POLITECNICO DI MILANO

SCHOOL OF CIVIL, ENVIRONMENTAL
AND LAND MANAGEMENT ENGINEERING
Master of Science in Civil Engineering
Track Structures

Damage identification in thin plates through hybrid Kalman-particle filtering and reduced-order modeling

Supervisors: Prof. Stefano MARIANI
Ph.D. Saeed EFTEKHAR AZAM
Ph.D. Francesco CAIMMI

Master of Science Thesis by:
Giovanni CAPELLARI
ID number 782706

Academic Year 2012 - 2013

Acknowledgements

These few lines are not adequate to express my sincere gratitude to everyone who made this thesis possible, in particular Professor Stefano Mariani for his brilliant ideas, suggestions and encouragements that allowed me to go ahead even during hard times, Ph.D. Saeed Eftekhar Azam for his patience and kindness in explaining me new interesting things and Ph.D. Francesco Caimmi for his essential and practical help.
I would like to thank everyone who supported me and stood my mood swings, in particular Giacomo Schierano and Roberto Peci for their technical advice, my roommates and all my friends.
Last but not least, this great achievement would not have been possible without the continuous presence of my family, that I owe everything I am.

Contents

List of Figures

List of Tables

Abstract

The aim of the present thesis is the description and the assessment of a damage identification method based on recursive Bayesian filters. Its main goal is the ability of detecting the damage indexes associated to any given structure, or in other words, estimating the actual local stiffness of the system, given a certain number of observations. Other two main objectives are to guarantee a reduced computational cost and the coupling with a commercial FE code. These requirements allow the method to be presented as a conceptual strategy that can be applied to a large variety of applications purposes and to any type of FE formulation. Since the main drawback of any identification method based on standard recursive filters is the excessive computational time, two solutions are adopted: a model order reduction and an improved filtering strategy, that uses a re-sampling technique and a modified particle filter. The model order reduction is obtained through a Galerkin-based projection of the original full model into a sub-space. The bases used to perform the projections corresponds to the so-called Proper Orthogonal Modes, calculated applying the Proper Orthogonal Decomposition in its snapshot-based version. The formulation is arranged in a way that it does not depend on which is the FE model used to discretize the structure and how it has been implemented, but only on some damage indexes and on the stiffness matrices of appropriate reference structures. Moreover, a dual estimation of reduced states and parameters, together with the update of the subspace, allows to track both the dynamic evolution of the system and the damage parameters. In order to assess the aforementioned identification procedure, the estimation strategy is applied to a thin plate. It is shown that the filter is able to identify and localize the damage even using a very reduced system. Finally, a proposal for a smart embedded data acquisition system, that could be used in coupling with the identification strategy, is presented. The choice of MEMS type accelerometers allow to make the system applicable to lightweight and small structures.

Sommario

Negli ultimi anni, grazie allo sviluppo di nuove tecnologie, una crescente attenzione è stata posta sul tema delle cosiddette città intelligenti. In questo senso, in ogni campo dell'ingegneria, si è osservato un crescente interesse per lo sviluppo di sistemi e metodologie di monitoraggio, inteso come il controllo e l'osservazione della realtà attraverso una rete di sensori. Il monitoraggio strutturale si riferisce a tutte quelle procedure che consentano di ricavare infromazioni sullo stato di un certo sistema meccanico o strutturale. Gli scopi principali sono la rilevazione, la localizzazione e l'identificazione del danno, nonchè, in alcuni casi, la stima della vita residua della struttura stessa. Dal punto di vista metodologico, tre sono i passi da seguire: innanzitutto lo stato del sistema viene osservato per un certo periodo di tempo attraverso una rete di sensori; successivamente, è necessario scegliere quei parametri che permettano di identificare il danno; infine gli stessi vengono stimati attraverso un appropriato modello matematico.
I primi tentativi atti a sviluppare un sistema di monitoraggio strutturale furono fatti nell'ambito dell'ingegneria petrolifera e delle piattaforme off-shore. Successivamente le prime vere applicazioni su larga scala riguardarono solamente un limitato numero di infrastrutture, in particolare ponti, dove il costo di un sistema di monitoraggio poteva essere giustificato dall'importanza dell'opera e dalla possibilità di ridurre gli elevati costi di manutenzione, grazie ad un miglioramento della gestione. Solamente in anni recenti e grazie al rapido sviluppo di sensori e reti di dati a basso costo e di dimensioni ridotte, le pratiche del monitoraggio strutturale si sono diffuse anche in altri campi quali le applicazioni aeronautiche e le opere civili di importanza non strategica. Parallelamente all'aspetto tecnico, sono stati presentati nuovi metodi matematici di identificazione che si sono affiancati a quelli tradizionali. I più utilizzati sono quelli basati sul confronto delle frequenze, della forma dei modi, degli spostamenti o di altri parametri modali.
Lo scopo di questa tesi è l'utilizzo e l'applicazione di un nuovo metodo di identificazione strutturale basato sui filtri ricorsivi Bayesiani. Lo scopo finale è quello di riuscire a comprendere se in una certa struttura si sia verificata una variazione delle caratteristiche di rigidezza e quantificare questa variazione. In altri termini, a partire da un certo numero di misurazioni, si vuole risalire alle caratteristiche del sistema e quindi stimare l'evoluzione dei parametri di danno.

Il metodo proposto risponde ad altri due requisiti. Innanzitutto il costo computazionale deve essere ridotto, in modo tale da poter effettuare l'identificazione in tempo reale o quasi reale. In secondo luogo, la formulazione del modello deve essere svolta in modo tale che dipenda solamente dai parametri di danno da stimare e non da come è stato costruito il modello agli elementi finiti. In questo modo la procedura può rappresentare una strategia teoricamente applicabile a qualsiasi tipo di struttura. Dal momento che lo svantaggio principale dei filtri Bayesiani è correlato al tempo di elaborazione dei dati, due tecniche sono state utilizzate al fine di ridurre il costo computazionale: innanzitutto è stato applicato un metodo di riduzione del modello, al fine di diminuire il numero dei gradi di libertà necessari per descrivere esaustivamente lo stato del sistema; in secondo luogo, si sono utilizzati metodi di ricampionamento e una versione modificata di un filtro particellare. La riduzione del modello è stata effettuata proiettando lo spazio completo in cui il sistema strutturale evolve in un sottospazio di dimensioni inferiori. La base necessaria per effettuare la proiezione alla Galerkin viene calcolata utilizzando la decomposizione ortogonale propria (POD) che consente di ottenere i modi propri che minimizzano in modo ottimale l'errore dovuto all'approssimazione.
La tesi si sviluppa su due piani: uno teorico, l'altro applicativo.
La prima parte riguarda l'introduzione e spiegazione dei metodi di riduzione del modello, con particolare riferimento alla POD. Successivamente viene ampiamente spiegato il concetto di stima, di inferenza Bayesiana e dei relativi metodi risolutivi, in particolare il filtro di Kalman, il filtro di Kalman esteso, adottato nel caso di problemi non lineari, e il filtro particellare. Dopo aver accennato ad ulteriori concetti quali la stima duale e la stima di modelli ridotti, la prima parte si conclude con la spiegazione dell'algoritmo che verrà in seguito utilizzato per la fase di simulazione, e che permette la stima duale del modello ridotto e dei parametri di danno con aggiornamento del sottospazio.
La seconda parte riguarda la valutazione degli algoritmi esposti in precedenza nell'applicazione ad una semplice struttura di riferimento: una piastra semplicemente appoggiata caricata con una forza sinusoidale, con o senza danno. Attraverso il confronto della risposta strutturale nei casi di modello completo e ridotto si dimostra che il metodo di riduzione del modello può approssimare adeguatamente il problema anche utilizzando un esiguo numero di gradi di libertà residui. Successivamente il funzionamento del filtro ricorsivo viene valutato verificandone la capacità di stimare i parametri di danno, note un certo numero di misurazioni delle rotazioni nei nodi. Il comportamento del metodo di stima viene valutato assumendo varie differenti ipotesi: la dimensione del modello ridotto, le incertezze associate alle misurazioni e al modello, il numero di sensori e le condizioni iniziali.

Si descrive infine una proposta di sviluppo di un sistema di acquisizione dei dati che possa essere applicato al monitoraggio di strutture leggere e di piccole dimensioni, nonchè con costi ridotti. La soluzione proposta prevede l'utilizzo di accelerometri MEMS e di un linguaggio di programmazione commerciale che permetta la massima flessibilità.

Chapter 1

Introduction

In recent years, due to the growing development of the information technology and electronic engineering, the concept of smart city is starting to play a key role in the modern society. In these terms, in every field of engineering a newly interest in system monitoring, meant as the control and observation of the reality through a network of sensing devices, has been revived. Structural Health Monitoring (SHM) represents the direct application of the aforementioned concept to mechanical and structural fields. SHM could be referred to all those procedures that can provide information about the state of a certain mechanical or structural system. The main goals of SHM can be summarized as follows:

- load detection;
- damage detection;
- damage localization;
- damage identification;
- remaining lifetime prognosis.

Regarding the method needed to reach the latter goals, any strategy should always deal with basically three conceptual stages [7]:

1. The system state is observed through periodically spaced measurements collected for a certain period of time. We can refer to long-time observations when the user is interested in slow variations of the system, such as those induced by corrosion and erosion phenomena or short-time observations when impacts or high-frequency events have to be investigated.

2. A certain number of features or indexes are chosen in order to identify the damage, i.e. any variation in the geometric or material characteristics of the system that can affect its behavior. The sources of damage basically depend both on the types of material and on structure of the system, for instance delamination and debonding for composite materials [8].

3. The measured data is analyzed and elaborated in order to identify the damage. This step could be performed either using on-line or off-line procedures.

In Figure 1.1 a conceptual block diagram representation of the SHM paradigm is shown.

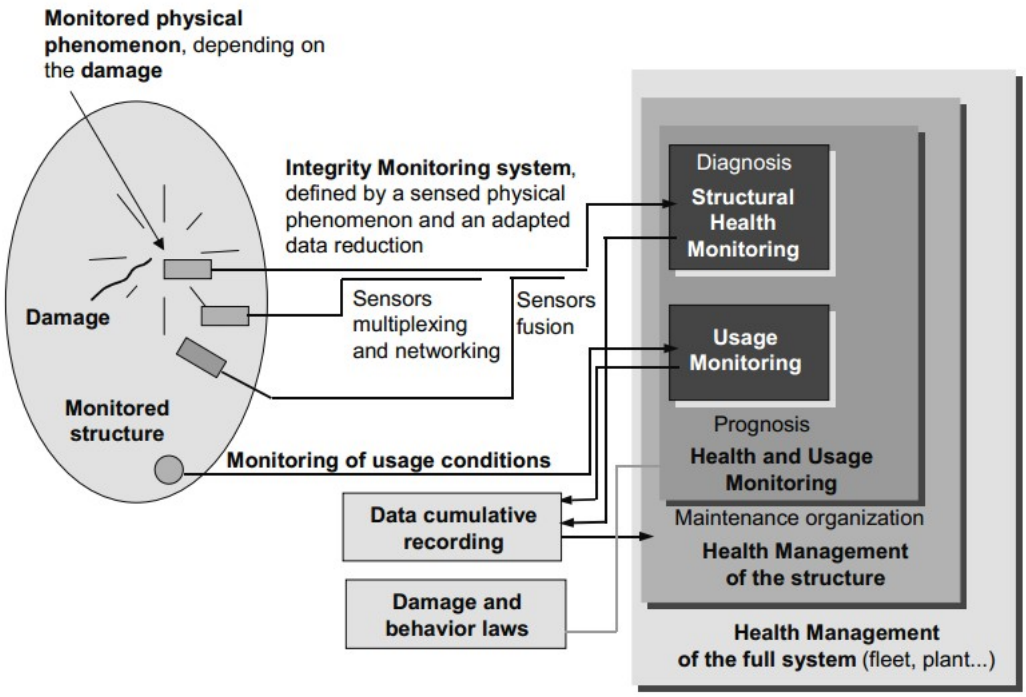


Figure 1.1: Block diagram representation of a SHM system [1]

The first efforts to try to develop some SHM systems were made by the oil industry during the 1980s in off-shore structures. The first large-scale applications in civil engineering were strategic infrastructures (WASHMS Hong Kong [9]) and bridges [10] where the investments were justified by the relevant social costs of a possible failure. It has been shown that a continuous control of the behavior of an infrastructure represents not only a way to understand and check the actual reliability of the system, but can also lead to a smarter management of the maintenance programs and hence to an effective decrease in overall lifetime costs [1]. Thanks to the development of off-the-shelf electronic devices and the reduction of cost of both the sensors [11] and the network systems [12], smart and cheap SHM systems can be now adopted for a wider range of structure and problems, such as aeronautical or low-scale civil applications.
Over the last years, several procedures and methods [13] have been developed in order to detect and estimate the structural damage. Some of these are natural frequency-based methods [14], mode shape-based methods [15], curvature/strain mode shape-based methods [16] or other modal parameters based methods [17].

Inverse methods, defined as procedures that allow to localize and detect damage using measured data, can be classified into two main groups [18]: sensitivity-type models and direct updating methods. While the first ones regards the estimation of some model parameters of the structure through the minimization of a penalty function, the second ones leads to a direct changing of the stiffness or mass matrix of a FE model, such that the measurement data are fitted.
The concepts of Bayesian estimation have been recently applied to system and damage identification of structural system. Some researchers have focused on the estimation of the parameters of simple mechanical systems or non-linear constitutive laws of materials. The identification of hysteretic mechanical models has been performed either through the unscented Kalman filter [19] or the extended Kalman filter [20]. Other works regard the estimation of the most likely structural model that guarantees the best accordance with the experimental data by using some Bayesian probabilistic frameworks [21, 22]. An other approach is the detection and localization of damage through the calculation of the probability that a certain set of model stiffness parameters is under a certain threshold [23, 24].
The aim of the present work is the assessment of a damage identification method based on recursive Bayesian filters. The algorithm is a modified version of the one proposed in [25]. The main goal of this identification method is the ability of detecting the damage indexes associated to any given structure, or in other words, estimating the actual local stiffness of the system, given a certain number of observations. Other two main objectives are to guarantee a reduced computational cost, such that the estimation could be performed in an on-line or nearly on-line way, and the coupling with a commercial state-of-the-art Finite Element (FE) code. These requirements allow the method to be presented as a conceptual strategy that could be applied to a large variety of applications purposes and to any type of FE discretization of the structure. Since the main drawback of any identification method based on standard recursive filters is exactly represented by the excessive computational time, two solutions will be adopted:

- a model order reduction that allows to reduce the dimensionality of the system, still maintaining a proper level of information;
- an improved particle filtering strategy, that uses a re-sampling technique and a modified particle filter.

The model order reduction will be obtained through a Galerkin projection of the original full model into a sub-space. The bases used to perform the projections corresponds to the so-called Proper Orthogonal Modes (POM), calculated applying the Proper Orthogonal Decomposition (POD) in its snapshot-based version. Regarding the use of a commercial code in order to generate the structural model, the formulation will be arranged in a way that it does not depend on which is the model used to discretize the structure and how it has been implemented. Rather, the model depends only on some damage indexes associated to a certain number

of regions in which the structure is divided to and the definition of some appropriate reference structures. Therefore, in order to be able to estimate both the dynamical evolution of the system and the damage parameters, a dual estimation of reduced states and parameters, together with the update of the subspace, will be employed. In order to assess the aforementioned identification procedure, the identification strategy will be applied to a simple thin plate. It will be shown that the filter is able to identify and localize the damage even using a very reduced system. For instance, the behavior of a 10×10 elements thin plate, having 726 degrees of freedom, is very well approximated by a reduced model with only three degrees of freedom. Finally, a proposal for a smart embedded data acquisition system, that could be used in couple with the identification strategy, will be presented. Since we wanted the system to be applicable even to light and small structures, the main requirements were a low cost, lightweight and tiny sizes of the devices and a flexible deployment and control of the acquired data. For these reasons, both a Micro Electro-Mechanical Systems (MEMS) type accelerometer and a commercial Labview programming interface will be chosen.

1.1 Outline of the thesis

The thesis is divided into two main parts: first of all a theoretical explanation of the identification procedure is given, then the results of a simulated benchmark analysis are shown.
Chapter 2 regards the theoretical framework of the model order reduction method and its different implementation strategies. A particular attention is paid to the Proper Orthogonal Decomposition (POD), that will be used in the simulations. Its basic formulation is explained showing that the optimal projection needed to reduce the original space is the one that minimize the overall error of the relative approximation. Moreover, the projecting bases calculated are called Proper Orthogonal Modes (POM). Afterwards, a brief review of the two methods used to perform POD are described: the Principal Component Analysis (PCA) and the Singular Value Decomposition (SVD). Since the latter one will be used, a conceptual interpretation is given and the link between POD and SVD is underlined. Later a energy-based methodology that allows to choice the appropriate order of the reduced system is explained. Finally the Galerkin projection-based formulation of the structural dynamical equation is furnished, as well as its related errors. At the end the physical interpretation of the POMs is given, considering some basic loading conditions.
Chapter 3 deals with the explanation of the recursive Bayesian filter concept. A first look at the general estimation problem formulation is given, explaining the meaning of state of a system. Then, the process and measurement models are defined: while the first one is needed to describe the evolution of the system over time, the latter one relates the state and the measurements collected through a

network of sensors. The concept of Bayesian inference is furnished, explaining the three basic steps of any estimation procedure: definition of the initial conditions, prediction stage and update stage. An optimal solution of the estimation problem is the Kalman filter (KF): for completeness, its demonstration and the related procedure is here reported. Whenever the model could be no more considered as linear, a sub-optimal method is required: hence, a linearization is necessary and the extended Kalman filter (EKF) can be applied. When the linearization is not a viable solution, a new strategy should be used: here the concept of Particle filtering (PF) and re-sampling procedures are introduced.
Chapter 4 collects the concepts previously explained, by introducing the final framework that will be used in the simulations. The notions of dual estimation and reduced model estimation are required since both the dynamical evolution of the system and the damage parameters will be estimated. Finally the Hybrid extended Kalman particle filter (HEKPF) is explained: the idea is to use a Kalman filter to speed up the particle filter algorithm by updating the samples considering the actual observations. Gathering all the latter concepts into a unique estimation scheme, the dual estimation of reduced states and parameters with subspace updating algorithm (EK-PF-KF) is described. The final step is the application of this procedure to the structural problem, implementing a new way of building the process model such that a commercial FE code can be used.
In chapter 5 a benchmark analysis has been shown, applying the procedure to a simply supported thin plate subjected to a smooth force, considering both the cases of an undamaged and damaged structure. The assessment of the model order reduction is given, showing that even a very reduced model can catch the main behavior of the full model. It is then proved that, chosen a certain level of information retained, the damaged structure requires a higher number of degrees of freedom with respect to the undamaged case. The convergence and stability of the POMs with respect to the number of snapshots that have been selected is shown as well. Afterwards, the capability of the filter of detecting and estimating the damage parameters, given a certain number of rotational observations, is shown. Its sensitivity with respect to the choice of the model size, the measurement noise, the process noise, the initial conditions, the number of sensors and the convergence of the POMs over time are discussed. Finally, it is shown that the operation of the damage identification procedure is guaranteed for systems with large number of degrees of freedom as well.
In chapter 6 a proposal for an embedded SHM data acquisition system is given. The choice of MEMS accelerometers and a ready-to-use commercial programming language is made in order to develop a very low cost system. Finally, some considerations about the optimal placement of the accelerometers are given.
Chapter 7 deals with some concluding remarks and possible future developments and researches.

Chapter 2

Model Order Reduction

2.1 Overview

Let us first of all consider a generic structure subjected to a set of variable forces and boundary conditions. The Principle of Virtual Work states that the work done by externally applied forces is equal to the sum of the work done by inertial, dissipative and internal forces for any virtual displacement, i.e. for any small motion that satisfies compatibility and essential boundary conditions. Moreover, in order the problem to be manageable, the structure can be discretized through the application of the Finite Element Method (FEM). Applying this latter procedure to the aforementioned theorem, we are lead to the dynamical equation of a given structural problem:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{D}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{F}(t) \tag{2.1}$$

where $\mathbf{u}(t)$ is the vector that contains the kinematic quantities, either displacements or rotations, at a certain instant t and associated to each node of the structure, $\mathbf{M}$ is the mass matrix, $\mathbf{D}$ is the damping matrix, $\mathbf{K}$ is the stiffness matrix, $\mathbf{F}$ is the vector of the external applied forces. The notations $\ddot{\mathbf{u}}(t)$ and $\dot{\mathbf{u}}(t)$ stand for respectively the first and second derivative of vector $\mathbf{u}(t)$ with respect to t. The dimension of the problem depends directly on the type and number of elements that have been used to space-discretize the system itself.
One of the goals of the present work is the description of a procedure that can work in real-time or near real-time: despite the use of a powerful and fast filtering scheme and of a simple geometry, since we want the system to be effective even for other and more complicated type of structures, a model reduction is performed. Model reduction, sometimes also called model order reduction (MOR), refers to all those mathematical methods that allow to reduce the dimensionality of a complex high-dimension model into another simplified model. Any method used to perform the reduction should:

- minimize the error between the full and reduced order models;

- maintain an acceptable level of information;
- describe exhaustively the behavior of the system.

This reduction is usually obtained through the projection of the original space, in which the mechanical system can evolve by means of $\mathbf{u}(t)$, $\dot{\mathbf{u}}(t)$ and $\ddot{\mathbf{u}}(t)$, onto a new lower order subspace, trying to lose the less possible information. For instance, considering a dynamic problem, equation (2.1) will become:

$$\mathbf{M}_r\ddot{\mathbf{v}}(t) + \mathbf{D}_r\dot{\mathbf{v}}(t) + \mathbf{K}_r\mathbf{v}(t) = \mathbf{F}_r(t) \tag{2.2}$$

where $\mathbf{u} \in \mathbb{R}^n$ with n being the dimension of the full model (i.e. the number of degrees of freedom for a FE model) and $\mathbf{v} \in \mathbb{R}^l$ with $l << n$ being the dimension of the reduced model; of course, similarly, $\mathbf{M}, \mathbf{D}, \mathbf{K} \in \mathbb{R}^{n\times n}$ and $\mathbf{M}_r, \mathbf{D}_r, \mathbf{K}_r \in \mathbb{R}^{l\times l}$. Since $\mathbf{v}$ is an approximate solution that leads to a certain grade of error with respect to the exact solution $\mathbf{u}$ of the problem, the main aim of all the reduction methods is to find an optimal subspace such that the approximation is still acceptable.
The most famous projection method is based on the Krylov sub-spaces [26], which led to several algorithms, such as Arnoldi, Lanczos, Conjugate gradient ones.
Two types of procedures are used nowadays:

- proper orthogonal decomposition based methods (POD);
- proper generalized decomposition based methods (PGD);
- moment matching based methods.

The proper generalized decomposition consists in building up the solution of a full order problem as a sum of tensorial product of functions defined in some sub-spaces with reduced dimension [27, 28].
Regarding the latter ones, having redefined equation (2.1) in a state-space form in terms of its state matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ (related to matrices $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ [29]) and its transfer function $H(s)$, the general idea is basically to impose the first l moments of $H(s)$ to be equal to those of the reduced order transfer function in a certain set of points $\mathbf{s}_k$. Since any direct calculation is numerically unstable [30], the calculation is made by using an iterative procedure based on Krylov sub-spaces.
In the present work, the model reduction will be performed using the so-called singular value decomposition that can be considered as a POD based method. As it will be explained in details in the following sections, the steps performed to reduce the model can be summarized as follows:

1. selection of a certain ensemble of data;
2. calculation of the proper orthogonal modes (POM) of the structure through a POD-based method;

3. projection of the model in a new low-dimension subspace.

2.2 POD

The concept of POD were first introduced by Kosambi (1943, [31]), Loeve (1941, [32]), Karhunen (1946, [33]), Pougachev (1953, [34]) and Obukhov (1954, [35]). The POD was developed in different formulations among different fields of research; it is called Principal Component Analysis (PCA) in statistical mathematics (Jolliffe, 1986, [36]), Karhunen-Loeve Decomposition (KLD) in signal processing and Singular Value Decomposition in mechanical engineering (Mees, Rapp & Jennings, 1987, [37]).
The main idea of the POD is to find a subspace approximating a given set of data in an optimal least-squares sense [38, 39, 40].
Consider the vector $\mathbf{u}(t)$ as the state evolution defined in the time interval $0 \leq t \leq T$ and obtained either through experimental data or as a solution of the full model. We want to find an appropriate orthogonal projection $\mathbf{\Pi}_l : \mathbf{V} \to \mathbf{V}_l$ of the data space $\mathbf{V}$, such that the following L^2 norm is minimized:

$$J(\mathbf{\Pi}_l) = \int_0^T ||\mathbf{u}(t) - \mathbf{\Pi}_l \mathbf{v}(t)||_2^2 \, dt \tag{2.3}$$

with $\mathbf{\Pi}_l \in \mathbb{R}^{n \times l}$ s.t. $\mathbf{\Pi}_l \mathbf{\Pi}_l^T = \mathbf{I}$. In other words, the aim is to find the appropriate l-sized sub-space that approximates the original complete space: the task is reached minimizing the error between these ones.
In order to find the solution of the problem, we first need to introduce the correlation matrix $\mathbf{K} \in \mathbf{R}^{n \times n}$ as [41]:

$$\mathbf{K} = \int_0^T \mathbf{u}(t)\mathbf{u}(t)^T \, dt \tag{2.4}$$

It can be prooved that $\mathbf{K}$ is a symmetric positive semi-definite matrix and therefore its eigenvalues are real and nonnegative.
The POD states then that the optimal subspace $\mathbf{V}_l$ representing the data is given by:

$$\mathbf{V}_l = span\{\boldsymbol{\phi}_i, ..., \boldsymbol{\phi}_l\} \tag{2.5}$$

The vectors $\boldsymbol{\phi}_k$ are called either Proper Orthogonal Modes or POD Modes and they are defined as the eigenvectors of $\mathbf{K}$, i.e. the solution of the associated eigenvalue problem (further details in [42]):

$$\mathbf{K}\boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i \tag{2.6}$$

Considering real applications, the dimension n of the space can be very large, therefore the calculation of the correlation matrix $\mathbf{K}$ can become computationally

inapplicable. In order to overcome the problem, a new method has been proposed by Sirovich [43]: the snapshot version of POD consists of a simplified solution of the previous eigenvalue problem by solving only a $m \times m$ system, with $m < n$.
First of all, we have to build the snapshot matrix, by collecting a certain number m of states in an evenly spaced time interval. The integral used in equation (2.4) to calculate the correlation matrix can be therefore approximated through:

$$\mathbf{K} = \sum_{i=1}^{N_{snap}} \mathbf{u}(t_i)\mathbf{u}(t_i)^T \tag{2.7}$$

Let us now define the snapshot matrix:

$$\mathbf{U} = (\mathbf{u}(t_i), ..., \mathbf{u}(t_{N_{snap}})) \in \mathbb{R}^{n \times N_{snap}} \tag{2.8}$$

Using the vectorial form, the correlation matrix can be simply written as $\mathbf{K} = \mathbf{U}\mathbf{U}^T$, however $\mathbf{K}$ is still defined in $\mathbb{R}^{n \times n}$. At this point, it is important to notice that the non-zero eigenvalues associated to $\mathbf{U}\mathbf{U}^T$ and the ones for $\mathbf{U}^T\mathbf{U}$ are the same; this trick is useful because now it is possible to switch to a side simpler eigenvalue formulation and thus reduce the computational cost:

$$\left(\mathbf{U}^T\mathbf{U}\right)\boldsymbol{\psi}_i = \lambda_i \boldsymbol{\psi}_i \tag{2.9}$$

where $\mathbf{U}^T\mathbf{U} \in \mathbb{R}^{N_{snap} \times N_{snap}}$ and $N_{snap} \ll n$.
Denoting $r = rank(\mathbf{U}^T\mathbf{U})$, the first r POMs can be calculated in the following way:

$$\boldsymbol{\phi}_i = \frac{1}{\sqrt{\lambda_i}}\mathbf{U}\boldsymbol{\psi}_i, \ i = 1, ..., r \tag{2.10}$$

In order to sum up the previous procedure, we can say that the POD is a method used to project a certain model onto a lower order sub-space built using the eigenvectors associated to the correlation matrix calculated from the set of given data.
As the theoretical explanation of POD has been presented, the final step is the calculation of the optimal orthogonal basis that solves the problem. Several different methods exists: the most used ones are the Principal Component Analysis (PCA), the Singular Value Decomposition (SVD) and the Karhunen-Loeve Decomposition (KLD). In [44] and [45] the equivalence of all these methods was shown.
In the following paragraphs a brief review of the first two procedures will be given.

2.2.1 PCA

The PCA is a statistical method that allows to condensate big amounts of data into few principal components; it was proposed by Pearson (1901, [46]) and developed by Hotelling (1933, [47]).
It is important to point out that a generic given set of data consists of a large number of interrelated variables, so the basic idea of the procedure is to build a new set of data that consists of ordered uncorrelated variables, obtained by looking for those derived variables that preserve most of the information [36]. In these terms, the method could be even seen as an optimization problem, such as either a variance maximization or least-mean-square problem.
Let $\mathbf{u} \in \mathbb{R}^n$ be a random vector, as defined previously, and $y_1, y_2, ..., y_n \in \mathbb{R}$ be the principal components, defined as follows:

$$y_1 = \sum_{i=1}^{n} \alpha_{i1} u_i = \boldsymbol{\alpha}_1^T \mathbf{u} \tag{2.11}$$

with $\boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{21}, ..., \alpha_{n1})^T$ is a constant vector and the variance of y_1 is:

$$\sigma_{y_1}^2 = \boldsymbol{\alpha}_1^T \boldsymbol{\Sigma}_{\mathbf{u}} \boldsymbol{\alpha}_1 \tag{2.12}$$

$\boldsymbol{\Sigma}_{\mathbf{u}} \in \mathbb{R}^{n \times n}$ is the semi-definite covariance matrix of the random vector $\mathbf{u}$, defined as ($E(\mathbf{u})$ is the expectation of $\mathbf{u}$):

$$\boldsymbol{\Sigma}_{\mathbf{u}} = E[(\mathbf{u} - E[\mathbf{u}])(\mathbf{u} - E[\mathbf{u}])^T] \tag{2.13}$$

As previously explained, since the aim is to maximize the variation of the values, we seek for that value of $\boldsymbol{\alpha}_1$ for which the variance $\sigma_{y_1}^2$ is maximum, assuming $\boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_1 = 1$. The normalization condition is needed, otherwise the stand-alone optimization problem would be indeterminate.
In order to solve this contrained problem, we use the method of Lagrange multipliers:

$$L(\boldsymbol{\alpha}_1, \lambda_1) = \boldsymbol{\alpha}_1^T \boldsymbol{\Sigma}_{\mathbf{u}} \boldsymbol{\alpha}_1 + \lambda_1 (1 - \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_1) \tag{2.14}$$

The derivative with respect to $\boldsymbol{\alpha}_1$ is then calculated and enforced to zero, obtaining:

$$\boldsymbol{\Sigma}_{\mathbf{u}} \boldsymbol{\alpha}_1 = \lambda_1 \boldsymbol{\alpha}_1 \tag{2.15}$$

The latter equation is then a simple eigenvalue problem, where λ_1 and $\boldsymbol{\alpha}_1$ are respectively the eigenvalue and the corresponding eigenvector of the matrix $\boldsymbol{\Sigma}_{\mathbf{u}}$. Since the variance should be as large as possible, then the maximum eigenvalue is chosen.
Avoiding the case of a zero value eigenvalue, that would lead to consider $\mathbf{u}$ no more as a random vector, but as a constant vector, we can procede to the calculation of $\boldsymbol{\alpha}_2$, in order to find the second principal component. The procedure

is basically the same used for the first one, even if a further condition should be enforced in order for second component to be uncorrelated with respect to the first component:

$$\boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_1 = 0 \tag{2.16}$$

Extending the technique to any principal component, we find that the i-th principal component of $\mathbf{u}$ is $y_i = \boldsymbol{\alpha}_i^T \mathbf{u}$ where $\boldsymbol{\alpha}_i$ is the i-th eigenvector of the covariance matrix of $\mathbf{u}$ and the variance of y_i is the i-th eigenvalue of $\boldsymbol{\Sigma}_{\mathbf{u}}$. Furthermore, the maximization process is done in a way that all the principal components and the respective variables are ordered following the eigenvalue arrangement.
Once all the $\boldsymbol{\alpha}$ coefficients are given, it can easily be proved that the random variable $\mathbf{u}$ can be expressed as a linear combination of the principal components:

$$\mathbf{u} = \sum_{i=1}^{n} \boldsymbol{\alpha}_i y_i \tag{2.17}$$

Comparing this equation with the POD, it is straightforward to understand that the orthonormal basis of the POD has been found and corresponds to the eigenvectors of the covariance matrix of the random vector $\mathbf{u}$. In other words the PCA is simply a method used to find the basis that satisfy the POD requirements.
The Karhunen-Loeve Decomposition [33] extends the PCA to the case of infinite-dimensional spaces, such as the space of continuous-time functions [44].

2.2.2 SVD

First of all, a brief theoretical explanation and definition of the SVD will be given; in the next paragraph the link between SVD and POD will be shown.
Consider the matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$; it can be proved that there exist two orthogonal matrices $\mathbf{V} \in \mathbb{R}^{m \times m}$ and $\mathbf{W} \in \mathbb{R}^{n \times n}$ such that:

$$\mathbf{A} = \mathbf{V} \boldsymbol{\Sigma} \mathbf{W}^T \tag{2.18}$$

with $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ defined as a pseudo-diagonal matrix, with:

$$\boldsymbol{\Sigma}_{ii} = \sigma_i \tag{2.19}$$

wherein $\{\sigma_i\}_{i=1}^{min(m,n)}$ are called singular values of $\mathbf{A}$ and they are ordered decreasingly. In the same way, the columns $\mathbf{v}_i \in \mathbb{R}^m$ and $\mathbf{w}_i \in \mathbb{R}^n$ of the matrices $\mathbf{V}$ and $\mathbf{W}$ are respectively the left and right singular vectors of $\mathbf{A}$.
The decomposition provides useful information about the matrix $\mathbf{A}$ [44]:

- σ_i^2 are the eigenvalues of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$;
- $\mathbf{A}\mathbf{w}_i = \sigma_i \mathbf{v}_i$ and $\mathbf{A}^T\mathbf{w}_i = \sigma_i \mathbf{w}_i$, so $\mathbf{v}_i$ and $\mathbf{w}_i$ are eigenvectors of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$;

- the rank of $\mathbf{A}$ is equal to the r index of the smallest non-zero singular value;
- being $\mathbf{V}_r = [\mathbf{v}_1 \cdots \mathbf{v}_r]$ and $\mathbf{W}_r = [\mathbf{w}_1 \cdots \mathbf{w}_r]$ the singular vectors associated to the non-zero singular values and $\mathbf{V}_{m-r} = [\mathbf{v}_{r+1} \cdots \mathbf{v}_m]$ and $\mathbf{W}_{n-r} = [\mathbf{w}_{r+1} \cdots \mathbf{w}_n]$ the singular vectors associated to the zero singular vectors, then $\mathbf{A} = \sigma_1 \mathbf{v}_1 \mathbf{w}_1^T + ... + \sigma_r \mathbf{v}_r \mathbf{w}_r^T = \sum_{i=1}^{r} \sigma_i \mathbf{v}_i \mathbf{w}_i^T$.

As it will be explained later on for the POD, the SVD can be applied to optimization problems. The aim is to minimize the 2-norm $||\mathbf{A} - \mathbf{X}||_2$, where $\mathbf{X} \in \mathbb{R}^{m \times n}$ and $\text{rank}(\mathbf{X}) = k < r = \text{rank}(\mathbf{A})$. As explained in [41], using the Schmidt-Eckart-Young-Mirsky theorem, it is possible to say that the matrix $\mathbf{X}$ such that the quantity:

$$||\mathbf{A} - \mathbf{X}||_2 = \sigma_{k+1}(\mathbf{A}) \tag{2.20}$$

is minimum, with $\sigma_k(\mathbf{A}) > \sigma_{k+1}(\mathbf{A})$, is:

$$\mathbf{X} = \sum_{i=1}^{k} \sigma_i \mathbf{v}_i \mathbf{w}_i^T \tag{2.21}$$

Interpretation of SVD The SVD of $\mathbf{A}$ provides useful information about the oriented energy of the matrix itself [48]:

1. the energy associated to the columns of $\mathbf{A}$ is equal to the energy in its singular spectrum $\sigma_1^2, ..., \sigma^2_{\min(m,n)}$:

$$En(\mathbf{A}) = ||\mathbf{A}||_F^2 = \sum_{i=1}^{n} \sum_{i=1}^{m} x_{ij}^2 = \sum_{k=1}^{\min(m,n)} \sigma_k^2 \tag{2.22}$$

2. the oriented energy of a vector sequence:

$$En_k(\mathbf{A}) = \sum_{i=1}^{m} (\mathbf{e}_k^T \mathbf{a}_i)^2 \tag{2.23}$$

   is maximum at each left singular vector direction [49].

This means that the oriented energy in the direction of the k-th left singular vector is equal to square of the k-th singular value. As it will be explained in the following section, the POMs are equal to the left singular vectors and therefore it is possible to conclude that they represent the optimal set of bases in terms of energy content; in other words, it does not exist another set of vectors that can retain more energy per each mode [48].

2.2.3 POD and SVD

The SVD can be used to perform the POD and hence reduce the model. Let us first of all recall the snapshot matrix $\mathbf{U}$ defined in (2.8); let us retrieve from the previous paragraph the eigenvalues associated to the correlation matrix $\mathbf{K} = \mathbf{U}\mathbf{U}^T$ of $\mathbf{U}$ and order them decreasingly:

$$\lambda_1 \geq ... \geq \lambda_r > \lambda_{r+1} = ... = \lambda_n = 0 \tag{2.24}$$

If we recall the SVD pattern (2.18), we can replace all the matrices and parameters defined for this specific case [41]:

$$\mathbf{A} = \mathbf{U}$$
$$\sigma_i = \sqrt{\lambda_i}$$
$$\mathbf{V} = \mathbf{\Phi}$$

The equation (2.10) then becomes:

$$\boldsymbol{\phi}_i = \frac{1}{\sigma_i}\mathbf{U}\boldsymbol{v}_i, \ i = 1, ..., r \tag{2.25}$$

Define $\mathbf{\Phi} = [\mathbf{\Phi}_1, \mathbf{\Phi}_2]$ with $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ being respectively the collections of the eigenvectors associated to the first r non-zero eigenvalues $\lambda_1, \lambda_2 \geq ... \geq \lambda_r$ and to the remaining ones $\lambda_{r+1}, \lambda_{r+2} \geq ... \geq \lambda_n$. From equation (2.9), the problem can be written as an eigenvalues formulation:

$$\mathbf{U}^T\mathbf{U}\mathbf{\Phi} = \mathbf{\Phi}\mathbf{\Lambda} \tag{2.26}$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose principal components are the eigenvalues of $\mathbf{K}$. According to the eigenvalue formulation and the equivalences shown before, the singular values of $\mathbf{U}$ are equal to the square roots of the eigenvalues of $\mathbf{U}\mathbf{U}^T$ and $\mathbf{U}^T\mathbf{U}$. Moreover, the left and right singular vectors of $\mathbf{U}$ are the eigenvectors of $\mathbf{U}\mathbf{U}^T$ and $\mathbf{U}^T\mathbf{U}$ respectively.
Multiplying the equation by $\mathbf{\Phi}^T$ and recalling the definition of matrix $\mathbf{\Sigma}$ from the previous paragraph, we get [44]:

$$[\mathbf{\Phi}_1, \mathbf{\Phi}_2]^T\mathbf{U}^T\mathbf{U}[\mathbf{\Phi}_1, \mathbf{\Phi}_2] = \begin{bmatrix} \mathbf{\Sigma}_r^2 & 0 \\ 0 & 0 \end{bmatrix} \tag{2.27}$$

Having defined $\boldsymbol{W}_1 = \mathbf{U}^T\mathbf{\Phi}_1{\mathbf{\Sigma}_r}^{-1}$ it is straightforward to prove that $\boldsymbol{W}_1^T\boldsymbol{W}_1 = 1$ and therefore the columns of $\boldsymbol{W}_1$ are mutually orthogonal. It follows that there exists n-r vectors which are orthonormal with respect to both each other and the columns of $\boldsymbol{W}_1$. Let us store these n-r vectors into a new orthonormal matrix $\boldsymbol{W}_2$ and calculate the following expression:

$$\boldsymbol{W}^T\mathbf{U}^T\mathbf{\Phi} = \begin{bmatrix} \boldsymbol{W}_1^T\mathbf{U}^T\mathbf{\Phi}_1 & \boldsymbol{W}_1^T\mathbf{U}^T\mathbf{\Phi}_2 \\ \boldsymbol{W}_2^T\mathbf{U}^T\mathbf{\Phi}_1 & \boldsymbol{W}_2^T\mathbf{U}^T\mathbf{\Phi}_2 \end{bmatrix} \tag{2.28}$$

Thanks to simple algebraic considerations, (2.28) becomes:

$$\boldsymbol{W}^T\mathbf{U}^T\boldsymbol{\Phi} = \begin{bmatrix}\boldsymbol{\Sigma}_r^2 & 0\\ 0 & 0\end{bmatrix} \tag{2.29}$$

and therefore:

$$\mathbf{U}^T = \boldsymbol{W}\begin{bmatrix}\boldsymbol{\Sigma}_r^2 & 0\\ 0 & 0\end{bmatrix}\boldsymbol{\Phi}^T \tag{2.30}$$

Transposing the previous equation, we get the SVD formulation:

$$\mathbf{U} = \boldsymbol{\Phi}\begin{bmatrix}\boldsymbol{\Sigma}_r^2 & 0\\ 0 & 0\end{bmatrix}\boldsymbol{W}^T \tag{2.31}$$

Comparing (2.31) with (2.18) the link between POD and SVD is displayed.
The two methods basically leads to the same results, but, since the snapshot-based POD requires the calculation of the eigenspectrum of the correlation matrix $\mathbf{K}$, its condition number is $\text{cond}(\mathbf{K}) = \text{cond}(\mathbf{U})^2 > \text{cond}(\mathbf{U})$ and thus it can lead to computational issues [41]. High values of condition number means that for a small variation of the input data, i.e. the snapshot matrix, a large variation in the POMs obtained can occur.
It is interesting to point out that naming the columns of $\begin{bmatrix}\boldsymbol{\Sigma}_r^2 & 0\\ 0 & 0\end{bmatrix}\boldsymbol{W}^T$ as $\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_n$, it follows:

$$\mathbf{u}_i = \boldsymbol{\Phi}\mathbf{d}_i$$

From the previous equations it is possible to see that the $r+1, ..., n$ components of each column $\mathbf{d}_i$ are equal to zero if the $r+1, ..., n$ singular values of $\mathbf{U}^T$ are zero; therefore the samples can be represented using only the first r singular vectors.
It can be proved that the basis chosen using the SVD is the optimal one, that is the error due to the decrease of the space order is as small as possible [44]. The error of all samples is defined as:

$$\epsilon^2(l) = \sum_{i=1}^{N_{snap}} ||\mathbf{u}_i - \mathbf{u}_i(l)||^2 = ||\mathbf{U}^T\boldsymbol{\phi}_{n-1}||_r^2 \tag{2.32}$$

The optimal basis vectors is the one that minimize the error:

$$\min\ \epsilon^2(l) = \sum_{j=l+1}^{m} \boldsymbol{\phi}_j^T\mathbf{U}\mathbf{U}^T\boldsymbol{\phi}_j^T \ \ s.t.\ \boldsymbol{\phi}_i^T\boldsymbol{\phi}_j = \delta_{ij} \tag{2.33}$$

where δ_{ij} is the Kronecker delta defined as follows:

$$\delta_{ij} = \begin{cases}0, & \text{if } i \neq j\\ 1, & \text{if } i = j\end{cases} \tag{2.34}$$

Applying the Lagrangian method to solve the extreme value problem, we lead to [44]:

$$\mathbf{U}\mathbf{U}^T\boldsymbol{\phi}_{n-1}\mathbf{P} = \boldsymbol{\phi}_{n-1}\mathbf{P}\boldsymbol{\Lambda} \tag{2.35}$$

where $\mathbf{P}$ is an appropriate orthogonal matrix. Suppose to collect $\boldsymbol{\phi}_{n-1}\mathbf{P}$ in matrix, the expression is identical to (2.26) and then the SVD provides excatly the optimal subspace.
The error due to the truncation of the POD basis can be actually estimated with the formula:

$$\epsilon^2(l) = tr((\mathbf{U}^T\boldsymbol{\phi}_{n-1}\mathbf{P})^T\mathbf{U}^T\boldsymbol{\phi}_{n-1}\mathbf{P}) = tr(\Lambda) \tag{2.36}$$

In the following section more details on this topic will be given.

