the variance of the measured restoring force. Experience shows that an NMSE value of less than 5% indicates good agreement while a value of less than 1% reflects an excellent fit.

1.3 Data processing

From the foregoing developments, it appears that the method requires to measure displacement, velocity, acceleration and force time histories at each degree-of-freedom (DOF). A pragmatic approach to the procedure requires that only one signal should be measured, and the other two should be estimated from it. Numerical integration and/or differentiation may be adopted [180].

Since acceleration signals are mostly measured in practice, the solution considered in the context of this study is to numerically integrate the acceleration to find velocity and displacement. Various methods for achieving integration exist, i.e., trapezium rule, Simpson's rule, frequency domain integration, and so forth. There are two main problems associated with the integration, namely the introduction of low- and high-frequency components in the signals. The trapezium rule only suffers from the introduction of low-frequency components and does not require the use of a low-pass filter. In addition, it is the simplest integration process and offers considerable saving of computation time. For these reasons, the trapezium rule is considered here.

It can be shown [65] that the transfer function of the trapezium rule is:

$$H(\omega) = \frac{FT(\text{estimated results})}{FT(\text{true results})} = \cos(\omega/2)\,\frac{\omega/2}{\sin(\omega/2)} \tag{1.7}$$

where FT is the Fourier transform and ω the normalised frequency, i.e., the frequency of interest divided by the sampling frequency. Equation (1.7) means that the trapezium rule only integrates constant signals without error and underestimates the integral at all other frequencies. Therefore, the sampling frequency must be chosen high enough in order that the highest frequency of interest is characterised by a low normalised frequency. A sampling frequency ten or twenty times higher than the highest frequency of interest seems to be a reasonable choice.

Since the trapezium rule basically acts as an amplifier of the low-frequency components, the integrated signals are to be high-pass filtered. High-pass filtering with cut-off $k/2N\Delta t$ is equivalent to a polynomial trend removal of order k where N is the number of points and Δt the sampling interval (cf. reference [180]). Choosing a cut-off frequency higher than 0 Hz imposes the filtered signals to be zero mean because a polynomial trend of order 0, i.e., a constant, is removed. This leads to inaccurate estimation of velocity and displacement of asymmetrical systems, and particularly when the non-linearity is bi-linear (see section 1.5.3).

1.4 Model selection

The problem of model selection is one of the common issues in the field of identification of non-linear systems. The purpose is to model the relationship between the response variable and a subset of predictor variables, possibly with interactions between these latter. Generally speaking, there is uncertainty about which subset to use.

With the aim of determining an appropriate functional form for the restoring force $f_{NL}(x(t), \dot{x}(t))$, it may be plotted in a three-dimensional graph against the displacement $x(t)$ and the velocity $\dot{x}(t)$. However, a more suitable representation of the restoring force is obtained by taking the slice of this three-dimensional plot along the axes where either $x(t)$ or $\dot{x}(t)$ is equal to zero. If the slice where $x(t) = 0$ ($\dot{x}(t) = 0$) is considered, the plot is called the stiffness curve (damping curve) and gives the form of the stiffness (damping) non-linearity. Therefore, prior to identification, the RFS approach offers a very interesting way of visualising the non-linearity.

1.4.1 Significance factor

Several models may seem equally likely, and the analyst has to go through the process of fitting each model and assessing their goodness-of-fit through the NMSE. A common means of deciding which terms may be safely discarded in equation (1.3) is by making use of the significance factor s_θ:

$$s_\theta = 100\,\frac{\sigma_\theta^2}{\sigma_m^2} \tag{1.8}$$

where σ_m^2 corresponds to the variance of the sum of all variables in the model, and σ_θ^2 is the variance of the predictor variable. Roughly speaking, the significance factor represents the percentage of the contribution of the variable to the model variance.

1.4.2 Bayesian model screening

Another important issue is to estimate the posterior probability of a particular model as opposed to simply relying on the goodness-of-fit. Posterior probability and goodness-of-fit complement each other because the former indicates whether the analyst's prior opinion of the form of the model is consistent with the evidence.

Model screening identifies the most probable models based on a family of models defined by the analyst and reference data that the predictions must reproduce with the highest possible fidelity. Model screening does not necessarily identify the best model but rather ranks potential models according to their posterior probability of occurrence. The present study exploits a technique based on the Bayesian definition of probability for identifying promising subset of predictor variables [27, 30, 56, 187]. While the frequentist interpretation defines probability strictly as the number of occurrences of an event amongst a collective of possibilities, the Bayesian approach defines probability as

the subjective opinion of the analyst. The work presented in this section results from a collaboration with Dr. François Hemez [96].

Bayesian model screening starts by defining a family of models. This is achieved by defining various effects (e.g., a cubic stiffness x^3) and how these effects are allowed to interact to form the population of potential models. The model-forming rule is that linear and quadratic interactions are allowed between the effects. The second step of the procedure is to assign the prior probability of occurrence of each effect. The priors can reflect empirical observations, experience or the analyst's knowledge of the system investigated. When there is no a priori knowledge of the system, probabilities of occurrence are set to a uniform level for all effects.

The next step is to let the Bayesian screening method find the most appropriate models among all possible combinations of effects. The most probable models are those that best reproduce the reference data. This can be assessed using a conventional root mean square error between data and predictions. Assuming Gaussian distributions, the root mean square error becomes proportional to the likelihood function $L(f_{NL}|\alpha)$ that estimates the probability that the model is appropriate given the available data:

$$L(f_{NL}|\alpha) = \sum_{k=1}^{N} \left(f_{NL(k)} - \hat{f}_{NL(k)} \right)^2 \tag{1.9}$$

Once the likelihood of a particular model has been estimated, the posterior probabilities of the model effects can be updated according to the Bayes theorem. This theorem states that the posterior probability density function $PDF(\alpha|f_{NL})$ is equal to the likelihood function $L(f_{NL}|\alpha)$ multiplied by the prior probability $PDF(\alpha)$ and divided by the probability of the data $PDF(f_{NL})$:

$$PDF(\alpha|f_{NL}) = \frac{L(f_{NL}|\alpha)\, PDF(\alpha)}{PDF(f_{NL})} \tag{1.10}$$

The probability of the observed data $PDF(f_{NL})$ is generally kept constant and omitted in the updating of equation (1.10). Because the procedure is iterative in nature, the Bayes update is repeated and posteriors of the ith iteration become the priors of the $(i+1)$th iteration. All models visited are kept in memory and, once enough samples have been drawn, the probability of occurrence of each model is estimated by the frequency of occurrence, i.e., the ratio between the number of times each model is visited and the total number of models visited.

It remains to explain how the unknown likelihood function $L(f_{NL}|\alpha)$ is sampled. The problem of exploring an unknown PDF is solved with the Markov Chain Monte Carlo (MCMC) algorithm. The MCMC sampling is advantageous because it can sample any distribution, whether it is Gaussian or not. The MCMC sampling can be viewed conceptually as an optimisation solver that performs a random walk through the optimisation space. More appropriate solutions are guaranteed more frequent visits because the acceptance criterion of a given solution is based on its likelihood function. Each candidate

point is accepted or rejected based on its value of the likelihood function and a Chi-square test. This particular acceptance criterion implies that inappropriate models have a small chance of being accepted just like appropriate models have a small chance of rejection. If rejected, a new point is randomly selected in the neighborhood of the last accepted point. The sampling procedure used in this study is the Gibbs sampling, the simplest of the many variants of the MCMC algorithm. Gibbs algorithm samples one direction of the design space at a time, which makes for simpler numerical implementation.

In summary, Bayesian model screening provides the probabilities of occurrence of the most appropriate members of a user-defined family of models, their goodness-of-fit indicators and the posterior probabilities of effects involved in the most likely models. This Bayesian inference approach should thus be useful for giving insight into the models defining the non-linearity.

To illustrate the overall performance of the model screening procedure, consider an output variable y defined by the following input-output model:

$$y = 2\sin(2t) + 3\cos(t) - 1.5\sin(3t)\cos(2t) \tag{1.11}$$

where t is an input variable that varies from zero to fifty with increments of $\Delta t = 0.05$. It is assumed that the model form shown in equation (1.11) is unknown. Instead, observations $y_{(k)} = y(k\Delta t)$, for $k = 0...100$, are obtained and the problem consists in identifying the numerical model that best matches the observed data.

Consider a set of candidate predictors:

$$\begin{cases} x_1 = \sin(t) \\ x_2 = \cos(t) \\ x_3 = \sin(2t) \\ x_4 = \cos(2t) \\ x_5 = \sin(3t) \\ x_6 = \cos(3t) \end{cases} \tag{1.12}$$

In addition to the six predictors of equation (1.12), six other predictors labelled x_7, x_8, x_9, x_{10}, x_{11} and x_{12} are defined as random functions. It can be observed that, if the functional form of the output variable y was known, it could be written as:

$$y = 3x_2 + 2x_3 - 1.5x_4x_5 \tag{1.13}$$

Clearly, y does not depend on predictors x_1, x_6, x_7, x_8, x_9, x_{10}, x_{11} and x_{12}. The objective of model screening is to identify the model form (1.13). Equivalently, it can be stated that the objective of model screening is to identify the linear effects x_2, x_3 and the linear interaction effect $x_4\,x_5$ from all the potential combinations defined by the family of models considered.

The family of models defined for this illustration is composed of the linear models that include the twelve linear effects x_i and the linear interaction models, defined as the

Model	Effects	Posterior probability
1	x_2, x_3, x_4x_5	52.0%
2	x_2, x_3, x_4x_5, x_{10}	3.0%
3	x_2, x_3, x_4x_5, x_3x_4	2.0%
4	$x_2, x_3, x_4x_5, x_2x_{10}$	2.0%
5	x_2, x_3, x_4x_5, x_2x_5	2.0%

Table 1.1: Top five models and number of appearances.

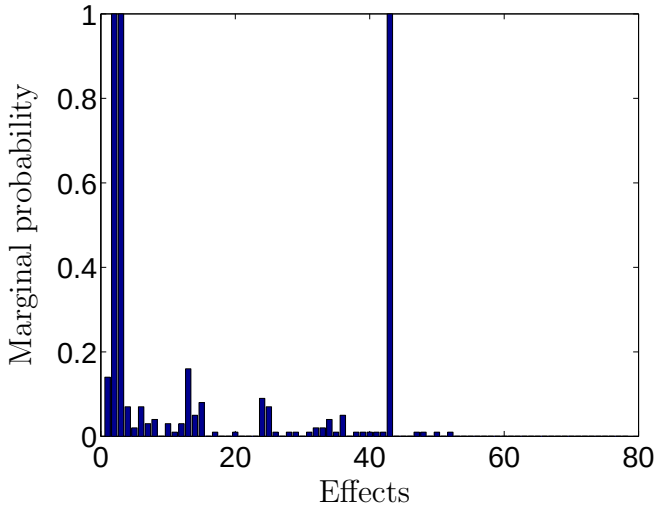


Figure 1.1: Marginal posterior probability of each effect included in the family of models.

previous models augmented with the 66 interaction effects x_ix_j. The total number of different effects x_i and $x_i\,x_j$ with twelve predictors is therefore equal to 78. The total number of different models that can be defined belonging to this family by combining the 78 effects is in excess of $3.02\,10^{23}$ models. Clearly, exploring such a large number of combinations without focusing on the models of highest likelihood would not be feasible.

The procedure described in the foregoing section is applied to the data using 50 samples dedicated to the initialisation of the Gibbs sampler and 100 samples for the computation. Initialising the Markov chain is referred to as "burn-in" and guarantees that the remainder of the chain is not biased due to a particular choice of starting point. The samples drawn during burn-in are disregarded and only the 100 samples drawn during the optimisation itself are kept to estimate the final probability of occurrence of each model in the family. The top five models are listed in Table 1.1. It can be observed that the best model in terms of posterior model probability is the actual model. The NMSE for the top five models is about 0.003%. This means that it is not necessary to include other terms than the ones present in the best model.

Figure 1.1 represents the marginal posterior probability of each effect being in a

particular model. The prior probabilities are set to 25% for each linear effect x_i, 10% for the interaction effects $x_i x_j$ if one of the parent effect x_i or x_j is selected in the model, and 1% only for the interaction effects $x_i x_j$ when neither x_i nor x_j are considered in the model. These uniform probabilities reflect the fact that the analyst has little prior knowledge about the form of the model. It can be observed that effects 2, 3 and 43, namely x_2, x_3 and $x_4 x_5$, are associated with a probability of 100% while the other effects may be ignored because their posterior probabilities are reduced to insignificant levels.

In conclusion, the Bayesian model screening clearly suggests a model that includes the three effects x_2, x_3 and $x_4 x_5$. The identified coefficients corresponding to these effects are equal to 2.99, 2.02 and -1.52, respectively, and they are in good agreement with the actual coefficients shown in equation (1.13).

1.5 Application to an impacting beam

The purpose of the present section is to realise the identification of an experimental structure consisting of an impacting beam by making use of the RFS method. This study results from a collaboration with Professor Keith Worden, and the set-up was prepared with the help of Andreas Kyprianou. Even if the beam is an MDOF system, the study focuses on the vibrations of the first mode. Indeed, the beam can be reduced to an SDOF system while keeping its non-linear characteristics.

A numerical analysis was also performed in reference [89]. In particular, the influence of the noise on the accuracy of the identification was discussed. For the sake of brevity, only the experimental results are presented in this dissertation.

1.5.1 Experimental set-up

The beam is mounted vertically with a clamped end and a free end as shown in Figure 1.2(a). The non-linearity is realised by the presence of bushes. If the amplitude of the transverse motion exceeds the clearance value, the beam makes contact with a steel bush [see Figure 1.2(b)]. When contact occurs, the effective length of the beam is lowered and there is an increase in stiffness. The vibrations of this structure are investigated for two different kinds of non-linearity, i.e., a bi-linear or piece-wise linear stiffness depending on the presence of one or two bushes respectively.

1.5.2 Piece-wise linear beam

The first mode of the beam is located around 18 Hz. If a band-limited white noise centred on this frequency is used, the beam is expected to behave as an SDOF system. With this assumption in mind, a white noise sequence band-limited in the 10-25 Hz range is produced [Figure 1.3(a)]. The excitation point is represented in Figure 1.2(c). The shaker is attached to the beam by a rigid link and the input force is measured using a

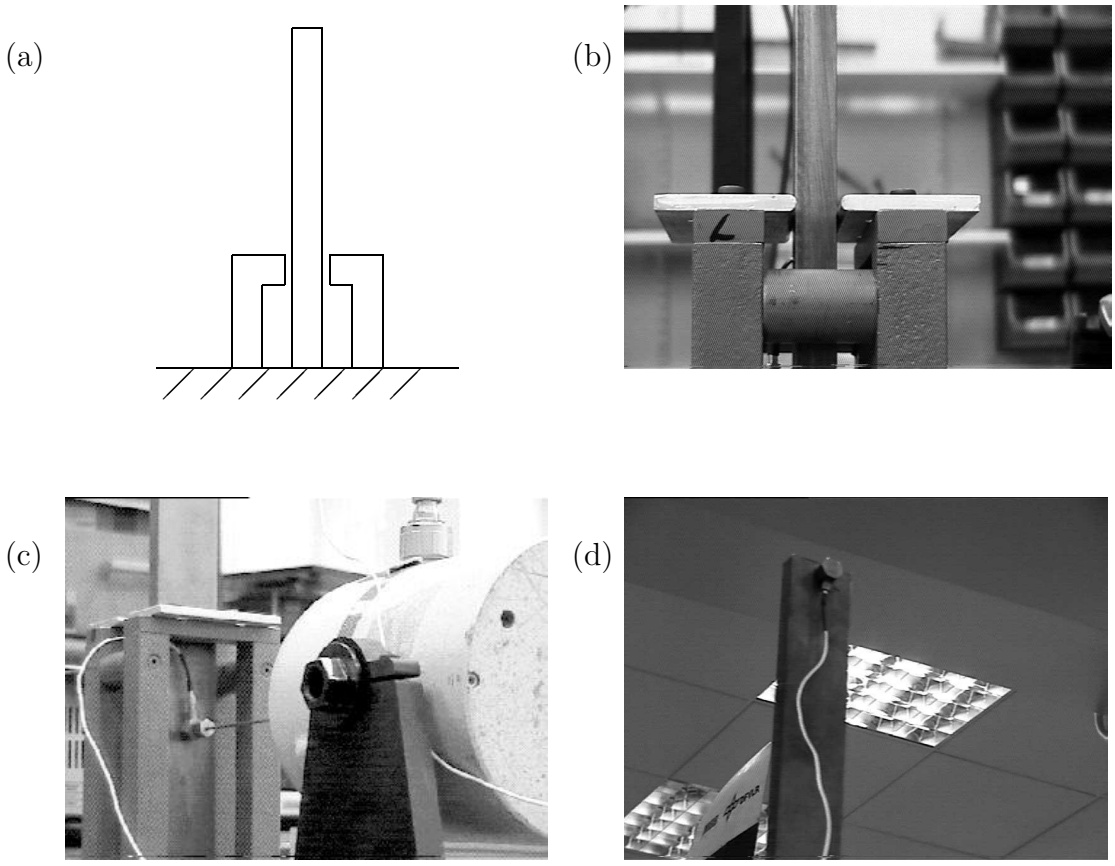


Figure 1.2: (a) The piece-wise linear beam; (b) the two bushes; (c) excitation point; (d) accelerometer location.

force transducer. The response is measured with an accelerometer at the beam tip [see Figures 1.2(d) and 1.3(b)]. Finally, the data is acquired with a sampling frequency set to 1000 Hz. The sampling may appear too high while the frequencies of interest are below 100 Hz. It is thus worth recalling that the sampling frequency should be chosen at least ten times higher than the highest frequency of interest in order to give an accurate integration.

The acceleration is integrated once to give velocity and twice to obtain displacement. The integrated signals are filtered using a high-pass Butterworth filter with cut-off at 10 Hz. The use of a low-pass filter is also necessary, as shown in Figure 1.4(a). This plot illustrates the evolution of the NMSE as a function of the cut-off frequency of the low-pass filter. It can be noticed that the NMSE greatly increases from frequencies around 100 Hz. This results from the participation of the second mode (around 115 Hz) in the system response.

Despite the band-limited excitation signal, it is thus difficult to excite only the first mode. The contributions of the higher modes are to be discarded keeping in mind that it is important to capture as many harmonics of the first mode as possible. In this example,

the choice of a 70 Hz cut-off frequency is found to be the best compromise. Figure 1.4(b) compares the power spectral density (PSD) of the acceleration before and after filtering. The contribution of the second mode is well filtered out while the harmonic of the first mode is still present.

An estimation of the mass should be provided in order to compute the restoring force. To overcome this problem, a solution consists in including the mass in the model and instead of fitting the restoring force, the applied force is fitted. Another solution is to perform a linear test, i.e., to reduce the excitation level in such a way that the beam does not make contact with the bushes. This latter alternative is preferred because it is much better to limit the number of parameters to be estimated in the non-linear model. The mass is found to be equal to 24.96 kg.

At this stage, the Bayesian approach can be exploited in order to determine an appropriate model for the restoring force. However, in this case, the system dynamics clearly suggests a piece-wise linear model for the stiffness. For the sake of comparison, a polynomial model up to order 3 is also considered during the identification. These models are described by the following equations:

Piece-wise linear model:

$$f_{NL}(x,\dot{x}) = \begin{cases} kx + c\dot{x} + k_+(x-d) + c_+(\dot{x} - \dot{x}_{x=d}) & \text{if } x \geq d \\ kx + c\dot{x} & \text{if } |x| < d \\ kx + c\dot{x} + k_-(x+d) + c_-(\dot{x} - \dot{x}_{x=-d}) & \text{if } x \leq -d \end{cases} \tag{1.14}$$

Polynomial model:

$$f_{NL}(x,\dot{x}) = \sum_{i=1}^{3} \alpha_i \, x^i + \beta_i \, \dot{x}^i \tag{1.15}$$

where d is the clearance value that needs to be determined. Inspection of the stiffness curve points out that the increase in stiffness occurs around 0.0004 m. To increase the accuracy, the NMSE is computed for a hundred values of d regularly spaced between 0.0003 and 0.0005 m, and the optimum value turns out to be 0.00039 m. For further details about this procedure, the reader is referred to reference [89].

The NMSE is equal to 1.70 % and 1.80 % for the polynomial and piece-wise linear models respectively which is an indication of an accurate identification. Surprisingly, the polynomial model provides better accuracy than the piece-wise linear model. This indicates that the change in stiffness is smoother in practice than believed. However, it should be kept in mind that the polynomial model is input dependent which represents a major drawback of this model.

The identified parameters and their significance factors are listed in Tables 1.2 and 1.3. The damping coefficients are badly estimated, i.e., some have a negative sign, although damping should actually be in the model. The reason is probably because damping in the beam is very low. Hopefully, this is not a problem because the corresponding significance

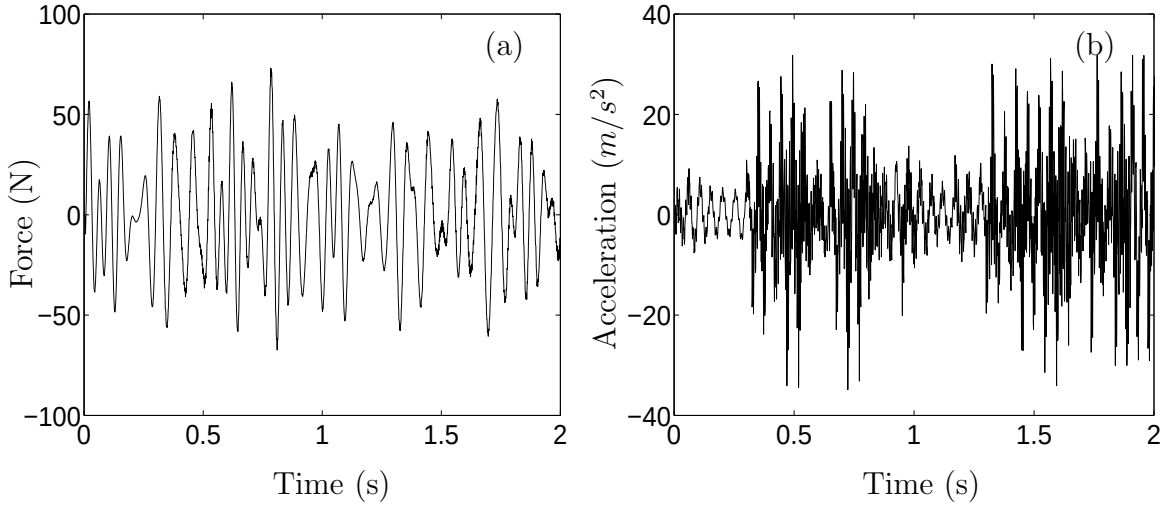


Figure 1.3: Measured signals. (a) Force time history; acceleration time history.

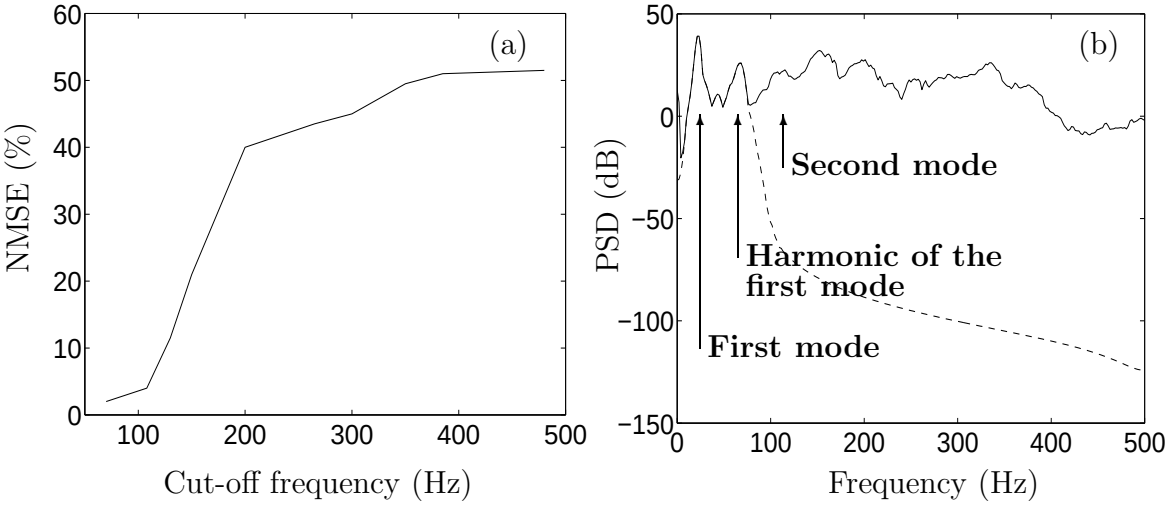


Figure 1.4: Piece-wise linear set-up. (a) NMSE vs. cut-off frequency; (b) PSD of the acceleration: ——, before filtering; - - - -, after filtering.

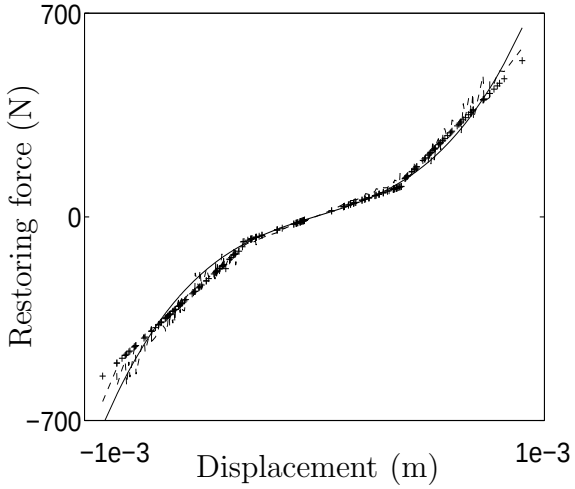


Figure 1.5: Measured and reconstructed stiffness curves: - - - -, measured; ——, reconstructed (polynomial model); +, reconstructed (piece-wise linear model).

	k_-	k	k_+	c_-	c	c_+
Parameters	885 783	281 011	871 787	-77.99	71.46	11.32
Significance factors	99.87	99.51	99.99	0.13	0.46	$2.78\ 10^{-3}$

Table 1.2: Identification results (piece-wise linear model, piece-wise linear beam).

	α_1	α_2	α_3	β_1	β_2	β_3
Parameters	248 903	$-2.14\ 10^7$	$5.89\ 10^{11}$	-32.09	1487	12 949
Significance factors	27.89	0.037	25.13	$8.4\ 10^{-3}$	0.052	0.064

Table 1.3: Identification results (polynomial model, piece-wise linear beam).

	k	k_+	c	c_+
Parameters (estimated displacement)	356 813	765 733	55.08	40.45
Parameters (improved displacement)	312 521	846 863	52.27	-62.93

Table 1.4: Identification results (bi-linear beam).

factors are negligible. It is also emphasised that for the piece-wise linear model, $k_+ \cong k_-$ as it should be for a symmetric system. Figure 1.5 presents the measured and reconstructed stiffness curves. Both reconstructed curves provide a close match to the measured curve, which confirms that the identification has given reliable results.

1.5.3 Bi-linear beam

The piece-wise linear set-up was completely dismantled in order to built the bi-linear set-up for which a single bush is present. Section 1.3 has shown that it is better to measure the displacement when considering an asymmetric system. Unfortunately, it was not possible to do so when the measurements were acquired and only the acceleration and the input force were measured. Note also that most of vibration measurements on mechanical structures are usually performed by accelerometers rather than velocity or displacement transducers.

The acceleration is integrated once and twice to give velocity and displacement respectively. Both signals are then filtered using a Butterworth filter with cut-off at 10 Hz. As for the piece-wise linear set-up, the use of a low-pass filter is necessary to filter out the contributions of the higher modes. Figure 1.6(a) compares the PSD of the acceleration signal before and after filtering. The bi-linear characteristic of the system introduces a sequence of harmonics at $n\omega$ where $n = 2, 3, 4$ and ω is the frequency of the first mode.

A critical issue for this set-up is the estimation of the displacement from the acceleration. Indeed, since the system is no longer symmetric, the displacement has not a

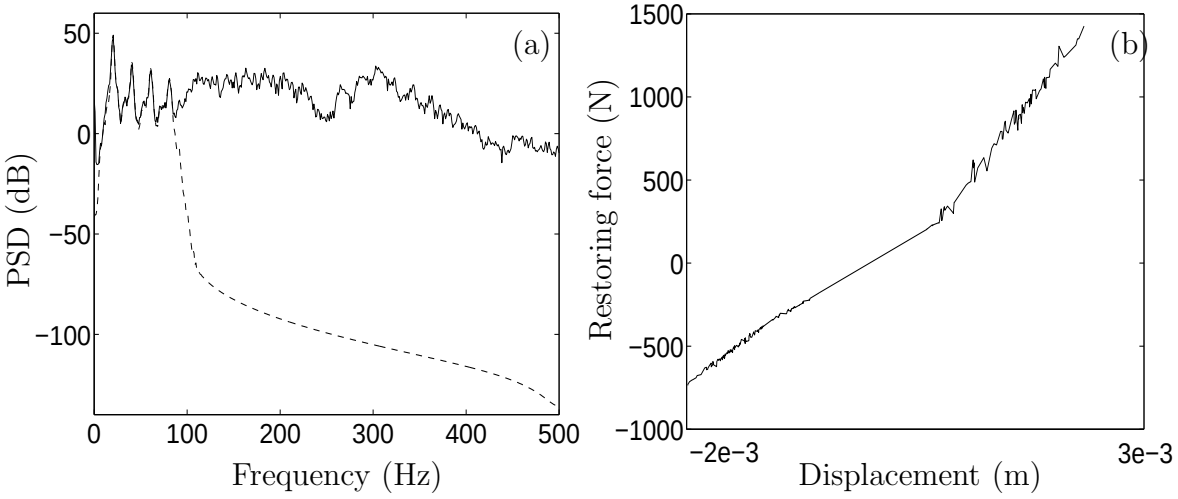


Figure 1.6: Bi-linear set-up. (a) PSD of the acceleration: ——, before filtering; - - - -, after filtering; (b) stiffness curve.

zero mean and the high-pass filter removes this component. In reference [89], we have proposed a means of improving the estimation of the displacement in such case. The first step of the procedure consists in adding a constant signal to the displacement obtained from the integration and filtering procedures. The evolution of the NMSE as a function of the constant signal amplitude is then examined, and the amplitude that realises the minimum is selected.

For the sake of comparison, the identification is carried out in two cases. On the one hand, a constant signal is not added to the estimated displacement and on the other hand, the proposed procedure is applied. The restoring force is represented using a bi-linear model:

$$f_{NL}(x,\dot{x}) = \begin{cases} kx + c\dot{x} + k_+(x-d) + c_+(\dot{x} - \dot{x}_{x=d}) & \text{if } x \geq d \\ kx + c\dot{x} & \text{else} \end{cases} \tag{1.16}$$

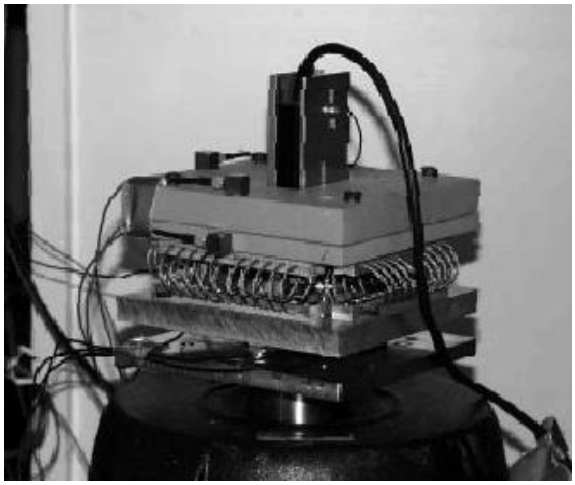
The clearance value is computed using the same procedure as for the piece-wise linear case and is found to be equal to 0.000762 m. This value is in good agreement with the stiffness curve represented in Figure 1.6(b). The identified parameters are listed in Table 1.4. The NMSE decreases from 2.56 % to 1.66 % when the displacement is improved by the addition of a constant signal. This confirms the efficiency of the proposed procedure.

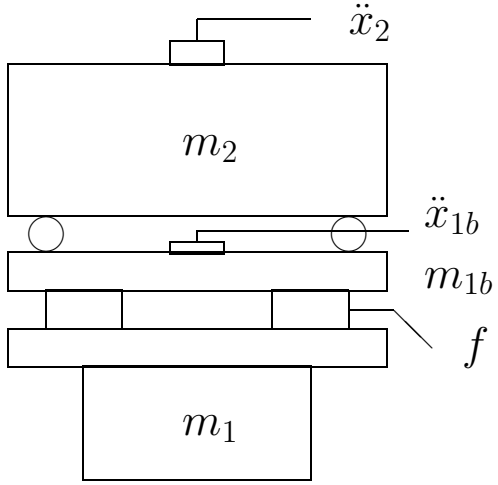
1.6 Application to wire rope isolators

The analysed data were chosen from those proposed by VTT Technical Research Centre of Finland within the framework of the European COST Action F3. The aim of this benchmark is to identify the dynamic properties of resilient mounts used for vibration isolation in industrial applications.

1.6.1 Experimental set-up

The structure investigated consists of wire rope isolators mounted between a load mass m_2 and a base mass m_{1b}, as shown in Figure 1.7. The motion and forces experienced by the isolators are measured. In particular, the acceleration responses $\ddot{x}_2$ and $\ddot{x}_{1b}$ of the load mass and bottom plate, the applied force f and the relative displacement x_{12} between the top and bottom plates are measured. The excitation produced by an electro-dynamic shaker corresponds to a white noise sequence, low-pass filtered at 400 Hz. Four excitation levels are considered ranging from 0.5 Volt up to 8.0 Volt. A nominal series of five tests are performed with a load mass of 2.2 kg. A sixth test is also carried out with a heavier load mass of 5.8 kg. Table 1.5 defines the different tests from which data sets have been collected.





System characteristics: Load mass: 2.2 kg or 5.8 kg;

Bottom plate: 1.1 kg;

Shaker base: 3.7 kg total;

Wire rope isolators: wire diameter 2 mm, 10 loops,

length 110 mm, loop diameter 30 mm.

Figure 1.7: Wire rope isolators.

1.6.2 Identification results

The formulation of the RFS method has been introduced in section 1.1 for the identification of SDOF systems. However, the structure considered here is clearly an MDOF system. Writing Newton's second law for the load mass m_2 yields:

$$m_2\,\ddot{x}_2 + f_{NL}(x_2 - x_{1b}, \dot{x}_2 - \dot{x}_{1b}) = 0 \tag{1.17}$$

Forcing level	Mass 1 (2.2 kg)	Mass 2 (5.8 kg)
0.5 Volt	Test 1	—
1 Volt	Test 2 (rejected)	—
2 Volt	Test 3	—
4 Volt	Test 4	Test 6
8 Volt	Test 5	—

Table 1.5: Different tests of the VTT benchmark.

Introducing the relative displacement $x_{12} = x_2 - x_{1b}$, this equation becomes:

$$f_{NL}(x_{12}, \dot{x}_{12}) = -m_2 \, \ddot{x}_{1b} - m_2 \, \ddot{x}_{12} \tag{1.18}$$

Equation (1.18) can be viewed as describing the response of an SDOF system subjected to a base acceleration. Because the acceleration signals are measured and the mass is known, it is possible to compute the restoring force at each sampling instant.

With the aim of determining an appropriate functional form for the restoring force, Figure 1.8 illustrates the stiffness curves of the different tests considered. The data corresponding to the 1 Volt level (test 2) is not further considered because the stiffness curve appears to be noisy. At low excitation level, the system behaviour is predominantly linear because the restoring force varies linearly with the displacement, as can be observed for the 0.5 Volt and 2 Volt levels (tests 1 and 3). As the excitation level is increased, a softening stiffness non-linearity appears, as can be seen from the 4 Volt and 8 Volt levels (tests 4-6). It should also be noted that the stiffness curves are symmetric which means that the non-linear effect can be characterised by an odd exponent. The damping curves for tests 4 and 5 are represented in Figure 1.9. These curves are quite noisy, essentially due to the slight participation of damping in the system response, and a linear damping model seems to be the most appropriate in this case.

The foregoing procedure has helped in the model selection in the sense that a softening symmetric non-linearity in stiffness has been detected. Accordingly, the restoring force is modelled as the sum of a linear stiffness term, a viscous damping term and a non-linear stiffness contribution:

$$f_{NL}(x_{12}, \dot{x}_{12}) = k_{lin} \, x_{12} + c_{lin} \, \dot{x}_{12} + k_{nl} \, |x_{12}|^{\alpha} \, \text{sign}(x_{12}) \tag{1.19}$$

where the coefficients k_{lin}, c_{lin}, k_{nl} and the non-linearity exponent α have to be estimated.

The exercise of identifying the VTT benchmark structure was first carried out using the singular value decomposition (SVD) [90]. In the context of this dissertation, the performance of the Bayesian model screening is assessed. While the principle of Bayesian inference has previously been applied to various problems in structural dynamics (e.g., references [102, 188]), no attempt has been made, to the best of our knowledge, to adapt the Bayes updating rule to the screening of model form during non-linear system identification.

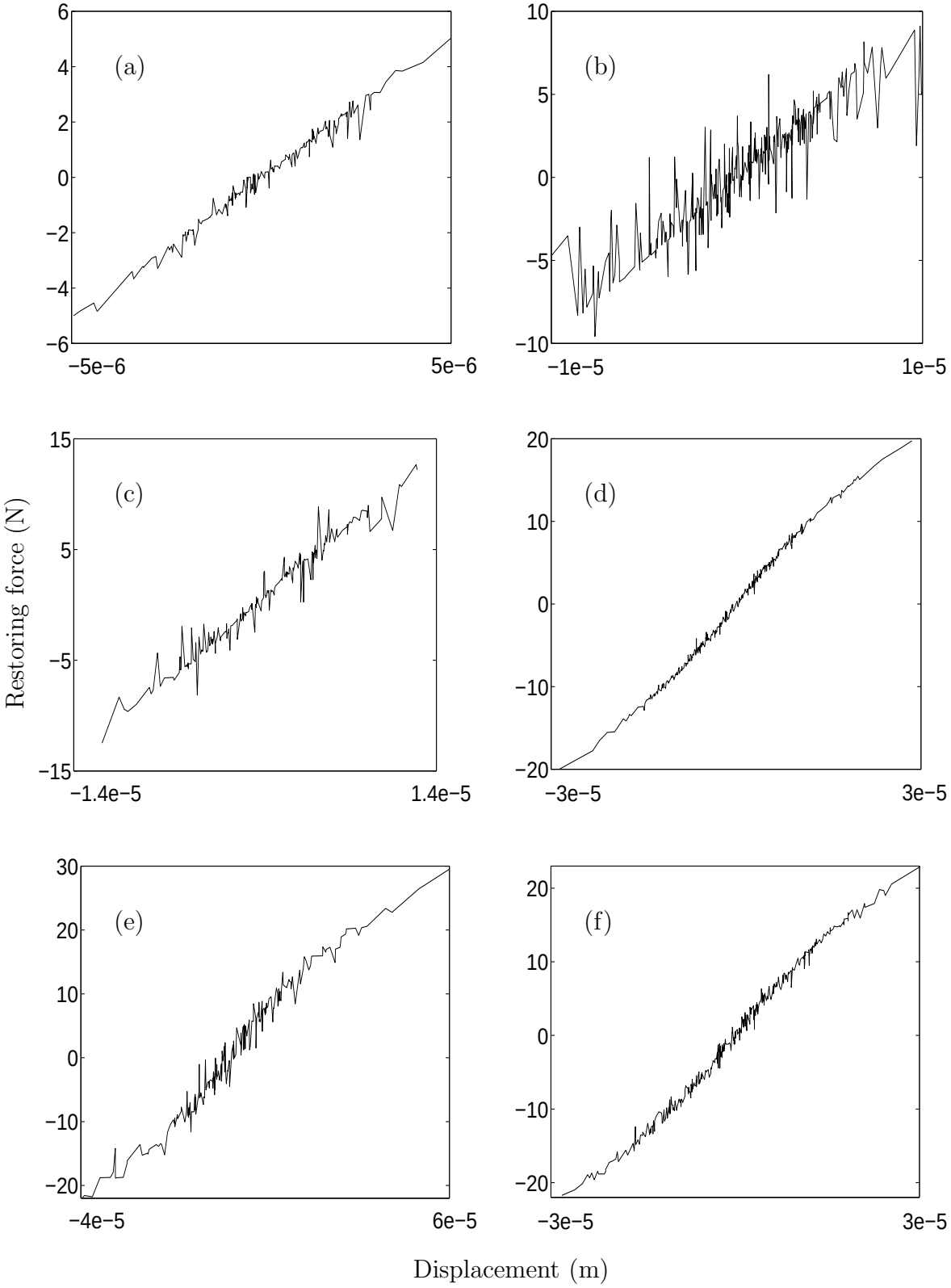


Figure 1.8: Stiffness curves. (a)-(f) Tests 1-6.

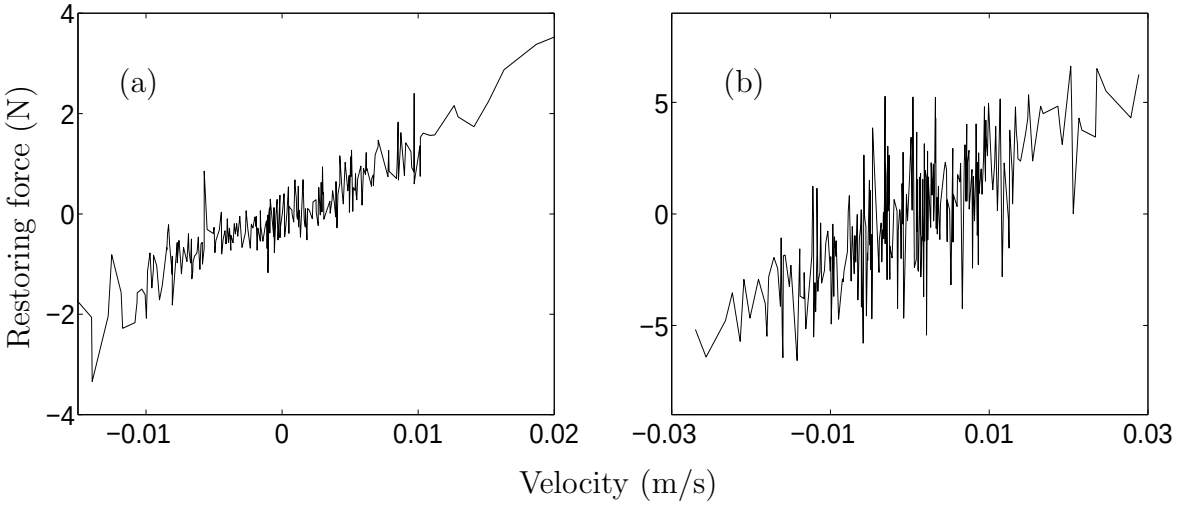


Figure 1.9: Damping curves. (a) Test 4; (b) test 5.

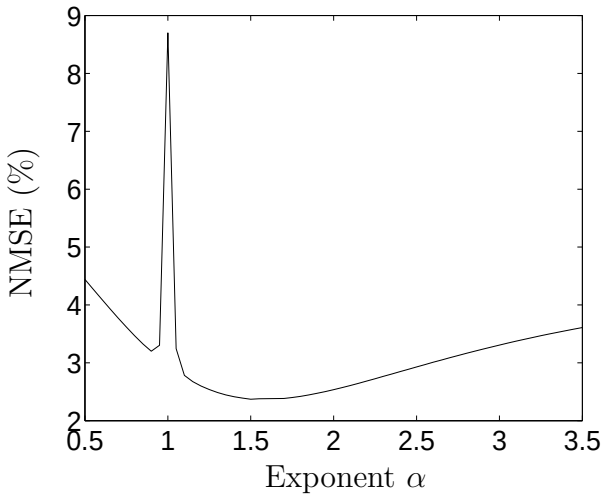


Figure 1.10: Evolution of the NMSE as a function of the non-linearity exponent.

The main effects described by equation (1.19) and linear interactions between these effects are allowed. This means that a total of six different model forms are allowed. Such combinatorial complexity is trivial compared to the numerical example discussed previously. The complexity here stems from the fact that real data sets are analysed with all the risk of erroneous identification caused by measurements and signal conditioning issues.

It should be noted that only four of the five data sets (tests 1, 3, 5, 6) are considered for identification, and the remaining data set (test 4) is exploited to assess the predictive accuracy of the identified model. Because the exponent α is unknown, the Bayesian model screening is repeated for several assumed values of α. The value that leads to the smallest NMSE is retained. For this application, an initial chain of length 50 is dedicated

to burn-in and a chain of length 300 is used for the optimisation. It has been verified that requesting more samples does not improve the quality of the final results. Figure 1.10 shows the evolution of the NMSE as a function of the exponent α. Two comments can be made regarding this figure:

1. A peak appears for a value of α equal to 1. Thus, a linear model is not sufficient to capture the dynamics because the NMSE is much smaller for other values of α;

2. The minimum value is obtained for α=1.5, the same value as the one identified in reference [90] with the RFS method and the SVD.

Table 1.6 and Figure 1.11 display the top five models and the marginal posterior probability of each effect respectively. The NMSE for each of the top five models is around 2.37%. Such NMSE values indicate that the agreement with experimental data meets the expected level of accuracy. The most likely model in Table 1.6 includes only the main effects, namely, x_{12}, $\dot{x}_{12}$ and $|x_{12}|^{1.5}\text{sign}(x_{12})$, and appears 86% of the time in the Markov chain.

The main effects, labelled 1-3 in Figure 1.11, have a posterior probability of 100% while interaction effects, labelled 4-6, may be considered negligible. For this application, the prior probabilities are set to a uniform 20% for the main effects; 10% for an interaction effect when at least one of the "parent" main effect is selected in the model; and 1% only for an interaction effect when none of the parents are selected. The increase in probability for effects 1-3 in Figure 1.11 and the reduction for effects 4-6 are therefore significant. From these results, it can be concluded that a suitable model for the restoring force is given by equation (1.19) with an exponent equal to $\alpha = 1.5$.

The coefficients identified with the Bayesian model screening are listed in Table 1.7 and compared with the values identified using the SVD. The last column of this table gives the relative deviation between the results obtained by the two approaches. The negative sign of the non-linear stiffness corroborates the presence of a softening effect. Although the solution is unknown, it can be stated that both methods provide consistent results because the maximum difference is less than 10%. The small differences witnessed may be attributed to the different data sets used. Indeed, in reference [90], the identification was conducted using the five combinations of input levels and load masses. The Bayesian model screening is restricted to four of the five cases to provide a validation of the model predictive accuracy. This is discussed in the next section.

1.6.3 Validation of the identified model

The data sets collected during tests 1, 3, 5 and 6 have been used for model screening and identification while the data collected during test 4 is now used for model validation. Essentially, equation (1.19) is evaluated with the coefficients of Table 1.7 to predict the restoring force. The time evolution of the predicted restoring force is then compared to

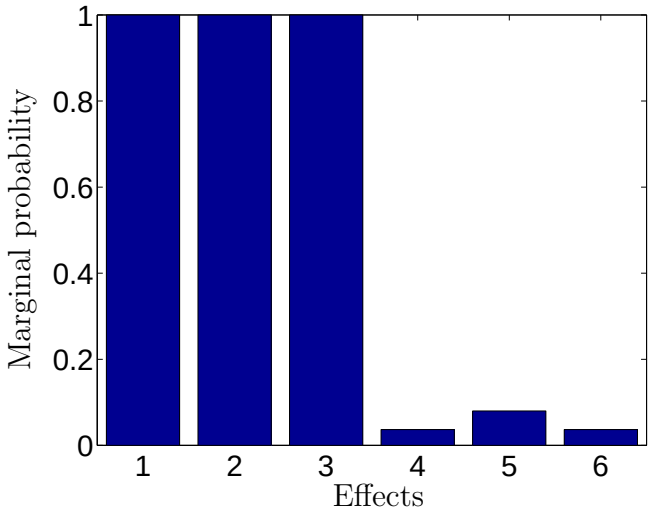


Figure 1.11: Marginal posterior probability of each effect included in the family of models.

Model	Effects	Posterior Probability
1	$x_{12}, \dot{x}_{12}, \lvert x_{12}\rvert^{1.5}\,\mathrm{sign}(x_{12})$	86.0%
2	Model 1 + $x_{12}\,\lvert x_{12}\rvert^{1.5}\,\mathrm{sign}(x_{12})$	7.0%
3	Model 1 + $\dot{x}_{12}\,\lvert x_{12}\rvert^{1.5}\,\mathrm{sign}(x_{12})$	3.0%
4	Model 1 + $x_{12}\,\dot{x}_{12}$	2.6%
5	Model 2 + $\dot{x}_{12}\,\lvert x_{12}\rvert^{1.5}\,\mathrm{sign}(x_{12})$	0.3%

Table 1.6: Top five models and number of appearances.

Coefficient	Value (Bayes)	Value (SVD)	Difference
k_{lin} (N/m)	$1.12\ 10^6$	$1.09\ 10^6$	2.75 %
c_{lin} (N.s/m)	198.19	183.44	8.04 %
k_{nl} $(N/m^{1.5})$	$-9.07\ 10^7$	$-8.52\ 10^7$	6.46 %
α	1.5	1.5	0 %

Table 1.7: Results of the Bayesian identification.

the restoring force measured during test 4. The comparison is displayed in Figure 1.12. It can be observed that the measured and predicted forces agree to the point where the difference between signals is not visible. The NMSE is equal to 0.77 %. This analysis leads to the conclusion that a reliable identification has been performed over the operational range of interest, i.e., within the ranges of 2.2 to 5.8 kg load mass and 0.5 to 8.0 Volt excitation level.

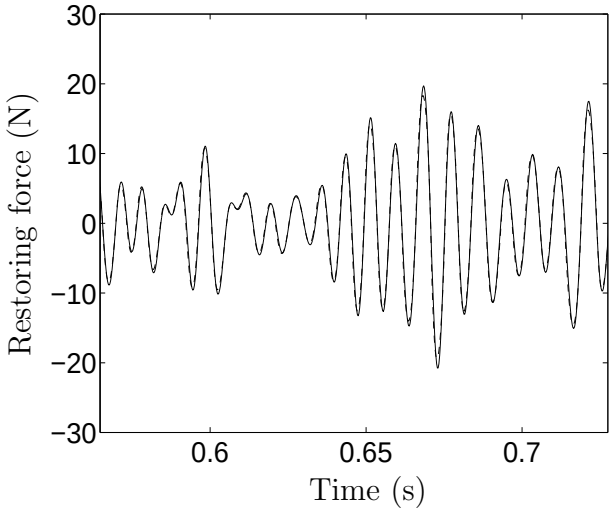


Figure 1.12: Measured and predicted restoring forces (test 3): - - - -, predicted; ——, measured.

1.6.4 Comparison with the conditioned reverse path method

The identification of the wire rope isolators has also been performed by Marchesiello and Fasana from the University of Torino using the conditioned reverse path (CRP) method [124], and the results have been compared in reference [90]. Although the CRP and RFS methodologies are in essence very different, similar results have been obtained:

1. Having specified the same functional form of the non-linearity for both methods, the same kind of non-linearity has been identified, i.e., a softening $|x_{12}|^{1.5}\text{sign}(x_{12})$ non-linearity;

2. On the one hand, the CRP method identified a value equal to $-5.51\,10^7 N/m^{1.5}$ that is slightly higher than the ones found with the RFS method ($-8.52\,10^7 N/m^{1.5}$ using the SVD and $-9.07\,10^7 N/m^{1.5}$ using the Bayesian approach). On the other hand, the frequency of the underlying linear system is slower in the CRP method (108 Hz versus 112 Hz). It seems thus that the softening effect due to a lower linear frequency in the CRP method compensates for the stiffening effect due to a less negative non-linear stiffness.

If the linear stiffness k_{lin} is imposed in order to have a frequency equal to 108 Hz in the RFS method, then the non-linear stiffness goes down to $-6.73\,10^7 N/m^{1.5}$ and $-7.03\,10^7 N/m^{1.5}$ using the SVD and Bayesian approach respectively. This is closer to the value identified in the CRP method. The increase in the NMSE is about 0.16 % in both cases which reflects that the accuracy of the procedure is not very sensitive to the change in linear stiffness.

1.7 Concluding remarks

A time domain technique for the identification of non-linear vibrating structures has been presented in this chapter. The RFS method is appealing by its simplicity and its reliability for identification of SDOF systems. The method can be extended to MDOF systems but by losing an important advantage that lies in the two-dimensional representation of the restoring force for SDOF systems (stiffness and damping curves). The need for numerical integration and filtering may introduce errors in the estimation of signals, and careful signal processing is required.

The method has been applied to an impacting beam and to an industrial structure consisting of wire rope isolators. In each case, accurate structural models have been obtained which confirms the efficiency of the RFS approach for the identification of SDOF non-linear systems (or non-linearity localised between two DOF).

Although an SDOF model is sometimes a good approximation of more physically realistic MDOF systems, multi-mode models, however, are necessary to predict the dynamics of complex systems accurately and to describe their behaviour over a wide frequency range more fully. In this context, the RFS method should be superseded by other methodologies.

Finally, we hope that this chapter has highlighted the primary importance of non-linearity characterisation prior to identification. Any a priori information on the system under investigation should be exploited in order to help the identification technique to capture the intrinsic dynamic characteristics accurately. To this end, the stiffness and damping curves are very useful tools. We have also proposed the use of a Bayesian inference approach in order to provide insight into the form of the non-linearity. The key advantage of the procedure is that a collection of potential models together with their posterior probability is obtained instead of the single best model. It allows for more flexibility in deciding the most appropriate model of the non-linearity.

Chapter 2

Frequency Response Function-Based Identification Methods

2.1 Introduction

The introductory chapter has highlighted that most of the current methodologies are limited to relatively simple non-linear systems. The development of techniques capable of dealing with continuous systems for which a lumped parameter representation is inappropriate is thus a central issue. To progress in this direction, there has been recently a growth in interest in the development of frequency response function (FRF)-based approaches, and the conditioned reverse path (CRP) and non-linear through feedback of the outputs (NIFO) formulations have been introduced. These methodologies aim at eliminating the distortions in FRFs caused by the presence of a non-linear effect and allow for the estimation of non-linear coefficients together with the FRFs of the underlying linear system.

The purpose of this chapter is to assess the overall performances of the CRP and NIFO methods for the identification of the dynamic characteristics of continuous multi-degree-of-freedom (MDOF) non-linear structures. With this aim in mind, an experimental structure consisting of a beam with a geometric non-linearity is investigated. This benchmark has been proposed by the Ecole Centrale de Lyon (France) within the framework of the European COST Action F3 working group on "Identification of non-linear systems". Although it is difficult to draw firm conclusions from a single experiment, a discussion regarding the respective merits of each technique is proposed at the end of the present chapter.

2.2 Conditioned reverse path method

The properties of MDOF linear systems are typically identified using frequency domain modal parameter estimation techniques, as explained in references [45, 69]. The frequency

domain techniques extract modal parameters from H_1 and H_2 estimated FRFs:

$$H_1 = \frac{G_{XF}}{G_{FF}}, \quad H_2 = \frac{G_{XX}}{G_{XF}} \tag{2.1}$$

where G_{XX}, G_{FF} and G_{XF} contain the power spectral density (PSD) of the response (e.g., acceleration signal), the PSD of the applied force and the cross PSD between the response and the applied force respectively.

In the presence of non-linear forces, the H_1 and H_2 estimators should not be used because non-linearities corrupt the underlying linear characteristics of the response. To address this issue, the CRP method takes advantage of the spectral conditioning techniques to remove the effects of non-linearities before computing the FRFs of the underlying linear system.

The concept of a reverse path (RP) model was introduced by Bendat [15, 16] and adapted to MDOF systems by Rice and Fitzpatrick [147]. However, this technique requires the external force to be applied at the location of the non-linearity. The CRP formulation proposed by Richards and Singh in reference [148] extends the application of the RP algorithm to systems characterised by non-linearities away from the location of the applied force. This method was developed by generalising the concepts introduced by Bendat [15, 16]. Conditioned frequency responses are first computed and yield the underlying linear properties without influence of non-linearities. The non-linear coefficients are then identified in a second step.

The vibrations of an MDOF non-linear system are assumed to be governed by the following equation:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \sum_{j=1}^{n} \mathbf{A}_j \mathbf{y}_j(t) = \mathbf{f}(t) \tag{2.2}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices respectively;
$\mathbf{x}(t)$ is the vector of displacement co-ordinates;
$\mathbf{y}_j(t)$ is a non-linear function vector;
$\mathbf{A}_j$ contains the coefficients of the term $\mathbf{y}_j(t)$;
$\mathbf{f}(t)$ is the applied force vector.

By way of illustration, in the case of a grounded cubic stiffness at the ith degree-of-freedom (DOF), the non-linear function vector is:

$$\mathbf{y}_j(t) = [0 \ldots x_i(t)^3 \ldots 0]^T \tag{2.3}$$

In the frequency domain, equation (2.2) becomes:

$$\mathbf{B}(\omega)\mathbf{X}(\omega) + \sum_{j=1}^{n} \mathbf{A}_j \mathbf{Y}_j(\omega) = \mathbf{F}(\omega) \tag{2.4}$$

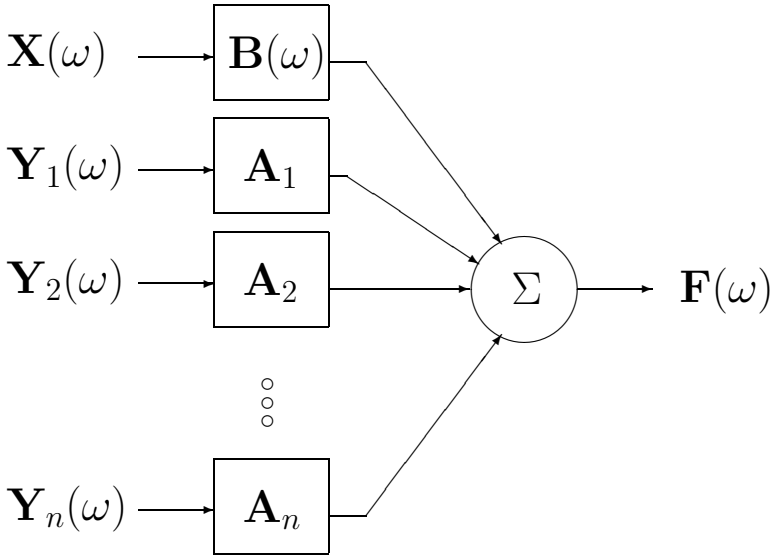


Figure 2.1: Reverse path model.

where $\mathbf{X}(\omega), \mathbf{Y}_j(\omega)$ and $\mathbf{F}(\omega)$ are the Fourier transform (FT) of $\mathbf{x}(t), \mathbf{y}_j(t)$ and $\mathbf{f}(t)$;
$\mathbf{B}(\omega) = -\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K}$ is the linear dynamic stiffness matrix.