2.2.4 Choice of dimension l

A critical point of all the reduction methods is the choice of the new sub-space size l; this parameter determines whether a method is usable in terms of accuracy and computational cost. Moreover, the most important issue is the level of information retained into the residual data furnished by the subspace, such that it can be possible to decide if the approximation is "good" and acceptable.
A first hint on this topic can be found in equation (2.36) where an estimation of the overall error associated to the reduction of order is given; nevertheless, $\epsilon^2(l)$ is an a posteriori information and it cannot be used to determine l.
A general idea to overcome the problem is to employ the 2-norm (also called Frobenius norm) of matrix $\mathbf{U}$ defined as:

$$||\mathbf{U}||_F = \sqrt{\sum_{i=1}^{n} \lambda_i(\mathbf{U})} \tag{2.37}$$

wherein $\lambda_i(\mathbf{U})$ are the eigenvalues of the correlation matrix. The error associated to the projection onto the sub-space is:

$$||\mathbf{U} - \Pi_l\mathbf{U}||_F = \sqrt{\sum_{i=1}^{l} \lambda_i(\mathbf{U})} \tag{2.38}$$

Therefore the relative error, sometimes even called relative information content $I(l)$ [50], is:

$$I(l) = 1 - \epsilon_r(l) = \frac{||\mathbf{U} - \Pi_l\mathbf{U}||_F^2}{||\mathbf{U}||_F^2} = \frac{\sum_{i=1}^{l} \lambda_i(\mathbf{U})}{\sum_{i=1}^{n} \lambda_i(\mathbf{U})} \tag{2.39}$$

Remembering that the singular values are $\sigma_i = \sqrt{\lambda_i}$, then:

$$I(l) = \frac{\sum_{i=1}^{l} \sigma_i^2}{\sum_{i=1}^{n} \sigma_i^2} \tag{2.40}$$

This means that if the required grade of information given by the projected sub-space were a percentage p, then l should be chosen such that $I(l) \geq \frac{p}{100}$. This type of criterion can be even seen as an energy-based method, in fact the Frobenius norm represents an index of the energy of the vector sequence collected in the matrix $\mathbf{U}$ and therefore $I(l)$ is the portion of the total energy captured by the first l POMs. For example, the total energy of a mechanical system (2.1) is:

$$E = \frac{1}{2}\mathbf{u}^T\mathbf{K}\mathbf{u} + \frac{1}{2}\ddot{\mathbf{u}}^T\mathbf{M}\ddot{\mathbf{u}} \tag{2.41}$$

Similarly, if the model is reduced, we get:

$$E_r = \frac{1}{2}\mathbf{v}^T\mathbf{K}_r\mathbf{v} + \frac{1}{2}\ddot{\mathbf{v}}^T\mathbf{M}_r\ddot{\mathbf{v}} \tag{2.42}$$

Moreover, in a dynamical system, large eigenvalues are always associated to large perturbations of the system itself. The approach based on the use of the eigenvalues is supported by the idea that the smaller are the values of the eigenvalues, the less is their influence on the behavior of the dynamic system.
There is another important remark: the POMs optimally approximate a certain set of data in a least-squares sense, but they are not built in order to optimally approximate the dynamics that generate that data. For example, in some types of dynamic systems, there could be low-energy POMs associated to very effective phenomena [42]; in those cases, other approaches are required.

2.3 Projection

Once the orthogonal bases have been calculated, the full model can be projected onto the sub-space. Let us retrieve the aforementioned notation, the time history which describes the evolution of the dynamic system can be written as a linear combination of vectors defined in full n dimensional space:

$$\mathbf{u}(t) = y_1(t)\boldsymbol{\phi}_1 + ... + y_n(t)\boldsymbol{\phi}_n = \sum_{i=1}^{n} \boldsymbol{\phi}_i y_i = \boldsymbol{\Phi}\mathbf{y} \tag{2.43}$$

where $\mathbf{u}(t), \mathbf{y}(t) \in \mathbb{R}^n$ and $\boldsymbol{\Phi} \in \mathbb{R}^{n\times n}$.
In order to decrease the computational cost, the state vector can be even expressed through a linear combination of a sub-space, whose size l is smaller than n:

$$\mathbf{u}_l(t) \approx \alpha_{l_1}(t)\boldsymbol{\phi}_1 + ... + \alpha_{l_l}(t)\boldsymbol{\phi}_l = \sum_{i=1}^{l} \boldsymbol{\phi}_i \alpha_{l_i} = \boldsymbol{\Phi}_l\boldsymbol{\alpha} \tag{2.44}$$

where $\mathbf{u}_l(t) \in \mathbb{R}^n$, $\boldsymbol{\alpha}(t) \in \mathbb{R}^l$ and $\boldsymbol{\Phi}_l \in \mathbb{R}^{n\times l}$ collects only the first l columns of $\boldsymbol{\Phi}$.
In this manner, we have defined the projection of the system using the optimal

basis found through the POD. α_i and y_i are called generalized coordinates and $\mathbf{u}_l(t)$ reduced state vectors. Since the projection used is supposed to be a linear operator, the derivatives of the state vector can be expressed in the same way:

$$\dot{\mathbf{u}}_l(t) \approx \dot{\alpha}_{l_1}(t)\boldsymbol{\phi}_1 + ... + \dot{\alpha}_{l_l}(t)\boldsymbol{\phi}_l = \sum_{i=1}^{l} \boldsymbol{\phi}_i \dot{\alpha}_{l_i} = \boldsymbol{\Phi}_l \dot{\boldsymbol{\alpha}} \tag{2.45}$$

$$\ddot{\mathbf{u}}_l(t) \approx \ddot{\alpha}_{l_1}(t)\boldsymbol{\phi}_1 + ... + \ddot{\alpha}_{l_l}(t)\boldsymbol{\phi}_l = \sum_{i=1}^{l} \boldsymbol{\phi}_i \ddot{\alpha}_{l_i} = \boldsymbol{\Phi}_l \ddot{\boldsymbol{\alpha}} \tag{2.46}$$

2.3.1 Model equation

Let us replace the expression of the reduced state vector into equation (2.1), having then:

$$\mathbf{M}\boldsymbol{\Phi}_l\ddot{\boldsymbol{\alpha}}(t) + \mathbf{D}\boldsymbol{\Phi}_l\dot{\boldsymbol{\alpha}}(t) + \mathbf{K}\boldsymbol{\Phi}_l\boldsymbol{\alpha}(t) \approx \mathbf{F}(t) \tag{2.47}$$

Of course the equation cannot be used because of the error introduced by choosing a dimension $l < n$. A residual function has to be introduced:

$$\mathbf{M}\boldsymbol{\Phi}_l\ddot{\boldsymbol{\alpha}}(t) + \mathbf{D}\boldsymbol{\Phi}_l\dot{\boldsymbol{\alpha}}(t) + \mathbf{K}\boldsymbol{\Phi}_l\boldsymbol{\alpha}(t) - \mathbf{F}(t) = \mathbf{r}(t) \tag{2.48}$$

At this stage, since $\mathbf{r}(t)$ makes the problem undefined, we need an additional equation, that is the orthogonality between the residual and a test basis $\mathbf{W} \in \mathbb{R}^{n\times l}$:

$$\mathbf{W}^T\mathbf{r}(t) = \mathbf{0} \tag{2.49}$$

Left-multiplying the dynamical equation by $\mathbf{W}^T$ we get the so-called Petrov-Galerkin projection-based formulation of the problem:

$$\mathbf{W}^T\mathbf{M}\boldsymbol{\Phi}_l\ddot{\boldsymbol{\alpha}}(t) + \mathbf{W}^T\mathbf{D}\boldsymbol{\Phi}_l\dot{\boldsymbol{\alpha}}(t) + \mathbf{W}^T\mathbf{K}\boldsymbol{\Phi}_l\boldsymbol{\alpha}(t) - \mathbf{W}^T\mathbf{F}(t) = \mathbf{W}^T\mathbf{r}(t) = \mathbf{0} \tag{2.50}$$

If we suppose $\mathbf{W} = \boldsymbol{\Phi}_l$, this procedure is simply called Galerkin projection and leads to the final formulation:

$$\boldsymbol{\Phi}_l^T\mathbf{M}\boldsymbol{\Phi}_l\ddot{\boldsymbol{\alpha}}(t) + \boldsymbol{\Phi}_l^T\mathbf{D}\boldsymbol{\Phi}_l\dot{\boldsymbol{\alpha}}(t) + \boldsymbol{\Phi}_l^T\mathbf{K}\boldsymbol{\Phi}_l\boldsymbol{\alpha}(t) - \boldsymbol{\Phi}_l^T\mathbf{F}(t) = \mathbf{0} \tag{2.51}$$

and finally, redefining the reduced state matrices of the system, we get:

$$\mathbf{M}_l\ddot{\boldsymbol{\alpha}}(t) + \mathbf{D}_l\dot{\boldsymbol{\alpha}}(t) + \mathbf{K}_l\boldsymbol{\alpha}(t) = \mathbf{F}_l(t) \tag{2.52}$$

The error due to the reduction of the model is:

$$\boldsymbol{\epsilon}_r(t) = \mathbf{u}(t) - \boldsymbol{\Phi}_l\boldsymbol{\alpha}(\boldsymbol{t}) \tag{2.53}$$

Using the Galerkin projection previously described, the error can be decomposed into two orthogonal components:

$$\boldsymbol{\epsilon}_r(t) = \boldsymbol{\epsilon}_\perp(t) + \boldsymbol{\epsilon}_\parallel(t) \tag{2.54}$$

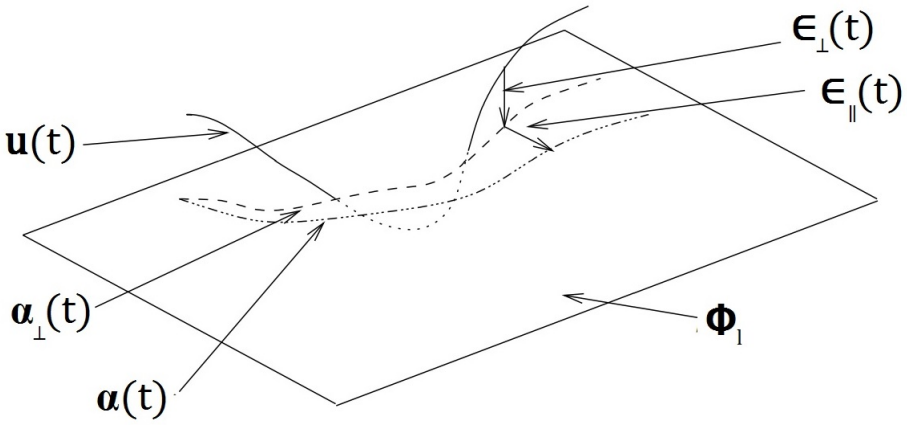


Figure 2.1: POD projection error
Figure adapted from [2]

In Figure 2.1 the projection and the respective errors are shown in a simple geometrical analogy (with $\boldsymbol{\alpha}_\perp(t) = \boldsymbol{\Phi}_l \boldsymbol{\Phi}_l^T \mathbf{u}(t)$). The figure helps to understand why a discrepancy arises when the model is reduced: in this case a 3 dimensional full model system is reduced in a 2 dimensional reduced model system, therefore the full state evolution $\mathbf{u}(t)$ is projected onto it.

2.3.2 Initial conditions

Suppose that the initial conditions applied to the full system are:

$$\mathbf{u}(t_0) = \mathbf{u}_0 \in \mathbb{R}^n \tag{2.55}$$

The reduced initial conditions are still obtained projecting the equation in the subspace:

$$\boldsymbol{\Phi}_l \boldsymbol{\alpha}(0) - \mathbf{u}_0 \approx \mathbf{r}_0 \in \mathbb{R}^l \tag{2.56}$$

$$\mathbf{W}^T \boldsymbol{\Phi}_l \boldsymbol{\alpha}(0) - \mathbf{W}^T \mathbf{u}_0 = \mathbf{W}^T \mathbf{r}_0 = \mathbf{0}$$

$$\boldsymbol{\alpha}(0) = (\mathbf{W}^T \boldsymbol{\Phi}_l)^{-1} \mathbf{W}^T \mathbf{u}_0$$

Since a Galerkin projection has been used and the basis is orthonormal, the final result is:

$$\boldsymbol{\alpha}(0) = (\boldsymbol{\Phi}_l^T \boldsymbol{\Phi}_l)^{-1} \boldsymbol{\Phi}_l^T \mathbf{u}_0 = \boldsymbol{\Phi}_l^T \mathbf{u}_0$$

2.4 Physical interpretation of POMs

Several papers have been written on trying to find a physical interpretation of the POMs or at least link them with the vibration eigenmodes associated to the system. First of all, it is important to emphasize the fact that from a theoretical point of view POMs and eigenmodes could not be compared because of their different nature: the POMs are linked to the evolution of the system and originates and depends from that certain set of data chosen and, hence, on how the structure is loaded, while the eigenmodes are related to the system itself and to its free vibration, that is a generic and ideal state. For all these reasons it is somehow impossible to find a theoretical correlation between them two.
Nevertheless, some remarkable conclusion can be said in case of some particular cases and later on they are summarized for the case of a mechanical dynamic system.

2.4.1 Free-vibration case

Let us consider an undamped and unforced dynamic linear system:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{0} \tag{2.57}$$

In the ideal case of a mass matrix proportional to the identity matrix and an infinite number of snapshots, then [51]:

- the proper orthogonal modes are equal to the eigenmodes;
- the first N_{snap} right singular vectors are the normalized time combination of the modes.

Therefore, if we still consider a mass matrix proportional to the identity matrix, but a finite number of samples N_{snap}, we can conclude that the POMs converge to the eigenmodes; if the mass matrix is not proportional to the identity matrix, the convergence is no more fulfilled. Nevertheless still a link between them holds, in fact, replacing $\mathbf{u} = \mathbf{M}^{-1/2}\mathbf{q}$ and multiplying by $\mathbf{M}^{-1/2}$, the equation (2.57) becomes:

$$\ddot{\mathbf{q}} + \mathbf{M}^{-1/2}\mathbf{K}\mathbf{M}^{-1/2}\mathbf{q} = \mathbf{0} \tag{2.58}$$

In this way, we have obtained a system where the previous ideal hypothesis holds: then, the POMs converge to the eigenmodes of this auxiliary system and it is possible to go back to the original eigenmodes through the simple operator $\mathbf{M}^{-1/2}$. Let us now consider instead a damped unforced dynamic system:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{D}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{0} \tag{2.59}$$

Since $\mathbf{u} \to \mathbf{0}$ for $t \to \infty$ due to the damping effect, it can be proved that the set of POMs is no more converging to the eigenmodes. However, if the damping causes little effects on the system and the number of samples N_{snap} is large enough, the POMs are still a good approximation of the eigenmodes.

2.4.2 Forced case

The system considered is:

$$\mathbf{M\ddot{u}} + \mathbf{Ku} = \mathbf{F}\sin(\omega t) \tag{2.60}$$

The previous properties still hold, but a new ausiliary system has to be found. A new snapshot matrix $\mathbf{U}' \in \mathbb{R}^{n\times(N_{snap}+1)}$ will collect N_{snap} components taken from the original system response and a new component equal to the harmonic response of the system. The first POM of $\mathbf{U}'$ is related to the eigenmodes through [52]:

$$\boldsymbol{\phi}_i = \mathbf{F}\sum_j^i \frac{\boldsymbol{\pi}_i\boldsymbol{\pi}_i^T}{(\omega_i^2-\omega^2)\mu_i} \tag{2.61}$$

where $\boldsymbol{\pi}_i$ and ω_i are respectively the i-th eigenmode and eigenvalue of the original system and μ_i is an appropriate constant described in [53]. Again if the mass matrix is proportional to the identity matrix, the first N_{snap} components of the remaining POMs are simply the eigenmodes of the original system.
More important results can be underlined when we consider only the part of the response that is synchronous with the excitation [52]:

- one single POM is able to retrieve all the information;
- that single POM can be calculated using an analytical expression, simply knowing the eigenmodes and eignevalues of the original system without calculate the evolution of the system:

$$\boldsymbol{\phi}_1 = \frac{\mathbf{f}\sum_j^n \frac{\boldsymbol{\pi}_i\boldsymbol{\pi}_i^T}{(\omega_i^2-\omega^2)\mu_i}}{||\mathbf{f}\sum_j^n \frac{\boldsymbol{\pi}_i\boldsymbol{\pi}_i^T}{(\omega_i^2-\omega^2)\mu_i}||_{L^2}} \tag{2.62}$$

- whenever the angular velocity ω of the load is close to an eigenvalue of the system, the POM is proportional to the relative eigenmode and thus equal to the resonant shape.

Let us now consider an undamped system subjected to a random excitation:

$$\mathbf{M\ddot{u}} + \mathbf{Ku} = \mathbf{F}(t) \tag{2.63}$$

If the modal frequencies of the system are distinct, the modal correlation matrix defined as a snapshot matrix of N_{snap} generalized coordinate vectors is diagonal and the elements are the mean-squared values of the coordinates. If the system is damped but the damping term does not play an important role on the solution, it can be proved that, for a large value of N_{snap}, the correlation matrix has predominant diagonal elements. Since that correlation matrix is diagonal, the

POMs of the system converge to the eigenmodes of the system [54].
It is interesting to point out that the latter result is valid only in the case of a convergent Fourier transform: for instance, if we consider a non-zero mean excitation, such as a constant force, the response will be constant and the first POM will catch exactly that configuration even if it can not be considered as an eigenmode [54]. The same problem arises whenever the number of snapshots is so small that the mean is not zero due to the numerical approximation. The simplest way to overcome this problem is to subtract the mean from the response values.

2.5 Summary

In the present chapter, the basic concepts about the model order reduction have been presented.
First of all a brief review of the most used methods has been described, then we focused on the Proper Orthogonal Decomposition explaining in details its mathematical meaning. The formulation of the optimization formulation associated to POD has been shown, describing then the two most famous procedure that allow to solve that problem: Principal Component Analysis and Singular Value Decomposition. Later, we dealt with the error associated to the reduction of the model and consequently a way to chose its appropriate reduced dimension l. Once the theoretical construct of the proper orthogonal modes have been displayed, it has been shown how to use the latter ones in order to project the full model space into the reduced model subspace. At the end, a review of the most recent findings on the physical interpretation of the POMs has been presented.
The following chapter will introduce the Bayesian Recursive filters: the model order reduction will be then applied to the latter ones, leading then to a faster estimation procedure.

Chapter 3

Recursive Bayesian Filters

3.1 Introduction

Bayesian inference can be referred to the family of all those methods that allow to estimate the probability distribution of some beliefs, known a certain set of available observations and the link between them two.
A wide variety of possible applications of this concept has been presented in the past in different research fields: economics [55, 56], robotics [57], biology [58], medicine [59], computer science [60], decision theory [61] etc.. In the present work we are interested in the application of these concept to the structural damage identification and localization [62, 63, 64, 65].
First of all a theorethical review of the problem is presented, then the concepts of Kalman filter and Particle filter will be introduced.

3.2 The estimation problem

First of all let us introduce the general estimation problem and details of its basic components. We will first deal with the theoretical formulation and only in a second moment it will be applied to the specific case of structural applications. Given a certain system, the first conceptual step is to reproduce it through a mathematical model: for this reason, a certain set of random variables should be chosen in order to describe exhaustively the system itself. The choice of these variables depends directly on the model used: for instance, considering a dynamic structural problem, a reasonable choice could be the displacements or rotations at each node of the FEM model. All these random variables are collected into the so-called state vector $\mathbf{x}(t) \in \mathbb{R}^n$. Similarly, the measurements available are collected in an other random variable $\mathbf{y}(t) \in \mathbb{R}^{N_{obs}}$. Practically, since for a real model the solution of the estimation problem in an analytical form is not feasible, the evolution of the system over time is described in a discrete way. From now on, the notation $\mathbf{x}_k$ and $\mathbf{y}_k$ will respectively stand for $\mathbf{x}(t_k)$ and $\mathbf{y}(t_k)$, where t_k

is a certain time instant; the index $k \in \mathbb{N}$ is defined in $[1, N]$, where N is the number of intervals in which the time is discretized.
The next steps are:

1. the description of the evolution of the system over time;
2. the description of the relationship between the state and measurement vectors.

These goals are respectively fulfilled through the so-called process model and measurement model.

1. **Process model**
   The evolution of a system within the generic time interval $[t_{k-1}\ t_k]$ can be written in the following way:

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_k) \tag{3.1}$$

   The functional $\mathbf{f}_k$ allows to relate the state of the system at the previous instants t_{k-1} with the one at the actual t_k. The latter function derives basically from the time discretization model exploited: for instance, as it will be explained in Table 4.4, in the structural case a Newmark integration scheme can be used.
   Considering the particular case of a linear system, the same model can be rearranged as follows:

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{w}_k \tag{3.2}$$

   In this case $\mathbf{F}_k \in \mathbb{R}^{n \times n}$ is a matrix called transition matrix. This notation is called state-space rapresentation and it has been developed to make the problems tractable and basically adopted for every kind of estimation problem. $\mathbf{w}_k$ is called process noise and its covariance $\mathbf{W}_k$ matrix is defined as:

$$\boldsymbol{W}_k = E[(\mathbf{w}_k - E[\mathbf{w}_k])(\mathbf{w}_k - E[\mathbf{w}_k])^T] \tag{3.3}$$

   Considering the particular case of a zero-mean white process noise, the elements of $\mathbf{W}_k$ become:

$$W_k^{ij} = \begin{cases} \sigma_{w_i}^2 = E[(w_i - E[w_i])^2], & \text{for } i = j \\ 0, & \text{for } i \neq j \end{cases} \tag{3.4}$$

   and its auto-correlation matrix is equal to the identity matrix.
   The process noise takes into account all the uncertainties deriving from the discretization procedure.

2. **Measurement model**
   Another important equation required to use the filter is:

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k) \tag{3.5}$$

The functional $\mathbf{h}_k$ relates the observations and the state vector at the actual time instant t_k.
Whenever $\mathbf{h}_k$ is linear, equation (3.5) becomes:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \tag{3.6}$$

As it will be later explained, considering the specific application presented in this work, the matrix $\mathbf{H}_k$ is a boolean matrix whom elements are 1 or 0 arranged in a way that:

$$y_k^i = x_k^j + v_k^j \tag{3.7}$$

where $j = 1, ..., n$ is the index of the observed degree of freedom that corresponds to the i-th element of the observation vector, with $i = 1, ..., N_{obs}$. In the same way presented before for the process model, even in this case an appropriate noise $\mathbf{v}_k$, called measurement noise, is considered. Intuitively, the noise level is defined as the level below which a signal cannot reliably be detected and the fluctuations due to uncertainties are too large compared to the signal itself. From a physical point of view, it reproduces the uncertainties related to the measurement system and could be due to several different effects [66]:

- degradation of signals due to the physical medium of the communication devices;
- random electrical noise related to sensors and electrical circuits;
- quality of the measurement devices;
- environmental noise.

The level of $\mathbf{v}_k$ can be lowered either improving the data acquisition system or using some filtering procedures, such as signal averaging [67]. Its covariance matrix is $\mathbf{V}_k$ defined as:

$$\mathbf{V}_k = E[(\mathbf{v}_k - E[\mathbf{v}_k])(\mathbf{v}_k - E[\mathbf{v}_k])^T] \tag{3.8}$$

Considering the particular case of a zero-mean white process noise, the elements of matrix $\mathbf{V}_k$ become:

$$V_k^{ij} = \begin{cases} \sigma_{v_i}^2 = E[(v_i - E[v_i])^2], & \text{for } i = j \\ 0, & \text{for } i \neq j \end{cases} \tag{3.9}$$

and its auto-correlation matrix is equal to the identity matrix.

Table 3.1 summarizes what has been explained so far, linking in a block diagram the concepts of model and measurement processes. In this figure $z^{-1}\mathbf{I}$ is the Z-transform of the function:

$$\delta[n-1] = \begin{cases} 1, & \text{if } n = 1 \\ 0, & \text{if } n \neq 0 \end{cases} \tag{3.10}$$

The notation $z^{-1}\mathbf{I}$ represents a time shifting from time instant t_{k+1} to t_k; in other words $\mathbf{x}_{k+1} \rightarrow \mathbf{x}_k$.

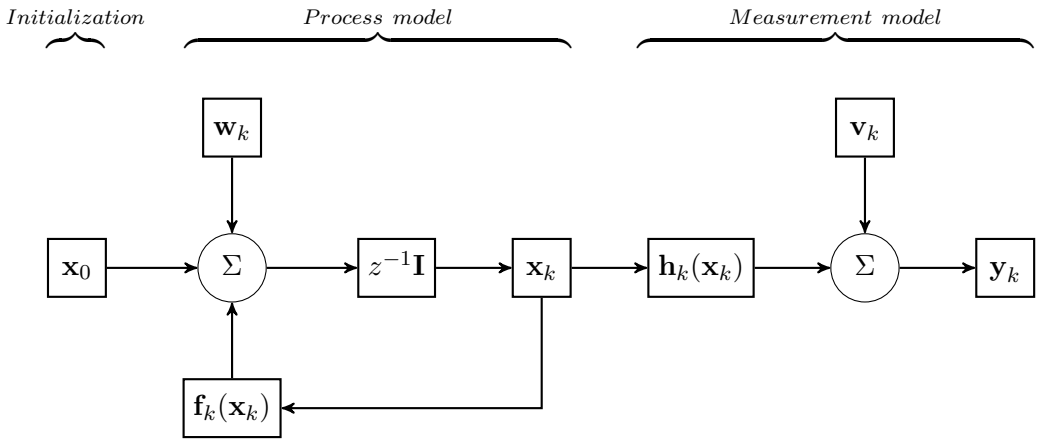


Figure 3.1: Block diagram representation of a discrete-time dynamical system Figure adapted from [3]

3.3 Bayesian Inference

Let us recall the main goal of Bayesian inference: we want to estimate the state vector of the system $\mathbf{x}_k$ at time instant t_k, given a certain set of available measurements $\mathbf{y}_j$ at time instants $t_1, ..., t_k$.
Before going ahead with the explanation of the procedure, some topics on Bayesian statistics should be recalled. Considering two events D and H, the conditional probability of event H, given that the event D is occurring, is defined as:

$$P(H|D) = \frac{P(H \cap D)}{P(D)} \tag{3.11}$$

Suppose now the events H and D are respectively the occurrence of a certain hypothesis and of an evidence, i.e. a certain set of data. The Bayes' rule [68] states that the conditional probability of H given D is:

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)} \tag{3.12}$$

The conditional probability $P(H|D)$ is called posterior probability, while $P(H)$ is named prior probability because it is associated to the occurrence of H before D is available. $P(D|H)$ is the likelihood, namely the compatibility of D with the hypothesis H and $P(D)$ is the so-called marginal likelihood, that is a neutral factor with respect to different choices of H. Considering the notations used in this

case, the hypothesis H and the evidence D refers respectively to the occurrence of the state vector $\mathbf{x}_k$ and the observations $\mathbf{y}_k$.
Having defined the Bayes' rule and linked the estimation problem within the Bayesian frame, one last theorem should be recalled. Consider a stochastic process in which the random variables f_1, f_2 and f_3 are defined. Suppose then that the conditional probability distribution of future states of those random variables depend only on the present state and not on the sequence of previous events [69]. The Chapman-Kolmogorov equation [70] allow to find the conditional probability distribution:

$$p(f_3|f_1) = \int_{-\infty}^{+\infty} p(f_3|f_2)p(f_2|f_1)df_2 \tag{3.13}$$

Let us now go back to our estimation problem: the basic idea is to estimate the expected value $E[\mathbf{x}_k|\mathbf{y}_{1:k}]$ of the random variable $\mathbf{x}$ at each time instant t_k, given a set of observed data $\mathbf{y}_1, ..., \mathbf{y}_k$. Since we want $\mathbf{x}_k$ to be estimated not only as conditioned by the observations but also by the previous states of the state vector, we are basically dealing with the estimation of the evolution of the random variable $\mathbf{x}$ in a recursive way. The problem can be even described in a more complete way as follows [71]:

1. the initial state of the system, or rather the probability density function (from now on pdf) of the state vector $p(\mathbf{x}_0|\mathbf{y}_0) = p(\mathbf{x}_0)$, is known and does not depend on any observation;

2. in the prediction stage the conditional pdf of the state vector $\mathbf{x}_k$ at time t_k is predicted, given the past observations, i.e. the data available until the previous instant t_{k-1}; recalling equation (3.13), the sought pdf can be calculated using the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1} \tag{3.14}$$

   where the probabilistic model of $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is basically defined by the process model and the respective noise.

3. in the update stage, the pdf of $\mathbf{x}_k$ at time t_k is updated taking into account the last available observations $\mathbf{y}_k$ at time t_k; the pdf $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ can be calculated by recalling the Bayes' rule (eq. (3.12)):

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \tag{3.15}$$

   where the denominator is a constant and can be evaluated using the Chapman-Kolmogorov equation (eq. (3.13)):

$$p(\mathbf{y}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{y}_k \tag{3.16}$$

Therefore, in order to summarize the latter scheme, it is possible to conclude that the estimation problem is basically a recursive procedure where the density function of the state is first of all estimated knowing the previous probability, then it is modified in order to take account of the new information available.
This formulation of the problem is merely conceptual, in fact considering general non-linear problems the integral in eq. (3.14) can not be analytically solved and hence some approximated procedures are required. On the other hand, under some specific hypotheses, such as linearity of the state space equation and uncorrelated white Gaussian noises, a solution can be found. In 1960, Kalman [72, 73, 74] introduced a method that allows to calculate that solution of the problem in an optimal way and under the aforementioned hypotheses.

3.3.1 Optimal estimate: Kalman filter

The basic assumption of the Kalman filter (KF) is to suppose the posterior pdf at every step to be Gaussian; therefore, the pdf can be always expressed through its mean and covariance as $p(\mathbf{x}_k) = \mathcal{N}\{E[\mathbf{x}_k], (\mathbf{x}_k - E[\mathbf{x}_k])(\mathbf{x}_k - E[\mathbf{x}_k])^T\}$. Summarizing, the Kalman filter represents an optimal solution of the estimation problem if the following hypotheses hold true:

- the posterior density is Gaussian at each step;
- the process and measurement noises $\mathbf{w}_k$ and $\mathbf{v}_k$ are white, uncorrelated and Gaussian;
- $\mathbf{f}_k$ is a linear function;
- $\mathbf{h}_k$ is a linear function.

As proved in [75], if $p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})$ is Gaussian and the latter assumptions are verified, then $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is Gaussian as well; the direct consequence is that the calculation of the Chapman-Kolmogorov integral can be always performed.
In order to achieve this goal, it can be proved that the problem can be even handled as an optimization problem [3, 76], where the appropriate loss function is:

$$\mathbf{J}_k = \text{tr}\left\{E[(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T]\right\} \tag{3.17}$$

where $\mathbf{x}_k$ is referred to the actual true state vector at time t_k, while the main goal is to find its estimation $\hat{\mathbf{x}}_k$. In these terms, the solution is represented by that estimated state vector $\hat{\mathbf{x}}_k$ for which the mean-square error $\mathbf{J}_k$ is minimum.
In order to find the optimal estimation, the following theorems have to be used [77, 78, 79]:

- if the stochastic processes $\mathbf{x}_k$ and $\mathbf{y}_k$ are jointly Gaussian, then the optimum estimate such that $\mathbf{J}_k$ is minimum is the conditional mean estimator $\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathbf{y}_1, ..., \mathbf{y}_k]$;

- if $\mathbf{x}_k$ and $\mathbf{y}_k$ are stochastic processes of zero means, then they are jointly Gaussian;
- if the optimal estimate is a linear function of the observations, then $\hat{\mathbf{x}}_k$ is the orthogonal projection of $\mathbf{x}_k$ onto the space spanned by the observations.

Let us consider a measurement $\mathbf{y}_k$ at time t_k and suppose an a priori estimation $\hat{\mathbf{x}}_k^-$ is available. Practically, $\hat{\mathbf{x}}_k^-$ is the estimation obtained in the prediction stage, using the aforementioned process model (eq. (3.2)):

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_k \hat{\mathbf{x}}_{k-1} \tag{3.18}$$

i.e. the estimated state vector at time t_k, given the estimated state vector at time t_{k-1}.
Then a better estimate of the state at time t_k can be expressed as a function of both the measurements and the a priori estimation:

$$\hat{\mathbf{x}}_k = \mathbf{G}_k^1 \hat{\mathbf{x}}_k^- + \mathbf{G}_k \hat{\mathbf{y}}_k \tag{3.19}$$

where $\mathbf{G}_k$ and $\mathbf{G}_k^1$ are appropriate matrices. In order to find them, using the previous statements, the solution of the minimization problem described in equation (3.17) can be found by simply imposing:

$$E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)\mathbf{y}_i^T] = \mathbf{0} \text{ for } i = 1, ..., k-1 \tag{3.20}$$

and replacing the expression of equations (3.5) and (3.19), we get:

$$E[(\mathbf{x}_k - \mathbf{G}_k^1 \hat{\mathbf{x}}_k^- - \mathbf{G}_k \mathbf{H}_k \mathbf{x}_k - \mathbf{G}_k \mathbf{w}_k)\mathbf{y}_i^T] = \mathbf{0} \text{ for } i = 1, ..., k-1 \tag{3.21}$$

Let us suppose the process and measurement noises are uncorrelated, then:

$$E[\mathbf{w}_k \mathbf{y}_i^T] = \mathbf{0} \text{ for } i = 1, ..., k-1 \tag{3.22}$$

Collecting all the previous equations and rearranging the elements, we get:

$$(\mathbf{I} - \mathbf{G}_k \mathbf{H}_k - \mathbf{G}_k^1) E[\mathbf{x}_k \mathbf{y}_i^T] = \mathbf{0} \text{ for } i = 1, ..., k-1 \tag{3.23}$$

For arbitrary variables, the expression holds true whenever:

$$\mathbf{I} - \mathbf{G}_k \mathbf{H}_k - \mathbf{G}_k^1 = \mathbf{0} \tag{3.24}$$

At this point we have then a link between $\mathbf{G}_k^1$ and $\mathbf{G}_k$, that makes equation (3.19) become:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{G}_k(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-) \tag{3.25}$$

The matrix $\mathbf{G}_k$ is called Kalman gain and it allows us to relate the previous state and observation with the actual estimate of the state itself. Practically, the

prediction is updated taking into account its respective error, i.e. the difference between the observed and predicted values.
The next step is basically to find an appropriate expression of $\mathbf{G}_k$. Using the aforementioned theorems, it can be stated that:

$$E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)\mathbf{y}_k^T] = \mathbf{0} \tag{3.26}$$

and therefore:

$$E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)\hat{\mathbf{y}}_k^T] = \mathbf{0} \tag{3.27}$$

where $\hat{\mathbf{y}}_k$ is an estimate of $\mathbf{y}_k$ given the previous measurements. Define the so called innovations process:

$$\tilde{\mathbf{y}}_k = \mathbf{y}_k - \hat{\mathbf{y}}_k = \mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^- \tag{3.28}$$

Therefore, subracting equations (3.26) and (3.27), we get:

$$E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)\tilde{\mathbf{y}}_k^T] = \mathbf{0} \tag{3.29}$$

Using (3.28) and (3.25), the latter equation becomes:

$$E[[(\mathbf{I} - \mathbf{G}_k\mathbf{H}_k)\tilde{\mathbf{x}}_k^- - \mathbf{G}_k\mathbf{v}_k](\mathbf{H}_k\tilde{\mathbf{x}}_k^- + \mathbf{v}_k)] = \mathbf{0} \tag{3.30}$$

where $\mathbf{v}_k$ is the measurement noise and $\tilde{\mathbf{x}}_k^- = \mathbf{x}_k^- - \hat{\mathbf{x}}_k^-$ is the state vector error. Since the measurement noise is independent of the vector state, it is possible to expand the expression:

$$(\mathbf{I} - \mathbf{G}_k\mathbf{H}_k)E[\hat{\mathbf{x}}_k^-\hat{\mathbf{x}}_k^{-^T}]\mathbf{H}_k^T - \mathbf{G}_kE[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{0} \tag{3.31}$$

where it has been considered that $E[\hat{\mathbf{x}}_k^-\mathbf{v}_k^T] = E[\mathbf{v}_k\hat{\mathbf{x}}_k^{-^T}] = \mathbf{0}$. Let us now define the covariance matrices:

$$\mathbf{P}_k^- = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T] = E[\tilde{\mathbf{x}}_k^-\tilde{\mathbf{x}}_k^{-^T}] \tag{3.32}$$

where $\tilde{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ is the state error vector.
Using the latter definition into the previous equation, we get:

$$(\mathbf{I} - \mathbf{G}_k\mathbf{H}_k)\mathbf{P}_k^-\mathbf{H}_k^T - \mathbf{G}_k\mathbf{W}_k = \mathbf{0} \tag{3.33}$$

This equation can be therefore solved to find the Kalman gain:

$$\mathbf{G}_k = \mathbf{P}_k^-\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{W}_k)^{-1} \tag{3.34}$$

where $\mathbf{W}_k$ is the covariance matrix of the process noise previously defined (eq. (3.3)).
Once the Kalman gain is obtained, there is one last stage to complete the description of the estimation procedure: the covariance matrix associated to the estimation error has to be predicted at a given time t_k.