By imposing the applied force to be the output and the measured responses to be the inputs, a RP model is constructed, as shown in Figure 2.1. For the sake of conciseness, the RP formulation is not recalled here, and the reader is referred to reference [148] for a detailed description of this technique.

2.2.1 Estimation of the underlying linear system properties

The purpose of the CRP method is, first, to estimate the FRFs of the underlying linear system contained in the dynamic compliance matrix $\mathbf{H}(\omega) = \mathbf{B}^{-1}(\omega)$. The key idea of the formulation is the separation of the non-linear part of the system response from the linear part and the construction of uncorrelated response components in the frequency domain.

Let us assume that a single non-linear term $\mathbf{Y}_1$ is present. The spectrum of the measured responses $\mathbf{X}$ can be decomposed into a component $\mathbf{X}_{(+1)}$ correlated with the spectrum of the non-linear vector $\mathbf{Y}_1$ through a frequency response matrix $\mathbf{L}_{1X}$, and a component $\mathbf{X}_{(-1)}$ uncorrelated with the spectrum of the non-linear vector. In what follows, the minus (plus) sign signifies uncorrelated (correlated) with. Likewise, the spectrum of the external force $\mathbf{F}$ can be decomposed into a component $\mathbf{F}_{(+1)}$ correlated with the spectrum of the non-linear vector $\mathbf{Y}_1$ through a frequency response matrix $\mathbf{L}_{1F}$, and a component $\mathbf{F}_{(-1)}$ uncorrelated with the spectrum of the non-linear vector. Since both vectors $\mathbf{X}_{(-1)}$ and $\mathbf{F}_{(-1)}$ are uncorrelated with the non-linear vector, the path between them is the linear dynamic stiffness matrix $\mathbf{B}$. The whole procedure is illustrated in

Figure 2.2.

The generalisation to multiple non-linearities is straightforward. In this case, the spectra of the response and the force need to be uncorrelated with all n non-linear function vectors:

$$\begin{cases} \mathbf{X}_{(-1:n)} = \mathbf{X} - \sum_{j=1}^{n} \mathbf{X}_{(+j)} = \mathbf{X} - \sum_{j=1}^{n} \mathbf{L}_{jX} \mathbf{Y}_{j(-1:j-1)} \\ \\ \mathbf{F}_{(-1:n)} = \mathbf{F} - \sum_{j=1}^{n} \mathbf{L}_{jF} \mathbf{Y}_{j(-1:j-1)} \end{cases} \tag{2.5}$$

The RP model represented in Figure 2.1 can thus be replaced by a RP model for which the inputs are uncorrelated. Figure 2.3 represents this modified RP model and shows that the path between $\mathbf{X}_{(-1:n)}$ and $\mathbf{F}_{(-1:n)}$ is the linear dynamic stiffness matrix $\mathbf{B}$:

$$\mathbf{F}_{(-1:n)}(\omega) = \mathbf{B}(\omega)\mathbf{X}_{(-1:n)}(\omega) \tag{2.6}$$

By transposing equation (2.6), pre-multiplying by the complex conjugate of $\mathbf{X}$, i.e., $\mathbf{X}^*$, taking the expectation and multiplying by $2/T$, the underlying linear system can be identified without corruption from the non-linear terms:

$$\begin{aligned} \mathbf{G}_{XF(-1:n)} &= \frac{2}{T} E[\mathbf{X}^* \mathbf{F}_{(-1:n)}^T] = \frac{2}{T} E[\mathbf{X}^* (\mathbf{B}\mathbf{X}_{(-1:n)})^T] \\ &= \frac{2}{T} E[\mathbf{X}^* \mathbf{X}_{(-1:n)}^T \mathbf{B}^T] = \mathbf{G}_{XX(-1:n)} \mathbf{B}^T \end{aligned} \tag{2.7}$$

where $\mathbf{G}_{XF(-1:n)}$ and $\mathbf{G}_{XX(-1:n)}$ are conditioned PSD matrices. Calculation of these matrices is laborious and involves a recursive algorithm. For the sake of conciseness, only the final formulae are given here. Reference [14] shows that:

$$\mathbf{G}_{ij(-1:r)} = \mathbf{G}_{ij(-1:r-1)} - \mathbf{G}_{ir(-1:r-1)} \mathbf{L}_{rj}^T \tag{2.8}$$

where

$$\mathbf{L}_{rj}^T = \mathbf{G}_{rr(-1:r-1)}^{-1} \mathbf{G}_{rj(-1:r-1)} \tag{2.9}$$

The dynamic compliance matrix $\mathbf{H}$ that contains the FRFs of the underlying linear system takes the following form:

$$H_{c2} : \mathbf{H}^T = \mathbf{G}_{XF(-1:n)}^{-1} \mathbf{G}_{XX(-1:n)} \tag{2.10}$$

This expression is known as the conditioned H_{c2} estimate. If relation (2.6) is multiplied by the complex conjugate of $\mathbf{F}$ instead of $\mathbf{X}$, the conditioned H_{c1} is obtained:

$$H_{c1} : \mathbf{H}^T = \mathbf{G}_{FF(-1:n)}^{-1} \mathbf{G}_{FX(-1:n)} \tag{2.11}$$

When FRFs of linear systems are estimated, H_1 always produces better estimates when there is measurement noise on the outputs, and H_2 produces better estimates when the noise is on the input measurements. Intuition may lead us to expect the H_{c2} estimate to perform better than the H_{c1} estimate in the presence of uncorrelated noise only in the excitation. Likewise, the H_{c1} estimate is expected to perform better than the H_{c2} estimate in the presence of uncorrelated noise only in the response. However, experience shows that the H_{c2} estimate tends to overcome the H_{c1} estimate in both situations. This may be a result of the conditioning required to calculate these estimates.

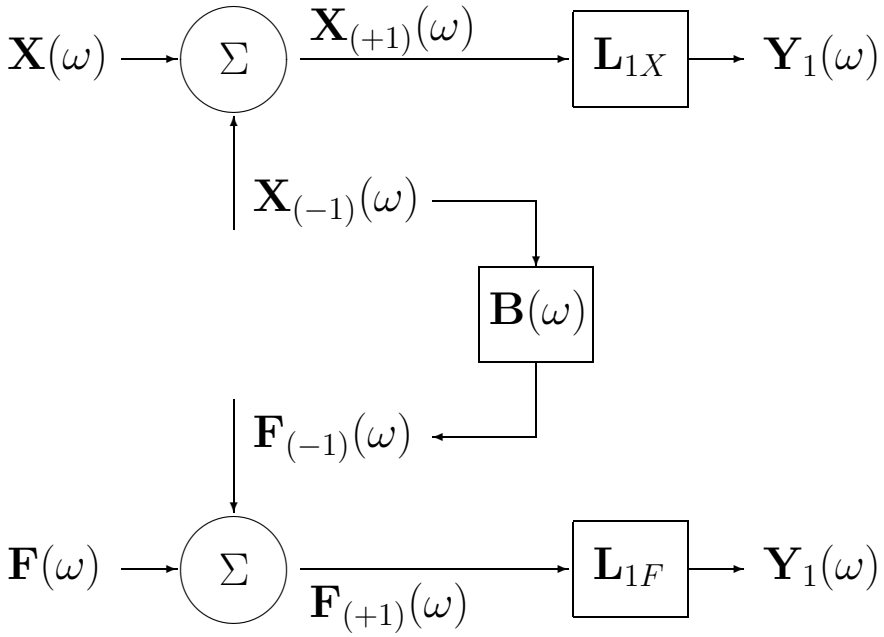


Figure 2.2: Decomposition of the force and response spectra.

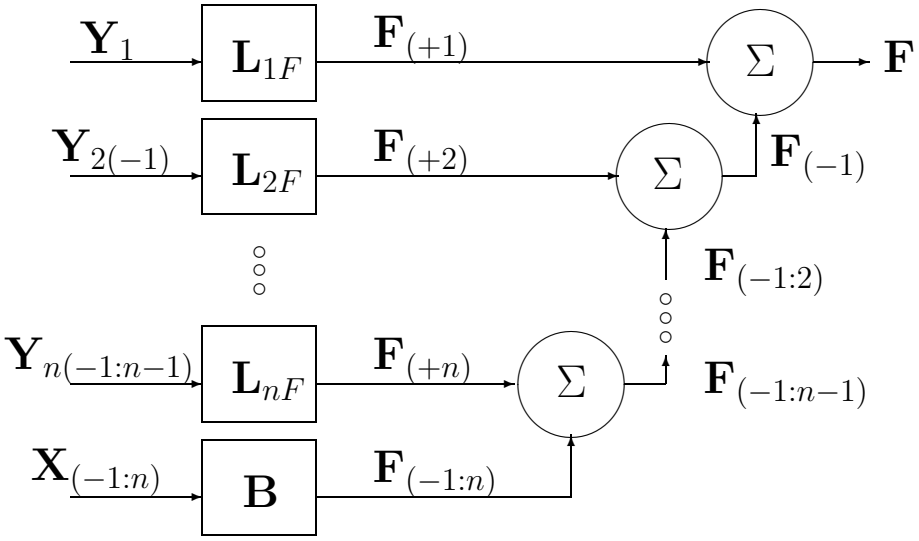


Figure 2.3: Reverse path model with uncorrelated inputs.

2.2.2 Estimation of the non-linear coefficients

Once the linear dynamic compliance $\mathbf{H}$ has been computed, the non-linear coefficients $\mathbf{A}_j$ can be estimated. By applying the same procedure as for equation (2.7) to equation (2.4), the following relationship is obtained:

$$\mathbf{G}_{iF(-1:i-1)} = \mathbf{G}_{iX(-1:i-1)}\mathbf{B}^T + \sum_{j=1}^{n} \mathbf{G}_{ij(-1:i-1)}\mathbf{A}_j^T \tag{2.12}$$

It should be noted that $\mathbf{G}_{ij(-1:i-1)} = E\left[\mathbf{Y}^*_{i(-1:i-1)}\mathbf{Y}_j^T\right] = \mathbf{0}$ for $j < i$ since $\mathbf{Y}^*_{i(-1:i-1)}$ is uncorrelated with the spectrum of the non-linear function vectors $\mathbf{Y}_1$ through $\mathbf{Y}_{i-1}$. If equation (2.12) is pre-multiplied by $\mathbf{G}^{-1}_{ii(-1:i-1)}$, the first term in the summation is $\mathbf{A}_i^T$. Equation (2.12) is then transformed into:

$$\mathbf{A}_i^T = \mathbf{G}^{-1}_{ii(-1:i-1)}\left(\mathbf{G}_{iF(-1:i-1)} - \mathbf{G}_{iX(-1:i-1)}\mathbf{B}^T - \sum_{j=i+1}^{n} \mathbf{G}_{ij(-1:i-1)}\mathbf{A}_j^T\right) \tag{2.13}$$

Because the expression of the linear dynamic compliance has been computed, equation (2.13) is rewritten in a more suitable form:

$$\mathbf{A}_i^T\mathbf{H}^T = \mathbf{G}^{-1}_{ii(-1:i-1)}\left(\mathbf{G}_{iF(-1:i-1)}\mathbf{H}^T - \mathbf{G}_{iX(-1:i-1)} - \sum_{j=i+1}^{n} \mathbf{G}_{ij(-1:i-1)}\mathbf{A}_j^T\mathbf{H}^T\right) \tag{2.14}$$

The identification process starts with the computation of $\mathbf{A}_n$ working backwards to $\mathbf{A}_1$. At this stage, it is emphasised that the non-linear coefficients are frequency dependent. However, by taking the spectral mean, the actual value of the coefficients can be retrieved.

If excitations are not applied at each response location, only the columns corresponding to the DOF where excitations are applied can be determined. By taking advantage of the reciprocity principle, the vector containing the non-linear coefficients $\mathbf{A}$ can then be identified. Let us illustrate this on an example that consists of a three DOF system with a non-linear force between the second and the third DOF (see Figure 2.4). The external force is applied on the first DOF. In this case, equation (2.14) becomes:

$$\begin{bmatrix} 0 & \alpha & -\alpha \end{bmatrix} \begin{bmatrix} H_{11} & H_{21} & H_{31} \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} =$$

$$G_{11}^{-1}\left(\begin{bmatrix} G_{1F_1} & 0 & 0 \end{bmatrix} \begin{bmatrix} H_{11} & H_{21} & H_{31} \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} - \begin{bmatrix} G_{1X_1} & G_{1X_2} & G_{1X_3} \end{bmatrix}\right) \tag{2.15}$$

Equation (2.15) clearly illustrates that coefficient α cannot be identified because it is multiplied by unknown quantities. However, the reciprocity principle allows to write that $H_{12} = H_{21}$ and $H_{13} = H_{31}$, and equation (2.15) is rewritten as below:

$$\begin{bmatrix} 0 & \alpha & -\alpha \end{bmatrix} \begin{bmatrix} H_{11} & H_{21} & H_{31} \\ H_{12} & ? & ? \\ H_{13} & ? & ? \end{bmatrix} =$$

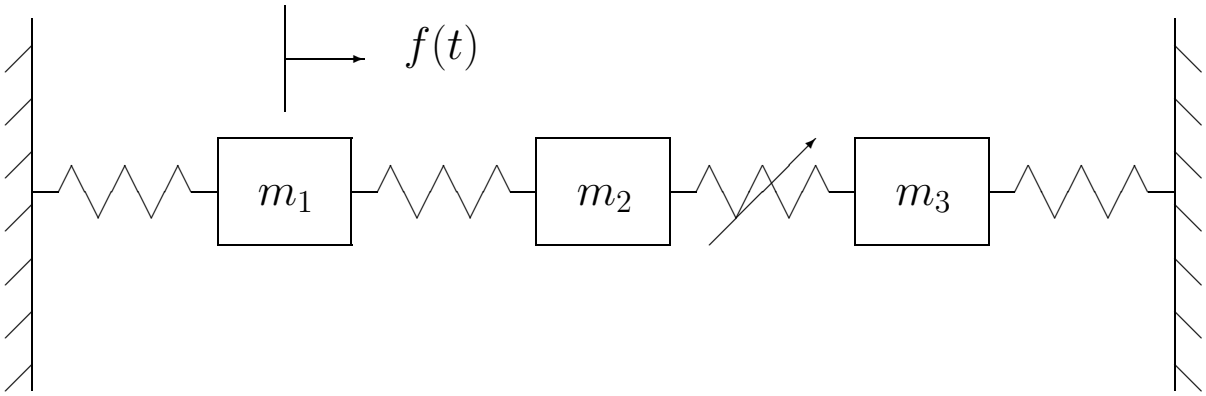


Figure 2.4: Three DOF system with a non-linear force between the second and third DOF.

$$G_{11}^{-1}\left([G_{1F_1} \quad 0 \quad 0]\begin{bmatrix} H_{11} & H_{21} & H_{31} \\ H_{12} & ? & ? \\ H_{13} & ? & ? \end{bmatrix} - [G_{1X_1} \quad G_{1X_2} \quad G_{1X_3}]\right) \tag{2.16}$$

and the non-linear coefficient can be calculated at each frequency:

$$\alpha = \frac{G_{1F_1}H_{11} - G_{1X_1}}{G_{11}(H_{12} - H_{13})} \tag{2.17}$$

2.2.3 Coherence functions

For linear systems, the ordinary coherence function is an indicator that gives a measure of the linear dependence between the applied force and the response as a function of frequency:

$$\gamma^2(\omega) = \frac{|G_{XF}(\omega)|^2}{G_{FF}(\omega)G_{XX}(\omega)} = \frac{H_1}{H_2} \tag{2.18}$$

The value of the coherence function is 1 when there is a perfect linear relationship between the two signals, and 0 when there is no relationship. It can be used to detect any departure from linearity or to detect the presence of uncorrelated noise on one or both of the excitation and response signals.

For a multiple input model with correlated inputs, the sum of ordinary coherences between the inputs and the output may be greater than unity. To address this problem, the ordinary coherence function has been replaced in reference [149] by the cumulative coherence function γ_{Mi}^2:

$$\gamma_{Mi}^2(\omega) = \gamma_{X_iF(-1:n)}^2(\omega) + \gamma_{YF}^2(\omega) = \gamma_{X_iF(-1:n)}^2(\omega) + \sum_{j=1}^{n} \gamma_{jF(-1:j-1)}^2(\omega) \tag{2.19}$$

where

- $\gamma^2_{X_iF(-1:n)}$ is the ordinary coherence function between the ith element of $X_{(-1:n)}$ and excitation F

$$\gamma^2_{X_iF(-1:n)} = \frac{\left|G_{X_iF(-1:n)}\right|^2}{G_{X_iX_i(-1:n)}G_{FF}} \tag{2.20}$$

  It indicates the contribution from the linear spectral component of the response of the ith output.

- $\gamma^2_{jF(-1:j-1)}$ is the ordinary coherence function between the conditioned spectrum $Y_{j(-1:j-1)}$ and excitation F

$$\gamma^2_{jF(-1:j-1)} = \frac{\left|G_{jF(-1:j-1)}\right|^2}{G_{jj_(-1:j-1)}G_{FF}} \tag{2.21}$$

  and $\sum_{j=1}^{n} \gamma^2_{jF(-1:j-1)}$ indicates the contribution from the non-linearities.

The cumulative coherence function is always between 0 and 1 and may be considered as a measure of the model accuracy.

2.3 Non-linear identification through feedback of the outputs method

The NIFO formulation introduced by Adams and Allemang in reference [1] is a new spectral approach for identification of MDOF non-linear systems. As for the CRP formulation, the central issue is to eliminate the distortions caused by the presence of non-linearities in FRFs. It exploits the spatial information and treats the non-linear forces as internal feedback forces in the underlying linear model of the system. The advantage of this method lies in its ability to estimate the FRFs of the underlying linear system and the non-linear coefficients in a single step. This is carried out in a least-squares system of equations through averaging.

Let us write equation (2.4) in the following form:

$$\mathbf{B}(\omega)\mathbf{X}(\omega) = \mathbf{F}(\omega) - \sum_{j=1}^{n} \mathbf{A}_j(\omega)\mathbf{Y}_j(\omega) \tag{2.22}$$

The non-linear forces may be considered as internal feedback forces and may be evaluated from the measured outputs. For instance, for a cubic stiffness between the mth and $(m+1)$th DOF, the non-linear force is:

$$\mathbf{Y}(\omega) = FT[(x_m(t) - x_{m+1}(t))^3] \tag{2.23}$$

Pre-multiplying equation (2.22) by the dynamic compliance matrix $\mathbf{H}(\omega)$ yields:

$$\mathbf{X}(\omega) = \mathbf{H}(\omega)\mathbf{F}(\omega) - \mathbf{H}(\omega)\sum_{j=1}^{n} \mathbf{A}_j(\omega)\mathbf{Y}_j(\omega) \tag{2.24}$$

and,

$$\mathbf{X}(\omega) = [\mathbf{H}(\omega) \quad \mathbf{H}(\omega)\mathbf{A}_1(\omega) \quad ... \quad \mathbf{H}(\omega)\mathbf{A}_n(\omega)] \begin{bmatrix} \mathbf{F}(\omega) \\ -\mathbf{Y}_1(\omega) \\ \vdots \\ -\mathbf{Y}_n(\omega) \end{bmatrix} \tag{2.25}$$

If the external force $\mathbf{F}(\omega)$ and the system response $\mathbf{X}(\omega)$ are measured, the system described by equation (2.25) may be solved at each frequency. This gives the FRFs of the underlying linear system $\mathbf{H}(\omega)$ together with the non-linear coefficients $\mathbf{A}_i(\omega)$. As for the CRP method, the non-linear coefficients are functions of frequency, but by taking the spectral mean, the actual values of these coefficients can be retrieved.

It is emphasised that equation (2.25) is not considered in its present form. A "PSD version" of this equation obtained by using the same procedure as for equation (2.7) is preferred. The use of PSDs reduces the degree to which linearly correlated terms corrupt the conditioning of the data matrices. An orthogonal least-squares solution [167] is also used to reduce the level of ill-conditioning.

2.4 Application to a beam with a geometric non-linearity

The purpose of this section is to assess the capability of the CRP and NIFO methods to identify the behaviour of a continuous MDOF non-linear structure.

2.4.1 Experimental set-up

The experimental structure consists of a clamped beam with a thin beam part at the right end of the main beam (cf. Figure 2.5). The geometric and mechanical properties of the set-up are listed in Table 2.1. Seven accelerometers that span the beam regularly are used to measure the response, and, in addition, a displacement sensor is located at the end of the main beam, i.e., at position 7. The external force is a white-noise sequence band-limited in the 0-500 Hz range. It should be noted that two different set-up were considered:

- The "horizontal set-up" in which the thin beam is horizontal and the shaker, located at position 2, excites the structure in a vertical plane [see Figure 2.6(a)];
- The "vertical set-up" in which the thin beam is vertical and the shaker, located at position 3, excites the structure in a horizontal plane [see Figure 2.6(b)].

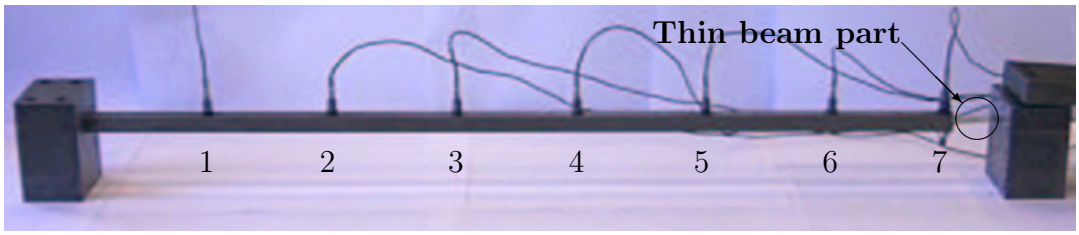


Figure 2.5: Experimental set-up.

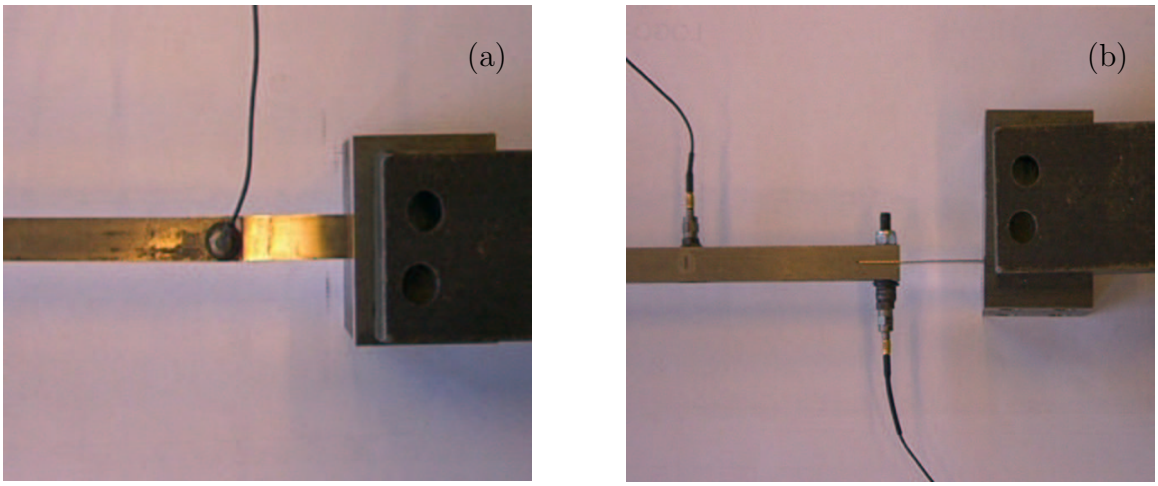


Figure 2.6: Experimental set-up, above view. (a) Horizontal set-up; (b) vertical set-up.

	Length (m)	Width (m)	Thickness (m)	Material
Main beam	0.7	0.014	0.014	Steel
Thin beam	0.04	0.014	0.0005	Steel

Table 2.1: Geometric and mechanical properties of the set-up.

2.4.2 Identification results with the CRP method: horizontal set-up

Conventional frequency response estimation (H_2 estimate)

The CRP method is exploited for the identification of the dynamic characteristics of the structure. Four excitation levels are considered ranging from 0.3 N up to 21 N.

In order to evaluate the participation of the non-linearity in the system response, the FRFs are first estimated using the linear H_2 estimate. For example, Figure 2.7 illustrates the magnitude of H_{72} for the lowest (0.3 N) and highest (21 N) excitation levels. As can be observed from this figure, distortions appear in the FRF when the excitation level is increased. The corresponding ordinary coherence functions are represented in Figure 2.8. Apart from frequencies below 20 Hz, the coherence is close to unity for the 0.3 N

level which is an indication of a linear relationship between the excitation force and the response acceleration. For the 21 N level, severe drops of the coherence function are observed and may be assigned to the presence of a non-linearity.

From these results, it appears that:

- The behaviour of the structure may be assumed to be linear for the lowest excitation level. The FRF computed in this case thus represents the FRF of the underlying linear system and is referred to as "reference FRF" in the remainder of this section.
- As the excitation level is increased, the behaviour of the structure becomes more and more non-linear. This is expected because the thin beam tends to be excited with a large deflection, and a geometric non-linearity is activated.

Model selection

In this example, it can be stated that the type and location of the non-linearity are perfectly identified. The next step in the identification procedure is the model selection. To this end, the cumulative coherence function γ^2_{Mi} as defined in section 2.2.3 is exploited. To model the stiffening effect of the thin beam, a grounded symmetric non-linearity whose functional form is described by $|x|^\alpha \operatorname{sign}(x)$ is introduced in the model at the end of the main beam (location 7). Due to the presence of a second harmonic of the first mode in the FRF of Figure 2.7(b), an asymmetric non-linearity is also included in the model. The functional form of the non-linearity is given by:

$$f_{NL}(x, \dot{x}) = A\,|x|^\alpha \operatorname{sign}(x) + B\,|x|^\beta \tag{2.26}$$

where x is the displacement at the end of the main beam.

Exponents α and β are determined by seeking the maximum value of the spectral mean of the averaged cumulative coherence of all the seven sensors:

$$accuracy = \frac{1}{N} \sum_{\omega=10}^{500} \left(\frac{1}{7} \sum_{i=1}^{7} \gamma^2_{Mi}(\omega) \right) \tag{2.27}$$

where N is the number of frequencies considered in the range from 10 to 500 Hz. The maximum value is found for $\alpha = 3$ and $\beta = 2$ and is equal to 0.9834. Figure 2.9 illustrates the cumulative coherence γ^2_{M7} for the 21 N level. Apart from a few drops (particularly around the first resonance), the cumulative coherence is unity indicating that the model is quite accurate.

Conditioned frequency response estimation (H_{c2} estimate)

The H_{c2} estimate (2.10) proposed by the CRP method is now used to compute the FRFs of the underlying linear system. Figure 2.10 presents the comparison between the reference FRF, i.e., the FRF for the 0.3 N level, and the FRF obtained using the H_{c2} estimate for

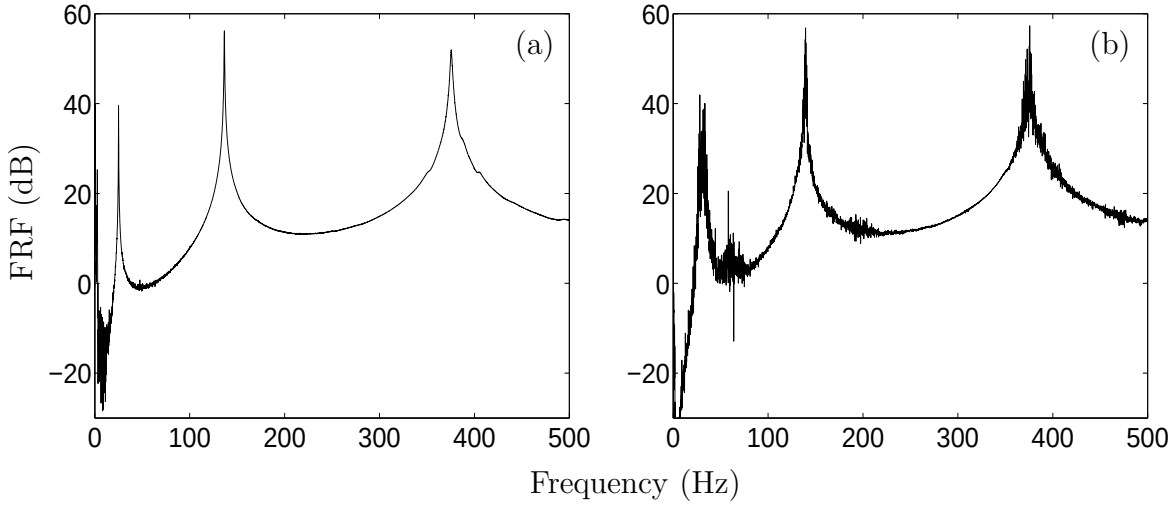


Figure 2.7: Horizontal set-up, magnitude of H_{72} (H_2 estimate). (a) 0.3 N; (b) 21 N.

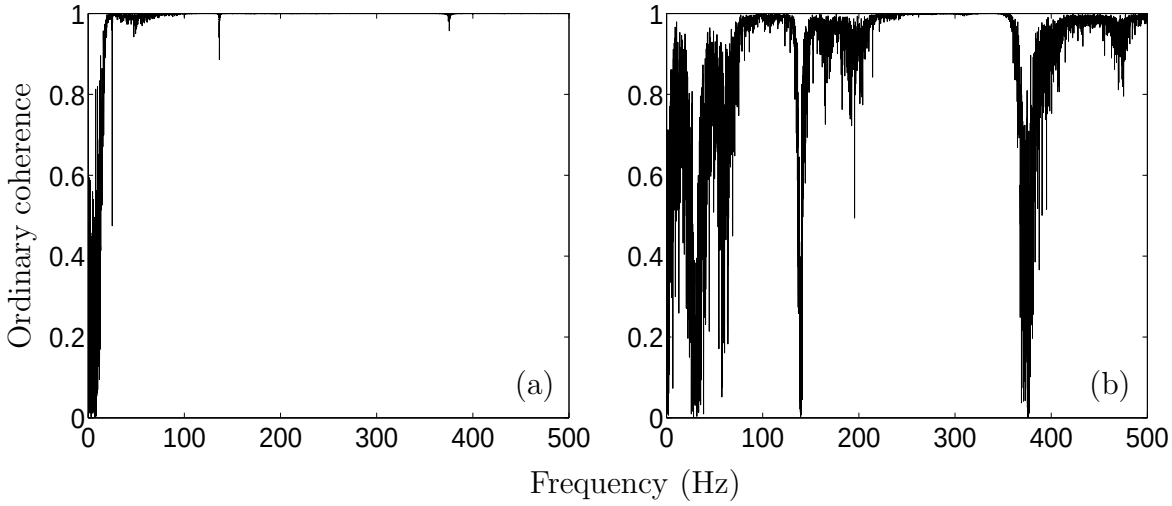


Figure 2.8: Horizontal set-up, ordinary coherence functions $\gamma^2_{X_7F}$. (a) 0.3 N; (b) 21 N.

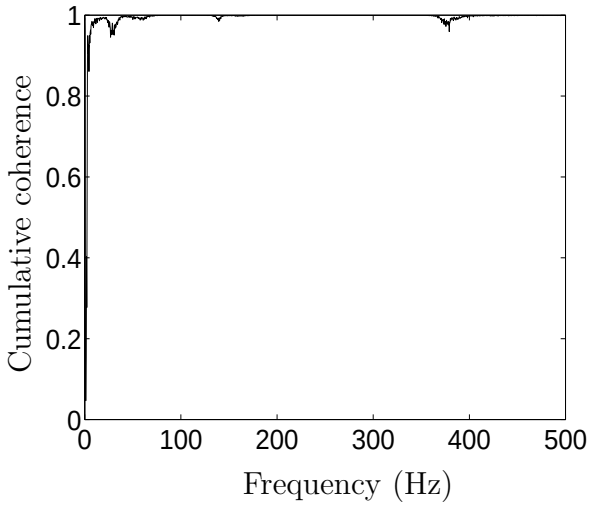


Figure 2.9: Horizontal set-up, cumulative coherence γ^2_{M7} (21 N level).

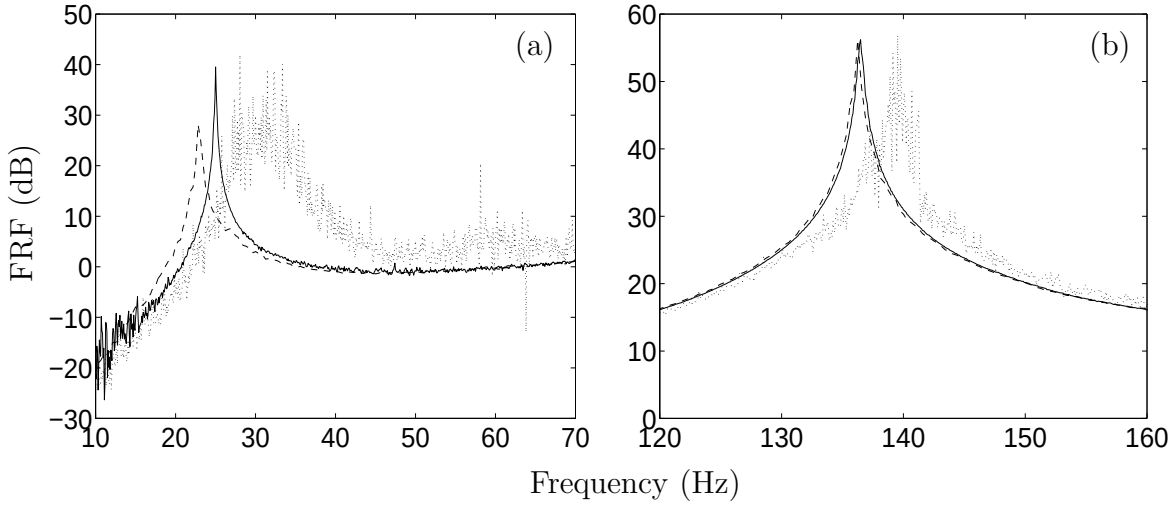


Figure 2.10: Horizontal set-up, magnitude of H_{72}: ——, Reference FRF (0.3 N); $\cdots$, H_2 estimate (21 N); - - -, H_{c2} estimate (21 N). (a) First resonance; (b) second resonance.

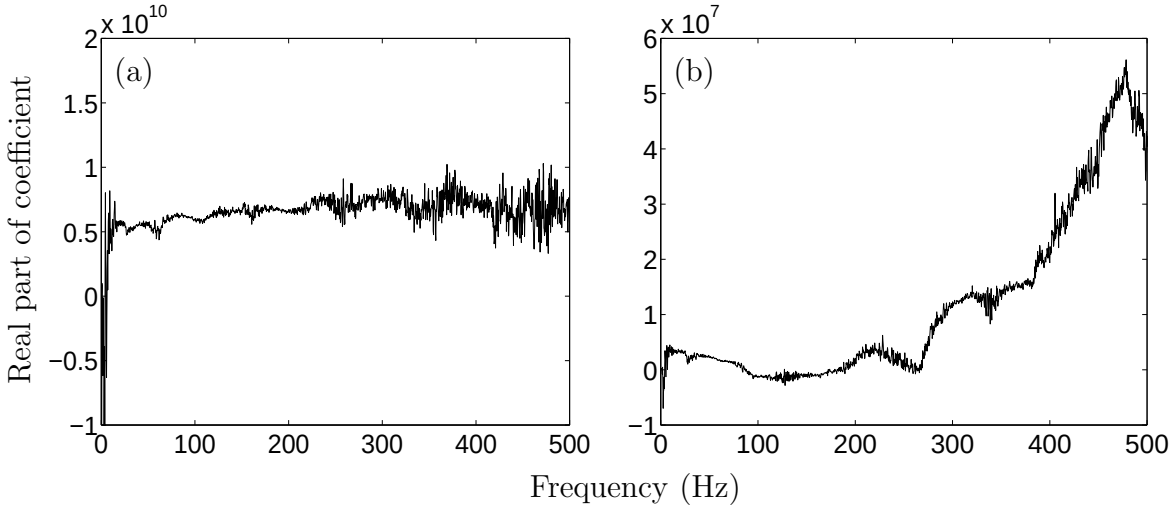


Figure 2.11: Horizontal set-up, real part of the non-linear coefficients (21 N). (a) A; (b) B.

the 21 N level. For the sake of comparison, the FRF computed from the H_2 estimate is also illustrated. It can clearly be seen that the H_{c2} estimate provides a closer match to the reference FRF than the H_2 estimate. However, if the second resonance is accurately predicted by the CRP method, it is not the case for the first resonance whose frequency is largely underestimated.

The results for the intermediate excitation levels, i.e., 8 N and 16 N, are similar to those of the 21 N level and are not presented.

Excitation level (N)	A (N/m^3)	B (N/m^2)
8	$6.65\ 10^9 - i\ 5.30\ 10^7$	$2.53\ 10^6 - i\ 1.45\ 10^4$
16	$6.41\ 10^9 + i\ 1.26\ 10^7$	$1.25\ 10^6 + i\ 2.64\ 10^4$
21	$6.29\ 10^9 + i\ 5.31\ 10^7$	$7.85\ 10^5 + i\ 5.91\ 10^4$

Table 2.2: Spectral mean (10-250 Hz) of the non-linear coefficients.

Estimation of the non-linear coefficients

Since the FRFs of the underlying linear system have been computed, the non-linear coefficients can be calculated. Figure 2.11 displays the real part of the non-linear coefficients A and B [see equation (2.26)] for the 21 N level. As pointed out previously, the coefficients are frequency dependent and a spectral mean has to be performed to obtain a single value for the coefficients. From the superposition of the reference FRF and the H_2 estimate, it can be observed that the non-linear behaviour mainly occurs between 10 and 250 Hz. Accordingly, the spectral mean of the non-linear coefficients is realised in this frequency range. Table 2.2 gives the spectral mean of A and B for the 8, 16 and 21 N levels. The real part of coefficient A is stable while the real part of coefficient B decreases as the level increases. The imaginary part of the coefficients, without any physical meaning, is several orders of magnitude below the real part.

Discussion of the results

From the previous results, it can be concluded that the H_{c2} estimate allows to remove, at least partially, the distortions introduced by non-linearities in FRFs. However, the results can be substantially improved. Although gravity is often ignored in structural dynamics, its role in this example should not be neglected. Indeed, gravity is at the origin of a static deflection of the two beams that imposes a non-negligible pre-stress in the thin beam.

In order to investigate the influence of gravity, it was decided to perform a numerical analysis through the definition of a finite element model of the beam. It should be noted that the model was not updated in order to closely match the dynamic characteristics of the experimental structure. For instance, the damping matrix was assumed to be proportional to the mass matrix. This is not a critical issue because qualitative results concerning the influence of gravity are sought.

The excitation level is similar to that corresponding to 22 N. The structural response is simulated using Newmark's algorithm and the simulation is carried out in two cases, i.e., with and without a gravity load. The FRFs are computed using the H_2 estimate and are represented in Figure 2.12. Several harmonics can be observed when the effect of gravity is taken into account which is not the case if gravity is neglected. In particular, a second harmonic of the first mode appears in the FRF, exactly as for the experimental measurements.

The second harmonics of the first mode may thus be assigned to gravity and cannot be modelled with an asymmetric non-linearity as in equation (2.26). This is the reason why coefficient B is found to vary significantly with frequency [cf. Figure 2.11(b)] and why its spectral mean decreases when the excitation level increases. This may also explain why the frequency of the first resonance is not well predicted by the CRP method.

A simple means of reducing drastically the influence of gravity is to built a set-up where the thin beam is vertical and where the shaker excites the structure in a horizontal plane. Indeed, for a vertical set-up, the stiffness of the thin beam in the vertical direction is much higher than for a horizontal set-up, and the initial pre-stress can be neglected.

2.4.3 Identification results with the CRP method: vertical set-up

Conventional frequency response estimation (H_2 estimate)

In this test configuration, six excitation levels are considered ranging from 1.4 N up to 22 N. In order to evaluate the influence of the non-linearity, the FRFs are first estimated using the linear H_2 estimate. Figure 2.13 displays the magnitude of H_{73} for the lowest (1.4 N) and highest (22 N) excitation levels. The corresponding ordinary coherence functions are represented in Figure 2.14. As for the horizontal set-up, the behaviour of the structure may be assumed to be linear for the lowest excitation level. The FRF in this case is again referred to as "reference FRF". The comparison of Figures 2.7 and 2.13 reveals that the second harmonic of the first mode has now disappeared which is an indication that the effect of gravity has been significantly reduced.

The natural frequencies are estimated in the 0-500 Hz range using the least squares complex exponential (LSCE) method [45, 69], and the results are listed in Table 2.3. It can be observed that the first two natural frequencies are shifted towards higher frequencies when the excitation level is increased. This is expected due to the stiffening effect of the thin beam. In the range of excitation levels considered here, the third natural frequency is not affected by the presence of the non-linearity.

Model selection

Because the effect of gravity is now negligible, the asymmetric non-linearity is no longer included in the model. The thin beam is thus modelled using a grounded symmetric non-linearity whose functional form is given by:

$$f_{NL}(x, \dot{x}) = A\,|x|^{\alpha}\,\mathrm{sign}(x) \tag{2.28}$$

The best results in terms of the spectral mean of the averaged cumulative coherence of all the sensors are obtained for $\alpha = 2.8$ (accuracy=0.9873). Figure 2.15 represents the cumulative coherence γ^2_{M7} for the 22 N level. It is emphasised that the value of

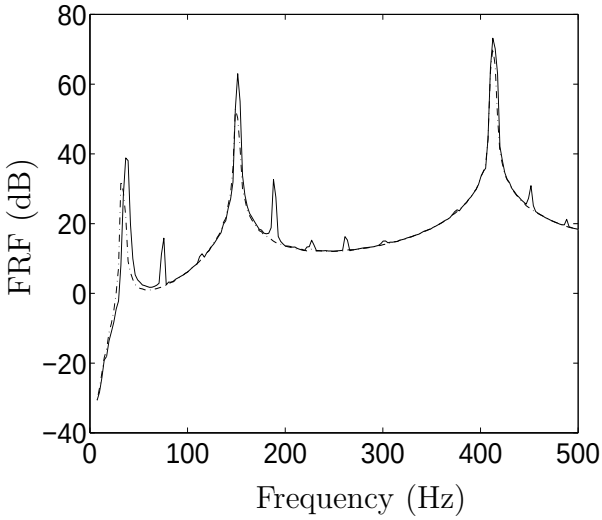


Figure 2.12: Numerical simulation of the horizontal set-up: – · –, without gravity; ——, with gravity.

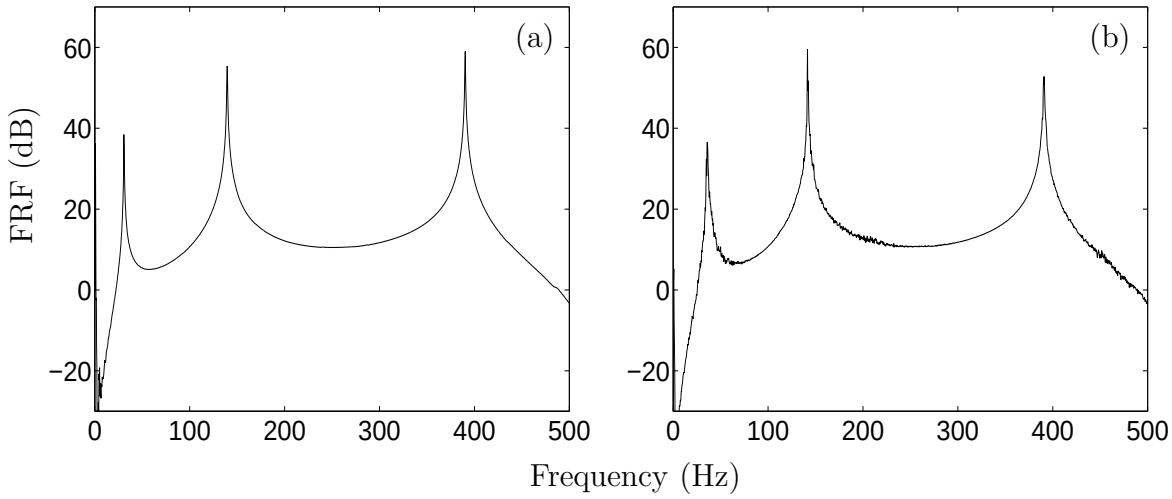


Figure 2.13: Vertical set-up, magnitude of H_{73} (H_2 estimate). (a) 1.4 N; (b) 22 N.

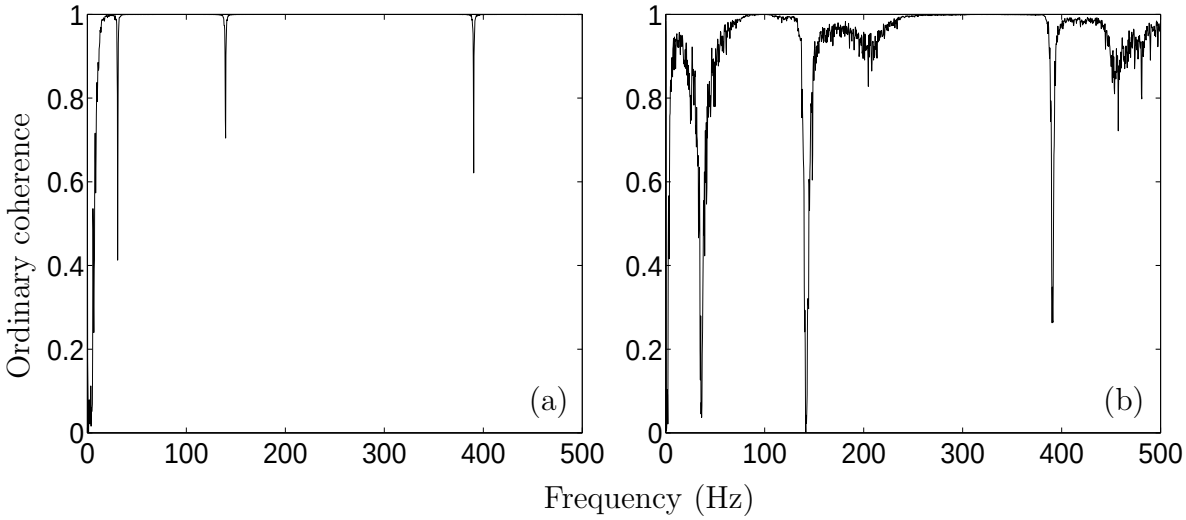


Figure 2.14: Vertical set-up, ordinary coherence functions $\gamma^2_{X_7F}$. (a) 1.4 N; (b) 22 N.

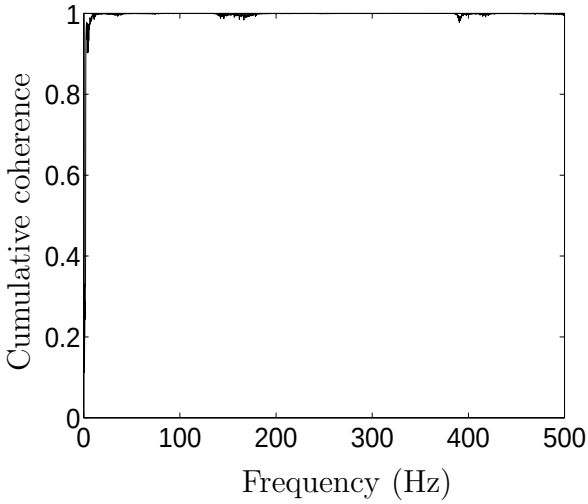


Figure 2.15: Vertical set-up, cumulative coherence γ^2_{M7} (22 N level).

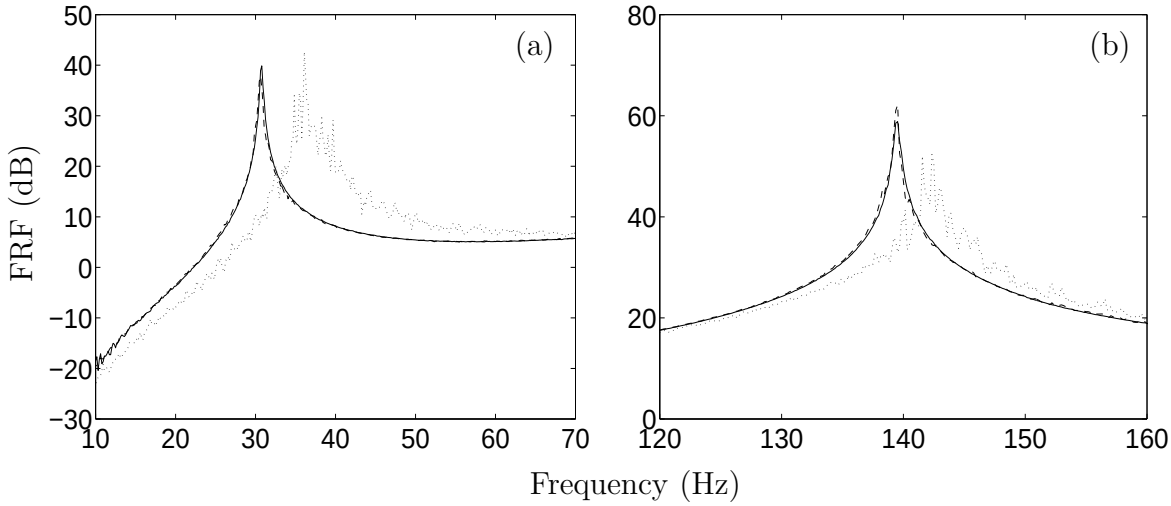


Figure 2.16: Vertical set-up, magnitude of H_{73}: ——, Reference FRF (1.4 N); $\cdots$, H_2 estimate (22 N); - - -, H_{c2} estimate (22 N). (a) First resonance; (b) second resonance.

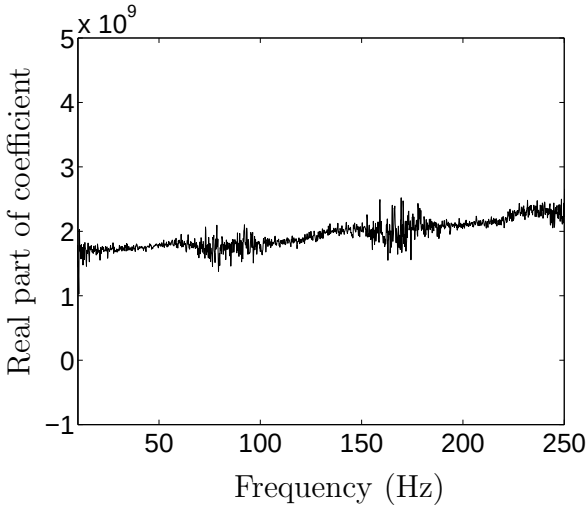


Figure 2.17: Vertical set-up, real part of the non-linear coefficient (22 N).

Excitation level (N)	1^{st} freq. (Hz)	2^{nd} freq. (Hz)	3^{rd} freq. (Hz)
1.4	30.74	139.47	390.22
2.8	30.99	139.49	390.16
5.5	31.60	139.64	390.12
11	33.13	140.29	390.17
16	34.81	140.84	390.34
22	36.80	141.80	390.54

Table 2.3: Natural frequencies (H_2 estimate).

Excitation level (N)	1^{st} freq. (Hz)	2^{nd} freq. (Hz)	3^{rd} freq. (Hz)
reference (1.4)	30.74	139.47	390.22
2.8	30.65	139.41	390.12
5.5	30.69	139.41	390.02
11	30.62	139.34	389.87
16	30.62	139.35	389.85
22	30.51	139.33	389.78

Table 2.4: Natural frequencies (H_{c2} estimate).

Excitation level (N)	A ($N/m^{2.8}$)
2.8	$2.69\ 10^9 - i\ 2.76\ 10^7$
5.5	$2.08\ 10^9 - i\ 6.07\ 10^7$
11	$1.94\ 10^9 + i\ 1.09\ 10^6$
16	$1.96\ 10^9 - i\ 6.20\ 10^6$
22	$1.96\ 10^9 + i\ 1.55\ 10^7$

Table 2.5: Spectral mean (10-250 Hz) of the non-linear coefficient.

the coherence function is higher than for the horizontal set-up although the asymmetric non-linearity has been removed from the model.

Conditioned frequency response estimation (H_{c2} estimate)

The FRFs of the underlying linear system are now computed using the H_{c2} estimate. Figure 2.16 presents the comparison between the reference FRF, i.e., the FRF for the 1.4 N level, and the FRF obtained using the H_{c2} estimate for the 22 N level. The FRF computed from the H_2 estimate is also illustrated. The H_{c2} estimate and the reference FRF agree to the point where the difference between signals is barely not visible. The natural frequencies estimated from the H_{c2} estimate are listed in Table 2.4. It can be observed that the first three natural frequencies are now well estimated whatever the

excitation level is.

The results for the intermediate excitation levels are similar to those of the 22 N level and are not presented.

Estimation of the non-linear coefficient

The last step of the identification procedure is the computation of the non-linear coefficient A. Figure 2.17 represents the real part of this coefficient, and its spectral mean is listed in Table 2.5. Apart from the 2.8 and 5.5 N levels for which the non-linearity does not participate sufficiently in the system response, a stable value for the non-linear coefficient is observed. Again, the imaginary part of the coefficient is several orders of magnitude below the real part.

2.4.4 Identification results with the NIFO method: vertical set-up

The NIFO method is now applied for the identification of the beam. Since better results have been obtained on the vertical set-up, the identification of the horizontal set-up is not considered here.

The functional form of the non-linearity is described in equation (2.28). As in the CRP method, the exponent α needs to be determined. Because the NIFO method does not involve the computation of the conditioned spectra, the cumulative coherence function cannot be calculated and the optimal exponent cannot be determined. This is the reason why the non-linearity exponent is set to be equal to the exponent found with the CRP method, i.e., $\alpha = 2.8$.

Figure 2.18 shows the FRFs given by the NIFO and H_2 estimates for the first two resonances (22 N level). These FRFs are compared with the reference FRF. It can be observed that the NIFO estimate closely matches the reference FRF. Nevertheless, the comparison of Figures 2.16 and 2.18 reveals that some distortions are introduced in the FRFs by the NIFO estimate. Hopefully, these distortions do not influence the natural frequencies, as shown in Table 2.6.

The real part of the non-linear coefficient is illustrated in Figure 2.19, and its spectral mean is listed in Table 2.7. The comparison between Figures 2.17 and 2.19 indicates that much larger deviations occur with the NIFO method. However, the spectral mean of the coefficient contained in Table 2.7 does not seem to be affected.

2.4.5 Comparison of the results

CRP and NIFO methods

The CRP and NIFO methods are similar in spirit in the sense that they both compute the FRFs of the underlying linear system together with the non-linear coefficients. These

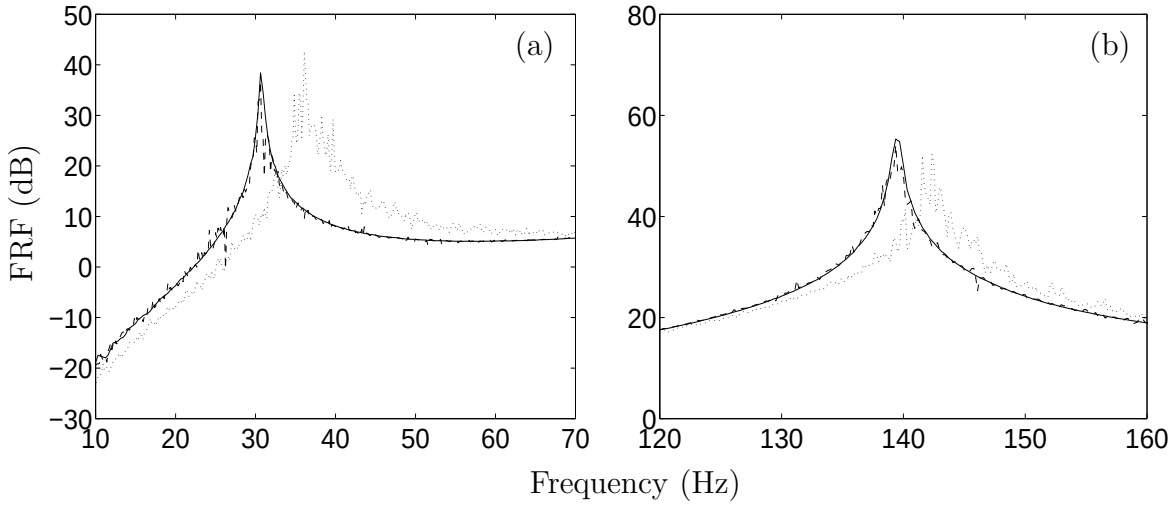


Figure 2.18: Vertical set-up, magnitude of H_{73}: ——, Reference FRF (1.4 N); $\cdots$, H_2 estimate (22 N); - - -, NIFO estimate (22 N). (a) First resonance; (b) second resonance.

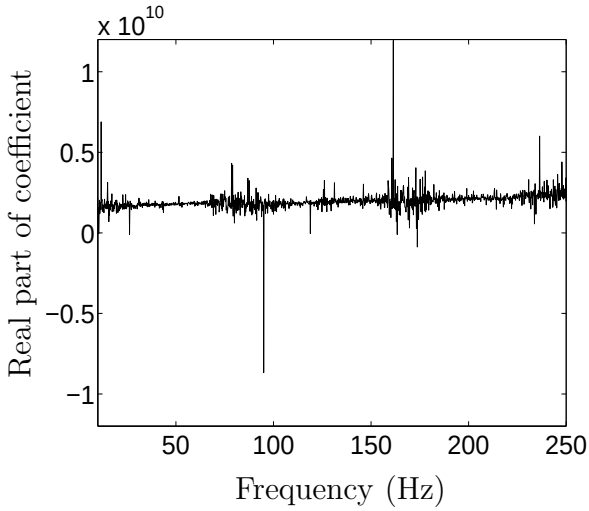


Figure 2.19: Vertical set-up, real part of the non-linear coefficient (22 N, NIFO).

techniques are compared in reference [92]. In summary, it can be stated that the NIFO technique is attractive for its simplicity and its capability to estimate the linear and non-linear coefficients in a single step. The correlation between the linear and non-linear terms is a critical issue in the NIFO method, and care must be taken to achieve a good conditioning of the data matrices. To this end, an orthogonal least-squares procedure has appeared to be useful. Nevertheless, from the results obtained on the non-linear beam, it can be observed that the NIFO method cannot offer the good conditioning that is naturally present in the CRP method. Numerical applications have also been performed to study this issue more carefully, and the same conclusion has been reached. Another advantage of the CRP technique is that it offers an efficient means of characterising the type of non-linearity through the use of the cumulative coherence function. This latter is also useful in the sense that it may be viewed as a measure of the model accuracy.

Excitation level (N)	1^{st} freq. (Hz)	2^{nd} freq. (Hz)	3^{rd} freq. (Hz)
reference (1.4)	30.74	139.47	390.22
2.8	30.74	139.43	390.15
5.5	30.75	139.38	390.04
11	30.75	139.26	389.92
16	30.67	139.35	389.81
22	30.63	139.33	389.77

Table 2.6: Natural frequencies (NIFO method).

Excitation level (N)	A ($N/m^{2.8}$)
2.8	$2.69\ 10^9 - i\ 1.46\ 10^8$
5.5	$1.99\ 10^9 - i\ 7.35\ 10^7$
11	$1.91\ 10^9 - i\ 3.21\ 10^7$
16	$1.95\ 10^9 + i\ 1.80\ 10^6$
22	$1.96\ 10^9 + i\ 2.28\ 10^7$

Table 2.7: Spectral mean (10-250 Hz) of the non-linear coefficients (NIFO method).

COST Action F3

Since the foregoing example was defined as benchmark within the framework of the European COST Action F3, it was also studied by five other research groups:

- Ecole Centrale de Lyon (France): Gibert *et al.* built a frequency response model based on the concept of non-linear normal mode (NNM) [60];
- Laboratoire d'Analyse des Matériaux et Identification ENPC-LCPC (Paris, France): Argoul *et al.* made use of the Wavelet transform [5];
- Laboratoire de Mécanique et d'Acoustique (Marseille, France): Bellizzi *et al.* defined an equivalent linear system with random parameters [12].
- University of Kassel (Germany): Meyer *et al.* linearised the non-linear equilibrium equations of the structure using the harmonic balance method and obtained a linear model in the frequency domain [129];
- Politecnico di Torino (Italy): Marchesiello *et al.* exploited the CRP method [125] and performed model selection using the partial and multiple coherence functions defined by Bendat in reference [14];

More details about all these methods are given in the introductory chapter. The proper orthogonal decomposition (POD) was also used for the identification but the discussion of the results is delayed until chapter 3.

A common issue in the study of this benchmark was the prediction of the properties of the first resonance frequencies of the beam. From the results obtained, it can be concluded that each method was able to reproduce the modifications of the properties of the first resonance frequency quite accurately. For the second frequency, discrepancies in the results were observed:

- Gibert *et al.* and Bellizzi *et al.* obtained a slight decrease in the frequency;
- Meyer *et al.* found that the second resonance frequency is not much affected and did not consider it during the identification;
- Marchesiello *et al.* and Argoul *et al.* pointed out an increase in frequency when the excitation level was increased.

However, it should be borne in mind that two different set-ups were built, one in Lyon (France) and one in Liège (Belgium), and that even for the same set-up, the methods used different kinds of excitation. The properties of the third resonance frequency were never found to be significantly influenced by the presence of the non-linearity.

2.5 Concluding remarks

Two frequency domain techniques called the CRP and NIFO methods have been investigated in this chapter.

The CRP formulation is based on the concept of a RP model but generalises its application to structures where non-linearities are away from the location of the applied force. The H_{c1} and H_{c2} estimates proposed by the method enable to eliminate the distortions introduced in the FRFs by non-linearities. In a second step, the non-linear parameters are estimated by taking advantage of the knowledge of the FRFs.

The NIFO method exploits spatial information and treats the non-linear forces as internal feedback forces in the underlying linear model of the system. The purpose is to derive a formulation that leads to obtain a much simpler parameter estimation method.

Both methods assume that the type and physical location of the non-linearity are known a priori. This represents a limitation on the practical use of the methods. However, this is a critical issue for many identification schemes in the field of non-linear structural dynamics. It is thus highly recommended to first characterise the underlying structure of the non-linear system.

Their overall performances have been assessed by means of an experimental structure consisting of a beam with a geometric non-linearity. This example was challenging in the sense that the structure is continuous and that several modes participate in the response. In addition, the gravity load is responsible for an initial pre-stress that has been neglected in references [5, 12, 60, 125, 129]. This is the reason why we have proposed to build a

vertical set-up for which the influence of gravity was drastically reduced. This enabled us to obtain accurate results with both methods.

Notwithstanding its ability to simultaneously identify linear and non-linear coefficients, the accuracy given by the NIFO method can by no means equal the one provided by the CRP formulation. It is our opinion that this latter method should be preferred even if its implementation requires a thorough knowledge of the spectral conditioning techniques that are complex by nature.

The CRP method thus appears as an efficient and reliable technique for the identification of the dynamic characteristics of continuous MDOF non-linear structures. In addition, the cumulative coherence function is a very useful tool for the model selection. The main drawback of the CRP approach is that it requires the measurements of the signals at the location of the non-linearity which is not always possible in practice. Another difficulty is that it identifies the non-linear coefficients and a modal model of the underlying linear system rather than a structural model of the structure. A particular class of identification techniques, i.e., finite element model updating techniques, overcomes this limitation and is described in the subsequent chapters.

Chapter 3

Proper Orthogonal Decomposition

3.1 Introduction

Many identification techniques cannot provide an estimation of the structural matrices, i.e., the mass, damping and stiffness matrices. This may represent a severe limitation when a structural model of the structure is needed, e.g., for simulation purposes. This is the reason why there has been recently a growth in interest in finite element model updating techniques of non-linear vibrating structures.

Model updating was first introduced in the context of linear structural dynamics. Some of its main characteristics have been briefly recalled in the introductory chapter. For a deeper discussion of the work in this field, we invite the interested reader to refer to the literature (e.g., [54, 115] and references therein).

The success of linear model updating mainly relies on two factors:

1. The global minimum of the objective function can accurately be located because the initial finite element model may often be assumed to be powerful enough to avoid being too far from this minimum.

2. Natural frequencies and mode shapes have the capability to reduce the amount of information that has to be analysed and provide useful insight into the dynamics of interest.

When investigating non-linear structures, serious problems are introduced which prevents us from using the strategies developed in the linear case.

A priori knowledge about the non-linearity is often limited, and the initial model cannot generally be assumed to be near the actual model. In that case, a local optimiser is likely to be trapped in a local optimum resulting in a poor solution. Global optimisation algorithms have been introduced but are far less developed than their linear counterpart.

The validity of modal analysis is limited to structures showing a linear behaviour. Accordingly, natural frequencies and mode shapes can no longer be exploited. An extension of the concept of mode shapes to discrete conservative non-linear systems was proposed

by Rosenberg in references [150, 151]. He defined a non-linear normal mode (NNM) as a synchronous periodic oscillation where all points pass through the equilibrium position at the same time. Other methodologies [156, 157] have been developed for the study of non-conservative, gyroscopic and continuous systems. These methods define NNMs in terms of motion that occurs on two-dimensional invariant manifolds in the system phase space. While the NNMs provide a valuable theoretical tool for understanding certain dynamic phenomena such as mode bifurcations or non-linear mode localization [175, 176], their computation still remains difficult for practical applications. It is therefore a little too early to tell whether they may be of substantial help in a model updating context.

This explains why model updating of non-linear vibrating structures is recognised to be a challenging task. In order that it may be handled in a efficient manner, the two aforementioned problems need to be addressed.

Firstly, the capability of optimisation strategies to locate the global minimum should be investigated. This is the main topic of chapter 5.

Secondly, the definition of features that accurately characterise the structural behaviour should help the structural dynamicist to investigate non-linear systems. In this context, the proper orthogonal decomposition (POD) investigated in the present chapter and its non-linear generalisations discussed in chapter 4 deserve our attention.

The POD is a multi-variate statistical method that aims at obtaining a compact representation of the data. This method may serve two purposes, namely data compression by projecting high-dimensional data into a lower-dimensional space and feature extraction by revealing relevant but unexpected structure hidden in the data. The key idea of the POD is to reduce a large number of interdependent variables to a much smaller number of uncorrelated variables while retaining as much as possible of the variation in the original variables. An orthogonal transformation to the basis of the eigenvectors of the sample covariance matrix is performed, and the data is projected onto the subspace spanned by the eigenvectors corresponding to the largest eigenvalues. This transformation decorrelates the signal components and maximises variance.

The most striking property of the POD is its optimality in the sense that it minimises the average squared distance between the original signal and its reduced linear representation. Although it is frequently applied to non-linear problems, it should be borne in mind that the optimality of the POD only holds within the class of linear methods. The linear nature of the method is appealing because the theory of linear operators is available but it also represents its major limitation when the data lies on a non-linear manifold. As stated in reference [19], the POD is "*a safe haven in the intimidating world of non-linearity; although this may not do the physical violence of linearisation methods*".

The POD, also known as the Karhunen-Loève decomposition, was proposed independently by several scientists including Karhunen [87], Kosambi [99], Loève [118], Obukhov [135] and Pougachev [143] and was originally conceived in the framework of continuous second-order processes (see reference [18] for a recent survey). When restricted to a finite

dimensional case and truncated after a few terms, the POD is equivalent to principal component analysis (PCA). This latter methodology originated with the work of Pearson [141] as a means of fitting planes by orthogonal least squares but was also proposed by Hotelling [79]. It is emphasised that the method appears in various guises in the literature and is known by other names depending on the area of application, namely PCA in the statistical literature, empirical orthogonal function in oceanography and meteorology, and factor analysis in psychology and economics.

The POD has emerged as a useful technique in numerous fields. For instance, several authors including Cazemier [28], Holmes [78] and Lumley [119] used the POD in the context of turbulence to extract coherent structures. Newman and Krishnaprasad [134] applied it to formulate reduced-order models for temperature field dynamics in a chemical reactor. Wax and Kailath [177] suggested its use to characterise signals in signal processing. Cottrell and Metcalfe [32] worked on the classification of faces and emotions in image processing. Barnston and Ropelewski [10] employed the method for the forecasting in meteorology. Leen *et al.* [106] introduced it as a means of classifying speech data.

The first applications of the POD in the field of structural dynamics date back to the early 1990s. Cusumano *et al.* [36] exploited the technique to estimate the intrinsic dimensionality of the dynamics of an impacting beam. Azeez and Vakakis [7] applied it to create low-dimensional models of an overhung rotor. Ma and Vakakis [122] studied non-linear effects by considering energy transfers between leading modes. Kreuzer and Kust [101] used the POD to control self-excited vibrations of long torsional strings. Kappagantu and Feeny [84] worked on the modal reduction of a frictionally excited system. References [66, 71, 109, 111] proposed the POD for model updating of non-linear structures. Other studies include the works of Benguedouar [17], Cusumano and Bai [35], Fitzsimons and Rui [50], Georgiou and Schwartz [57], Kappagantu and Feeny [85, 86], Kerschen *et al.* [93], Ma *et al.* [120, 121], Riaz and Feeny [146].

The chapter is organised as follows. The next section briefly reviews the mathematical formulation of the POD while the different means of computing the POD are presented in section 3.3. Section 3.4 provides insight into the physical interpretation of the modes extracted from the decomposition. The application of the technique to model reduction and model updating of non-linear structures is then explained. Finally, the method's strengths as well as its limitations are discussed in section 3.7.

3.2 Mathematical formulation

Let $\theta(x,t)$ be a random field on a domain Ω. This field is first decomposed into mean $\mu(x)$ and time varying parts $\vartheta(x,t)$:

$$\theta(x,t) = \mu(x) + \vartheta(x,t) \tag{3.1}$$

At time t_k, the system displays a snapshot $\vartheta^k(x) = \vartheta(x,t_k)$. The POD aims at

obtaining the most characteristic structure $\varphi(x)$ of an ensemble of snapshots of the field $\vartheta(x,t)$. This is equivalent to finding the basis function $\varphi(x)$ that maximises the ensemble average of the inner products between $\vartheta^k(x)$ and $\varphi(x)$:

$$\text{Maximise} \quad \langle |(\vartheta^k, \varphi)|^2 \rangle \qquad \text{with } \|\varphi\|^2 = 1 \tag{3.2}$$

where $(f,g) = \int_\Omega f(x)g(x)\, d\Omega$ denotes the inner product in Ω;
$\langle \cdot \rangle$ denotes the averaging operation;
$\| \cdot \| = (\cdot,\cdot)^{\frac{1}{2}}$ denotes the norm;
$| \cdot |$ denotes the modulus.

Expression (3.2) means that if the field ϑ is projected along φ, the average energy content is greater than if the field is projected along any other basis function.

The constraint $\|\varphi\|^2 = 1$, imposed to make the computation unique, can be taken into account by the use of a Lagrange multiplier:

$$J[\varphi] = \langle |(\vartheta, \varphi)|^2 \rangle - \lambda \left(\|\varphi\|^2 - 1 \right) \tag{3.3}$$

The extremum is reached when the functional derivative is equal to zero. Reference [78] shows that this condition reduces to the following integral eigenvalue problem:

$$\int_\Omega \langle \vartheta^k(x) \vartheta^k(x') \rangle \varphi(x')\, dx' = \lambda \varphi(x) \tag{3.4}$$

where $\langle \vartheta^k(x) \vartheta^k(x') \rangle$ is the averaged auto-correlation function.

The solution of the optimisation problem (3.2) is thus given by the orthogonal eigenfunctions $\varphi_i(x)$ of the integral equation (3.4), called the proper orthogonal modes (POMs). The corresponding eigenvalues λ_i ($\lambda_i \geq 0$) are the proper orthogonal values (POVs). The POMs may be used as a basis for the decomposition of the field $\vartheta(x,t)$:

$$\vartheta(x,t) = \sum_{i=1}^{\infty} a_i(t) \varphi_i(x) \tag{3.5}$$

where the coefficients $a_i(t)$ are uncorrelated, i.e., $\langle a_i(t) a_j(t) \rangle = \delta_{ij}\, \lambda_i$, and are determined by $a_i(t) = (\vartheta(x,t), \varphi_i(x))$.

The POM associated with the greatest POV is the optimal vector to characterise the ensemble of snapshots. The POM associated with the second greatest POV is the optimal vector to characterise the ensemble of snapshots but restricted to the space orthogonal to the first POM, and so forth. The energy E contained in the data is defined as the sum of the POVs, i.e., $E = \sum_j \lambda_j$, and the energy percentage captured by the ith POM is given by $\lambda_i / \sum_j \lambda_j$.

3.3 Practical computation of the proper orthogonal decomposition

In practice, the data is discretised in space and time. Accordingly, m observations of a n-dimensional vector $\mathbf{x}$ are collected, and an $(n \times m)$ response matrix is formed:

$$\mathbf{X} = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_m] = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \tag{3.6}$$

The purpose of this section is to explain how the POD can be computed in practical applications.

3.3.1 Eigensolutions of the sample covariance matrix

Since the data is now discretised and since it does not necessarily have a zero mean, the averaged auto-correlation function is replaced by the covariance matrix $\mathbf{\Sigma} = E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T]$ where $E[\cdot]$ is the expectation and $\boldsymbol{\mu} = E[\mathbf{x}]$ is the mean of the vector $\mathbf{x}$. Under the assumption that the process is stationary and ergodic and that the number of time instants is large, a reliable estimate of the covariance matrix is given by the sample covariance matrix:

$$\mathbf{S} = \frac{1}{m} \begin{bmatrix} \{\sum_{j=1}^{m}(x_{1j} - \frac{1}{m}\sum_{k=1}^{m} x_{1k})^2\} & \cdots & \{\sum_{j=1}^{m}(x_{1j} - \frac{1}{m}\sum_{k=1}^{m} x_{1k}) \\ & & (x_{nj} - \frac{1}{m}\sum_{k=1}^{m} x_{nk})\} \\ \cdots & \cdots & \cdots \\ \{\sum_{j=1}^{m}(x_{nj} - \frac{1}{m}\sum_{k=1}^{m} x_{nk}) & \cdots & \{\sum_{j=1}^{m}(x_{nj} - \frac{1}{m}\sum_{k=1}^{m} x_{nk})^2\} \\ (x_{1j} - \frac{1}{m}\sum_{k=1}^{m} x_{1k})\} & & \end{bmatrix} \tag{3.7}$$

The POMs and POVs are thus characterised by the eigensolutions of the sample covariance matrix $\mathbf{S}$. If the data is zero mean, the sample covariance is merely given by the following expression:

$$\mathbf{S} = \frac{1}{m}\mathbf{X}\mathbf{X}^T \tag{3.8}$$

It should be noted that when the number of degree-of-freedom (DOF) is much higher than the number of snapshots (e.g., in turbulence theory), the computation of the sample covariance matrix may become expensive. In this context, the POD modes are most

easily computed using the method of snapshots proposed by Sirovich in reference [159]. However, the description of this technique is beyond the scope of this dissertation.

Interpretation of the eigenvalue problem

An interesting interpretation of the eigenvalue problem is that if a matrix is real, symmetric and positive definite, then the eigenvectors of the matrix are the principal axes of the associated quadratic form that is an n-dimensional ellipsoid centred at the origin of the Euclidean space [128]. Accordingly, the POMs as eigenvectors of the sample covariance matrix $\mathbf{S}$ are the principal axes of the family of ellipsoids defined by:

$$\mathbf{z}^T\mathbf{S}\mathbf{z} = c \tag{3.9}$$

where $\mathbf{z}$ is a real non-zero vector and c is a positive constant.

The result is statistically important if the data has a multi-variate normal distribution. In this case, the ellipsoids given by (3.9) define contours of constant probability [133].

3.3.2 Singular value decomposition of the response matrix

For any real $(m \times n)$ matrix $\mathbf{X}$, i.e., for our purpose, the response matrix measured about its mean, there exists a real factorisation called the singular value decomposition (SVD) that can be written:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{3.10}$$

where $\mathbf{U}$ is an $(m \times m)$ orthonormal matrix. Its columns form the left singular vectors.

$\mathbf{\Sigma}$ is an $(m \times n)$ pseudo-diagonal and semi-positive definite matrix with diagonal entries containing the singular values σ_i.

$\mathbf{V}$ is an $(n \times n)$ orthonormal matrix. Its columns form the right singular vectors.

It is important to point out that the matrix $\mathbf{\Sigma}$ defined in this context is not the covariance matrix. It is fairly standard notation to use $\mathbf{\Sigma}$ as well for the covariance matrix as for the matrix containing the singular values. However, it will always be obvious from the context whether one or the other is considered.

Reliable algorithms have been developed to compute the SVD [62]. The SVD may also be calculated by means of solving two eigenvalue problems, or even one if only the left or the right singular vectors are required. Indeed,

$$\begin{aligned} \mathbf{X}\mathbf{X}^T &= \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^T \\ \mathbf{X}^T\mathbf{X} &= \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T \end{aligned} \tag{3.11}$$

According to equations (3.11), the singular values of $\mathbf{X}$ are found to be the square roots of the eigenvalues of $\mathbf{X}\mathbf{X}^T$ or $\mathbf{X}^T\mathbf{X}$. In addition, the left and right singular vectors of $\mathbf{X}$ are the eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\mathbf{X}^T\mathbf{X}$ respectively. The POMs, defined as the eigenvectors of the sample covariance matrix $\mathbf{S}$ (3.8), are thus equal to the left singular vectors of $\mathbf{X}$.

The POVs, defined as the eigenvalues of matrix $\mathbf{S}$, are the square of the singular values divided by the number of samples m.

The main advantage in considering the SVD to compute the POD instead of the eigenvalue problem described in section 3.3.1 is that additional information is obtained through the matrix $\mathbf{V}$. The column $\mathbf{v}_i$ of matrix $\mathbf{V}$ contains the time modulation of the corresponding POM $\mathbf{u}_i$, normalised by the singular value σ_i. This information provides important insight into the system dynamics and plays a prominent role in the model updating of non-linear systems.

Interpretation of the singular value decomposition

The SVD of a matrix $\mathbf{X}$, seen as a collection of column vectors, provides important insight into the oriented energy distribution of this set of vectors. It is worth recalling that:

1. The energy of a vector sequence $\mathbf{x}_i$ building an $(n \times m)$ matrix $\mathbf{X}$ is defined via the Frobenius norm:

$$E(\mathbf{X}) = \|\mathbf{X}\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}^2 = \sum_{k=1}^{p} \sigma_k^2 \quad \text{where } p = \min(m,n) \tag{3.12}$$

   so that the energy of a vector sequence is equal to the energy in its singular spectrum $\sigma_1^2, ..., \sigma_p^2$;

2. The oriented energy of a vector sequence in some direction k with unit vector $\mathbf{e}_k$ is the sum of squared projections of the vectors onto direction k:

$$E_k(\mathbf{X}) = \sum_{i=1}^{m} (\mathbf{e}_k^T \, \mathbf{x}_i)^2 \tag{3.13}$$

One essential property of SVD is that extrema in this oriented energy distribution occur at each left singular direction [138, 163]. The oriented energy measured in the direction of the kth left singular vector is equal to the kth singular value squared. Since the POMs are equal to the left singular vectors, it can be stated that they are optimal with respect to energy content in a least square sense, i.e., they capture more energy per mode than any other set of basis functions.

3.4 Physical interpretation of the proper orthogonal modes

The connection between the modes obtained from the POD, i.e., the POMs, that are statistical by nature, and the underlying dynamics of the system under consideration is investigated in this section. The purpose is to reveal whether a physical interpretation

can be attributed to the POMs. In particular, it is inquired under which circumstances the POMs are related to vibration eigenmodes.

Two different ways of interpreting the POMs were proposed in the literature [47, 91]. Feeny and Kappagantu [47] sketched some relationships between the POMs and the modes of vibration by considering the POMs as eigenvectors of the sample covariance matrix. In the present study, the emphasis is shifted towards the SVD of the response matrix [91]. The physical interpretation of the POMs of discrete linear systems for the free response in the undamped and damped cases, for the harmonic response, and for the stationary random response to a white noise excitation is first studied. A geometric approach to the comparison between the vibration eigenmodes and the POMs is then proposed. Finally, the relationships between the NNMs and the POMs are investigated.

3.4.1 Undamped and unforced linear systems

The equation of motion of an undamped and unforced linear system with n DOF may be written as follows:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = 0 \qquad \text{with } \mathbf{x}(0) = \mathbf{x}_0 \text{ and } \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0 \tag{3.14}$$

where $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices respectively;
$\mathbf{x}$ is the vector of displacement co-ordinates.

The system response due to initial conditions may be expressed as:

$$\mathbf{x}(t) = \sum_{i=1}^{n} (C_i \cos \omega_i t + D_i \sin \omega_i t)\, \mathbf{p}_{(i)} = \sum_{i=1}^{n} a_i(t)\, \mathbf{p}_{(i)} \tag{3.15}$$

where $\omega_i, \mathbf{p}_{(i)}$ are the natural frequencies (in rad/s) and eigenmodes of the system;
C_i and D_i are constants depending on the initial conditions;
$a_i(t) = C_i \cos \omega_i t + D_i \sin \omega_i t$ represents the time modulation of mode $\mathbf{p}_{(i)}$.

The time discretisation of the system response leads to m sampled values of the time functions that form an $(n \times m)$ matrix whose columns are the members of the data ensemble:

$$\begin{aligned} \mathbf{X} &= \left[\begin{array}{ccc} \mathbf{x}(t_1) & \cdots & \mathbf{x}(t_m) \end{array} \right] \\ &= \left[\begin{array}{ccc} \displaystyle\sum_{i=1}^{n} a_i(t_1)\, \mathbf{p}_{(i)} & \cdots & \displaystyle\sum_{i=1}^{n} a_i(t_m)\, \mathbf{p}_{(i)} \end{array} \right] \end{aligned} \tag{3.16}$$

which can also be written as:

$$\begin{aligned}
\mathbf{X} &= \begin{bmatrix} \mathbf{p}_{(1)} & \cdots & \mathbf{p}_{(n)} \end{bmatrix} \begin{bmatrix} a_1(t_1) & \cdots & a_1(t_m) \\ \cdots & & \cdots \\ a_n(t_1) & \cdots & a_n(t_m) \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{p}_{(1)} & \cdots & \mathbf{p}_{(n)} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^T \\ \cdots \\ \mathbf{a}_n^T \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{p}_{(1)} & \cdots & \mathbf{p}_{(n)} \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}^T \\
&= \mathbf{P}\mathbf{A}^T \\
&= \mathbf{P} \begin{bmatrix} \mathbf{I} & \mathbf{Z} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}^T
\end{aligned} \tag{3.17}$$

where $\mathbf{a}_i = \begin{bmatrix} a_i(t_1) & \cdots & a_i(t_m) \end{bmatrix}^T$;

$\mathbf{P}$ is the $(n \times n)$ modal matrix whose columns are the eigenmodes of the system;

$\mathbf{I}$ is an $(n \times n)$ identity matrix;

$\mathbf{Z}$ is an $(n \times (m-n))$ matrix full of zeros;

$\mathbf{A}$ is an $(m \times n)$ matrix whose columns are the functions $a_i(t)$ at times $t_1, \cdots, t_m$;

$\mathbf{B}$ is an $(m \times (m-n))$ matrix.

It is worthwhile noticing that $\mathbf{B}$ does not influence $\mathbf{X}$ since it is multiplied by the matrix $\mathbf{Z}$ that is full of zeros. Equation (3.17) can be expressed in a more familiar form:

$$\mathbf{X} = \underbrace{\begin{bmatrix} \mathbf{P} \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} \mathbf{I} & \mathbf{Z} \end{bmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}^T}_{\mathbf{V}^T} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{3.18}$$

Accordingly, the above decomposition of $\mathbf{X}$ may be thought of as the SVD of this matrix. However, this decomposition requires matrices $\mathbf{U}$ and $\mathbf{V}$ to be orthonormal as mentioned in section 3.3.2. The next step is thus to determine when the columns of $\mathbf{U}(\equiv \mathbf{P})$ and $\mathbf{V}(\equiv [\mathbf{A}\ \ \mathbf{B}])$ are orthogonal.

1. The columns of $\mathbf{U}$ are formed by the eigenmodes of the structure. The eigenmodes are orthogonal to each other in the metrics of the mass and stiffness matrices. If the mass matrix is proportional to the identity matrix, it turns out that $\mathbf{p}_{(i)}^T \mathbf{p}_{(j)} = \delta_{ij}$. Consequently, $\mathbf{U}$ is orthogonal if the mass matrix is proportional to the identity matrix.

2. It remains to determine when the columns of $\mathbf{V}$ are orthogonal. For this purpose, equation (3.17) may be written as follows:

$$\begin{aligned}
\mathbf{X} &= \mathbf{P}\begin{bmatrix}\mathbf{I} & \mathbf{Z}\end{bmatrix}\begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}^T \\
&= \mathbf{P}\begin{bmatrix}\operatorname{diag}(\|\mathbf{a}_i\|) & \mathbf{Z}\end{bmatrix}\begin{bmatrix}\operatorname{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A} & \mathbf{B}\end{bmatrix}^T \\
&= \begin{bmatrix}\mathbf{p}_{(1)} \cdots \mathbf{p}_{(n)}\end{bmatrix}\begin{bmatrix}\|\mathbf{a}_1\| & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \|\mathbf{a}_2\| & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \|\mathbf{a}_n\| & 0 & \cdots & 0\end{bmatrix}\begin{bmatrix}\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}\cdots\frac{\mathbf{a}_n}{\|\mathbf{a}_n\|}\mathbf{B}\end{bmatrix}^T
\end{aligned} \tag{3.19}$$

If the natural frequencies ω_i are distinct, it can be easily argued that the columns of $\operatorname{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A}$ are orthogonal if an infinite set of sampled values is considered, i.e.,

$$\frac{\mathbf{a}_i}{\|\mathbf{a}_i\|}\frac{\mathbf{a}_j}{\|\mathbf{a}_j\|} \to 0 \quad \text{if } m \to \infty,\ i \neq j \tag{3.20}$$

Since $\mathbf{B}$ does not have an influence on $\mathbf{X}$, its columns can be computed in order that they are orthogonal to those of $\operatorname{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A}$. As can also be seen from equation (3.19), the POD is a bi-orthogonal decomposition that uncouples the spatial and time information contained in the data.

In conclusion, if the mass matrix is proportional to the identity matrix and if the number of samples is infinite, the SVD of $\mathbf{X}$ is such that:

1. The columns of $\mathbf{U}$ are the eigenmodes;

2. The first n columns of $\mathbf{V}$ are the normalised time modulations of the modes.

As stated in section 3.3.2, the POD basis vectors are the columns of the matrix $\mathbf{U}$ in the SVD of the displacement matrix. Therefore, it can be concluded that the POMs converge to the eigenmodes of an undamped and unforced linear system whose mass matrix is proportional to identity whether a sufficient number of samples is considered. This result is in concordance with the conclusion obtained by Feeny and Kappagantu [47]. As already mentioned, their demonstration relies on the fact that the POMs are the eigenvectors of the covariance matrix.

In the case of a mass matrix not proportional to identity, the POMs no longer converge to the eigenmodes since the former are orthogonal to each other while the latter are orthogonal with respect to the mass matrix. However, if the mass matrix is assumed to be known, it is still possible to retrieve the eigenmodes from the POMs. Applying the co-ordinate transformation $\mathbf{x} = \mathbf{M}^{-\frac{1}{2}}\,\mathbf{x}^*$ to equation (3.14) yields:

$$\ddot{\mathbf{x}}^* + \mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{M}^{-\frac{1}{2}}\mathbf{x}^* = 0 \tag{3.21}$$

In equation (3.21), the system matrices are still symmetric while the effective mass matrix is equal to the identity. Therefore, the left singular vectors of $\mathbf{X}^* = [\mathbf{x}^*(t_1) \cdots \mathbf{x}^*(t_m)]$, i.e., the POMs, converge to the eigenmodes $\mathbf{p}^*_{(i)}$ of this system. It is a simple matter

to demonstrate that the eigenmodes $\mathbf{p}_{(i)}$ of system (3.14) are related to those of system (3.21) by the following relationship:

$$\mathbf{p}_{(i)} = \mathbf{M}^{-\frac{1}{2}}\mathbf{p}^*_{(i)} \tag{3.22}$$

This section has investigated the discrete case. A detailed study of distributed systems can be found in reference [46] and underlines that the previous conclusions hold if the distributed system is uniformly discretised.

3.4.2 Damped and unforced linear systems

Consider now a damped but still unforced linear system with n DOF for which the equation of motion is given as follows:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = 0 \qquad \text{with } \mathbf{x}(0) = \mathbf{x}_0 \text{ and } \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0 \tag{3.23}$$

If the structure is lightly damped or with the assumption of modal damping, the system response can be readily written:

$$\mathbf{x}(t) = \sum_{i=1}^{n} C_i \, \exp^{-\epsilon_i\omega_i t} \, \cos(\sqrt{1-\epsilon_i^2}\,\omega_i t + \alpha_i)\,\mathbf{p}_{(i)} = \sum_{i=1}^{n} a_i(t)\,\mathbf{p}_{(i)} \tag{3.24}$$

where C_i, α_i are constants depending on the initial conditions;
ϵ_i is the modal damping coefficient.

Using the same procedure as in the previous section yields:

$$\begin{aligned}
\mathbf{X} &= \left[\begin{array}{ccc} \mathbf{x}(t_1) & \cdots & \mathbf{x}(t_m) \end{array}\right] \\
&= \left[\begin{array}{ccc} \displaystyle\sum_{i=1}^{n} a_i(t_1)\,\mathbf{p}_{(i)} & \cdots & \displaystyle\sum_{i=1}^{n} a_i(t_m)\,\mathbf{p}_{(i)} \end{array}\right] \\
&= \left[\begin{array}{ccc} \mathbf{p}_{(1)} & \cdots & \mathbf{p}_{(n)} \end{array}\right] \left[\begin{array}{c} \mathbf{a}_1^T \\ \cdots \\ \mathbf{a}_n^T \end{array}\right] \\
&= \mathbf{P}\mathbf{A}^T \\
&= \mathbf{P}\left[\begin{array}{cc} \mathbf{I} & \mathbf{Z} \end{array}\right]\left[\begin{array}{cc} \mathbf{A} & \mathbf{B} \end{array}\right]^T \\
&= \mathbf{P}\left[\begin{array}{cc} \mathrm{diag}(\|\mathbf{a}_i\|) & \mathbf{Z} \end{array}\right]\left[\begin{array}{cc} \mathrm{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A} & \mathbf{B} \end{array}\right]^T \\
&= \left[\mathbf{p}_{(1)} \cdots \mathbf{p}_{(n)}\right]\left[\begin{array}{ccccccc} \|\mathbf{a}_1\| & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \|\mathbf{a}_2\| & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & & & & & & \cdots \\ 0 & 0 & \cdots & \|\mathbf{a}_n\| & 0 & \cdots & 0 \end{array}\right]\left[\frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}\frac{\mathbf{a}_2}{\|\mathbf{a}_2\|}\cdots\frac{\mathbf{a}_n}{\|\mathbf{a}_n\|}\mathbf{B}\right]^T \\
&= \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T
\end{aligned} \tag{3.25}$$

where

$$\mathbf{a}_i = \left[C_i \exp^{-\epsilon_i \omega_i t_1} \cos(\sqrt{1-\epsilon_i^2}\,\omega_i t_1 + \alpha_i) \quad \cdots \quad C_i \exp^{-\epsilon_i \omega_i t_m} \cos(\sqrt{1-\epsilon_i^2}\,\omega_i t_m + \alpha_i) \right]^T \tag{3.26}$$

Again, the columns of $\mathbf{U}(\equiv \mathbf{P})$ are orthogonal if the mass matrix is proportional to the identity matrix. The main difference with the undamped case is that the time modulations $a_i(t) \to 0$ if $t \to \infty$ since the system returns to the equilibrium position in a finite time. Consequently, it can no longer be assumed that $\|\mathbf{a}_i\| \to \infty$ if $m \to \infty$ and that the columns of $\mathrm{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A}$ are orthogonal to each other. One might reasonably expect the POMs to be different from the eigenmodes. However, if the damping is low and if a sufficient number of points are considered, $\mathrm{diag}(\|\mathbf{a}_i\|^{-1})\,\mathbf{A}$ is almost orthogonal.

In conclusion, the POMs of a lightly damped and unforced linear system are a very good approximation of the eigenmodes of this system. This is in concordance with the result obtained in reference [47] using the eigensolution perspective.

3.4.3 Response of a linear system to harmonic excitation

This section is divided into two parts. Firstly, the global response of a linear system to harmonic excitation is considered. By global response, we mean the combination of the free and forced responses. Secondly, attention is focused only on the forced response of the linear system.

Global response to harmonic excitation

The equation of motion of a linear system with n DOF excited by an harmonic force with a constant amplitude is:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f} \sin \omega_e t \tag{3.27}$$

Equation (3.27) may be transformed by introducing a new variable $s = \sin \omega_e t$ that accounts for the harmonic force:

$$\begin{cases} \mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}s \\ \ddot{s} + \omega_e^2 s = 0 \end{cases} \quad \text{with } s(0) = 0, \dot{s}(0) = \omega_e \tag{3.28}$$

which yields:

$$\overbrace{\begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}}^{\mathbf{M}^*} \begin{bmatrix} \ddot{\mathbf{x}} \\ \ddot{s} \end{bmatrix} + \overbrace{\begin{bmatrix} \mathbf{K} & -\mathbf{f} \\ \mathbf{0} & \omega_e^2 \end{bmatrix}}^{\mathbf{K}^*} \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 0 \end{bmatrix} \tag{3.29}$$

For the sake of clarity, note $(\omega_i^2, \mathbf{p}_{(i)})$ the eigensolutions of the initial system (3.27) and $(\omega_i^{*2}, \mathbf{p}^*_{(i)})$ the eigensolutions of the transformed system (3.28). This latter system may be viewed as an unforced system with $n+1$ DOF [see equation (3.29)]. If the mass matrix is proportional to identity and if the number of samples is large enough, section

3.4.1 allows to conclude that the POMs of the transformed system response converge to the eigenmodes of that system.

Let us now compute the eigenmodes of the transformed system. These are the solutions of:

$$(\mathbf{K}^* - \omega_i^{*2}\mathbf{M}^*)\mathbf{p}^*_{(i)} = 0 \tag{3.30}$$

where ω_i^{*2} is a root of the algebraic equation:

$$\det(\mathbf{K}^* - \omega^2\mathbf{M}^*) = \det\left(\begin{bmatrix} \mathbf{K} & -\mathbf{f} \\ \mathbf{0} & \omega_e^2 \end{bmatrix} - \omega^2 \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}\right) = 0 \tag{3.31}$$

This equation becomes:

$$\det\left(\begin{bmatrix} \mathbf{K} - \omega^2\mathbf{M} & -\mathbf{f} \\ \mathbf{0} & \omega_e^2 - \omega^2 \end{bmatrix}\right) = \left(\omega_e^2 - \omega^2\right) \det\left(\mathbf{K} - \omega^2\mathbf{M}\right) = 0 \tag{3.32}$$

As can be seen from equation (3.32), the transformed system has $n+1$ eigenvalues. n eigenvalues are equal to those of the initial system (3.27):

$$\omega_i^{*2} = \omega_i^2 \quad \text{with } i = 1, ..., n \tag{3.33}$$

and the additional eigenvalue is equal to the square of the excitation frequency:

$$\omega_{n+1}^{*2} = \omega_e^2 \tag{3.34}$$

The eigenmodes corresponding to these eigenvalues may now be calculated. As illustrated in equation (3.30), the eigenmodes are the solutions of:

$$(\mathbf{K}^* - \omega_i^{*2}\mathbf{M}^*)\mathbf{p}^*_{(i)} = \begin{bmatrix} \mathbf{K} - \omega_i^{*2}\mathbf{M} & -\mathbf{f} \\ \mathbf{0} & \omega_e^2 - \omega_i^{*2} \end{bmatrix} \mathbf{p}^*_{(i)} = 0 \tag{3.35}$$

For $\omega_i^{*2} = \omega_i^2$, an obvious solution of system (3.35) is:

$$\mathbf{p}^*_{(i)} = \begin{bmatrix} \mathbf{p}_{(i)} \\ 0 \end{bmatrix} \tag{3.36}$$

Accordingly, the eigenmodes of the transformed system corresponding to ω_i^2 have the first n components equal to the eigenmodes of the initial system. The last component is equal to 0.

It remains to evaluate the eigenmode corresponding to ω_e^2, i.e., $\mathbf{x}^*_{(n+1)}$. With this aim, finding the eigensolutions of the transformed system is also equivalent to finding the eigensolutions of the following matrix:

$$\mathbf{M}^{*-1}\mathbf{K}^* = \begin{bmatrix} \mathbf{M}^{-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{K} & -\mathbf{f} \\ \mathbf{0} & \omega_e^2 \end{bmatrix} = \begin{bmatrix} \mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{f} \\ \mathbf{0} & \omega_e^2 \end{bmatrix} \tag{3.37}$$

and,

$$\begin{bmatrix} \mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{f} \\ \mathbf{0} & \omega_e^2 \end{bmatrix} \begin{bmatrix} \mathbf{p}^*_{(1)} & \cdots & \mathbf{p}^*_{(n)} & \mathbf{p}^*_{(n+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^*_{(1)} & \cdots & \mathbf{p}^*_{(n)} & \mathbf{p}^*_{(n+1)} \end{bmatrix} \begin{bmatrix} \mathrm{diag}(\omega_1^{*2}, \cdots, \omega_n^{*2}) & \mathbf{0} \\ \mathbf{0} & \omega_{n+1}^{*2} \end{bmatrix} \tag{3.38}$$

$$\left[\begin{array}{c|c} \mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{f} \\ \hline \mathbf{0} & \omega_e^2 \end{array}\right] \left[\begin{array}{c|c} \mathbf{P} & \mathbf{p}^*_{(n+1),1:n} \\ \hline \mathbf{0} & p^*_{(n+1),n+1} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{P} & \mathbf{p}^*_{(n+1),1:n} \\ \hline \mathbf{0} & p^*_{(n+1),n+1} \end{array}\right] \left[\begin{array}{c|c} \mathbf{\Omega} & \mathbf{0} \\ \hline \mathbf{0} & \omega_e^2 \end{array}\right] \tag{3.39}$$

where $\mathbf{P} = [\mathbf{p}_{(1)} \cdots \mathbf{p}_{(n)}]$ and $\mathbf{\Omega} = \mathrm{diag}(\omega_1^2, \cdots, \omega_n^2)$ are the eigensolutions of the initial system. It follows from equation (3.39) that:

$$\mathbf{M}^{-1}\mathbf{K}\mathbf{P} = \mathbf{P}\mathbf{\Omega} \tag{3.40}$$

$$\mathbf{M}^{-1}\mathbf{K}\mathbf{p}^*_{(n+1),1:n} - \mathbf{M}^{-1}\mathbf{f}p^*_{(n+1),n+1} = \mathbf{p}^*_{(n+1),1:n}\omega_e^2 \tag{3.41}$$

$$\mathbf{0} = \mathbf{0} \tag{3.42}$$

$$\omega_e^2 p^*_{(n+1),n+1} = p^*_{(n+1),n+1}\omega_e^2 \tag{3.43}$$

Equation (3.41) allows us to calculate the first n components of the last eigenmode $\mathbf{p}^*_{(n+1)}$:

$$\mathbf{p}^*_{(n+1),1:n} = \left[\mathbf{M}^{-1}\mathbf{K} - \omega_e^2\mathbf{I}\right]^{-1}\mathbf{M}^{-1}\mathbf{f}p^*_{(n+1),n+1} = \left[\mathbf{K} - \omega_e^2\mathbf{M}\right]^{-1}\mathbf{f}p^*_{(n+1),n+1} \tag{3.44}$$

$[\mathbf{K} - \omega_e^2\mathbf{M}]^{-1}$ is the dynamic compliance matrix and its spectral expansion is given in reference [58]:

$$\left[\mathbf{K} - \omega_e^2\mathbf{M}\right]^{-1} = \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}^T_{(i)}}{(\omega_i^2 - \omega_e^2)\mu_i} \tag{3.45}$$

Let us now introduce the spectral expansion (3.45) in equation (3.44):

$$\mathbf{p}^*_{(n+1),1:n} = \left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}^T_{(i)}}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f}p^*_{(n+1),n+1} \tag{3.46}$$

Since an eigenmode is defined to a scale factor and since $p^*_{(n+1),n+1}$ is a scalar, the final expression for the eigenmode corresponding to ω_e^2 is:

$$\mathbf{p}^*_{(n+1),1:n} = \left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}^T_{(i)}}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f} \tag{3.47}$$

In conclusion, let us consider a matrix that contains the response of the transformed system (3.28), i.e., its first n rows contain the response of the initial system (3.27) and its $(n+1)$th row is the applied force:

$$\mathbf{X}^* = \begin{bmatrix} \mathbf{x}(t_1) & \cdots & \mathbf{x}(t_m) \\ s(t_1) & \cdots & s(t_m) \end{bmatrix} \tag{3.48}$$

This matrix has $n+1$ POMs that have $n+1$ components. The dominant POM is related to the forced harmonic response of the system as shown in the next section, and its first n components are given by equation (3.47). In addition, if the mass matrix is proportional to identity, the first n components of the remaining POMs are merely the eigenmodes of the linear system. This perspective should be useful in the context of modal analysis because the POD requires only the measurement of the structural responses. In addition, it is much easier to implement compared to experimental modal analysis.

Forced harmonic response

The forced response is defined as the stationary part of the response synchronous to the excitation:

$$\mathbf{x}(t) = \mathbf{x}_f \sin \omega_e t \tag{3.49}$$

The forced response amplitude is given by the following expression [58]:

$$\mathbf{x}_f = \left\{ \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)} \mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i} \right\} \mathbf{f} \tag{3.50}$$

and,

$$\mathbf{x}(t) = \left\{ \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)} \mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i} \right\} \mathbf{f} \sin \omega_e t \tag{3.51}$$

The displacement matrix $\mathbf{X}$ becomes:

$$\begin{aligned} \mathbf{X} &= \left[\begin{array}{ccc} \mathbf{x}(t_1) & \cdots & \mathbf{x}(t_m) \end{array} \right] \\ &= \left[\begin{array}{ccc} \left\{ \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)} \mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i} \right\} \mathbf{f} \sin \omega_e t_1 & \cdots & \left\{ \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)} \mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i} \right\} \mathbf{f} \sin \omega_e t_m \end{array} \right] \end{aligned} \tag{3.52}$$

Equation (3.52) may be expressed in the form:

$$\begin{aligned} \mathbf{X} &= \left[\left\{ \sum_{i=1}^{n} \frac{\mathbf{p}_{(i)} \mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i} \right\} \mathbf{f} \right] \left[\begin{array}{c} \sin \omega_e t_1 \\ \cdots \\ \sin \omega_e t_m \end{array} \right]^T \\ &= \mathbf{x}_f \mathbf{a}^T \\ &= \left[\begin{array}{cc} \mathbf{x}_f & \mathbf{B}_1 \end{array} \right] \left[\begin{array}{ccccccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & & & \vdots & & \vdots \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array} \right] \left[\begin{array}{cc} \mathbf{a} & \mathbf{B}_2 \end{array} \right]^T \end{aligned}$$

$$\begin{aligned}\mathbf{X} &= \left[\begin{array}{cc} \frac{\mathbf{x}_f}{\|\mathbf{x}_f\|} & \mathbf{B}_1 \end{array}\right] \left[\begin{array}{ccccccc} \|\mathbf{x}_f\|\,\|\mathbf{a}\| & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & & & \vdots & & \vdots \\ \vdots & & \ddots & & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array}\right] \left[\begin{array}{cc} \frac{\mathbf{a}}{\|\mathbf{a}\|} & \mathbf{B}_2 \end{array}\right]^T \\ &= \underbrace{\left[\begin{array}{cc} \frac{\mathbf{x}_f}{\|\mathbf{x}_f\|} & \mathbf{B}_1 \end{array}\right]}_{\mathbf{U}} \underbrace{\left[\begin{array}{c} \mathbf{I_1} \end{array}\right]}_{\mathbf{\Sigma}} \underbrace{\left[\begin{array}{cc} \frac{\mathbf{a}}{\|\mathbf{a}\|} & \mathbf{B}_2 \end{array}\right]^T}_{\mathbf{V^T}} \end{aligned} \tag{3.53}$$

where $\mathbf{B}_1$ is an $(n \times (n-1))$ matrix;

$\mathbf{I_1}$ is an $(n \times m)$ matrix containing only one non-zero element $\|\mathbf{x}_f\|\,\|\mathbf{a}\|$;

$\mathbf{B_2}$ is an $(m \times (m-1))$ matrix.

Matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ do not influence equation (3.53) since they are both multiplied by zero elements. If both matrices are chosen in order that $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, then equation (3.53) is the SVD of the matrix $\mathbf{X}$.

In conclusion, the following points may be noted:

1. Since there is only one non-zero singular value, the forced harmonic response of a linear system is captured by a single POM whatever the number of DOF is. Nevertheless, all the eigenmodes are necessary to reconstruct the response. This property is independent of the mass distribution and underlines the optimality of the POMs.

2. An analytical expression of the POM is obtained:

$$POM = \frac{\mathbf{x}_f}{\|\mathbf{x}_f\|} = \frac{\left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f}}{\left\|\left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f}\right\|} \tag{3.54}$$

   If the structural matrices and the spatial discretisation of the excitation are assumed to be known, the POM may be calculated without first simulating the system response as required in the definition of the POMs.

3. The expression of the POM given by equation (3.54) is equal, except for the norm, to the last eigenmode of the transformed system for the global response to harmonic excitation (3.47). This last eigenmode is thus related to the forced harmonic response.

4. The convergence of the dominant POM to an eigenmode is no longer guaranteed. The POM appears now to be a combination of all the eigenmodes. However, if the excitation frequency ω_e is close to a natural frequency of the system (e.g., ω_j), then

the denominator $\omega_j^2 - \omega_e^2$ of the jth term of combination (3.54) tends to zero. It is thus observed that this term has a much larger amplitude than the others:

$$POM = \frac{\left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f}}{\left\|\left\{\sum_{i=1}^{n} \frac{\mathbf{p}_{(i)}\mathbf{p}_{(i)}^T}{(\omega_i^2 - \omega_e^2)\mu_i}\right\}\mathbf{f}\right\|} \simeq \frac{\frac{\mathbf{p}_{(j)}\mathbf{p}_{(j)}^T}{(\omega_j^2 - \omega_e^2)\mu_j}\mathbf{f}}{\left\|\frac{\mathbf{p}_{(j)}\mathbf{p}_{(j)}^T}{(\omega_j^2 - \omega_e^2)\mu_j}\mathbf{f}\right\|} = \alpha\,\mathbf{p}_{(j)} \quad \text{if } \omega_e \to \omega_j \tag{3.55}$$

Since $\mathbf{p}_{(j)}^T\mathbf{f}$ represents a scalar product, the POM has the same direction as the eigenmode $\mathbf{p}_{(j)}$ which means that the POM is equal to the resonating mode shape. This is consistent with what was obtained in reference [47] using the eigensolution perspective. It is worth pointing out that the non-resonating mode shapes should not be revealed by the POD.

3.4.4 Stationary random response of a linear system to a white noise excitation

The equation of motion of a linear system excited with a white noise sequence is:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}(t) \tag{3.56}$$

where $\mathbf{f}(t)$ is a white noise process such that:

$$E[\mathbf{f}(t)] = 0 \text{ and } E[\mathbf{f}(t)\mathbf{f}(\tau)^T] = \nu\delta(t-\tau) \tag{3.57}$$

Equation (3.56) is recast in the state variable form:

$$\dot{\mathbf{r}} = \mathbf{D}_1\mathbf{r} + \mathbf{D}_2\mathbf{f} \tag{3.58}$$

$$\mathbf{x} = \mathbf{D}_3\mathbf{r} \tag{3.59}$$

where $\mathbf{D}_1 = \begin{bmatrix} \mathbf{Z} & \mathbf{I} \\ \mathbf{M}^{-1}\mathbf{K} & \mathbf{M}^{-1}\mathbf{C} \end{bmatrix}$ is the system matrix;

$\mathbf{D}_2$ is the input matrix;

$\mathbf{D}_3$ is the output matrix.

It is assumed that the system is stable and time invariant, and that all processes are Gaussian. In this context, it can be shown [25] that the covariance matrix of the steady state response $\mathbf{\Sigma}_r = E[\mathbf{r}(t)\mathbf{r}(t)^T]$ satisfies the Lyapunov equation:

$$\mathbf{D}_1\mathbf{\Sigma}_r + \mathbf{\Sigma}_r\mathbf{D}_1^T + \nu\mathbf{D}_2\mathbf{D}_2^T = 0 \tag{3.60}$$

It is worth pointing out that $\mathbf{\Sigma}_r$ also corresponds, with the exception of constant ν, to the definition of the controllability Grammian $\mathbf{W}_c$ of the system.

If only the displacements are considered, then the covariance matrix of the system response becomes:

$$\mathbf{\Sigma}_x = E[\mathbf{x}(t)\mathbf{x}(t)^T] = \mathbf{D}_3\mathbf{\Sigma}_r\mathbf{D}_3^T \tag{3.61}$$

The POMs may be evaluated without first simulating the system. Indeed, if the structural matrices are assumed to be known, the Lyapunov equation (3.60) may be solved in order to compute the covariance matrix $\mathbf{\Sigma}_r$, and equation (3.61) allows to calculate $\mathbf{\Sigma}_x$. By definition, the POMs are the eigenvectors of $\mathbf{\Sigma}_x$. The analytical relationship between the POMs and the eigenmodes is now obscured. If all states (displacement and velocity) are measured, the POMs are merely the eigenvectors of the controllability Grammian $\mathbf{W}_c$.

The POMs may also be geometrically interpreted. In the case of a Gaussian white noise excitation, the response of a linear system is also Gaussian and is characterised at each DOF by a probability density function (PDF) equal to:

$$\text{PDF}(x) = \frac{1}{\sqrt{2\pi}\sigma}\ \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \tag{3.62}$$

where $\mu = E[x]$ is the mean and $\sigma = E[(x-\mu)^2]$ is the standard deviation. The joint PDF reads:

$$\text{PDF}(x_1, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma_1 \cdots \sigma_n}\ \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right] \tag{3.63}$$

It can be demonstrated [133] that the contours of the joint PDF consists of n-dimensional ellipsoids and that the POMs are the principal axes of these ellipsoids.

3.4.5 Linear normal modes, non-linear normal modes and proper orthogonal modes: a geometric approach

For the sake of clarity, the eigenmodes of a linear system are called here linear normal modes (LNMs). The determination of LNMs is reduced to the equivalent problem of computing the eigensolutions of linear transformations. Obviously, such an approach is inadmissible for non-linear systems. The concept of synchronous NNM for discrete conservative oscillators was introduced for non-linear systems by Rosenberg [150]: "*A non-linear system vibrates in normal modes when all masses execute periodic motions of the same period, when all of them pass through the equilibrium position at the same instant, and when, at any time t, the position of all the masses is uniquely defined by the position of any one of them.*"

The objective of this section is to examine the geometric interpretation of LNMs, NNMs and POMs.

Linear systems

First, let us consider a linear system consisting of masses and springs (Figure 3.1). If the displacement of the ith mass from its equilibrium position is noted x_i, then the equations

of motion of the system are:

$$m_i\ddot{x}_i = k_i(x_{i-1} - x_i) - k_{i+1}(x_i - x_{i+1}) \quad \text{where} \quad i = 1, 2, \cdots, n \qquad x_0 = x_{n+1} \equiv 0 \tag{3.64}$$

If the co-ordinates are normalised using the transformations $\xi_i = m_i^{1/2} x_i, \quad i = 1, \cdots, n$, equation (3.64) becomes:

$$\ddot{\xi}_i = \frac{k_i}{m_i^{1/2}}\left(\frac{\xi_{i-1}}{m_{i-1}^{1/2}} - \frac{\xi_i}{m_i^{1/2}}\right) - \frac{k_{i+1}}{m_i^{1/2}}\left(\frac{\xi_{i+1}}{m_{i+1}^{1/2}} - \frac{\xi_i}{m_i^{1/2}}\right) \quad \text{where} \quad m_0 = m_{n+1} \equiv \infty \tag{3.65}$$

The transformed equations of motion (3.65) may be regarded as those of a unit mass that moves in a n-dimensional space. The right-hand side of equation (3.65) derives from a potential function:

$$\ddot{\xi}_i = \frac{\partial U}{\partial \xi_i}, \quad \text{with} \quad U = -\sum_{i=1}^{n+1} \frac{k_i}{2}\left(\frac{\xi_{i-1}}{m_{i-1}^{1/2}} - \frac{\xi_i}{m_i^{1/2}}\right)^2 \tag{3.66}$$

If no external force is present and if the motion is due to an initial displacement, the system occupies at time $t = 0$ a position of maximum potential $U = -U_0$. This latter equation defines an ellipsoid that is symmetric with respect to the origin. This ellipsoid is called the bounding ellipsoid because all solutions must lie in this domain.

In its definition of a normal mode for a linear system, Rosenberg [150] stated that it is a straight line in the $(\xi_1, \cdots, \xi_n)$ space that passes through the origin of that space and that intersects the bounding ellipsoid orthogonally. It follows from the definition that the LNMs are the principal axes of the bounding ellipsoid in the $(\xi_1, \cdots, \xi_n)$ space.

This result can also be obtained using the geometric interpretation of the eigenvalue problem (see section 3.3.1). Indeed, the LNMs are the eigenvectors of the matrix $\mathbf{M}^{-1}\mathbf{K}$. In order that the LNMs be the principal axes of the ellipsoid $\mathbf{x}^T\mathbf{M}^{-1}\mathbf{K}\mathbf{x} = 1$, the matrix $\mathbf{M}^{-1}\mathbf{K}$ must be real, positive definite and symmetric. This is the case if the mass matrix is proportional to identity, i.e., $\mathbf{M} = \alpha\mathbf{I}$. Accordingly, the LNMs are the principal axes of the ellipsoid $\frac{1}{\alpha}\mathbf{x}^T\mathbf{K}\mathbf{x} = 1$. This expression is, with the exception of a constant, the expression of the potential energy in the $(x_1, \cdots, x_n)$ space. Since the mass matrix is assumed to be proportional to identity, this is also the expression, except for a constant, of the potential energy in the $(\xi_1, \cdots, \xi_n)$ space. This is an alternative way of demonstrating that the LNMs are the principal axes of the bounding ellipsoid in the $(\xi_1, \cdots, \xi_n)$ space.

If the mass matrix is proportional to identity, the LNMs are also the principal axes of the bounding ellipsoid in the $(q_1, \cdots, q_n)$ space whose expression is:

$$U_0 = \sum_{i=1}^{n+1} \frac{k_i}{2}\left(x_{i-1} - x_i\right)^2 = \frac{1}{2}\mathbf{x}^T\mathbf{K}\mathbf{x} \tag{3.67}$$

As far as the POMs are concerned, they are the principal axes of the ellipsoid $c = \mathbf{x}^T\mathbf{\Sigma}\mathbf{x}$ where $\mathbf{\Sigma}$ is the covariance matrix (cf. section 3.3.1). Since for an unforced system with

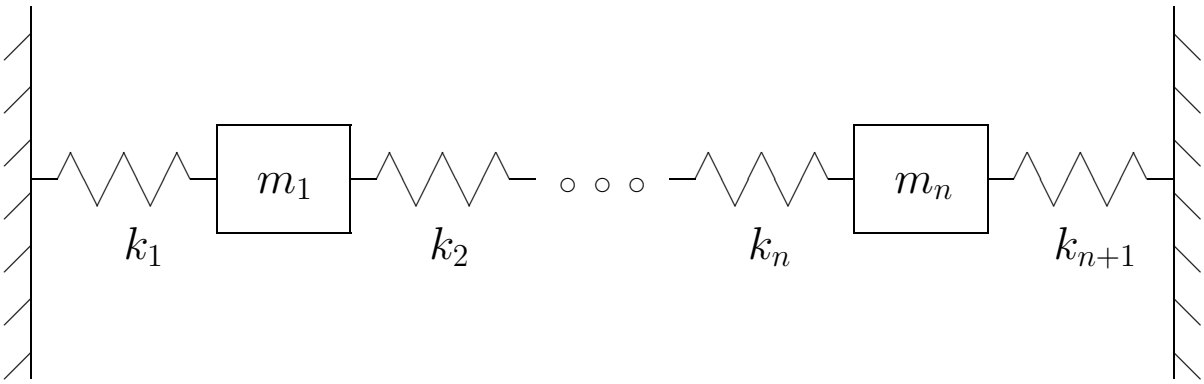


Figure 3.1: Linear system consisting of masses and springs.