Table 3.1: Kalman Filter algorithm (KF)

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$$

$$\mathbf{P}_0 = E[(\mathbf{x}_0 - E[\mathbf{x}_0])(\mathbf{x}_0 - E[\mathbf{x}_0])^T]$$

2. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_k \hat{\mathbf{x}}_{k-1}^-$$

$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{W}_k$$

   (b) Updating stage

$$\mathbf{G}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{V}_k)^{-1}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{G}_k(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-)$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{G}_k \mathbf{H}_k \mathbf{P}_k^-$$

Actually, the covariance matrix has to be modified twice within each step:

1. The covariance matrix $\mathbf{P}_k^-$ is predicted knowing the covariance matrix at the previous step t_{k-1}.
   As previously explained, the process model is:

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_k \hat{\mathbf{x}}_{k-1}^- \tag{3.35}$$

   The a priori estimation error is then:

$$\tilde{\mathbf{x}}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^- = \mathbf{F}_k \tilde{\mathbf{x}}_{k-1}^- + \mathbf{w}_{k-1} \tag{3.36}$$

   Knowing that $E[\mathbf{w}_k \tilde{\mathbf{x}}_{k-1}^T] = \mathbf{0}$ and using equations (3.32) and (3.36) we get:

$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{V}_k \tag{3.37}$$

2. The predicted covariance matrix $\mathbf{P}_k^-$ is updated, considering the state estimation at that step.
   Using equations (3.25) and (3.32) we get:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{H}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{G}_k \mathbf{H}_k)^T + \mathbf{G}_k E[\mathbf{v}_k \mathbf{v}_k^T] \mathbf{G}_k^T \tag{3.38}$$

   Expanding the latter equation and replacing the definition of Kalman gain (eq. (3.34)) we have:

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{G}_k \mathbf{H}_k \mathbf{P}_k^- \tag{3.39}$$

At this point, all the phases and components of the Kalman filter has been explained, and therefore a brief schematic summary of the overall procedure is shown in Table 3.1.
As previously explained, the Kalman filter is an optimal solution of the estimation problem only if the model is linear. As it will be shown in the next section, whenever this hypothesis does not hold true, the Extended Kalman Filter (EKF) can be used.

3.3.2 Suboptimal estimate: Extended Kalman Filter

One of the main assumptions of the Kalman filter is the linearity of the functions $\mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})$ and $\mathbf{h}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$. In mechanical systems, the source of non-linearity could be due to the material constitutive law, loading or geometric effects. In Section 4.2.1, we will show that even the dual estimation problems are non-linear because some bi-linear terms arise. If the system is not linear but the pdf of the state vector at time t_{k-1} is Gaussian, then it is no more possible to prove that the pdf at the following step t_k is still Gaussian. The Kalman filter is no more an optimal solution of the Bayesian inference and a new procedure should be used.

Table 3.2: Extended Kalman Filter algorithm (EKF)

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$$
$$\mathbf{P}_0 = E[(\mathbf{x}_0 - E[\mathbf{x}_0])(\mathbf{x}_0 - E[\mathbf{x}_0])^T]$$

2. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage

$$\mathbf{F}_k = \nabla_{\mathbf{x}} \mathbf{f}_k(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_{k-1}}$$
$$\mathbf{H}_k = \nabla_{\mathbf{x}} \mathbf{h}_k(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_{k-1}}$$
$$\hat{\mathbf{x}}_k^- = \mathbf{F}_k \hat{\mathbf{x}}_{k-1}^-$$
$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{W}_k$$

   (b) Updating stage

$$\mathbf{G}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{V}_k)^{-1}$$
$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{G}_k \mathbf{y}_k - \mathbf{h}_k \hat{\mathbf{x}}_k^-$$
$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{G}_k \mathbf{H}_k \mathbf{P}_k^-$$

In order to cope with the problem, the basic idea is to linearize the system through a truncated Taylor expansion of the functions $\mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})$ and $\mathbf{h}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$ centered at the point $\mathbf{x}_k$ [80]. The new transition and measurement matrices can be straightforwardly calculated as the Jacobian matrix of $\mathbf{f}_x$ and $\mathbf{h}_k$ as:

$$\mathbf{F}_k = \begin{bmatrix} \frac{\partial \mathbf{f}_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}_1(\mathbf{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{f}_N(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}_N(\mathbf{x})}{\partial x_N} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_{k-1}} = \nabla_{\mathbf{x}} \mathbf{f}_k(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_{k-1}} \tag{3.40}$$

$$\mathbf{H}_k = \begin{bmatrix} \frac{\partial \mathbf{h}_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial \mathbf{h}_1(\mathbf{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{h}_N(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial \mathbf{h}_N(\mathbf{x})}{\partial x_N} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_k} = \nabla_{\mathbf{x}} \mathbf{h}_k(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_k} \tag{3.41}$$

where the notation $\nabla_{\mathbf{x}}$ represents the gradient of a certain function with respect to the elements of a given vector $\mathbf{x}$. It is important to point out that, since the method relies on a truncated approximation of the system, the algorithm can no more be considered as an optimal solution of the problem. As a consequence, the matrix $\mathbf{P}_k$ does not represent the true covariance of the estimate, but tends to understimate it. The main drawback of the EKF is that a linearization error is always introduced as the difference between the nonlinear functions and their tangential approximations. This error is not always negligible: it has been shown that in some severe nonlinear systems a biased or even divergent estimation can arise [81, 82, 83]. Some modified EKF algorithms have been developed in order to take into account the errors [84]; on the other hand the use of second-order EKF is prevented due to the high computational cost [85]. Whenever these remedies are not sufficient different methods, such as Unscented Kalman filters [86] or Particle filters, can be used.

3.4 Particle Filter

As previously explained, the Kalman filter is considered an optimal solution of the estimation problem only when a certain set of assumptions holds true. The extended version can be used for non-linear problems, but, since it is basically based on an approximation, it fails in two cases [87]:

- the linearization is not enough accurate;
- the process model cannot be written in an explicit form or the calculation of the Jacobian matrix is too computationally expensive.

Regarding this last issue, the implementation of the EKF can become very difficult whenever a non-holonomic model is considered: in this case, in order to

calculate the Jacobian, one should know if the system is evolving towards a loading or an unloading condition [88, 89, 25]. In order to overcome these issues, a more general solution has been presented in [90]. According to the different fields in which it is used, the method is known as particle filtering (PF) [91], bootstrap filtering [90], condensation algorithm [92], interacting particle approximations [93] and survival of the fittest [94].
Since the main goal of the estimation problem is to find the posterior density function of the state variables, we will no more try to find an analytical solution of the Chapman-Kolmogorov equation, but instead we will directly approximate the posterior density function through the generation of a certain number of samples and relevant weights. In other words, given the observations from time t_0 to t_k, while before we were looking for the conditional probability $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ of the vector state $\mathbf{x}_k$ at time t_k, now we concentrare on $p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$ considering all the instants at a glance.
It is important to point out that neither the hypothesis on the linearity of the model nor on the Gaussianity are required. The samples are created using a Monte Carlo method, therefore, increasing the number of samples, the solution converges to the optimal Bayesian estimate [95].
According to [96], the wanted pdf can be always expressed as:

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) = \int f(\mathbf{x}_{0:k})p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})d\mathbf{x}_{0:k} \tag{3.42}$$

Consider a set of samples $\mathbf{x}_k^i$ with $i = 1, ..., N_s$ generated from the posterior distribution. The Monte Carlo approximation reads:

$$\int f(\mathbf{x}_{0:k})p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})d\mathbf{x}_{0:k} \approx \frac{1}{N_s}\sum_{i=1}^{N_s} f(\boldsymbol{x}_k^i) \tag{3.43}$$

The main trick of the method is to use the Dirac delta as function f [97], obtaining:

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) \approx \frac{1}{N_s}\sum_{i=1}^{N_s} \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^i) \tag{3.44}$$

According to the law of large numbers (LLN), the higher is the number of samples N_s, the more accurate is the approximation [95]. In order to reduce the number of samples needed to still retain a high degree of accuracy, it is possible to use the so-called importance sampling [96, 98]. In this case the samples are drawn from an arbitrary chosen distribution $\pi(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})$ called importance function (or density), such that:

$$p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) \approx \frac{1}{N_s}\sum_{i=1}^{N_s} \omega_k^{*i}\delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^i) \tag{3.45}$$

where ω_k^{*i} are called importance weights and are equal to [99]:

$$\omega_k^{*i} = \frac{p(\mathbf{y}_{1:k}|\mathbf{x}_{0:k}^i)p(\mathbf{x}_{0:k}^i)}{p(\mathbf{y}_{1:k})\pi(\mathbf{x}_{0:k}^i|\mathbf{y}_{1:k})} \tag{3.46}$$

The weights can be recursively calculated using the following expression:

$$\omega_k^{*i} = \omega_{k-1}^{*i} \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{\pi(\mathbf{x}_k^i|\mathbf{x}_{0:k-1}^i, \mathbf{y}_{1:k})} \tag{3.47}$$

It can be proved that if only the full posterior probability density $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ is required, these coefficients become [71]:

$$\omega_k^i = \frac{p(\mathbf{y}_{1:k}|\mathbf{x}_{0:k}^i)p(\mathbf{x}_{0:k}^i)}{\pi(\mathbf{x}_k^i|\mathbf{y}_{1:k})} \tag{3.48}$$

Hence, even the recursive weights formulation (3.47) is modified:

$$\omega_k^{*i} = \omega_{k-1}^{*i} \frac{p(\mathbf{y}_k|\mathbf{x}_k^i)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{\pi(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i, \mathbf{y}_k)} \tag{3.49}$$

The weights are then normalized in order to make them independent from the number of samples used, having then:

$$\tilde{\omega}_k^i = \frac{\omega_k^i}{\sum_{j=1}^{N_s} \omega_k^j} \tag{3.50}$$

Equation (3.45) then becomes:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k}) \approx \sum_{i=1}^{N_s} \tilde{\omega}_k^i \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^i) \tag{3.51}$$

where, of course [95]:

$$\lim_{N_s \to \infty} \left\{ \sum_{i=1}^{N_s} \tilde{\omega}_k^i \delta(\mathbf{x}_{0:k} - \mathbf{x}_{0:k}^i) \right\} = p(\mathbf{x}_k|\mathbf{y}_{1:k}) \tag{3.52}$$

3.4.1 Degeneracy

One of the problems that affects the particle filtering procedures is the degeneracy of the weights: it has been proved in [100] and shown in [96] that, since the variance of the importance weights can only increase throughout the iterations, at a certain step only one weight will be not negligible, while all the other ones will converge to zero. The direct consequence is that part of the computational efforts required by the procedure are basically useless because of those negligible

contributions.
The degeneracy effect can be measured using the effective sample size [101], theoretically defined as:

$$N_{s,eff} = \frac{N_s}{1 + \sigma^2_{\boldsymbol{\omega}_k}} \tag{3.53}$$

where $\sigma^2_{\boldsymbol{\omega}_k}$ is the variance of the weights and it is defined as:

$$\sigma^2_{\boldsymbol{\omega}_k} = E[(\boldsymbol{\omega}_k - E[\boldsymbol{\omega}_k])^2] = \frac{1}{N_s} \sum_{i=1}^{N_s} \left(\omega_k^i - \frac{\sum_{j=1}^{N_s} (\omega_k^j)}{N_s} \right)^2 \tag{3.54}$$

Practically only an approximated version of this index can be calculated:

$$N_{s,eff} \approx \frac{1}{\sum_{i=1}^{N_s} (\omega_k^i)^2} \tag{3.55}$$

The smaller is $N_{s,eff}$, the quicker is the degeneracy. In order to overcome this issue, it is possible to increase the number of samples, choose a suitable importance function such that the variance is minimized or use some resampling method. These two methods will be explained in Sections 3.4.2 and 3.4.3.

3.4.2 Importance function

Several versions of the distribution π have been proposed. The degeneracy of the weights can be lowered optimally using the following importance density [96]:

$$\pi(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^i, \mathbf{y}_k) \tag{3.56}$$

and

$$\omega_k^i = \omega_{k-1}^i \int p(\mathbf{y}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{x}_{k-1}^i) d\mathbf{x}_k \tag{3.57}$$

In [96] it has been shown that this particular choice of importance density is the optimal one, since it makes the variance $\sigma^2_{\boldsymbol{\omega}_k} = 0$ and hence, recalling equation (3.53), the degeneracy effect is lowered as well. The major drawbacks of the latter solution are that, in order to update the weights, it requires the ability of drawing samples from the respective pdf and to calculate the integral. These tasks can be performed only in two particular cases, i.e. when $\mathbf{x}_k$ is a member of a finite set or the measurement model is linear. Some approximating methods that use localized linearizations have been proposed [96]. Since they require an additional computational cost that can overcompensate the choice of an optimal π, usually a different version of π is used:

$$\pi(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^i) \tag{3.58}$$

and therefore [102]:

$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k | \mathbf{x}_k^i) \tag{3.59}$$

Despite the fact this choice is not the optimal one, it guarantees a simple weights updating, since it requires the evaluation of $p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$, which does not depend on the observations.

3.4.3 Resampling

The basic idea of the resampling methods is to select only those particles to which is associated an higher probability of occurence and then larger weights. The easier procedure is the systematic resampling: a new set of samples is created such that their weights are all equal to $\omega_k^i = 1/N_s$. The m-th sample $\mathbf{x}_k^m$ is duplicated if the random value $u_i \in [0, 1]$ lies between the values of the cumulative density function of *(m-1)*-th and m-th samples:

$$\text{find} \quad m \quad \text{s.t.} \sum_{i=1}^{m-1} \omega_k^i < u_i \leq \sum_{i=1}^{m} \omega_k^i \tag{3.60}$$

In order to find m, the systematic re-sampling algorithm [103] is used:

1. a set of N_s samples $\mathbf{x}_k^i$ with $i = 1, ..., N_s$ is given;
2. the cumulative density function is calculated by simply sequentially adding the weights:
$$F(\mathbf{x}_K^m) = \sum_{i=1}^{m} \omega_k^i \tag{3.61}$$
3. generate $u_i = U[0, 1]$, where U is the uniform probability distribution on the interval $[0, 1]$;
4. increase m until $F(\mathbf{x}_K^m) \geq u_i$;
5. select the sample and the respective weight for the resempled set:
$$\begin{cases} \bar{\mathbf{x}}_k^i = \mathbf{x}_k^m \\ \bar{\omega}_k^{*i} = \frac{1}{N_s} \end{cases} \tag{3.62}$$

Despite the fact the re-sampling methods reduce the degeneracy problems of the filter, there is still a drawback called sample impoverishment [71]: since the samples with high weights are selected many times, the new set of samples will have many repeated particles. Especially in the case of small process noise, after a certain number of iterations all samples will converge to a point. For the extreme case of a zero process noise, the particle filter is not the appropriate way to estimate static states. In order to overcome sample impoverishment, several methods based on genetic algorithms have been used, such as evolutionary particle filter [104] or sequential importance evolutionary particle filter [105].

Having explained the procedure and the main remarks regarding the particle filter, the overall algorithm is summarized in Table 3.3.

Table 3.3: Particle filter algorithm (PF)

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$$
$$\mathbf{x}_0^i = \hat{\mathbf{x}}_0$$
$$\omega_0^i = p(\mathbf{y}_0|\mathbf{x}_0)$$

2. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage ($i = 1, ..., N_s$)

$$\mathbf{x}_k^i \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$$
$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k|\mathbf{x}_k^i)$$

   (b) Resampling stage ($i = 1, ..., N_s$)

$$u_i \sim U[0, 1]$$
$$\text{find} \quad m \quad s.t. \sum_{i=1}^{m-1} \omega_k^i < u_i \leq \sum_{i=1}^{m} \omega_k^i$$
$$\bar{\mathbf{x}}_k^i = \mathbf{x}_k^m$$
$$\bar{\omega}_k^{*i} = \frac{1}{N_s}$$

   (c) Updating stage

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N_s} \bar{\omega}_k^i \bar{\mathbf{x}}_k^i$$

3.5 Summary

In the present chapter, the concept of estimation problem has been explained, focusing on the description of the process and measurement models, both for linear and non-linear cases. Moreover, the definitions of process and measurement noises have been given and their physical interpretation have been discussed. The Bayesian inference has been presented as a framework used to estimate the states of a system at a certain time t_k, given the observations of that system

at the previous times t_{k-1}. After the preliminary definitions of Bayes' rule and Chapman-Kolmogorov equation have been given, the main steps of the Bayesian inference have been explained. Two methods can be used to solve the latter problem: Kalman filter (KF) for linear models, extended Kalman (EKF) filter for nonlinear models. The mathematical proof of the first one is shown, while the EKF is presented as a linearization of the KF. Finally the concept of particle filter is presented, underlining that this method should be used whenever the EKF is not enough accurate or inapplicable. The problem of degeneracy of samples is introduced and its remedies are shown, namely the choice of an appropriate importance function and the resampling technique.
In the next chapter, we will focus on the explanation of the algorithm that will be used in the damage identification. In Chapter 3 an overview on the basic concepts of estimation problem and the methods used to perform it has been presented. Besides, in order to be able to introduce our particular identification method, some further specific tools are needed and they will be explained in Chapter 4. These are the concepts of dual estimation, reduced model estimation hybrid extended Kalman particle filter and dual estimation of reduced states and parameters with subspace updating for damage identification.

Chapter 4

POD Kalman Observer

4.1 Introduction

The basic objective of the present thesis is the description and assessment of a particular recursive Bayesian filter used to identify structural damages. Since in the previous chapter the concepts of estimation problem and recursive Bayesian filters have been presented in the most general way and for any particular application, here we want to specialize a method for our specific task. Section 4.2 collects the explanations of some additional tools that are required in what follows: dual estimation, reduced model estimation and hybrid extended Kalman particle filter. In section 4.3 the latter procedures are used in the final version of the identification algorithm, i.e. the dual estimation of reduced states and parameters with subspace updating algorithm. Finally, in Section 4.4 the application of the latter procedure to the structural case is explained.

4.2 Preliminary tools

As explained in Chapter 1, the damage identification method should be able to estimate not only the dynamical evolution of the system, but also the damage parameters associated to the damaged zones. In what follows, a procedure that allow to modify the filters taking into account both the estimations is explained. We use here other two procedures that will be applied to speed up the recursive calculation. The reduced model estimation allow to reduce the number of the dynamic variables we want to estimate, while an hybrid version of the particle filter is employed to decrease the number of samples needed.

4.2.1 Dual estimation

As previously mentioned, we want now to extend the concepts described in the previous sections to the dual estimation problem.
In other words the method should be able to estimate both the dynamic response of the system and some parameters of the system itself. In order to achieve this purpose, the state vector is composed by two different parts [63]:

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{u}_k \\ \boldsymbol{\theta}_k \end{bmatrix} \tag{4.1}$$

where $\mathbf{u}_k \in \mathbb{R}^n$ and $\boldsymbol{\theta}_k \in \mathbb{R}^{N_p}$ collects respectively the dynamical n random variables that describes the evolution of system and the N_p parameters that will be chosen to estimate.
The concept of process model and measurement model described in the previous chapter still holds, but due to the different nature of the variables, a more detailed description should be made.
The process model of the parameters is modified using the following equation:

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \mathbf{w}_k^{\boldsymbol{\theta}} \tag{4.2}$$

The basic idea is to allow the parameters to change at each time step, leading to final estimation values. The noise $\mathbf{w}_k^{\theta}$ is conceptually different from the process and measurement noise: since it is a fictitious noise introduced just for sake of computational reasons, there is any physical meaning and hence it should be tuned in order for the filter to work properly.
The process model (3.2) is modified in order to take into account equation (4.2), leading to:

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{u}_k \\ \boldsymbol{\theta}_k \end{bmatrix} = \mathbf{f}_k(\mathbf{x}_{k-1}, \boldsymbol{\theta}_{k-1}) + \mathbf{w}_k \tag{4.3}$$

Regarding the measurement model, equation (3.5) is modified as follows:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k = \begin{bmatrix} \mathbf{H}_k^{\mathbf{u}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{k-1} \\ \boldsymbol{\theta}_{k-1} \end{bmatrix} + \mathbf{v}_k \tag{4.4}$$

where the boolean matrix $\mathbf{H}_k$ is changed in order to take account the fact that only the dynamic variables can be measured, while the parameters can only be estimated.
All the other components of the filter maintain the definitions previously given and their expressions in dual estimation derive straightforwardly from equations (4.3) and (4.4).
In [62], it has been shown that whatever linear system will lead to a non-linear dual estimation problem. Let us consider a generic model:

$$\mathbf{u}_k = \mathbf{f}_k^{\mathbf{u}}(\mathbf{u}_{k-1}, \boldsymbol{\theta}_{k-1}) + \mathbf{w}_k^{\mathbf{u}} \tag{4.5}$$

Replacing the latter equation and (4.2) into expression (4.3), we get:

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{f}_k^{\mathbf{u}}(\mathbf{u}_{k-1}, \boldsymbol{\theta}_{k-1}) \\ \boldsymbol{\theta}_{k-1} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_k^{\mathbf{u}} \\ \mathbf{w}_k^{\boldsymbol{\theta}} \end{bmatrix} \tag{4.6}$$

Even considering a simple linear model, there will be always at least one bi-linear term, consisting in the product between a parameter and a dynamical variable. Therefore, it is obvious that equation (4.6) can never be written in a linear appearance, i.e. it can not be found a matrix $\mathbf{F}_k$ such that the process model could be rearranged in form (3.2). For these reasons, in practice, dual estimation can be performed only with the filters developed for non-linear systems, such as EKF or PF.

4.2.2 Reduced model estimation

Let us now describe how the model reduction affects the formulation of the filter. Recalling what has been explained in Chapter 2, all the components of the full model will be replaced by the components of the reduced model:

$$\mathbf{u}_k \rightarrow \mathbf{u}_{r,k} \tag{4.7}$$

$$\mathbf{f}_k(\mathbf{u}_{k-1}) \rightarrow \mathbf{f}_{r,k}(\mathbf{u}_{r,k-1}) \tag{4.8}$$

and of course all the noises will be referred to these new variables.
From equation (4.3), the process model is then:

$$\mathbf{x}_{r,k} = \begin{bmatrix} \mathbf{u}_{r,k} \\ \boldsymbol{\theta}_k \end{bmatrix} = \mathbf{f}_{r,k}(\mathbf{u}_{r,k-1}, \boldsymbol{\theta}_{k-1}) + \mathbf{w}_k \tag{4.9}$$

Regarding the measurement model, since the signals are of course collected in the full model environment but the state vector lies in the reduced space, equation (4.4) is modified as follows:

$$\mathbf{y}_k = \mathbf{H}_k^{\mathbf{u}} \mathbf{L}_k \mathbf{x}_{r,k} + \mathbf{v}_k \tag{4.10}$$

The matrix $\mathbf{L}_k$ is an appropriate matrix needed to switch from the reduced subspace of estimations to the full space of the measurement:

$$\mathbf{L}_k = \begin{bmatrix} \boldsymbol{\Phi}_{l,k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{4.11}$$

The matrix $\boldsymbol{\Phi}_{l,k} = [\boldsymbol{\phi}_{1,k} \cdots \boldsymbol{\phi}_{l,k}]$ collects the bases needed to project and reduce the original full order space. For instance, if the method used to reduce the system is the POD (see Section 2.2), then $\boldsymbol{\Phi}_{l,k}$ collects the first l POMs. It is interesting to note that if the system is not reduced, we have:

$$\mathbf{L}_k = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{4.12}$$

Table 4.1: Kalman filter algorithm for sub-space updating

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{\Phi}}_{l,0} = E[\mathbf{\Phi}_{l,0}]$$

$$\mathbf{P}_0^{ss} = E[(\mathbf{\Phi}_{l,0} - E[\mathbf{\Phi}_{l,0}])(\mathbf{\Phi}_{l,0} - E[\mathbf{\Phi}_{l,0}])^T]$$

2. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage

$$\hat{\mathbf{\Phi}}_{l,k}^{-} = \hat{\mathbf{\Phi}}_{l,k-1}$$

$$\mathbf{P}_k^{ss,-} = \mathbf{P}_{k-1}^{ss} + \mathbf{W}_k^{ss}$$

   (b) Updating stage

$$\mathbf{G}_k^{ss} = \mathbf{P}_k^{ss,-}\mathbf{H}_k^{ssT}(\mathbf{H}_k^{ss}\mathbf{P}_k^{i-}\mathbf{H}_k^{ssT} + \mathbf{V}_k)^{-1}$$

$$\hat{\mathbf{\Phi}}_{l,k} = \hat{\mathbf{\Phi}}_{l,k}^{-} + \mathbf{G}_k^{ss}(\mathbf{y}_k - \mathbf{H}_k^{ss}\hat{\mathbf{\Phi}}_{l,k}^{-})$$

$$\mathbf{P}_k^{ss} = \mathbf{P}_k^{ss,-} - \mathbf{G}_k^{ss}\mathbf{H}_k^{ss}\mathbf{P}_k^{ss,-}$$

that leads of course to equation (4.4).
In the case of systems whose characterisctics do not change in time, the measurement and process matrices are stationary and thus they do not depend on the step k in which are calculated:

$$\left.\begin{array}{c}\mathbf{F}_k = \mathbf{F} \\ \mathbf{H}_k = \mathbf{H}\end{array}\right\} \quad \forall k \in [1, N] \tag{4.13}$$

On the other hand, considering a non-stationary system, the latter matrices should be updated at each step. Since also the sub-space has to be modified, the estimation of $\mathbf{\Phi}_{l,k}$ should be included and a new equation needed to allow the variation of the sub-space is introduced, as done for the parameters in (4.2), according to:

$$\mathbf{\Phi}_{l,k} = \mathbf{\Phi}_{l,k-1} + \mathbf{w}_k^{ss} \tag{4.14}$$

where $\mathbf{w}_k^{ss}$ is a fictitious zero mean gaussian noise, whose covariance should be tuned in order to reach the best estimation. Having defined the process model of the sub-space, the measurement model can be obtained. Using equation (4.10) it is possible to derive the following expression:

$$\mathbf{y}_k = \mathbf{H}_k^{\mathbf{u}}\mathbf{\Phi}_{l,k}\mathbf{u}_{r,k} + \mathbf{v}_k \tag{4.15}$$

Rearranging the terms it is possible to lead to:

$$\mathbf{y}_k = \mathbf{H}_k^{ss}\mathbf{\Phi}_{l,k} + \mathbf{v}_k = [u_{r,k}^1\mathbf{H}_k^{\mathbf{u}} \ \cdots \ u_{r,k}^{N_r}\mathbf{H}_k^{\mathbf{u}}]\mathbf{\Phi}_{l,k} + \mathbf{v}_k \tag{4.16}$$

where N_r is the the dimension of the reduced system and $u_{r,k}^{1,..,N_r}$ are the elements of the reduced vector $\mathbf{u}_{r,k}$. In order for the matricial products to be consistent, the matrix $\mathbf{\Phi}_{l,k}$ is rearranged as follows:

$$\mathbf{\Phi}_{l,k} = \begin{bmatrix} \boldsymbol{\phi}_{1,k} \\ \vdots \\ \boldsymbol{\phi}_{l,k} \end{bmatrix} \tag{4.17}$$

Since the sub-space process model described in equation (4.14) is basically an identity transformation, for less than the fictitious noise, and therefore it is linear, a Kalman filter can be used to estimate $\mathbf{\Phi}_{l,k}$.
Recalling what has been explained in the present section, Table (4.1) shows the application of the Kalman filter to the sub-space updating.

4.2.3 Hybrid extended Kalman particle filter

The particle filter described in Section 3.4 presents some disadvantages: since the filter is not optimal considering the latest observations, a high number of samples is required to describe the probability distribution and therefore, in some cases, the computational cost can be high. In [62], a modified version of the particle filter has been presented in order to speed up the procedure. The basic idea is to use an extended Kalman filter in order to update the set of samples generated, taking into account the last observation available. Comparing Tables 4.2 and 3.3, the only difference from the standard Particle filter is the exploitation of a sample updating stage:

- first of all the particles are generated from the the conditional pdf $p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$;
- in order to push the samples towards higher probability zones, they are modified taking into account the latest observations at time t_k, through an extended Kalman filter;
- finally the re-sampling technique is used, leading to the updating of the estimated state vector.

The employment of this type of modified particle filter, called hybrid extended Kalman particle filter (HEKPF), guarantees a higher accuracy than the standard particle filter [106]. We should remember, in fact, that the accuracy of the particle filter is lowered by the non optimal choice of the importance function: as explained in Section (3.4.2) the Gaussian distribution is the most used distribution, even if it is not the optimal one. Moreover, in [106] it has been shown that HEKPF leads to a lower computational cost.

Table 4.2: Hybrid extended Kalman particle filter algorithm (HEKPF)

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0]$$
$$\mathbf{P}_0 = E[(\mathbf{x}_0 - E[\mathbf{x}_0])(\mathbf{x}_0 - E[\mathbf{x}_0])^T]$$
$$\mathbf{x}_0^i = \hat{\mathbf{x}}_0$$
$$\omega_0^i = p(\mathbf{y}_0|\mathbf{x}_0)$$

2. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage ($i = 1, ..., N_s$)

$$\mathbf{x}_k^i \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$$
$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k|\mathbf{x}_k^i)$$

   (b) EKF updating stage ($i = 1, ..., N_s$)

$$\mathbf{P}_k^{i-} = \mathbf{F}_k \mathbf{P}_{k-1}^i \mathbf{F}_k^T + \mathbf{W}_k$$
$$\mathbf{G}_k^i = \mathbf{P}_k^{i-} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^{i-} \mathbf{H}_k^T + \mathbf{V}_k)^{-1}$$
$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^{i-} + \mathbf{G}_k^i \mathbf{y}_k - \mathbf{h}_k \hat{\mathbf{x}}_k^{i-}$$
$$\mathbf{P}_k^i = \mathbf{P}_k^{i-} - \mathbf{G}_k^i \mathbf{H}_k \mathbf{P}_k^{i-}$$

   (c) Resampling stage ($i = 1, ..., N_s$)

$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k|\mathbf{x}_k^i)$$
$$u_i \sim U[0, 1]$$
$$\text{find} \quad m \quad \text{s.t.} \quad \sum_{i=1}^{m-1} \omega_k^i < u_i \leq \sum_{i=1}^{m} \omega_k^i$$
$$\bar{\mathbf{x}}_k^i = \mathbf{x}_k^m$$
$$\bar{\omega}_k^{*i} = \frac{1}{N_s}$$

   (d) Updating stage

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N_s} \bar{\omega}_k^i \bar{\mathbf{x}}_k^i$$

4.3 Dual estimation of reduced states and parameters with subspace updating

At this point, since all the required computational tools and procedures have been previously introduced, it is now possible to explain the algorithm that will be used for damage identification. The method is characterized by the following features:

- the model order reduction guarantees a low computational cost, since the number of estimates is reduced (see Section (4.2.2));
- the projection $\mathbf{\Phi}_{l,k}$ needed to perform the model order reduction is updated taking into account the latest observations, through a Kalman filter (see Section (4.2.2)); therefore, the basis $\mathbf{\Phi}_{l,k}$ is modified accordingly with the possible variations of the system;
- the performances of the particle filter employed to estimate the reduced state vector $\mathbf{x}_k$ are improved using two remedies:
  - an extended Kalman filter pushes the samples towards high probabilty zones (as explained in Section (4.2.3));
  - a resampling method speeds up the procedure (sse Section (3.4.3)).

In summary, the dual estimation of reduced states and parameters with subspace updating algorithm is shown in Table 4.3.

Table 4.3: EK-PF-KF algorithm

1. Initialization ($t_k = t_0$)

$$\hat{\mathbf{x}}_{r,0} = L_0^T E[\mathbf{x}_0]$$

$$\mathbf{P}_{r,0} = L_0^T E[(\mathbf{x}_0 - E[\mathbf{x}_0])(\mathbf{x}_0 - E[\mathbf{x}_0])^T] L_0$$

$$\mathbf{x}_{r,0}^i = \hat{\mathbf{x}}_{r,0}$$

$$\omega_0^i = p(\mathbf{y}_0 | \mathbf{x}_{r,0})$$

$$\hat{\mathbf{\Phi}}_{l,0} = E[\mathbf{\Phi}_{l,0}]$$

$$\mathbf{P}_0^{ss} = E[(\mathbf{\Phi}_{l,0} - E[\mathbf{\Phi}_{l,0}])(\mathbf{\Phi}_{l,0} - E[\mathbf{\Phi}_{l,0}])^T]$$

1. Recursive computation ($t_k = t_1, ..., t_N$)

   (a) Prediction stage ($i = 1, ..., N_s$)

$$\mathbf{x}_k^i \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}^i)$$

$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k|\mathbf{x}_k^i)$$

   (b) EKF updating stage ($i = 1, ..., N_s$)

$$\mathbf{P}_k^{i-} = \mathbf{F}_k \mathbf{P}_{k-1}^i \mathbf{F}_k^T + \mathbf{W}_k$$

$$\mathbf{G}_k^i = \mathbf{P}_k^{i-} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^{i-} \mathbf{H}_k^T + \mathbf{V}_k)^{-1}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}^i}_k^{i-} + \mathbf{G}_k^i \mathbf{y}_k - \mathbf{h}_k \hat{\mathbf{x}}_k^{i-}$$

$$\mathbf{P}_k^i = \mathbf{P}_k^{i-} - \mathbf{G}_k^i \mathbf{H}_k \mathbf{P}_k^{i-}$$

   (c) Resampling stage ($i = 1, ..., N_s$)

$$\omega_k^i = \omega_{k-1}^i p(\mathbf{y}_k|\mathbf{x}_k^i)$$

$$u_i \sim U[0, 1]$$

$$\text{find} \quad m \quad \text{s.t.} \quad \sum_{i=1}^{m-1} \omega_k^i < u_i \leq \sum_{i=1}^{m} \omega_k^i$$

$$\bar{\mathbf{x}}_k^i = \mathbf{x}_k^m$$

$$\bar{\omega}_k^{*i} = \frac{1}{N_s}$$

   (d) Updating stage

$$\hat{\mathbf{x}}_k = \sum_{i=1}^{N_s} \bar{\omega}_k^i \bar{\mathbf{x}}_k^i$$

   (e) Subspace prediction and updating stage

$$\mathbf{\Phi}_{l,k}^- = \mathbf{\Phi}_{l,k-1}$$

$$\mathbf{P}_k^{ss,-} = \mathbf{P}_{k-1}^{ss} + \mathbf{W}_k^{ss}$$

$$\mathbf{G}_k^{ss} = \mathbf{P}_k^{ss,-} \mathbf{H}_k^{ssT} (\mathbf{H}_k^{ss} \mathbf{P}_k^{i-} \mathbf{H}_k^{ssT} + \mathbf{V}_k)^{-1}$$

$$\mathbf{\Phi}_{l,k} = \mathbf{\Phi}_{l,k}^- + \mathbf{G}_k^{ss} (\mathbf{y}_k - \mathbf{H}_k^{ss} \mathbf{\Phi}_{l,k}^-)$$

$$\mathbf{P}_k^{ss} = \mathbf{P}_k^{ss,-} - \mathbf{G}_k^{ss} \mathbf{H}_k^{ss} \mathbf{P}_k^{ss,-}$$

4.4 Damage detection and localization

4.4.1 Damage parameters

As the theoretical framework has been explained, the procedure used to detect and localize the damage in a mechanical system is now introduced.
The structural problem is described by equation (2.1), here recalled for sake of completeness:

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{D}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = \mathbf{F}(t) \tag{4.18}$$

where the meaning of the matrices was given in Section 2.1.
In order to be able to assess the damage state of the structure, we need to introduce some damage parameters into the model. The idea is to divide the structure in N_p (number of parameters) regions and express the complete stiffness matrix as a linear combination of some appropriate stiffness matrices associated to those regions. Let us first of all consider a reference "undamaged" structure, in which all the regions are characterized by a Young modulus E. Then, consider N_p auxiliary structures defined as follows: the i-th structure is characterized by having a Young modulus $\bar{E}$ on the i-th region, while all the other regions have the original Young modulus E. In other words, we will have to build N_p different auxiliary structures, such that each of them is identical to the reference structure, apart from one element. The definition of $\bar{E}$ is purely arbitrary and can be expressed as:

$$\bar{E} = (1 - \bar{\kappa})E \tag{4.19}$$

where the factor $\bar{\kappa}$ can be any number $\in (0, 1)$.
Let us now focus on the real damaged structure that has to be identified: each region will be characterized by its own Young modulus E_i that can be expressed as a function of the original Young modulus E as follows:

$$E_i = (1 - d_i)E \tag{4.20}$$

The coefficients $d_i \in [0, 1)$ are the damage parameters that will be estimated by the filter.
Define now the stiffness matrices of the aforementioned reference and auxiliary structures repectively as $\mathbf{K}_{und} \in \mathbb{R}^{N_{dof} \times N_{dof}}$ and $\mathbf{K}_i \in \mathbb{R}^{N_{dof} \times N_{dof}}$. Therefore the stiffness matrix of the real damaged structure is equal to:

$$\mathbf{K} = \sum_{i=1}^{N_p} E_i \frac{\mathbf{K}_{und} - \mathbf{K}_i}{E - \bar{\kappa}E} \tag{4.21}$$

From equation (4.21) it is straightforward to understand that the choice of $\bar{\kappa}$ can be arbitrary and it is basically delegated to the user; since $\bar{\kappa}$ is required only to somehow normalize the difference $(\mathbf{K}_{und} - \mathbf{K}_i)$, its value does not affect the

results.
Replacing equation (4.20) in (4.21), we get:

$$\mathbf{K} = \sum_{i=1}^{N_p} E(1-d_i)\frac{\mathbf{K}_{und} - \mathbf{K}_i}{E - \bar{\kappa}E} \tag{4.22}$$

and hence we lead to the final formulation of the stiffness matrix:

$$\mathbf{K} = \sum_{i=1}^{N_p} \frac{1-d_i}{1-\bar{\kappa}}(\mathbf{K}_{und} - \mathbf{K}_i) \tag{4.23}$$

In this way, we managed to express the stiffness matrix as a function of the damage parameters that we want to estimate:

$$\mathbf{K} = \mathbf{K}(d_1, ..., d_{N_p}) \tag{4.24}$$

and therefore this expression can be applied to the filtering procedure that will be explained afterwards.
Let us now discuss the reason this type of formulation has been developed and used and which are its main advantages. It is interesting to point out that, considering the definition of the stiffness matrix given in equation (4.23), no internal parameters of the finite element (FE) model are introduced. In other words, the damage parameters are externally defined and hence we do not need to estimate any parameter that depend on how each stiffness matrix has been built. This choice presents two main advantages:

- in order to build the stiffness matrices, the method can be coupled to any FE commercial code, even if the exact formulation used to build them is unknown; in other words we do not require to know a parametric version of the stiffness matrix, but just a numerical one is needed.

- the method is very flexible, in fact, since it is not important how the stiffness matrix has been built, the type of finite elements and discretization used does not affect the definition of the damage parameters;

- this strategy could be applied potentially to any type of structure.

At this point, it is now possible to apply the latter concepts to the framework previously defined in (4.3) and therefore describe the algorithm used to identify the damage parameters d_i.

4.4.2 Procedure description

The state vector is defined as:

$$\mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \dot{\mathbf{u}} \\ \ddot{\mathbf{u}} \\ \mathbf{d} \end{bmatrix} \tag{4.25}$$

where $\mathbf{u} \in \mathbb{R}^{N_{dof}}$ and $\mathbf{d}$ is defined as:

$$\mathbf{d} = \begin{bmatrix} d_1 \\ \ddots \\ d_{N_p} \end{bmatrix}$$

is the vector that contains the local stiffness parameters and its size is equal to the number of regions N_p.
For sake of computational cost, the model is reduced using a POD procedure. The equation (4.23) is projected onto the actual subspace, where $\mathbf{\Phi} \in \mathbb{R}^{N_{dof} \times l}$ is the matrix that collects the first l actual proper orthogonal modes of the structure:

$$\mathbf{\Phi}^T \mathbf{K} \mathbf{\Phi} = \mathbf{\Phi}^T \left[\sum_{i=1}^{N_p} \frac{1 - d_i}{1 - \bar{\kappa}} (\mathbf{K}_{und} - \mathbf{K}_i)) \right] \mathbf{\Phi} \tag{4.26}$$

i.e.:

$$\mathbf{K}_r = \sum_{i=1}^{N_p} \frac{1 - d_i}{1 - \bar{\kappa}} (\mathbf{K}_{und,r} - \mathbf{K}_{i,r}) \tag{4.27}$$

In the same way, the cinematic components of the state vector are projected as well, becoming:

$$\mathbf{x}_r = \begin{bmatrix} \boldsymbol{\alpha} \\ \dot{\boldsymbol{\alpha}} \\ \ddot{\boldsymbol{\alpha}} \\ \mathbf{d} \end{bmatrix} \tag{4.28}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^l$.
It has been shown in [64] that the operation of a filter is not affected by damping. On the other hand, the only recommendation required for a damped system is to have the system continuously excited. Here, for sake of simplicity, an undamped case is considered; hence, assuming $\mathbf{D} = \mathbf{0}$, the reduced dynamic equation is:

$$\mathbf{M}_r \ddot{\boldsymbol{\alpha}}(t) + \mathbf{K}_r \boldsymbol{\alpha}(t) = \mathbf{F}_r(t) \tag{4.29}$$

Once the system has been described, let us concentrate on the process model:

$$\mathbf{x}_{r,k} = \mathbf{f}_{r,k}(\mathbf{x}_{r,k-1}) + \mathbf{w} \tag{4.30}$$

where the process noise $\mathbf{w}$ is supposed to be a white gaussian noise.
In order to solve the differential equation 2.1, the Newmark explicit integration method [107] is used (Table 4.4).

Table 4.4: Newmark explicit integration method

1. Initialization

$$\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}(t_0)$$
$$\dot{\boldsymbol{\alpha}}_0 = \dot{\boldsymbol{\alpha}}(t_0)$$
$$\ddot{\boldsymbol{\alpha}}_0 = \mathbf{M}^{-1}(-\mathbf{C}\dot{\boldsymbol{\alpha}}_0 - \mathbf{K}(\chi_k)\boldsymbol{\alpha}_0 + \mathbf{F}(t_0))$$

2. Recursive computation (k=1,...,N)

   (a) Predictor stage

$$\tilde{\boldsymbol{\alpha}}_k = \boldsymbol{\alpha}_{k-1} + \Delta t\dot{\boldsymbol{\alpha}}_{k-1} + \Delta t^2\left(\frac{1}{2} - \beta\right)\ddot{\boldsymbol{\alpha}}_{k-1}$$

$$\dot{\tilde{\boldsymbol{\alpha}}}_k = \dot{\boldsymbol{\alpha}}_{k-1} + \Delta t(1-\gamma)\ddot{\boldsymbol{\alpha}}_{k-1}$$

   (b) Explicit integration

$$\ddot{\boldsymbol{\alpha}}_k = (\mathbf{M} + \gamma\Delta t\mathbf{C} + \beta\Delta t^2\mathbf{K}(\boldsymbol{\chi}_k))^{-1}(-\mathbf{C}\dot{\tilde{\boldsymbol{\alpha}}}_k - \mathbf{K}(\chi_k)\tilde{\boldsymbol{\alpha}}_k + \mathbf{F}_k)$$

   (c) Correction stage

$$\dot{\boldsymbol{\alpha}}_k = \dot{\tilde{\boldsymbol{\alpha}}}_k + \Delta t\gamma\ddot{\boldsymbol{\alpha}}_k$$
$$\boldsymbol{\alpha}_k = \tilde{\boldsymbol{\alpha}}_k + \Delta t^2\beta\ddot{\boldsymbol{\alpha}}_k$$

Let us now define:

$$\bar{\boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{\alpha} \\ \dot{\boldsymbol{\alpha}} \\ \ddot{\boldsymbol{\alpha}} \end{bmatrix} \tag{4.31}$$

The previous expressions are collected all together and rearranged in a matricial form, obtaining:

$$\bar{\boldsymbol{\alpha}}_k = \bar{\mathbf{f}}(\bar{\boldsymbol{\alpha}}_{k-1}) = \mathbf{A}_{k-1}\bar{\boldsymbol{\alpha}}_{k-1} + \mathbf{B}_{k-1} \tag{4.32}$$

where

$$\mathbf{A}_{k-1} = \begin{bmatrix} \mathbf{I} - \beta\Delta t^2\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & \Delta t\mathbf{I} - \beta\Delta t^3\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & \\ -\Delta t\gamma\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & \mathbf{I} - \Delta t^2\gamma\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & \cdots \\ -\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & -\Delta t\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} & \end{bmatrix} \tag{4.33}$$

$$\begin{matrix} & \begin{matrix} \Delta t^2(1/2-\beta)[\mathbf{I}-\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1}\Delta t^2\beta] \end{matrix} \\ \cdots & \begin{matrix} (1-\gamma)dt\mathbf{I}-\Delta t^3\gamma(1/2-\beta)\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} \\ -\Delta t^2(1/2-\beta)\mathbf{M}_{k-1}^{-1}\mathbf{K}_{k-1} \end{matrix} \end{matrix} \Bigg]$$

$$\mathbf{B}_{k-1} = \begin{bmatrix} \Delta t^2\beta\mathbf{M}_{k-1}^{-1}\mathbf{F}_k \\ \Delta t\gamma\mathbf{M}_{k-1}^{-1}\mathbf{F}_k \\ \mathbf{M}_{k-1}^{-1}\mathbf{F}_k \end{bmatrix} \tag{4.34}$$

and

$$\mathbf{K}_{k-1} = \mathbf{\Phi}_{k-1}^T \left[\sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und} - \mathbf{K}_i)) \right] \mathbf{\Phi}_{k-1} =$$

$$= \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r) \tag{4.35}$$

$$\mathbf{M}_{k-1} = \mathbf{\Phi}_{k-1}^T \mathbf{M}_{und} \mathbf{\Phi}_{k-1} = \mathbf{M}_{k-1}^r \tag{4.36}$$

$$\mathbf{F}_k = \mathbf{\Phi}_{k-1}^T \mathbf{F}_k^f = \mathbf{F}_k^r \tag{4.37}$$

Equation (4.32), together with $\mathbf{d}_k = \mathbf{d}_{k-1}$, provides the function $\mathbf{f}_{r,k}$.

$$\mathbf{x}_k = \begin{bmatrix} \bar{\boldsymbol{\alpha}}_j \\ \mathbf{d}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\mathbf{d}_{k-1})\bar{\boldsymbol{\alpha}}_{k-1} + \mathbf{B}(\mathbf{d}_{k-1}) \\ \mathbf{d}_{k-1} \end{bmatrix} = \mathbf{f}_{r,k}(\mathbf{x}_{r,k-1}) \tag{4.38}$$

The three components of $\bar{\boldsymbol{\alpha}}_j$ are:

$$\begin{aligned} \boldsymbol{\alpha}_k = \mathbf{I} - \beta\Delta t^2[\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\boldsymbol{\alpha}_{k-1} + \\ + \Delta t\mathbf{I} - \beta\Delta t^3[\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\dot{\boldsymbol{\alpha}}_{k-1} + \\ + \Delta t^2(1/2-\beta)\{\mathbf{I} - [\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\Delta t^2\beta\}\ddot{\boldsymbol{\alpha}}_{k-1} \end{aligned} \tag{4.39}$$

$$\begin{aligned} \dot{\boldsymbol{\alpha}}_k = -\gamma\Delta t[\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\boldsymbol{\alpha}_{k-1} + \\ + \mathbf{I} - \gamma\Delta t^2[\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\dot{\boldsymbol{\alpha}}_{k-1} + \\ + \Delta t(1-\gamma)\mathbf{I} - \Delta t^3\gamma(1/2-\beta)[\mathbf{M}_{k-1}^r]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}_{und,k-1}^r - \mathbf{K}_{i,k-1}^r)\ddot{\boldsymbol{\alpha}}_{k-1} \end{aligned} \tag{4.40}$$

$$\ddot{\boldsymbol{\alpha}}_k = -[\mathbf{M}^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}^r_{und,k-1} - \mathbf{K}^r_{i,k-1}) \boldsymbol{\alpha}_{k-1} +$$
$$- \Delta t [\mathbf{M}^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}^r_{und,k-1} - \mathbf{K}^r_{i,k-1}) \dot{\boldsymbol{\alpha}}_{k-1} +$$
$$- \Delta t^2 (1/2-\beta) [\mathbf{M}^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (\mathbf{K}^r_{und,k-1} - \mathbf{K}^r_{i,k-1}) \ddot{\boldsymbol{\alpha}}_{k-1} \quad (4.41)$$

The computation of the jacobian matrix of $\mathbf{f}_{r,k}(\mathbf{x})$ is required:

$$\mathbf{F}_{r,k} = \begin{bmatrix} \frac{\partial \mathbf{f}_{r,1}(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial \mathbf{f}_{r,1}(\mathbf{x})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{f}_{r,n}(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial \mathbf{f}_{r,n}(\mathbf{x})}{\partial x_n} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_{k-1}} = \nabla_{\mathbf{x}} \mathbf{f}_{r,k}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_{k-1}} \quad (4.42)$$

The dimension of the state vector is $n = 3l + N_p$, where l is the dimension of the reduced order model, $3l$ is required because even the first and second derivatives are considered in the state vector and N_p is the number of the damage parameters, i.e. the number of the regions in which the structure is divided.
Equations from (4.43) to (4.58) show the elements of the Jacobian matrix $\mathbf{F}_{r,k} = \nabla_x \mathbf{f}_{r,k}(\mathbf{x})$, for the case of only one single POM retained. The selection of only one POM makes the mass and stiffness matrices scalars and hence simplifies the exposition.

$$\frac{\partial \alpha_k}{\partial \alpha_{k-1}} = 1 - \beta \Delta t^2 [M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (K^r_{und,k-1} - K^r_{i,k-1}) \quad (4.43)$$

$$\frac{\partial \alpha_k}{\partial \dot{\alpha}_{k-1}} = \Delta t - \beta \Delta t^3 [M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (K^r_{und,k-1} - K^r_{i,k-1}) \quad (4.44)$$

$$\frac{\partial \alpha_k}{\partial \ddot{\alpha}_{k-1}} = \Delta t^2 (1/2-\beta) \{1 - [M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (K^r_{und,k-1} - K^r_{i,k-1}) \Delta t^2 \beta\} \quad (4.45)$$

$$\frac{\partial \alpha_k}{\partial d_{i,k-1}} = \Delta t^2 \frac{K^r_{und,k-1} - K^r_{i,k-1}}{1-\bar{\kappa}} [M^r_{k-1}]^{-1} (\beta \alpha_{k-1} + \Delta t \dot{\alpha}_{k-1} + \Delta t^2 (1/2-\beta) \beta \ddot{\alpha}_{k-1}) \quad (4.46)$$

$$\frac{\partial \dot{\alpha}_k}{\partial \alpha_{k-1}} = -\gamma \Delta t [M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (K^r_{und,k-1} - K^r_{i,k-1}) \quad (4.47)$$

$$\frac{\partial \dot{\alpha}_k}{\partial \dot{\alpha}_{k-1}} = 1 - \gamma \Delta t^2 [M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}} (K^r_{und,k-1} - K^r_{i,k-1}) \quad (4.48)$$

$$\frac{\partial \dot{\alpha}_k}{\partial \ddot{\alpha}_{k-1}} = \Delta t(1-\gamma) - \Delta t^3 \gamma(1/2-\beta)[M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}}(K^r_{und,k-1} - K^r_{i,k-1}) \tag{4.49}$$

$$\frac{\partial \dot{\alpha}_k}{\partial d_{i,k-1}} = \gamma \Delta t \frac{K^r_{und,k-1} - K^r_{i,k-1}}{1-\bar{\kappa}}[M^r_{k-1}]^{-1}(\alpha_{k-1} + \Delta t \dot{\alpha}_{k-1} + \Delta t^2(1/2-\beta)\beta \ddot{\alpha}_{k-1}) \tag{4.50}$$

$$\frac{\partial \ddot{\alpha}_k}{\partial \alpha_{k-1}} = -[M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}}(K^r_{und,k-1} - K^r_{i,k-1}) \tag{4.51}$$

$$\frac{\partial \ddot{\alpha}_k}{\partial \dot{\alpha}_{k-1}} = -\Delta t[M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}}(K^r_{und,k-1} - K^r_{i,k-1}) \tag{4.52}$$

$$\frac{\partial \ddot{\alpha}_k}{\partial \ddot{\alpha}_{k-1}} = -\Delta t^2(1/2-\beta)[M^r_{k-1}]^{-1} \sum_{i=1}^{N_p} \frac{1-d_{i,k-1}}{1-\bar{\kappa}}(K^r_{und,k-1} - K^r_{i,k-1}) \tag{4.53}$$

$$\frac{\partial \ddot{\alpha}_k}{\partial \chi_{i,k-1}} = \frac{K^r_{und,k-1} - K^r_{i,k-1}}{1-\bar{\kappa}}[M^r_{k-1}]^{-1}(\alpha_{k-1} + \Delta t \dot{\alpha}_{k-1} + \Delta t^2(1/2-\beta)\beta \ddot{\alpha}_{k-1}) \tag{4.54}$$

$$\frac{\partial d_{i,k}}{\partial \alpha_k} = 0 \tag{4.55}$$

$$\frac{\partial d_{i,k}}{\partial \dot{\alpha}_k} = 0 \tag{4.56}$$

$$\frac{\partial d_{i,k}}{\partial \ddot{\alpha}_k} = 0 \tag{4.57}$$

$$\frac{\partial d_{i,k}}{\partial d_{i,k-1}} = 1 \tag{4.58}$$

Having specialized all the components of the method for the structural case, we can now summarize the procedure in Figure 4.1. The process, measurement and fictitious noises are all supposed to be white zero-mean Gaussian and uncorrelated to each other. Some further explanations are needed.
First of all, the matrix $\mathbf{L}_k$ is equal to:

$$\mathbf{L}_k = \begin{bmatrix} \boldsymbol{\Phi}_{l,k} & & & \\ & \boldsymbol{\Phi}_{l,k} & & \\ & & \boldsymbol{\Phi}_{l,k} & \\ & & & \mathbf{0} \end{bmatrix} \tag{4.59}$$

This particular formulation of matrix $\mathbf{L}_k$ has been proposed in [25]: each sub-matrix $\boldsymbol{\Phi}_{l,k}$ collects the first l POMs considered and basically can project each dynamic variable. In other words, three concatenated sub-matrices are required because the state vector is built considering both the reduced dynamic variables

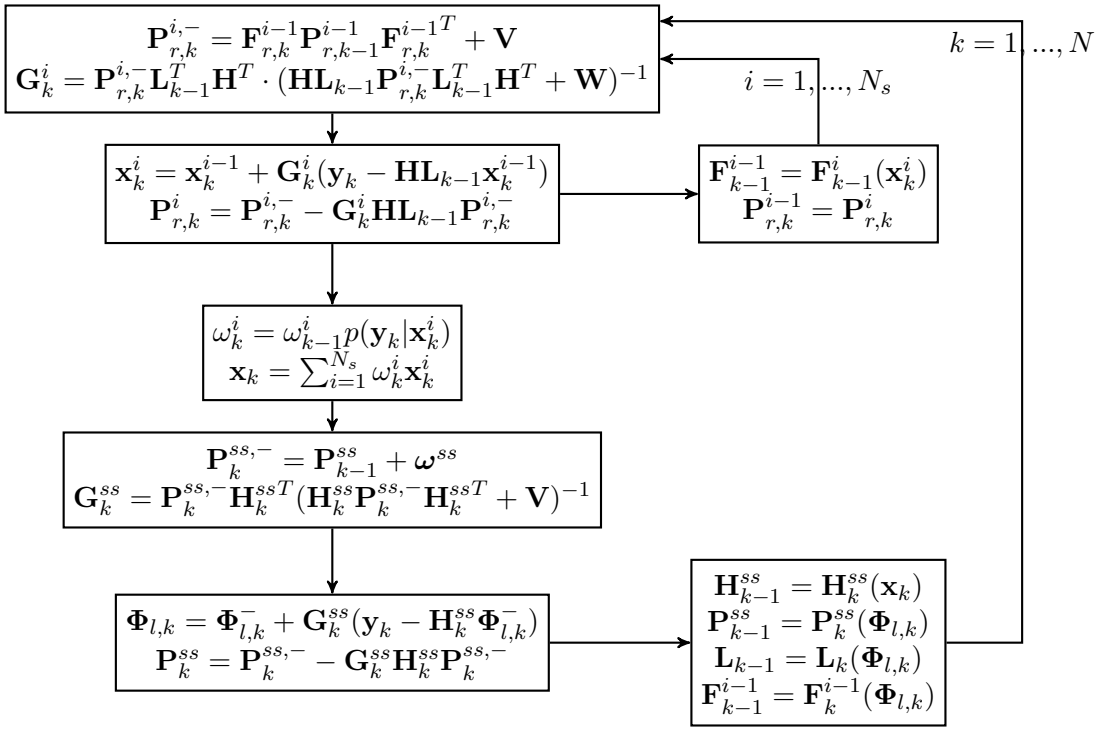


Figure 4.1: Flow chart of the procedure used for dual estimation of the reduced model and sub-space update

$\boldsymbol{\alpha}_k$ and their first and second derivatives $\dot{\boldsymbol{\alpha}}_k$ and $\ddot{\boldsymbol{\alpha}}_k$. Since the damage parameters are not projected into the reduced sub-space, a null matrix is added. Let us now consider the measurement model:

$$\mathbf{y}_k = \mathbf{H}\mathbf{L}_{k-1}\mathbf{x}_k + \mathbf{w} \tag{4.60}$$

The measurement matrix $\mathbf{H} \in \mathbb{R}^{N_{obs} \times 3N_{dof}+N_p}$ links the N_{obs} observations and the full order state vector, that has dimension $3N_{dof} + N_p$. Remember that $\mathbf{x}_k$ is the reduced state vector: the transformation $\mathbf{L}_{k-1}$ is required because we want to link the measurement vector, that lies in the full order space, and the state vector, that lies in the reduced sub-space and, therefore, $\mathbf{x}_k$ should be re-projected into the full space.
In order to evolve the weights ω^i_k, the following multivariate Gaussian distribution has been used to generate samples from $p(\mathbf{y}_k|\mathbf{x}^i_k)$:

$$p(\mathbf{y}_k|\mathbf{x}^i_k) = \frac{1}{\sqrt{(2\pi)^n|\mathbf{P}_{k-1}|}}\exp\left\{-\frac{1}{2}(\mathbf{y}_k - \mathbf{H}\mathbf{L}_{k-1}\mathbf{x}^i_k)^T\mathbf{P}^{-1}_{k-1}(\mathbf{y}_k - \mathbf{H}\mathbf{L}_{k-1}\mathbf{x}^i_k)\right\} \tag{4.61}$$

Regarding the subspace updating stage, the matrix $\boldsymbol{\omega}^{ss}$ is defined as follows:

$$\boldsymbol{\omega}^{ss} = \begin{bmatrix} \mathbf{W}^{ss} & & \\ & \mathbf{W}^{ss} & \\ & & \mathbf{W}^{ss} \end{bmatrix} \tag{4.62}$$

where $\mathbf{W}^{ss}$ are the covariances of the fictitious noise associated to the subspace. The repetition of sub-matrices $\mathbf{W}^{ss}$ is required in order to take into account the first and second derivatives of the dynamical variables as well. Finally the matrix $\mathbf{H}_k^{ss}$ is:

$$\mathbf{H}_k^{ss} = [\alpha_k^1 \mathbf{H}^{\mathbf{u}} \ \cdots \ \alpha_k^l \mathbf{H}^{\mathbf{u}}] \boldsymbol{\Phi}_{l,k} + \mathbf{v}_k \tag{4.63}$$

basically following the idea explained in equation (4.16). In equation (4.63), the matrix $\mathbf{H}^{\mathbf{u}}$ is defined in $\mathbb{R}^{N_{obs} \times N_{dof}}$, therefore in order for the matricial products to be consistent, the matrix $\boldsymbol{\Phi}_{l,k}$ is defined in $\mathbb{R}^{l \cdot N_{dof} \times 1}$. In other words, the POMs are concatenated as:

$$\boldsymbol{\Phi}_{l,k} = \begin{bmatrix} \phi_{1,k} \\ \vdots \\ \phi_{l,k} \end{bmatrix} \tag{4.64}$$

4.5 Summary

In the present chapter the damage identification algorithm has been described in details.
In Section 4.2 the concepts of dual estimation, reduced order model and hybrid extended Kalman particle filter has been given. Then in Section 4.3 the algorithm used to perform the dual estimation of reduced states and parameters with subspace updating has been reviewed. Finally in Section 4.4 all the previous tools have been used to describe the damage identification procedure. Furthermore, the damage parameters d_i have been defined and the formulation of the stiffness matrix is given: the particular strategy described in Section 4.4.1 allows to release the estimation method from the finite element formulation used to discretize the system. Therefore, the method can be easily applied to any type of structure and type of finite element used.
In Chapter 5 the performances of the damage identification procedure will evaluated, performing a benchmark analysis on a simple thin plate.

Chapter 5

Results

5.1 Introduction

In the present chapter the performances of the method described in Table (4.1) will be investigated and some remarks about the choice and influence of the settings of the filter will be described.
The structure considered is a simple square plate, whose dimensions are 200 mm$\times$ 200 mm and thickness of 5 mm. The material chosen is Aluminum 6061-T6 [108]. Since the forces applied are very low, the material is supposed to always remain inside the elastic domain. The Young modulus is equal to $E = 68.9$ MPa and the density is $\rho = 2.5 \cdot 10^3$ kg/m^3.
The plate is modeled through the commercial finite element (FE) code Abaqus (Abaqus/CAE 6.10-1 ©Dassault Systemes, 2010), using the S4R general-purpose conventional shell elements, that take into account transverse shear deformations. For further details on the elements, the reader may refer to [109]. This type of elements is characterized by having 6 degrees of freedom per node, i.e. the displacements and rotations in the three directions:

$$\mathbf{u} = \begin{bmatrix} u_x \\ u_y \\ u_z \\ \varphi_x \\ \varphi_y \\ \varphi_z \end{bmatrix} \tag{5.1}$$

Figure 5.1: Numbering of regions and points of the structure

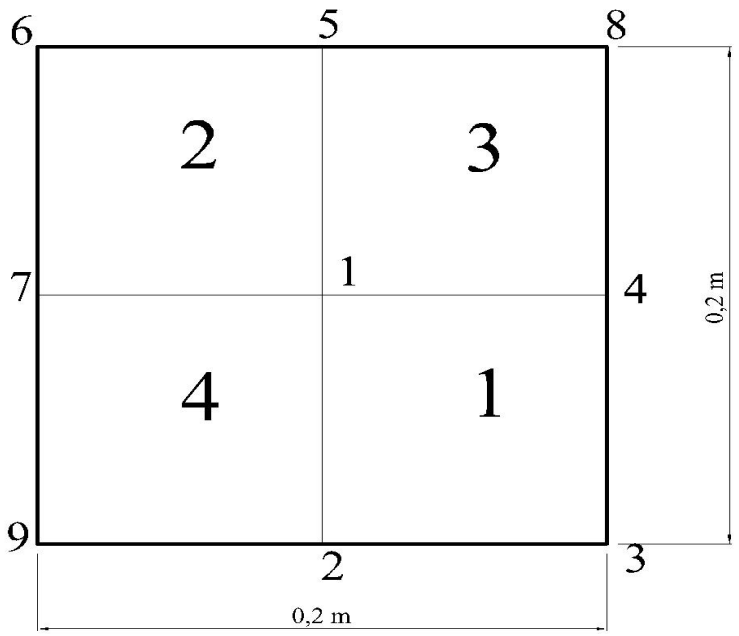
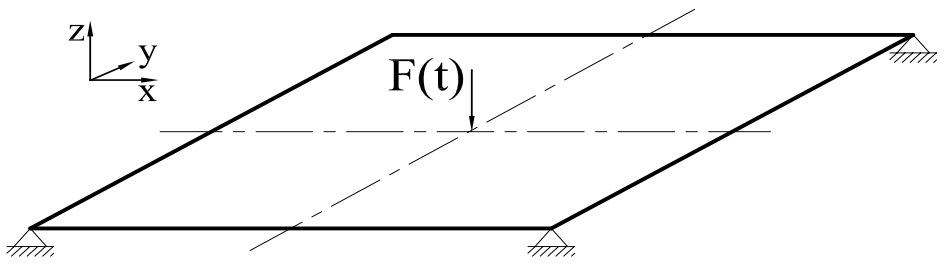
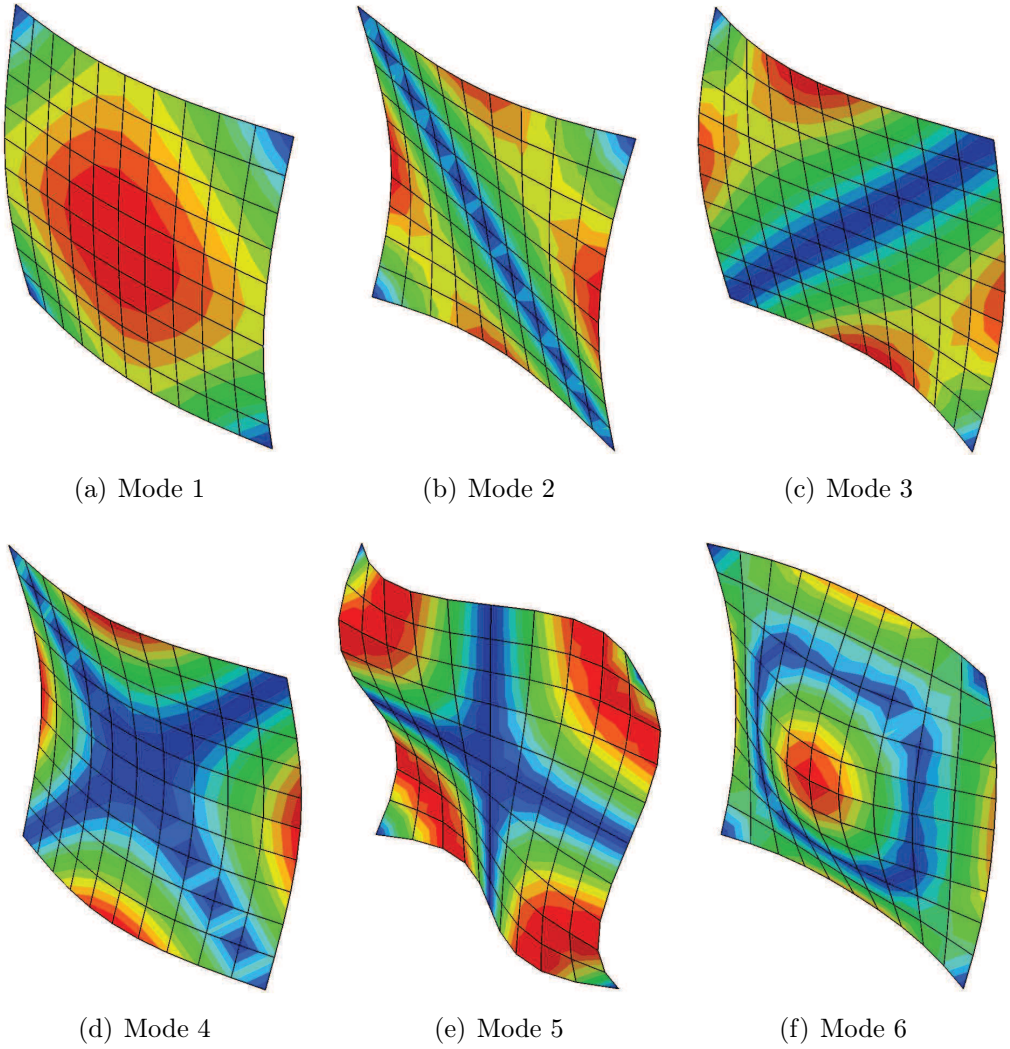


Figure 5.2: Boundary conditions and load





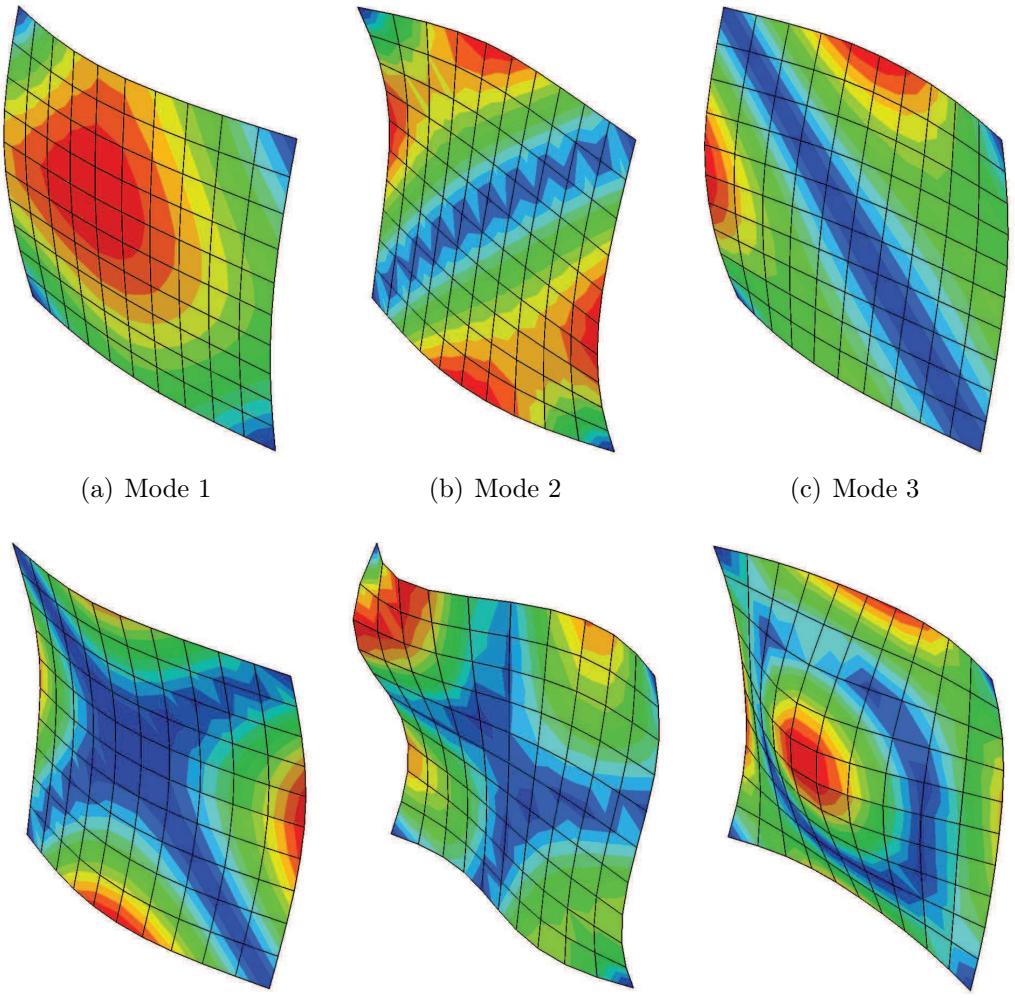
(a) Mode 1 (b) Mode 2 (c) Mode 3

(d) Mode 4 (e) Mode 5 (f) Mode 6

Figure 5.3: Undamaged case - Shapes of the first six natural modes

In Figure 5.1, the numbering of some remarkable points and the regions in which the structure has been divided are shown. Regarding the boundary conditions, the plate is simply supported at the four corners, such that $u_z^3 = u_z^6 = u_z^8 = u_z^9 = 0$. The plate is subjected to a sinusoidal force applied to point 1, in the center of the plate and along the direction z:

$$\mathbf{F}(t) = \begin{bmatrix} 0 \\ 0 \\ A\sin(\omega t) \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{5.2}$$



(a) Mode 1 (b) Mode 2 (c) Mode 3

(d) Mode 4 (e) Mode 5 (f) Mode 6

Figure 5.4: Damaged case - Shapes of the first six natural modes

Figure 5.2 shows the boundary conditions and the load applied. In [110], it has been shown that the choice of the load does not affect the operation of the filter, therefore only one type of load will be used here. The load amplitude is equal to $A = 100$ N and the angular velocity is $\omega = 500$ rad/s.
Two configurations are used:

1. an undamaged structure characterized by a Young modulus E associated to all the regions;

2. a damaged structure, characterized by having a Young modulus reduced to $E/2$ within region 2 and a Young modulus E in all the other ones.

In Section 5.2 a benchmark analysis is performed applying the method to the aforementioned thin plate modeled through a 4-elements discretization, both in the undamaged and damaged case. The performances of the plate will be evaluated as a function of the number of POMs, the initial conditions, the measurement noise, the process noise and the number of sensors deployed.
Afterwards, the damage identification method will be applied to the same damaged thin plate, modeled using a more refined mesh (10×10 elements mesh) and therefore a higher number of degrees of freedom.

5.2 Benchmark analysis

In order to benchmark the method, the most simple discretization of the structure has been used: since the FE model should be able to allow the deformation of the structure, the minimum number of elements that can do it is 4. The nodes are therefore 9 and the number of total degrees of freedom is 54. In the following subsections, first of all a comparison between the full order model and the reduced order model will be shown, evaluating the performances of the proper orthogonal decomposition in terms of the dynamic response. Then, the estimation method will be tested on the damaged structure, evaluating the ability of tracking both the unknown damage parameters d_i and the dynamic evolution of the system.
All the simulations are performed using the Newmark integration scheme described in Table 4.4, with $\beta = 1/4$ and $\gamma = 1/2$. This particular choice of β and γ guarantees the unconditional stability of the integration method [111]. All the following simulations are performed using $\Delta t = 5 \cdot 10^{-5}$ s.
The relevant first natural frequencies of the damaged structure using the 4-elements aforementioned discretization are shown in Table 5.1.

Table 5.1: Natural frequencies of the damaged structure

vibration mode index	1	2	3	4	5	6
natural frequency [Hz]	281	497	500	507	515	649

The relevant first natural frequencies of the undamaged system are shown in Table 5.2.

Table 5.2: Natural frequencies of the undamaged structure

vibration mode index	1	2	3	4	5	6
natural frequency [Hz]	315	540	540	552	552	695

The natural frequencies of the damaged structure are of course lower than the ones associated to the undamaged structure, because of the corresponding

reduction of stiffness at constant mass properties. It is interesting to note that since the undamaged structure is not only geometrical but also structurally symmetrical, then some modes of vibration are symmetrical and therefore the natural frequencies associated are identical.

5.2.1 Model order reduction

5.2.1.1 Undamaged structure

Let us first of all consider the undamaged structure.
The first step is the calculation of the POMs associated to the structure subjected to the load previously defined. The POMs are calculated through the proper orthogonal decomposition described in Section 2.2.3, considering a certain number of snapshots N_{snap} of the dynamic evolution of the system. The choice and influence of N_{snap} on the results will be discussed later on. Once the POMs and their respective singular values have been found, a first way to look at the performances of the model order reduction is to calculate the index $I(l)$ defined in equation (2.39), where l is the number of POMs retained and thus the order of the reduced model. This index gives a direct information about the energy accuracy of the reduction, without requiring the comparison of the time histories.

Table 5.3: Undamaged case - $I(l)$

l	$I(l)$
1	0.981632
2	0.999999
3	≈ 1

From table 5.3 we can see that for example if the projection onto the subspace is performed using only one single POM, the level of information retained by the reduced model is still very high. This is due to the symmetry of the model and to the type of the load applied, that basically stimulate the first modes. We will see later on that using 2 POMs a perfect match of the full model is obtained and therefore the POMs of higher order are basically useless. The level of information associated to the 3-POMs model is not reported because physically speaking can be assumed as 1 and it goes beyond the machine epsilon. The choice of N_{snap} depends on the convergence of the POMs: in order to measure this issue, in Figure 5.5 it is possible to see the evolution of the L^2 relative error between the POM calculated for a certain number of snapshots and the one calculated for 200 snapshots. The corresponding time interval of the training stage is approximately 3 times the fundamental period of vibration of the structure:

$$t_{train}^{200} = N_{snap} \cdot \Delta t = 200 \cdot 5 \cdot 10^{-5}\text{s} = 0.01 \text{ s} \approx 3 \cdot \frac{1}{f_1} = 3 \cdot \frac{1}{315.3} \text{ s} \qquad (5.3)$$

The relative error is calculated as:

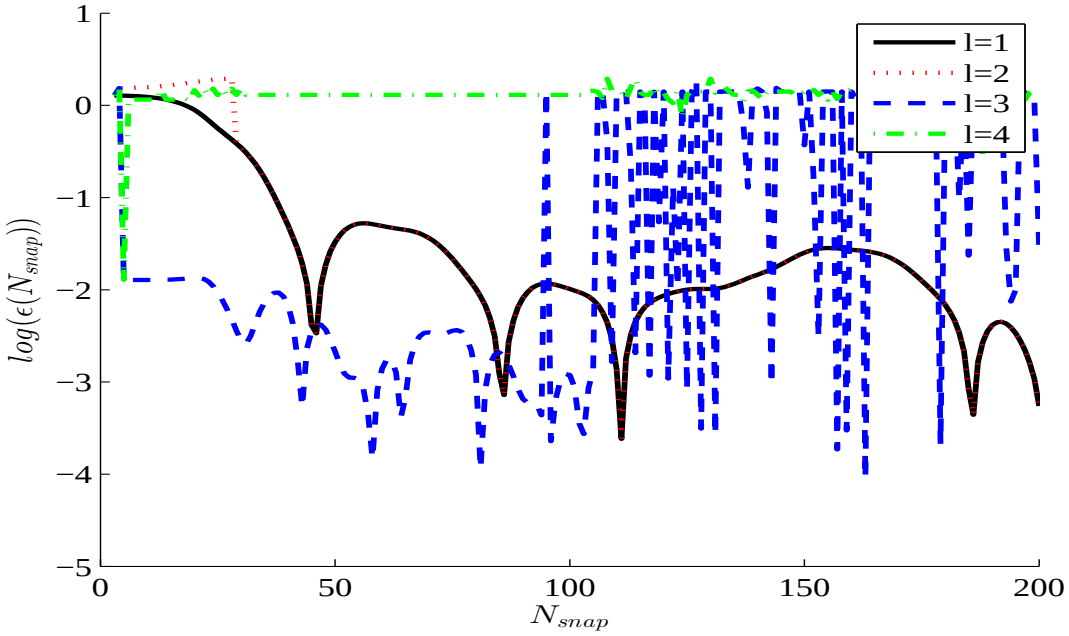
$$\epsilon_l(N_{snap}) = \frac{||\boldsymbol{\phi}_l(N_{snap}) - \boldsymbol{\phi}_l(200)||_{L^2}}{||\boldsymbol{\phi}_l(200)||_{L^2}} \tag{5.4}$$

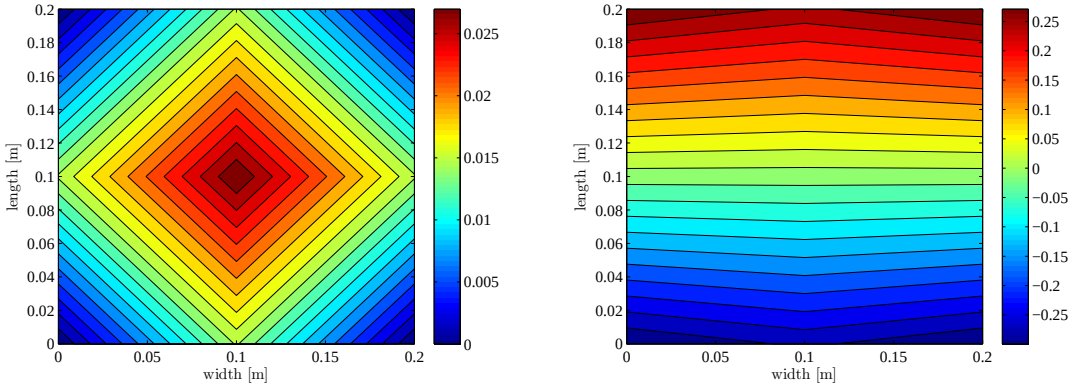
where the L^2 norm of an arbitrary vector $\mathbf{v}$ is defined as:

$$||\mathbf{v}||_{L^2} = \sqrt{\sum_{i=1}^{size(\mathbf{v})} |v_i|} \tag{5.5}$$

It is interesting to point out that the first two POMs converge very quickly, in fact just after approximately 40 snapshots considered, the relative error is under 10%. In [112] it has been shown that the convergence of the POMs depends not only on the structure, but also on the excitation used. In this case, as it is possible to see from Figure 5.5, a training time t_{train} approximately equal to the fundamental period of vibration ($N_{snap} \approx 60$) guarantees an acceptable convergence. The convergence of the third and fourth POM is not reached or there are continuous fluctuations over the number of snapshots: in accordance with table 5.3 we can see that the increase of the number of POMs over $l = 2$ carries a negligible fraction of information (or energy). This means that the exploitation of all the $l > 2$ POMs is basically useless, since the additional level of information retained is practically zero. For this reason, from now on, in all the comparisons associated to the undamaged structure, we will retain only the first two POMs.
Figures 5.6 and 5.7 shows the displacement and rotation maps of the first two POMs. In other words, the values of one degrees of freedom, for instance the displacements u_z, for each node are selected from the POM and reported to the figure. Since both the structures, the load and the boundary conditions are symmetric, only the non-zero components of $\mathbf{u}$ are shown: therefore the displacements u_x and u_y and rotation φ_z does not appear. For the same reason, the maps shown are symmetric.
The last remark regards the comparison between the POMs and the natural modes of the structure: as shown in figure 5.8, the shapes does not perfectly match. This statement is in accordance with what has been explained in Section 2.4.2, i.e. only if the structure is subjected to a resonant load, the POMs converge to the relevant natural modes.