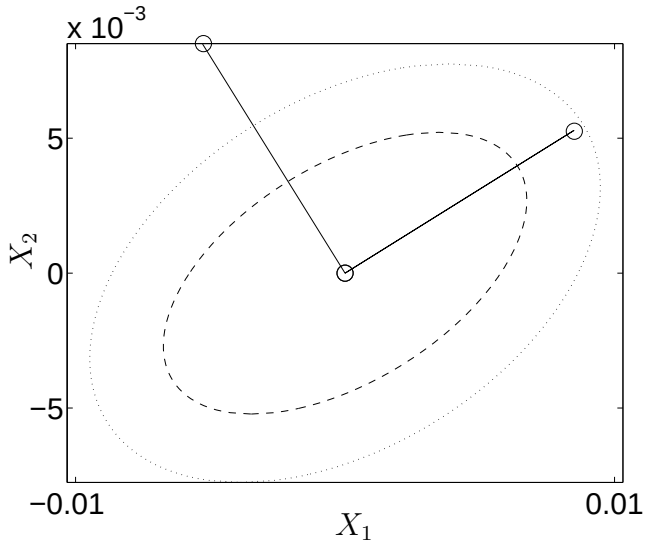


Figure 3.2: LNMs and POMs as principal axes of similar and similarly placed ellipsoids: - - -, U_0; · · · · ·, c; ——, POMs; o o, LNMs.

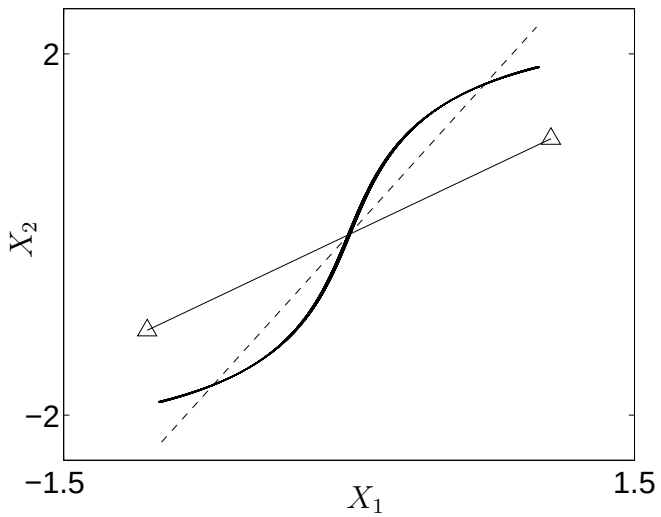


Figure 3.3: Comparison of LNM, POM and NNM on a two DOF example: ——, NNM; - - - -, POM; △——△, LNM.

a mass matrix proportional to identity, the POMs and the LNMs coincide (see section 3.4.1), it can be concluded that $U_0 = \frac{1}{2}\mathbf{x}^T\mathbf{K}\mathbf{x}$ and $\mathbf{x}^T\mathbf{\Sigma}\mathbf{x} = c$ are similar and similarly placed ellipsoids. This is illustrated in Figure 3.2 for a two DOF system with an initial displacement.

Non-linear systems

If an LNM is a straight line in the co-ordinate space, a NNM is represented by a line, straight (similar NNM) or curved (non-similar NNM). But generally, NNMs are non-similar and the POMs, characterised by straight lines in the co-ordinate space, cannot be merged with NNMs. However, due to their optimality property [cf. equation(3.2)], if the motion is a single, synchronous NNM, the resonant POM minimises the square of the distance with the NNM under the constraint that it passes through the origin of the co-ordinate system and as stated in reference [47], the POM can be considered as the best linear representation of the NNM.

This is illustrated using a two DOF example in which the first mass is connected to a wall through a cubic spring. This system was analysed in details by Shaw and Pierre [156]. For particular initial conditions, the system undergoes a non-linear modal motion in the first mode. Figure 3.3 compares the NNM, the dominant POM and the first LNM. It underlines that the POM is the best linear approximation of a single synchronous NNM and that the LNM is no longer efficient to capture the dynamics of a non-linear system.

3.5 Model reduction of non-linear systems

3.5.1 Motivation

In many domains of applied sciences and in structural dynamics particularly, dealing with large-scale dynamical structures is a central issue. In the presence of non-linearities, seeking for the solution by use of mathematical modelling and simulation (e.g., finite element method) may be computationally intensive. Accordingly, due to the complexity of such a numerical approach, it is worth reducing the dimensionality of the system while retaining its intrinsic properties.

Conventionally, one of the methods used to analyse non-linear systems is through projection of the equations of motion onto the eigenmodes of the linearised system. This method often requires the concurrent simulation of many modes to achieve a satisfactory accuracy, and the resulting model may be computationally expensive.

The present study is thus motivated by the fact that, in the field of non-linear systems, new features must be defined to better represent the system dynamics. In reference [8], lower-dimensional models of non-linear vibrating systems are created using the POMs in order to prove their efficiency and their superiority over the mode shapes. The approach presented herein conjectures that a reduced-order model built using the POMs of a chaotic

orbit allows for much more efficient analysis of the original system than any other set of POMs extracted from a non-chaotic response. Indeed, due to the "random nature" of chaos, the POMs of a chaotic orbit are expected to better capture the system dynamics.

This study results from a collaboration with Professor Brian Feeny.

3.5.2 Reduced-order models

Even though the original model of a continuous structure typically has an infinite number of DOF to start with, it is assumed that the equations of motion have already been discretised in space by using a finite element model. For a non-linear mechanical structure, the spatial discretisation usually yields the following equations:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \mathbf{f}_{NL}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \mathbf{f}(t) \tag{3.68}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, damping and stiffness matrices respectively;

$\mathbf{x}$ is the vector of displacement co-ordinates;

$\mathbf{f}_{NL}$ is the non-linear force vector;

$\mathbf{f}$ is the external applied force vector.

However, this discretised model may lead to a large number n of DOF to sufficiently approximate the true physical phenomenon. In order to reduce the complexity, model reduction techniques are to be considered. They may be viewed as a projection of the n-dimensional vector of displacement co-ordinates $\mathbf{x}(t)$ onto a vector $\mathbf{a}(t)$ of lower dimension r. The co-ordinate transformation considered in this study is defined by an $(n \times r)$ projection matrix $\mathbf{P}$ whose columns are the POMs corresponding to the r greatest POVs:

$$\mathbf{x}(t) = \mathbf{P}\mathbf{a}(t) + \boldsymbol{\mu} \tag{3.69}$$

where $\boldsymbol{\mu}$ accounts for the mean of vector $\mathbf{x}$. Since the POMs cannot be calculated from the structural matrices, it is emphasised that it is necessary either to first simulate the system or to measure experimentally the system response in order to estimate matrix $\mathbf{P}$.

The equations of motion expressed in the lower-dimensional setting are:

$$\ddot{\mathbf{a}}(t) + \mathbf{C}^*\dot{\mathbf{a}}(t) + \mathbf{K}^*\mathbf{a}(t) + (\mathbf{MP})^\dagger\mathbf{K}\boldsymbol{\mu} + (\mathbf{MP})^\dagger\mathbf{f}_{NL} = (\mathbf{MP})^\dagger\mathbf{f}(t) \tag{3.70}$$

where $\mathbf{C}^* = (\mathbf{MP})^\dagger\mathbf{CP}$, $\mathbf{K}^* = (\mathbf{MP})^\dagger\mathbf{KP}$ and symbol † denotes the pseudo-inverse of the matrix. The solution in the original coordinate system can be retrieved using equation (3.69). This of course does not yield the correct solution but an approximation whose quality depends on the severity of the reduction.

It remains to determine the order of reduction, i.e., the number of POMs to be included in matrix $\mathbf{P}$. The number r of POMs is chosen based on a prescription, in this case such that:

$$\frac{\sum_{i=1}^{r}\lambda_i}{\sum_{i=1}^{n}\lambda_i} \geq 0.9999 \tag{3.71}$$

where λ_i is the ith POV. In other words, assuming that the POVs are sorted in descending order, the first r POMs that account for 99.99 % of the total energy are considered.

3.5.3 Numerical application

Description of the structure

The application under investigation is a clamped beam made of steel under the action of a sinusoidal force as shown in Figure 3.4. Magnets are placed to the left and right of the free end to buckle the beam. The magnetic field creates forces that are assumed to be concentrated at the free end of the beam. In the literature, numerous studies were realised on the single-degree-of-freedom (SDOF) oscillator (e.g., references [77, 131]) but only a few studies tackled the multi-degree-of-freedom (MDOF) problem [171].

In the absence of excitation, and depending on the initial conditions, the beam settles on one of the two stable equilibria. In addition, the beam has an unstable central equilibrium position. A possible model for the magnetic potential energy is given by a two-well potential (see Figure 3.5):

$$V(x) = K_1 \frac{x^2}{2} + K_2 \frac{x^4}{4} \qquad \text{with } K_1 < 0 \tag{3.72}$$

where K_1 and K_2 are the magnet stiffness terms. Their values are set to be equal to those in reference [171], i.e., $K_1 = -11.29\ N/m$ and $2.74\,10^4\ N/m^3$.

The beam is discretised with seven beam elements and is simulated with a forcing frequency equal to 5 Hz and for amplitudes F_0 going up to 3.5 N. To illustrate the varied dynamics of the system, a bifurcation diagram is presented in Figure 3.6. This plot represents the horizontal displacement at the end of the beam (at the moment when the sine force is equal to F_0) as a function of the amplitude F_0. For a particular amplitude, the number of points is related to the period. The initial conditions at each amplitude are equal to zero and the bifurcation diagram is obtained by checking for any periodicity within the first 800 s. In the absence of periodicity, the transients are not yet died out or the dynamics is non-periodic. It is worth pointing out that the beam can oscillate equally around either a magnet or the other depending on the initial conditions. This is the reason why points in the diagram may be associated with a positive or negative displacement. For some amplitudes, particularly between 1 and 2.5 N, the number of points is large which may be interpreted as a non-periodic oscillation within the time allowed for the simulation.

Figure 3.7 shows phase portraits of the translational DOF at the end of the beam for various amplitudes. For amplitudes equal to 0.1 N and 3.5 N, the orbit is periodic while for the two selected intermediate amplitudes, i.e., 0.8 N and 2 N, the orbit appears to be chaotic.

Reduced-order model based on the POMs of a periodic orbit

A reduced-order model based on the POMs of a periodic orbit, i.e., for $F_0 = 0.01N$, is first considered. In this case, the first three POMs capture more than 99.99 % of the total energy and are exploited to build the lower-dimensional model. The POMs are illustrated

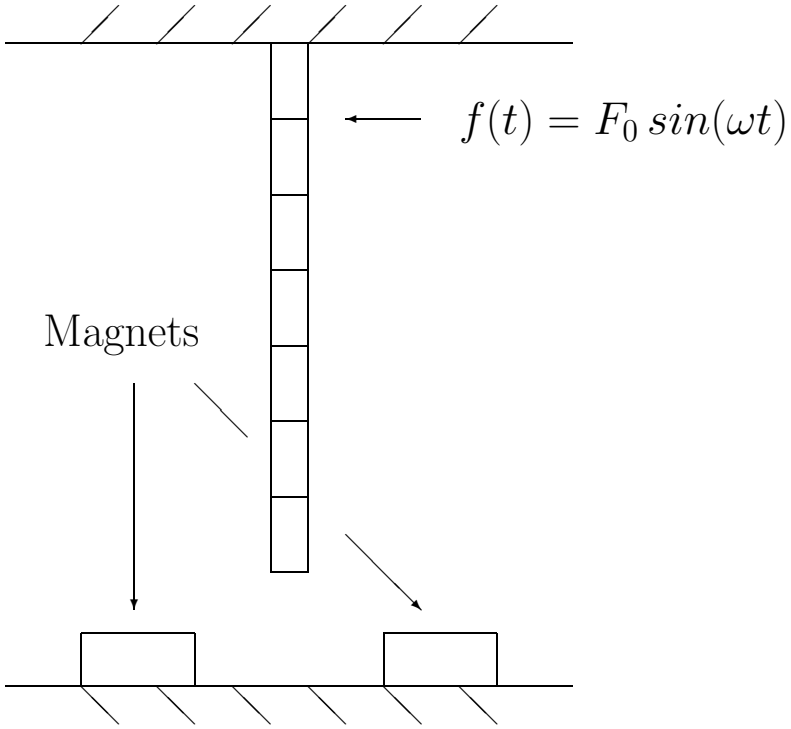


Figure 3.4: Clamped beam with two permanent magnets.

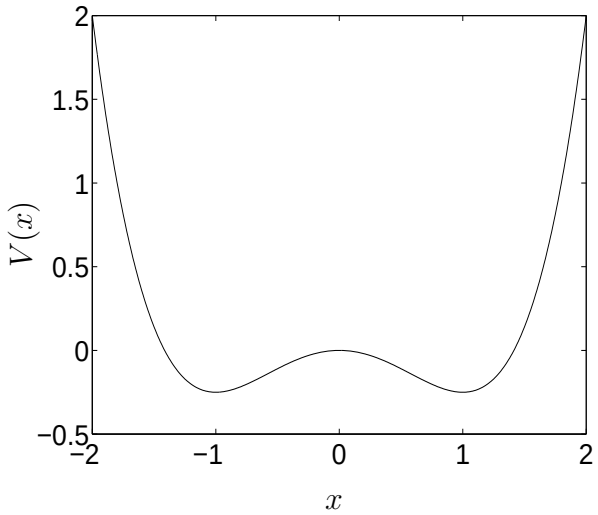


Figure 3.5: Magnetic potential energy.

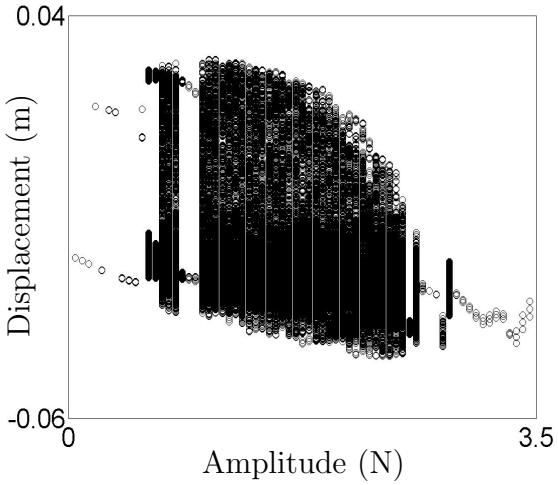


Figure 3.6: Bifurcation diagram of the full model.

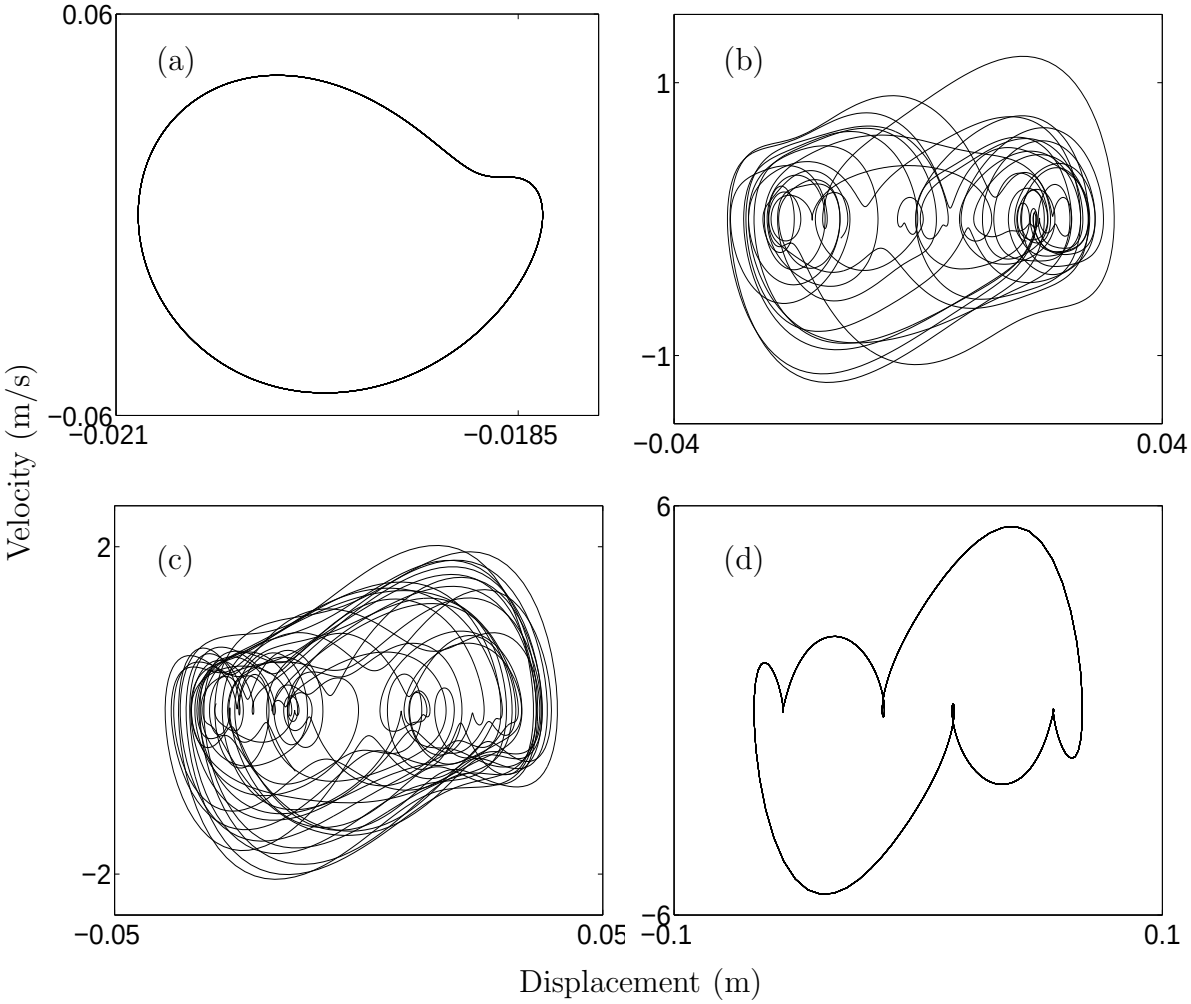


Figure 3.7: Phase portraits computed from the full model. (a) 0.1 N; (b) 0.8 N;(c) 2 N; (d) 3.5 N.

	1st POV	2nd POV	3rd POV	4th POV	5th POV
POVs (0.01 N)	0.9980	0.0014	0.0005	$7\,10^{-7}$	$3\,10^{-11}$
POVs (1 N)	0.6970	0.2702	0.0295	0.0031	0.0002

Table 3.1: Normalised POVs for 0.01 N and 1 N levels.

	Full model	Five-mode model	Three-mode model
0.8 N	2.85	2.76	10^{-13}
2 N	3.03	2.95	0.71

Table 3.2: Greatest Lyapunov exponent for the full and reduced-order models.

	Full model	Five-mode model	Three-mode model
0.8 N	2.65	2.52	—
2 N	2.85	2.73	1.93

Table 3.3: Correlation dimension for the full and reduced-order models.

in Figure 3.8, and the corresponding POVs are listed in Table 3.1. All the DOF are used in the POD but only the translational DOF are illustrated in this figure.

Figure 3.9 displays the phase portraits computed from the three-mode model and is to be compared with the phase portraits of the full model, i.e., the model discretised by the finite element modelling, represented in Figure 3.7. However, the brute analysis of time series is not a suitable means of characterising chaotic signals. In this context, indicators such as the greatest Lyapunov exponent and the correlation dimension that are defined in appendix A are helpful and are given in Tables 3.2 and 3.3 for the full and three-mode models.

Apart from the amplitude $F_0 = 0.1\ N$, the orbits are not well reconstructed by the three-mode model. In particular, it can be observed that:

- The orbit predicted for $F_0 = 0.8\ N$ is periodic while the actual orbit is chaotic;
- The greatest Lyapunov exponent and the correlation dimension for $F_0 = 2\ N$ differ widely from those given by the full model.

Another possible means of assessing the results is to compare the occurrence of bifurcations. Indeed, bifurcations tend to be sensitive to modelling errors. To this end, the bifurcation diagram of the three-mode model is presented in Figure 3.10(a), and the comparison with Figure 3.6 confirms that the predictive accuracy of this model is rather poor.

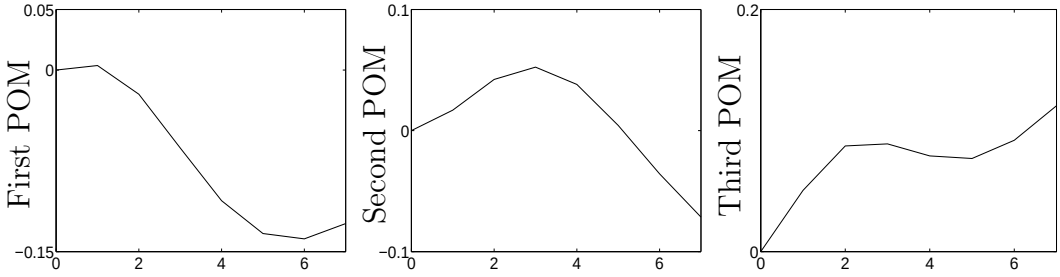


Figure 3.8: POMs of 0.01 N level.

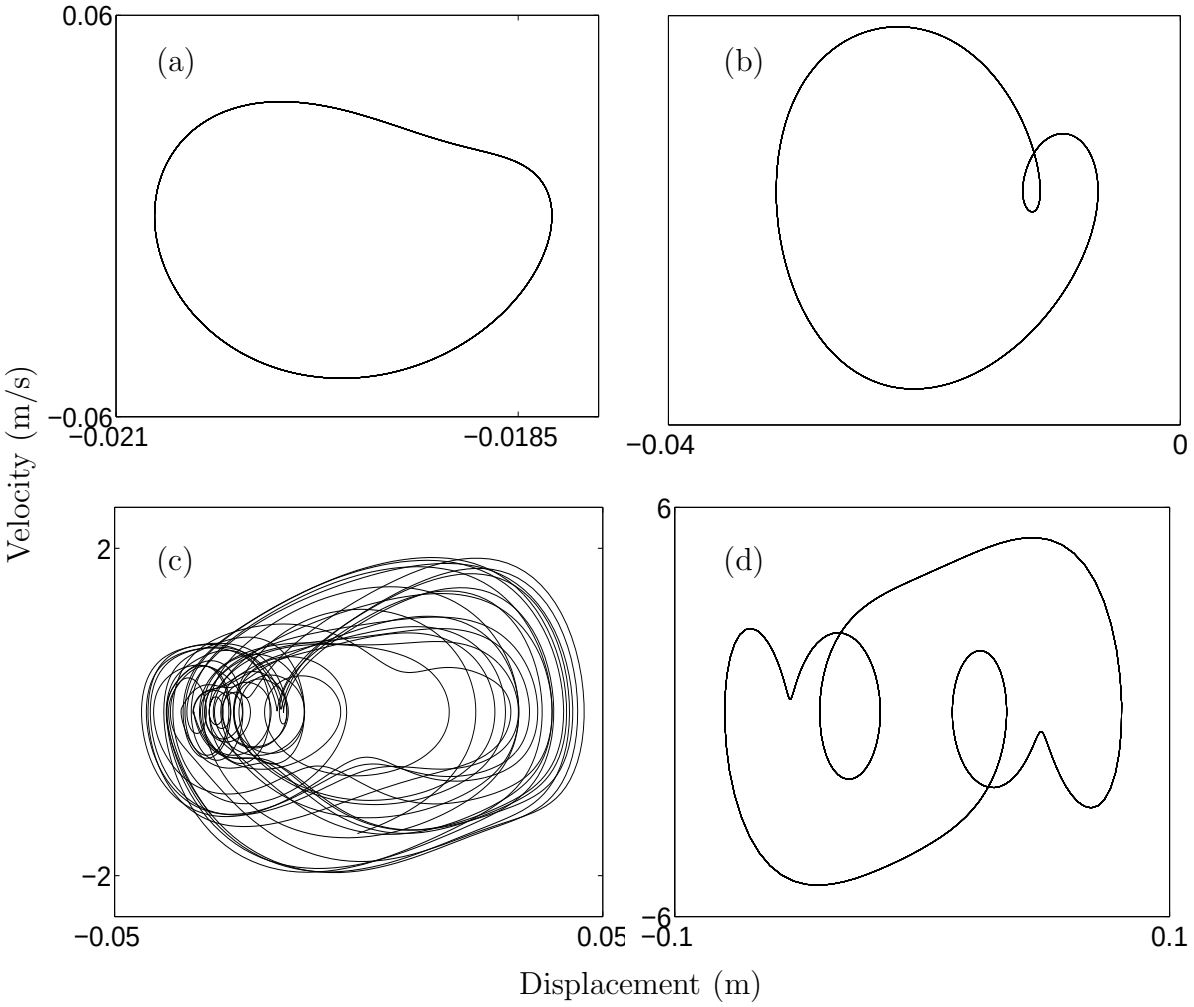


Figure 3.9: Phase portraits computed from the three-mode model. (a) 0.1 N; (b) 0.8 N;(c) 2 N; (d) 3.5 N.

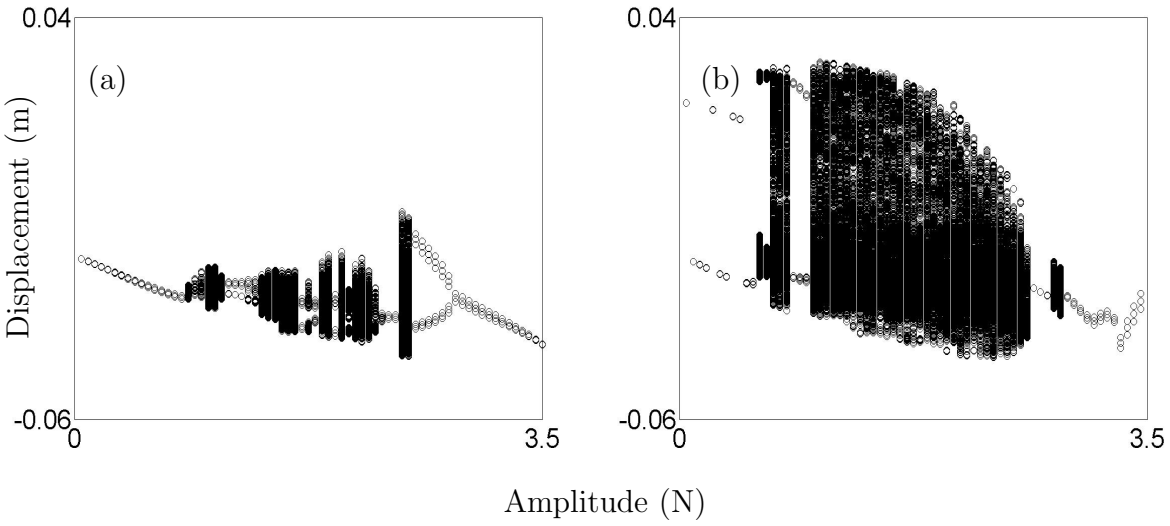


Figure 3.10: Bifurcation diagram of the reduced-order models. (a) Three-mode model; (b) five-mode model.

It should be noted that the inclusion of more POMs in the model was also considered but the results were not significantly improved.

Reduced-order model based on the POMs of a chaotic orbit

The purpose of this section is to show that the POMs extracted from a chaotic orbit represent a convenient choice for the construction of a low-dimensional model of the beam. However, the system is chaotic for a wide range of amplitudes, and it remains to determine the value of the amplitude for which the modes are extracted. Interestingly enough, it was observed that the POMs do not vary significantly as long as the response is chaotic. The first five modes corresponding to an amplitude F_0=1 N constitute 99.99 % of the total signal energy, and are thus chosen. These modes are displayed in Figure 3.11, and the corresponding POVs are listed in Table 3.1.

Figure 3.12 depicts the phase portraits computed from the five-mode model and is to be compared with Figure 3.7. A convincing correspondence between both figures is obtained but the chaotic orbits, namely the orbits for 0.8 N and 2 N levels, need further verification. The greatest Lyapunov exponent and the correlation dimension are listed in Tables 3.2 and 3.3. It can clearly be seen that the indicators of the five-mode model are in good concordance with those computed from the full model. The bifurcation diagram of the reduced-order model is presented in Figure 3.10(b) and is in satisfactory agreement

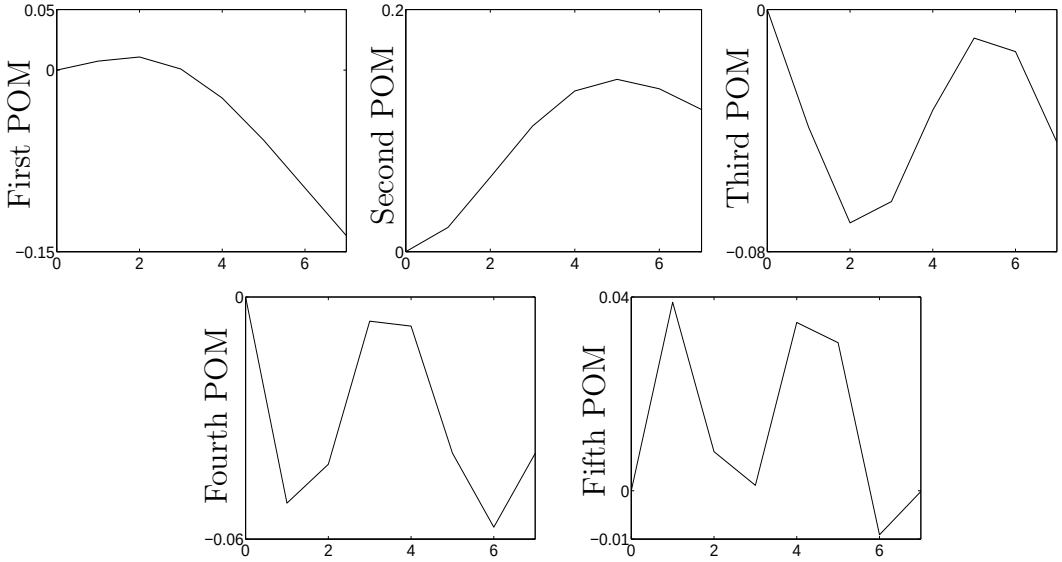


Figure 3.11: POMs of 1 N level.

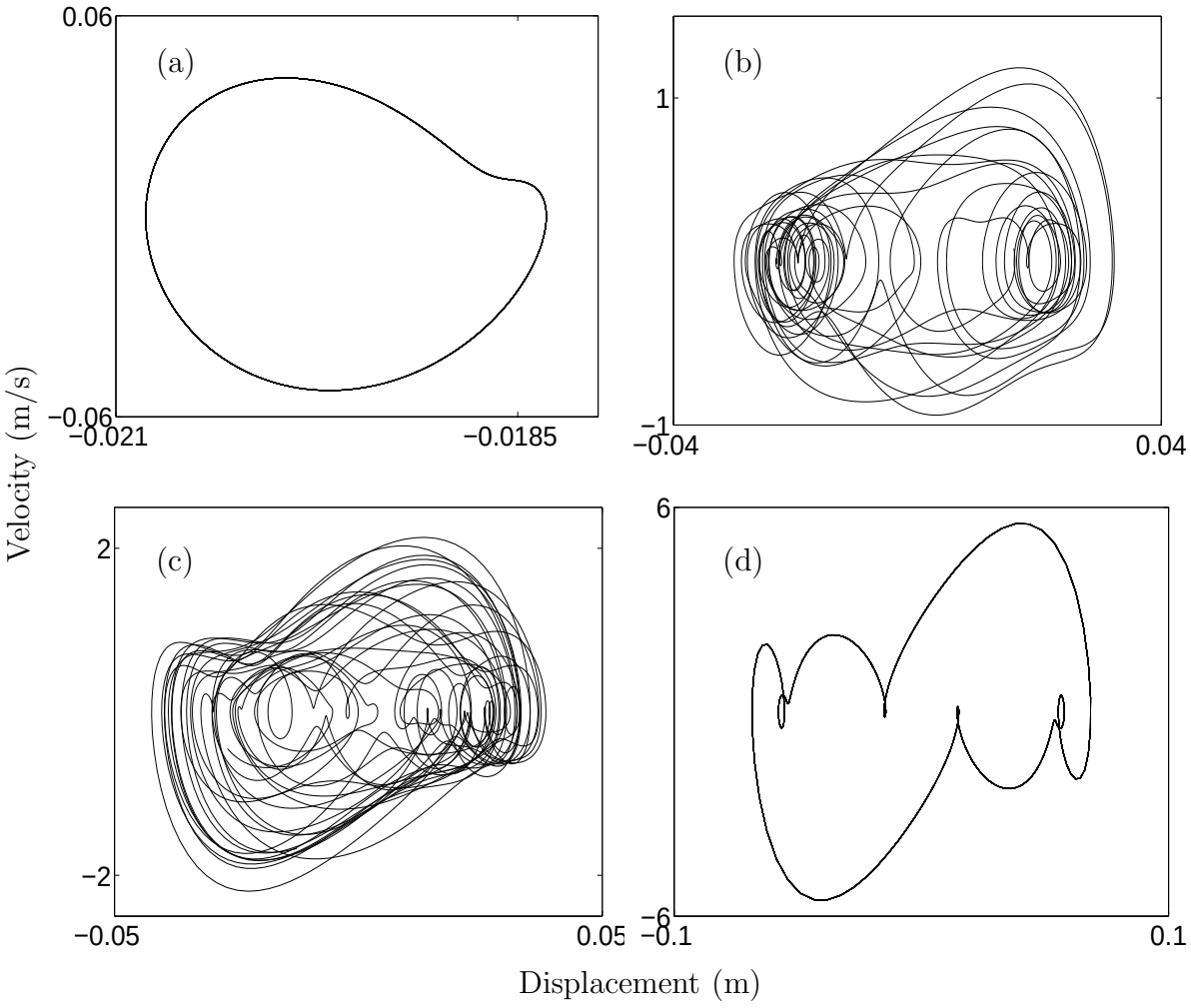


Figure 3.12: Phase portraits computed from the five-mode model. (a) 0.1 N; (b) 0.8 N;(c) 2 N; (d) 3.5 N.

with the diagram obtained from the full model.

In conclusion, the previous results illustrate that the POMs of a chaotic orbit are much more representative of the dynamics of a non-linear system than any other set of modes extracted from a non-chaotic response. The five-mode model allows for a reduction from an order 28 (14 DOF $\times$ 2) to an order 10 (5 DOF $\times$ 2), and requires less than half floating operations in comparison with the full model.

3.6 Model updating of non-linear systems

Referring to the discussion of non-linear model updating in the introductory chapter, we see that only a few studies are available in this field of research.

Among the most promising methods that have been developed, the approach based on the POD must be cited. Indeed, as shown in the previous sections, the POD enables to obtain useful spatial information through the POMs (matrix $\mathbf{U}$) and time information through their time modulations (matrix $\mathbf{V}$). In addition, the optimality of the POD for the reconstruction of the system dynamics has been highlighted. This explains why the POMs have been exploited as a promising alternative to the mode shapes in a model updating context [66, 71, 109, 111].

In reference [109] and in a companion thesis [111], the proposed model updating procedure relies on the solution of an optimisation problem that consists in minimising the difference between the POMs of the measured and predicted data. It should be noted that the full decomposition is not included in the objective function. Only the POMs corresponding to the highest singular values are taken into account since the smallest singular values are known to be associated with the noise present in the data. Unfortunately, it is not possible to give definitive guidance on the number of terms to include in the objective function. Some rules have been proposed (e.g., the sum of the r highest singular values must exceed 99 % of the sum of all singular values) but their efficiency is largely application dependent.

In order to retain the information contained in matrix $\mathbf{V}$ but without the drawback of its oscillatory nature, the computation of the Wavelet transform or the windowed discrete-time Fourier transform of matrix $\mathbf{V}$ has also been proposed. This gives a time-frequency representation of the energy contained in the signal. The instantaneous frequencies and envelopes corresponding to the maxima in the time-frequency plane are extracted and included into the objective function instead of the right singular vectors. This enables to smooth the objective function as it transforms the oscillatory nature of the time information into a more suitable form. The usefulness of the Wavelet transform for model updating purposes is discussed at length in reference [111].

Experimental validation

In a first step, the procedure was validated on a numerical example consisting of a beam with a cubic stiffness [109]. The results were promising and the investigation of the experimental structure considered in chapter 2 was then performed. Recall that this structure consists of a beam with a geometric non-linearity on its extremity. The experimental set-up is described in section 2.4.1.

The main beam is modelled with seven beam finite elements and the thin beam with four beam finite elements. An additional element is used to model the junction between the two beams. This element connects the rotational and translational DOF of the beams by rotational and translational stiffnesses. The non-linearity is modelled using a grounded spring at the junction and has the same functional form as the one used in the conditioned reverse path (CRP) method, i.e., $f_{NL}(x, \dot{x}) = A\,|x|^{\alpha}\,\mathrm{sign}(x)$. The finite element model is displayed in Figure 3.13.

Several parameters need to be identified, namely, the rotational and translational stiffnesses at the clamped ends and at the junction, the coefficient and exponent of the non-linearity, and the damping ratio of the first four modes of the underlying linear system. A first test is performed at a low excitation level for which the structure is assumed to be linear. During this test, the rotational and translational stiffnesses at the clamped ends and at the junction are updated which reduces the number of parameters to estimate during the non-linear analysis.

A second test is performed for which the behaviour of the structure is now non-linear. The excitation level is similar to the highest level considered in the CRP method. After updating, the non-linearity exponent is $\alpha = 2.8$, and the corresponding coefficient is equal to $A = 1.65\,10^9\,N/m^{2.8}$.

The main results are illustrated in Figures 3.14, 3.15 and 3.16. Figure 3.14 shows the comparison between the experimental POMs, the POMs of the linear model of the structure (first test), and the POMs of the non-linear model (second test). It can be observed that the POMs of the non-linear model provide a close match to the experimental POMs. The POMs of the linear model illustrates that the non-linearity mainly acts on the first two modes.

Figures 3.15 and 3.16 represent the instantaneous frequency and envelope of the first and second POMs time decomposition respectively. It can be seen that the instantaneous frequencies of the non-linear model and of the experimental measurements are in good agreement. They both decrease with time to reach the natural frequency of the underlying linear system. This is expected because the amplitude of the vibration decreases with time, and the stiffening effect of the non-linearity becomes less important. A detailed discussion of these results and of the use of the POD for model updating purposes is available in references [110, 111].

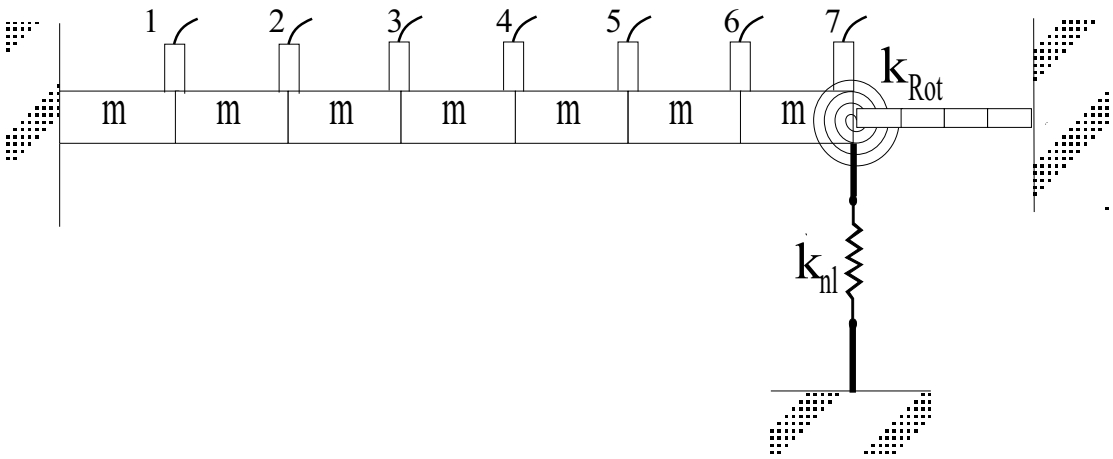


Figure 3.13: Finite element model of the structure.

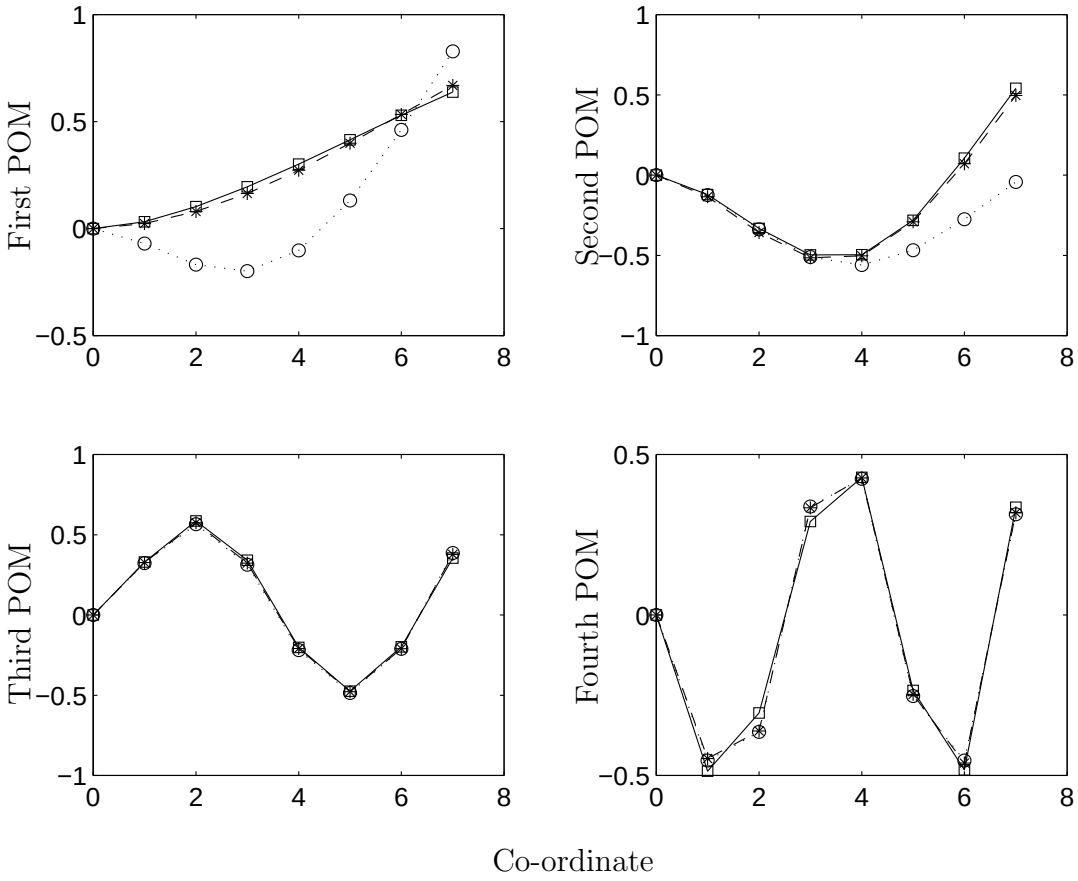


Figure 3.14: First four POMs (Vertical set-up). $-\ \square\ -$, Experimental POM; $\cdots \mathrm{o} \cdots$, linear model; $-\ *\ -$, non-linear model.

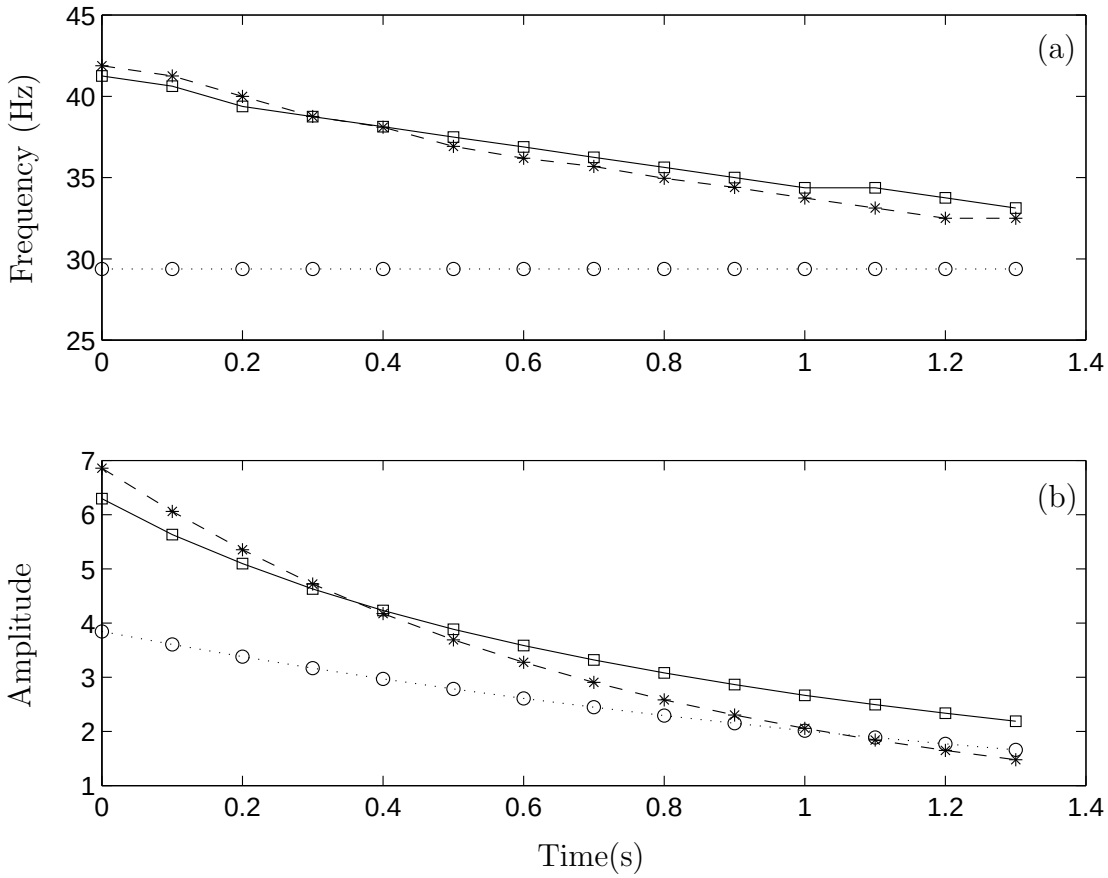


Figure 3.15: Time-frequency properties of the 1st POM (Vertical set-up). $-\,\square\,-$, Experimental system; $\cdots$o$\cdots$, linear model; $-\,*\,-$, non-linear model. (a) Instantaneous frequency; (b) envelope.

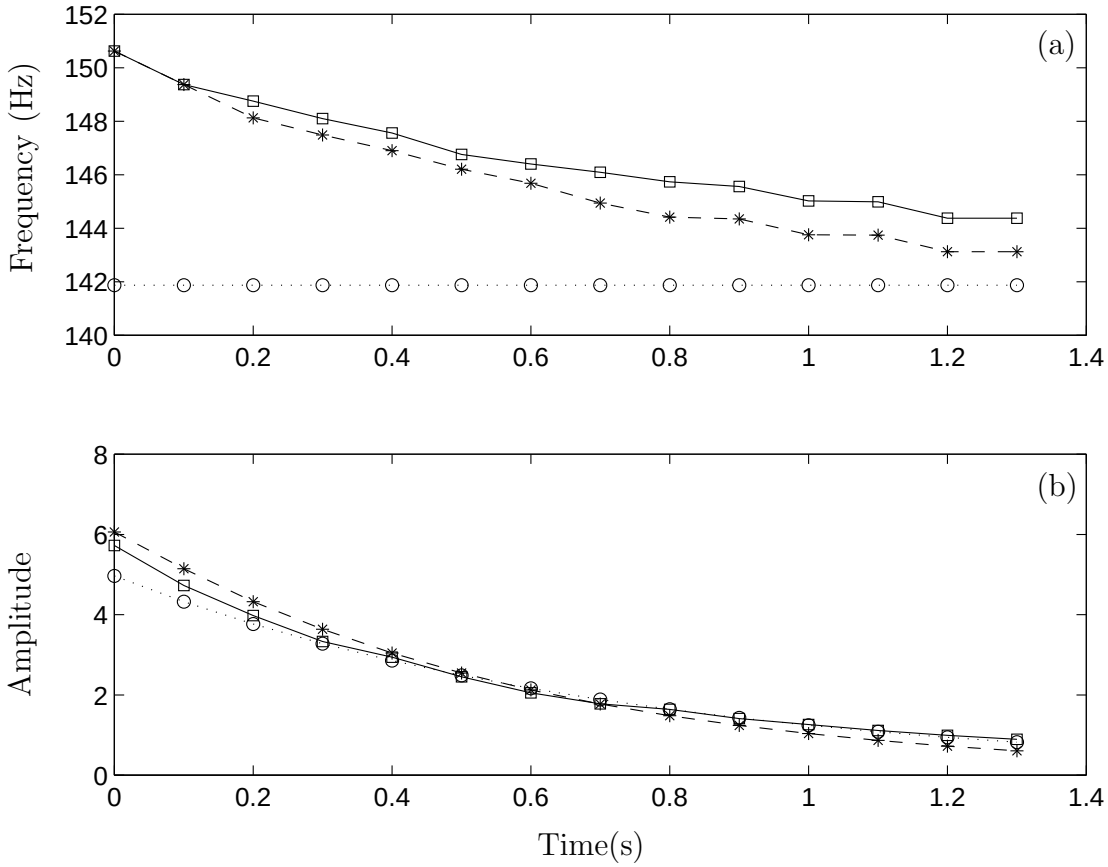


Figure 3.16: Time-frequency properties of the 2nd POM (Vertical set-up). $-\ \square\ -$, Experimental system; $\cdots \mathrm{o} \cdots$, linear model; $-\ *\ -$, non-linear model. (a) Instantaneous frequency; (b) envelope.

Comparison of the results with the CRP method

The comparison of the results obtained with the CRP and POD methods is available in references [95, 110]. It can be stated that they have provided consistent results:

- Having specified the same functional form of the non-linearity for both methods, the same kind of non-linearity has been identified, i.e., a $|x|^{2.8}\,\mathrm{sign}(x)$ non-linearity;
- The non-linear coefficient identified with the POD method is slightly lower than the coefficient obtained with the CRP method for the highest level of excitation, i.e., $1.65\ 10^9$ vs. $1.96\ 10^9\ N/m^{2.8}$. The difference may be due to the different means of excitation used (hammer impact for POD and shaker for CRP) and to the signal processing that is strongly different from one method to the other.

3.7 Concluding remarks

The purpose of this chapter was to introduce the POD and highlight its utility in the field of non-linear structural dynamics. It appears that the POD is a powerful technique for model reduction and model updating purposes for several reasons:

- The POMs may be considered as an alternative to the classical mode shapes for non-linear systems;
- The POVs give clear indications about the participation of the corresponding POMs in the system response;
- The SVD filters out the presence of measurement noise by discarding the smallest singular values;
- The spatial and time information contained in matrices $\mathbf{U}$ and $\mathbf{V}$ respectively is explicitly separated;
- The computation of the POD is not too intensive and does not require the knowledge of the structural matrices;
- The POD is the optimal linear algorithm since, among all linear techniques, it obtains the minimum expected squared distance between the original signal and its dimension-reduced representation.

This chapter has also utilised the SVD for the physical interpretation of the POMs. This study has underlined some features of the POD that might be useful in the future. Since the POMs are related to the vibration eigenmodes in some cases, the POD should be an alternative way of extracting the mode shapes of a structure.

Despite its widespread use, the POD may sometimes be too simple for dealing with real-world data. Since it builds a global, linear model of the data, i.e., an r-dimensional

hyperplane if the r leading eigenvectors of the covariance matrix are retained, two weaknesses should be emphasised:

1. The linear nature of the method may be viewed as a key advantage in the sense that the theory of linear operators is available but it may also represent a severe restriction for some data sets. If the data lies on a non-linear manifold, the method then overestimates the intrinsic dimensionality. For instance, the covariance matrix of data sampled from a helix in $\mathbb{R}^3$ has full-rank, and the POD requires the use of three variables for the description of the data. The helix however is a one-dimensional manifold and can be smoothly parametrised with a single variable. In this case, the method represents the data in a space of higher dimension than the number of intrinsic DOF.

2. The POD tries to describe all the data using the same global features. Generally, a rich data set has often varying characteristics in different regions of the space. This suggests the use of local features for efficient representation of qualitatively different domains of the data.

The POD can thus determine an appropriate embedding space for a low-dimensional structure but it cannot provide the most efficient description of a data set where non-linear dependencies are present. Since there is no a priori reason to believe that any lower-dimensional structure underlying a multi-variate data set is linear, an important issue is to develop techniques that can eliminate dependencies not removed by the POD. To this end, researchers in the field of statistics and neural networks have proposed non-linear alternatives to the POD. This is discussed in further details in the next chapter.

Finally, it is emphasised that the POD may be interpreted from a statistical point of view. Indeed, the transformation removes linear correlations among variables, i.e., diagonalises the covariance matrix. It is only sensitive to second-order statistics and does not necessarily yield statistical independence. Decorrelation implies statistical independence only under the assumption that the variables are Gaussian. This limitation has been recently addressed by the introduction of a method called independent component analysis in reference [80]. However, the investigation of this technique is beyond the scope of this dissertation.

Chapter 4

Non-linear Proper Orthogonal Decomposition

4.1 Introduction

The proper orthogonal decomposition (POD) transformation is limited by virtue of being a linear technique. It is therefore unable to capture complex non-linear correlations. Although the term 'POD' is in common usage in the field of structural dynamics, it should be mentioned that other terminology may be encountered for the same technique. In particular, the POD is better known as PCA in the statistical literature (see for instance references [81, 83, 100]).

Recognising the limitations of the POD, researchers in the statistics and neural network communities have developed non-linear extensions of this method. Principal curves have been defined by Hastie and Stuetzle [67, 68]. A principal curve may be viewed as a non-linear generalisation of a principal component and is a smooth curve that passes through the "middle" of the data. It can provide an insightful summary of high-dimensional data not typically attainable by classical linear methods. The notion of principal curve has been extended to that of principal surface [67, 105]. But due to the difficulty in implementing principal curves and surfaces, the applications of these techniques have been mostly confined to two-dimensional problems. It is only recently that high-dimensional problems have been addressed with the work of Chang [29].

In the early 1990s, a neural network-based generalisation of the POD was introduced in the chemical engineering literature by Kramer [100]. Referred to as non-linear principal component analysis (NLPCA), Kramer implemented the solution using an auto-associative neural network with five layers. Since then, it has been applied to problems in psychology [52] and in biomedical signal processing [164], to find low-dimensional representations of face images [38] or to analyse climate data [130]. In the field of structural dynamics, NLPCA has been exploited by Sohn *et al.* [161] to separate the structural changes from the changes caused by the environment (e.g., temperature, input force level)

for damage detection purposes.

The POD and NLPCA try to describe all the data using the same global features. An alternative paradigm is to capture data complexity by a combination of local linear POD projections. A local model implementation of the POD involves a two-step procedure, i.e., a clustering of the data space into disjoint regions and the estimation of the proper orthogonal modes (POMs) within each region. The first application of a local POD method dates back to 1971 and is due to Fukunaga and Olsen [55]. Since then, it has been applied successfully for the identification of intrinsic dimension of data [24, 55], for handwritten character recognition [75] and for dimension reduction of speech [83] and images [4, 83].

The chapter is organised as follows. The next two sections introduce non-linear extensions of the POD, namely a global approach called NLPCA and a local approach called vector quantisation principal component analysis (VQPCA). In section 4.4, these methodologies are applied to a synthetic data set sampled from a NNM motion. Synthetic data is used to develop intuition about their implementation. Finally, the respective merits of each technique are highlighted in section 4.5.

4.2 Non-linear proper orthogonal decomposition: a global approach

4.2.1 Proper orthogonal decomposition and neural networks

Before describing NLPCA itself, it is interesting to show that the POD can be computed using neural networks. Although neural networks offer no direct advantage over the classical means of computing the POD described in section 3.3, they suggest an interesting non-linear generalisation, i.e., NLPCA.

Neural networks

An artificial neural network is a computational system inspired by the learning characteristics and the structure of biological neural networks. From beginnings about fifty years ago, neural network theory has matured to the point that the artificial neural network has become an irreplaceable tool in many applications and disciplines. The most important advantage of neural networks is in solving problems that are too complex for conventional technologies, i.e., problems that do not have an algorithmic solution or for which an algorithmic solution is too complex to be found.

The type of neural network considered here is composed of a series of parallel layers, each of which contains a number of processing elements analogous to neurons. The neurons are tied together with weighted connections that are analogous to synapses. The network has distinguished input and output neurons, and all other neurons are called hidden

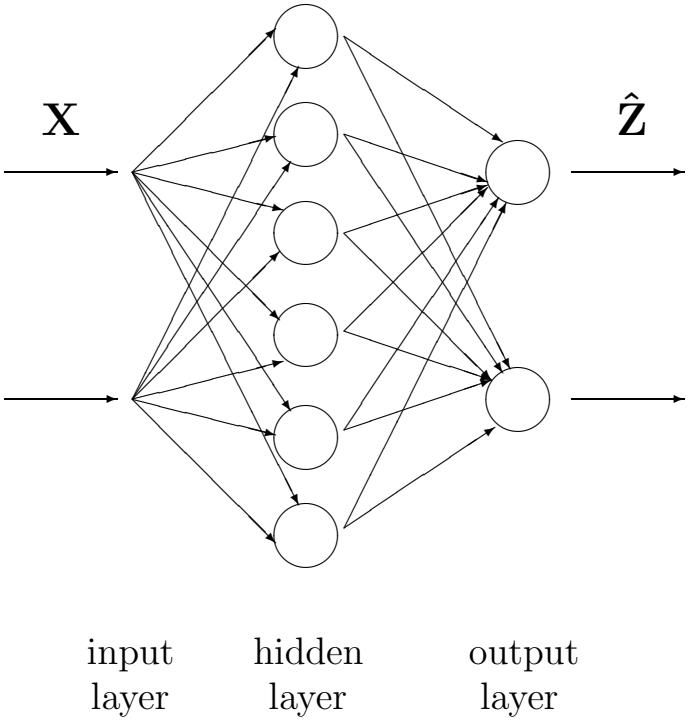


Figure 4.1: Representation of a neural network with input data $\mathbf{X}$ and target data $\mathbf{Z}$.

neurons. Each neuron computes a weighted sum of its inputs, and its resulting output is a well defined function of the weighted sum. The outputs of the ith layer are used as inputs to the $(i+1)$th layer. If $y_j^{(i)}$ is the output of the jth neuron of the ith layer, then

$$y_k^{(i+1)} = \sigma^{(i+1)} \left(\sum_j w_{kj}^{(i+1)} y_j^{(i)} + b_k^{(i+1)} \right) \tag{4.1}$$

is the output of the kth neuron of the $(i+1)$th layer. The elements w_{kj} and b_k are referred to as weights and biases respectively, and $\sigma(\cdot)$ is a linear or non-linear transfer function. Neural networks are often represented as in Figure 4.1 with open circles representing the neurons and straight lines the weights. It should be noted that the input neurons are not represented since their transfer function is merely the identity map.