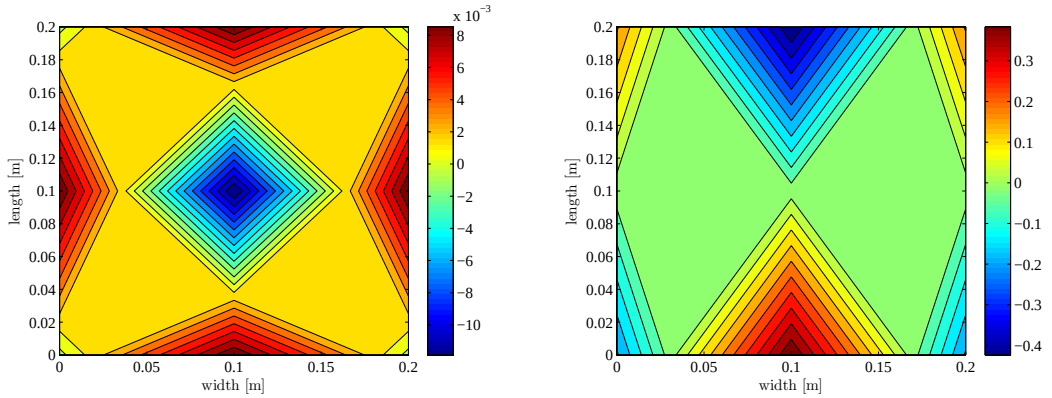
Figure 5.5: Undamaged case - POM convergence $\epsilon_l(N_{snap}) = \frac{||\boldsymbol{\phi}_l(N_{snap}) - \boldsymbol{\phi}_l(200)||_{L^2}}{||\boldsymbol{\phi}_l(200)||_{L^2}}$





(a) u_z (b) φ_x

Figure 5.6: Undamaged case - POM 1 components maps



(a) u_z (b) φ_x

Figure 5.7: Undamaged case - POM 2 components maps

Let us now compare the time histories of the rotations and displacements obtained using the full and reduced order models. Since the dynamic evolution of the degrees of freedom of point 1 is not interesting because it is directly correlated with the applied load, we focus on a side middle point, i.e. node 2. In this case, since the structure keeps the symmetry even when it is subjected to the load, the choice of the node is not important. Having previously discussed about the convergence of the POMs, in all the simulation we will use $N_{snap} = 200$.
Figures from 5.9 to 5.17 show the evolution of displacement u_z and rotation φ_x in node 2, comparing the results obtained with the full or reduced models with 1 or 2 POMs retained. The index of accuracy for the third POM is so high that it does not worth the results to be reported in the figures. Considering Figures from 5.9 to 5.14, some remarks can be emphasized. First of all, as previously shown in Table 5.3, in all the figures the 2-POM order reduction can be considered as a very good approximation of the full model, in fact both the displacements, the rotations and their derivatives have been catched very well. Regarding the 1-POM order reduction, it can be stated that only the displacements and the rotations can fit the real evolution in an acceptable way. This is due to the fact that the POD is performed using displacements and rotations and therefore the POMs are based on them. Nevertheless, only the lower frequency of the signal is reproduced correctly, while the fluctuations associated to the other ones do not suit the real one. This behavior can be underlined calculating the Fourier transform F of the signals (figure 5.15): the second peak of the 1-POM signal is basically wrong, while the first frequency corresponds perfectly. Looking at the second derivative, the discrepancy regards not only the frequencies, but also the amplitude, that is lower then the real one; this is basically due to the reduction of flexibility of the reduced order model. Also in this case, as shown in Figures 5.16 and 5.17, the reduced 1-POM model signal matches only the first frequency.

Figure 5.8: Undamaged case - Comparison between 1^{st} (red solid line) POM and 1^{st} (blue dashdot line) natural mode

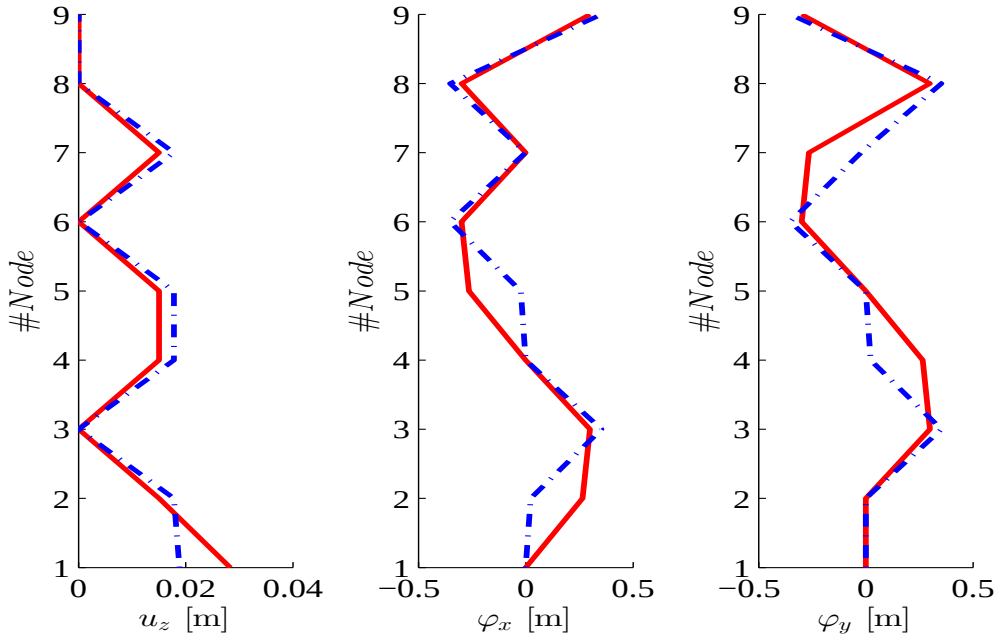


Figure 5.9: Undamaged case - Point 2 - Displacement $u_z(t)$

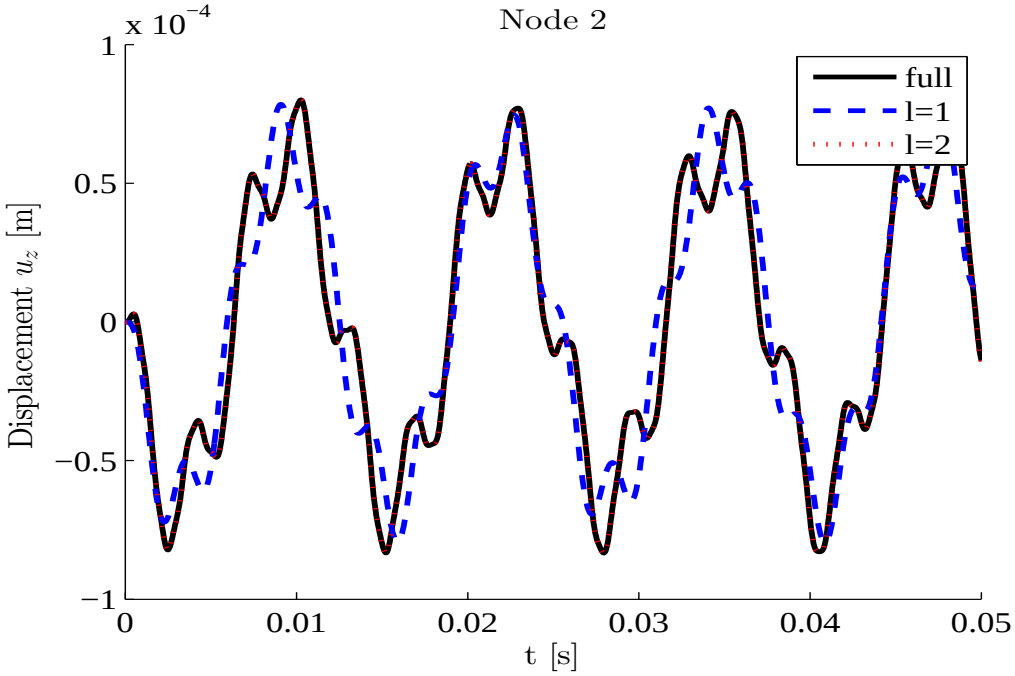


Figure 5.10: Undamaged case - Point 2 - Rotation $\varphi_x(t)$

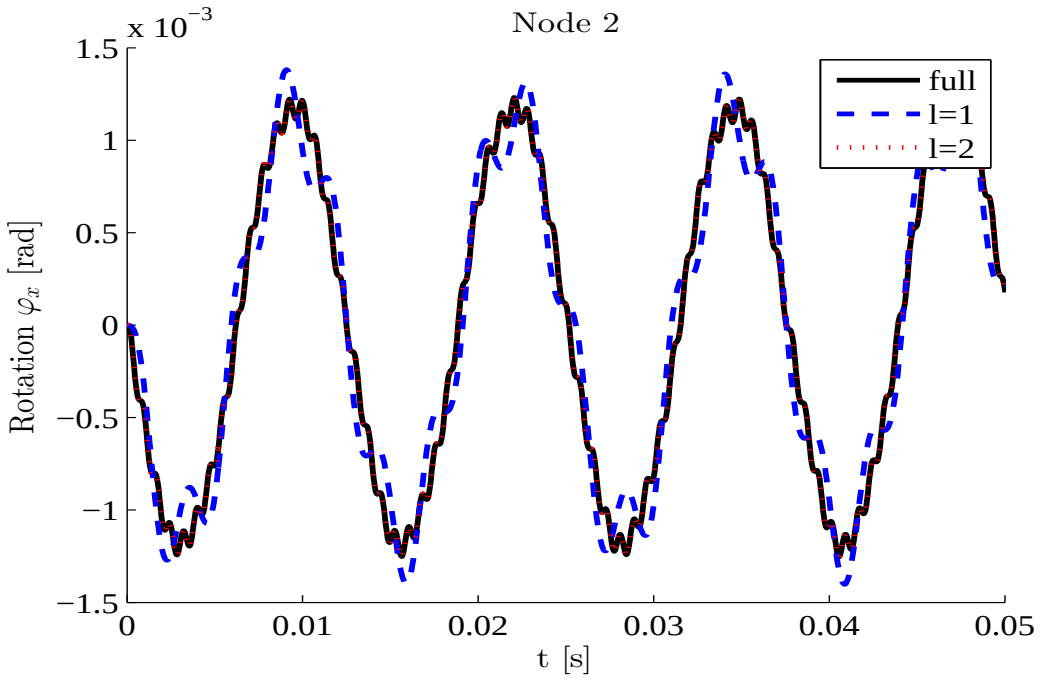


Figure 5.11: Undamaged case - Point 2 - $\dot{u}_z(t)$

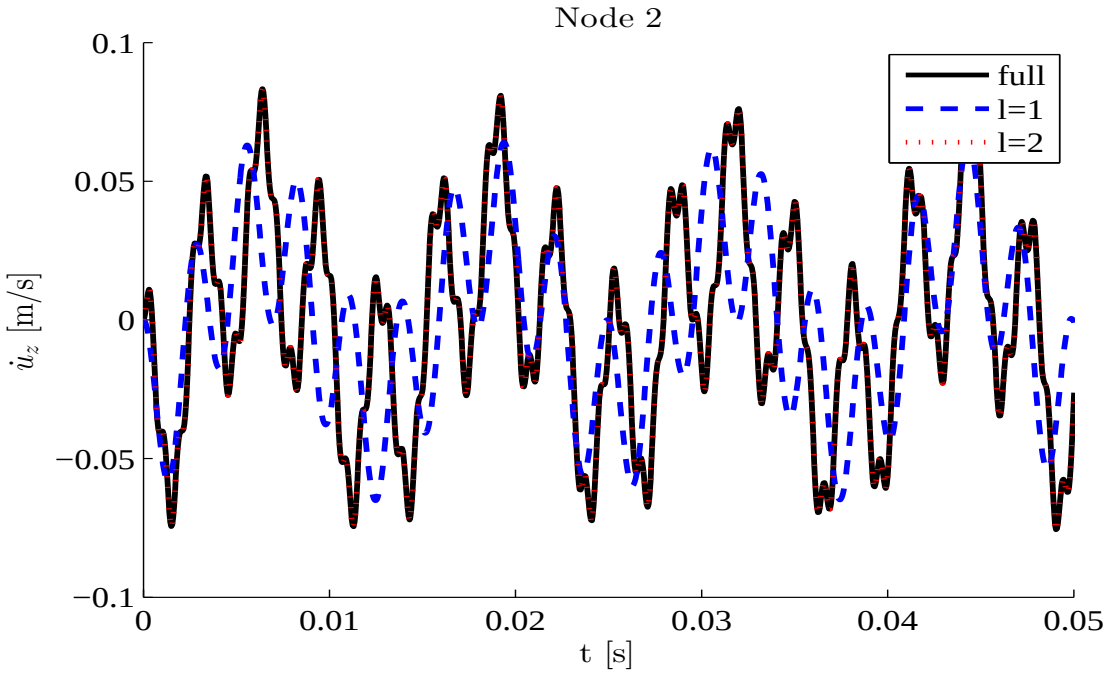


Figure 5.12: Undamaged case - Point 2 - $\dot{\varphi}_x(t)$

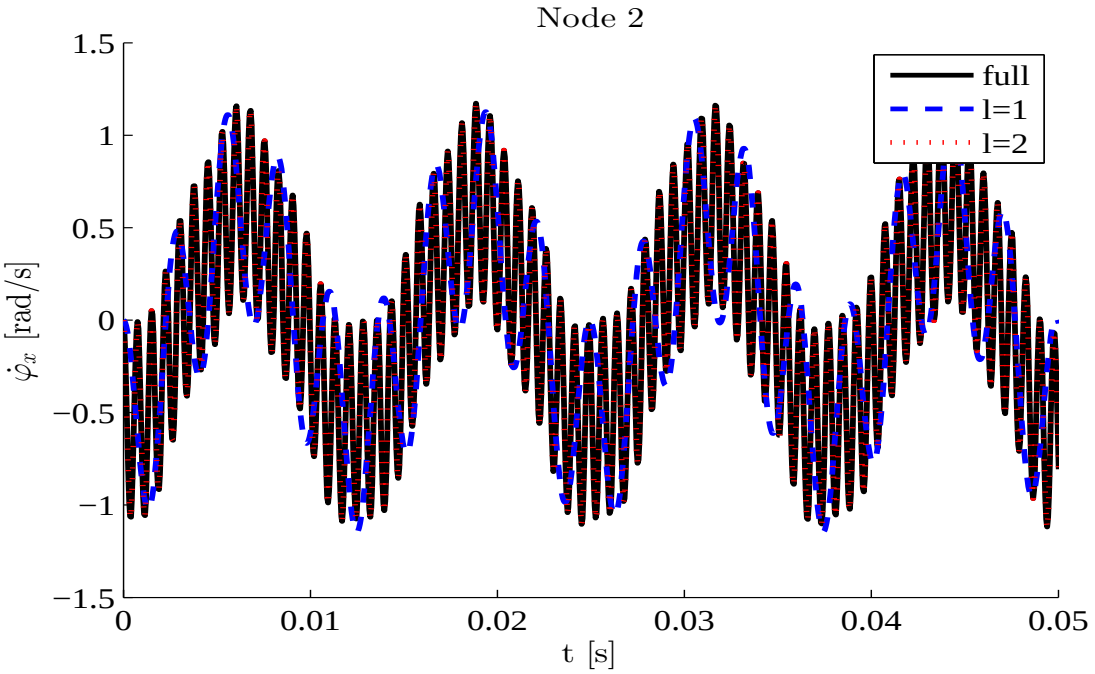


Figure 5.13: Undamaged case - Point 2 - $\ddot{u}_z(t)$

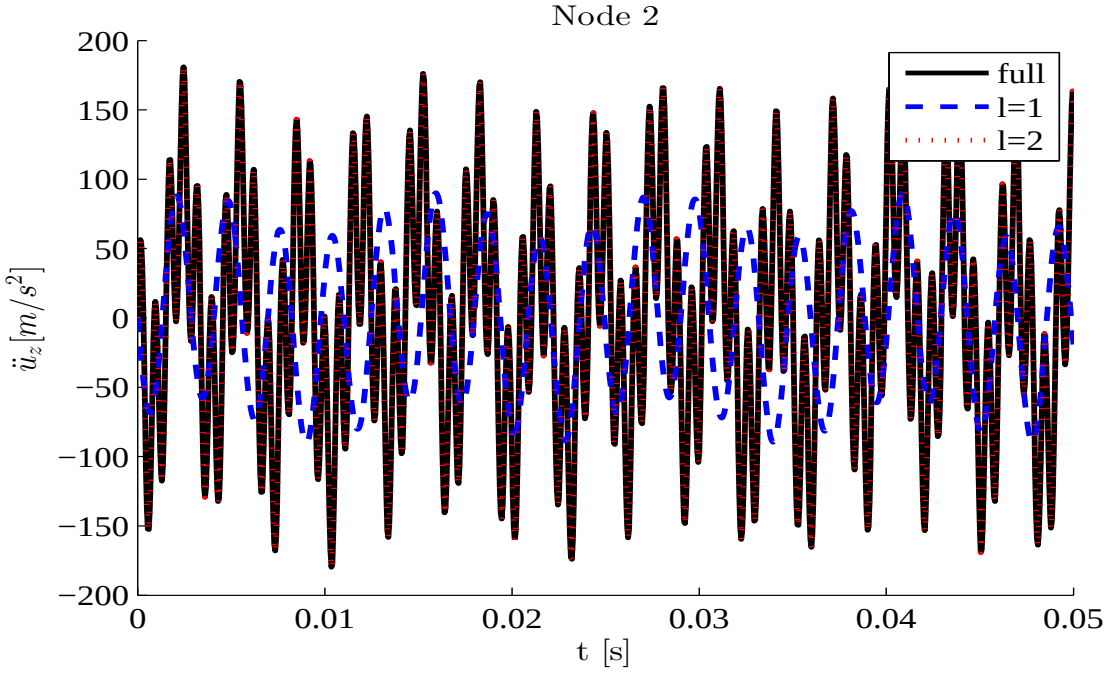


Figure 5.14: Undamaged case - Point 2 - $\ddot{\varphi}_x(t)$

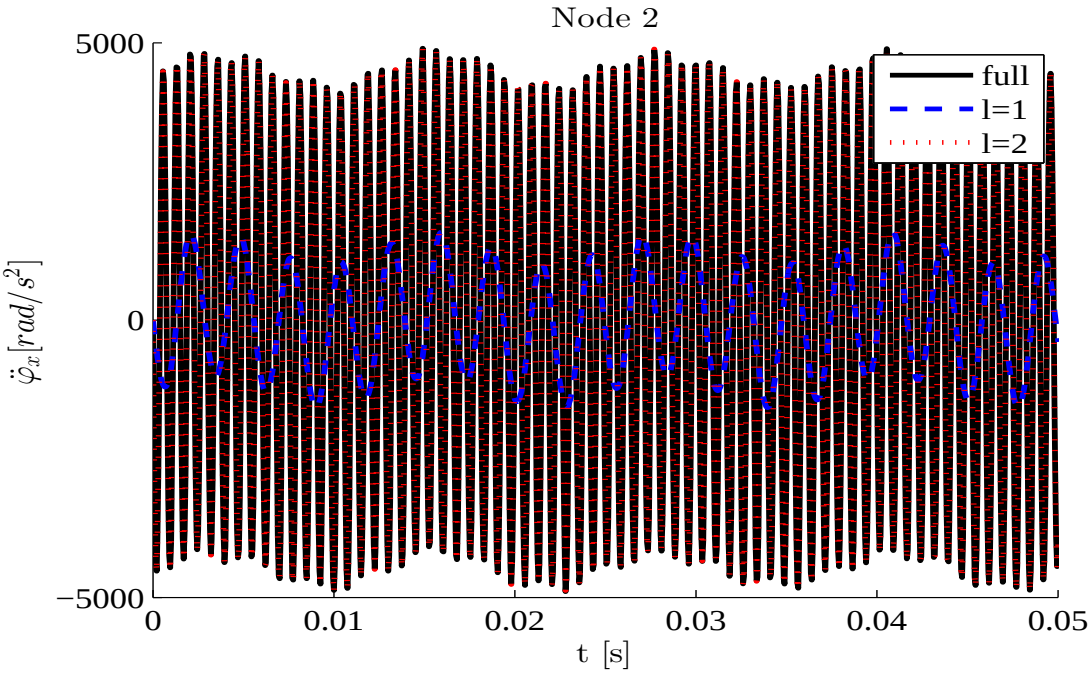


Figure 5.15: Undamaged case - Point 2 - $\mathcal{F}[u_z(t)]$

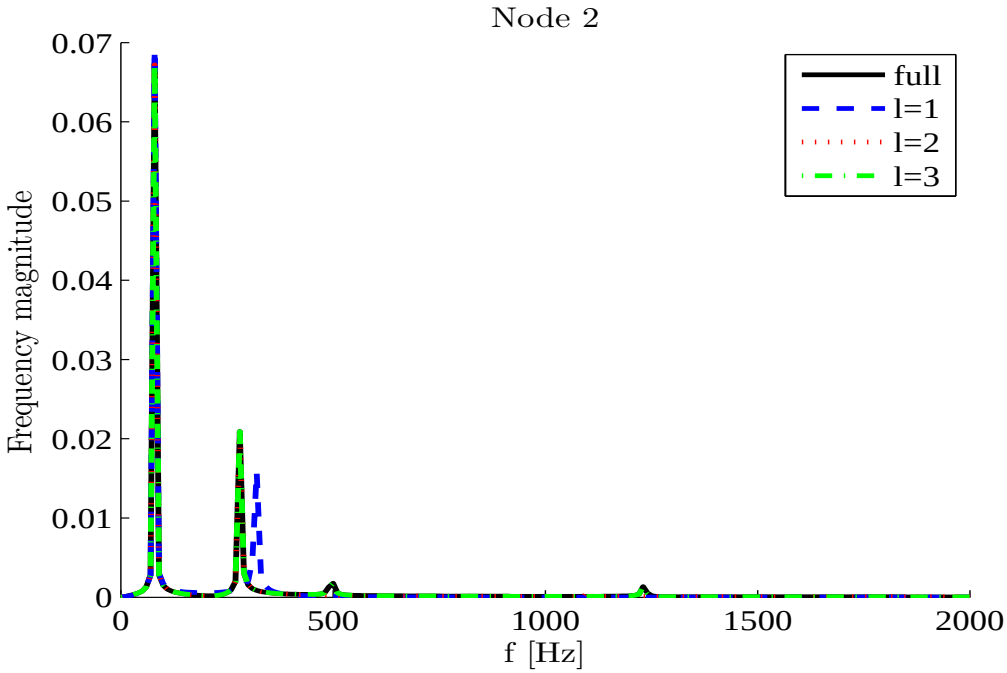


Figure 5.16: Undamaged case - Point 2 - $\mathcal{F}[\dot{u}_z(t)]$

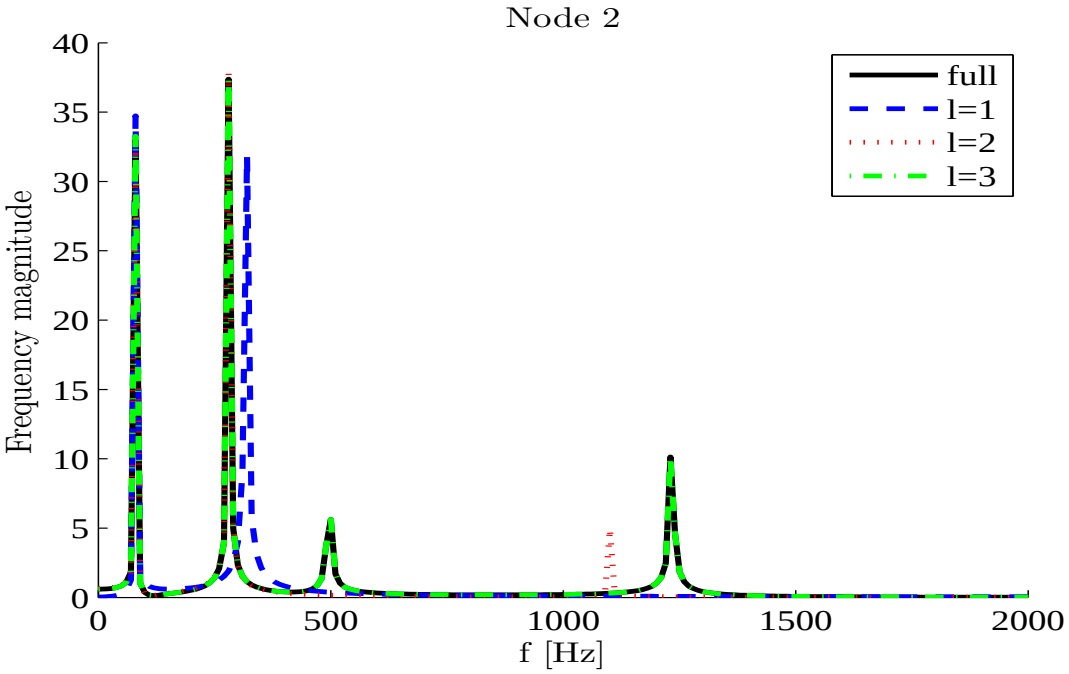
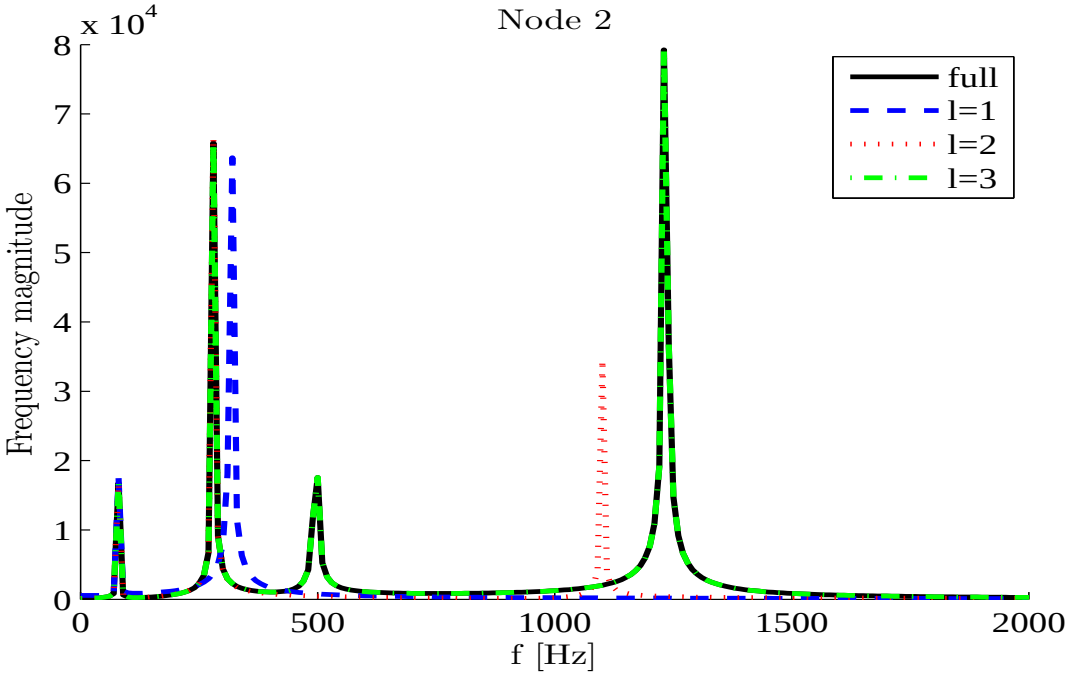


Figure 5.17: Undamaged case - Point 2 - $\mathcal{F}[\ddot{u}_z(t)]$



Let us now consider in Figure 5.18 the relative error between the exact evolution of the system and the reduced ones:

$$\epsilon_l(t) = \frac{||\mathbf{u}_l(t) - \mathbf{u}(t)||_{L^2}}{||\mathbf{u}(t)||_{L^2}} \tag{5.6}$$

First of all, it is possible to note that, in accordance with what has been previously explained, $\epsilon_1(t)$ is almost six order of magnitude higher than $\epsilon_2(t)$, while there is not a substantial difference between the latter one and $\epsilon_3(t)$. Another important remark is that the relative error is almost stationary; in order to better explain this concept, let us consider the cumulative error $\bar{\epsilon}_l(t)$, calculated as:

$$\bar{\epsilon}_l(t) = \int_o^t \epsilon_l(\tau) d\tau \tag{5.7}$$

In figure 5.19 we can see that the cumulative error $\bar{\epsilon}_1(t)$ associated to the first POM grows over time. It is important to point out that, since the slope of the cumulative error decreases over time, the accuracy of the approximation increases. For the same reason, since the cumulative errors $\bar{\epsilon}_2(t)$ and $\bar{\epsilon}_3(t)$ are already very low compared to $\bar{\epsilon}_1(t)$, the variation of the accuracy over time in these cases is low.

Figure 5.18: Undamaged case - Relative error - $\epsilon_l(t) = \frac{||\mathbf{u}_l(t)-\mathbf{u}(t)||_{L^2}}{||\mathbf{u}(t)||_{L^2}}$

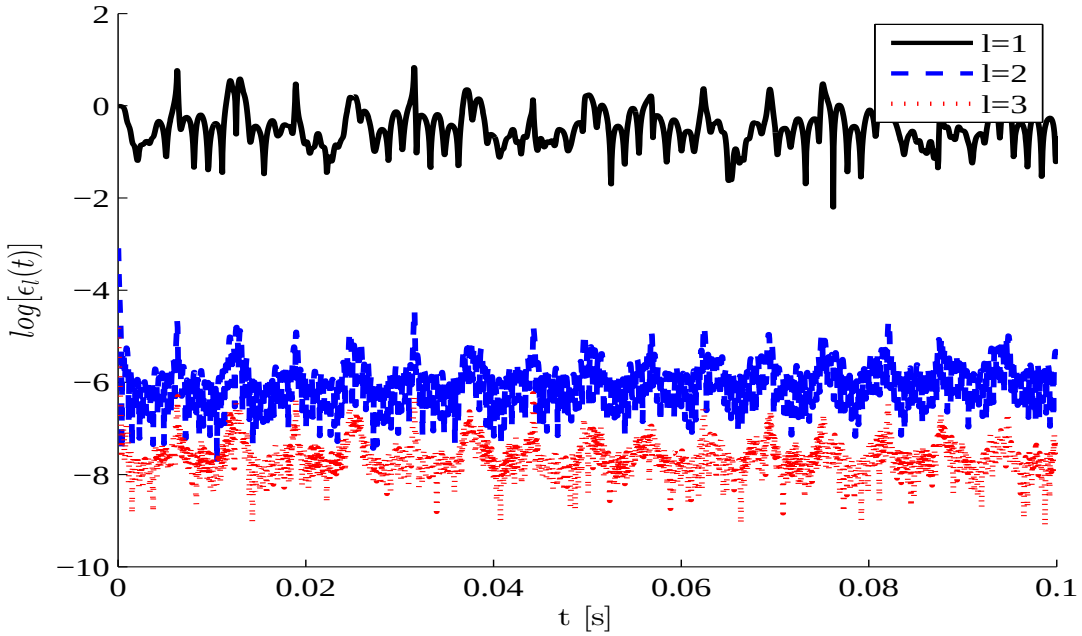
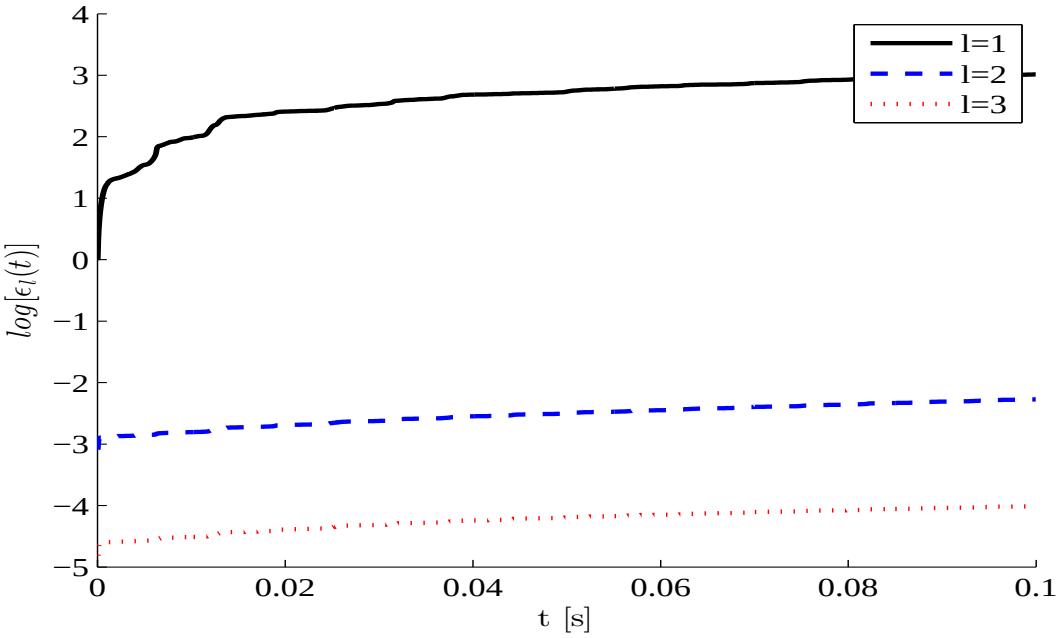


Figure 5.19: Undamaged case - Cumulative error - $\bar{\epsilon}_l(t)$



5.2.1.2 Damaged structure

Once the performances of the model order reduction method, namely the proper orthogonal decomposition, have been shown for the undamaged structure, let us focus on the damage structure defined in 5.1.
Once again, the first step is the calculation of the POMs associated to the structure subjected to the load previously defined. Similarly to what has been explained for the undamaged case, we can first of all calculate the level of information $I(l)$ retained for each reduced model (Table 5.4).

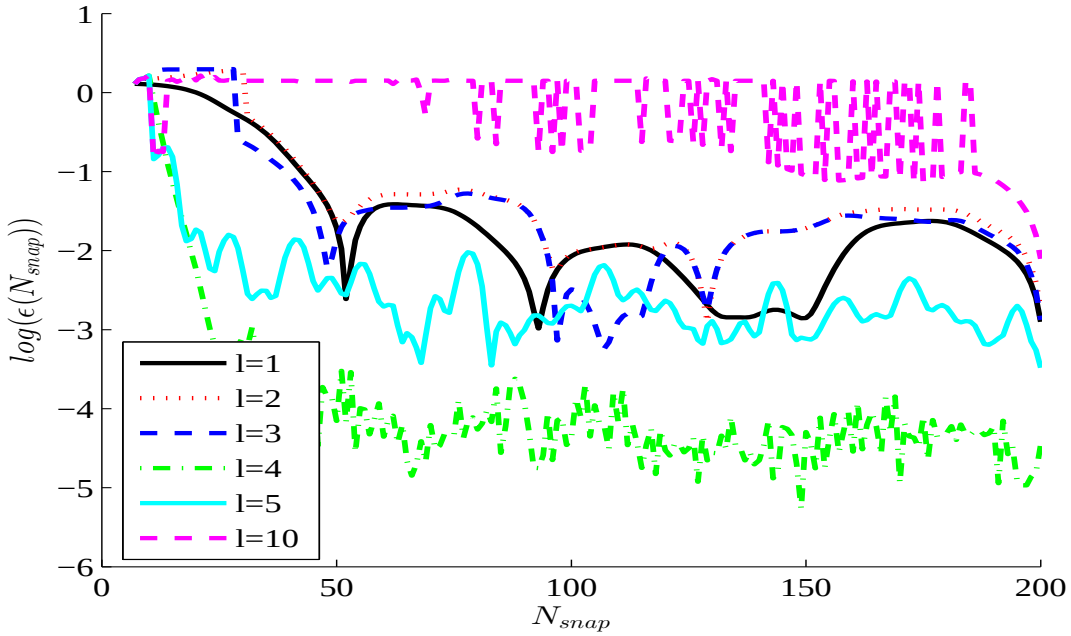
Table 5.4: Damaged case - $I(l)$

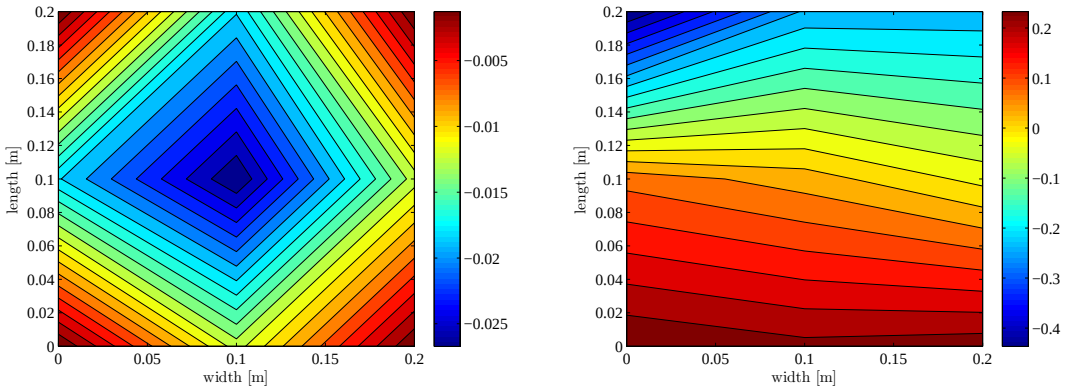
l	$I(l)$
1	0.980145
2	0.999443
3	0.999999
4	≈ 1

Comparing Tables 5.4 and 5.3, on equal number of POMs l retained, the level of information is lower; a direct consequence is that, in order to reach the same level of information, a larger number of POMs should be used. An explanation of this behavior has to be searched in the characteristics of the structure: un-

like the undamaged case, the damaged structure is geometrically non-symmetric; therefore a higher number of POM shapes are required to reproduce exhaustively the dynamics of the full model. We will see later on that using 3 POMs a perfect match of the full model is obtained. The level of information $I(l)$ associated to the 4-POMs model is not reported because physically speaking it can be assumed as 1 and it goes beyond the machine epsilon.
As explained in the previous paragraph, the convergence of POMs varying the number of snapshots N_{snap} is shown in Figure 5.20. We can see that the first POMs converge very quickly, in fact, just after approximately 50 snapshots considered, the relative error is under 10%. The convergence of the higher POMs have basically the same behavior noticed for the undamaged structure: since the additional amount of information brought is negligible, these POMs do not converge. Clearly, according to what has been previously explained about the difference between the undamaged and damaged structure, in this case this unstable behavior can be seen only for $l \geq 10$, while in the undamaged case for $l \geq 4$.
Figures from 5.21 to 5.22 shows the displacement and rotation maps of the first two POMs. Unlike the undamaged case, since the structure is not symmetric, also the maps shown are non symmetric.

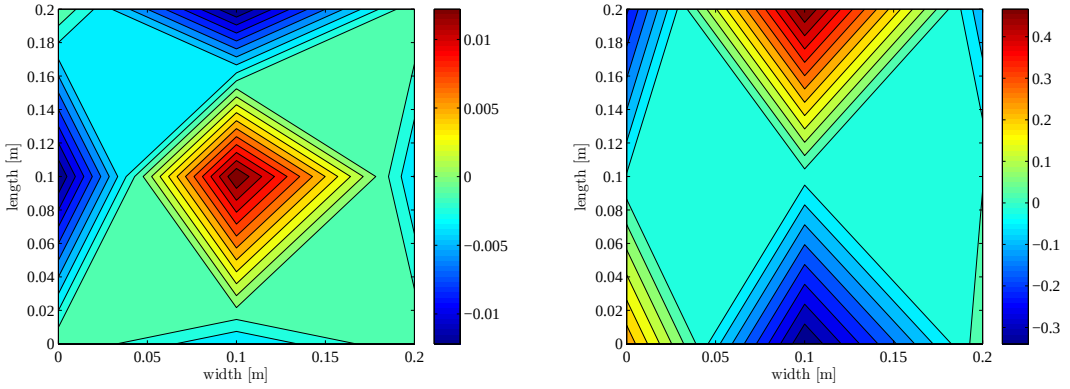
Figure 5.20: Damaged case - POM convergence $\epsilon_l(N_{snap}) = \frac{||\phi_l(N_{snap}) - \phi_l(200)||_{L^2}}{||\phi_l(200)||_{L^2}}$





(a) u_z (b) φ_x

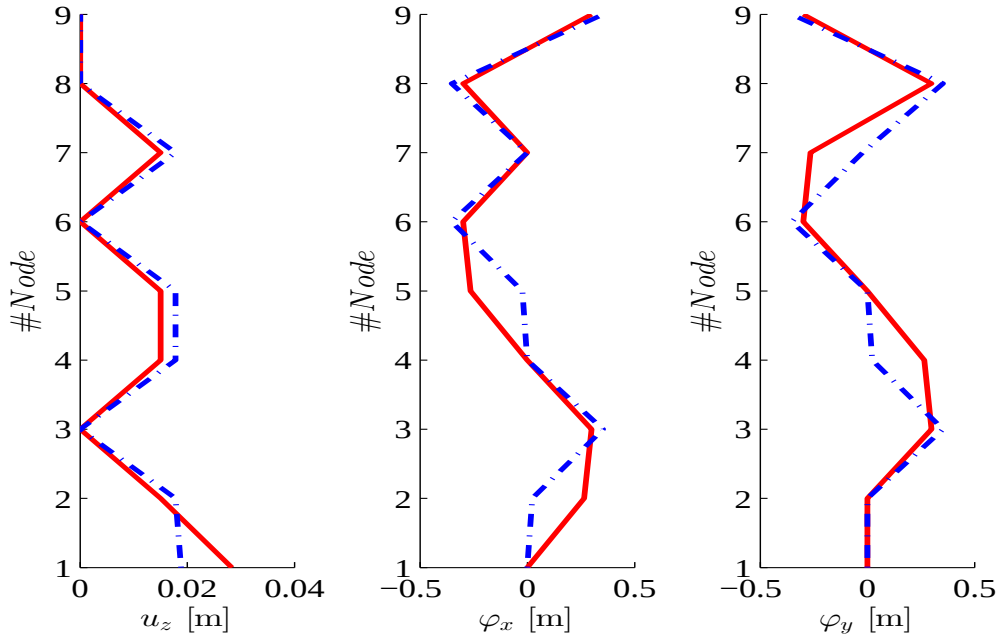
Figure 5.21: Damaged case - POM 1 components maps



(a) u_z (b) φ_x

Figure 5.22: Damaged case - POM 2 components maps

Figure 5.23: Damaged case - Comparison between 1^{st} (red solid line) POM and 1^{st} (blue dashdot line) natural mode



The last remark regards the comparison between the POMs and the natural modes of the structure: as shown in figure 5.23 and likewise with what has been shown for the undamaged case, no perfect match exists between them two.
Let us now compare the time histories obtained through the full model and the reduced models, still considering node 2. Having previously discussed about the convergence of the POMs, in all the simulation we will use $N_{snap} = 200$.
Figures from 5.24 to 5.33 show the evolution of node 2 comparing the results obtained with the full model and the reduced models with 1, 2 or 3 POMs retained. Considering figures from 5.24 to 5.30, some remarks can be emphasized. First of all, as previously shown in table 5.3, in all figures the 3-POM order reduction can be considered as a very good approximation of the full model, in fact both the displacements, the rotations and their derivatives has been catched very well.
Regarding the 1-POM order reduction, it can be stated that only the displacements and the rotations can fit the real evolution in an acceptable way; nevertheless only the lower frequency of the signal is reproduced correctly, while the fluctuations associated to the other ones does not suit the real one. As shown in the undamaged case and displayed in figures 5.31, 5.32 and 5.33, the l-order reduced model can reproduce correctly only the first l frequencies of the signal produced by the full model.