The first step in using a network is the training phase in which the appropriate values for the connection weights are evaluated. Let us assume that it is desired to fit the data set $\mathbf{X}$ to the data set $\mathbf{Z}$ that is often referred to as target. Denoting the function computed by the network as f, the weights and biases are updated until the following cost function, known as the mean square error (MSE), is minimised:

$$MSE = E[\|\mathbf{Z} - f(\mathbf{X}; \mathbf{w}, \mathbf{b})\|^2] \tag{4.2}$$

where $E[\cdot]$ denotes the expectation. When the transfer functions are non-linear, the MSE must be minimised using an iterative algorithm. Once the error is reduced to an

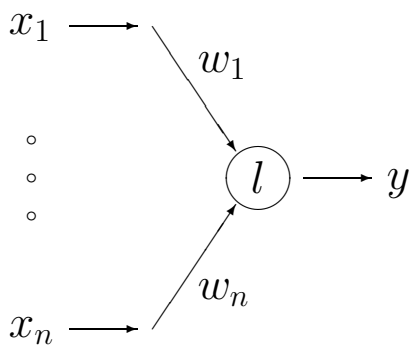


Figure 4.2: Representation of a simple linear neuron model.

acceptable level over the whole training set, the training phase ends and the network is established.

For a deeper description of neural networks and their properties, we invite the interested reader to refer to the literature (e.g., reference [22]).

Oja's learning rule

Consider a simple linear neuron model where the inputs $x_1, ..., x_n$ are connected by the weights $w_1, ..., w_n$ to the output $y = \mathbf{x}^T\mathbf{w}$. The network diagram is shown in Figure 4.2. Oja [136] proposed the stochastic gradient ascent algorithm based on a modified Hebbian rule for which the output neuron is able to extract the POMs:

$$\Delta\mathbf{w} = \mathbf{w}_{(j+1)} - \mathbf{w}_{(j)} = \alpha(y_{(j)}\mathbf{x}_{(j)} - {y_{(j)}}^2\mathbf{w}_{(j)}) \tag{4.3}$$

where α is the learning rate controlling the speed of convergence and $..._{(j)}$ refers to the result of the jth iteration. This learning rule does not contain any explicit normalisation for the weight vector $\mathbf{w}$ but this effect is an automatic result of the non-linear feedback term $y^2\mathbf{w}$.

After convergence, the expectation of the weight update $\Delta\mathbf{w}$ is equal to zero:

$$E[\Delta\mathbf{w}] = E[\alpha(y\mathbf{x} - y^2\mathbf{w})] = \alpha E[\mathbf{x}\mathbf{x}^T\mathbf{w} - \mathbf{w}^T\mathbf{x}\mathbf{x}^T\mathbf{w}\mathbf{w}] = \alpha(\mathbf{\Sigma}\mathbf{w} - \mathbf{w}^T\mathbf{\Sigma}\mathbf{w}\mathbf{w}) = 0 \tag{4.4}$$

and,

$$\mathbf{\Sigma}\mathbf{w} = (\mathbf{w}^T\mathbf{\Sigma}\mathbf{w})\mathbf{w} = \lambda\mathbf{w} \tag{4.5}$$

where $\mathbf{\Sigma} = E[\mathbf{x}\mathbf{x}^T]$ is the covariance matrix of $\mathbf{x}$ that is assumed to be zero mean. As can be seen from equation (4.5), the weights converge to an eigenvector of the covariance matrix, i.e., a POM. In fact, it can be shown that they approach the eigenvector with the largest eigenvalue.

An obvious generalisation of this simple unit to extract the first k POMs is a network of k linear units whose weights vary again according the Hebbian law with internal

feedback [137]:

$$\Delta \mathbf{W} = \mathbf{W}_{(j+1)} - \mathbf{W}_{(j)} = \alpha(\mathbf{y}_{(j)}\mathbf{x}_{(j)}^T - \mathbf{y}_{(j)}\mathbf{y}_{(j)}^T\mathbf{W}_{(j)}) \tag{4.6}$$

In this case, after convergence, the matrix $\mathbf{W}$ containing the weight vectors of all the units span the same subspace as the first k POMs. Sanger [152] also proposed a learning rule called the generalised Hebbian algorithm:

$$\Delta \mathbf{W} = \alpha(\mathbf{y}_{(j)}\mathbf{x}_{(j)}^T - LT(\mathbf{y}_{(j)}\mathbf{y}_{(j)}^T)\mathbf{W}_{(j)}) \tag{4.7}$$

where LT stands for low-triangular and sets the matrix entries above the diagonal to zero. This differs from Oja's learning rule (4.6) in the sense that the POMs are extracted in order, with respect to the output unit ordering.

Auto-associative neural networks

Auto-associative neural networks are neural networks in which the target output pattern is identical to the input pattern. The aim is thus to approximate as closely as possible the input data itself. When used with a hidden layer smaller than the input and output layers, a perfect reconstruction of all input data is generally not possible. However, it can be shown [9] that an auto-associator with linear activation functions performs a compression scheme equivalent to the POD.

To this end, let us consider an auto-associative neural network with a single hidden unit with a linear activation function as shown in Figure 4.3. The sum of squares error (SSE) function E for this network is given by the following equation:

$$E = \frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^T(\mathbf{x} - \hat{\mathbf{x}}) = \frac{1}{2}(\mathbf{x} - y\mathbf{w})^T(\mathbf{x} - y\mathbf{w}) \tag{4.8}$$

The derivative of the error function is:

$$\frac{\partial E}{\partial \mathbf{w}} = -(\frac{\partial y}{\partial \mathbf{w}}\mathbf{w} + y\frac{\partial \mathbf{w}}{\partial \mathbf{w}})(\mathbf{x} - y\mathbf{w}) = -(\mathbf{x}^T\mathbf{w} + y)(\mathbf{x} - y\mathbf{w}) = -2y(\mathbf{x} - y\mathbf{w}) \tag{4.9}$$

and the weight update becomes:

$$\Delta \mathbf{w} = -\alpha\frac{\partial E}{\partial \mathbf{w}} = 2\alpha(y\mathbf{x} - y^2\mathbf{w}) \tag{4.10}$$

This latter expression is equivalent to learning rule (4.3) which demonstrates the ability of the auto-associative network to perform the POD.

In addition, Baldi and Hornik [9] gave a complete description of the landscape of the error function E. They showed that the error function has, up to equivalence, a unique minimum, and that at this minimum, the network with k hidden units performs a projection onto the k-dimensional subspace spanned by the first k POMs of the data.

By way of illustration, the data set represented in Figure 4.4(a) is considered. The landscape of the error function is calculated and is displayed in Figure 4.4(b). Surprisingly

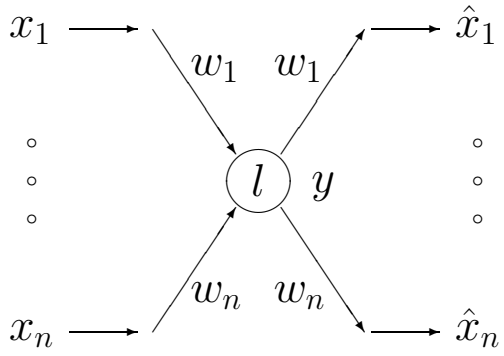


Figure 4.3: An auto-associative neural network with a single hidden unit with a linear activation function.

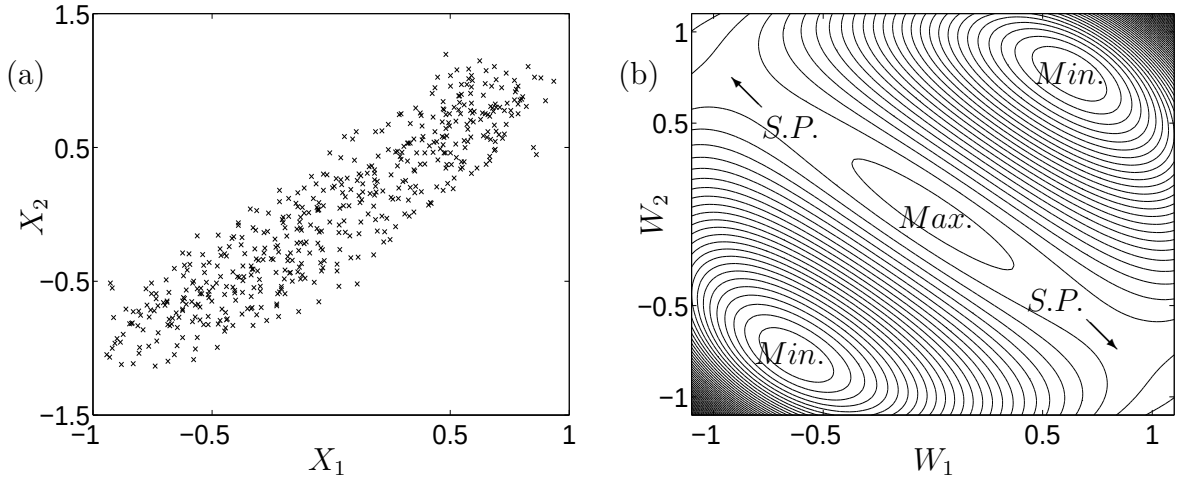


Figure 4.4: Illustrative example. (a) Data set; (b) contour plot of the error function (Min.=minimum, Max.=maximum, S.P.=saddle point).

enough, it can be observed that two distinct minima are present. However, they only differ by a negative sign, and the weights value realises the first POM.

The advantage of neural networks over the eigenvalue problem for the computation of the POD is that there is no need to compute the full covariance matrix. Apart from that, for practical applications, there would appear to be no direct advantage in using neural networks.

4.2.2 Five-layer auto-associative neural networks

The response of a mechanical structure can be expressed as a n-dimensional spatial vector $\mathbf{x}(t)$ that evolves in time. Generally speaking, a dimension reduction technique provides an approximation $\hat{\mathbf{x}}(t)$ to $\mathbf{x}(t)$ that is the composition of two functions f and g:

$$\mathbf{x}(t) = \hat{\mathbf{x}}(t) + \boldsymbol{\epsilon}(t) = g(f(\mathbf{x}(t))) + \boldsymbol{\epsilon}(t) \tag{4.11}$$

The projection function $f : \mathbb{R}^n \rightarrow \mathbb{R}^r$ projects the original n-dimensional data $\mathbf{x}(t)$ onto an r-dimensional subspace while the expansion function $g : \mathbb{R}^r \rightarrow \mathbb{R}^n$ defines a mapping from an r-dimensional space back into the original n-dimensional space with $\boldsymbol{\epsilon}$ as the residue. The feature extraction problem involves the determination of adequate functions f and g. The MSE in reconstructing the original data is:

$$MSE = E[\|\mathbf{x} - g(f(\mathbf{x}))\|^2] \tag{4.12}$$

The POD is the algorithm that obtains the smallest MSE among all techniques with linear projection and expansion functions f and g. To circumvent this limitation, functions that allow for non-linear dependencies between the input variables should be considered. To this end, it is recalled that the POD can be computed using a three-layer auto-associative neural network with linear activation functions. Thus, the most obvious generalisation of the POD would be to consider non-linear transfer functions in the hidden layer. But a much better compression can be obtained by exploiting a theorem due to Cybenko [37]. It states that a three-layer neural network with n input neurons, non-linear transfer functions in the second layer and linear transfer functions in the third layer of r neurons can approximate to arbitrary accuracy any continuous functions from $\mathbb{R}^n$ to $\mathbb{R}^r$. This is true provided that the number of neurons in the second layer is sufficiently large. This theorem suggests the use of a three-layer network to recover projection function f and the use of another three-layer network to recover expansion function g. Both networks are shown in Figure 4.5.

In training the network representing the projection function, the input $\mathbf{x}$ is known but the target $\mathbf{z}$ is unknown. Conversely, in training the network representing the expansion function, the target, i.e., the input data $\mathbf{x}$, is known but the corresponding input $\mathbf{z}$ is not. Therefore, direct supervised training of these networks is not feasible. To alleviate this problem, it should be noted that $\mathbf{z}$ is both the target and the input of the networks representing functions f and g respectively. The networks can thus be combined in series to build a five-layer auto-associative neural network since the input and the target are identical and set to be the system response $\mathbf{x}$.

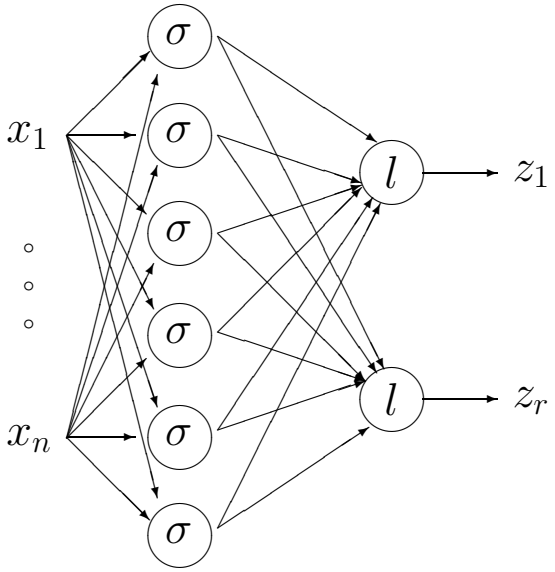
Figure 4.6 illustrates the architecture of a feed-forward auto-associative neural network with five layers:

- The first and fifth layers are referred to as input and output layers respectively. They contain a number of neurons equal to the data dimensionality.

- The second and fourth layers are called the mapping and demapping layers respectively. They are not required to have the same number of neurons. In this thesis, it is fixed to be the same in order to reduce the number of free parameters in the model architecture. The transfer function of the mapping and demapping layers is the hyperbolic tangent represented in Figure 4.7:

$$\sigma(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{4.13}$$

Projection function f:



Expansion function g:

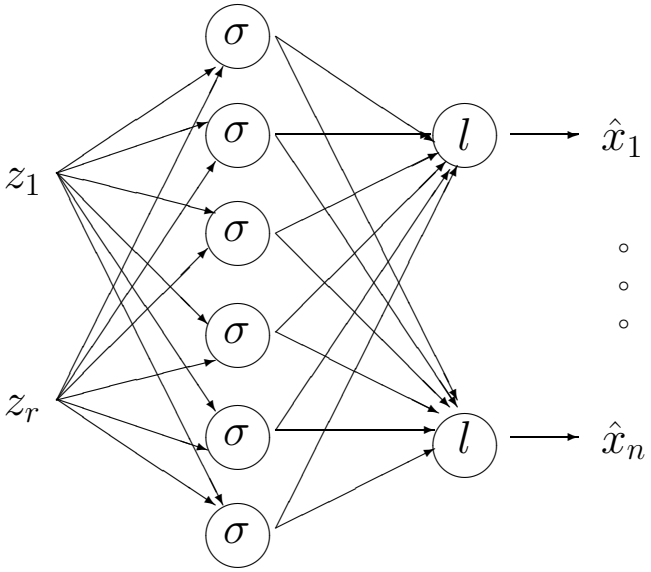


Figure 4.5: Networks implementing projection and expansion functions (σ=sigmoidal transfer function, l=linear transfer function).

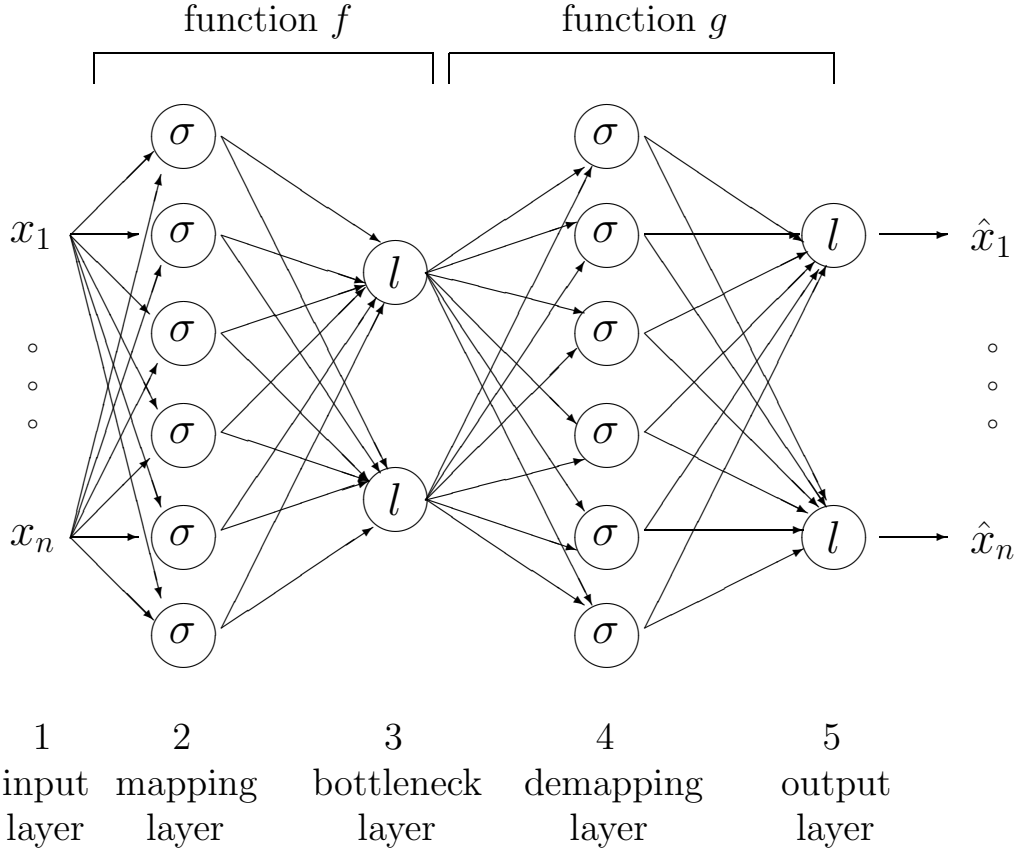


Figure 4.6: Network architecture for implementation of NLPCA (σ=sigmoidal transfer function, l=linear transfer function).

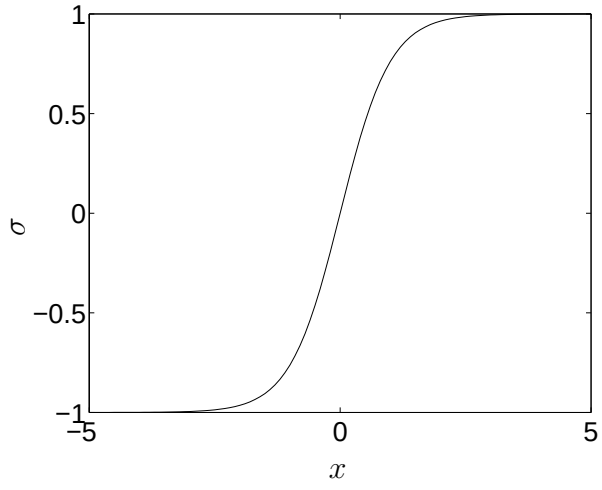


Figure 4.7: Hyperbolic tangent.

The characteristic of this function is that it is bounded between -1 and 1, it is monotonically increasing, and it is continuous and differentiable everywhere. The main advantage of such function lies in its flexibility. Indeed, if the weights are chosen to be small, the transfer function is approximately linear while if the weights are higher, the transfer function is non-linear. This provides the capability for modelling arbitrary functions f and g.

- The third layer is referred to as the bottleneck layer. It must have a smaller dimension than the input and output layers. Since there are fewer neurons in the bottleneck layer, a perfect reconstruction of all input data is not in general possible. This forces the network to optimally encode the input data and to represent significant features in the data. The outputs of the bottleneck layer are called non-linear principal components that may be viewed as a non-linear generalisation of principal components. The bottleneck layer can have either linear or non-linear activation functions. Generally speaking, linear transfer functions are chosen unless bounded response is desired.

4.2.3 Bias/variance trade-off

Generalisation is a fundamental property of a neural network. A reasonable approximation to the data is sought but the model should also give the best possible predictions for novel data. These goals are somewhat conflicting, and the point of best generalisation is determined by the trade-off between bias and variance.

This can be interpreted from a mathematical point of view. The complete mathematical developments are explained in reference [22]. Suppose the aim is to recover a function $x(t)$ for which only noisy data $y(t)$ is available. If the size of the data set goes to infinity, it can be shown that the SSE (or equivalently the MSE) between the network output $\hat{x}(t)$ and the observed data $y(t)$ is the sum of two terms. The first term measures how close the network output $\hat{x}(t)$ is to the desired function $x(t)$. It should vanish for the optimal network. The second term is independent of the network weights and is the result of the intrinsic noise in the data.

It is of interest to have a closer look at the first term. Since its value depends on the training set used, it is better to consider the average over the ensemble of data sets:

$$E_D[\{\hat{x}(t) - E[y|t]\}^2] = \underbrace{\{E_D[\hat{x}(t)] - E[y|t]\}^2}_{(\text{bias})^2} + \underbrace{E_D[\{\hat{x}(t) - E_D[\hat{x}(t)]\}^2]}_{\text{variance}} \tag{4.14}$$

where $E[y|t] = x(t)$ is the conditional expectation of the target data, and E_D is the ensemble average over all possible training sets D. Equation (4.14) consists of the sum of two positive components:

- The bias measures the difference between the average network output $E_D[\hat{x}(t)]$ and the function $x(t)$. This can be viewed as the approximation error.

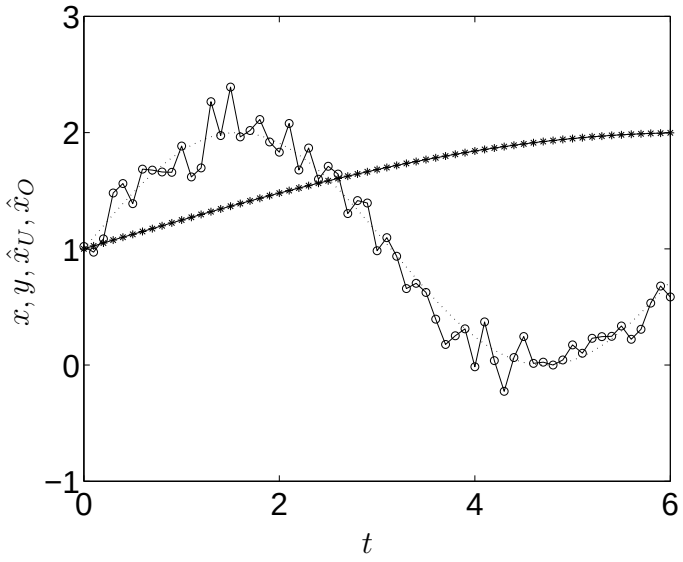


Figure 4.8: An illustration of the bias and variance: $\cdots$, underlying function $x(t)$; o , noisy data $y(t)$; $*$, under-fitted approximation $\hat{x}_U(t)$; —— , over-fitted approximation $\hat{x}_O(t)$.

- The variance is the variance of function $\hat{x}(t)$ over all the training sets D. It represents the sensitivity of the results to the particular choice of a data set.

To have the best generalisation possible, the sum of the bias and variance terms must be minimised. A possible strategy for reducing the bias is to increase the number of hidden units in the network. This allows for more flexibility in the function the network can learn but also increases the variance at the same time. Conversely, a possible strategy for reducing the variance is to limit the number of hidden units but this increases the bias. There is thus a natural trade-off between bias and variance. One way for reducing both terms is to consider more data points.

Let us illustrate these concepts using the example of Figure 4.8. The goal is to model the function $x(t) = sin(t) + 1$ from data corrupted with noise $y(t)$. A neural network that is not sufficiently complex can fail to learn the underlying function x, leading to an under-fitted approximation $\hat{x}_U$ of the data. The resulting model has a large bias. A network that is too complex may fit the noise, not just the underlying function, leading to an over-fitted approximation $\hat{x}_O$. The resulting model has too much flexibility in relation to the particular data set and has a large variance.

4.2.4 Training the network

The first step of using a neural network is the training phase. The weights and biases are updated iteratively until the MSE or the SSE is minimised and the output $\hat{\mathbf{x}}$ approximates the input $\mathbf{x}$ as closely as possible. Minimising the MSE during the network training results

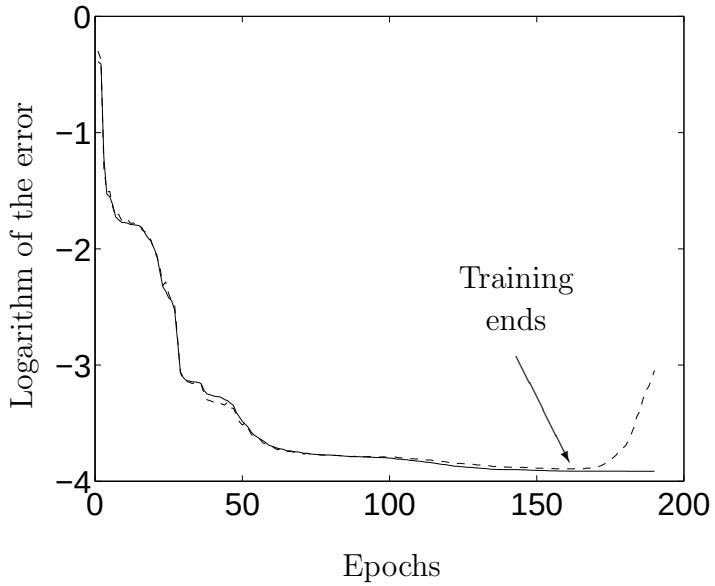


Figure 4.9: Error over the training and validation sets as a function of epochs: ——, training set; - - -, validation set.

in minimum information loss in the same sense as the POD [see equation (4.12)].

The error over the training set generally decreases as a function of the number of iterations also called epochs in the neural network community. However, as suggested by the previous section, it is not desirable to learn the particular details of the training set which may lead to over-training. To have the best generalisation performance, a procedure called early stopping is considered in which the training phase is terminated before the error function is minimised. In essence, the idea behind early stopping is that the training is allowed to continue sufficiently long to fit the structure underlying the data, but not long enough to fit the noise. A fraction of data points selected randomly is held aside in a validation set not used to train the network. While network training proceeds, the network performance on the validation set is monitored, and training is stopped when this performance begins to degrade. This procedure is illustrated in Figure 4.9.

In this dissertation, the training phase is carried out using the conjugate gradient method. In the case of a quadratic function, this algorithm finds mutually conjugate search directions in the weight space, and these directions are such that minimisation along one of them does not spoil the previous minimisation along the other directions. In the case of a non-quadratic function, the conjugacy of the generated directions may then progressively be lost. This is the reason why the algorithm is restarted every n iterations where n is the number of variables. A more detailed discussion about this technique is available in references [64, 132].

4.2.5 Selection of the number of neurons

A five-layer feed-forward auto-associative neural network has been chosen for the implementation of NLPCA but it still remains to determine the number of neurons in each layer. The input and output layers have obviously a number of neurons equal to the data dimensionality. This section presents some rules to set the number of neurons in the other layers, namely the mapping, bottleneck and demapping layers.

Mapping and demapping layers

An assumption that is made throughout this dissertation is that the number of neurons p in the mapping and demapping layers is set to be identical. This is not guided by any theoretical development, the only reason being to limit the number of free parameters in the network.

The number of neurons is related to the complexity of the non-linear functions that can be generated by the network. If too few mapping neurons are provided, accuracy might be low because the representational capacity of the network is limited. As the number of hidden neurons increases, there is more flexibility in the function the network can learn. However, care should be taken to avoid allowing the network to fit the noise which is referred to as over-fitting. Indeed, there is a number of hidden neurons above which the generalisation performance of the network begins to degrade. This is the optimum in terms of the trade-off between bias and variance.

Several rules have been proposed in the literature for over-fitting avoidance. The most widely used technique is likely the cross-validation technique in which the performances of networks with different number of neurons are compared on an data set, different from the training set. This data set can be the same as the one used for early stopping, i.e., the validation set. By gradually increasing the number of mapping neurons, the error over the training set decreases steadily but the error over the validation set should show a minimum. The number of neurons that realises this minimum is then selected. This procedure is somewhat similar to over-training avoidance and is illustrated in Figure 4.10.

It is also emphasised that the optimal number of neurons is related to the number of training samples. For a network with a single hidden layer, it increases with the number of training samples as shown by Barron in reference [11]:

$$p_{optimal} = \mathcal{C}\sqrt{\frac{N}{n \log N}} \tag{4.15}$$

where n and N are the number of inputs and the training set size respectively, and $\mathcal{C}$ is a measure of problem complexity as defined by Barron.

Bottleneck layer

The POD has an additive structure in the sense that the optimal linear r-dimensional substructure of $\mathbf{x}(t)$ can be found all at once, or POM by POM. In the iterative approach,

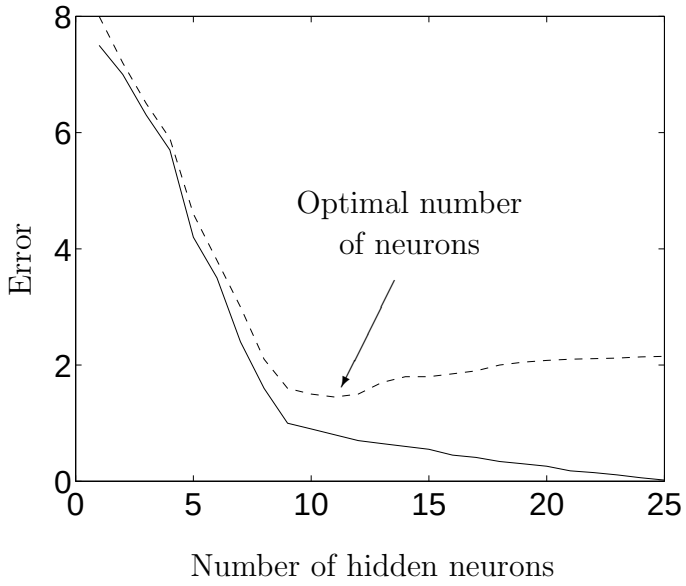


Figure 4.10: Error over the training and validation sets as a function of the number of hidden neurons: - - -, validation set; ——, training set.

the first POM of $\mathbf{x}(t)$ is calculated from the entire data set. The reconstructed dynamic response $\hat{\mathbf{x}}^{(1)}(t)$ is then computed from the POM. The second POM is calculated from the residual $(\mathbf{x}(t) - \hat{\mathbf{x}}^{(1)}(t))$, taking advantage of the fact that the second POM of $\mathbf{x}(t)$ is the first POM of the residual, and so forth.

Unlike the POD, NLPCA does not have an additive structure since an arbitrary smooth function f of r variables cannot be decomposed as a sum of smooth functions of a single variable:

$$f(x_1, ..., x_r) \neq f_1(x_1) + ... + f_r(x_r) \tag{4.16}$$

for some functions $f_1, ..., f_r$. Therefore, an auto-associative neural network with several neurons in the bottleneck layer gives different results in comparison with a series of separate networks with a single neuron in the bottleneck layer, the residual of the ith network being the input for the $(i+1)$th network. The network with several neurons in the bottleneck layer should be superior to the series of networks because it is drawn from a broader class of functions.

From a geometrical point of view, a neural network with r neurons in the bottleneck layer learns an r-dimensional surface that best represents the data. With a single neuron, the outputs of the network are forced to be aligned along a one-dimensional curve.

The selection of the number of neurons in the bottleneck layer plays a prominent role since it determines the order of reduction. To have an idea of the compression performed by the neural network, the fraction of explained variance (FEV) can be calculated:

$$FEV = \frac{E[\|\hat{\mathbf{x}}(t)\|^2]}{E[\|\mathbf{x}(t)\|^2]} \tag{4.17}$$

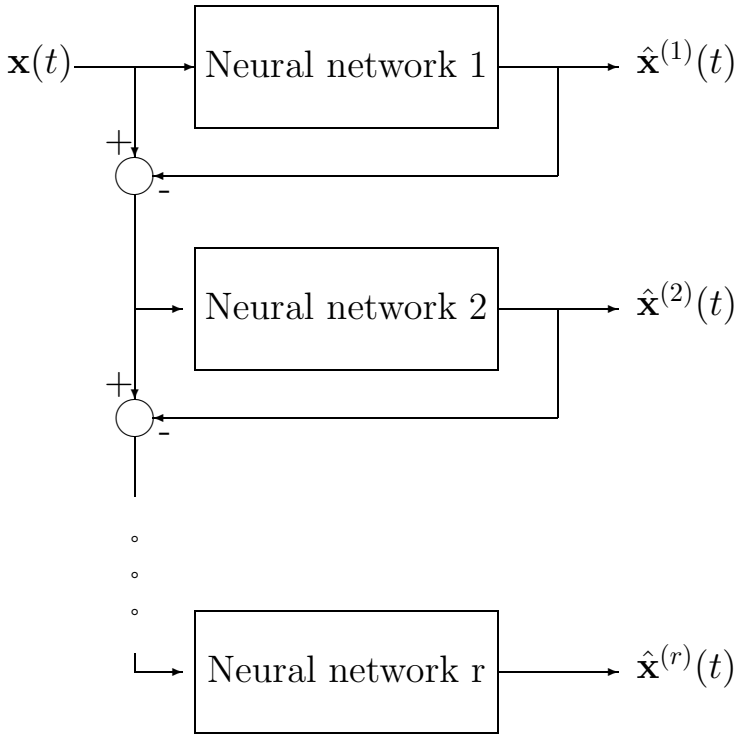


Figure 4.11: Architecture for a "modal" NLPCA analysis.

The FEV indicator is similar in spirit to the eigenvalues of the covariance matrix that explain the energy percentage captured by the POD. A possible strategy is thus to gradually increase the number of neurons until a prescribed value of the FEV has been attained.

However, it is worth recalling that our motivation is to obtain features that could be viewed as an alternative to the notion of mode shape in non-linear structural dynamics. A series of r separate networks with a single neuron in the bottleneck layer better fulfills this requirement. Indeed, they give a collection of r curves that are more amenable to physical interpretation than a single r-dimensional feature. In addition, the r features can be ranked according to their order of appearance, i.e., the feature extracted by the first network is the one with the most importance in the system response, and so forth.

This the reason why a "modal" NLPCA analysis illustrated in Figure 4.11 is considered throughout this dissertation.

4.2.6 Feature extraction

Reconstructed dynamic response

The reconstructed dynamic response $\hat{\mathbf{x}}(t)$ is given by the activations of the output layer and is a projection of the input data onto the hypersurface learnt by the neural network. In the case of a neural network with a single neuron in the bottleneck layer (modal NLPCA

analysis) and provided that the global minimum of the MSE or the SSE is found, the reconstructed dynamic response is the optimal curve characterising the data.

In a sense, this one-dimensional approximation can be compared to classical modes such as the ones obtained through modal analysis or the POD. This is the reason why the reconstructed dynamic response is referred to as non-linear proper orthogonal mode (NLPOM) in this thesis. Although the NLPOM should not be considered as an invariant manifold like the NNMs, it is our opinion that it can be regarded as an appealing feature for the purpose of test-analysis correlation.

Non-linear principal components

The non-linear principal components (NLPCs) $\mathbf{z}(t)$ are a non-linear generalisation of principal components and are given by the outputs of the bottleneck layer. In the case of a modal NLPCA analysis, the NLPC $z(t)$ is a scalar and contains the co-ordinates of the input data on the curve learnt by the network.

It should also be noted that a particular reconstruction $\hat{\mathbf{x}}(t)$ of the data $\mathbf{x}(t)$ can be realised by different values of the NLPC $z(t)$. Indeed, the projection and expansion functions f and g are only determined up to a continuous function h with a continuous inverse h^{-1} because:

$$\hat{\mathbf{x}} = g(z) = g(f(\mathbf{x})) = g(h^{-1}(h(f(\mathbf{x})))) \tag{4.18}$$

and,

$$z = f(\mathbf{x}) \quad \text{or} \quad z = h(f(\mathbf{x})) \tag{4.19}$$

By way of illustration, consider the unit circle that is a one-dimensional parameterised curve in $\mathcal{R}^2$:

$$\mathbf{x} = f(z) = \begin{pmatrix} \cos(z) \\ \sin(z) \end{pmatrix} \quad \text{with } z \,\epsilon\, [0, 2\pi] \tag{4.20}$$

There are many different functions f that define the same curve. For example, the unit circle can also be defined by:

$$\mathbf{x} = f^*(z) = \begin{pmatrix} \cos(az) \\ \sin(az) \end{pmatrix} \quad \text{with } z \,\epsilon\, [0, \frac{2\pi}{a}] \tag{4.21}$$

This issue is not critical for a curve because the parametrisation is determined up to the choice of the origin and scaling. The information present is not radically changed, and the NLPCs might be interesting features for the characterisation of non-linear systems.

Discussion

Two kinds of features have been investigated, i.e., the NLPOMs and the NLPCs (see Figure 4.12).

For a modal NLPCA analysis, the NLPOM is a vector whose dimension is equal to the data dimensionality while the NLPC is a scalar. It is also emphasised that a NLPOM

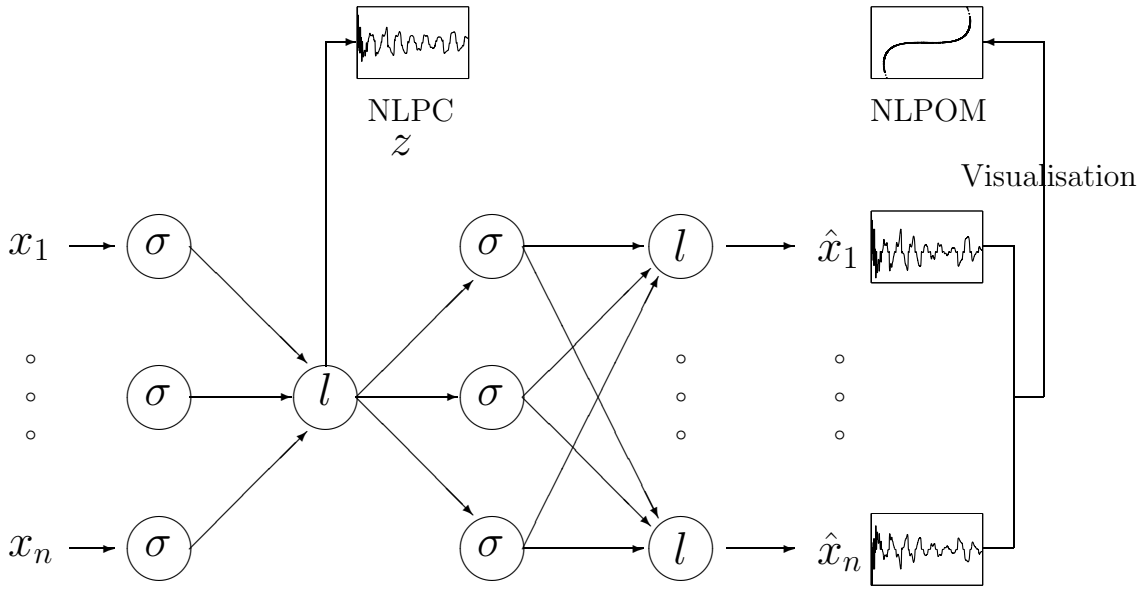


Figure 4.12: Feature extraction using NLPCA.

is the representation of the NLPC in the input data space. These two features contain thus the same information but in different spaces.

From this discussion, it appears that the NLPC should be preferred for the formulation of the objective function in model updating processes because it reduces the amount of data that has to be analysed. To solve the problem of the origin and scaling, the NLPC will be forced to have a zero mean and a unitary norm.

Since the NLPOM draws a curve in the data space, it facilitates the visual representation of the NLPC and will be exploited for visualisation purposes at the correlation step.

4.2.7 Limitations

Despite its aesthetic appeal, Malthouse [123] has pointed out three limitations of NLPCA:

- NLPCA is unable to model curves and surfaces that intersect themselves. This limitation originates from the fact that the bottleneck neuron is not an angular variable. Kirby and Miranda [98] have introduced the concept of a circular neuron and have shown that NLPCA with a circular node in the bottleneck layer is capable of extracting closed curve solutions. Instead of a single neuron, there are now two coupled neurons that realise a single angular variable.

- NLPCA is unable to parameterise curves with discontinuities. This drawback is due to the fact that the mapping and demapping projections are continuous functions. Monahan [130] has noticed that a four-layer neural network is better able to

approximate a discontinuous function than is a three-layer network. A seven-layer network where the first four and last four layers approximate the potentially discontinuous projection and expansion functions respectively is helpful to overcome this limitation.

- As discussed in the previous section, the projection and expansion functions f and g are only determined up to an arbitrary homeomorphism. This issue is not critical for a curve but for surfaces, the indeterminacy is more problematic because it can include rotations. This leads to a serious complication in the interpretation of the time series produced by the bottleneck layer when there is more than a single neuron in the layer.

Finally, as for the POD, NLPCA tries to describe all the data using the same global features. A rich data set may have varying characteristics in different regions of the space. In this case, the use of local features for efficient representation of qualitatively different domains of the data may be of great interest. This is discussed in the next section with the introduction of VQPCA.

4.3 Non-linear proper orthogonal decomposition: a local approach

4.3.1 Dimension reduction by local linear models

This section presents a local alternative to the POD, denoted as VQPCA, where VQ stands for vector quantisation. A local POD approach implies the integration of two procedures:

1. A clustering of the data space into distinct regions. VQPCA partitions the space by vector quantisation.

2. The construction of separate low-dimensional co-ordinate systems in each local region using the POD.

Let $\mathbf{X}$ be a set of observed n-dimensional data points $\mathbf{x}_i$ with $i = 1, ..., m$. In order to reduce the dimensionality of the vector $\mathbf{x}_i$, a projection function $f : \mathbb{R}^n \rightarrow \mathbb{R}^r$ such that $\mathbf{z}_i = f(\mathbf{x}_i)$ is a compact r-dimensional representation of $\mathbf{x}_i$ needs to be determined. Similarly, an expansion function $g : \mathbb{R}^r \rightarrow \mathbb{R}^n$ has to be calculated such that $\hat{\mathbf{x}}_i = g(f(\mathbf{x}_i))$ is the reconstruction of the initial vector $\mathbf{x}_i$.

In the present study, the purpose is to build low-dimensional co-ordinate systems in the q local regions defined by the vector quantiser. If the local regions are small enough, the data manifold is not curved much over the extent of the region and it may be locally approximated as a hyperplane. In other words, a separate POD model in each of the q

regions should be adequate. Instead of having single projection and expansion functions, a collection of functions $f_i(\cdot)$ and $g_i(\cdot)$ with $i = 1, ..., q$ is obtained.

4.3.2 Vector quantisation

Vector quantisation is a classical technique for signal coding and data compression [59]. A q-level vector quantiser is defined by a codebook $\mathcal{C} = (\boldsymbol{\mu}_1, ..., \boldsymbol{\mu}_q)$, a partition $\mathcal{S} = (S_1, ..., S_q)$ and a distortion function $d(\mathbf{x}, \boldsymbol{\mu})$. It is a mapping f that approximates each point $\mathbf{x}_i$ in the set $\mathbf{X}$ by a component $\boldsymbol{\mu}_j$ of the codebook $\mathcal{C}$: $f(\mathbf{x}_i) = \boldsymbol{\mu}_j$ if $\mathbf{x}_i \in S_j$. A q-level quantiser is said to be optimal if it minimises the averaged distortion $D = E[d(\mathbf{x}, \boldsymbol{\mu})]$.

In his implementation of VQPCA, Kambhatla [83] makes use of vector quantisation to define the regions for the local POD. The algorithm to design the vector quantiser is based on an approach proposed by Lloyd in reference [117] and is referred to as the generalised Lloyd algorithm [59]. The codebook vectors $\boldsymbol{\mu}_j$ and the regions S_j satisfy Lloyd's optimality conditions:

1. Each region S_j (with its corresponding codebook vector $\boldsymbol{\mu}_j$) corresponds to all $\mathbf{x}_i$ that lie closer to $\boldsymbol{\mu}_j$ than to any other codebook vector (nearest neighbor mapping). Mathematically, $S_j = \{\mathbf{x}_i \mid d(\mathbf{x}_i, \boldsymbol{\mu}_j) < d(\mathbf{x}_i, \boldsymbol{\mu}_k), \forall\, k \neq j\}$;

2. Each codebook vector $\boldsymbol{\mu}_j$ is placed at the centroid of the corresponding region S_j.

For a distortion function based on Euclidean distance, the regions are convex sets called Voronoi cells, and the centroid of a region is the mean of the data points in this region $\boldsymbol{\mu}_j = E[\mathbf{x}_i \mid \mathbf{x}_i \in S_j]$.

Accordingly, the generalised Lloyd algorithm is as follows:

1. Given q the number of regions, initialise the codebook $\mathcal{C}$ from randomly selected points in the data set $\mathbf{X}$;

2. Compute the corresponding optimal partition following the first optimality condition;

3. Compute the corresponding optimal codebook following the second optimality condition;

4. Iterate steps 2 and 3 until convergence.

The convergence is achieved when the fractional change in the averaged distortion D between the kth and $(k + 1)$th iterations is below some specified threshold. It can be argued that each iteration of the algorithm either reduces the distortion or leaves it unchanged.

It is worth noticing that several variants of this algorithm exist, e.g., tree-searched VQ and multistep VQ [59]. These variants aim at reducing the computation or memory

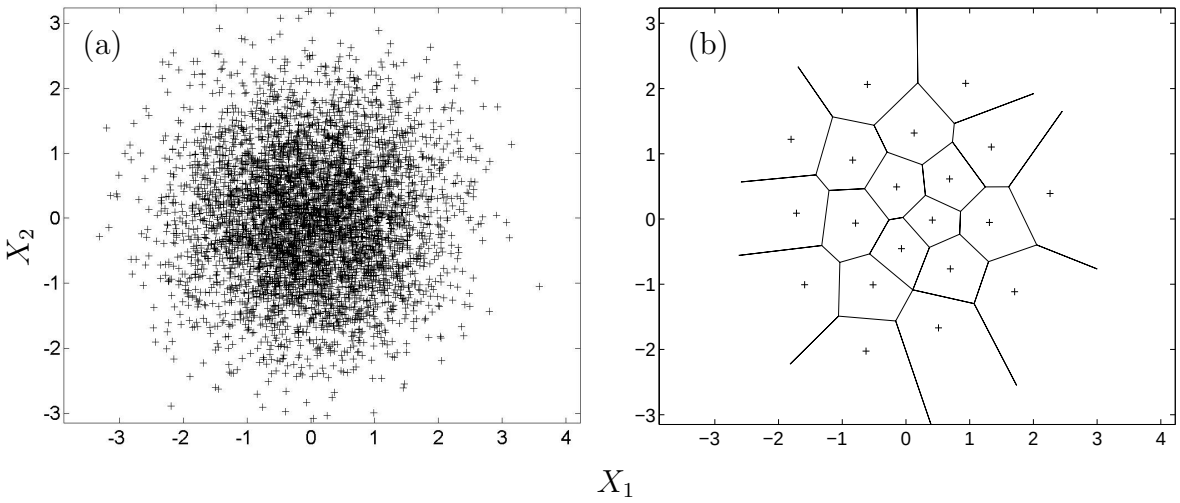


Figure 4.13: (a) Data distribution; (b) clustering of the data: —, Voronoi cells; +, centroids.

requirements but may compromise the performance relative to what could be achieved with a standard VQ. Since the computational aspects are not an issue in this study, the basic Lloyd algorithm is considered throughout this dissertation.

By way of illustration, the generalised Lloyd algorithm, with a distortion function based on the Euclidean distance and a number of regions equal to 20, is applied to a set of 5000 two-dimensional random vectors chosen from a normal distribution with mean zero and variance one. The distribution of the random vectors is displayed in Figure 4.13(a). Figure 4.13(b) depicts the Voronoi cells S_j together with their corresponding centroids $\boldsymbol{\mu}_j$.

4.3.3 Vector quantisation principal component analysis algorithm

The VQPCA algorithm is an extension of a standard vector quantiser. VQPCA partitions the input space into a set of regions and approximates each region by a hyperplane defined by the POD, while a standard vector quantiser approximates each region by a codebook vector. The VQPCA algorithm is:

1. Partition $\mathbb{R}^n$ into q disjoint regions $S_1, ..., S_q$ using the generalised Lloyd algorithm with Euclidean distance as the distortion function.

2. For each Voronoi cell S_j and its corresponding centroid $\boldsymbol{\mu}_j$, estimate the local covariance matrix:

$$\boldsymbol{\Sigma}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in S_j} (\mathbf{x} - \boldsymbol{\mu}_j)(\mathbf{x} - \boldsymbol{\mu}_j)^T \tag{4.22}$$

where N_j is the number of vectors mapped to S_j. Next, compute the eigenvectors $(\mathbf{p}_{j1}, ..., \mathbf{p}_{jr})$ of each matrix $\boldsymbol{\Sigma}_j$.

3. To reduce the dimension of any vector $\mathbf{x}_i$, determine the cell S_j that contains the vector and project $\mathbf{x}_i$ onto the r leading eigenvectors to obtain the local linear co-ordinates:

$$\mathbf{z}_i = f_j(\mathbf{x}_i) = \left[\mathbf{p}_{j1} \; ... \; \mathbf{p}_{jr}\right]^T (\mathbf{x}_i - \boldsymbol{\mu}_j) = \begin{bmatrix} \mathbf{p}_{j1}^T(\mathbf{x}_i - \boldsymbol{\mu}_j) \\ \vdots \\ \mathbf{p}_{jr}^T(\mathbf{x}_i - \boldsymbol{\mu}_j) \end{bmatrix} \qquad \text{if } \mathbf{x}_i \in S_j \tag{4.23}$$

The compressed representation of $\mathbf{x}_i$ consists of the index j of the Voronoi cell in which $\mathbf{x}_i$ lies and the r-dimensional vector $\mathbf{z}_i$. The data is reconstructed from this representation according to:

$$\hat{\mathbf{x}}_i = g_j(f_j(\mathbf{x}_i)) = g_j(\mathbf{z}_i) = \boldsymbol{\mu}_j + \left[\mathbf{p}_{j1} \; ... \; \mathbf{p}_{jr}\right] \mathbf{z}_i \tag{4.24}$$

4.4 An illustrative example

The aim of this section is to illustrate NLPCA and VQPCA on a simple example. This example is also used to highlight the potential benefits of these methods in comparison with classical POD. For visualisation purposes, the data set taken from this example is chosen to be two-dimensional.

The example consists of a non-linear, conservative system with two unit masses that are connected by linear springs k and k_c as shown in Figure 4.14. The first mass is also connected to the ground through a non-linear spring k_{nl} that behaves according to $f_{nl}(x_1) = k_{nl}\, x_1^3$. For $m = k = 1$, $k_c = k_{nl} = 15$ and initial conditions on the displacement vector $\mathbf{x} = [x_1 \; x_2]^T$ equal to $[1 \; 0.918]^T$, the motion corresponds to a single and synchronous non-linear normal mode (NNM), i.e., the displacements are confined to a curve represented in Figure 4.15(a). 2000 samples are considered from this NNM motion with a sampling frequency set to 50 Hz. White Gaussian noise is added to the data such that the noise contributes to 20 % of the signal root mean square value [see Figure 4.15(b)]. The data is then rescaled so that it has zero mean and unit standard deviation. This scaling weighs the two variables equally important.

4.4.1 Proper orthogonal decomposition

Before investigating NLPCA and VQPCA, the POD is first applied to the data of Figure 4.15(b). The mode given by this technique is shown in Figure 4.16 together with the synchronous NNM. This figure confirms the conclusions drawn in section 3.4.5, i.e., the POM can be considered as the best linear representation of the NNM.

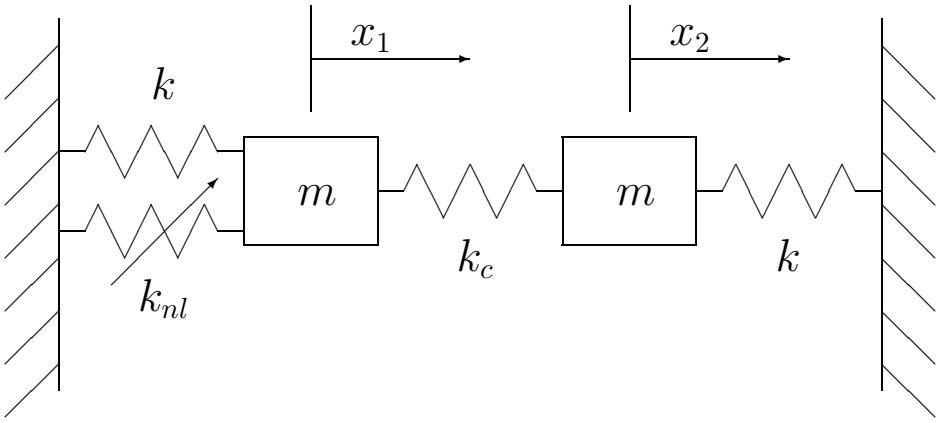


Figure 4.14: Model of the 2-d.o.f. example.

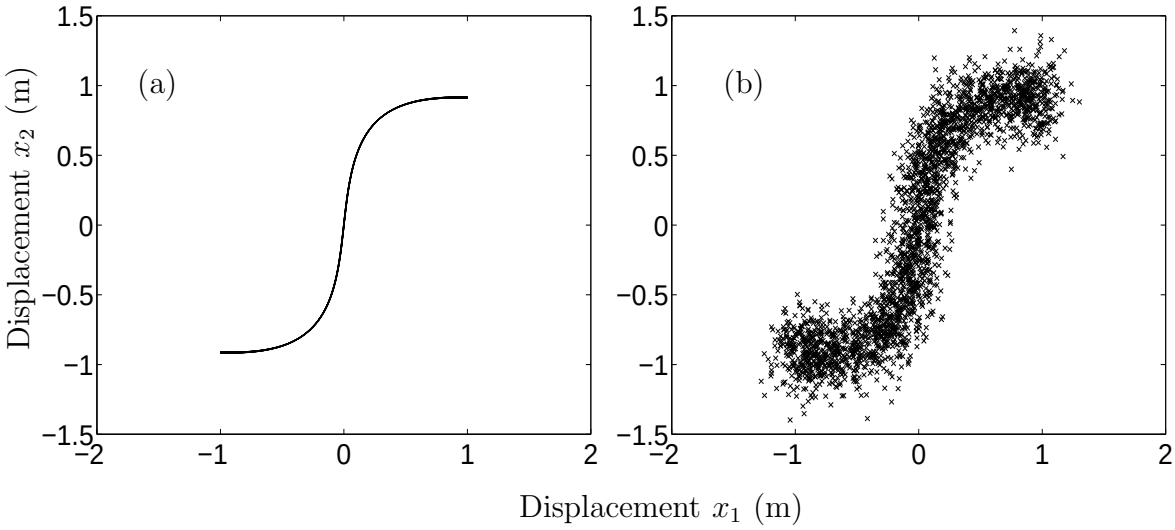


Figure 4.15: Synchronous NNM motion. (a) Without noise; (b) 20 % noise added.

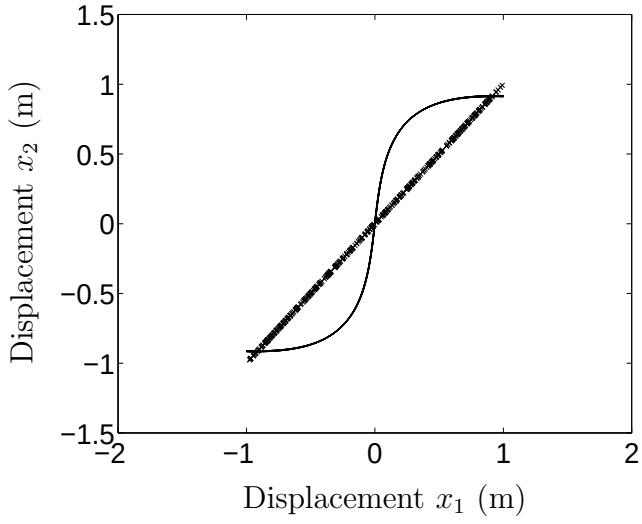


Figure 4.16: Comparison of the modes: ×, POM; ——, synchronous NNM.

The accuracy of the compressed representation is assessed using the normalised mean square error (NMSE):

$$NMSE = \frac{E[\|\mathbf{x} - \hat{\mathbf{x}}\|^2]}{E[\|\mathbf{x} - E[\mathbf{x}]\|^2]} \tag{4.25}$$

The NMSE is equal to 5.71 %, and the POM explains 94.29 % of the total variance.

4.4.2 Non-linear principal component analysis

NLPCA is now applied to the data of Figure 4.15(b). Several tests are first performed with a data set size varying from 250 samples up to 2000 samples. It is observed that beyond 500 samples (400 samples for the training set and 100 for the validation set), the NMSE is no longer improved.

Selection of the number of neurons

The evolution of the NMSE on the validation set with the number of neurons in the mapping and demapping layers is displayed in Figure 4.17. It turns out that a network with 10 neurons realises the minimum value for the NMSE. However, it is worth comparing the NLPOMs provided by the networks with 2 and 10 neurons (Figure 4.18). Apart from slight differences, these modes remain essentially the same.

However, the time needed for the training of the network with 10 neurons is more than three times as high than for the network with two neurons. The decrease in the NMSE (around 0.06 %) does not appear to be sufficient to justify the increase in complexity.

These results reveal that the cross-validation technique needs to be improved, at least for feature extraction purposes. The goal is to obtain the most accurate representation of a data set but the complexity of the network required to compute the solution should also be taken into account. In this context, the information theoretic criterion (AIC) and the final prediction error (FPE) [116] should be useful to adjust the number of neurons:

$$AIC = \overbrace{\ln(e)}^{error} + \overbrace{2\frac{N_w}{N}}^{complexity} \tag{4.26}$$

$$FPE = \overbrace{e}^{error} \frac{\overbrace{(1 + \frac{N_w}{N})}^{complexity}}{1 - \frac{N_w}{N}} \tag{4.27}$$

where $N = mn$ is the number of training vectors times their dimension;

$e = \|\mathbf{x} - g(f(\mathbf{x}))\|^2/(2N)$ is the average SSE on the training set;

$N_w = n + r + 2p(n + r + 1)$ is the number of weights and biases in the network;

r is the number of neurons in the bottleneck layer.