Figure 5.24: Damaged case - Point 2 - Displacement $u_z(t)$

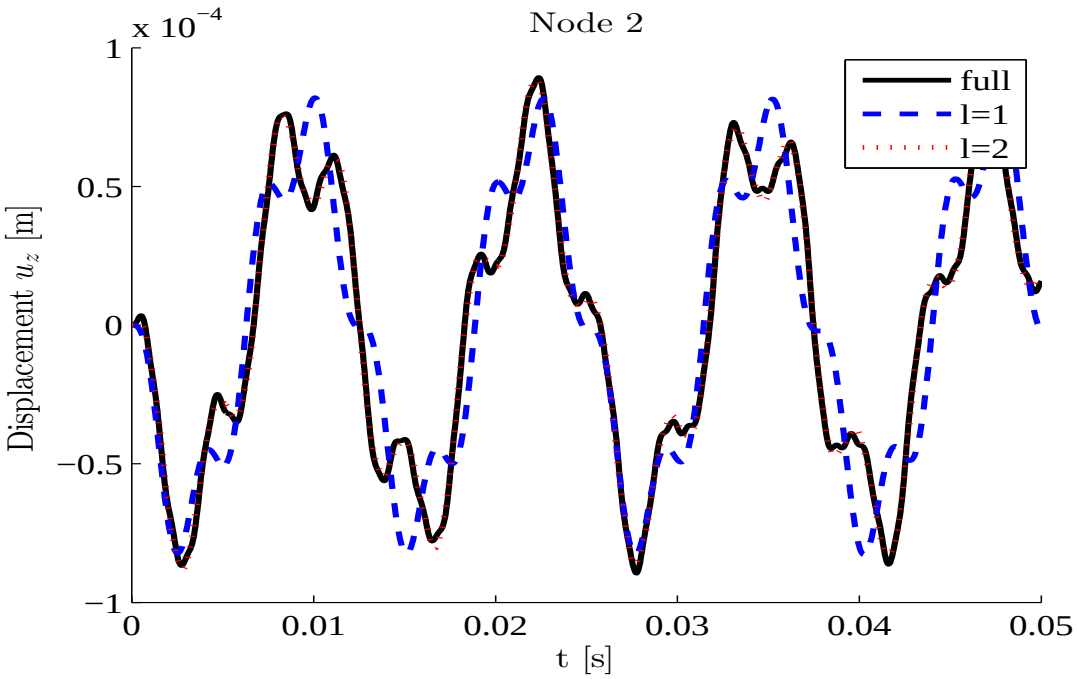


Figure 5.25: Damaged case - Point 2 - Displacement $u_z(t)$ - Detail of figure 5.24

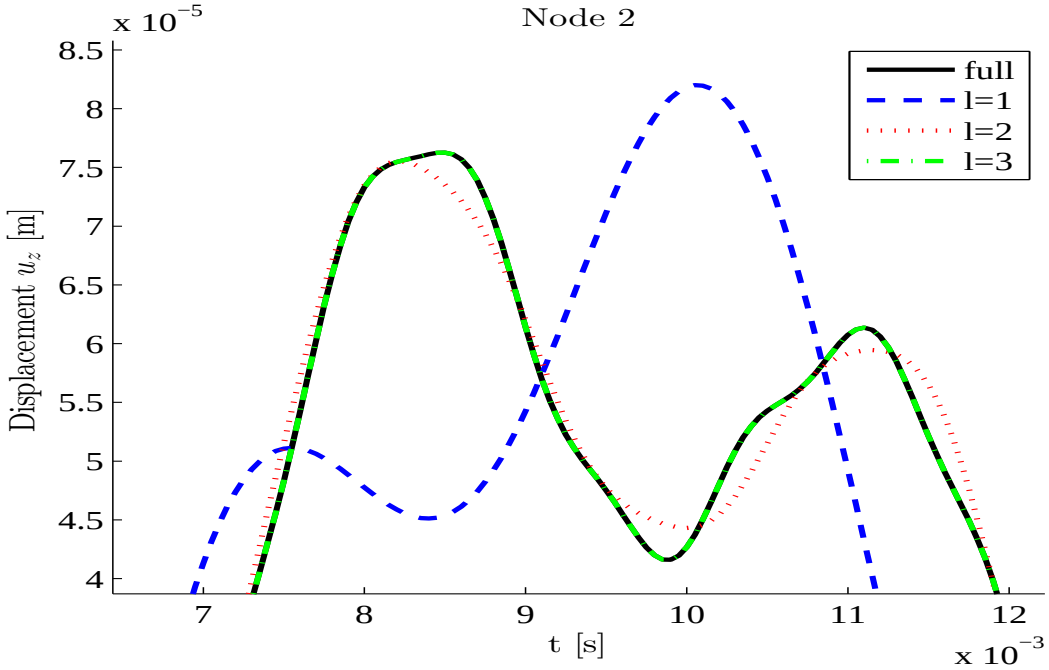


Figure 5.26: Damaged case - Point 2 - Rotation $\varphi_x(t)$

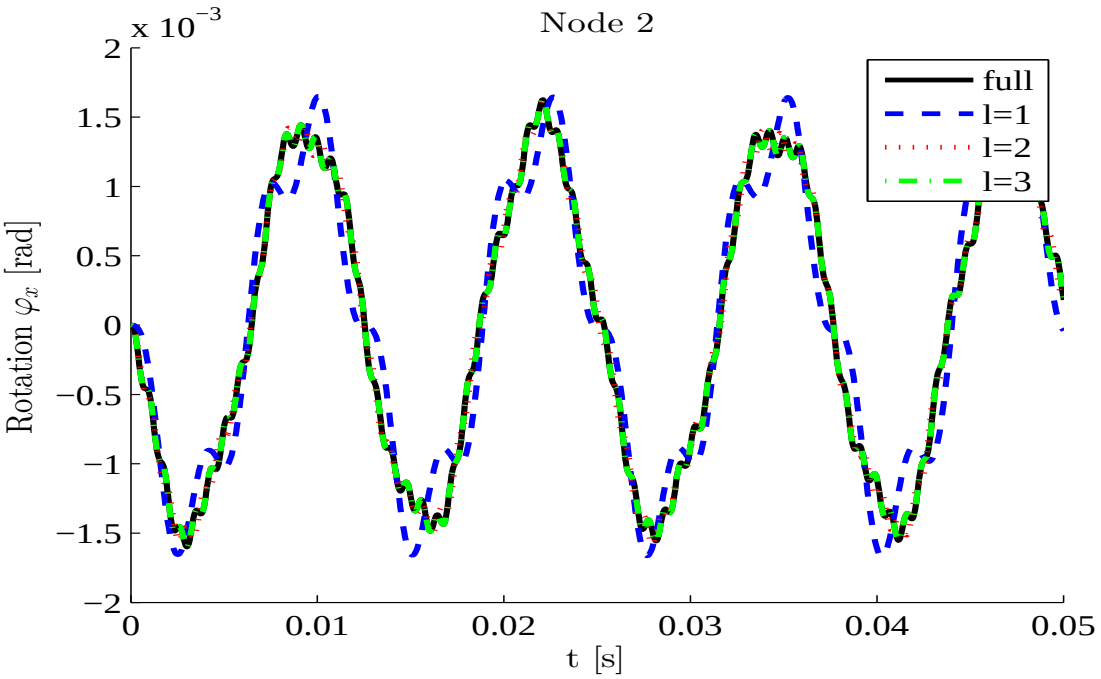


Figure 5.27: Damaged case - Point 2 - $\dot{u}_z(t)$

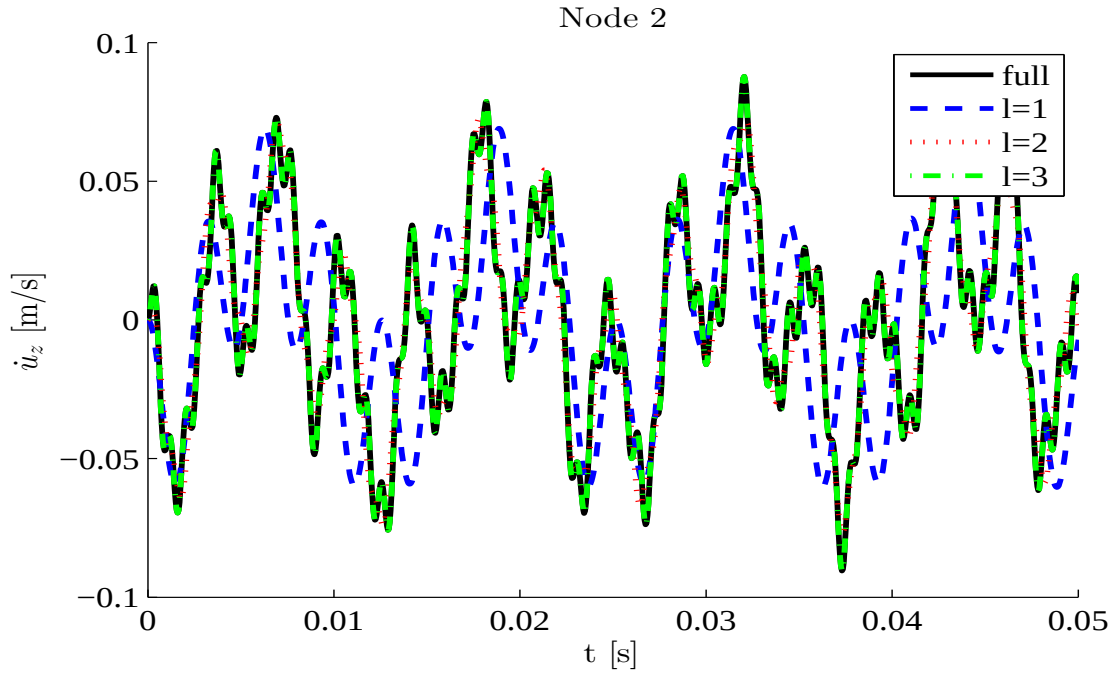


Figure 5.28: Damaged case - Point 2 - $\dot{\varphi}_x(t)$

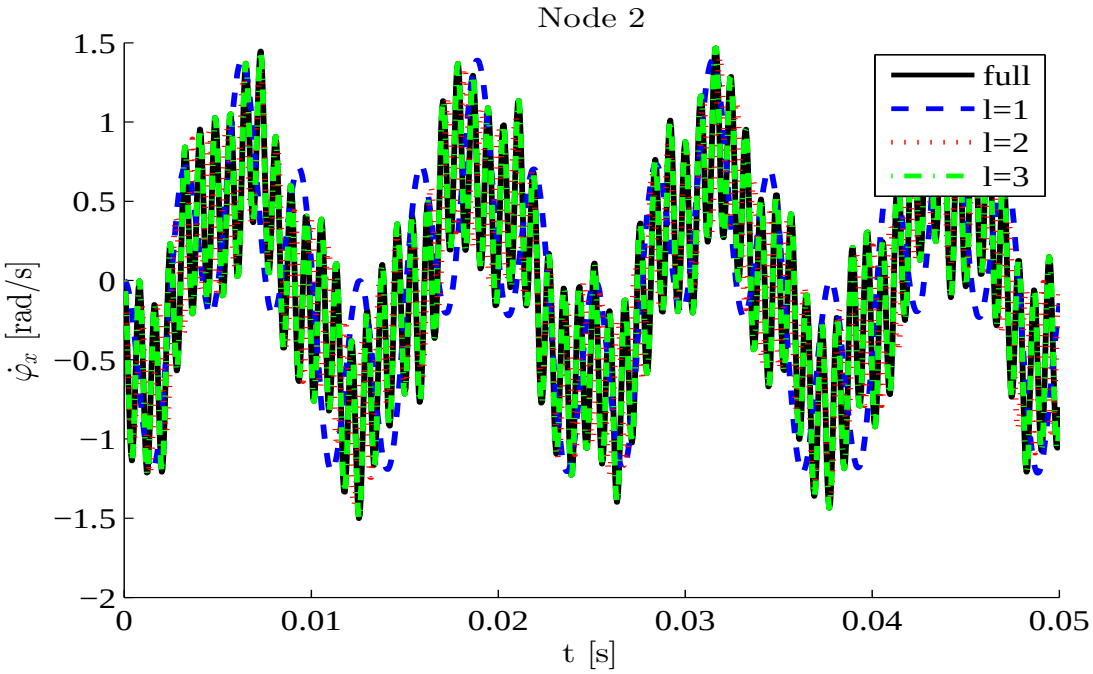


Figure 5.29: Damaged case - Point 2 - $\ddot{u}_z(t)$

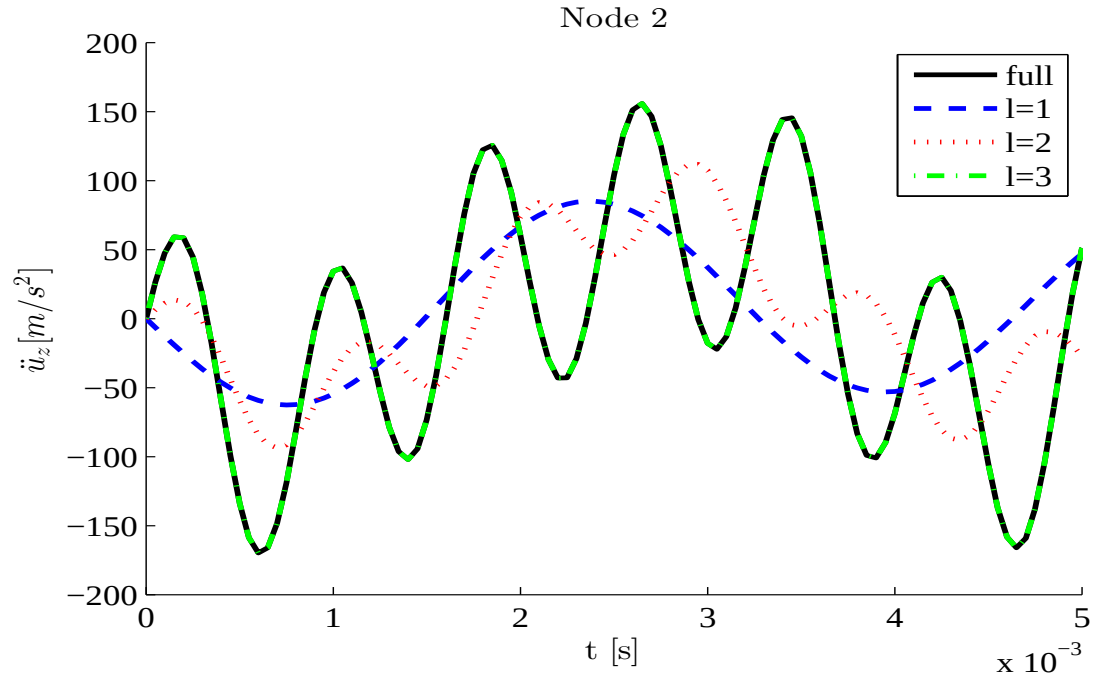


Figure 5.30: Damaged case - Point 2 - $\ddot{\varphi}_x(t)$

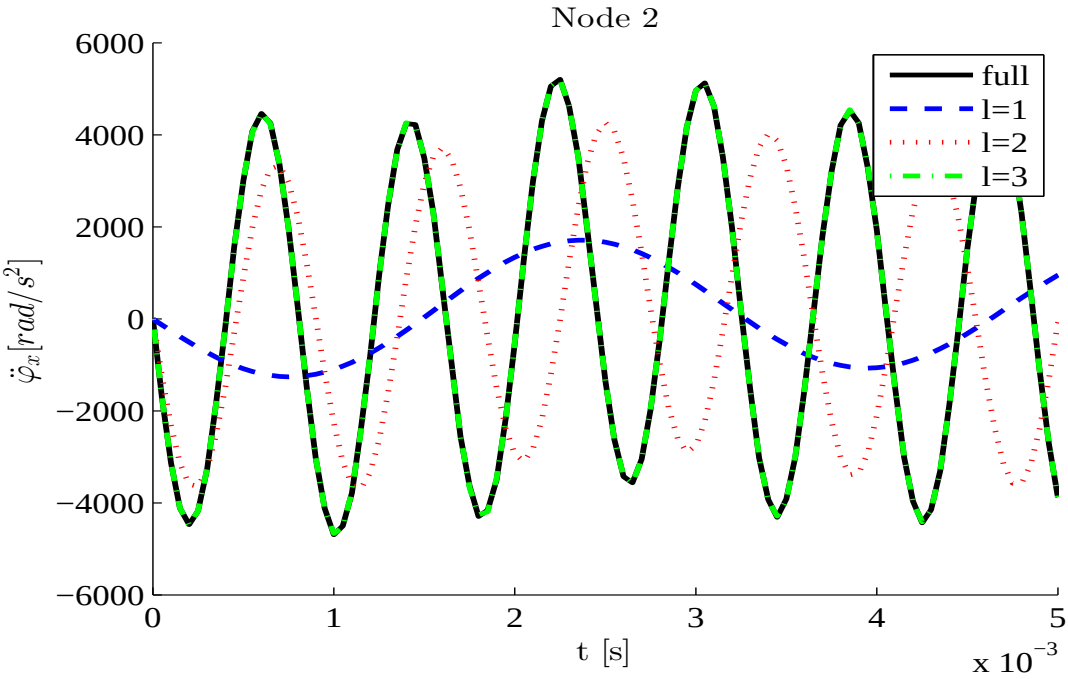


Figure 5.31: Damaged case - Point 2 - $\mathcal{F}[u_z(t)]$

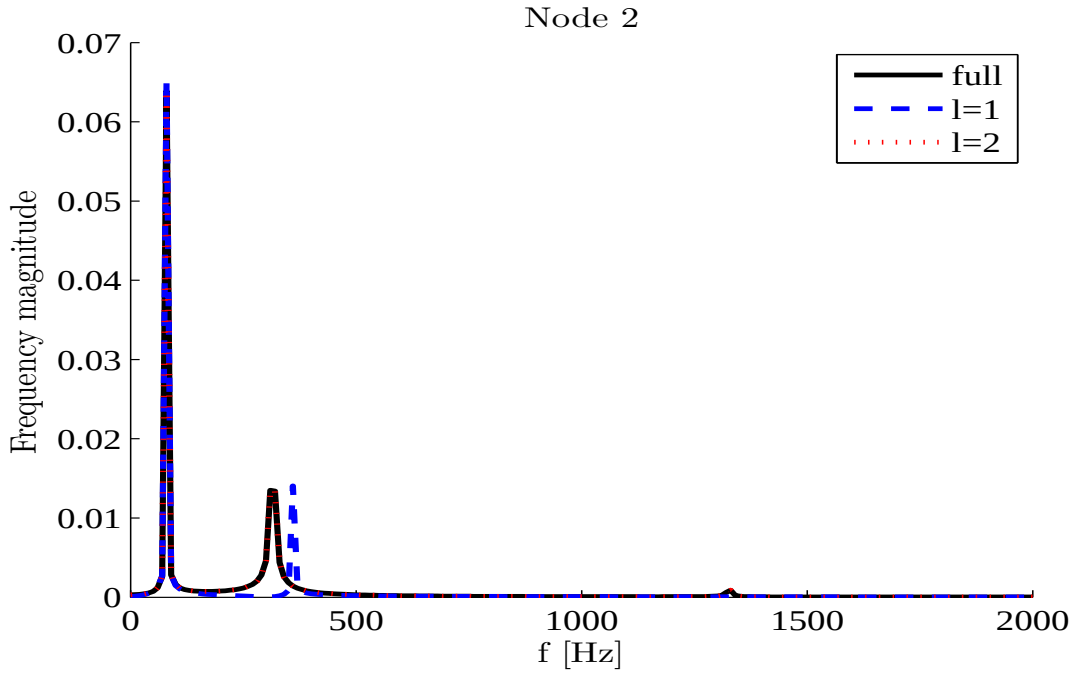


Figure 5.32: Damaged case - Point 2 - $\mathcal{F}[\dot{u}_z(t)]$

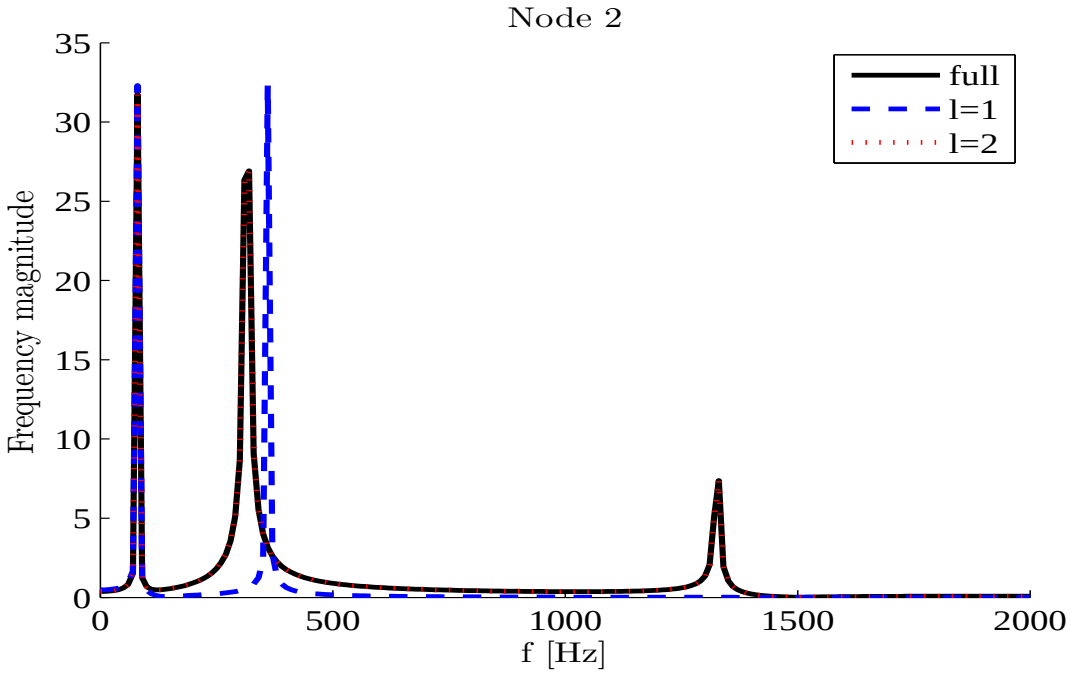
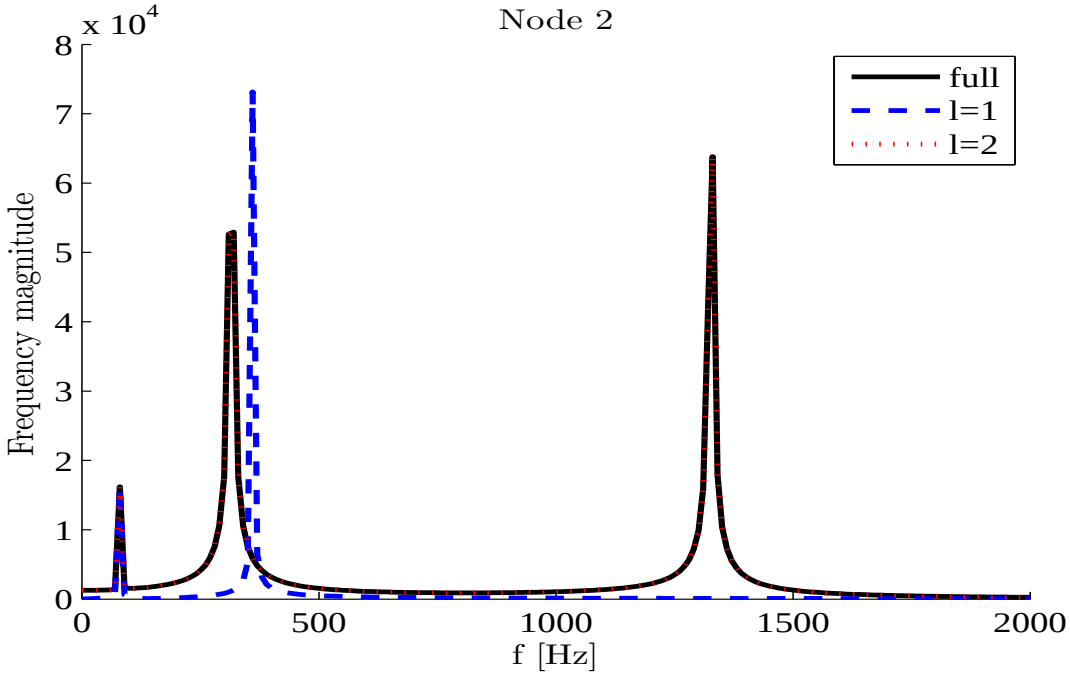


Figure 5.33: Damaged case - Point 2 - $\mathcal{F}[\ddot{u}_z(t)]$



Let us now consider in figure 5.34 the relative error between the exact evolution of the system and the reduced ones as defined in (5.6). First of all, it is possible to note that, in accordance with what has been previously explained, $\epsilon_1(t)$ and $\epsilon_2(t)$ are almost six order of magnitude higher than $\epsilon_3(t)$, while there is not a substantial difference between the latter one and the $\epsilon_4(t)$ error. An other important remark is that the relative error is almost stationary; in order to better explain this concept, let us consider the cumulative error $\bar{\epsilon}_l(t)$, defined in (5.7). In Figure 5.35 we can see that the cumulative errors $\bar{\epsilon}_1(t)$ and $\bar{\epsilon}_2(t)$ grows over time but, since the slope decreases, the accuracy of the approximation increases. On the other hand, $\bar{\epsilon}_3(t)$ and $\bar{\epsilon}_4(t)$ are basically constant, hence the accumulation of error is negligible.

Figure 5.34: Damaged case - Relative error - $\epsilon_l(t) = \frac{||\mathbf{u}_l(t)-\mathbf{u}(t)||_{L^2}}{||\mathbf{u}(t)||_{L^2}}$

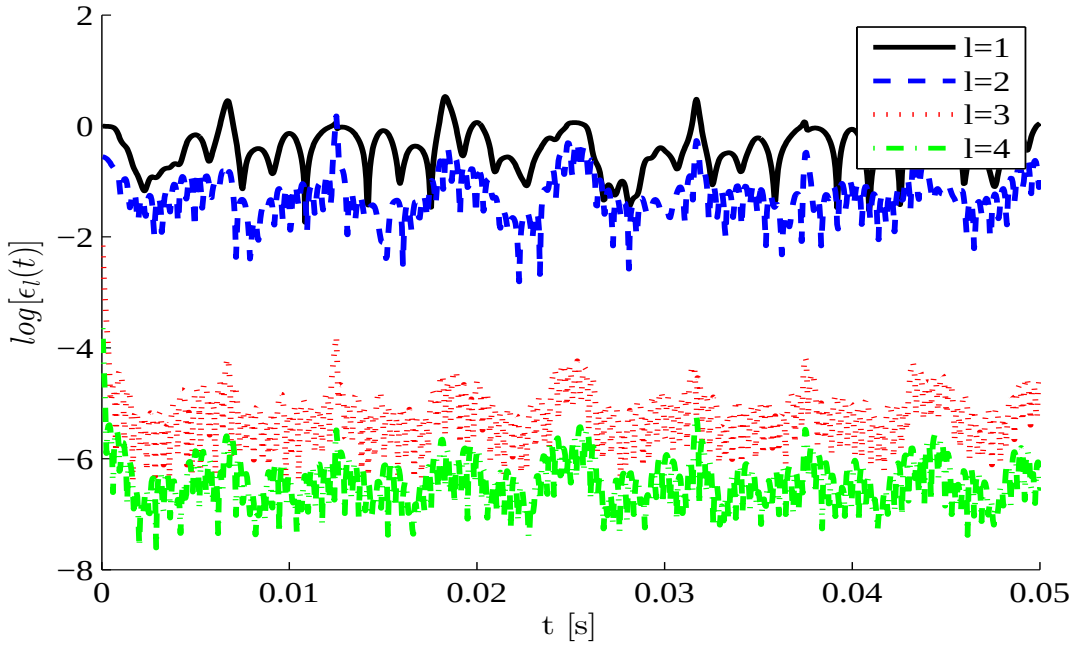
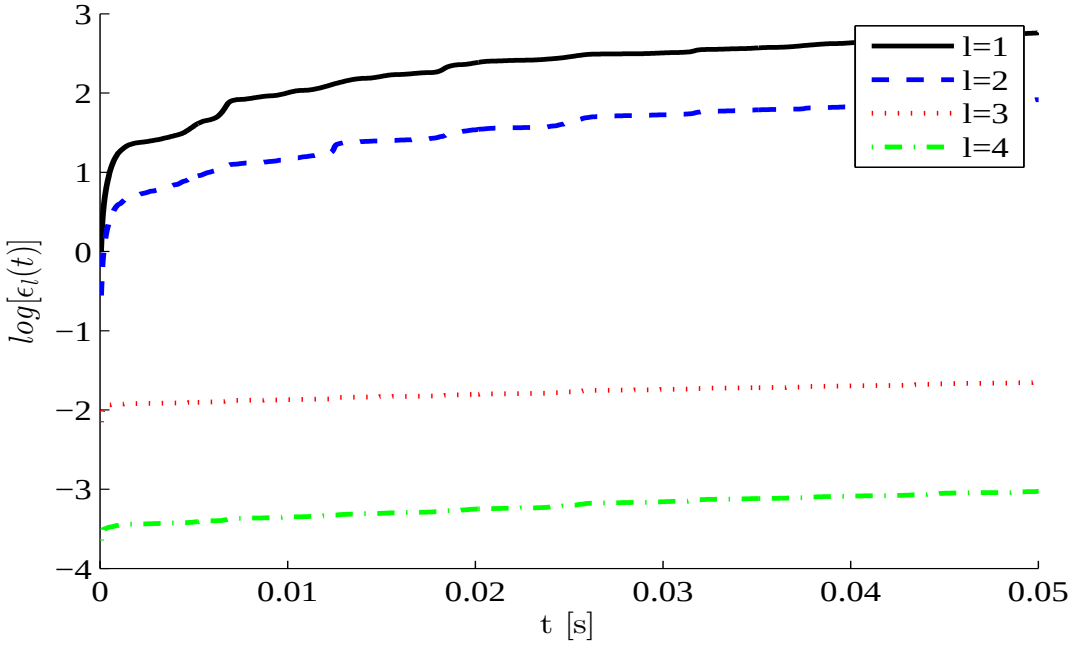


Figure 5.35: Damaged case - Cumulative error - $\bar{\epsilon}_l(t)$

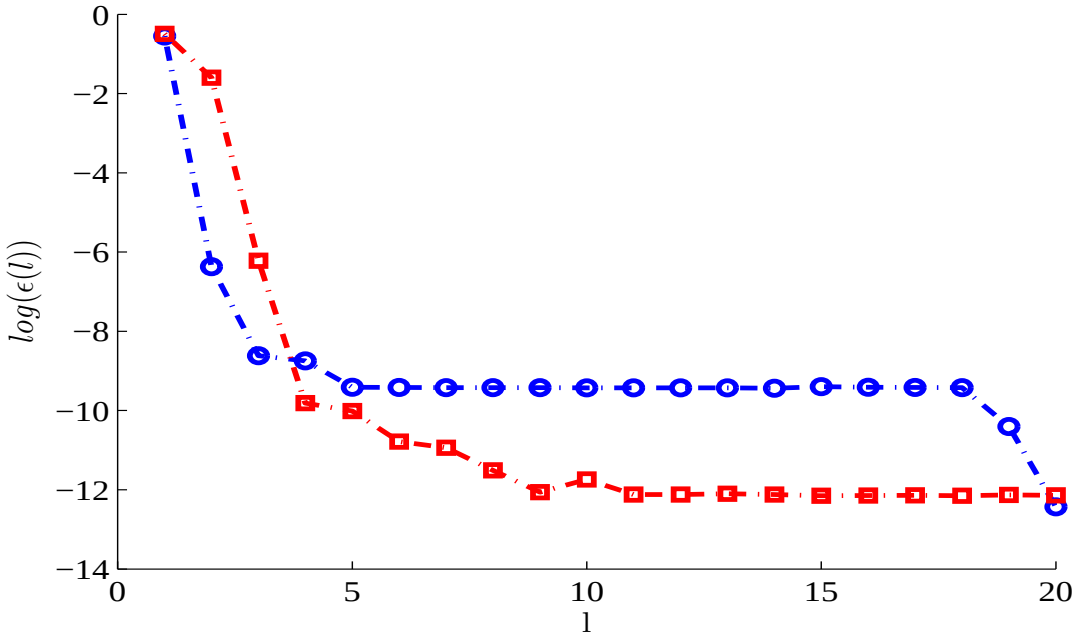


In order to sum up all the concepts previously explained and compare the model order reduction of the undamaged and damaged structure, let us consider the signal of the displacement $u_z(t)$ at point 2 and calculate the relative error between the reduced model and the full model:

$$\epsilon(l) = \frac{||\mathbf{u}_{l,z} - \mathbf{u}_z||_{L^2}}{||\mathbf{u}_z||_{L^2}} \tag{5.8}$$

Figure 5.36 shows the evolution of the error for different choices of l, for the damaged and undamaged case. As previously explained, we can see that the damaged structure requires a higher number of POMs in order to guarantee the same level of accuracy.

Figure 5.36: Damaged case - Point 2 - error u_z (red line: damaged, blue line: undamaged)



5.2.2 POD Kalman observer

In the previous Section, the performance of the model reduction method have been evaluated. It is possible to state that the POD can successfully reduce the dimension of the problem without loosing much information and hence it can be used in the identification phase.
We can now move on and evaluate the performance of the Bayesian recursive filter described in Section 4.4. The basic goal of the procedure is to estimate the damage parameters of the damaged structure described at the beginning of the present chapter. Since the structure has been divided into four regions, we want to estimate the four damage parameters d_i associated to their relevant regions. It has been previously explained that the damaged structure is characterized by having a reduced Young modulus $E/2$ in zone 2 and an original E in all the other regions. Therefore, the target parameters vector $\bar{\mathbf{d}}$ to be estimated is:

$$\bar{\mathbf{d}} = \begin{bmatrix} 0 \\ 0.5 \\ 0 \\ 0 \end{bmatrix} \tag{5.9}$$

where of course the damage in regions 1, 3, 4 is zero because no reduction of Young modulus and thus stiffness appears. We want now to see how the choice

of the order l, the levels of measurement and process noises and the number of observations can affect the estimation of both the damage parameters and the dynamic evolution of the system. In all the simulation, if not otherwise specified, the default settings are: $l = 3$, $\sigma_{\mathbf{v}} = 10^{-5}$, $\sigma_{\mathbf{w}} = 10^{-5}$, $\mathbf{d}_0 = \mathbf{0}$. In all the simulations we will suppose that the only available observations are the rotations along the x and y directions of nodes 2, 4, 5 and 7. The measurement vector is therefore:

$$\mathbf{y}(t_k) = \begin{bmatrix} \varphi_x^2(t_k) \\ \varphi_y^2(t_k) \\ \varphi_x^4(t_k) \\ \varphi_y^4(t_k) \\ \varphi_x^5(t_k) \\ \varphi_y^5(t_k) \\ \varphi_x^7(t_k) \\ \varphi_y^7(t_k) \end{bmatrix} \tag{5.10}$$

where $\mathbf{y}_k \in \mathbb{R}^{N_{obs}}$ with $N_{obs} = 8$. In a possible application, this choice prevents to deploy the sensors at the corners and the middle of the plate where the load is applied. Later on, the effect of the reduction of the number N_{obs} of degrees of freedom measured will be shown.

5.2.2.1 Number of POMs

Let us first of all compare the damage parameters estimation obtained using 1, 2 or 3 POMs. As shown in the previous Section, we can expect that the estimation obtained using one single POM will be not enough accurate and that increasing the number of POMs the quality of the estimation should grow. All the following simulations are performed using very low level of noises ($< 1\%$ of the signal amplitude) since in this phase the only aim is to compare the performance of the filter with different choices of l. Afterwards, a further explanation on the effects of the level of noises will be given.
Figures 5.37 to 5.39 show the parameter estimation, at varying number of POMs and hence dimension of the reduced system.
Considering the 1-POM reduced system, the filter is able to detect that there is a damage in the second region, in fact the estimated d_1 is higher than the other damage indexes. Nevertheless, as shown in the previous section, the accuracy is too low and therefore the estimated parameters are not acceptable compared to the target. The estimation obtained considering 2 POMs is better than the one with 1 POM, but still the damage indexes of the undamaged regions is not good, since the target values are not completely matched. If we now look at the 3-POMs system estimation, both the parameters d_i of the undamaged and damaged regions are almost perfectly estimated. It is possible to lead to the same conclusion if the calculation of the relative error between the target parameters vector and the estimated one is performed; as done in the previous section, the

relative error is defined as:

$$\epsilon(l) = \frac{||\mathbf{d}_l - \bar{\mathbf{d}}||_{L^2}}{||\bar{\mathbf{d}}||_{L^2}} \tag{5.11}$$

In figure 5.40, the decrease of the error associated to the increase in the umber l of POMs retained is shown. As explained in Section 5.2.1, the choice of l depends on the type of structure, and in this case $l = 3$ because we are considering the damaged case.
Regarding the speedup of the damage identification procedure due to the adoption of a reduced order model, in Table 5.5 the calculation time of the algorithm described in Table 4.1 is reported.

Table 5.5: Computational time of the damage identification algorithm

number of POMs retained	1	2	3	4	5
computational time [s]	11.13	23.88	36.16	55.28	61.96

The calculations have been performed using a personal computer Intel Core i7-3632QM, with 4 Gb of RAM, running Windows 7x64 as operating system. The algorithm has been implemented the MATLAB R2013a computing environment. It is straightforward to see that the reduction of the order guarantees an important decrease in the computational cost. Moreover, these results are even more remarkable if we consider that the full order model would have 54 degrees of freedom, leading to an unbearable time. The part of the overall procedure that require the most computational effort is the calculation of the Jacobian matrix described in equation (4.42).

Table 5.6: Jacobian matrix calculation time

number of POMs retained	1	2	3	4	5
computational time [s]	5.05	13.14	71.51	510.02	8163.15

While the computational time of the recursive filter increases approximately linearly with respect to the number of POMs retained, the time required to obtain the Jacobian matrix increases geometrically (Table 5.6). This behavior is basically due to the poor performances of the MATLAB programming language in treating the symbolic calculations. Nevertheless, since the Jacobian matrix is a preliminary calculation and it is always performed off-line, it does not affect the overall computational time of the damage identification algorithm.

Figure 5.37: Parameters d_i estimation - 1 POM

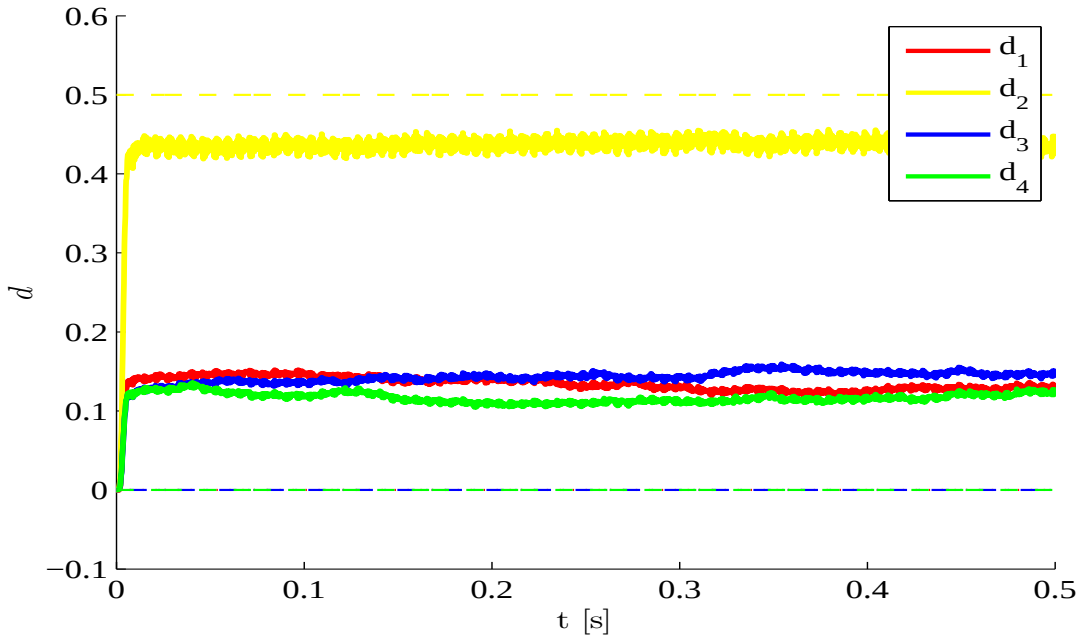


Figure 5.38: Parameters d_i estimation - 2 POMs

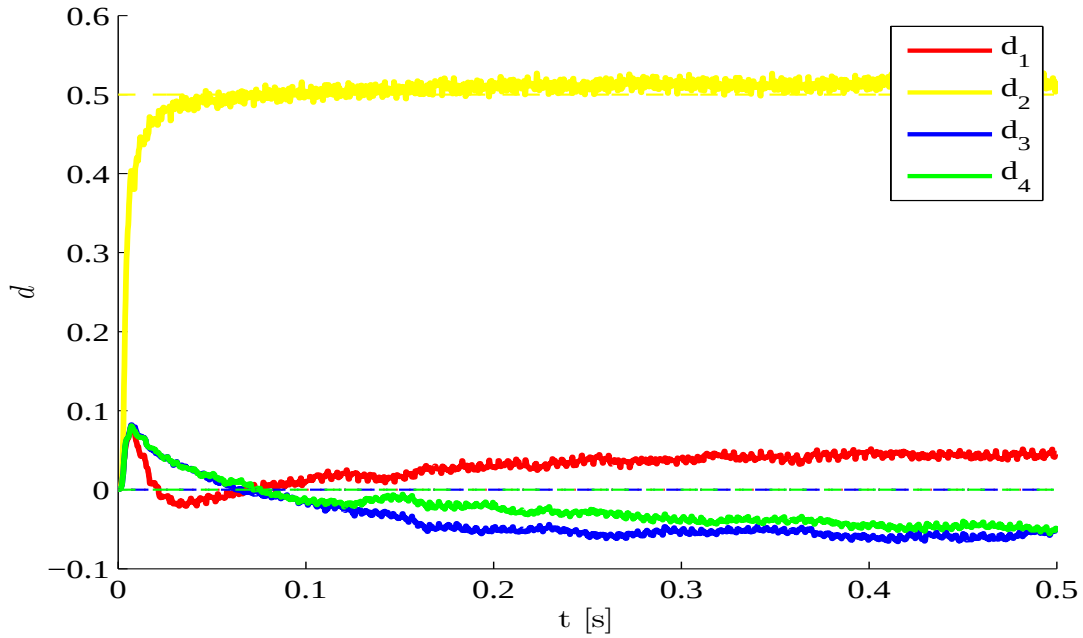


Figure 5.39: Parameters d_i estimation - 3 POMs

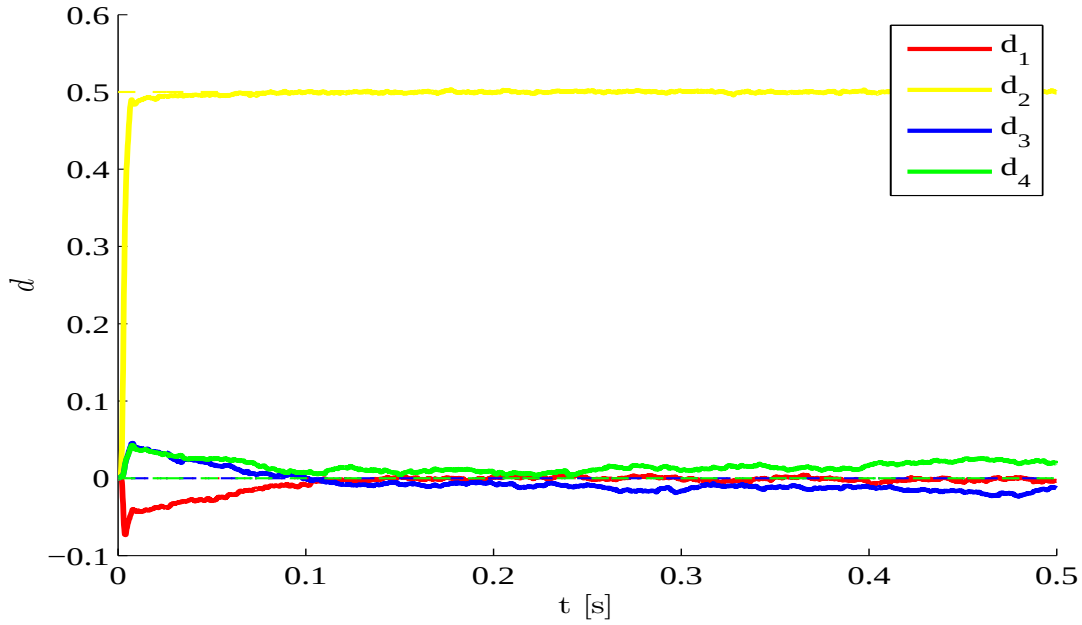
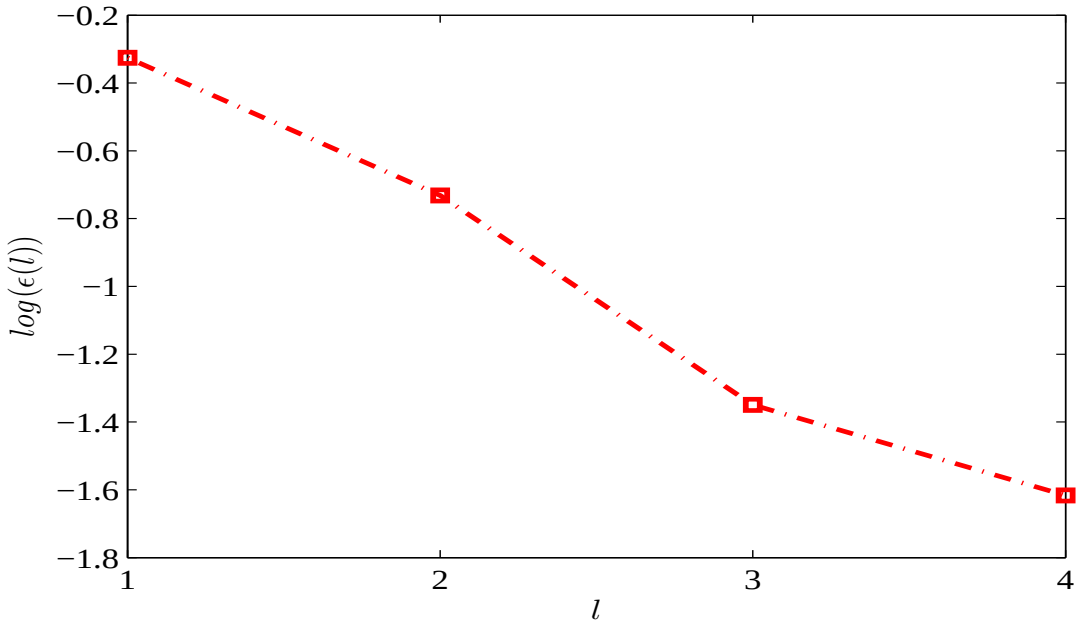


Figure 5.40: Relative error - $\epsilon(l) = \frac{||\mathbf{d}_l - \bar{\mathbf{d}}||_{L^2}}{||\bar{\mathbf{d}}||_{L^2}}$



Since the 3-POMs reduced order model gives a good approximation and estimation, from now on 3 POMs will be retained.

5.2.2.2 Initial conditions

Let us now consider how the filter reacts to a variation of the initial guess $\mathbf{d}_0$ of the parameters. The estimation problem can be seen either as a damage identification or as a system identification, in fact the damage index d_i can always be related to the Young modulus E_i, i.e. to the stiffness of that specific region.

$$E_i = (1 - d_i)E \tag{5.12}$$

The latter equation can be generalized to a generic local stiffness parameter:

$$k_i = (1 - d_i)k \tag{5.13}$$

For instance, considering a plate and remembering that the bending stiffness is $D = \frac{Eh^3}{12(1-\nu^2)}$ (where h is the thickness and ν is the Poisson's ratio), the equation (5.13) can be specialized for that specific case; using equation (5.12), we have:

$$\frac{12(1-\nu^2)}{h^3} D_i = (1 - d_i)\frac{12(1-\nu^2)}{h^3} D \tag{5.14}$$

$$D_i = (1 - d_i)D \tag{5.15}$$

Considering the damage identification, from a physical point of view the parameter d should always lies in $[0, 1)$ because a generic damage can only produce a reduction of the stiffness. Imagine now that the aim is the estimation of the stiffness of a structure: in this case the undamaged structure is no more considered as the initial situation, but it is only a reference system needed to make the method work properly. The direct consequence is that the initial conditions of the system are unknown and therefore it is important to see if it is still possible to identify the parameters whenever the initial guess is not given. Figures 5.41 and 5.42 correspond respectively to the cases with initial conditions equal to $E_i(t_0) = \frac{3}{2}E$ and $E_i(t_0) = 0$: the filter is still able to estimate the parameters. The variation of the initial values of the parameters does not affect the accuracy of the dynamic evolution estimation.
If the damage parameters of the initial guess are too far from the target damage parameters, the estimation can lead to instability and divergence: in Figure 5.43 the case of $E_i(t_0) = 2E$ is shown.

Figure 5.41: Parameters d_i estimation - $E_i(t_0) = \frac{3}{2}E$

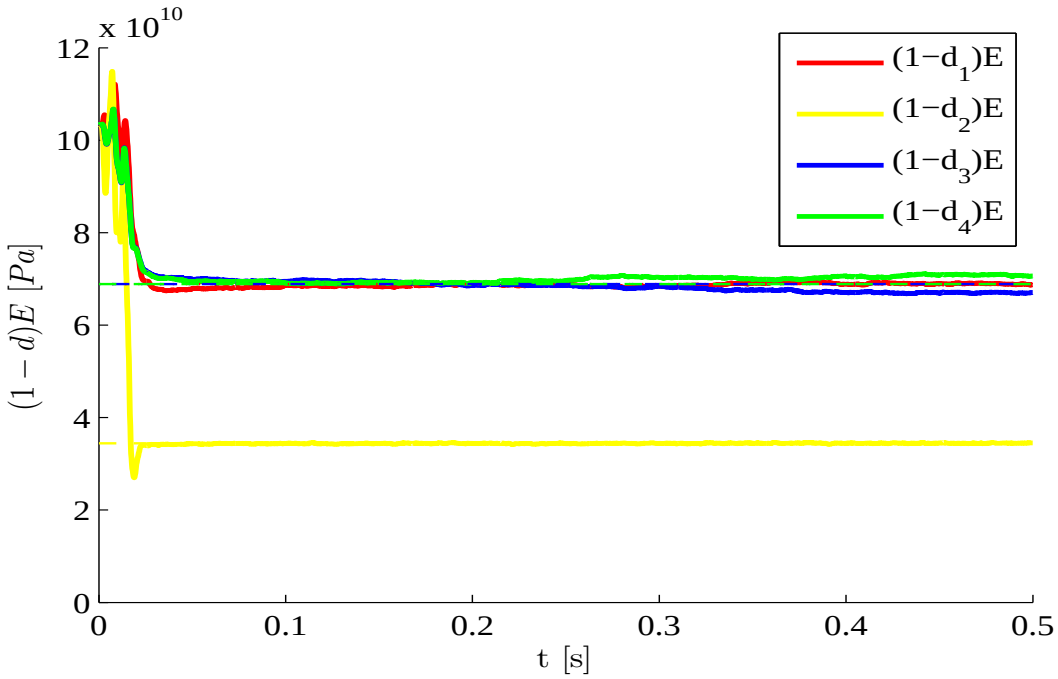


Figure 5.42: Parameters d_i estimation - $E_i(t_0) = 0$

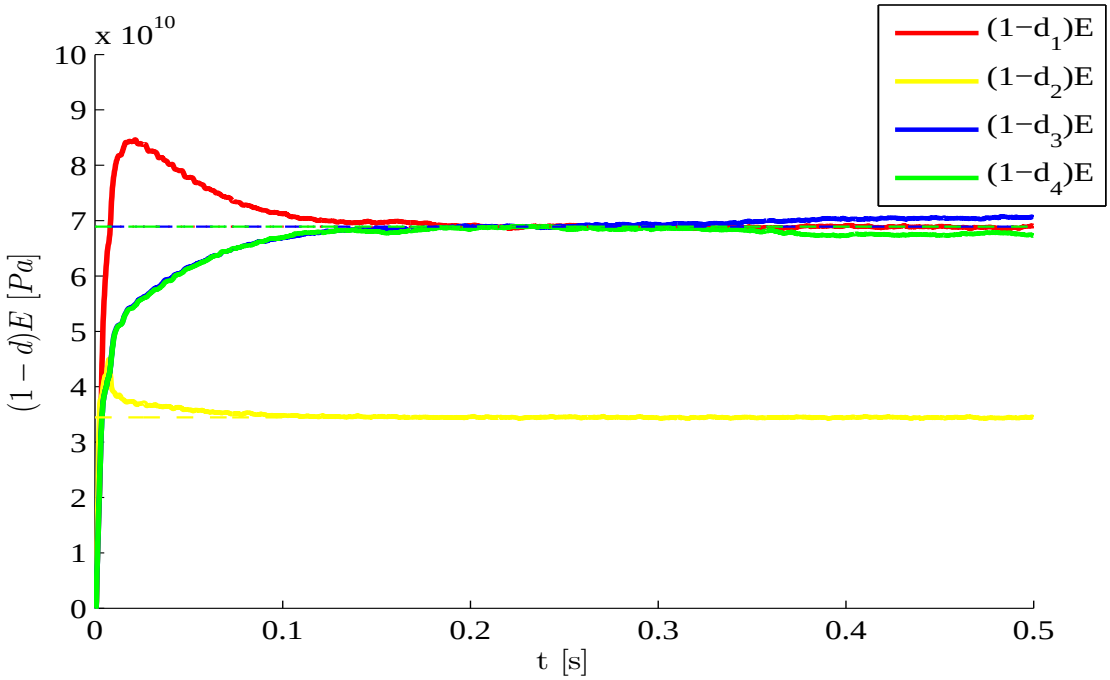
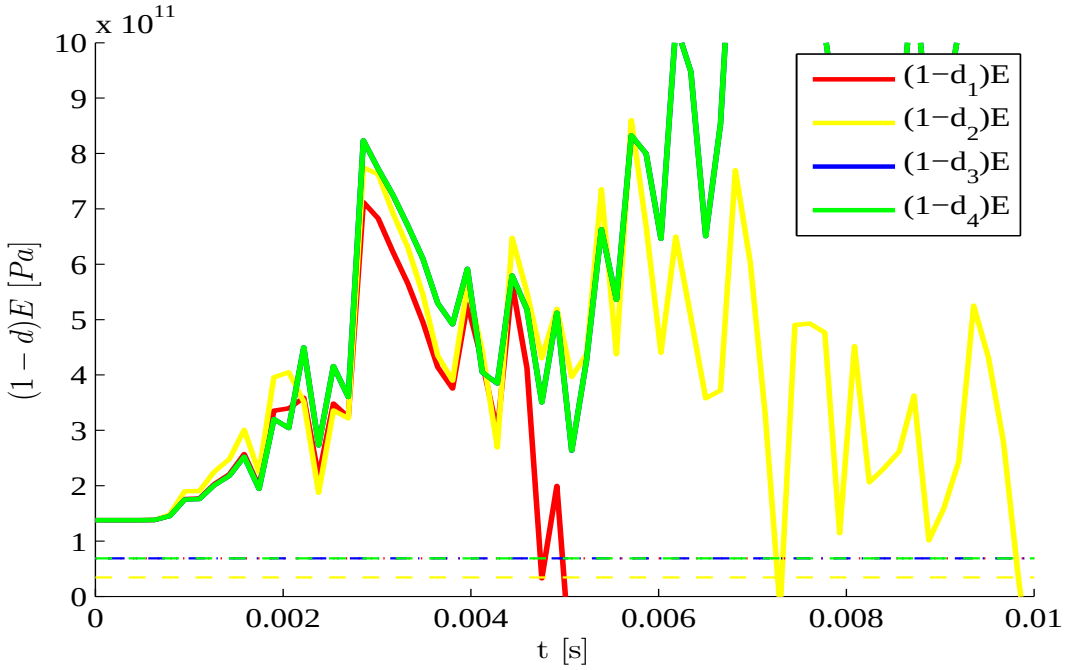


Figure 5.43: Parameters d_i estimation - $E_i(t_0) = 2E$



5.2.2.3 Measurement noise

Let us now consider how the dual estimation is affected by the measurement noise $\mathbf{v}$ associated to the observations. As explained in Section 3.2, $\mathbf{v}$ is generated as a white noise with standard deviation $\sigma_{\mathbf{v}}$ and covariance matrix:

$$\mathbf{V} = \sigma_{\mathbf{v}}^2 \mathbf{I} \tag{5.16}$$

As explained in Section 5.2, since we are assuming that only the rotational degrees of freedom are measured, it is required only one single value of $\sigma_{\mathbf{v}}$. On the other hand, assuming that the measurements come from mixed physical values, such as both rotations and displacements, different level of noises should be specified for each type of degree of freedom. It is important to remember that in a Gaussian distribution the probability in the interval $[\mu - 3\sigma, \mu + 3\sigma]$ is 0.9973 (μ is the mean); therefore, since we are considering a zero-mean white Gaussian noise, almost $\approx 100\%$ of the samples generated assuming a certain $\sigma_{\mathbf{v}}$ lie in the interval $[-3\sigma_{\mathbf{v}}, +3\sigma_{\mathbf{v}}]$. All the estimations shown until now have been performed using a very low level of noise in order to compare the results independently from the effect of the measurement noise ($\sigma_{\mathbf{v}} = 10^{-5}$). In order to have an intuitive idea of the magnitude of the measurement noise, it is possible to compare its standard deviation with the amplitude of the signal. For instance, recalling Figure 5.10, we can see that the maximum rotation is approximately $1.5 \cdot 10^{-3}$ rad and thus a standard deviation of $\sigma_{\mathbf{v}} = 10^{-5}$ is lower than 1% of its amplitude.

Let us now compare the estimated and real dynamic evolution at point 3: the choice of this point is particularly meaningful because neither the load is applied there nor the observations are collected at it, therefore the error is theoretically maximized. Considering Figures from 5.44 to 5.46 , it is possible to see that the estimation is accurate: nevertheless, at the beginning of the estimation, the estimated damage parameters are far from the target ones, hence also the estimated dynamic evolution is affected by this discrepancy. This behavior can be explained considering the fact that at a given instant t_k the estimated dynamics is calculated using the reduced stiffness matrix $\mathbf{K}_{r,k} = \mathbf{K}_{r,k}(d_1, ..., d_{N_p})$ calculated at time t_k, that depends on the estimated damage parameters at the same instant t_k (see equation 4.27). Therefore, until the the parameters have reached the target values, it looks like the system is stiffer than the real one.

Figure 5.44: Point 3 - $\varphi_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-5}$

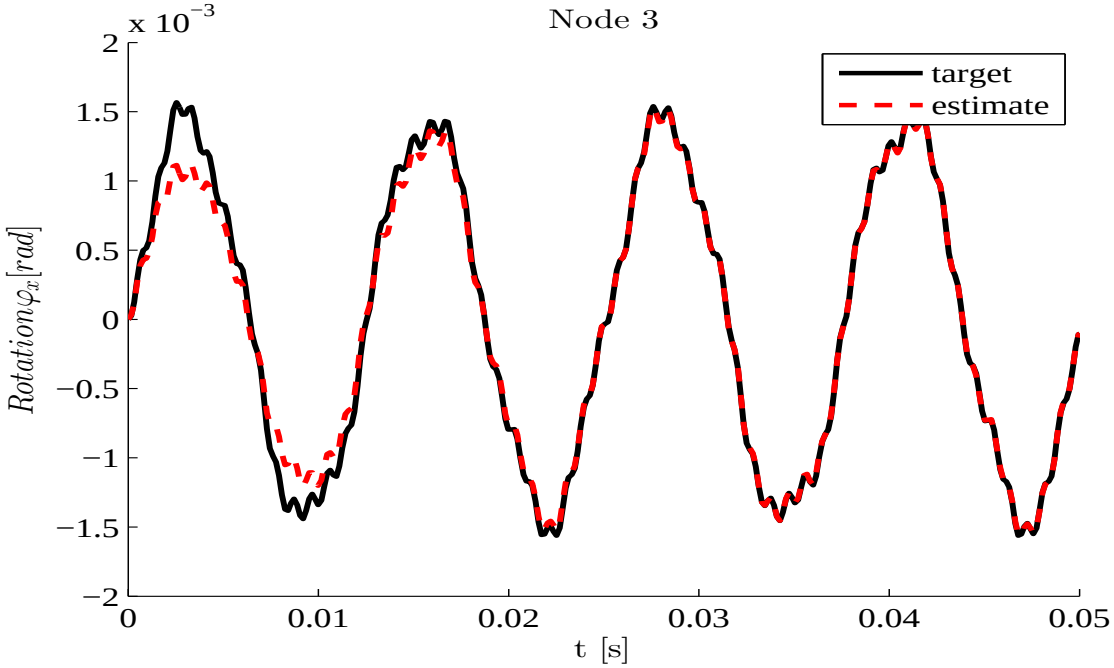


Figure 5.45: Point 3 - $\dot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-5}$

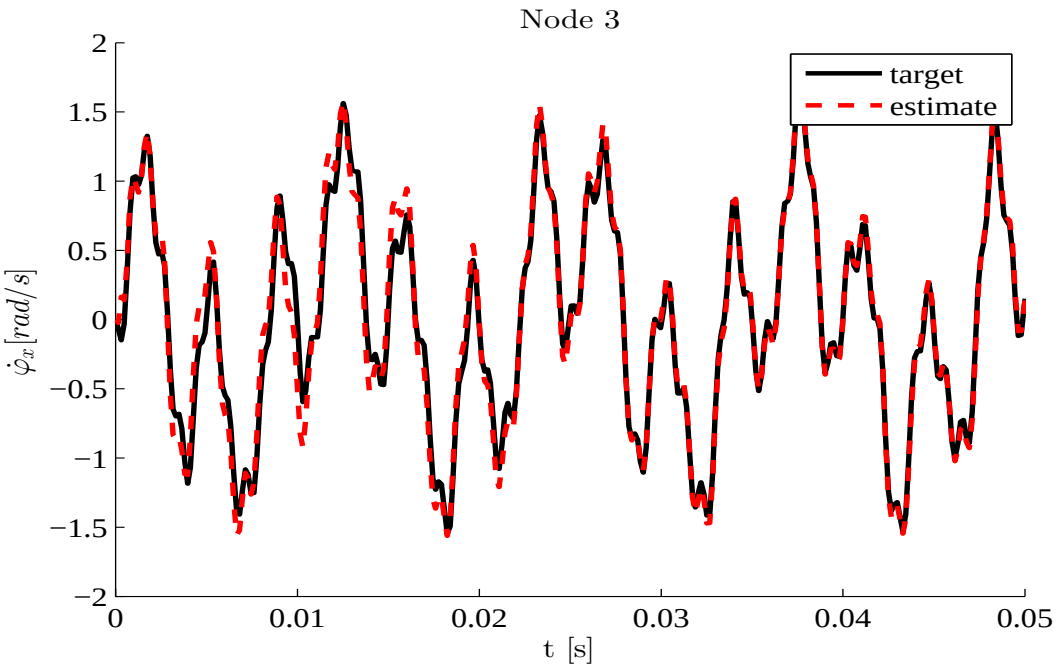
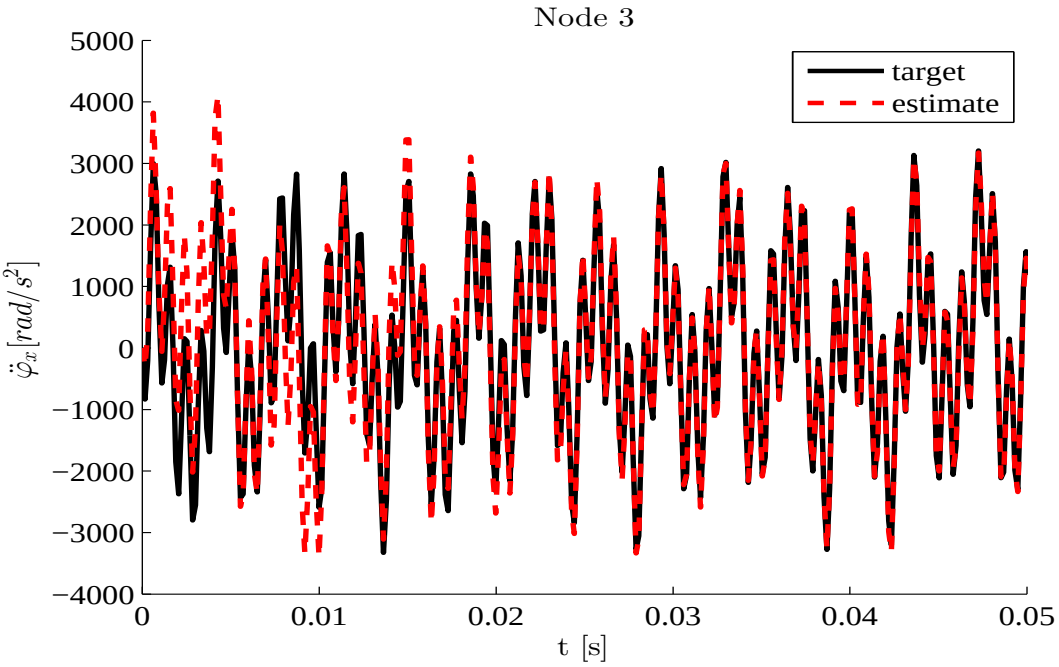


Figure 5.46: Point 3 - $\ddot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-5}$



Assume now we consider a measurement noise approximately equal to 10% of the amplitude, i.e. $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-4}$. Comparing Figures 5.40 and 5.47, it is possible to point out that the accuracy of the estimation is not affected so much by the measurement noise, but the filter requires a larger number of observation steps t_k in order for the parameters to reach the target. Analyzing Figures from 5.48 to 5.50, we can see that the quality of the estimation drops dramatically when the order of the time derivative increases; despite the fact the amplitude of the signal is not estimated correctly, the first frequency is catched. Similarly to what has been previously explained, the estimated evolution requires a certain time in order to converge to the target time history.

Figure 5.47: Parameters d_i estimation - $\sigma_{\mathbf{v}} = 1,5 \cdot 10^{-4}$

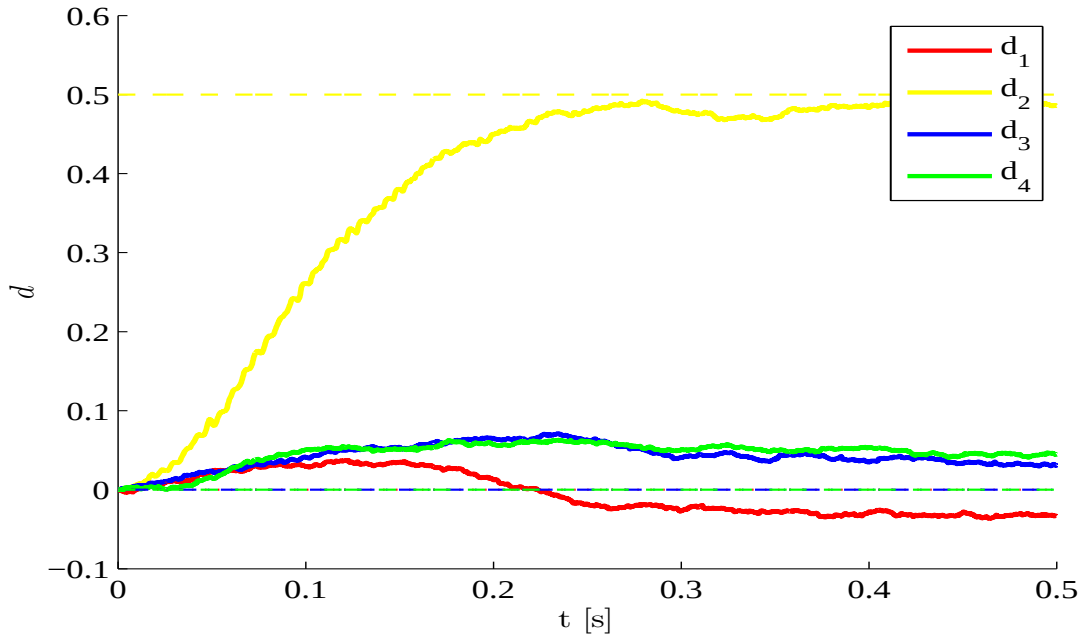


Figure 5.48: Point 3 - $\varphi_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-4}$

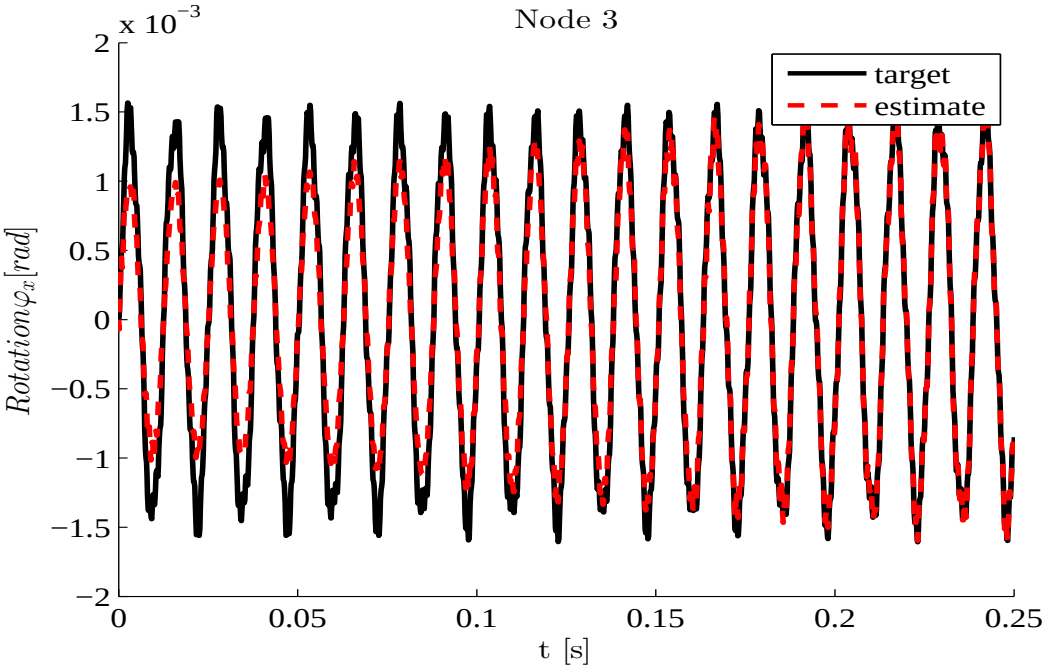


Figure 5.49: Point 3 - $\dot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-4}$

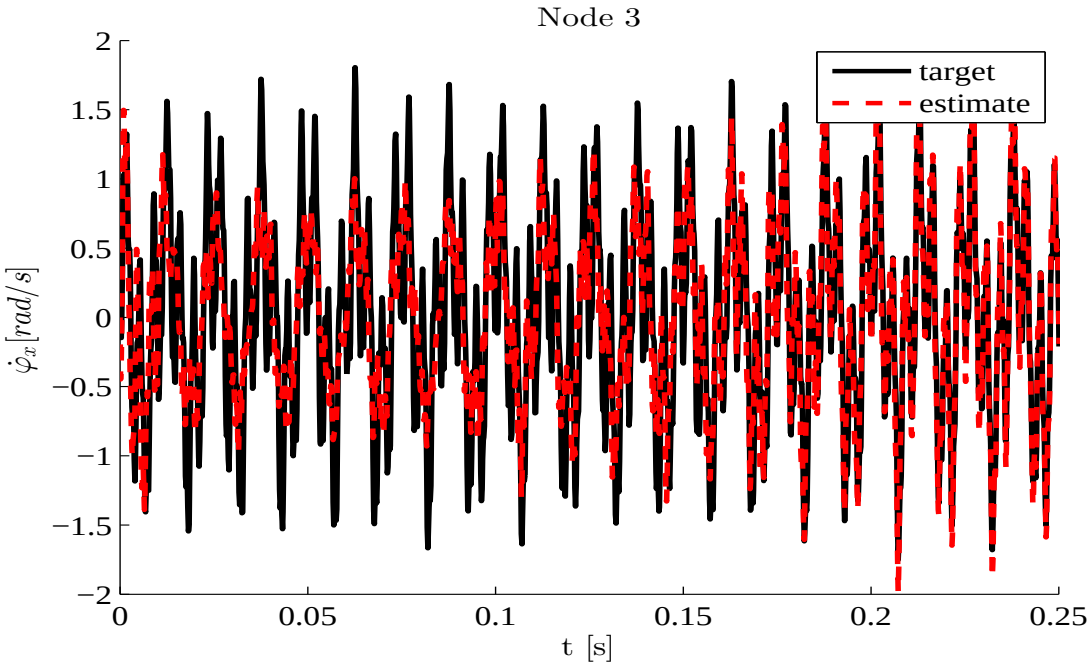


Figure 5.50: Point 3 - $\ddot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-4}$

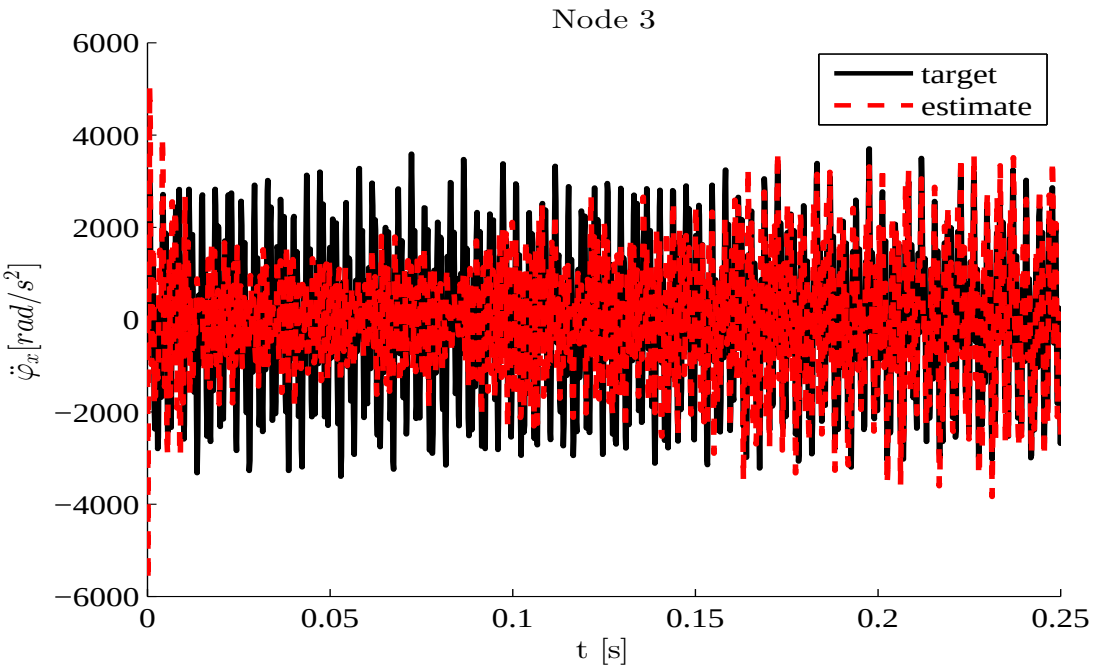
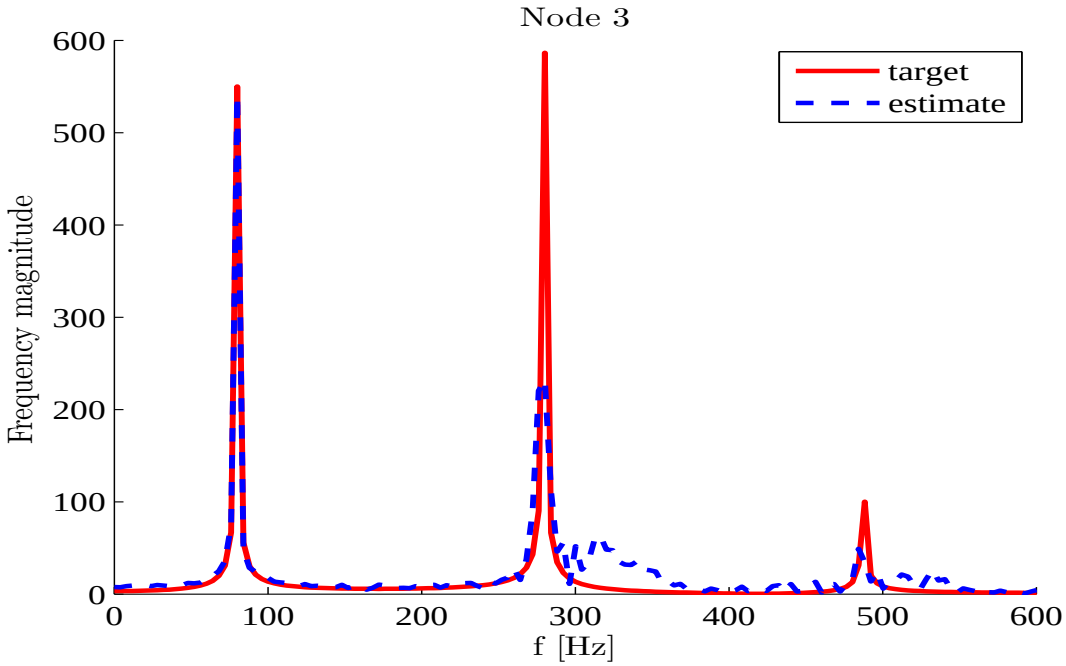


Figure 5.51: Point 3 - $\mathcal{F}[\dot{\varphi}_x(t)]$ - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-4}$



If we consider the extreme case of a measurement noise variance $\sigma_{\mathbf{v}}$ with a magnitude of equal order of the signal, some important remarks can be emphasized. This case can be basically referred to two eventualities:

- the quality of the measurement system is so poor that the noise produced hides the real mechanical response of the structure;
- the applied external force is so low that the response can not be detected by the sensors deployed.

First of all, Figure 5.52 suggests that the filter is not able to estimate the parameters or, in other words, the estimation requires so much time to reach convergence that is basically useless. Regarding the dynamics of the system (Figures 5.53 to 5.55), despite the fact the amplitude estimated is not accurate for the reasons previously explained, it seems that the frequency is correctly captured. This behavior is due to the fact that the load is supposed to be a given datum both for the real and the estimated systems. In order to better explain this concept, it is useful to remember that the Fourier transform of the output signal is equal to the product of the Fourier transform of the input signal, i.e. the sinusoidal load, and the Fourier transform of the transfer function. Let us now look at Figure 5.56: the only frequency the filter is able to estimate is the frequency associated with the applied force:

$$f = \frac{\omega}{2\pi} = \frac{500 \text{ rad/s}}{2\pi} = 79.6 \text{ Hz} \tag{5.17}$$

This frequency is captured because the input signal is known and it has not to be estimated. Moreover, the other spikes are not fitted because they are related to the natural frequencies of the structure, therefore it means that the filter is not estimating the parameters associated to the damaged structure. A further evidence is that the peaks estimated correspond to the natural frequencies of the undamaged structure (Table 5.2), therefore the estimated stiffness matrix has not been correctly updated towards the damaged case.