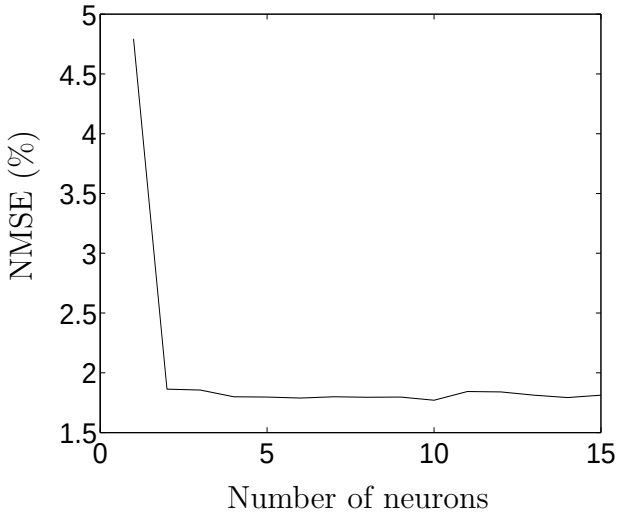


Figure 4.17: Evolution of the NMSE on the validation set with the number of neurons.

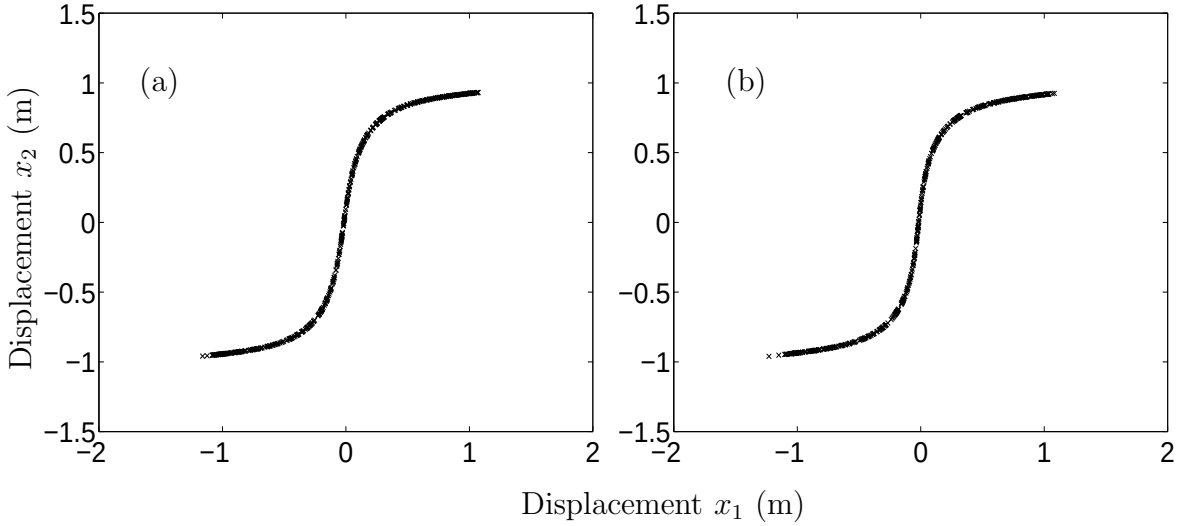


Figure 4.18: NLPOM. (a) 2 neurons; (b) 10 neurons.

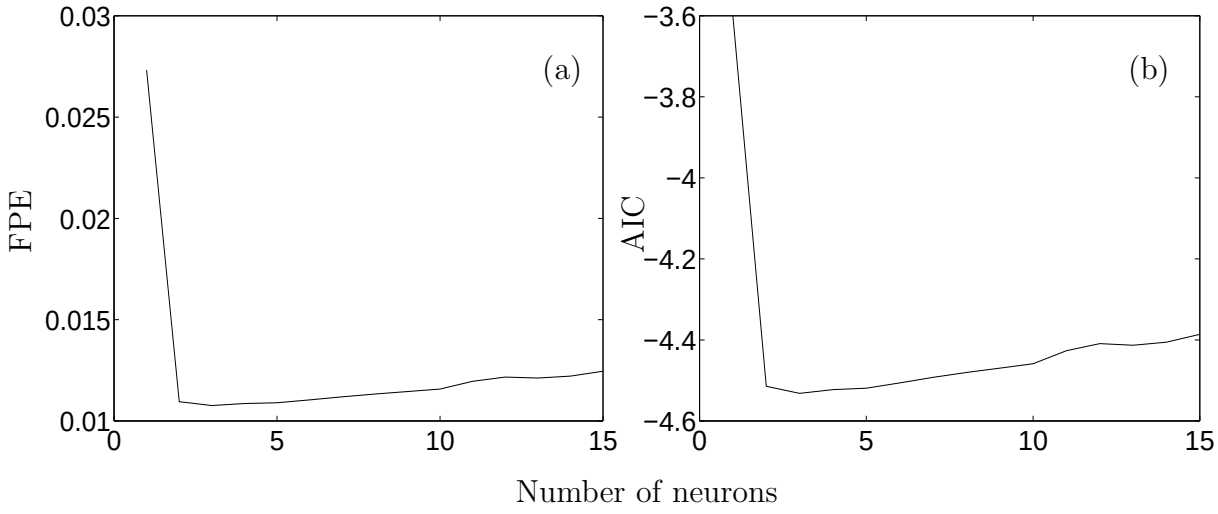


Figure 4.19: (a) FPE; (b) AIC.

The FPE et AIC criteria express the trade-off between goodness-of-fit and the number of adjustable parameters. Unlike the cross-validation technique, they penalise network complexity. Through their minimisation, a neural network with a reasonable number of neurons should be obtained. The evolution of the FPE and AIC indicators with the number of neurons is displayed in Figure 4.19. They both show a minimum for a neural network with 3 neurons in the mapping and demapping layers. However, one would have expected to obtain a number of neurons equal to 2 since the NLPOMs provided by the networks with 2 and 10 neurons are almost identical. In fact, the FPE et AIC criteria are only valid in the case of a linear regression, and they tend to overfit with neural networks.

Although the Schwarz information criterion (SIC) is also rigorously established for linear models, it is often employed in the neural network community because of its tendency to choose more parsimonious models than AIC and FPE. Schwarz [155] has established the criterion as an asymptotic approximation to a transformation of the Bayesian posterior probability of a candidate model:

$$SIC = \overbrace{\ln(e)}^{error} + \overbrace{\frac{N_w}{N}\ln(N)}^{complexity} \tag{4.28}$$

The result given by this indicator is shown in Figure 4.20. This figure shows a clear minimum for 2 neurons which tends to confirm the superiority of SIC over AIC and FPE.

Another interesting means of determining an appropriate number of neurons is by using the "1 standard error rule". This rule is often used for decision tree pruning. When applied to neural networks, it states that the optimal network is the smallest network whose error is within one standard deviation of the minimal error. This is illustrated in Figure 4.21. It can be observed that the error for two neurons is not larger than the minimal error (realised for 10 neurons) plus its standard deviation. As for the SIC indicator, the number of neurons suggested by this rule is equal to 2.

The "1 standard error rule" requires less assumptions than the SIC indicator and is thus considered in this thesis.

Optimal NLPOM

The optimal NLPOM for the "1 standard error rule" is represented in Figure 4.18(a). The comparison with the synchronous NNM points out that both curves agree to the point where the difference between them is not visible. Figure 4.22 shows that the time series produced by NLPCA are also in good agreement with those obtained before the addition of noise. The neural network can thus recover the initial NNM motion although 20 % of noise has been added to the data. The NMSE is equal to 1.86 %, and the NLPOM explains 98.60 % of the variance.

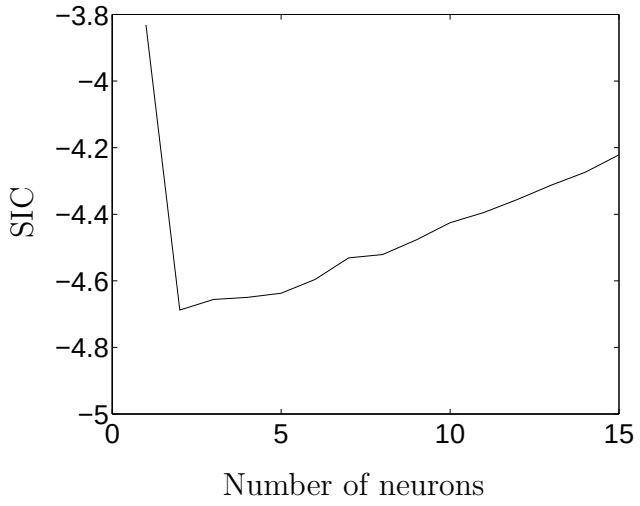


Figure 4.20: SIC.

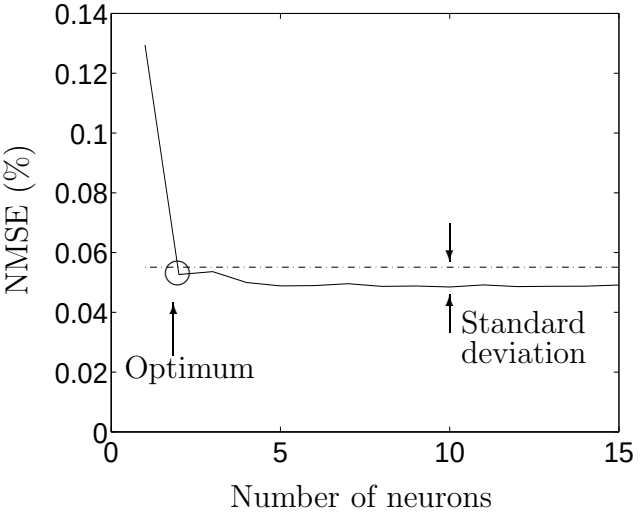


Figure 4.21: 1 standard error rule.

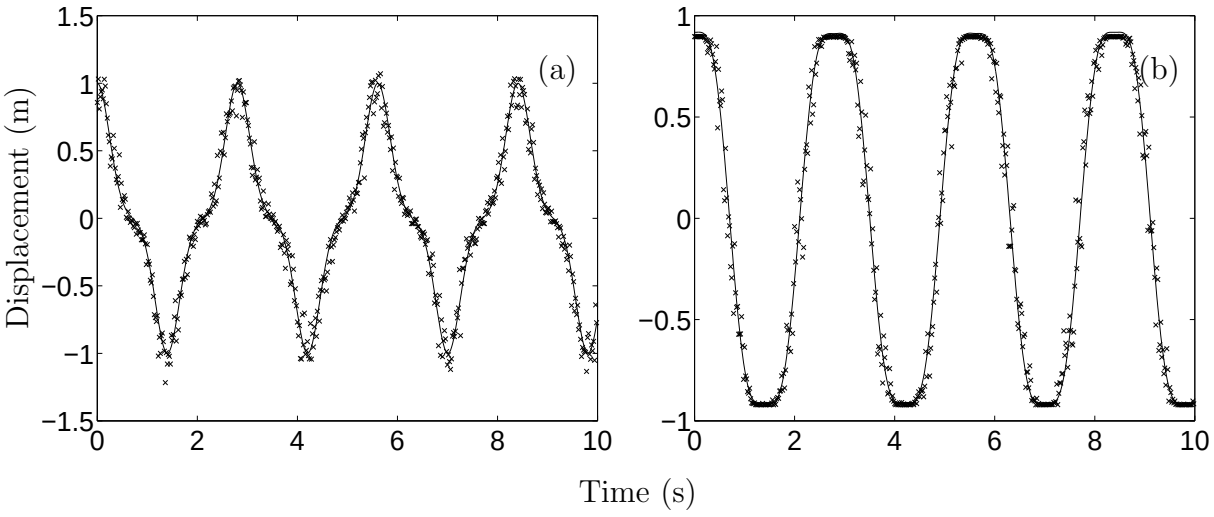


Figure 4.22: Comparison of the time series: ×, NLPCA approximation (2 neurons); ——, NNM motion (0 % noise).

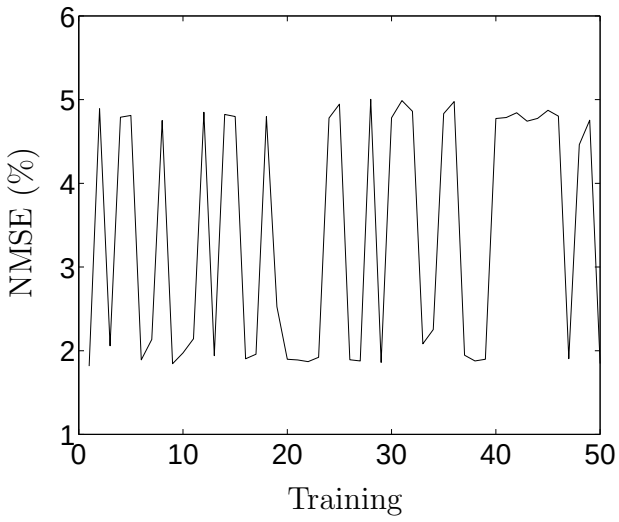


Figure 4.23: NMSE vs. training (3 neurons).

Bias/variance trade-off

The sensitivity of the results to the particular choice of the training set is also investigated. The 400 samples are shuffled to form 10 different training sets, and NLPCA is applied to each one of these data sets. It turns out that all the approximations are almost identical to that represented in Figure 4.18(a). This enables us to conclude that NLPCA provides excellent results in terms of the bias/variance trade-off [equation (4.14)]. Firstly, the data reconstruction given by the neural network is almost equal to the initial NNM motion, i.e., the bias is minimised. Secondly, the approximation is not sensitive to the particular data set used to train the network, i.e., the variance is reduced.

Minimisation of the NMSE

Finally, it is worth pointing out that the minimisation of the NMSE often ends up in one of the local minima. This is because the training of the neural network is a high-dimensional non-linear optimisation problem over the weights. This problem has been addressed by inspecting the minimisation results from an ensemble of neural networks with different initial weights. In this example, for a given network architecture, 50 trainings are carried out, and the best result in terms of the NMSE is selected among these solutions. Figure 4.23 shows the different NMSEs obtained for the 50 trainings of the network with 2 neurons. The NMSE mainly oscillates between two values that are around 2 % and 5 %. The former value corresponds to the actual solution of the problem while the latter is the value that would be obtained using the POD.

To assess the robustness of the solution, a statistical indicator called normalised mean square distance (NMSD) has been introduced in reference [130]. It is a quantitative measure of the difference between two NLPCA approximations $\hat{\mathbf{x}}^{(a)}(t)$ and $\hat{\mathbf{x}}^{(b)}(t)$ to $\mathbf{x}(t)$:

$$NMSD = \frac{E[\|\hat{\mathbf{x}}^{(a)} - \hat{\mathbf{x}}^{(b)}\|^2]}{E[\|\mathbf{x}\|^2]} \tag{4.29}$$

Experiments with this indicator reveal that two approximations are indistinguishable if their NMSD is about 2 % or less. For NMSD below 20 %, the approximations have the same gross features but may differ in details. When the NMSD between two NLPCA approximations exceeds about 40 %, they differ considerably. Despite that the NMSD lacks a theoretical understanding, it can serve as a useful measure of the distance between NLPCA approximations to a data set.

The NMSD is computed between the best three approximations out of the 50 given by NLPCA, and values equal to 0.06 %, 0.08 % and 0.11 % are found. The conclusion is that the approximations are identical and that the solution provided by NLPCA is robust.

4.4.3 Vector quantisation principal component analysis

This section presents the results obtained by performing VQPCA on the data set of Figure 4.15(b). The first step is to determine the number of regions using the "1 standard error rule". It should be noted that for a fixed number of regions, the VQPCA algorithm is run with different initialisations of the codebook. The approximation with the lowest NMSE is then selected. The number of regions is varied from 1 to 20, and the result is shown in Figure 4.24. The error attains its minimum for 20 regions. It can also be seen that the error for a smaller number of regions is always larger than the minimum error plus its standard deviation which tends to indicate that the number of regions should still be increased.

Figure 4.25 depicts the evolution of the NMSE with the number of adjustable parameters for VQPCA and NLPCA algorithms. Apart from a small number of parameters, the approximation given by NLPCA is always less accurate than the one computed by VQPCA. In addition, the NLPCA error remains constant although the number of parameters still increases. This is mainly due to the fact that the mapping and demapping projections must be continuous functions which limits the type of approximations that can be learnt. This restriction is not encountered with VQPCA, and the method can decrease the error more drastically. In terms of the bias/variance trade-off, it can be stated that the approximations given by VQPCA have a smaller bias but a larger variance. Conversely, the NLPCA approximations have a smaller variance but a larger bias.

This example shows the superiority of VQPCA over NLPCA from the point of view of data compression. However, it is again emphasised that the purpose of this study is to define features that give useful insight into the system dynamics. From Figure 4.26(a) that represents the VQPCA mode referred to as vector quantisation proper orthogonal mode (VQPOM) obtained using 20 regions, it appears that this mode has no clear physical interpretation unlike the feature given by NLPCA. In addition, the VQPOM is highly sensitive to the training set used, as shown in Figure 4.26(b) that represents the mode obtained with another training set.

From these results, it turns out that the number of regions should not be increased as suggested by the "1 standard error rule". It should rather be decreased drastically in

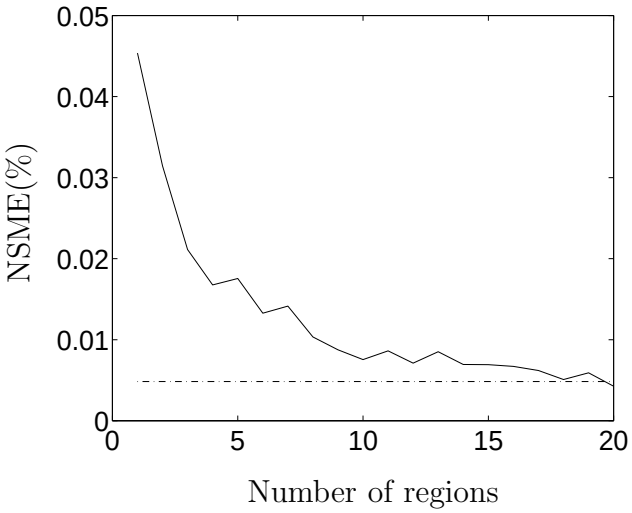


Figure 4.24: 1 standard error rule.

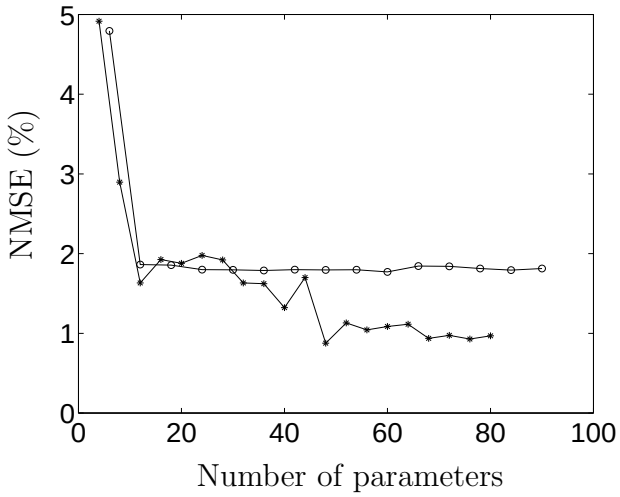


Figure 4.25: NMSE vs. the number of parameters: $-$ o $-$, NLPCA; $-$ $*$ $-$ VQPCA.

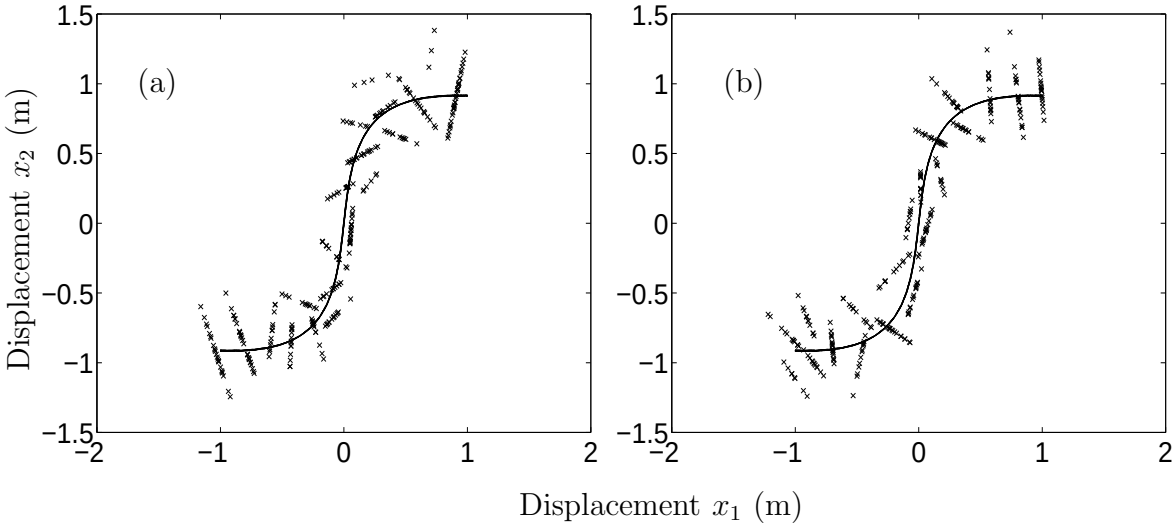


Figure 4.26: Comparison of the modes: $\times$, VQPOM (20 regions); ——, synchronous NNM. (a) Training set 1; (b) training set 2.

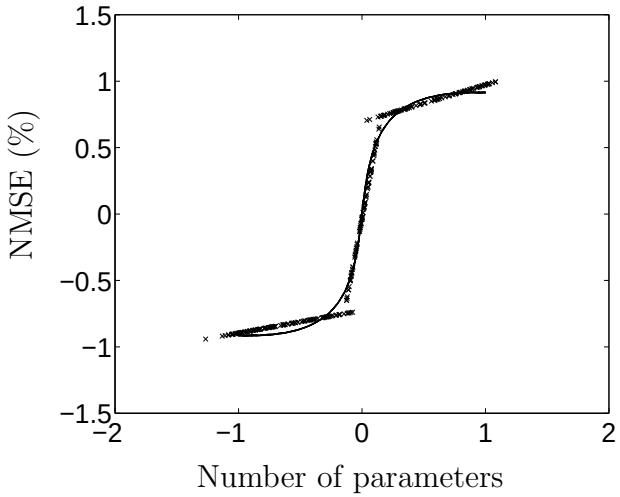


Figure 4.27: Comparison of the modes: ×, VQPOM (3 regions); ——, synchronous NNM.

order to reduce the variance of the VQPCA approximation and to avoid modelling the noise present in the system response. For instance, if the reconstruction is realised using 3 regions, the VQPOM provides a close approximation to the initial NNM motion. This is illustrated in Figure 4.27. In this case, the NMSE is equal to 2.21 %, and the mode explains 97.79 % of the variance. It has also been verified that the mode does no longer depend on the particular data set used for the training phase.

4.5 Concluding remarks

The aim of this chapter was to define features that give useful insight into the behaviour of a non-linear structure. With this aim in mind, two non-linear generalisations of the POD have been investigated.

A special attention has first been devoted to a global generalisation of the POD. Auto-associative neural networks have been exploited as a means to perform a NLPCA analysis. The inherent ability of such networks to capture non-linear correlation makes NLPCA superior to its linear counterpart.

For the purpose of model updating, we have focused on a "modal" NLPCA analysis that involves networks with a single neuron in the bottleneck layer. In this context, two features may be of great interest. The NLPOMs draw a curve in the data space that is optimal in terms of the reconstruction error. They may thus be regarded as a non-linear generalisation of the POMs and facilitate the visual representation. They might also be useful for the formulation of the objective function because they are more representative of the system dynamics and more sensitive to the non-linearity than the POMs. However, the NLPCs are preferred for this purpose because they contain the same information as the NLPOMs but in a smaller space.

It is emphasised that NLPCA has also disadvantages compared to the POD:

- NLPCA does not admit an analytical solution and its implementation requires the use of neural networks with multiple sigmoidal hidden layers. These networks have a large number of parameters even for moderate input dimensionality. Therefore, NLPCA is difficult to train and tends to require several orders more computation time than the POD.
- Training the neural network is a high-dimensional non-linear optimisation problem over the weights. Usually, the minimisation of the error ends up in one of the local minima, unlike the globally optimal solution returned by the POD. This problem can be partially addressed by inspecting the minimisation results from an ensemble of feed-forward neural networks with different initial weights.
- Although NLPCA provides interesting features, it does not have the theoretical foundations of the NNM theory. NLPCA thus lacks a well-developed and established theory and is mostly based on heuristic rules and general properties of neural networks (e.g., universal approximation).

In a second step, a local alternative to the POD denoted as VQPCA has been analysed. From the point of view of data compression, VQPCA overcomes NLPCA but seems to be much more sensitive to the noise present in the data. A lot of care should be taken when determining of the number of cells in this algorithm to avoid overfitting.

Chapter 5

Parameter Estimation using Optimisation Algorithms

5.1 Introduction

Feature extraction plays a central role in model updating of non-linear structures. This is the reason why the previous two chapters have suggested features for which the non-linear behaviour is enhanced.

Another crucial issue that still needs to be explored is the capability of optimisation methods to accurately locate the global minimum of the objective function. Indeed, in non-linear structural dynamics, the initial finite element model cannot generally assumed to be near the actual model because a priori knowledge about the non-linearity is limited. In that case, a local optimiser is likely to be trapped in a local optimum resulting in a poor solution.

The purpose of this chapter is to assess the performance of a global optimisation method called differential evolution. This method belongs to the class of evolutionary computation. Another means of tackling the minimisation of the objective function is also presented. It exploits the conditioned reverse path (CRP) method in order to alleviate our lack of a priori knowledge about the non-linearity.

Finally, a model updating strategy of non-linear vibrating structures based upon the developments made in the present chapter and in chapters 3 and 4 is proposed.

5.2 Differential evolution

The objective function may possess numerous local minima if the non-linear coefficient is allowed to vary in a wide range of amplitudes and/or the number of updating parameters is high. In this context, it is meaningful to inquire whether global optimisation algorithms may be of substantial help.

This section aims at presenting the differential evolution algorithm. The first appli-

cation of this algorithm in structural dynamics was realised for identification of hysteresis models and engine hydraulic mounts [103, 104]. The results obtained were very encouraging, and this is the reason why differential evolution is considered here. Another interesting feature is that all operations can be performed on real floating point numbers rather than strings of zeros and ones.

5.2.1 Evolutionary computation

Evolutionary computation techniques abstract evolutionary principles into algorithms that may be used to search for optimal solutions to a problem. For a search space with only a small number of possible solutions, all the solutions can be examined in a reasonable amount of time and the optimal solution is found. This exhaustive search quickly becomes impractical as the search space grows in size. The key aspect distinguishing an evolutionary search algorithm from such exhaustive search is that it is based on a population. All evolutionary algorithms share a common conceptual base of simulating the evolution of successive generations of individuals, and they perform an efficient directed search.

More precisely, evolutionary algorithms alter a population of individuals representing the solutions to a problem using operators derived from the mechanism of natural selection. These operators include mutation and recombination to alter individuals, and selection operators to choose the best individuals from the population. Each individual is evaluated using a fitness function that is specific to the problem being solved. Based upon their fitness values, a number of individuals are chosen to be parents. New individuals, i.e., children, are produced from these parents through recombination, that exchanges information between parents, and mutation, that further perturbs the children. The fitness values of the children are determined and the survival step decides who survives in the population. These steps are then repeated iteratively until a solution that satisfies a termination criterion is found. Although simplistic from a biologist's viewpoint, these algorithms are sufficiently complex to provide robust and powerful adaptive search mechanisms. A typical evolutionary algorithm is outlined in Figure 5.1.

A wide variety of algorithms has been proposed in the literature. They vary in the representation of individuals (from bit-strings to real-valued vectors), in the selection mechanisms they consider (roulette wheel selection, random selection, and so forth), and in the relative importance of recombination and mutation. Actually, there are three main streams of research into evolutionary methods, namely "evolutionary programming" [51], "evolution strategies" [145] and "genetic algorithms" [76]. The purpose, however, is not to do a comprehensive literature review of evolutionary algorithms. Considerably more detailed discussion is available in reference [162].

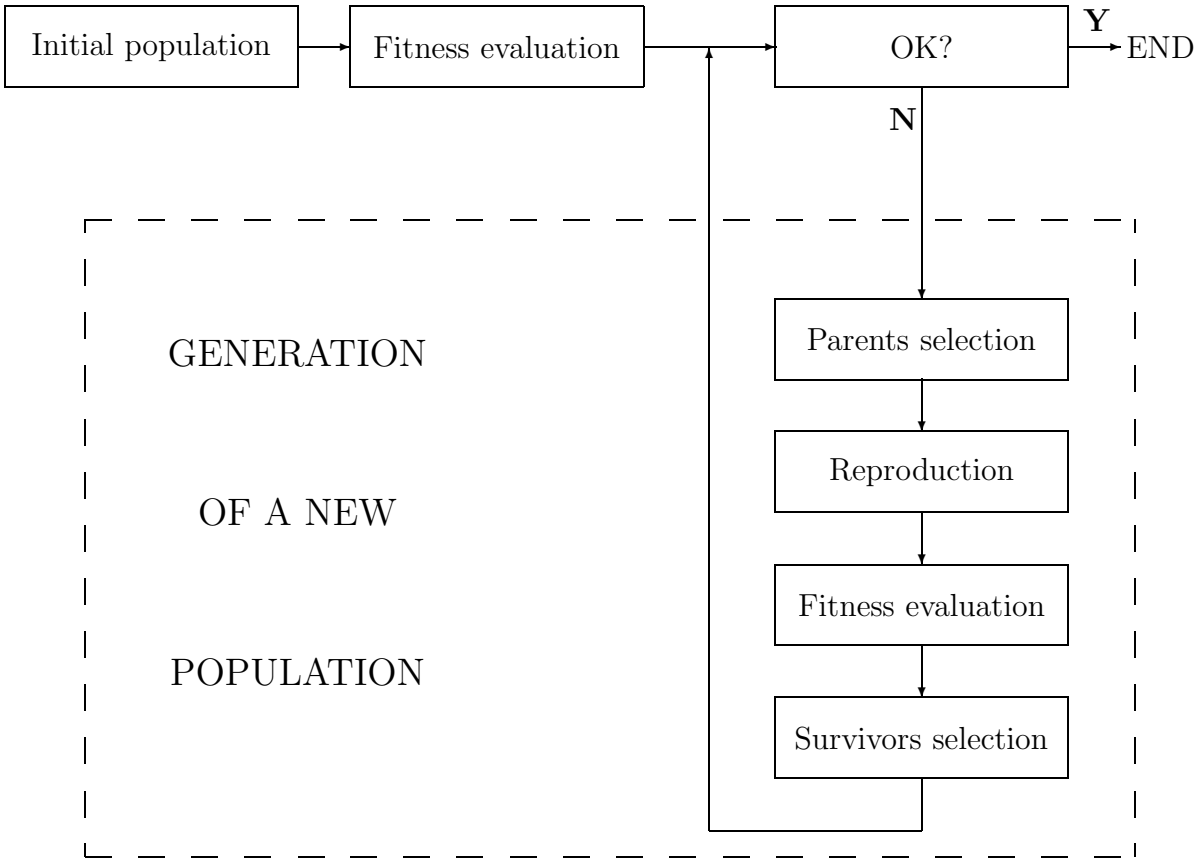


Figure 5.1: A typical evolutionary algorithm

5.2.2 Description of the algorithm

The differential evolution algorithm was introduced in the mid 1990s by Storn and Price [165, 166]. The schematic representation of this algorithm is illustrated in Figure 5.2.

The algorithm starts with an initial population of individuals for the parameter vector to be optimised, and interprets the value of the objective function at each of these individuals as a measure of their fitness. These individuals are chosen randomly, although a priori knowledge can be used to bias the search. It is emphasised that, since the minimisation of an objective function is considered, an individual survives if its fitness is lower than the one of the other individuals. Figure 5.2 involves a population of eight individuals composed of three variables, called genes, A, B, and C, and associated with a fitness value (e.g., 24 for the first individual).

The fitness of an individual in the population, called 'target vector' P_t, is compared to the fitness of a child, called 'trial vector' P_{tr}. The child is generated according to the following procedure. An individual P_p, different from P_t, is chosen randomly in the population. This individual undergoes a mutation that employs the difference of two

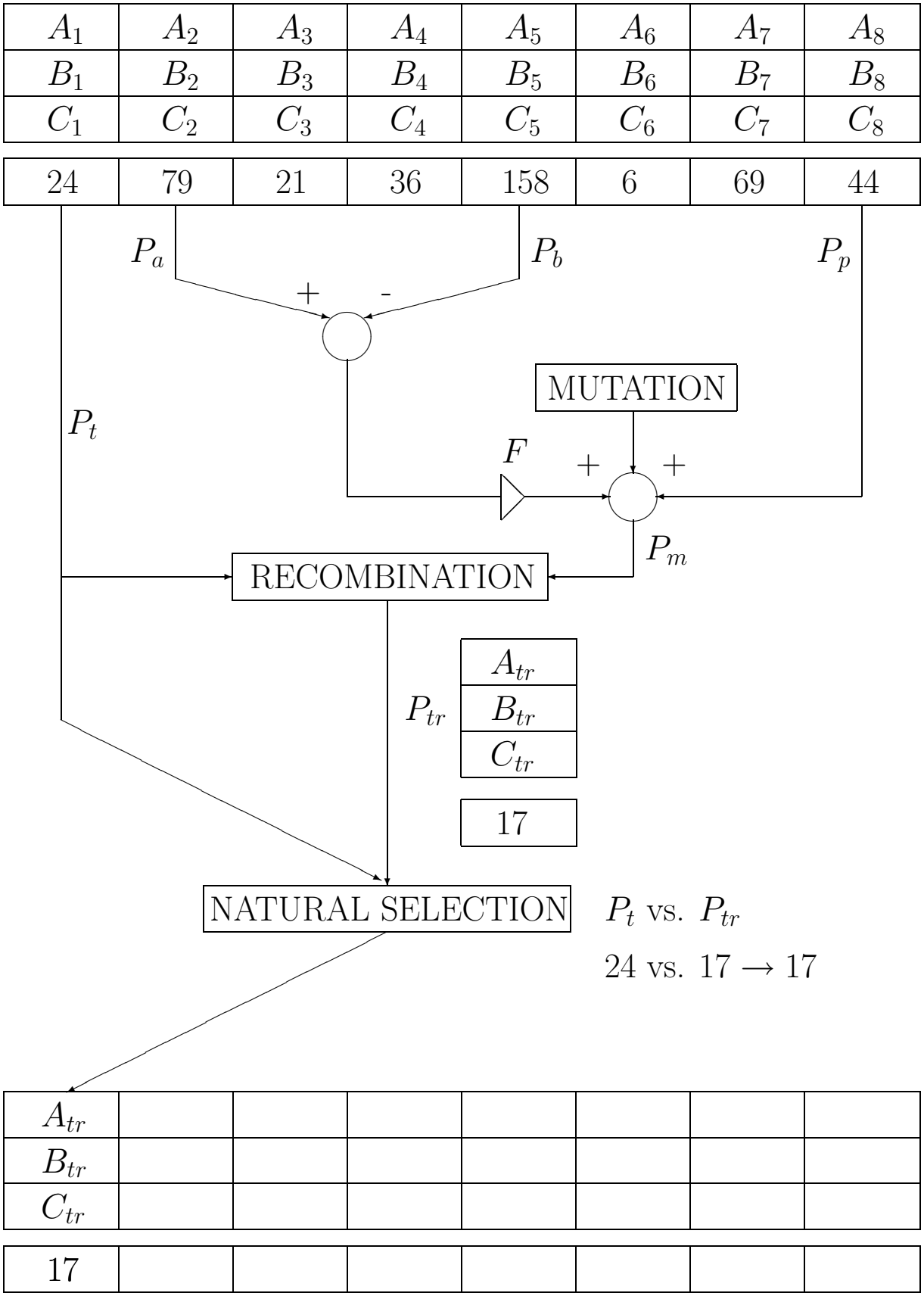


Figure 5.2: Differential evolution.

distinct individuals P_a and P_b, scaled by a constant parameter F, called scaling factor, as the source of random variation. P_a and P_b must be different from P_t and P_p. The resulting individual is called 'mutated vector' P_m.

The next step is the recombination between the mutated vector P_m and the target vector P_t in order to create the trial vector P_{tr}. The mutated and target vectors are the parents and the trial vector is the child. It remains to explain how the child gets the genes from his parents. For each gene, a random number between 0 and 1 is chosen. If this number is lower than a constant parameter referred to as the crossover ratio CR, the child then takes the gene from the target vector P_t, otherwise the child takes the gene from the mutated vector P_m. Once the trial vector is computed, its fitness is compared to the one of the target vector. The vector with the lowest fitness survives and belongs to the next generation.

The procedure is repeated for a new target vector. Once all the individuals of a generation have been chosen as a target vector, the algorithm is restarted with the next generation. The algorithm is terminated when the fitness of an individual is lower than a pre-defined threshold or when the maximum number of iterations is exceeded.

It should be noted that differential evolution lends itself very well to parallel computation. In this case, all the individuals of a generation become target vectors simultaneously.

Differential evolution differs widely from the other evolutionary algorithms essentially for four reasons:

- Usually, the parents are individuals from the actual population while here the mutated vector results from a mutation of an individual in the population;
- Mutation appears before recombination and thus involves a parent instead of a child;
- The perturbation is not a random number but the difference of two individuals multiplied by the scaling factor F;
- The crossover ratio allows to give more or less importance to a parent in the sense that a high (low) crossover ratio increases the chance to get the gene from the target vector (mutated vector).

Exploration versus exploitation

There are three factors that determine the evolutionary process of differential evolution, namely the population size, the scaling factor F and the crossover ratio CR. The main problem is finding a set of control parameters that optimally balances "exploration" and "exploitation". Exploration and exploitation of solutions in the search space are distinctive characteristics of an evolutionary algorithm, and are responsible for the success or failure of the search process.

On the one hand, one could imagine a population where children are an exact replica of their parents. The fittest individual would be expected to quickly dominate the popu-

lation. In this case, exploitation (exploration) is extreme (minimal), and it may lead to premature convergence. On the other hand, children could have a minimal resemblance with their parents which maximises exploration to the detriment of exploitation. The probability of finding the global minimum is high but the search is ineffective.

Ideally, the fittest individuals should be kept in the population while preventing the loss of diversity. The dilemma is that neither exploitation nor exploration can be pursued exclusively without failing at the task. Furthermore, the relative importance of both operators is largely application dependent. It is thus meaningful to understand the influence of the three factors on the exploration/exploitation trade-off:

- Population size: a population with a few (a lot of) individuals should favour exploitation (exploration). The convergence is quick (low) but the chance to find a local minimum is relatively high (low).
- Crossover ratio CR: with a high (low) value of the crossover ratio, the probability to get the gene from the target vector is high (low), i.e., children tends to be an exact replica of (different from) their parents. Exploitation (exploration) tends to dominate the process.
- Scaling factor F: since it operates during mutation, a high (low) value of the scaling factor gives great (little) diversity and thus increases exploration (exploitation).

5.2.3 An illustrative example

In this section, the performance of differential evolution is highlighted using Griewangk's function. A Griewangk's function with n variables is described by the following equation:

$$f(x) = \sum_{j=1}^{n} \frac{x_j^2}{4000} - \prod_{j=1}^{n} \cos(\frac{x_j}{\sqrt{j}}) + 1 \qquad \text{with } x_j \in [-400, 400] \tag{5.1}$$

It is known as a fairly difficult problem to solve due to the presence of numerous local minima. Figure 5.3 represents the function in the case of two variables x_1 and x_2.

A Griewangk's function with ten variables is considered here. The global minimum is equal to zero. The algorithm stops when at least one individual in the population has a fitness value below 10^{-6} or when the number of generations exceeds 2000. Several values of the scaling factor and of the crossover ratio are considered. For each pair of values, a hundred optimisations are carried out. The results are listed in Table 5.1.

This table illustrates that the results are strongly influenced by the values chosen for the control parameters. For $F = 0.1$ and $CR = 0.9$, none of the hundred runs can retrieve the global minimum while for $F = 0.5$ and $CR = 0.9$, only fourteen optimisations converge to a local minimum. If the population size is increased up to 25, for $F = 0.5$ and $CR = 0.7$, the hundred optimisations retrieve the global minimum.

Another interesting thing to note is that the results corroborate the observations in the previous section. For instance, when $F = 0.5$ and $CR = 0.9$, the number of iterations is

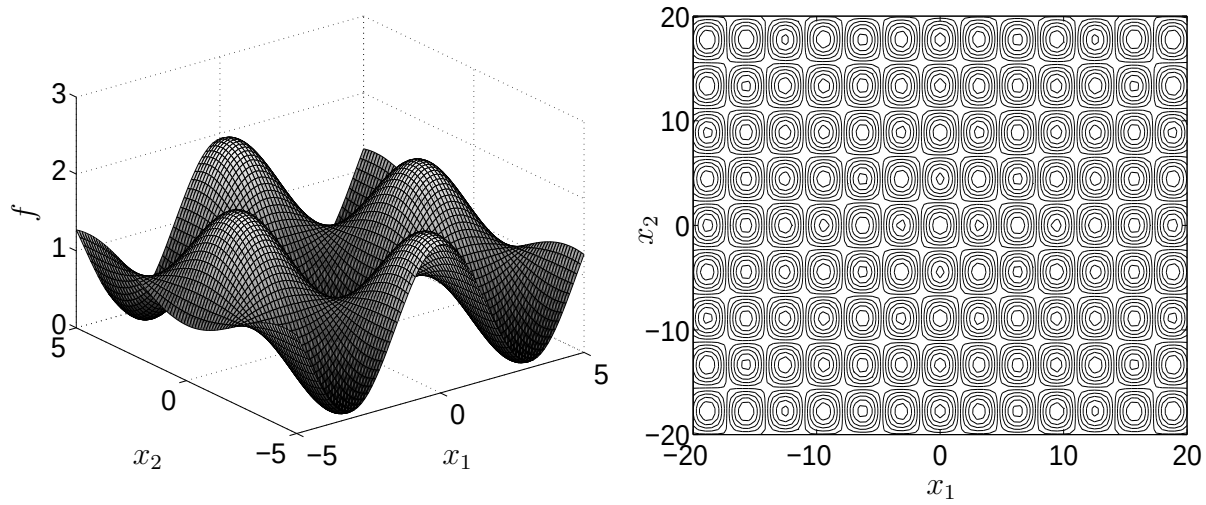


Figure 5.3: Griewangk's function (two variables)

Population size	Scaling factor F	Crossover ratio CR	Number of iterations	Global minimum reached
15	0.9	0.1	—	0/100
15	0.5	0.1	—	0/100
15	0.9	0.3	—	0/100
15	0.1	0.9	—	0/100
15	0.5	0.9	14881	86/100
15	0.9	0.9	22893	80/100
15	0.5	0.7	8678	80/100
25	0.5	0.7	14340	100/100

Table 5.1: Optimisation of Griewangk's function

equal to 14881. If F is increased up to 0.9, the exploration is greater than for $F = 0.5$. The number of iterations should then increase which is indeed the case (22893). Surprisingly enough, the performance of the algorithm decreases at the same time (80/100 instead of 86/100). The reason is that the algorithm stops when the number of generations exceeds 2000, which is maybe not enough in this particular case.

For the sake of comparison, two other algorithms, namely Nelder and Mead approach [26] and a pure random algorithm, have also been applied to Griewangk's function. Nelder and Mead approach only requires objective function evaluations and is known to perform well locally. Again, a hundred optimisations have been carried out. Neither of these algorithms could retrieve the global minimum.

5.2.4 Discussion

From the previous results, it appears that differential evolution is efficient and robust provided that the control parameters are carefully tuned. Differential evolution is also simple to understand and to implement, and a key advantage is that it only requires objective function evaluations. It can then be applied on problems where the gradient and/or Hessian is discontinuous which is not the case with classic methods.

The essential difficulty lies in the fact that numerous objective function evaluations are necessary to achieve a good convergence. This is a critical issue when the objective function evaluation requires significant CPU time which may be the case in a non-linear model updating context. Indeed, a numerical simulation using Newmark's algorithm may take hundreds of seconds depending on the frequency range of interest and the simulation time. Since differential evolution lends itself very well to parallel computation, a means of accelerating the convergence would be to execute it on a parallel machine or a network of computers.

It should however be kept in mind that the crucial decision and difficulty is what to optimise and not the optimisation method used. This is the reason why we believe that feature extraction plays a prominent role in model updating of non-linear vibrating structures.

5.3 Coupling local optimisers with the conditioned reverse path method

When there is little prior knowledge about the non-linearity, the use of a global optimisation algorithm is recommended. However, as mentioned in the previous section, differential evolution requires numerous objective function evaluations. The situation may be improved by adding constraints on the updating parameters and especially on the non-linear term.

The determination of these constraints still needs to be addressed. Our suggestion is to exploit, during a preliminary step, the CRP method studied in chapter 2 to compute the constraints. This choice is motivated essentially by two reasons:

- The work realised in chapter 2 has pointed out that the CRP formulation has the ability to deal with continuous multi-degree-of-freedom (MDOF) structures. In addition, the method gives an estimation of the non-linearity without building a finite element model of the structure. Problems such as modelling of boundary conditions or damping are thus avoided.

- The accuracy of the identification can be assessed using the cumulative coherence function. This represents an interesting piece of information because the higher the cumulative coherence is, the tighter the constraints can be.

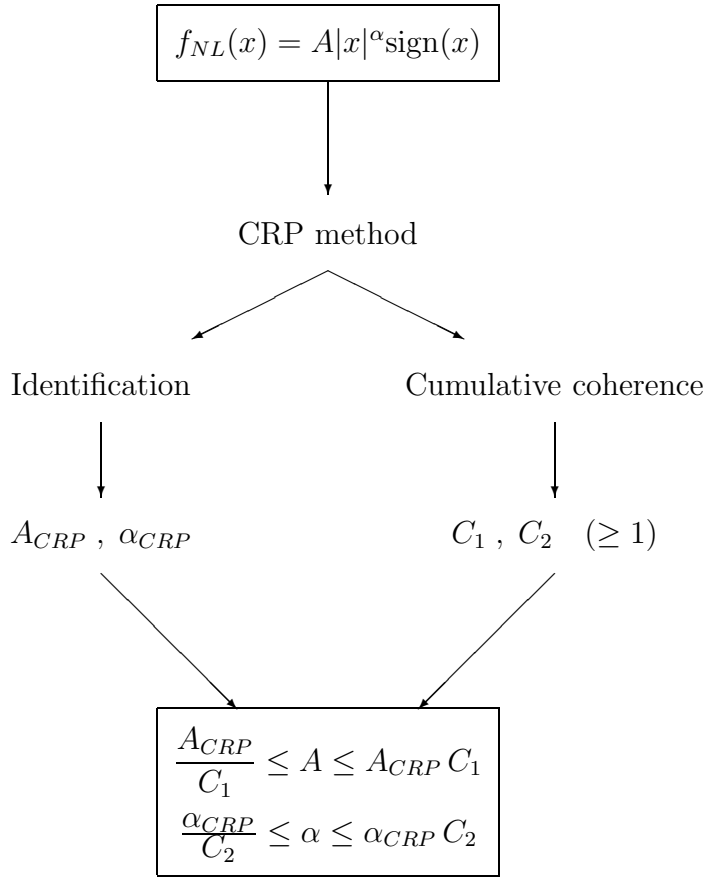


Figure 5.4: Estimation of the constraints using the CRP method.

In the case of an accurate identification, tight constraints on the non-linearity can be added which dramatically restricts the search space. The feasible set of solutions may then exclude many of the local minima, and it may be comparatively easy to pick the global minimum from those that remain. In this context, the use of a local optimiser seems to be appropriate.

The proposed procedure is depicted in Figure 5.4 for a non-linearity modelled using the following functional form $f_{NL}(x) = A|x|^{\alpha}\text{sign}(x)$. The first step is to compute the values of the coefficient A and the exponent α using the CRP method, i.e., A_{CRP} and α_{CRP}. From the value of the cumulative coherence function, constants C_1 and C_2 are then chosen, and the constraints are written as follows:

$$\frac{A_{CRP}}{C_1} \leq A \leq A_{CRP}\, C_1 \quad \text{and} \quad \frac{\alpha_{CRP}}{C_2} \leq \alpha \leq \alpha_{CRP}\, C_2 \tag{5.2}$$

5.4 A model updating strategy of non-linear vibrating structures

Considerable effort has been spent in order to address the two following problems encountered in non-linear model updating:

- Feature extraction (chapters 3 and 4);
- Minimisation of the objective function (present chapter).

Based upon these developments, we propose the model updating strategy illustrated in Figure 5.5 and summarised in what follows.

1. Experimental measurements and finite element modelling

Let us consider a real structure for which vibration tests have been performed using n sensors. A set of m samples of time-domain responses has been collected (e.g., acceleration data) and an $(n \times m)$ response matrix $\mathbf{X}$ is formed:

$$\mathbf{X} = [\mathbf{x}(t_1) \ \cdots \ \mathbf{x}(t_m)] = \begin{bmatrix} x_1(t_1) & \cdots & x_1(t_m) \\ \cdots & \cdots & \cdots \\ x_n(t_1) & \cdots & x_n(t_m) \end{bmatrix} \tag{5.3}$$

From the knowledge of the geometric and mechanical properties of the structure, a finite element model can be created. Assuming similar excitation conditions, the structural response can then be predicted using time-integration methods (e.g., Newmark's method [58]). A matrix $\hat{\mathbf{X}}$ is obtained and differs generally from matrix $\mathbf{X}$ due to the presence of parameter errors and imperfect boundary conditions for instance.

2. Feature extraction and correlation study

The next step is the correlation step. As explained in the introductory chapter, the brute comparison between matrices $\mathbf{X}$ and $\hat{\mathbf{X}}$ is to be avoided, and the comparison of features is preferred. Accordingly, the proper orthogonal decomposition (POD), vector quantisation principal component analysis (VQPCA) and non-linear principal component analysis (NLPCA) play a crucial role. Indeed, the features they provide should give useful insight into the dynamics of the structure in order to decide whether the initial model is accurate enough for the purposes for which it is intended.

3. Parameter selection

If correlation is not satisfactory, the finite element model needs to be corrected. The correction of the model begins with the selection of the updating parameters $\mathbf{p}$. In this thesis, it is assumed that some a priori knowledge about the non-linearity and the erroneous linear parameters is available. This should not be considered as a too severe

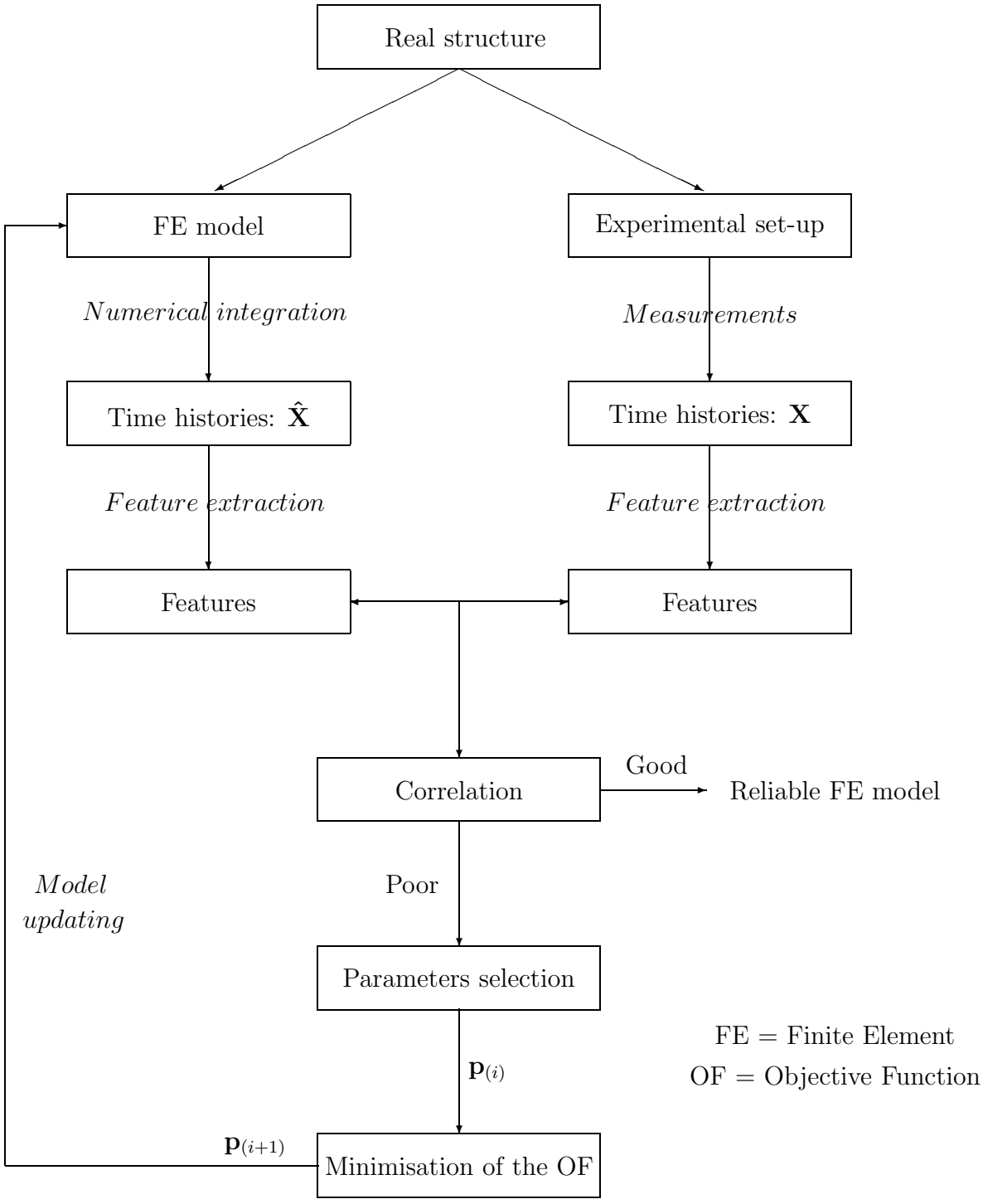


Figure 5.5: Model updating sequence of non-linear systems.

restriction because engineering judgment and system-specific knowledge will always be inherent in a model updating process.

4. Model updating

Finally, the finite element model is updated by minimising an objective function expressing the discrepancies between the features extracted from the real structure and those given by the model:

- Objective function for the POD:
  $\frac{-1}{k}\sum_{i=1}^{k} \text{MAC}(POM_{\mathbf{t}(i)}, POM_{\mathbf{s}(i)}) \in [-1, 0]$
- Objective function for VQPCA: $\frac{-1}{k}\sum_{i=1}^{k} \text{MAC}(VQPOM_{\mathbf{t}(i)}, VQPOM_{\mathbf{s}(i)}) \in [-1, 0]$
- Objective function for NLPCA: $\sum_{i=1}^{k} \text{NMSE}(NLPC_{\mathbf{t}(i)}, NLPC_{\mathbf{s}(i)}) \geq 0$

where $\ldots_{\mathbf{t}}$ and $\ldots_{\mathbf{s}}$ refer to the test and simulation results respectively;
k is the number of modes included in the objective function;
MAC is the modal assurance criterion [see equation (3)];
NMSE is the normalised mean square error [see equation (4.25)].

An important remark concerns the objective function for NLPCA. Ideally, a neural network (or several if more than one mode is desired) should be trained for each new value of the updating parameters, and the non-linear principal component obtained through this network should be compared to that of the experimental results. However, training a neural network at each iteration of the updating process is computationally intensive.

Our suggestion is to train a neural network (or several if more than one mode is desired) using the experimental data and to use the trained network throughout the updating procedure. By doing that, a curve that is optimal for the characterisation of the experimental data is computed, and the objective function expresses the difference between the projection of the simulated data and of the experimental data onto this curve.

The objective function can be minimised using the differential evolution algorithm or a local optimiser (sequential quadratic programming method in this thesis) coupled with the CRP method. In the latter case, the derivatives of the objective function are computed using finite differences. Again, feature extraction plays a key role. Features should guarantee a sufficient sensitivity of the objective function with respect to the updating parameters. In addition, the objective function should ideally show a smooth topological search space with a small number of local minima around the global minimum.

Some differences with the procedure used in the case of a linear system represented in Figure 2 of the introductory chapter should be noted:

- The features can no longer be computed directly from the structural matrices. It is necessary to perform a numerical integration of the equations of motion at each iteration which greatly increases the computational cost of the process.
- Expansion and/or reduction techniques are not implemented in order to solve the mesh incompatibility. In this thesis, the model is defined in order to have a node at the location of each sensor. A point-to-point comparison between the measured and predicted features is performed at the measured partition of the degree-of-freedom (DOF). The DOF of the numerical model that are not measured in practice are thus ignored.
- The selection of the updating parameters only relies on engineering judgment. Sensitivity analysis is also widely used in the literature but is not considered here. Our motivation is that the parameters with the highest sensitivity do not necessarily include those that present inaccuracies.

Chapter 6

Model Updating of a Three-dimensional Portal Frame

6.1 Introduction

The objective of this chapter is to assess the performance of the model updating strategy described in chapter 5 using a three-dimensional portal frame with simulated data. More precisely, it is intended to:

- Determine the respective merits of the features extracted from the proper orthogonal decomposition (POD), vector quantisation principal component analysis (VQPCA) and non-linear principal component analysis (NLPCA);
- Inquire whether the differential evolution algorithm is efficient in the particular context of model updating;
- Shed light on the potential benefits of the coupling of the conditioned reverse path (CRP) method with a local optimisation technique.

6.2 Description of the structure

The structure illustrated in Figure 6.1 is composed of beams with circular cross-sections and with the following characteristics:

$$\begin{aligned} A &= 1.13\,10^{-4}\,\mathrm{m}^2 & J_x &= 9.27\,10^{-9}\,\mathrm{m}^4 \\ I_y &= 4.64\,10^{-9}\,\mathrm{m}^4 & I_z &= 4.64\,10^{-9}\,\mathrm{m}^4 \end{aligned} \tag{6.1}$$

The material is aluminium:

$$E = 3\,10^{10}\,\mathrm{N/m}^2 \qquad \nu = 0.33 \qquad \rho = 2700\,\mathrm{kg/m}^3 \tag{6.2}$$

The column bases are clamped to the ground.

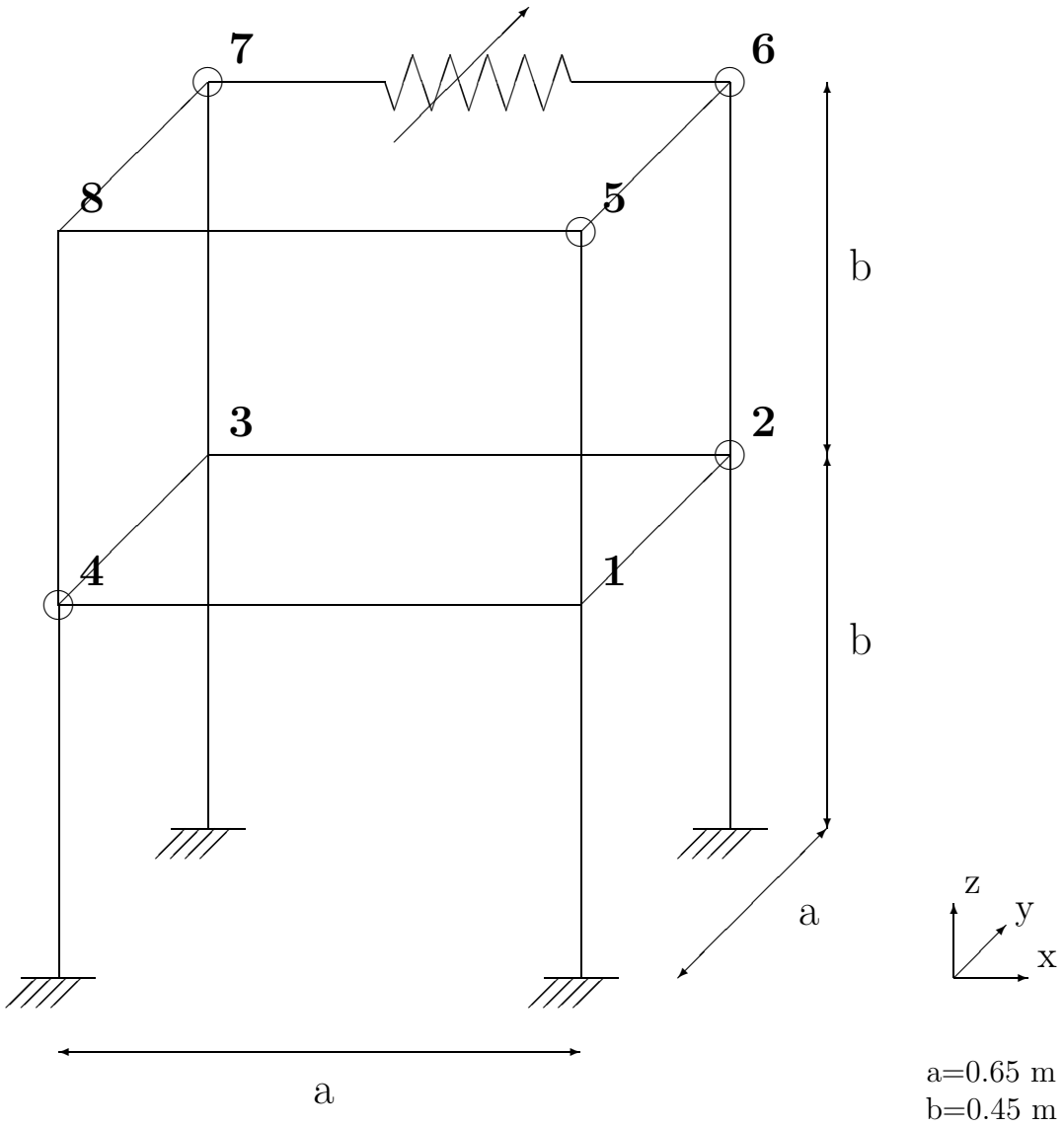


Figure 6.1: Three-dimensional portal frame (O=sensor).

DOF	Type	Node
$1 \rightarrow 3$	translation Ox,Oy,Oz	1
$4 \rightarrow 6$	rotation Ox,Oy,Oz	1
$7 \rightarrow 9$	translation Ox,Oy,Oz	2
$10 \rightarrow 12$	rotation Ox,Oy,Oz	2
$13 \rightarrow 15$	translation Ox,Oy,Oz	3
$16 \rightarrow 18$	rotation Ox,Oy,Oz	3
$19 \rightarrow 21$	translation Ox,Oy,Oz	4
$22 \rightarrow 24$	rotation Ox,Oy,Oz	4
$25 \rightarrow 27$	translation Ox,Oy,Oz	5
$28 \rightarrow 30$	rotation Ox,Oy,Oz	5
$31 \rightarrow 33$	translation Ox,Oy,Oz	6
$34 \rightarrow 36$	rotation Ox,Oy,Oz	6
$37 \rightarrow 39$	translation Ox,Oy,Oz	7
$40 \rightarrow 42$	rotation Ox,Oy,Oz	7
$43 \rightarrow 45$	translation Ox,Oy,Oz	8
$46 \rightarrow 48$	rotation Ox,Oy,Oz	8

Table 6.1: DOF numbering.

It is emphasised that no beam makes the connection between nodes 6 and 7. However, a non-linear stiffness element is added between the translational degree-of-freedom (DOF) along direction x:

$$f_{NL} = 10^7 \ |x_{31} - x_{37}|^{2.5} \operatorname{sign}(x_{31} - x_{37}) \tag{6.3}$$

where x_i refers to the displacement of the ith DOF. The DOF numbering is listed in Table 6.1.

The model updating is now achieved following the procedure detailed in chapter 5, i.e.:

1. Experimental measurements and finite element modelling (section 6.3);
2. Feature extraction and correlation study (section 6.4);
3. Parameter selection (section 6.5);
4. Model updating (sections 6.6 and 6.7). Section 6.6 assumes that a priori knowledge about the non-linearity is limited whereas section 6.7 takes the results given by the CRP method into account.

6.3 Experimental measurements and finite element modelling

Finite element modelling

The structure is modelled with three-dimensional beam elements. The model has 12 nodes, 4 of them being fixed. Each node possesses 6 DOF, i.e., 3 displacements and 3 rotations. The portal frame has thus a total of (12-4) $\times$ 6 = 48 DOF.

Two finite element models are considered in what follows. A model, referred to as "test model", is used for obtaining the "experimental data". Another finite element model, referred to as "initial model", is created without taking the non-linear behaviour into account. In addition, the initial model differs from the test model in the value of Young's modulus. A reduction of 5 % of the value of this parameter is considered ($2.85\,10^{10}$ vs. $3\,10^{10}\,\mathrm{N/m^2}$).

It should also be noted that modal damping is added in both models, i.e., the damping matrix is proportional to the stiffness matrix:

$$\mathbf{C} = 3\,10^{-5}\,\mathbf{K} \tag{6.4}$$

Experimental measurements

The structure is tested for two different excitations. A random excitation band-limited into the 0-100 Hz range is first applied. The driving point is the translational DOF along direction x of node 6, i.e., DOF 31. Secondly, an impulse force is applied at the same location.

The experimental data is obtained through the numerical simulation of the test model using Newmark's algorithm. The sampling frequency is set to 1500 Hz and 10000 Hz for the random and impact excitations respectively in order to guarantee an accurate simulation. Although the structural response is computed at each DOF, it is assumed that only nodes 2, 4, 5, 6 and 7 are instrumented with tri-axial accelerometers. There is thus a total of 5 $\times$ 3 measured DOF. The sensor placement is shown in Figure 6.1. It should also be noted that measurement noise equal to 1 % of the RMS value is introduced in all the acceleration signals.

At this stage, it should be noted that useful information can be gleaned from the inspection of frequency response functions (FRFs). The FRF at the driving point and the corresponding ordinary coherence function have been computed from the experimental data using the H_2 estimate (see section 2.2 for its definition) for three increasing amplitudes, namely 0.05, 0.5 and 5 Nrms. The plots are represented in Figure 6.2.

A first remark concerns the appearance of distortions in the FRFs at frequencies below 10 Hz. These distortions should not be attributed to some non-linear effect but rather to the noise added to the acceleration signals. Indeed, in this frequency range, the structural response is small, and is highly sensitive to the noise.

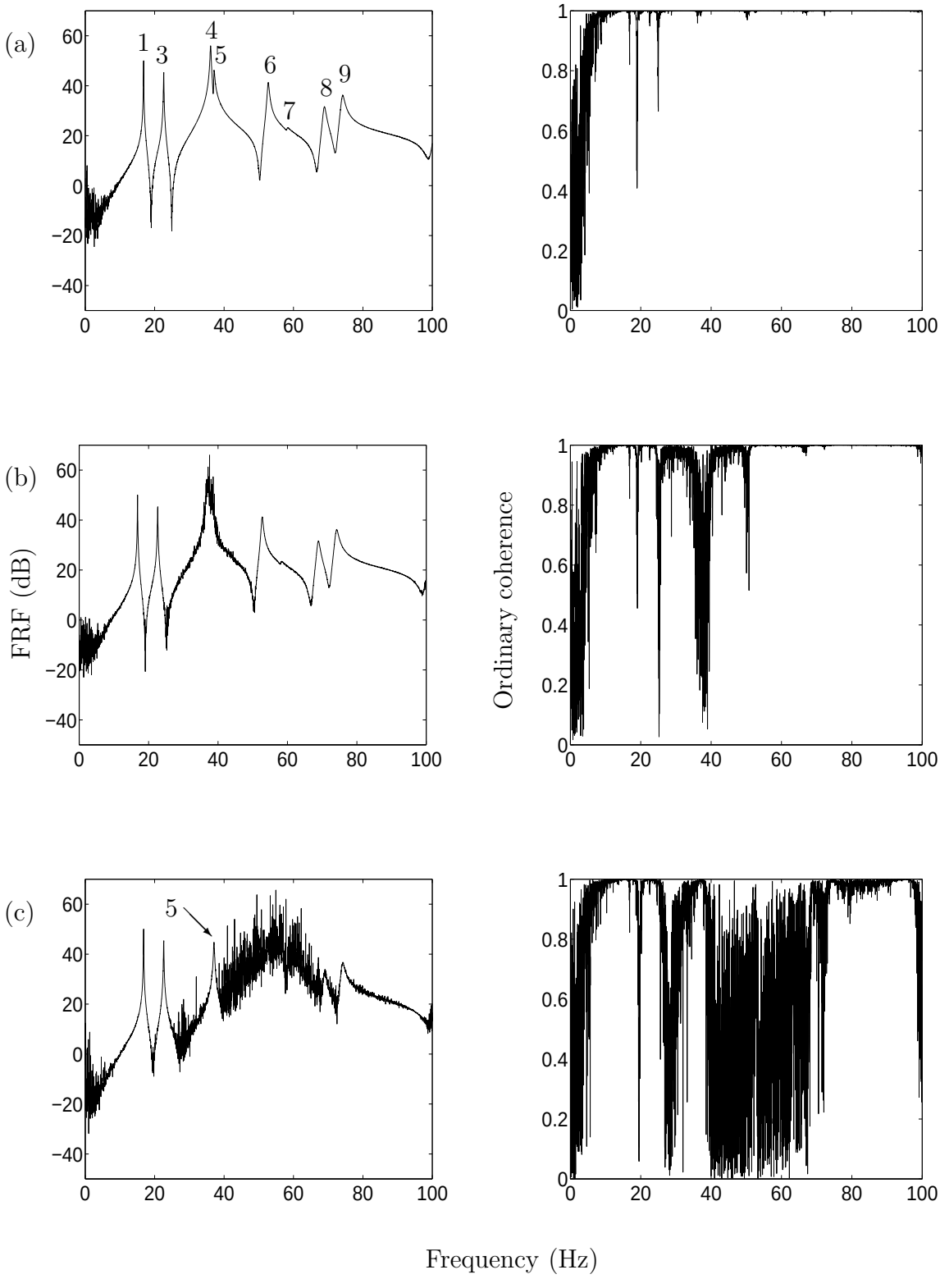
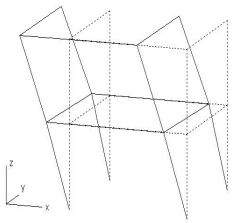
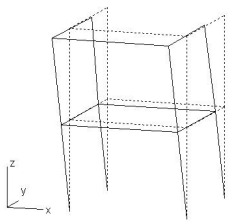


Figure 6.2: FRF at the driving point and ordinary coherence function (experimental data). (a) 0.05 Nrms; (b) 0.5 Nrms; (c) 5 Nrms.

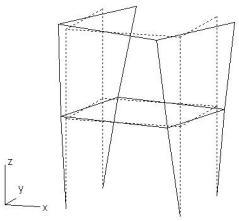
Mode 1 (16.38 Hz), bending Oxz



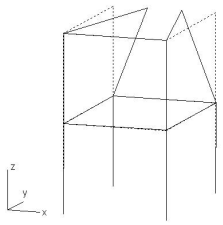
Mode 2 (17.74 Hz), bending Oyz



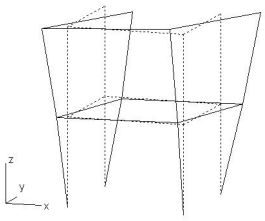
Mode 3 (22.04 Hz), torsion



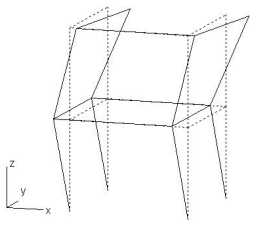
Mode 4 (35.22 Hz), local mode



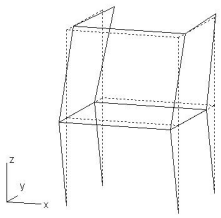
Mode 5 (36.18 Hz), torsion



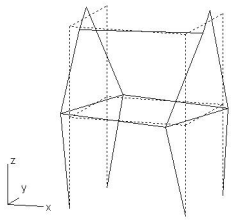
Mode 6 (51.42 Hz), bending Oxz



Mode 7 (56.76 Hz), bending Oyz



Mode 8 (67.23 Hz), torsion



Mode 9 (72.35 Hz), torsion

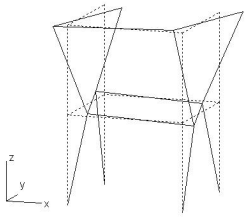


Figure 6.3: The first nine modes computed from the initial model.

At the lowest level (0.05 Nrms), apart from a few drops at resonances and anti-resonances, the ordinary coherence function is unitary which means that the structural behaviour may be assumed to be linear.

For the two remaining amplitudes, severe distortions in the FRFs together with important drops of the coherence function can be noticed. A non-linearity is thus activated when the amplitude of the applied force is increased. More precisely, it can be observed that:

- At the intermediate excitation level, the distortions mainly appear in the neighbourhood of the fourth and fifth resonance frequencies, i.e., around 40 Hz.
- At the highest level, the distortions are shifted towards higher frequencies which results from a stiffening effect of the non-linearity. A closer look at the FRF indicates that the fifth resonance frequency is not affected by the non-linearity.

From these results, it can confidently be assumed that a non-linear effect that mainly affects the fourth resonance peak is present. The non-linearity location can then be deduced from the visualisation of the fourth mode shape.

The mode shapes may be computed from the experimental FRFs. But due to the limited number of sensors, they are preferably estimated from the structural matrices of the initial model for which all the DOF are available. The first nine mode shapes are depicted in Figure 6.3. This figure clearly illustrates that the fourth mode is a local mode of the structure involving a relative translational displacement between DOF 31 and 37. This shows that the non-linearity is likely to be located between these two DOF.

6.4 Feature extraction and correlation study

The model updating of the portal frame is realised for two different excitations, namely an impulse force with an amplitude equal to 20 N and a random excitation band-limited into the 0-100 Hz range with an amplitude set to 2 Nrms. The non-linearity participation in the system response is more or less the same for these two excitations.

Now that the experimental measurements and the finite element modelling have been carried out, the next step in the updating procedure consists in correlating the features extracted from the experimental and simulated data. Three kinds of features are considered in what follows:

- The proper orthogonal modes (POMs);
- The modes given by VQPCA and referred to as vector quantisation proper orthogonal modes (VQPOMs);

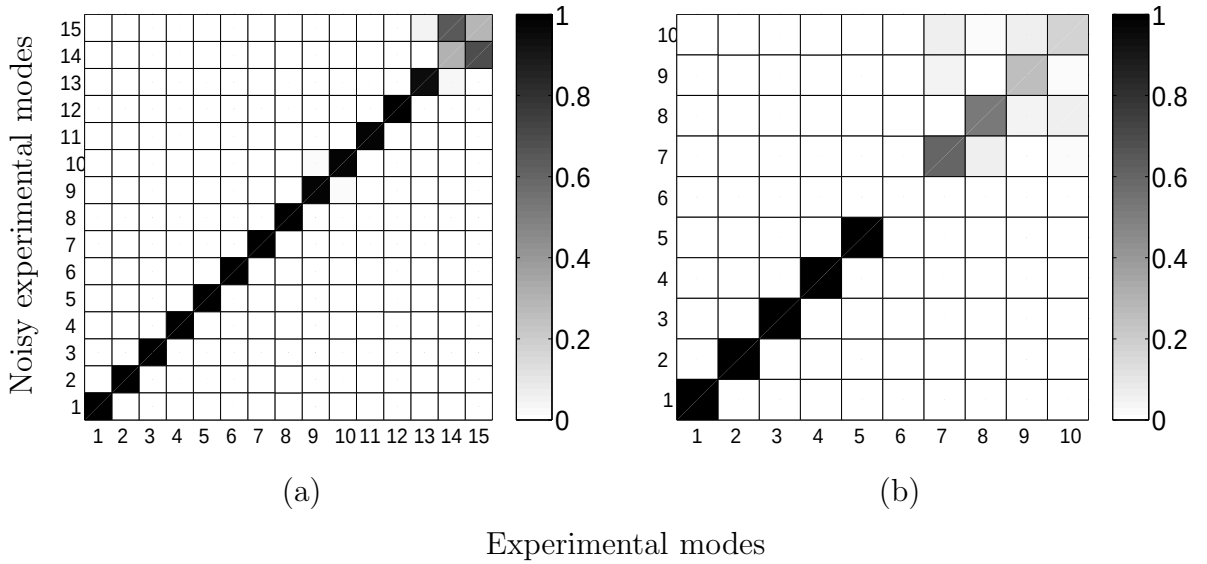


Figure 6.4: Correlation between experimental and "noisy experimental" modes (impact). (a) POMs; (b) VQPOMs (2 cells).

- The features given by NLPCA. Referring to the discussion made in section 4.2.6, a distinction is made between the outputs of the bottleneck layer, i.e., the nonlinear principal components (NLPCs) and those of the neural network, i.e., the nonlinear proper orthogonal modes (NLPOMs). They contain the same information but in different spaces. The NLPCs are useful for the correlation study and for the construction of the objective function whereas the NLPOMs are exploited for visualisation purposes.

Feature extraction

Before doing the correlation study itself, it remains to choose the modes to be included in the analysis. The usual way of determining the number of modes is based on a prescription. For instance, the first r modes that account for at least 99 % of the energy can be considered. The main drawback of this rule is that it does not take the relative importance of the noise present in the system response into account.

This is the reason why we propose to consider the following rule. In practice, by adopting rigorous testing procedures, it can confidently be assumed that the noise level is below 5 %. Let us add this percentage of noise to the experimental response. The idea is to correlate the features extracted from the experimental response with the features extracted from this data plus the noise. This enables us to study the sensitivity of the different modes to the noise and discard those that are the most affected.

Figure 6.4 represents the modal assurance criterion (MAC) matrix between the POMs and the VQPOMs of the experimental response and those of the "noisy" experimental

Number of cells	Number of modes not affected by the noise	Energy captured %
2	5	95.79
3	3	90.05
4	2	84.97
5	1	81.47

Table 6.2: Selection of the number of cells for the VQPCA algorithm (impact).

	POD (modes / energy)	VQPCA (modes / energy)	NLPCA (modes / energy)
Impact	7 / 99.29 %	5 (2 cells) / 95.79 %	3 / 87.37 %
Random	7 / 99.02 %	5 (2 cells) / 95.05 %	3 / 84.60 %

Table 6.3: Number of modes.

response for the impulse force. It should be recalled that the MAC is defined as:

$$\mathrm{MAC}(\boldsymbol{\phi}_t, \boldsymbol{\phi}_s) = \frac{(\boldsymbol{\phi}_t^T \boldsymbol{\phi}_s)^2}{(\boldsymbol{\phi}_t^T \boldsymbol{\phi}_t)(\boldsymbol{\phi}_s^T \boldsymbol{\phi}_s)} \tag{6.5}$$

where $\boldsymbol{\phi}_t$ and $\boldsymbol{\phi}_s$ represent the modes that need to be correlated (t=test, s=simulation). MAC values oscillate between 0 and 1, and a unitary value means perfect correlation. Figure 6.4(a) shows that the first thirteen POMs are well correlated which means that they are not too sensitive to the noise. However, only the first seven POMs are retained because they already capture more than 99 % of the energy.

The procedure is more complicated in the case of the VQPCA algorithm. Indeed, the illustrative example in chapter 4 has highlighted that the determination of the number of cells is a critical issue. To alleviate this drawback, the proposed procedure can be exploited. For each number of cell, the modes that are not much affected by the noise together with the energy they capture are calculated. The number of cell for which the most energy is captured is then selected. This information is given in Table 6.2 for the impact excitation. The best results are obtained for 2 cells and 5 modes in each cell [see Figure 6.4(b)]. In this case, 95.79 % of the total energy is captured.

For NLPCA, the MAC cannot be used in order to compare the NLPCs. The comparison is now performed using the normalised mean square error (NMSE):

$$\mathrm{NMSE}(NLPC_t, NLPC_s) = \frac{\mathrm{E}[\|NLPC_t - NLPC_s\|^2]}{\mathrm{E}[\|NLPC_t - E[NLPC_t]\|^2]} \tag{6.6}$$

The NMSE between the first five experimental and noisy experimental NLPCs is equal to 0.23, 3.08, 4.45, 11.42 and 19.72 % respectively (impact excitation). From this result,

it appears that the first mode is not sensitive to the noise while the fourth and fifth modes are. The second and third modes are slightly affected by the presence of the noise but it is decided to keep them.

It should be noted that the number of neurons is selected using the "1 standard error rule" (see section 4.4). 6, 7 and 8 neurons in the mapping and demapping layers have been considered for the first three modes. It is also verified that the number of weights in the neural network is smaller than the number of training samples divided by a constant factor set to five here. This is carried out to ensure enough redundancy in the training set and to further reduce over-fitting.

The selection of the number of modes also needs to be carried out for the random excitation. Since the procedure is the same as for the impulse force, only the final results are given here, and Table 6.3 summarises the number of modes considered for the two excitations.

Figures 6.5 and 6.7 show the first two POMs for the impact and random excitations respectively. The visualisation is not obvious because all the DOF are not measured. However, it can still be observed that the POMs for the impact and random excitation resemble each other. The first POM mainly involves a relative displacement between DOF 31 and 37 thereby being similar to the fourth mode of the underlying linear system (see Figure 6.3). As far as the second POM is concerned, it almost appears as the result of a static force applied along DOF 31. The relative displacement between DOF 31 and 37 is now very small.

For VQPCA and NLPCA, the visualisation is realised through the projection of the modes onto several two-dimensional planes. Figures 6.6 and 6.8 thus represent the first VQPOM and NLPOM projected onto planes $x_1 - x_{11}$, $x_5 - x_7$, $x_{11} - x_{14}$ and $x_4 - x_8$. For the sake of comparison, the first POM is also given. The first NLPC is represented in Figure 6.9 for both excitations.

Correlation study

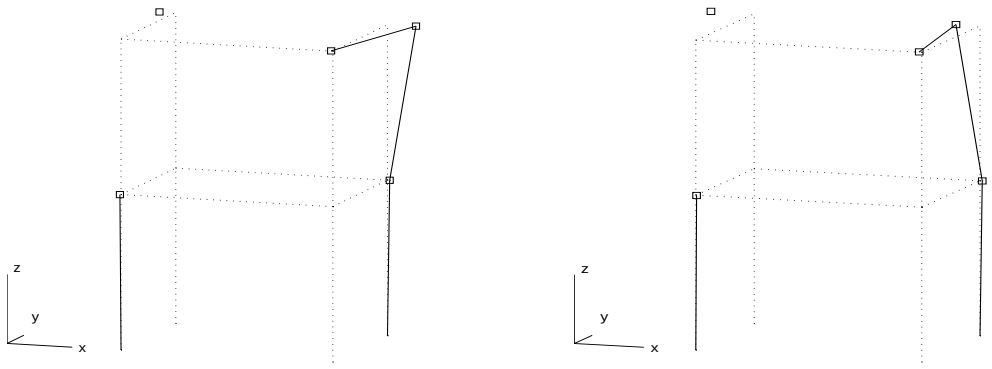
The inspection of the FRFs in section 6.3 has revealed that the portal frame is a non-linear structure. The initial finite element model should thus include a non-linearity whose functional form is given by:

$$f_{NL} = k_{NL} \, |x_{31} - x_{37}|^{\alpha} \operatorname{sign}(x_{31} - x_{37}) \tag{6.7}$$

The following values for the non-linear coefficient and exponent are chosen as an initial guess:

$$k_{NL} = 10^8 \, N/m^3 \quad \text{and} \quad \alpha = 3 \tag{6.8}$$

The correlation step aims at deciding whether or not these initial values are accurate enough to describe the structural behaviour. To this end, Figures 6.10 and 6.11 show the MAC matrix between the experimental and simulated POMs and VQPOMs respectively. The NMSE between the experimental and simulated NLPCs is listed in Table 6.4.



First POM Second POM

Figure 6.5: First two POMs.

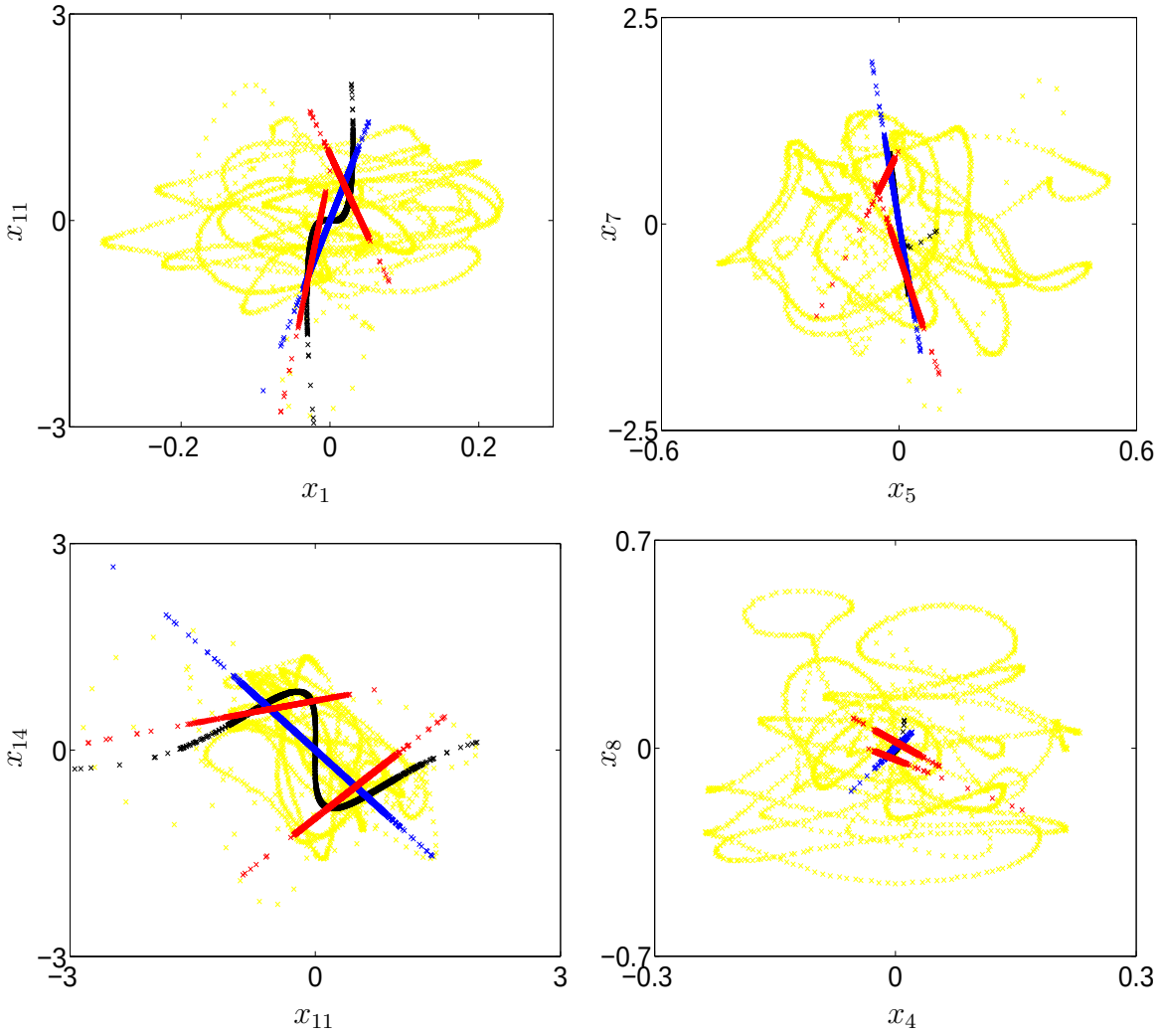
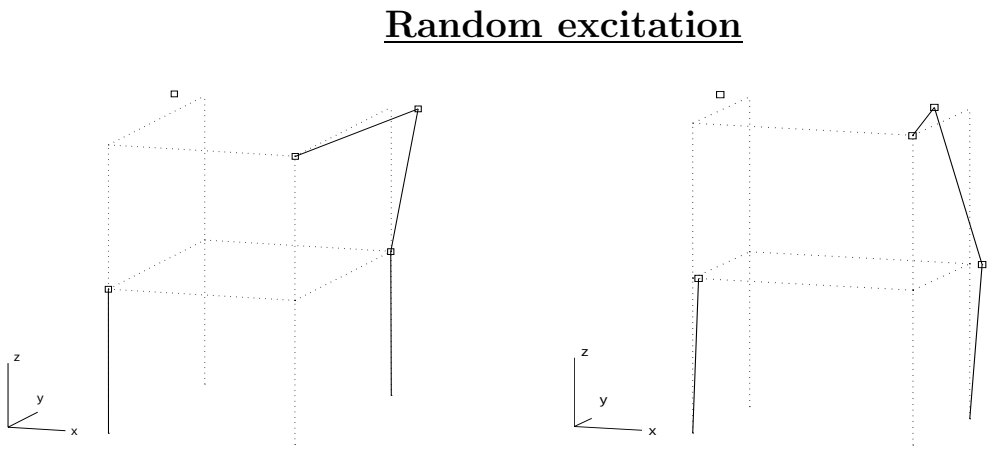


Figure 6.6: Four projections of the first mode. Yellow: experimental data; blue: POM; red: VQPOM; black: NLPOM.



First POM Second POM

Figure 6.7: First two POMs.

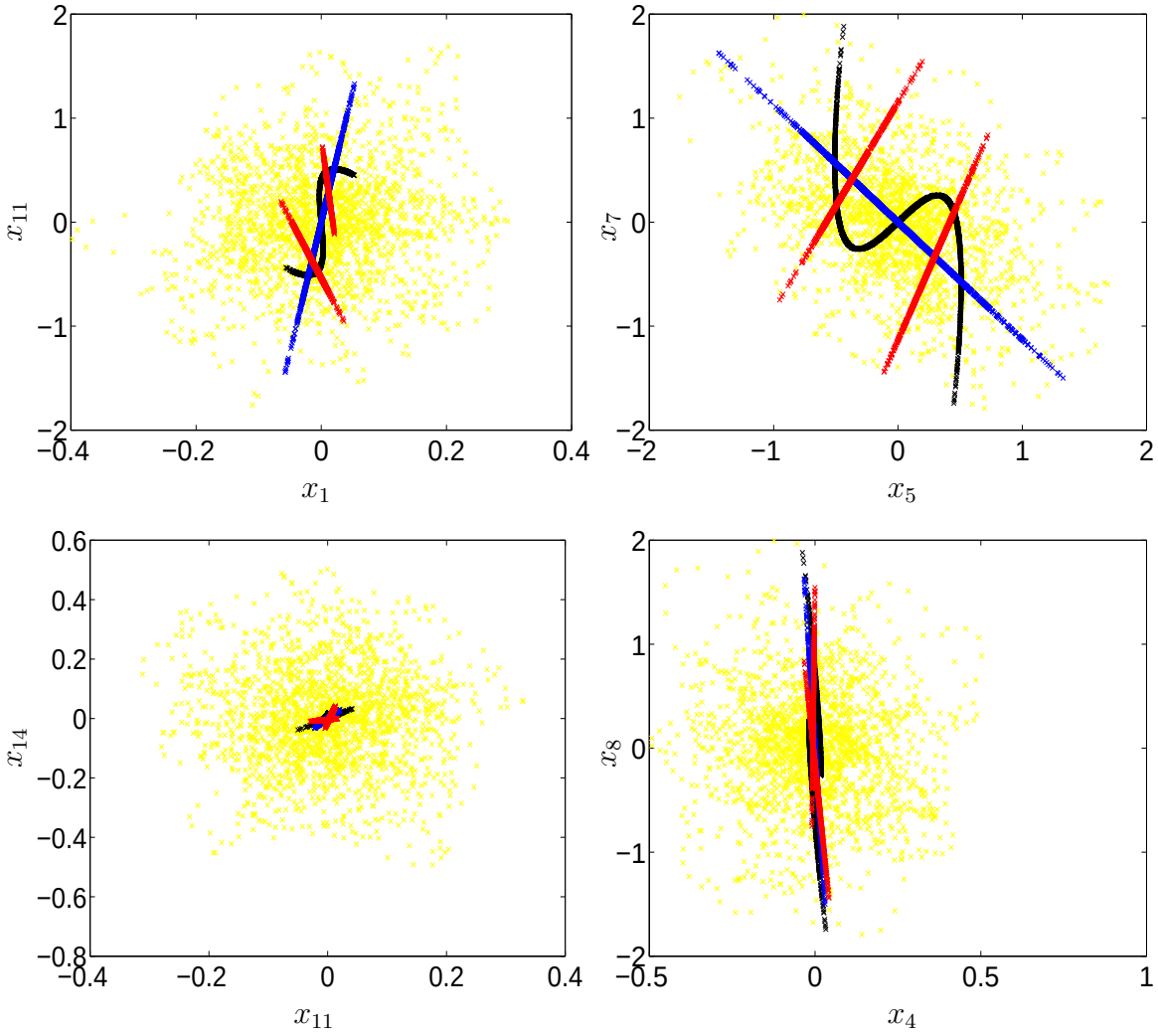


Figure 6.8: Four projections of the first mode. Yellow: experimental data; blue: POM; red: VQPOM; black: NLPOM.

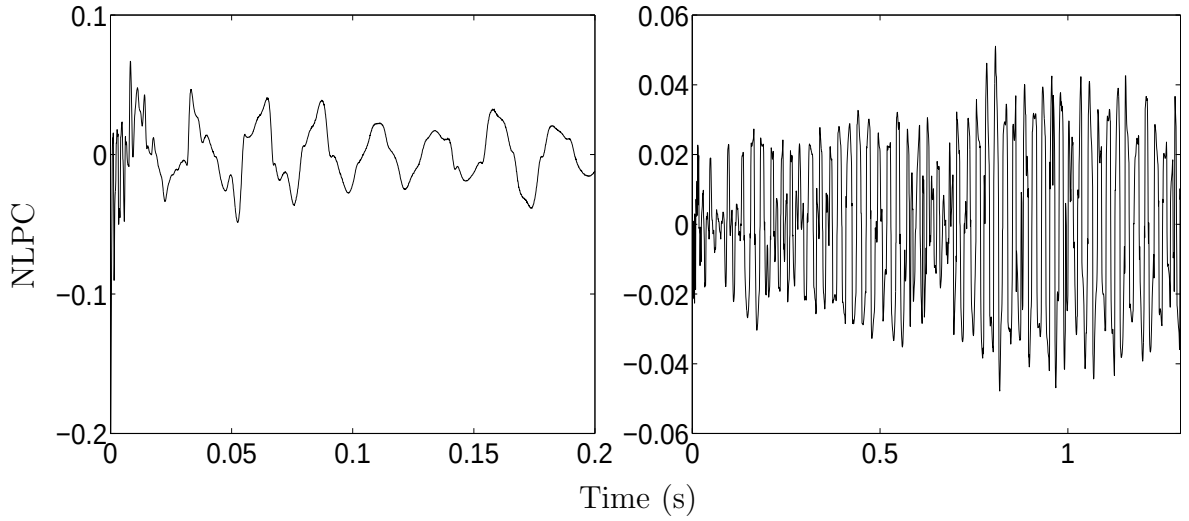


Figure 6.9: First NLPC. (a) Impact; (b) random.

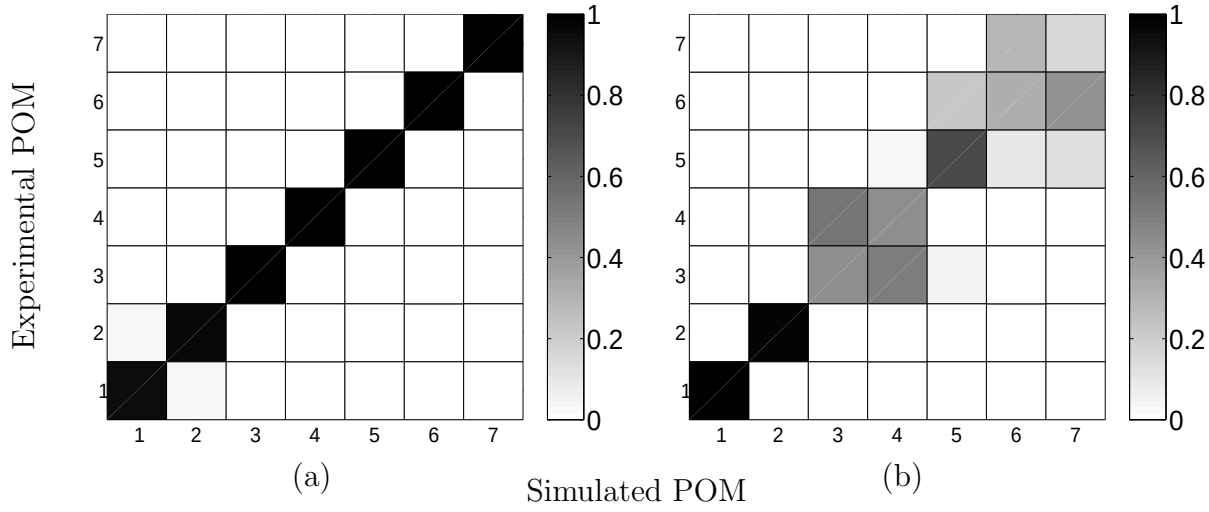


Figure 6.10: Correlation between experimental and simulated POMs. (a) Impact; (b) random.

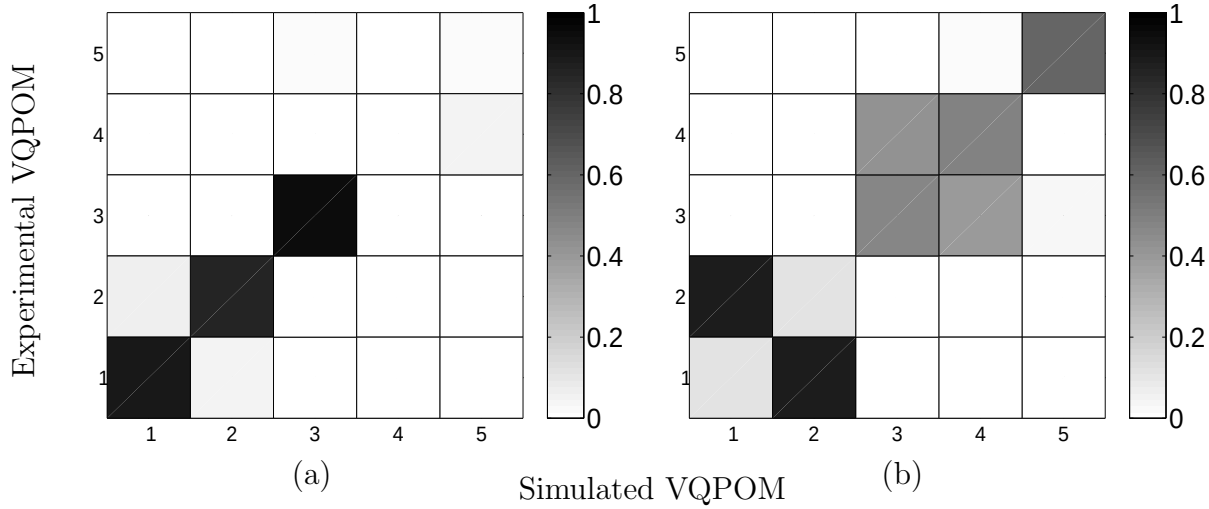


Figure 6.11: Correlation between experimental and simulated VQPOMs. (a) Impact; (b) random.

Mode	NMSE impact (%)	NMSE random (%)
1	137.72	61.07
2	152.95	141.97
3	72.16	158.22

Table 6.4: Correlation between experimental and simulated NLPCs

Updating parameter	Minimum value	Maximum value
Non-linear coefficient	10^5	10^9
Exponent	1.5	4
Young's modulus	$2.565\,10^{10}$	$3.135\,10^{10}$
Damping coefficient	10^{-5}	10^{-4}

Table 6.5: Updating parameters and their minimum and maximum values

For the POD, the correlation is almost perfect for the impulse force which does not allow us to conclude that the initial model needs to be improved. Fortunately, this is not the case for the random excitation for which the correlation is not satisfactory.

This problem is not encountered with VQPCA and NLPCA which clearly indicate for both types of excitation that the model needs to be updated.

6.5 Parameter selection

The correlation study has highlighted that the initial model needs to be corrected. To this end, the non-linear coefficient and exponent together with Young's modulus are updated. Since, in practice, a priori knowledge about damping is often limited, it is decided to consider the damping coefficient [see equation (6.4)] as an updating parameter.

The intervals in which these four parameters are allowed to vary are listed in Table 6.5. The non-linear coefficient and exponent may vary between 10^5 and 10^9, and 1.5 and 4 respectively which reflects our lack of a priori knowledge about the non-linearity. The material properties are often known with great confidence and variations of about 10 % may be considered as sufficient to include the global optimum of the problem. Finally, the damping coefficient ranges from 10^{-5} to 10^{-4}.

6.6 Model updating with limited knowledge about the non-linearity

The finite element model is now updated by minimising the objective functions based on the POD and its non-linear generalisations. In this section, it is assumed that there is no a priori knowledge about the non-linear coefficient and exponent.

In order to test the capability of a local optimisation method to find the problem solution, the minimisation of the objective functions is first realised using the sequential quadratic programming (SQP) algorithm. For a detailed description of this algorithm, the reader is referred to references [64, 132].

Although the minimisation of the three objective functions has been carried out using an ensemble of thirty different initial values of the updating parameters, the global optimum has never been reached neither for the impulse force nor for the random excitation. As it was suspected, it is not advisable to use a local optimiser when a priori knowledge about the updating parameters is limited, and it is meaningful to inquire whether the differential evolution algorithm may be of substantial help. This is investigated in what follows.

6.6.1 Results for the impulse force

The model updating using differential evolution is first realised for the impulse force. The population size is set to 12 which should favour exploration whereas the scaling factor F and the cross-over ratio CR are equal to 0.5 and 0.3 respectively which should favour exploitation. We have thus tried to strike a balance between exploration and exploitation. The minimisation is carried out using an ensemble of ten different initial values of the updating parameters.

- For the objective function based on the POD, all the minimisations have reached the optimum value, i.e., -1. The results of three of these minimisations are listed in Table 6.6. If Young's modulus and the damping coefficient are accurately identified, this is not the case for the non-linear coefficient and exponent. Their values vary between $1.57\,10^6$ and $1.03\,10^8$, and 2.21 and 2.87 respectively. Nevertheless, it is worth comparing all these solutions with the solution of the problem. For instance, Figure 6.12(a) displays the actual non-linearity together with the following non-linearity $1.57\,10^6\ |x_{31} - x_{37}|^{2.21}\,\mathrm{sign}(x_{31} - x_{37})$. It can clearly be seen that these two curves are almost identical. In fact, all the solutions are equivalent and offer a good approximation of the non-linearity. However, it should be borne in mind that if a higher excitation level was to be analysed, none of the solutions could however be used to predict the structural behaviour. This is illustrated in Figure 6.12(b) for which the relative displacement $x_{31} - x_{37}$ is five times as high as in Figure 6.12(a). The identified non-linearity now deviates greatly from the actual non-linearity.

Objective function	Best value of the objective function	Non-linear coefficient	Exponent	Young's modulus	Damping
POD	-1	$1.57\,10^{6}$	2.21	$3\,10^{10}$	$3\,10^{-5}$
	-1	$8.53\,10^{6}$	2.47	$3\,10^{10}$	$3\,10^{-5}$
	-1	$1.03\,10^{8}$	2.87	$3\,10^{10}$	$3\,10^{-5}$
Actual values	-1	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$
VQPCA	-1	$1.91\,10^{7}$	2.60	$3.03\,10^{10}$	$3.00\,10^{-5}$
	-1	$2.48\,10^{7}$	2.63	$3.03\,10^{10}$	$2.98\,10^{-5}$
	-1	$1.09\,10^{8}$	2.86	$3.06\,10^{10}$	$3.23\,10^{-5}$
Actual values	-0.997	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$
NLPCA	0.7568	$8.91\,10^{6}$	2.48	$3\,10^{10}$	$3.01\,10^{-5}$
Actual values	0.89	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$

Table 6.6: Optimisation results using the differential evolution algorithm (4 updating parameters, impulse force)

Objective function	Best value of the objective function	Non-linear coefficient	Exponent	Young's modulus	Damping
POD	-1	$1.02\,10^{7}$	2.50	$3\,10^{10}$	$3\,10^{-5}$
	-1	$1.10\,10^{7}$	2.52	$3\,10^{10}$	$3.01\,10^{-5}$
	-1	$1.47\,10^{7}$	2.56	$3\,10^{10}$	$3.06\,10^{-5}$
Actual values	-1	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$
VQPCA	-1	$9.86\,10^{6}$	2.49	$3\,10^{10}$	$2.99\,10^{-5}$
Actual values	-0.998	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$
NLPCA	1.44	$9.85\,10^{6}$	2.49	$3\,10^{10}$	$3.01\,10^{-5}$
Actual values	1.53	10^{7}	2.5	$3\,10^{10}$	$3\,10^{-5}$

Table 6.7: Optimisation results using the differential evolution algorithm (4 updating parameters, random excitation)

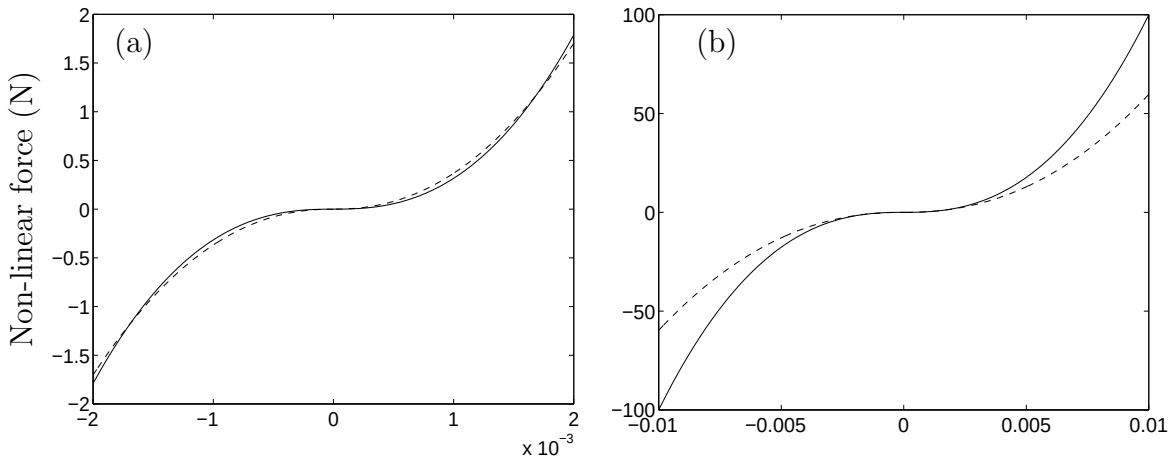


Figure 6.12: Comparison of actual and identified non-linearities: ——, actual non-linearity $10^7\ |x_{31} - x_{37}|^{2.5}\,\mathrm{sign}(x_{31} - x_{37})$; - - -, identified non-linearity $1.57\,10^6\ |x_{31} - x_{37}|^{2.21}\,\mathrm{sign}(x_{31} - x_{37})$.

- For VQPCA, 6 optimisations have reached -1. The results of three of these minimisations are presented in Table 6.6. This table shows that Young's modulus and the damping coefficient are not in very good agreement with the actual values. In addition, if the identified non-linearity was compared to the actual non-linearity as in Figure 6.12(a), it would be observed that the identification is less accurate than for the POD.

- For the objective function based on NLPCA, 7 optimisations have led to an objective function below 0.8. The minimum value of the objective function is equal to 0.7568, and the corresponding parameters are given in Table 6.6. For all the 7 optimisations, Young's modulus is properly identified and the damping coefficient is either $3\,10^{-5}$ or $3.01\,10^{-5}$. The difference with the results obtained with the POD and VQPCA is that the range in which the non-linear coefficient and exponent vary is much smaller. For instance, the exponent is always between 2.48 and 2.50 which tends to indicate that the objective function based on NLPCA is more sensitive to the non-linearity.

6.6.2 Results for the random excitation

Table 6.7 presents the results for the random excitation obtained using differential evolution. The same parameters as for the impulse force have been considered, i.e., F=0.5, CR=0.3, population size=12.

- For the POD, 6 iterations out of 10 have reached the optimum value. Young's modulus is always accurately estimated whereas slight variations in the damping coefficient and in the non-linear coefficient and exponent can be noticed.

- For VQPCA, a single optimisation has converged to -1. The non-linear coefficient is a bit lower than the actual value but this is due to the fact that the exponent is not 2.5 but rather 2.49.
- For NLPCA, 9 iterations out of 10 have reached 1.44. In this case, the identified parameters are in very good agreement with the actual parameters.

6.6.3 Discussion

Although the structural behaviour is very different for an impulse force or a random excitation, the following general observations regarding the performance of the three objective functions can be made:

- The objective function based on the POD has shown good convergence properties in the sense that the global minimum has often been attained. However, it has also appeared that it is not very sensitive to the non-linear coefficient and exponent. This may be attributed to the fact that the POD does not take the non-linear correlations between variables into account.
- The results obtained using VQPCA are relatively disappointing. Indeed, for the impulse force, the identified parameters were not as accurate as for the POD or NLPCA whereas for the random excitation, a single minimisation out of 10 has only converged to the solution. In addition, the number of function evaluations is much higher than for the two other objective functions.
- NLPCA has provided an accurate identification of the updating parameters thanks to its great sensitivity to the non-linearity.

It has also appeared that the differential evolution algorithm could almost always pick up the global minimum or a point very close to it which demonstrates its efficiency in a model updating context. The average number of function evaluations is presented in Table 6.8. This optimisation algorithm demands about a thousand function evaluations for the POD and NLPCA which might be excessive when the objective function evaluation requires significant CPU time. In the present example, 7.3 seconds are needed for a function evaluation (600 MHz processor) which means that a single optimisation using differential evolution requires about two hours. This still represents a reasonable computational burden.

In order to better study the topology of the objective functions, a contour plot can be generated but cannot be visualised for more than two variables. For instance, Figure 6.13 represents a contour plot of the objective functions for the impulse force when the non-linear coefficient and Young's modulus (left column), and the non-linear coefficient and damping coefficient (right column) vary. The objective function based on the POD appears to be smooth but much less sensitive to the updating parameters than VQPCA or NLPCA. These plots thus confirm our previous findings.

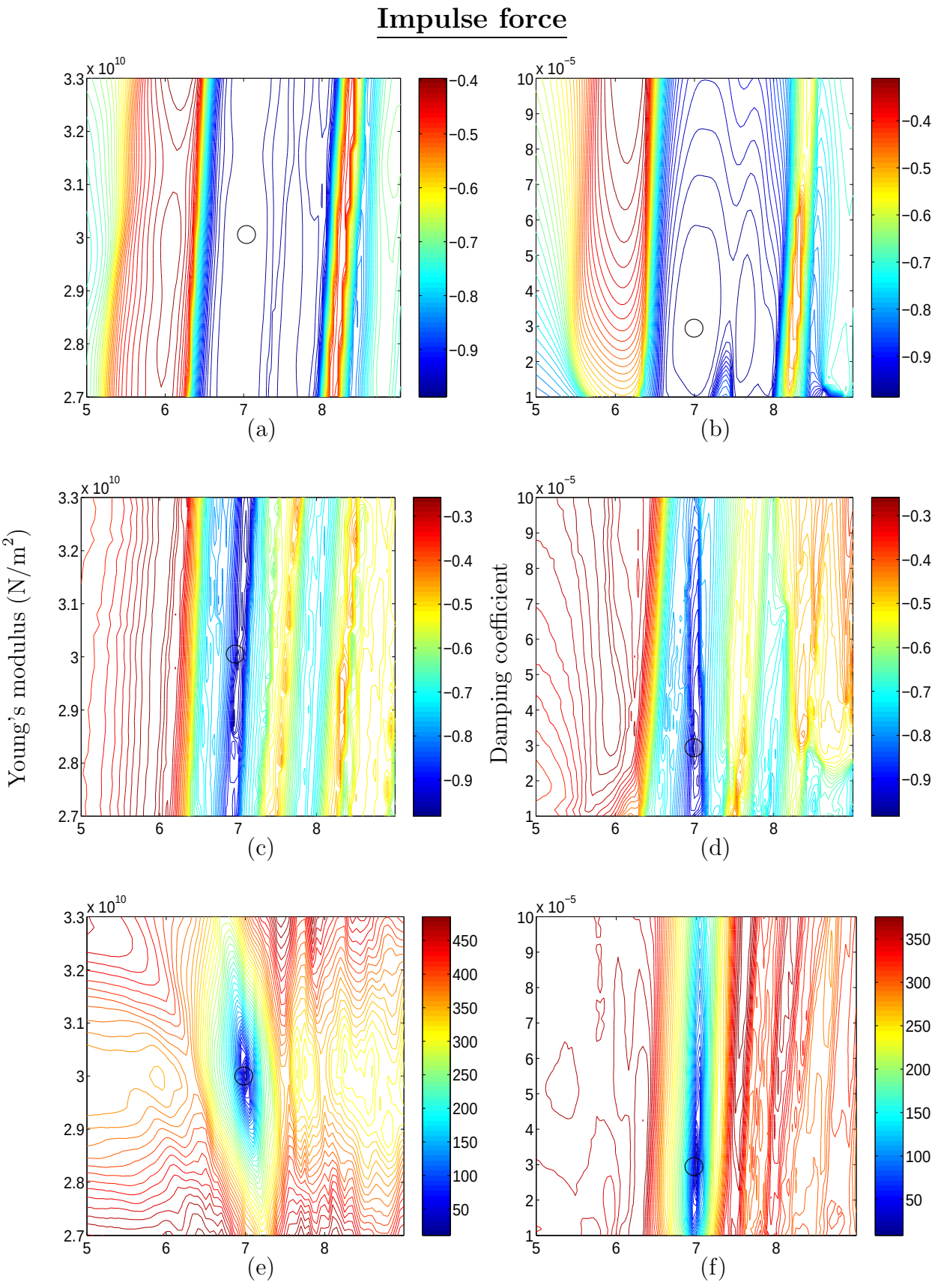


Figure 6.13: Contour plots of objective functions (◯, problem solution). Left column, Young's modulus vs. non-linear coefficient; right column, damping coefficient vs. non-linear coefficient. (a) POD; (b) POD; (c) VQPCA; (d) VQPCA; (e) NLPCA; (f) NLPCA.

Objective function	Differential evolution (4 parameters, impact)	Differential evolution (4 parameters, random)
POD	1044	1164
VQPCA	3075	3636
NLPCA	1128	750

Table 6.8: Average number of function evaluations for differential evolution

6.7 Model updating using the conditioned reverse path method

The model updating of the portal frame could be considered as completed since satisfactory results have been obtained in the previous section using NLPCA and differential evolution.

However, it is of great interest to inquire whether the CRP method may help to reduce the number of function evaluations needed for convergence by giving an estimation of the non-linear coefficient and exponent. Furthermore, it is also investigated whether the use of a local optimiser is appropriate in this case.

6.7.1 Application of the conditioned reverse path method

The non-linearity is modelled using the following functional form:

$$f_{NL} = k_{NL} \, |x_{31} - x_{37}|^{\alpha} \operatorname{sign}(x_{31} - x_{37}) \tag{6.9}$$

The first step is to determine the exponent α by seeking the maximum value of the spectral mean of the averaged cumulative coherence function of all the sensors:

$$accuracy = \frac{1}{N} \sum_{\omega=10}^{100} \left(\frac{1}{15} \sum_{i=1}^{15} \gamma_{Mi}^2(\omega) \right) \tag{6.10}$$

where N is the number of frequencies considered in the range from 10 to 100 Hz. Figure 6.14 depicts the evolution of the averaged cumulative coherence as a function of exponent α for two excitation levels (0.5 and 5 Nrms). It can be observed that the maximum value for the cumulative coherence is obtained in both cases for α equal to 2.5.

Now that the exponent has been determined, the FRFs of the underlying linear system together with the non-linear coefficient k_{NL} can be computed. The FRFs at the driving point are represented in Figure 6.15. The comparison with Figure 6.2 reveals that the distortions have disappeared and that the FRFs closely match the one obtained at low excitation level (0.05 Nrms) for which the structural behaviour is linear. In addition, the cumulative coherence functions represented in Figure 6.16 confirms that the identification is accurate.

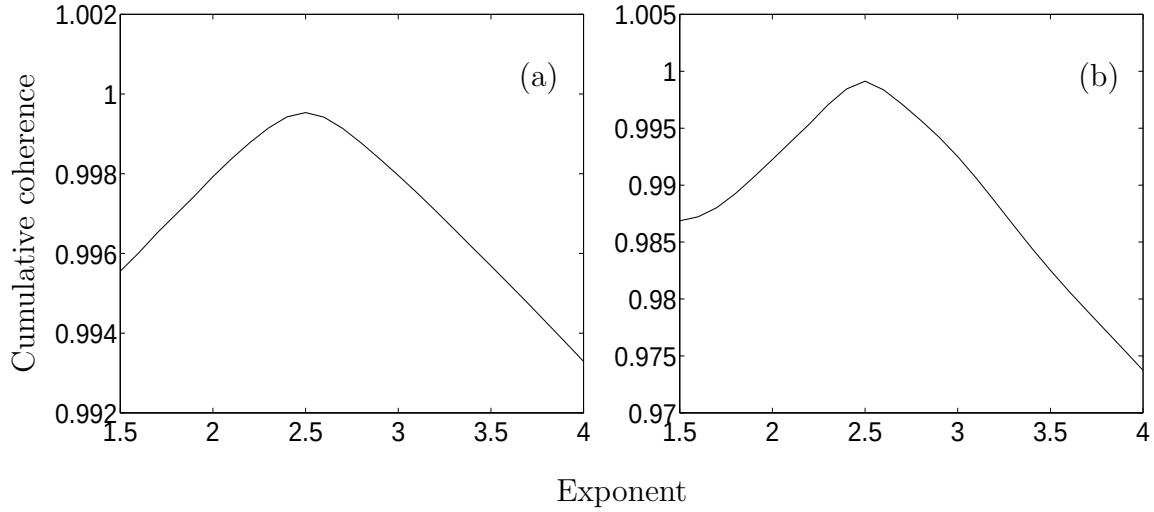


Figure 6.14: Cumulative coherence vs. exponent α. (a) 0.5 Nrms; (b) 5 Nrms.

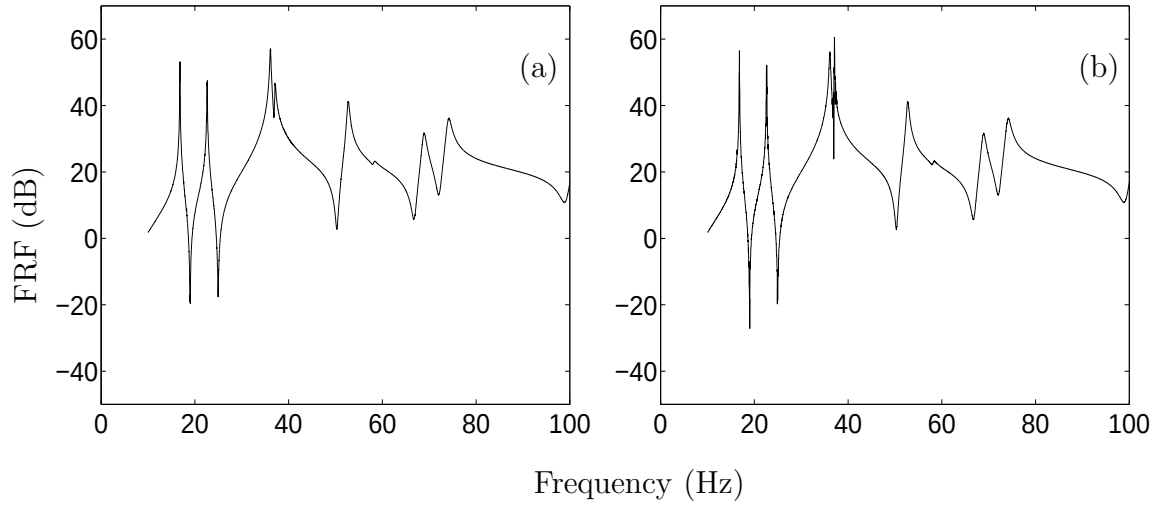


Figure 6.15: FRF at the driving point (H_{c2} estimate). (a) 0.5 Nrms; (b) 5 Nrms.

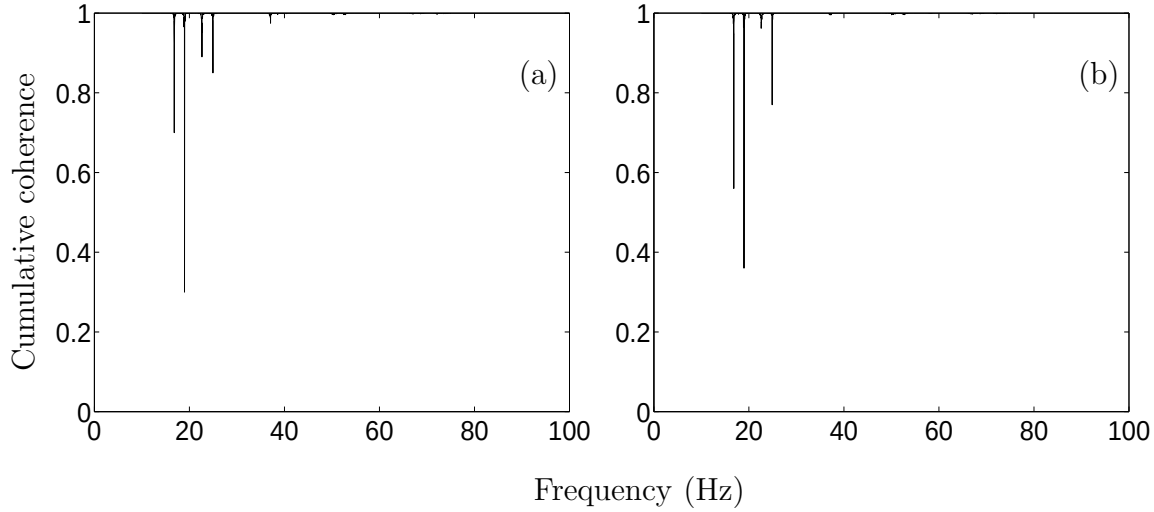


Figure 6.16: Cumulative coherence at the driving point. (a) 0.5 Nrms; (b) 5 Nrms.

The spectral mean of the non-linear coefficient is equal to $1.03\,10^7\,\mathrm{N/m^{2.5}}$ and $1.05\,10^7\,\mathrm{N/m^{2.5}}$ for 0.5 and 5 Nrms levels respectively. The corresponding imaginary parts are equal to $-2.56\,10^5\,\mathrm{N/m^{2.5}}$ and $1.21\,10^5\,\mathrm{N/m^{2.5}}$ and may be neglected.

6.7.2 Results for the impulse force

The model updating is now realised by taking the results given by the CRP method into account. Since the identification was accurate, our suggestion is to consider that the exponent remains equal to 2.5 during model updating in order to limit the number of updating parameters. As far as the non-linearity coefficient is concerned, it may vary between $5\,10^6\,\mathrm{N/m^{2.5}}$ and $2\,10^7\,\mathrm{N/m^{2.5}}$ which drastically restricts the search space.

The minimisation is carried out using the SQP algorithm and an ensemble of thirty different initial values of the updating parameters. Tables 6.9 and 6.10 contain the results and the average number of function evaluations respectively. It turns out that:

- For the objective function based on the POD, 11 minimisations have reached the global optimum for which the identified parameters can take different values. For instance, the non-linear coefficient vary from $9.89\,10^6$ to $1.11\,10^7\,\mathrm{N/m^{2.5}}$. For this latter value, the error is about 10% which is not negligible. It should also be noted that only 79 function evaluations are required for the convergence. The computational burden is thus very low compared to differential evolution.

- For VQPCA, a single run has converged to -1. The error on the corresponding parameters is about a few percent.

- For the objective function based on NLPCA, the global optimum has never been attained. The best value of the objective function is 1.54 for which the identified parameters are not so well estimated.

The error on the updating parameters never exceeds 10 % in Table 6.9 which might be considered as satisfactory. However, we have the feeling that the results may be improved by taking advantage of the respective merits of two features, namely the POMs and the NLPCs. Indeed, it appears that the objective function based on the POMs has good convergence properties but little sensitivity to the updating parameters. Unlike the POMs, the NLPCs are greatly sensitive to the updating parameters, and in particular to the non-linearity, but are responsible for less good convergence properties. Our suggestion is thus to build an objective function in which the POMs and the NLPCs are both included.

It still remains to determine their relative importance in the objective function assuming that they are both normalised between -1 and 0. Ideally, the good convergence properties of the POD should not be deteriorated too much. But once arrived near the optimum, NLPCA should help to increase the sensitivity to the updating parameters. From this reasoning, a much more importance to the POD should be given because NLPCA

Objective function	Best value of the objective function	Non-linear coefficient	Damping coefficient	Young's modulus
POD	-1	$9.89\,10^{6}$	$3.02\,10^{-5}$	$3\,10^{10}$
	-1	$1.09\,10^{7}$	$2.99\,10^{-5}$	$3\,10^{10}$
	-1	$1.11\,10^{7}$	$2.82\,10^{-5}$	$3.02\,10^{10}$
Actual values	-1	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
VQPCA	-1	$1.03\,10^{7}$	$3.05\,10^{-5}$	$3.02\,10^{10}$
Actual values	-0.997	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
NLPCA	1.54	$1.09\,10^{7}$	$2.86\,10^{-5}$	$3\,10^{10}$
Actual values	0.89	10^{7}	$3\,10^{-5}$	$3\,10^{10}$

Table 6.9: Optimisation results using the SQP algorithm (3 updating parameters, impulse force)

Objective function	SQP (3 parameters, impact)	SQP (3 parameters, random)
POD	79	85
VQPCA	197	257
NLPCA	163	159
POD-NLPCA (1 %)	129	—
POD-NLPCA (5 %)	180	—
POD-NLPCA (10 %)	229	—

Table 6.10: Average number of function evaluations for SQP algorithm

Objective function	Best value of the objective function	Non-linear coefficient	Damping coefficient	Young's modulus
POD-NLPCA (1 %)	-1.0106	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
Actual values	-1.0103	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
POD-NLPCA (5 %)	-1.053	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
Actual values	-1.0513	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
POD-NLPCA (10 %)	-1.097	10^{7}	$2.97\,10^{-5}$	$3\,10^{10}$
Actual values	-1.1026	10^{7}	$3\,10^{-5}$	$3\,10^{10}$

Table 6.11: Optimisation results using the SQP algorithm (3 updating parameters)

Objective function	Best value of the objective function	Non-linear coefficient	Damping coefficient	Young's modulus
POD	-1	$8.66\,10^{6}$	$3.01\,10^{-5}$	$3\,10^{10}$
	-1	$9.96\,10^{6}$	$3.00\,10^{-5}$	$3\,10^{10}$
	-1	$1.05\,10^{7}$	$3.00\,10^{-5}$	$3\,10^{10}$
Actual values	-1	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
VQPCA	-1	10^{7}	$3.02\,10^{-5}$	$3\,10^{10}$
Actual values	-0.998	10^{7}	$3\,10^{-5}$	$3\,10^{10}$
NLPCA	1.44	$9.96\,10^{6}$	$3.01\,10^{-5}$	$3\,10^{10}$
Actual values	1.53	10^{7}	$3\,10^{-5}$	$3\,10^{10}$

Table 6.12: Optimisation results using the SQP algorithm (3 updating parameters, random excitation)

should not influence the convergence before having reached very low values of the objective function.

We propose to investigate the performance of three objective functions that give a relative importance to NLPCA with respect to the POD equal to 1, 5 and 10 % respectively. Again, thirty minimisations starting from different initial values are carried out. The results are listed in Table 6.11. It can be observed that:

- For 1 %, three minimisations have converged to the global optimum for which the actual values of the parameters are retrieved. The average number of function evaluations is equal to 129 (see Table 6.10).
- For 5 %, a single minimisation has converged to the global optimum for which the actual values of the parameters are also retrieved. The average number of function evaluations is equal to 180.
- For 10 %, none of the minimisations has converged to the global optimum due to the fact that too much importance is given to NLPCA.

From these results, it appears that an objective function coupling the POD and NLPCA with a weight of about 1 % is very appealing because it has relatively good convergence properties and it gives a correct estimation of the parameters.

6.7.3 Results for the random excitation

The results for the random case are presented in Table 6.12. Table 6.10 shows the number of function evaluations. It can be noticed that:

- For the POD, 8 optimisations out of 30 have converged to -1, and the number of function evaluations is equal to 85. Young's modulus and the damping coefficient are accurately estimated but the non-linear coefficient takes again different values.
- For VQPCA, a single optimisation has reached -1. The non-linear coefficient and Young's modulus are equal to the actual values, and the error on the damping coefficient is below 1 %.
- For the objective function based on NLPCA, the global optimum has been attained five times. Contrary to the POD, the identified parameters take unique values at the optimum that are in very good agreement with the actual values.

Here, it is not necessary to build an objective function coupling the POMs and NLPCs because excellent results in terms of convergence and accuracy are obtained using NLPCA.

6.8 Analysis of the results and concluding remarks

The purpose of this chapter was to validate the proposed model updating strategy using simulated data from a three-dimensional portal frame. Excellent results have been obtained for the two types of excitation considered in the sense that the updating parameters have been accurately estimated.

A special attention has been devoted to the two following problems:

- Feature extraction;
- Minimisation of the objective function.

Feature extraction

From the results obtained in sections 6.6 and 6.7, some general conclusions can be drawn:

- The objective function based on the POD has very good convergence properties but is not enough sensitive to the non-linearity.
- NLPCA offers a very good characterisation of the non-linear behaviour. Accordingly, it has provided the best results in terms of accuracy.
- For VQPCA, the global optimum has often been difficult to attain. In addition, the average number of function evaluations is much higher than for the POD or NLPCA, and the updating parameters have not always been well estimated.

Although it is difficult to draw firm conclusions from a single experiment, our recommendation is to systematically use the objective function based on NLPCA because it enables us to obtain the most accurate results. If the convergence is problematic, then the

POD should be of great help due to its inherent smoothness. The coupling of these two features should improve the convergence while giving a reliable estimation of the updating parameters. The determination of their respective importance in the objective function is obviously application dependent. However, it can easily be tuned by observing whether the sensitivity to the parameters is too small or whether the global optimum is difficult to attain. In the former case, a greater importance to NLPCA should be given whereas in the latter case, a greater importance to the POD is advised.

The usefulness of NLPCA at the correlation step should also be enhanced. Indeed, for the impulse force, the POMs of the initial model were almost perfectly correlated with those of the experimental data [see Figure 6.10(a)] whereas the correlation between the two sets of NLPCs has clearly indicated that the initial model was not accurate.

Minimisation of the objective function

The coupling of a local optimiser with an identification technique such as the CRP method is to be recommended because it greatly reduces the computational burden in comparison with the use of differential evolution. In the present example, the average computational load was reduced by about a factor 10.

If for some reason this procedure cannot be applied, the utility of a local optimiser is questionable. In this context, it is advisable to use the differential evolution algorithm that has shown great efficiency in finding the global optimum.

Conclusions

The objective of this doctoral dissertation is to generate accurate and robust mathematical models of non-linear vibrating structures through the development and implementation of identification and finite element model updating techniques.

Three identification techniques have been assessed based upon the experience gained from the study of experimental applications. The restoring force surface method is appealing by its simplicity and efficiency for identification of single-degree-of-freedom systems. However, it suffers from a number of difficulties that make it unsuitable for the analysis of more complex non-linear structures. One fruitful approach to progress in this direction is the conditioned reverse path formulation. The non-linear identification through feedback of the outputs method has also been examined but, from our experience, it can by no means provide the same accuracy as the one obtained with the conditioned reverse path method.

As far as finite element model updating is concerned, the research has yielded two important findings. Firstly, the need for features that are representative of the non-linear behaviour and that give smooth objective functions is crucial. Table 7.1 compares the relative merits of the different features investigated herein. It is our opinion that decision making should not rely upon a single analysis based on one technique but should make use of the information provided by the proper orthogonal modes and the non-linear principal components. Indeed, these two features are very complementary in terms of smoothness and accuracy. Secondly, we cannot stress enough the importance of having a priori knowledge about the non-linearity in a model updating context. For this reason, identifications techniques such as the conditioned reverse path method should thus be exploited as a pre-processing step. If this is not feasible, the utility of local optimisers is questionable and, despite its computational load, we suggest the use of the differential evolution algorithm.

Table 7.2 gives the characteristics of the methods studied in this thesis whereas Table 7.3 summarises, according to our personal experience, their relative performances. It is not possible to give definitive guidance on which method is the best mainly because their characteristics and objectives are different. We however believe that the conditioned reverse path method and the proposed model updating strategy pave the way for the analysis of large-scale non-linear structures.

Features	Computational load	Non-linearity sensitivity	Smoothness of the objective function	Physical interpretability
POMs	low	low	high	high
VQPOMs	low	high	low	medium
NLPCs	high	high	medium	medium

Table 7.1: Performance of the features.

Method	Modal model	Structural model	SDOF	MDOF	Time domain	Frequency domain
RFS		×	×		×	
CRP	×			×		×
NIFO	×			×		×
Model updating		×		×	×	

Table 7.2: Characteristics of the methods.

Method	Computational load	Accuracy	Ease of implementation	Ease of use
RFS	low	high	medium	high
CRP	medium	high	medium	high
NIFO	low	medium	high	high
Model updating	high	high	low	low

Table 7.3: Performance of the methods.

Main contributions

As pointed out by Worden in his literature review [184]:
"*There are many techniques which are capable of dealing with non-linear systems with low dimensionality, but collapse if they are faced with a system with high model density. This can be partly attributed to concentrating too much on lumped parameter systems and a fruitful area of enquiry will be to consider continuous systems.*"

This thesis contributes to the development of methods capable of analysing continuous non-linear structures. The main achievements of this research work are:

- Three identification techniques have been implemented and compared using experimental applications [89, 90, 92, 95]. Accurate results have been obtained whereas experimental identification is known to be difficult.

- A relevant physical interpretation of the proper orthogonal modes has been presented using the singular value decomposition [91]. This allows us to justify the use of these modes in a model updating context.

- The thesis provides original contributions at two steps of the updating procedure:

  1. At the feature extraction level, non-linear generalisations of the proper orthogonal decomposition have been introduced [94, 97].
  2. At the optimisation level, the coupling of identification techniques with local optimisation algorithms has been proposed.

  Although the overwhelming majority of finite element model updating techniques is restricted to linear systems, these developments enable us to tackle non-linear systems. The model updating strategy has been successfully applied to a three-dimensional portal frame. Simulated measurements rather than actual experimental data have been used but this may be considered as a pertinent problem regarding the applications currently investigated in the literature.

Directions for future work

Although the results presented herein show promise in the difficult area of non-linear system identification, there is still much work to be done. The following discussion presents some of the key aspects that, we believe, will drive the development of non-linear system identification in the years to come.

- Feature extraction. Emphasis has been placed on finding new types of modes for non-linear systems. It has been shown that those given by the POD and NLPCA are well suited for model updating purposes. However, it is our opinion that no universal features are applicable to all situations in non-linear dynamics. Future research should thus continue in this direction. Particularly, a closer look at the non-linear normal modes would be of great interest. Features that accurately describe the frequency content of non-linear systems (e.g., higher order spectral techniques [13, 31]) should also be investigated. This has already been partially addressed with the use of the Wavelet transform in a companion thesis [111].

- Error localisation. Procedures for carefully selecting the locations where model errors are present have not been examined in the present study. Although engineering judgment and system-specific knowledge will always be inherent in the model updating process, the development of error localisation techniques in non-linear structural dynamics is yet to be performed in details.

- Optimisation of the objective function. Prior knowledge about the non-linear coefficient and/or exponent is not always available in practice. In this case, it is

advisable to use a global optimiser. Differential evolution has appeared to be useful for finding the global solution but may require thousands of objective function evaluations which may be prohibitive for real-life application. It is thus advantageous to inquire whether other global optimisers such as "Coupled Local Minimizers [168]" require much less iterations than evolutionary algorithms while providing the same efficiency.

- Meta-models/surrogate models. No attempt has been made to optimise the time needed to compute the objective function. In order that the updating procedure be viable for structures involving hundreds of degree-of-freedom, the objective function is to be obtained at low computational cost. To this end, meta-models also called surrogate models deserve our attention. The use of meta-model strategies is designed to reduce the number of computer simulations required during optimisation while maintaining the pertinent features of the design problem. Just as simulation models are simplifications of complex physical processes, meta-models are simplifications of expensive simulation models.

- Damping modelling. The process of modelling damping matrices is known to be challenging. This is the reason why the modal damping assumption is often used in structural dynamics although this assumption has no real physical background. Since the characterisation of damping is extremely important in making accurate predictions, a robust and practical procedure to identify damping matrices would be of particular interest.

- Non-linearity modelling. In order to limit the work, the use of polynomial approximations for modelling the non-linear behaviour was ubiquitous throughout the thesis. Although non-linearities in practical applications can often be modelled using such generic forms, the development of non-parametric models would be a useful contribution. These models have the advantage that an explicit functional form for the non-linearity is not necessary, it is rather determined from the system response by the method itself. In this context, the approach proposed by Peifer *et al.* [142] in which the non-linearity is modelled using additive models and smoothing splines is to be examined.

- Uncertainty quantification. All the structures considered in this thesis have been assumed to be deterministic that is, for given excitation conditions, the system response is always the same without any uncertainty. In practice, several parameters are known only within a certain band of accuracy. Possible future research should thus consider the natural dispersion as an integral part of the model. This problem is progressively being addressed in structural dynamics (see for instance references [41, 74]) but is still far from being resolved.

Appendix A

Quantitative Analysis of a Chaotic Signal

Analysis of chaotic processes is often performed by reconstructing the phase space using the method of delays [170]. Consider x(t) as a single scalar component of a process of unknown dimension. The embedding procedure aims at generating a vector time series. This is achieved by choosing an embedding dimension m and a time delay τ such that:

$$\mathbf{y}(t) = [x(t)\ x(t+\tau)\ ...\ x(t+(m-1)\tau)]^T \tag{A.1}$$

Part of the phase space is thus reconstructed from the time series $x(t)$ without even knowing the state space itself.

In order to reconstruct the attractor of a dynamic system, two problems need to be addressed. The first issue concerns the selection of the time delay τ. The time lag can be calculated from the first minimum of the average mutual information function:

$$I(\tau) = \sum_{x(t),x(t+\tau)} P(x(t), x(t+\tau)) \log_2 \frac{P(x(t), x(t+\tau))}{P(x(t))P(x(t+\tau))} \tag{A.2}$$

that evaluates the amount of information shared between two sets of data over a range of time lags [53].

The second problem deals with the computation the embedding dimension m of the system. The number of co-ordinates should provide a phase space in which the geometrical structure of the motion is completely unfolded. The false nearest neighbours technique can be exploited for this purpose [88]. For each point $\mathbf{y}(i)$ on the attractor in d dimensions, the nearest neighbour $\mathbf{y}^{NN}(i)$ is sought. The same vectors $\mathbf{y}(i)$ and $\mathbf{y}^{NN}(i)$ are then examined in dimension $d+1$. The specific test for false neighbours consists in evaluating the ratio:

$$\frac{|y_{d+1}(i) - y_{d+1}^{NN}(i)|}{R_d(i)} \tag{A.3}$$

where $R_d(i)$ is the distance in dimension d:

$$R_d(i)^2 = \frac{1}{d}\sum_{k=1}^{d}[x(i+(k-1)\tau) - x^{NN}(i+(k-1)\tau)]^2 \tag{A.4}$$

If the ratio is greater than some threshold ρ (typically, $15 < \rho < 20$), $\mathbf{y}(i)$ and $\mathbf{y}^{NN}(i)$ are deemed false nearest neighbours. By gradually increasing the dimension d until there are no false neighbours, a value for the necessary embedding dimension is established.

A.1 Correlation dimension

The main difference between a chaotic signal and a signal from a random process is that points on the chaotic attractor are spatially organised. A possible measure of this spatial organisation is given by the correlation integral [63]:

$$C(r) = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^{n} \theta(r - \|\mathbf{y}(i) - \mathbf{y}(j)\|) \tag{A.5}$$

where the Heaviside function $\theta(\cdot)$ is equal to one when its argument is positive and zero when its argument is negative. The expression counts the number of points out of the data set within an hypersphere of radius r and then divides by the square of the total number of points. As r becomes small, the correlation exponent ν is found to be characterised by:

$$C(r) \approx r^{\nu} \tag{A.6}$$

and is thus given by the slope of a log-log plot of the correlation integral $C(r)$ vs. the distances r:

$$\nu = \frac{\partial \log[C(r)]}{\partial \log(r)} \tag{A.7}$$

In practice, this plot usually shows three regions:

1. a linear region called the scaling region where the slope is equal to ν;
2. the lower region distorted by fluctuations due to the small number of points;
3. the upper region distorted due to the finite size of the attractor.

The correlation exponent has then to be plotted as a function of the embedding dimension. If there is deterministic dynamics in the system, the correlation exponent reaches the saturation value, i.e., it remains approximately constant when the embedding dimension is increased. The saturated correlation exponent is called the correlation dimension of the attractor, also known as D_2.

A.2 Lyapunov exponents

The Lyapunov exponents allow for an efficient diagnostic of chaotic systems. They quantify the average rate of convergence or divergence of nearby trajectories. A positive exponent is an indication of the global instability and sensitivity to initial conditions

and any system containing at least one positive exponent is defined to be chaotic. The magnitude of the exponent reflects the time scale on which system dynamics become unpredictable.

The Lyapunov spectrum is the set of numbers describing the rate of exponential deformation of an infinitesimal hypersphere in phase space over time. Due to the contracting or expanding nature of directions in phase space, the sphere is deformed into an ellipsoid with principal axis $p_i(t)$. The ith Lyapunov coefficient λ_i is then defined by:

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \log_2 \frac{p_i(t)}{p_i(0)} \tag{A.8}$$

The approach proposed by Wolf [178] can be used to estimate non-negative Lyapunov exponents from a reconstructed attractor. The discussion here is limited to the estimation of the largest Lyapunov exponent. The evolution of a point and its nearest neighbour, initially distant from L_i, is monitored. The points evolve with time on different trajectories and once the distance L_i' between both points exceeds some threshold, the procedure is restarted by considering a new pair of points. This is repeated until the trajectory has traversed the entire data file. The largest Lyapunov exponent is then computed from:

$$\lambda_1 = \frac{1}{t_n - t_0} \sum_{i=1}^{M} \log_2 \frac{L_i'}{L_i} \tag{A.9}$$

where M is the number of times the procedure is restarted.

Bibliography

[1] D.E. Adams and R.J. Allemang 2000 *Mechanical Systems and Signal Processing* **14**, 637-656. A frequency domain method for estimating the parameters of a non-linear structural dynamic model through feedback.

[2] M.A. Al-Hadid and J.R. Wright 1989 *Mechanical Systems and Signal Processing* **3**, 269-290. Developments in the force-state mapping technique for non-linear systems and the extension to the location of non-linear elements in a lumped parameter system.

[3] M.A. Al-Hadid and J.R. Wright 1992 *Mechanical Systems and Signal Processing* **6**, 383-401. Estimation of mass and modal mass in the identification of nonlinear single and multi degree-of-freedom systems using the force-state mapping approach.

[4] C. Archer and T.K. Leen 1999 *Proceedings of the International Joint Conference on Neural Networks, Washington DC, U.S.A.* Optimal dimension reduction and transform coding with mixture principal components.

[5] P. Argoul and T.P. Le 2003 *Mechanical Systems and Signal Processing* **17**, 243-250. Instantaneous indicators of structural behaviour based on the continuous Cauchy wavelet analysis.

[6] B. Armstrong-Helouvry, P. Dupont and C. Canudas de Wit 1994 *Automatica* **30**, 1083-1138. A survey of models, analysis tools and compensation methods for the control of machines with friction.

[7] M.F.A. Azeez and A.F. Vakakis 1997 *Proceedings of 1997 ASME Design Engineering Technical Conference, Sacramento, U.S.A.* Numerical and experimental analysis of the nonlinear dynamics due to impacts of a continuous overhung rotor.

[8] M.F.A. Azeez and A.F. Vakakis 2001 *Journal of Sound and Vibration* **240**, 859-889 . Proper orthogonal decomposition of a class of vibroimpact oscillations.

[9] P. Baldi and K. Hornik 1989 *Neural Networks* **2**, 53-58. Neural networks and principal component analysis: learning from examples without local minima.

[10] A.G. Barnston and C.F. Ropelewski 1992 *Journal of Climate* **5**, 1316-1345. Prediction of ENSO episodes using canonical correlation analysis.

[11] A.R. Barron 1993 *IEEE Transactions on Information Theory* **39**, 930-945. Universal approximation bounds for superposition of a sigmoidal function.

[12] S. Bellizzi and M. Defilippi 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 377-386. Nonlinear mechanical systems identification using linear system with random parameters.

[13] J.S. Bendat and A.G. Piersol 1980 *Engineering Applications of Correlation and Spectral Analysis.* New York, U.S.A.: John Wiley & Sons.

[14] J.S. Bendat 1986 *Random Data: Analysis and Measurement Procedures.* New York, U.S.A.: John Wiley-Interscience, second edition.

[15] J.S. Bendat 1990 *Nonlinear System Analysis and Identification from Random Data.* New York, U.S.A.: John Wiley & Sons.

[16] J.S. Bendat 1993 *Shock and Vibration* **1**, 21-31. Spectral techniques for nonlinear system analysis and identification.

[17] A. Benguedouar 1995 *Ph.D. Thesis, Boston University.* Proper orthogonal decomposition in dynamical modeling: a qualitative dynamic approach.

[18] G. Berkooz 1992 *Observations on the Proper Orthogonal Decomposition* in *Studies in Turbulence*, 229-247. New-York, U.S.A.: Springer-Verlag.

[19] G. Berkooz, P. Holmes and J.L. Lumley 1993 *Annual Review of Fluid Mechanics* **25**. The proper orthogonal decomposition in the analysis of turbulent flows.

[20] S.A. Billings and K.M. Tsang 1989 *Mechanical Systems and Signal Processing* **3**, 319-339. Spectral analysis for non-linear systems, Part I: Parametric non-linear spectral analysis.

[21] S.A. Billings and K.M. Tsang 1989 *Mechanical Systems and Signal Processing* **3**, 341-359. Spectral analysis for non-linear systems, Part II: Interpretation of non-linear fequency response functions.

[22] C.M. Bishop 1995 *Neural Networks for Pattern Recognition.* Oxford, U.K.: Oxford University Press.

[23] J.B. Bodeux and J.C. Golinval 2001 *Smart Materials and Structures* **10**, 479-489. Application of ARMAV models to the identification and damage detection of mechanical and civil engineering structures.

[24] D.S. Broomhead 1991 *Adaptive Signal Processing, SPIE Proceedings*, vol. 1565, 228-243. Signal processing for nonlinear systems.

[25] A.E. Bryson and Y.C. Ho 1975 *Applied Optimal Control (Optimization, Estimation and Control)*. New York, U.S.A.: Wiley.

[26] B.D. Bunday and G.R. Garside 1987 *Optimisation Methods in Pascal*. London, U.K.: Edward Arnold.

[27] B.P. Carlin and S. Chib 1995 *Journal of the Royal Statistical Society Series B* **77**, 473-484. Bayesian model choice via Markov Chain Monte Carlo.

[28] W. Cazemier 1997 *Ph.D. Thesis, Rijksuniversiteit, Groningen*. Proper orthogonal decomposition and low dimensional models for turbulent flows.

[29] K.Y. Chang 2000 *Ph.D. Thesis, University of Texas at Austin*. Nonlinear dimensionality reduction using probabilistic principal surfaces.

[30] H. Chipman, M. Hamada and C.F.J. Wu 1997 *Technometrics* **39**, 372-381. A Bayesian variable selection approach for analyzing designed experiments with complex aliasing.

[31] D. Choi, J. Chang, R.O. Stearman and E.J. Powers 1984 *Proceedings of the 2nd International Modal Analysis Conference, Orlando, U.S.A.*, 602-609. Bispectral interaction of nonlinear mode interactions.

[32] G.W. Cottrell and J. Metcalfe 1991 *Advances in Neural Information Processing Systems* 3. EMPATH: Face, emotion, and gender recognition using holons.

[33] E.F. Crawley and A.C. Aubert 1986 *AIAA Journal* **24**, 155-162. Identification of nonlinear structural elements by force-state mapping.

[34] E.F. Crawley and K.J. O'Donnel 1986 AIAA Paper 86-1013, 659-667. Identification of nonlinear system parameters in joints using the force-state mapping technique.

[35] J.P. Cusumano and B.Y. Bai 1993 *Chaos, Solitons and Fractals* **3**, 515-535. Period-infinity periodic motions, chaos and spatial coherence in a 10-degree-of-freedom impact oscillator.

[36] J.P. Cusumano, M.T. Sharkady and B.W. Kimble 1994 *Philosophical Transactions of the Royal Society of London*, **347**, 421-438. Experimental measurements of dimensionality and spatial coherence in the dynamics of a flexible-beam impact oscillator.

[37] G. Cybenko 1989 *Mathematics of Control, Signals, and Systems* **2**, 303-314. Approximation by superpositions of a sigmoidal function.

[38] D. DeMers and G. Cottrell 1993 *Advances in Neural Information Processing Systems 5, San Mateo, U.S.A.*, 580-587. Nonlinear dimensionality reduction.

[39] A. Deraemaeker 2001 *Ph.D. Thesis, Ecole Normale Supérieure de Cachan.* Sur la maitrise des modèles en dynamique des structures à partir de résultats d'essais.

[40] K.D. Dippery and S.W. Smith 1998 *Proceedings of the 16th International Modal Analysis Conference, Santa Barbara, U.S.A.*, 637-643. An optimal control approach to nonlinear system identification.

[41] S. Doebling, F.M. Hemez and W. Rhee 2000 *Proceedings of the European COST F3 Conference on System Identification & Structural Health Monitoring, Madrid, Spain*, 409-418. Statistical model updating and validation applied to nonlinear transient structural dynamics.

[42] S. Duym and J. Schoukens 1995 *Mechanical Systems and Signal Processing* **9**, 139-158. Design of excitation signals for the restoring force surface method.

[43] S. Duym, J. Schoukens and P. Guillaume 1995 *Proceedings of the 13th International Modal Analysis Conference, Nashville, U.S.A.*, 1392-1399. A local restoring force surface method.

[44] S. Duym 1998 *Ph.D. Thesis, Vrije Universiteit, Brussel.* Nonparametric identification of nonlinear mechanical systems.

[45] D.J. Ewins 2000 *Modal Testing: Theory, Practice and Application.* Hertfordshire, U.K.: Research Studies Press LTD, second edition.

[46] B.F. Feeny 1997 *Proceedings of ASME Design Engineering Technical Conferences, Sacramento, U.S.A.* Interpreting proper orthogonal modes in vibrations.

[47] B.F. Feeny and R. Kappagantu 1998 *Journal of Sound and Vibration* **211**, 607-616. On the physical interpretation of proper orthogonal modes in vibrations.

[48] M. Feldman 1994 *Mechanical Systems and Signal Processing* **8**, 119-127. Nonlinear system vibration analysis using the Hilbert transform - I. Free vibration analysis method 'FREEVIB'.

[49] M. Feldman 1994 *Mechanical Systems and Signal Processing* **8**, 309-318. Nonlinear system vibration analysis using the Hilbert transform - II. Forced vibration analysis method 'FORCEVIB'.

[50] P.M. Fitzsimons and C. Rui 1993 *Advances in Robust and Nonlinear Control Systems*, ASME DSC 53. Determining low dimensional models of distributed systems.

[51] L.J. Fogel, A.J. Owens and M.J. Walsh 1966 *Artificial Intelligence Through Simulated Evolution.* New York, U.S.A.: Wiley Publishing.

[52] D. Fotheringhmae and R. Baddeley 1997 *Biological Cybernetics* **77**, 282-288. Nonlinear principal component analysis of neuronal spike train data.

[53] A.M. Fraser and H.L. Swinney 1986 *Physical Review A* **33**, 1134-1140. Independant coordinates for strange attractors from mutual information.

[54] M.I. Friswell and J.E. Mottershead 1995 *Finite Element Model Updating in Structural Dynamics*. London, U.K.: Kluwer Academic Publishers.

[55] K. Fukunaga and D.R. Olsen 1971 *IEEE Transactions on Computers* **20**, 176-183. An algorithm for finding intrinsic dimensionality of data.

[56] E.I. George and R.E. McCulloch 1993 *Journal of the American Statistical Society* **88**, 881-889. Variable selection via Gibbs sampling.

[57] I.T. Georgiou and I.B. Schwartz 1999 *SIAM Journal of Applied Mathematics* **59**, 1178-1207. Dynamics of large scale coupled structural-mechanical systems: a singular perturbation proper orthogonal decomposition approach.

[58] M. Geradin and D. Rixen 1994 *Mechanical Vibrations, Theory and Application to Structural Dynamics*. Paris, France: Masson.

[59] A. Gersho and R.M. Gray 1992 *Vector Quantization and Signal Compression*. Norwell, U.S.A.: Kluwer Academic Publishers.

[60] C. Gibert and F. Thouverez 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 183-192. Identification of the non-linear modes applied to a structure including a non-linear component.

[61] S.J. Gifford and G.R. Tomlinson 1989 *Journal of Sound and Vibration* **135**, 289-317. Recent advances in the application of functional series to non-linear structures.

[62] G.H. Golub and C.F. Van Loan 1996 *Matrix Computations*. London, U.K.: The Johns Hopkins University Press, third edition.

[63] P. Grassberger and I. Procaccia 1983 *Physica D* **9**, 189-208. Measuring the strangeness of strange attractors.

[64] R.T. Hafka and Z. Gurdal 1992 *Elements of Structural Optimization*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

[65] R.W. Hamming 1989 *Digital filters*. Englewood Cliffs, U.S.A.: Prentice-Hall, third edition.

[66] T.K. Hasselman, M.C. Anderson and W.G. Gan 1998 *Proceedings of the 16th International Modal Analysis Conference, Santa Barbara, U.S.A.*, 644-651. Principal component analysis for nonlinear model correlation, updating and uncertainty evaluation.

[67] T. Hastie 1984 *Ph.D. Thesis, Stanford University*. Principal curves and surfaces.

[68] T. Hastie and W. Stuetzle 1989 *Journal of the American Statistical Association* **84**, 502-516. Principal curves.

[69] W. Heylen, S. Lammens and P. Sas 1997 *Modal Analysis Theory and Testing.* Leuven, Belgium: KUL Press.

[70] F.M. Hemez 1993 *Ph.D. Thesis, University of Colorado, Boulder.* Theoretical and experimental correlation between finite element models and modal tests in the context of large flexible space structures.

[71] F.M. Hemez and S.W. Doebling 1999 *Proceedings of the 17th International Modal Analysis Conference, Kissimmee, U.S.A.*, 1501-1510. Test analysis correlation and finite element model updating for nonlinear transient dynamics.

[72] F.M. Hemez and S.W. Doebling 2000 *Proceedings of the International Conference on Noise and Vibration Engineering, Leuven, Belgium.* Inversion of structural dynamics simulations: state-of-the-art and orientations of the research.

[73] F.M. Hemez and S.W. Doebling 2001 *Mechanical Systems and Signal Processing* **15**, 45-74. Review and assessment of model updating for non-linear, transient dynamics.

[74] F.M. Hemez and S.W. Doebling 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 13-35. Uncertainty, validation of computer models and the myth of numerical predictability.

[75] G.E. Hinton, M. Revow and P. Dayan 1995 *Advances in Neural Information Processing Systems 7, Cambridge, U.S.A.*, 1015-1022. Recognizing handwritten digits using mixtures of linear models.

[76] J.H. Holland 1975 *Adaptation in Natural and Artificial Systems.* Ann Arbor, U.S.A.: The University of Michigan Press.

[77] P. Holmes 1979 *Philosophical Transactions of the Royal Society of London* **292**, 419-448. A nonlinear oscillator with a strange attractor.

[78] P. Holmes, J.L. Lumley and G. Berkooz 1996 *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* New York, U.S.A.: Cambridge.

[79] H. Hotelling 1933 *Journal of Educational Psychology* **24**, 417-441, 498-520. Analysis of a complex of statistical variables into principal components.

[80] A. Hyvarinen, J. Karhunen, E. Oja 2001 *Independent Component Analysis.* New-York, U.S.A.: John Wiley & Sons.

[81] I.T. Jolliffe 1986 *Principal Component Analysis.* New York, U.S.A.: Springer-Verlag.

[82] A.A. Khan and N. S. Vyas 1999 *Journal of Sound and Vibration* **221**, 805-821. Non-linear parameter estimation using Volterra and Wiener theories.

[83] N. Kambhatla 1996 *Ph.D. Thesis, Oregon Graduate Institute of Science & Technology*. Local models and gaussian mixture models for statistical data processing.

[84] R. Kappagantu and B.F. Feeny 1999 *Journal of Sound and Vibration* **224**, 863-877. An optimal modal reduction of a system with frictional excitation.

[85] R. Kappagantu and B.F. Feeny 2000 *Nonlinear Dynamics* **22**, 317-333. Part 1: Dynamical characterization of a frictionally excited beam.

[86] R. Kappagantu and B.F. Feeny 2000 *Nonlinear Dynamics* **23**, 1-11. Part 2: Proper orthogonal modal modeling of a frictionally excited beam.

[87] K. Karhunen 1946 *Annals of Academic Science Fennicae, Series A1 Math. Phys.*, **37**. Uber lineare methoden in der wahrscheinlichkeitsrechnung.

[88] M.B. Kennel, R. Brown and H.D.I. Abarbanel 1992 *Physical Review A* **45**, 3403-3411. Determining embedding dimension for phase-space reconstruction using a geometrical construction.

[89] G. Kerschen, J.C. Golinval and K. Worden 2001 *Journal of Sound and Vibration* **244**, 597-613. Theoretical and experimental identification of a non-linear beam.

[90] G. Kerschen, V. Lenaerts, S. Marchesiello and A. Fasana 2001 *Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME*, **123**, 645-650. A frequency domain versus a time domain identification technique for nonlinear parameters applied to wire rope isolators.

[91] G. Kerschen and J.C. Golinval 2002 *Journal of Sound and Vibration* **249**, 849-865. Physical interpretation of the proper orthogonal modes using the singular value decomposition.

[92] G. Kerschen and J.C. Golinval 2002 *Proceedings of the 20th International Modal Analysis Conference, Los Angeles, U.S.A.*, 331-337. Frequency domain approaches for the identification of an experimental beam with a local non-linearity.

[93] G. Kerschen, B.F. Feeny and J.C. Golinval 2003 *Computer Methods in Applied Mechanics and Engineering* **192**, 1785-1795. On the exploitation of chaos to build reduced-order models.

[94] G. Kerschen and J.C. Golinval 2002 *Journal of Sound and Vibration* **254**, 867-876. Non-linear generalisation of principal component analysis: from a global to a local approach.

[95] G. Kerschen, V. Lenaerts and J.C. Golinval 2002 *Journal of Sound and Vibration*, in press. Identification of a continuous structure with a geometrical non-linearity, part I: conditioned reverse path method.

[96] G. Kerschen, J.C. Golinval and F.M. Hemez 2002 *Journal of Vibration and Acoustics*, in press. Bayesian model screening for the identification of non-linear mechanical structures.

[97] G. Kerschen and J.C. Golinval 2002 *Smart Materials and Structures*, submitted. Feature extraction using auto-associative neural networks.

[98] M.J. Kirby and R. Miranda 1996 *Neural Computations* **8**, 390-402. Circular nodes in neural networks.

[99] D. Kosambi 1943 *Journal of Indian Mathematical Society* **7**, 76-88. Statistics in function space.

[100] M.A. Kramer 1991 *A.I.Ch.E. Journal* **37**, 233-243. Nonlinear principal component analysis using autoassociative neural networks.

[101] E. Kreuzer and O. Kust 1996 *Nonlinear Dynamics and Control* ASME DE-Vol. 91, 105-110. Proper orthogonal decomposition - an efficient means of controlling self-excited vibrations of long torsional strings.

[102] P. Kulczycki 2001 *Journal of Dynamic Systems, Measurement and Control* **123**, 611-614. An algorithm for Bayes parameter estimation.

[103] A. Kyprianou 1999 *Ph.D. Thesis, University of Sheffield*. Non-linear parameter estimation of dynamic models using differential evolution: application to hysteretic systems and hydraulic engine mounts.

[104] A. Kyprianou, K. Worden, M. Panet 2001 *Journal of Sound and Vibration* **248**, 289-314. Identification of hysteretic systems using the differential evolution algorithm.

[105] M. Leblanc and R. Tibshirani 1994 *Journal of the American Statistical Association* **89**, 53-64. Adaptive principal surfaces.

[106] T.K. Leen, M. Rudnick and R. Hammerstrom 1990 *Proceedings of the IJCNN, IEEE, Piscataway, U.S.A.*, 51-56. Hebbian feature discovery improves classifier efficiency.

[107] V. Lenaerts, G. Kerschen and J.C. Golinval 2000 *Proceedings of the 18th International Modal Analysis Conference, San Antonio, U.S.A.*, 133-139. Parameter identification of nonlinear mechanical systems using proper orthogonal decomposition.

[108] V. Lenaerts, G. Kerschen, I. Trendafilova, J.C. Golinval, H. Van Brussel and W. Heylen 2000 *Proceedings of the International Conference on Noise and Vibration Engineering, Leuven, Belgium*, 431-438. Detection, localisation and identification of nonlinearities in structural dynamics.

[109] V. Lenaerts, G. Kerschen and J.C. Golinval 2001 *Mechanical Systems and Signal Processing* **15**, 31-43. Proper orthogonal decomposition for model updating of non-linear mechanical systems.

[110] V. Lenaerts, G. Kerschen and J.C. Golinval 2002 *Journal of Sound and Vibration*, in press. Identification of a continuous structure with a geometrical non-linearity, part II: proper orthogonal decomposition.

[111] V. Lenaerts 2003 *Ph.D. Thesis, Université de Liège*. Identification et recalage de modèles structuraux en dynamique non-linaire.

[112] I.J. Leontaritis and S.A. Billings 1985 *International Journal of Control* **41**, 303-328. Input-output parametric models for nonlinear systems, part I: deterministic nonlinear systems.

[113] I.J. Leontaritis and S.A. Billings 1985 *International Journal of Control* **41**, 329-344. Input-output parametric models for nonlinear systems, part II: stochastic nonlinear systems.

[114] R.I. Levin and N.A.J. Lieven 1998 *Mechanical Systems and Signal Processing* **12**, 91-120. Dynamic finite element model updating using simulated annealing and genetic algorithms.

[115] M. Link 1999 *Updating of Analytical Models - Basic Procedures and Extensions* in *Modal Analysis & Testing*, 281-304. J.M.M. Silva and N.M.M. Maia, eds.: The Netherlands: Kluwer Academic Publishers.

[116] L. Ljung 1987 *System Identification - Theory for the User*. Englewood Cliffs, U.S.A.: Prentice-Hall.

[117] S.P. Lloyd 1957 *Bell Laboratories Technical Note*. Least Squares Quantization in PCM. Published in 1982 in *IEEE Transactions on Information Theory* **28**, 129-137.

[118] M. Loeve 1945 *Compte Rend. Acad. Sci. (Paris)*. Fonctions aléaloires de second ordre.

[119] J.L. Lumley 1967 *The Structure of Inhomogeneous Turbulent Flows* in *Atmospheric Turbulence and Radio Wave Propagation*, 166-178. A.M. Yaglom and V.I. Tatarski, eds.: Moscow: Nauka.

[120] X. Ma, M.F.A. Azeez and A.F. Vakakis 1998 *Proceedings of the 7th Conference on Nonlinear Vibration, Stability and Dynamics of Structures, Blacksburg, U.S.A.* Nonparametric nonlinear system identification of a nonlinear flexible system using proper orthogonal decomposition.

[121] X. Ma, M.F.A. Azeez and A.F. Vakakis 2000 *Mechanical Systems and Signal Processing* **14**, 37-48. Non-linear normal modes and non-parametric system identification of non-linear oscillators.

[122] X. Ma and A.F. Vakakis 2001 *Journal of Vibration and Acoustics* **123**, 36-44. Nonlinear transient localization and beat phenomena due to backlash in a coupled flexible system.

[123] E.C. Malthouse 1998 *IEEE Transaction on Neural Networks* **9**, 165-173. Limitations of nonlinear PCA as performed with generic neural networks.

[124] S. Marchesiello, L. Garibaldi, J.R. Wright and J.E. Cooper 2000 *Proceedings of the European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain*, 429-438. Applications of the conditioned reverse path method to multi-degree-of-freedom non-linear systems.

[125] S. Marchesiello, L. Garibaldi and A. Fasana 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 139-150. CRP application to the non-linear beam benchmark of the COST F3 action.

[126] S.F. Masri and T.K. Caughey 1979 *Journal of Applied Mechanics* **46**, 433-447. A nonparametric identification technique for nonlinear dynamic problems.

[127] S.F. Masri, H. Sassi and T.K. Caughey 1982 *Journal of Applied Mechanics* **49**, 619-628. A nonparametric identification of nearly arbitrary nonlinear systems.

[128] L. Meirovitch 1980 *Computational Methods in Structural Dynamics*. Alphen a/d Rijn, The Netherlands: Sijthoff and Noordhoff.

[129] S. Meyer, M. Weiland and M. Link 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 151-160. Modelling and updating of local non-linearities using frequency response residuals.

[130] A.H. Monahan 2000 *Ph.D. Thesis, University of British Columbia*. Nonlinear principal component analysis of climate data.

[131] F.C. Moon 1980 *Journal of Applied Mechanics* **47**, 190-196. Experiments on chaotic motions of a forced nonlinear oscillator: strange attractors.

[132] A.J. Morris 1982 *Foundations of Structural Optimization: a Unified Approach*. Chichester, U.K.: John Wiley & Sons.

[133] D.F. Morrison 1967 *Multivariate Statistical Methods*. New York, U.S.A.: McGraw-Hill.

[134] A.J. Newman and P.S. Krishnaprasad 1998 *Proceedings of the 32th Conference on Information Sciences and Systems, Princeton, U.S.A.* Nonlinear model reduction for RTCVD.

[135] M.A. Obukhov 1954 *T. Geophys. Int. Akad. Nauk. USSR* **24**, 3-42. Statistical description of continuous fields.

[136] E. Oja 1982 *Journal of Mathematical Biology* **15**, 267-273. A simplified neuron model as a principal component analyzer.

[137] E. Oja 1989 *International Journal of Neural Systems* **1**, 61-68. Neural networks, principal components and subspaces.

[138] D. Otte 1994 *Ph.D. Thesis, Katholieke Universiteit, Leuven*. Development and evaluation of singular value analysis methodologies for studying multivariate noise and vibration problems.

[139] R. Pascual, J.C. Golinval and M. Razeto 1997 *Proceedings of the 15th International Modal Analysis Conference, Orlando, U.S.A.*, 587-592. A frequency domain correlation technique for model correlation and updating.

[140] R. Pascual 1999 *Ph.D. Thesis, Université de Liège*. Model based structural damage assessment using vibration measurements.

[141] K. Pearson 1901 *Philosophical Magazine* **2**, 559-572. On lines and planes of closest fit to systems of points in space.

[142] M. Peifer, J. Timmer and H.U. Voss 2002 *Journal of Sound and Vibration*, submitted. Nonparametric identification of nonlinear scillating systems.

[143] V.S. Pougachev 1953 *Izv. Akad. Nauk. USSR* **17**, 1401-1402. General theory of the correlations of random functions.

[144] T. Prajogo, F. Al-Bender and H. Van Brussel 1995 *Proceedings of the 8th International Precision Engineering Seminar, Compiègne, France*, 229-232. Identification of prerolling friction dynamics of rolling element bearings: modelling and application to precise positioning systems.

[145] I. Rechenberg 1973 *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution.* Stuttgart, Germany: Frommann-Holzboog.

[146] M.S. Riaz and B.F. Feeny 1999 *Proceedings of 1999 ASME Design Engineering Technical Conference, Las Vegas, U.S.A.* Proper orthogonal decomposition of an experimental cantilever beam.

[147] H.J. Rice and J.A. Fitzpatrick 1991 *Journal of Sound and Vibration* **149**, 397-411. A procedure for the identification of linear and non-linear multi-degree-of-freedom systems.

[148] C.M. Richards and R. Singh 1998 *Journal of Sound and Vibration* **213**, 673-708. Identification of multi-degree-of-freedom non-linear systems under random excitations by the reverse-path spectral method.

[149] C.M. Richards and R. Singh 1999 *Journal of Sound and Vibration* **220**, 413-450. Feasibility of identifying non-linear vibratory systems consisting of unknown polynomial forms.

[150] R.M. Rosenberg 1962 *Journal of Applied Mechanics* **29**, 7-14. The normal modes of nonlinear n-degree-of-freedom systems.

[151] R.M. Rosenberg 1966 *Advances in Applied Mechanics* **9**, 155-242. On nonlinear vibrations of systems with many degrees of freedom.

[152] T.D. Sanger 1989 *Neural Networks* **2**, 459-473. Optimal unsupervised learning in a single linear feedforward neural network.

[153] M. Schetzen 1980 *The Volterra and Wiener Theories of Nonlinear Systems.* New York, U.S.A.: John Wiley & Sons.

[154] J.F. Schultze, F.M. Hemez, S.W. Doebling and H. Sohn 2001 *Shock and Vibration* **8**, 325-337. Application of non-linear system model updating using feature extraction and parameter effect analysis.

[155] G. Schwarz 1978 *The Annals of Statistics*, **6**, 461-464. Estimating the dimension of a model.

[156] S.W. Shaw and C. Pierre 1993 *Journal of Sound and Vibration* **164**, 85-124. Normal modes for non-linear vibratory systems.

[157] S.W. Shaw and C. Pierre 1994 *Journal of Sound and Vibration* **169**, 319-347. Normal modes of vibration for non-linear continuous systems.

[158] M. Simon and G.R. Tomlinson 1984 *Journal of Sound and Vibration* **96**, 421-436. Use of the Hilbert transform in modal analysis of linear and non-linear structures.

[159] L. Sirovich 1987 *Quarterly of Applied Mathematics* **45**, 561-571. Turbulence and the dynamics of coherent structures, Part I: Coherent Structures.

[160] H. Sohn, C.R. Farrar, N.F. Hunter and K. Worden 2001 *Journal of Dynamic Systems, Measurement, and Control, Transactions of the ASME*, **123**, 706-711. Structural health monitoring using statistical pattern recognition techniques.

[161] H. Sohn, K. Worden and C.R. Farrar 2001 *Proceedings of SPIE's 8th Annual International Symposium on Smart Structures and Materials, Newport Beach, U.S.A.* Novelty detection under changing environmental conditions.

[162] W.M. Spears, K.A. De Jong, T. Back, D. Fogel and H. de Garis 1993 *Proceedings of the 1993 European Conference on Machine Learning, Vienna, Austria*, 442-459. An overview of evolutionary computation.

[163] J. Staar 1982 *Ph.D. Thesis, Katholieke Universiteit, Leuven.* Concepts for reliable modeling of linear systems with application to on-line identification of multivariate state space descriptions.

[164] T. Stamkopoulos, K. Diamantaras, N. Maglaveras and M. Strintzis 1998 *IEEE Transactions on Signal Processing* **46**, 3058-3067. ECG analysis using nonlinear PCA neural networks for ischemia detection.

[165] R. Storn 1996 *Proceedings of the 1996 North American Fuzzy Information Processing Society, Berkeley, U.S.A.*, 519-523. On the usage of differential evolution for function optimization.

[166] R. Storn and K. Price 1997 *Journal of Global Optimization* **11**, 341-359. Differential evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces.

[167] G. Strang 1986 *Introduction to Applied Mathematics.* Cambridge, U.S.A.: Wellesley-Cambridge Press.

[168] J.A.K. Suykens, J. Vandewalle and B. De Moor 2001 *International Journal of Bifurcation and Chaos* **11**, 2133-2144. Intelligence and cooperative search by coupled local minimizers.

[169] J. Swevers, F. Al-Bender, C. Ganseman and T. Prajogo 2000 *IEEE Transactions on Automatic Control* **45**, 675-686. An integrated friction model structure with improved presliding behavior for accurate friction compensation.

[170] F. Takens 1981 *Detecting strange attractors in turbulence.* In Rand, D. A. and Young, L.-S., editors, Dynamical Systems and Turbulence: Lecture Notes in Mathematics, Vol. 898, 336-381. Berlin, Germany: Springer-Verlag.

[171] D.M. Tang and E.H. Dowell 1988 *Journal of Applied Mechanics* **55**, 190-196. On the threshold force for chaotic motions for a forced buckled beam.

[172] F. Thouverez and L. Jezequel 1996 *Journal of Sound and Vibration* **189**, 193-213. Identification of NARMAX models on a modal base.

[173] I. Trendafilova and H. Van Brussel 1999 *Proceedings of the Second International Conference on Identification in Engineering Systems, Swansea, Wales*, 169-178. On feature extraction for condition monitoring using time series analysis and distance techniques.

[174] I. Trendafilova and H. Van Brussel 2001 *Mechanical Systems and Signal Processing* **15**, 1141-1164. Non-linear dynamics tools for the motion analysis and condition monitoring of robot joints.

[175] A.F. Vakakis, L.I. Manevitch, Y.V. Mikhlin, V.N. Pilipchuk and A.A. Zevin 1996 Normal Modes and Localization in Nonlinear Systems. New York, U.S.A.: John Wiley & Sons.

[176] A.F. Vakakis 1997 *Mechanical Systems and Signal Processing* **11**, 3-22. Non-linear normal modes and their applications in vibration theory: an overview.

[177] M. Wax and T. Kailath 1985 *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP **33**, 387-392. Detection of signals by information theoretic criteria.

[178] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano 1985 *Physica D* **16**, 285-317. Determining Lyapunov exponents from a time series.

[179] K. Worden and G.R. Tomlinson 1988 *Proceedings of the 6th International Modal Analysis Conference, Kissimmee, U.S.A.*, 1471-1479. Developments in Force State Mapping for Nonlinear Systems.

[180] K. Worden 1990 *Mechanical Systems and Signal Processing* **4**, 295-319. Data processing and experiment design for the restoring force surface method, part I: integration and differentiation of measured time data.

[181] K. Worden and G.R. Tomlinson 1991 *Proceedings of the 9th International Modal Analysis Conference, Las Vegas, U.S.A.*, 757-764. An experimental study of a number of nonlinear SDOF systems using the restoring force surface method.

[182] K. Worden and G. Manson 1998 *Journal of Sound and Vibration* **217**,781-789. Random vibrations of a duffing oscillator using the Volterra series.

[183] K. Worden and G. Manson 1999 *Journal of Sound and Vibration* **226**,397-405. Random vibrations of a multi-degree-of-freedom non-linear system using the Volterra series.

[184] K. Worden 2000 *Proceedings of the European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain*, 29-52. Nonlinearity in structural dynamics: the last ten years.

[185] K. Worden and G.R. Tomlinson 2001 *Nonlinearity in Structural Dynamics: Detection, Identification and Modelling*. Bristol, U.K.; Philadelphia, U.S.A.: Institute of Physics Publishing.

[186] J.R. Wright, M.F. Platten, J.E. Cooper and M. Sarmast 2001 *Proceedings of the International Conference on Structural System Identification, Kassel, Germany*, 49-68. Identification of multi-degree-of-freedom weakly non-linear systems using a model based in modal space.

[187] C.F.J. Wu and M. Hamada 2000 *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York, U.S.A.: Wiley.

[188] K.V. Yuen, L.S. Katafygiotis, C. Papadimitriou and N.C. Mickleborough 2001 *Journal of Dynamic Systems, Measurement and Control* **123**, 677-686. Optimal sensor placement methodology for identification with unmeasured excitation.

[189] O. Zienkiewicz 1977 *The Finite Element Method*. London, U.K.: McGraw-Hill.