Figure 5.52: Parameters d_i estimation - $\sigma_{\mathbf{v}} = 1,5 \cdot 10^{-3}$

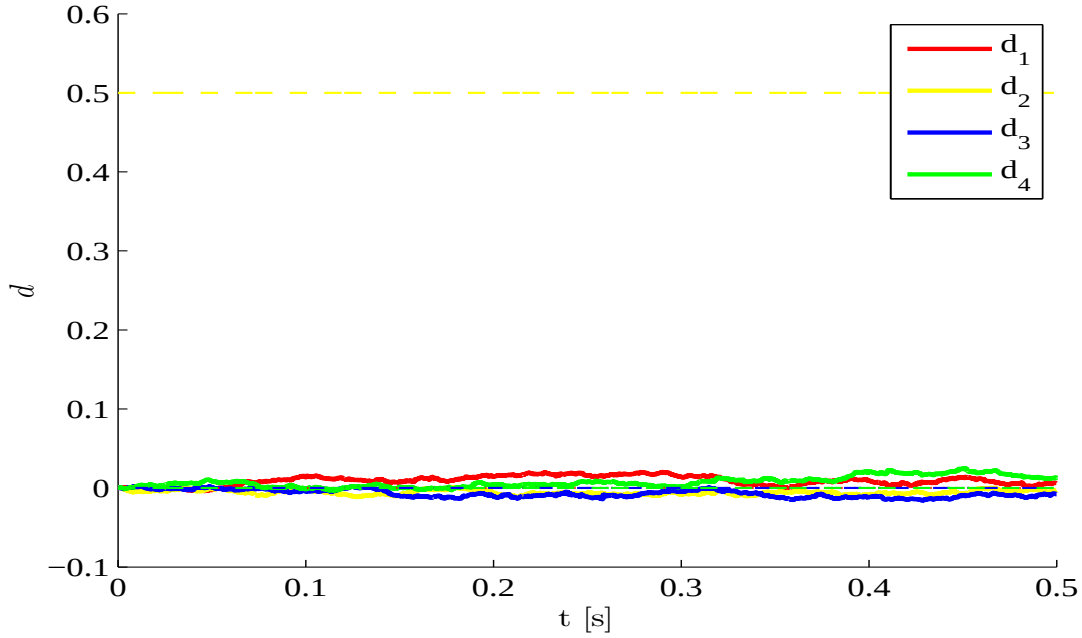


Figure 5.53: Point 3 - $\varphi_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-3}$

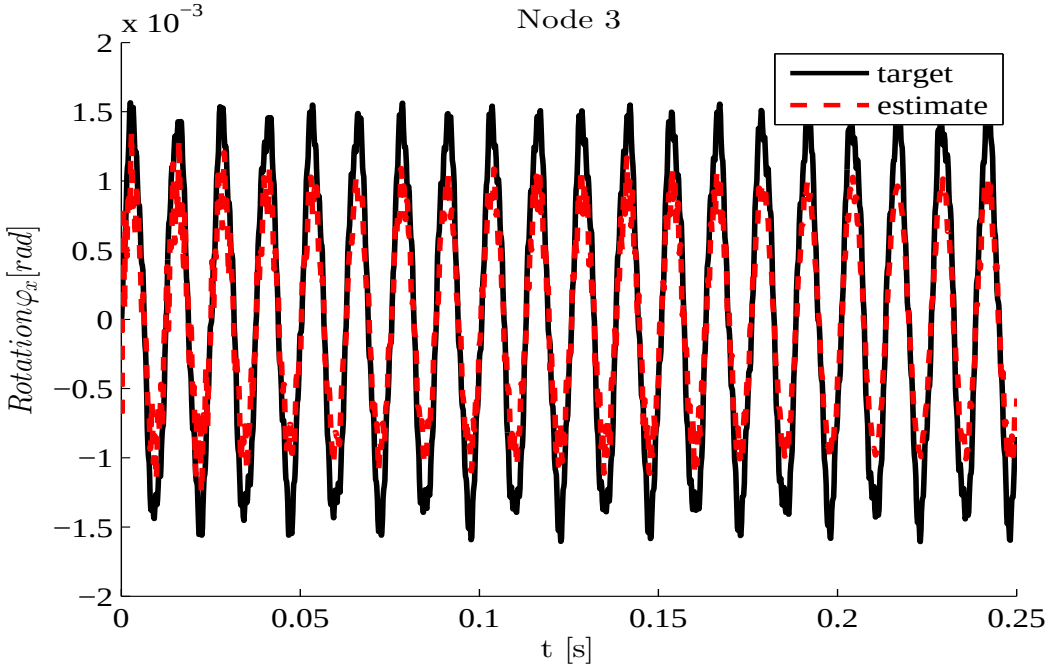


Figure 5.54: Point 3 - $\dot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-3}$

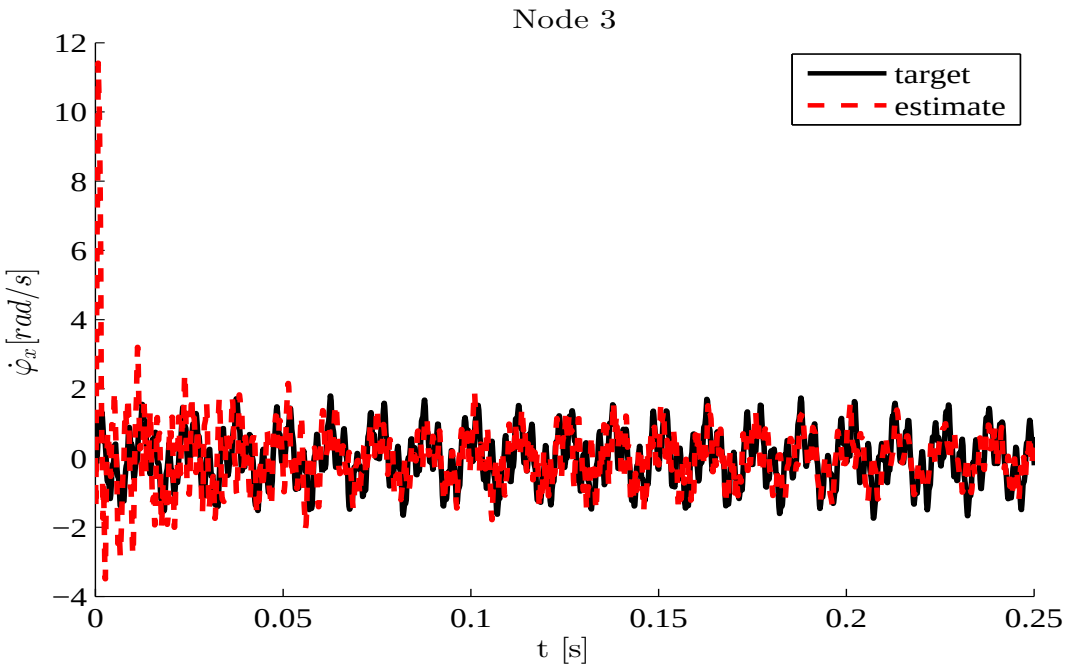


Figure 5.55: Point 3 - $\ddot{\varphi}_x(t)$ estimation - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-3}$

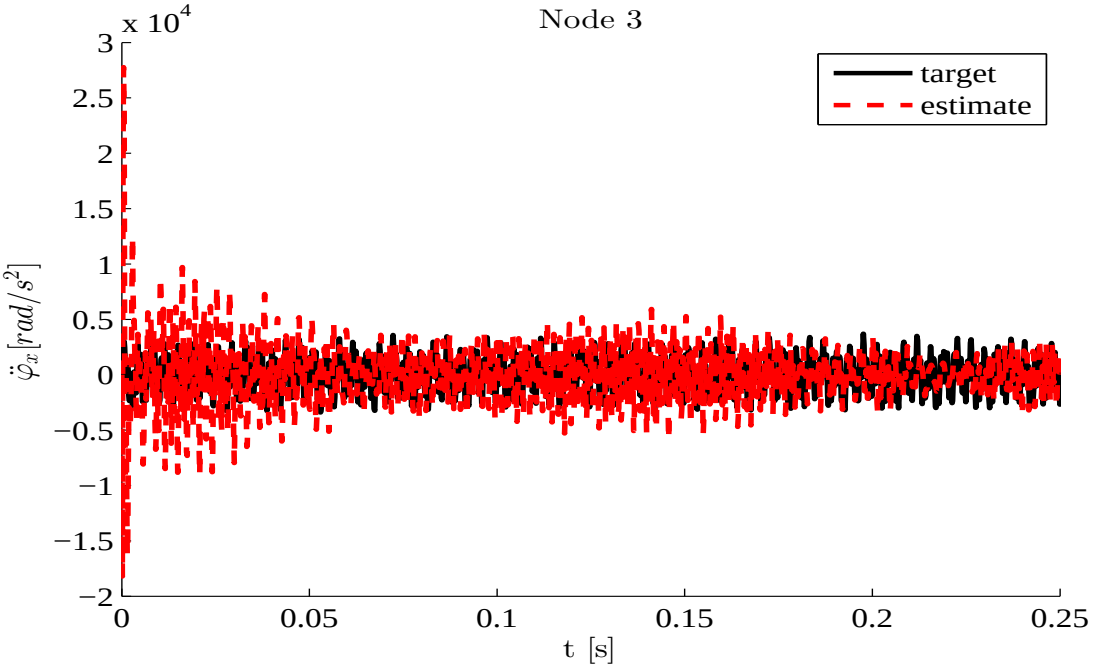
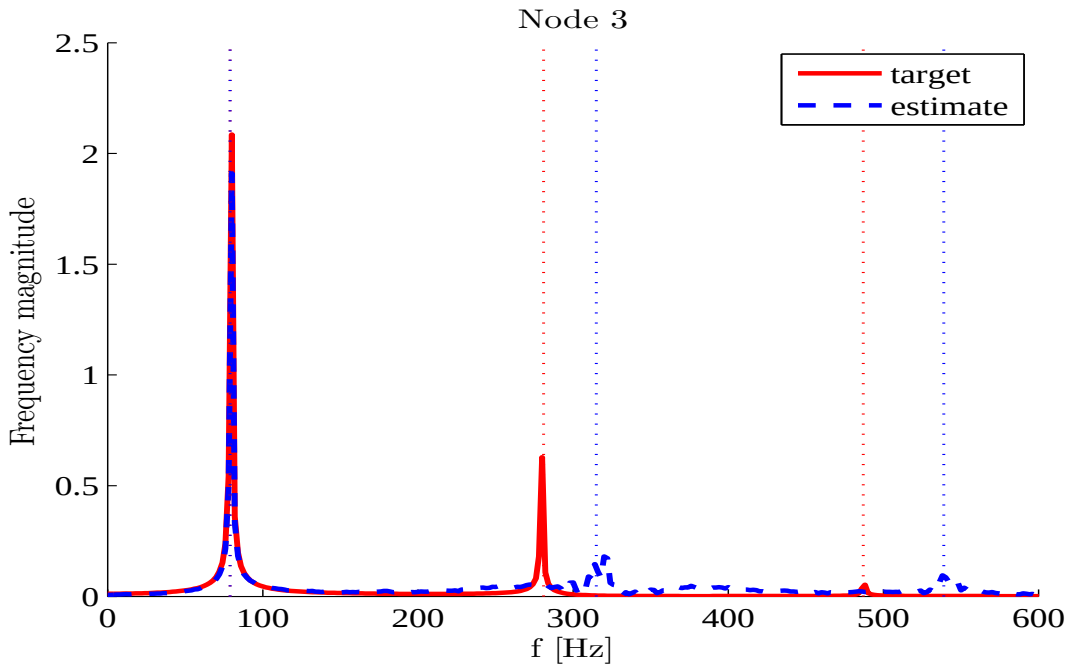


Figure 5.56: Point 3 - $\mathcal{F}[\dot{\varphi}_x(t)]$ - $\sigma_{\mathbf{v}} = 1.5 \cdot 10^{-3}$



5.2.2.4 Process noise

The next step is to consider the effect of the process noise $\mathbf{w}$ on the performance of the filter. Recalling what has been explained in section 3.2, the process noise is related to the quality of the model that describes the system considered and to all the approximations that have been introduced in the procedure. For instance, these approximations could be due to the time and spatial discretization methods employed, i.e. in this specific case respectively the Newmark integration scheme and the finite element method. A high process noise takes into account the high level of uncertainty introduced by the model. Figure 5.57 shows that considering $\sigma_{\mathbf{w}} = 10^{-2}$, the model quality is so poor that the filter is still able to reach the convergence but the accuracy is low and the estimates fluctuate in an unacceptable way. A lower level $\sigma_{\mathbf{w}} = 10^{-5}$ (figure 5.58) corresponds to little uncertainties, but on the other hand the variation of the parameters over time is prevented. The result obtained is a very smooth but slow variation of estimates. Recalling Figure 5.10, it is possible to note that the standard deviation of the process noise $\sigma_{\mathbf{w}} = 10^{-5}$ is approximately 5% of the maximum rotation ($1.5 \cdot 10^{-3}$ rad). In other words, since we are assuming a Gaussian distribution, the probability that the relative error of the simulated response with respect to real response is under 5% is almost 0.70.

Figure 5.57: Parameters d_i estimation - $\sigma_{\mathbf{w}} = 10^{-2}$

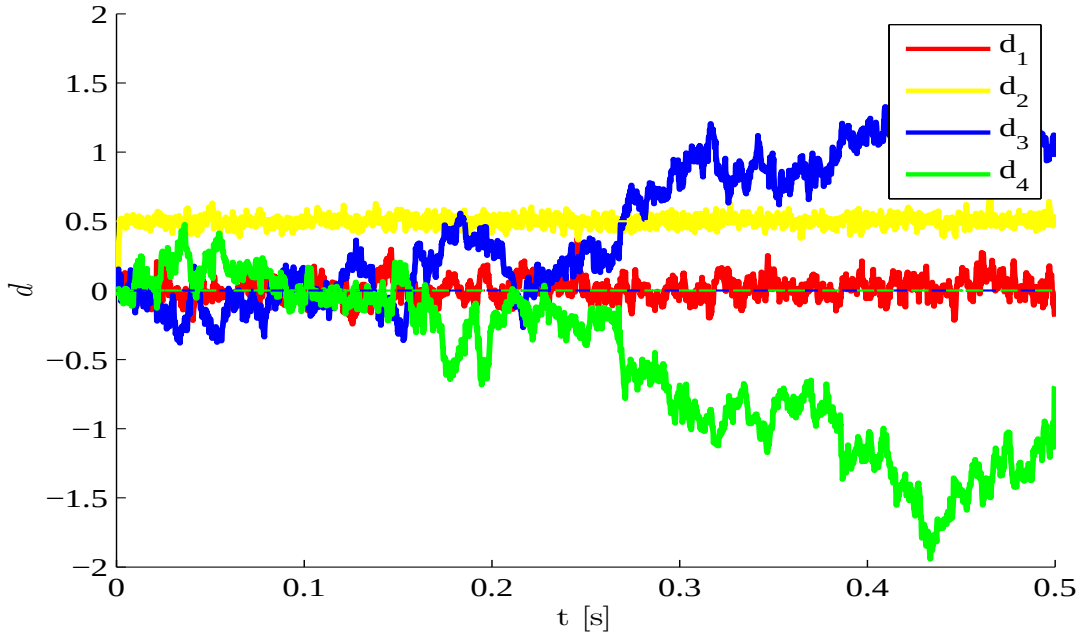
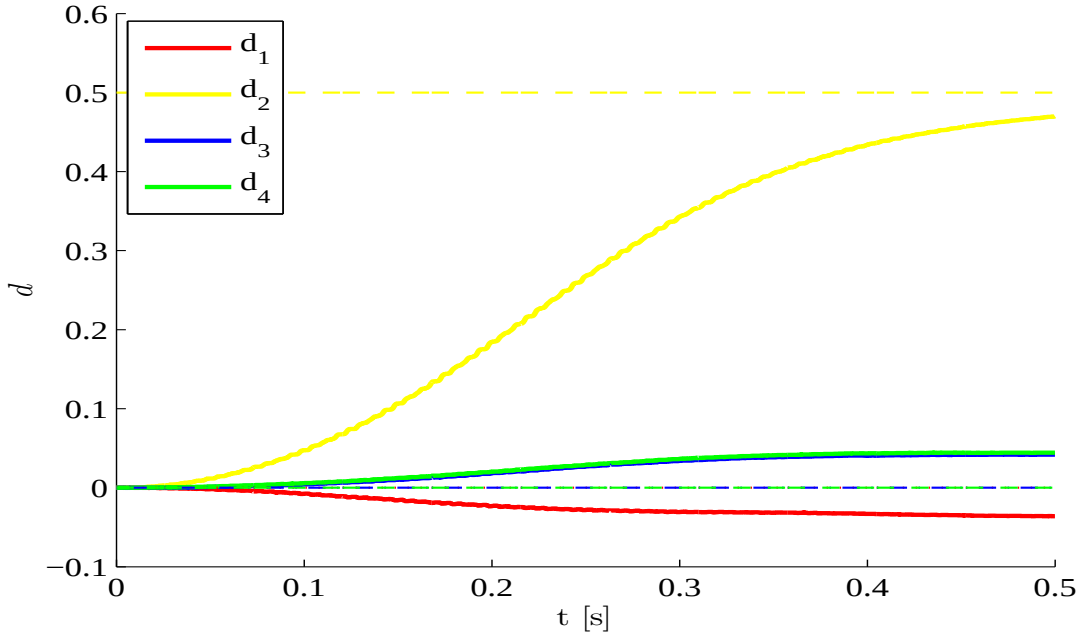


Figure 5.58: Parameters d_i estimation - $\sigma_{\mathbf{w}} = 10^{-5}$



5.2.2.5 Number of observations

As shown in equation (5.10), all the results obtained and shown so far refer to the case in which the measurements are collected from all the four mid-point nodes (2, 4, 5, 7).
At this point, it could be interesting to analyze whether the filter is able to detect the damage also for a lower number of sensors. Let us consider three different cases:

1. the rotations φ_x and φ_y at node 7 are measured;

2. the rotation φ_x at node 7 is measured;

3. the rotation φ_y at node 7 is measured.

From Figure 5.59 it is possible to deduce that the filter is still able to detect the parameters d_i, but, comparing with Figure 5.39, the accuracy of the estimation is somehow worse. The estimation ability is guaranteed also for a low number of observations, because the structure is very simple. Moreover, since we are basically only performing a benchmark of the method, the discretization of the plate is very coarse. For more complicated structures or with a higher number of elements, it can be shown that the placement of the sensors is a critical point. Some methods have been developed in order to optimally place a certain number of sensors, such that the level of information obtained is maximized [113, 110].
Considering Figure 5.60, the estimation is basically similar to the 1-POM case (Figure 5.37): the measurement is sufficient to detect the damage but it is not able to numerically evaluate it. The behavior is far worse considering the third case (Figure 5.61): none of the parameters has been correctly estimated and, moreover, the location of the damage has not been detected. Nevertheless, the performance of the filter can be considered still quite good, since actually we are estimating four parameters using a 3-POMs reduced order model and with only one observation.

Figure 5.59: Parameters d_i estimation - φ_x^7 and φ_y^7 measured

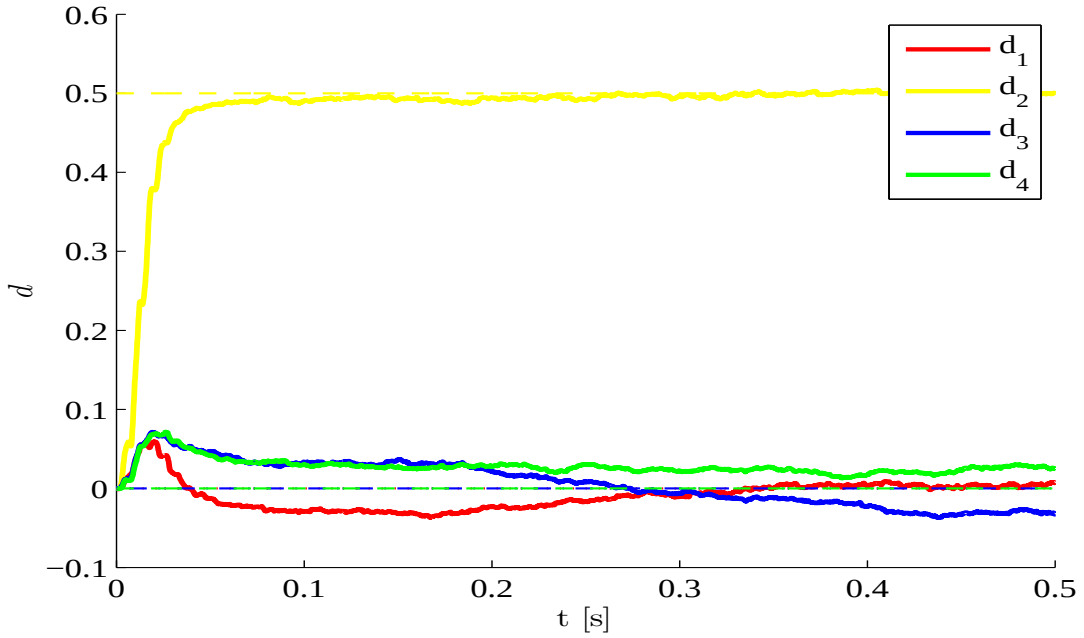


Figure 5.60: Parameters d_i estimation - φ_x^7 measured

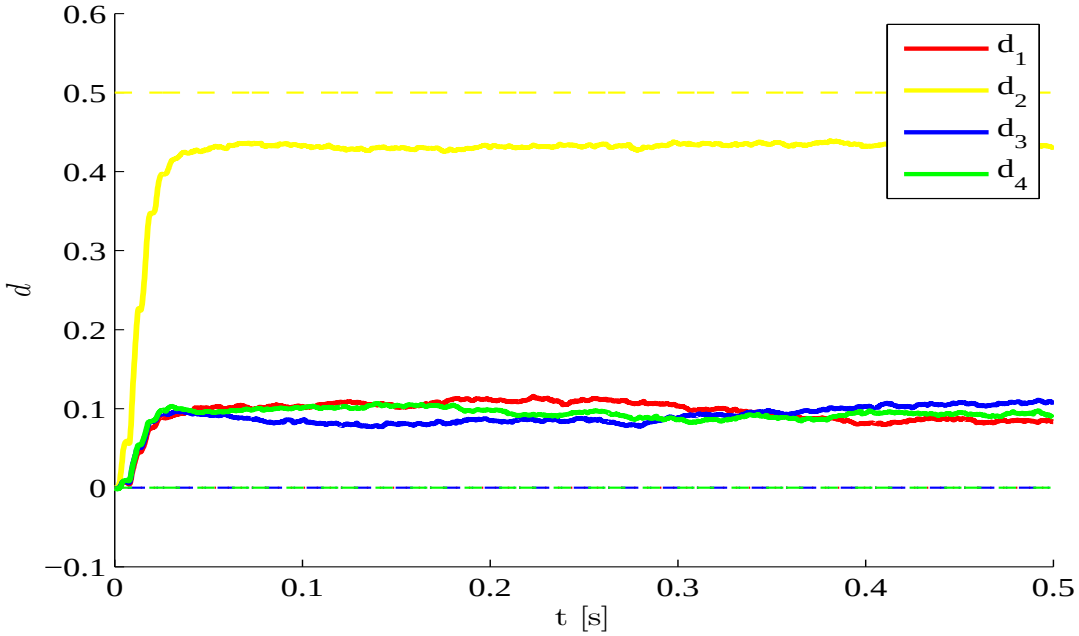
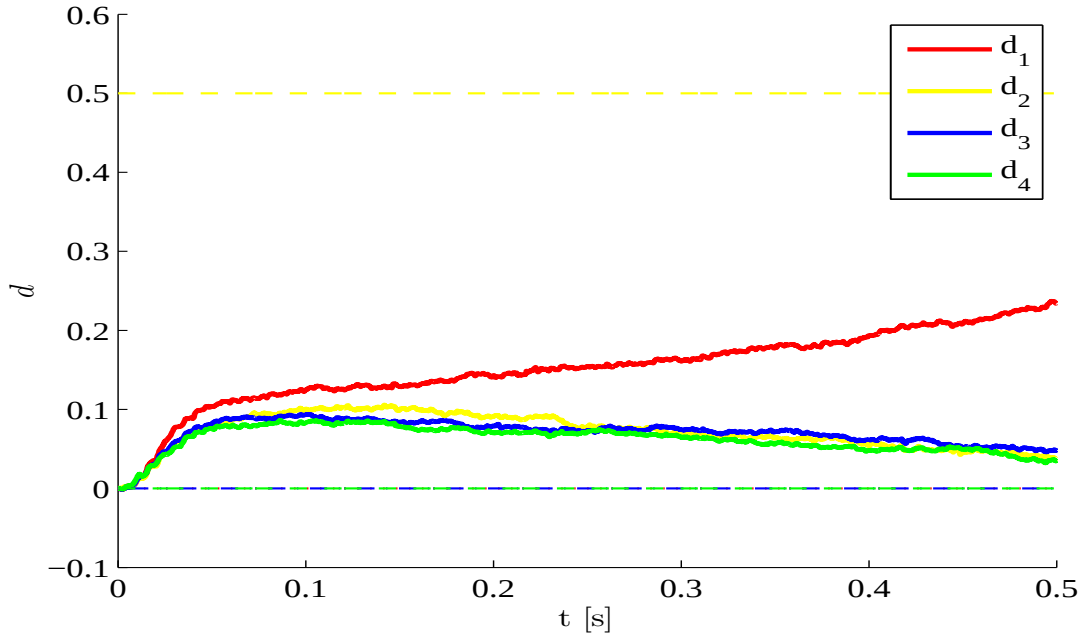


Figure 5.61: Parameters d_i estimation - φ_y^7 measured



5.2.2.6 POM convergence

In all the simulations performed so far, the system was supposed to be stationary, i.e. no variations of the full order stiffness and mass matrices were occurring over time. Consider now the case of a system that could change: for instance an impact or a fast variation in the stiffness conditions are happening. In this case, the updating stage of the POMs is fundamental in order to be able to detect and identify the damage. Let us assume the case in which a variation of the damage parameter is occurring at time $t = 0.25$ s, from $d_2 = 0.5$ to $d_2 = 0.3$.
Figure 5.62 shows what would happen assuming that the POMs are not updated and remain constant over time. As it was expected, a rapid variation of the damage parameters is evident, but the new parameters are not well estimated. This is due to the fact that a wrong bases is projecting the full order matrices in a subspace that is not coherent with the original full order space. In other words, the bases that we are using are not the optimal ones, i.e. the error is not optimally minimized. Despite the fact the filter can not identify the right values of the damage parameters, still the method can detect a whatever variation in the system and therefore it could be used as a sort of damage alert, warning the user that something unexpected is occurring to the system. Practically, a warning condition could be defined as follows:

$$\exists\ d_i \text{ for } i = 1, ..., N_p \text{ s.t. } |d_i(t_k) - d_i(t_{k+s})| > \zeta \tag{5.18}$$

In other words, the variation of the system is occurring if at least one of the damage parameters is changing unexpectedly within a certain time interval. Since the parameter estimation is always affected by uncertainties and fluctuations due to the measurement and process noises and to the random generation of the particles, the value of ζ should be tuned such that those aforementioned fluctuations are not considered as an effect of the system variation. The value of s should be tuned as well, depending on the type of structure, the time interval between each step of the procedure and the type of variation is expected to occur. Consider now the case in which the POMs are updated. As shown in Figure 5.63, after an interval in which the estimates changes, the damage parameters are identified correctly. Figure 5.64 shows the convergence of the first three POMs calculated using the L^2 relative error:

$$\epsilon(t) = \frac{||\hat{\Phi}_l(t) - \Phi_l(t)||_{L^2}}{||\Phi_l(t)||_{L^2}} \tag{5.19}$$

It can be stated that the convergence is monotonic and therefore the error associated to the POMs decreases over time.

Figure 5.62: Parameters d_i estimation - Not updated POMs

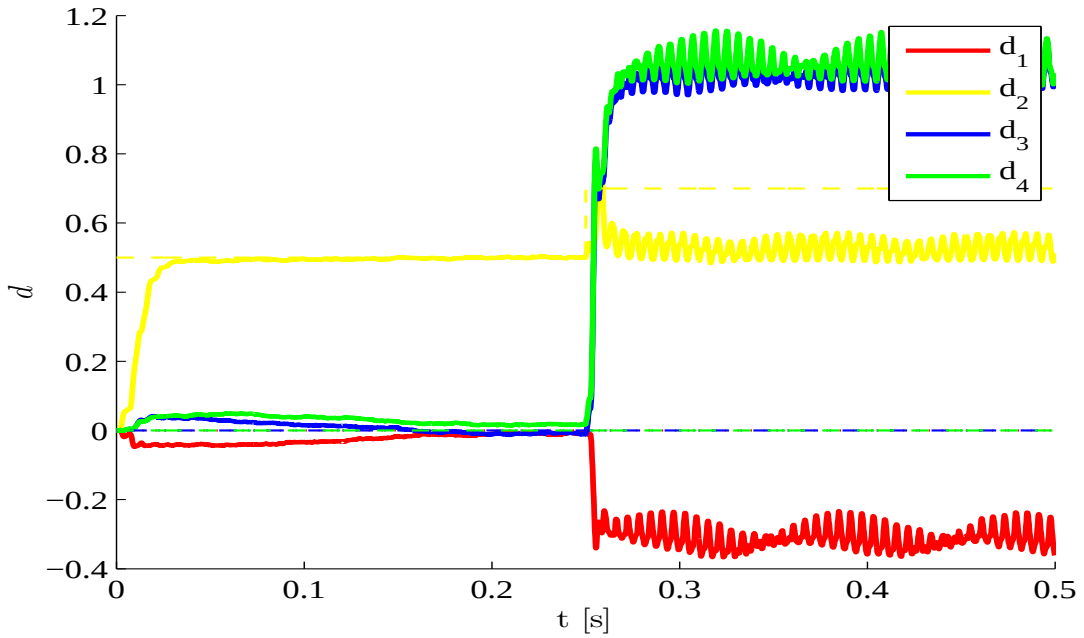


Figure 5.63: Parameters d_i estimation - Updated POMs

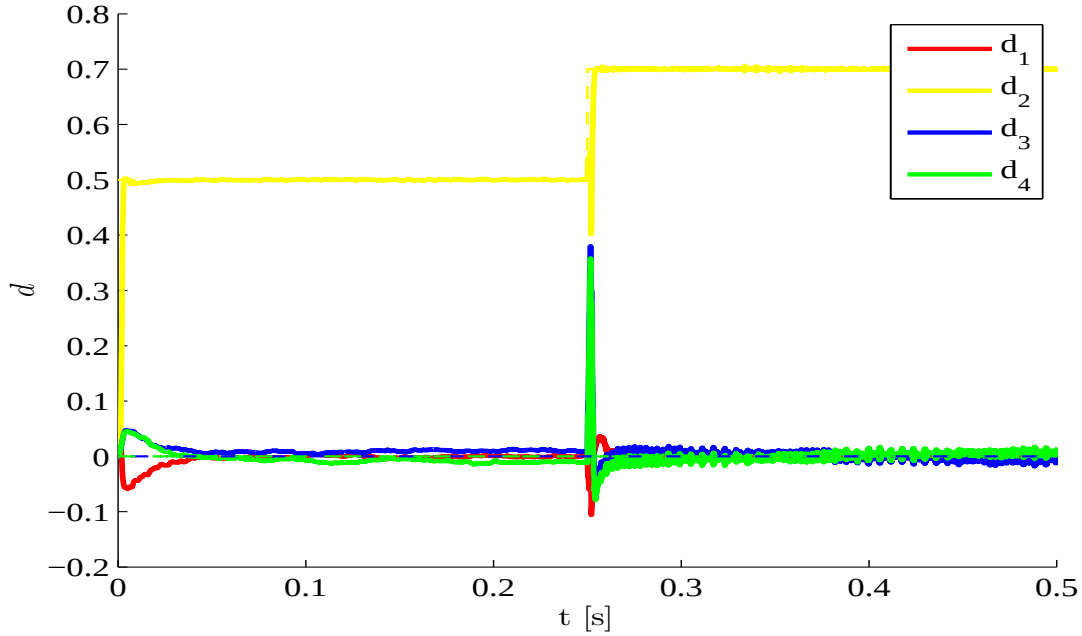
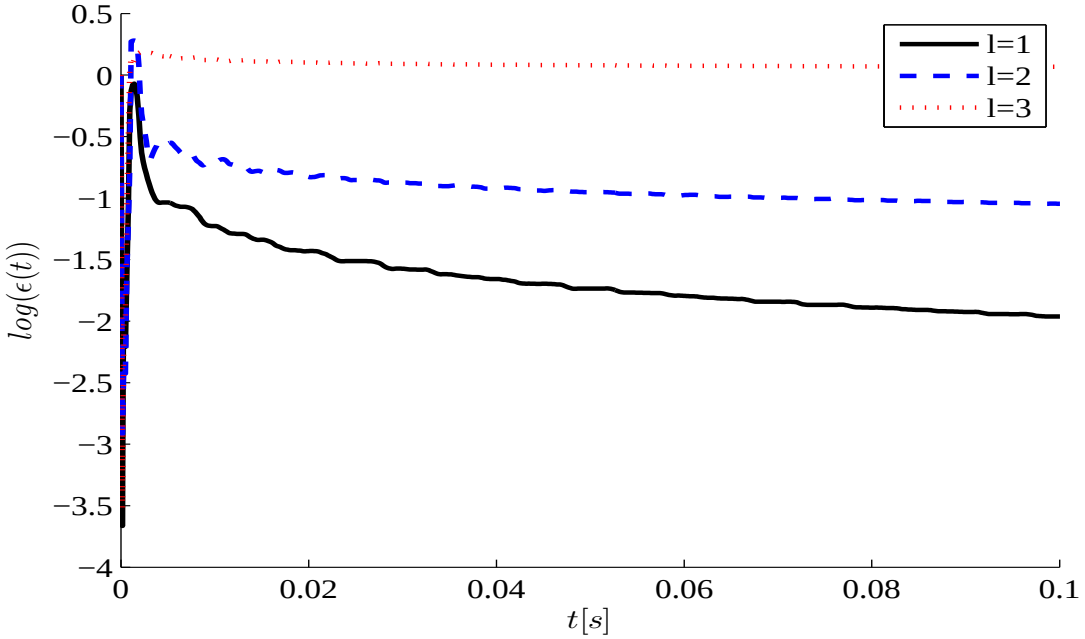


Figure 5.64: POMs convergence - $\epsilon(t) = \frac{||\hat{\Phi}_l(t) - \Phi_l(t)||_{L^2}}{||\Phi_l(t)||_{L^2}}$



5.3 Additional analyses

In the previous section a benchmark analysis has been performed in order to evaluate the performances of the damage identification algorithm. It has been therefore used a simple 4 elements square thin plate and the effects on the filter of the different settings conditions (such as number of POMs, measurement noise, process noise, etc.) have been compared.
We want now to apply the damage identification strategy to a more accurate FE model: we consider then the same structure described in Section (5.1) with a more refined mesh. It has been shown in [114] that the error of the maximum deflection and bending moment calculated with a 8×8 elements FE model with respect to the analytical solution of the simply supported plate is less than 1%. Therefore, here a 10×10 elements discretization is chosen, practically making the discretization error negligible. The full order model have, then, 100 elements and 121 nodes: recalling that 6 degrees of freedom are associated to each node, the overall order of the full model is 726.
The relevant first natural frequencies of the damaged and undamaged structure are respectively shown in Tables 5.7 and 5.8.

Table 5.7: Natural frequencies of the damaged structure

vibration mode index	1	2	3	4	5	6
natural frequency [Hz]	196.9	415.2	462.5	549.8	1090.1	1242.3

Table 5.8: Natural frequencies of the undamaged structure

vibration mode index	1	2	3	4	5	6
natural frequency [Hz]	219.9	477.3	477.3	597.3	1205.2	1359.8

The 10×10 elements model is of course more flexible than the benchmark model, therefore the relevant natural frequencies are lower (see Tables 5.1 and 5.2).

5.3.1 Model order reduction

Let us not consider the damaged case, as defined in Section 5.1. In Table 5.9 the level of information $I(l)$ is shown.

Table 5.9: Damaged case - $I(l)$

l	$I(l)$
1	0.997208
2	0.999865
3	0.999992
4	0.999998

Comparing Table 5.9 and Table 5.4 we can see that the level of information retained by the first two POMs in this case is higher than the one of the benchmark case. Nevertheless, as it will be shown later on, the higher order modes are not negligible and they play an effective role on the structural response.
From Figure 5.65 it is possible to underline that even the first two POMs are sufficient to catch the basic behavior of the displacement u_z time history. Moreover, considering the relative L^2 error (defined in equation (5.6)), it can be stated that the solutions found with 3 and 4 POMs are characterized by a non-negligible error. Comparing Table 5.66 and Table 5.34, it is possible to conclude that the higher modes can not be neglected, as it was possible in the benchmark case. This statement can be better understood by looking at the acceleration $\ddot{u}_z$ time history at point 2 (Table 5.67) and its respective Fourier transform (Table 5.68).

Figure 5.65: Damaged case - Point 2 - Displacement $u_z(t)$

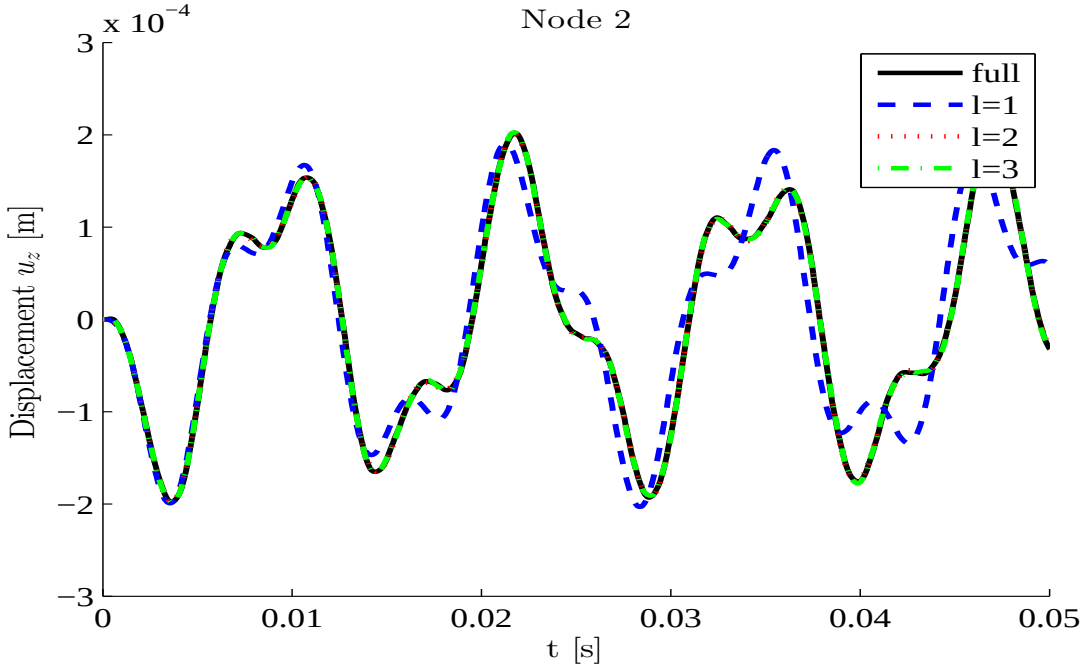


Figure 5.66: Damaged case - Relative error - $\epsilon_l(t) = \frac{||\mathbf{u}_l(t)-\mathbf{u}(t)||_{L^2}}{||\mathbf{u}(t)||_{L^2}}$

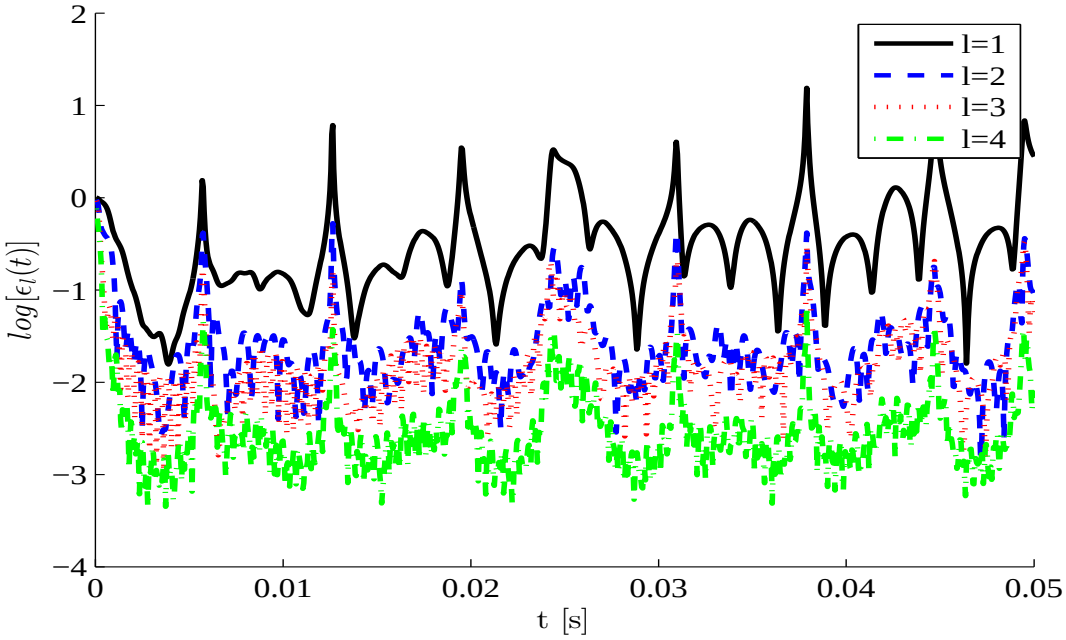


Figure 5.67: Damaged case - Point 2 - $\ddot{u}_z(t)$

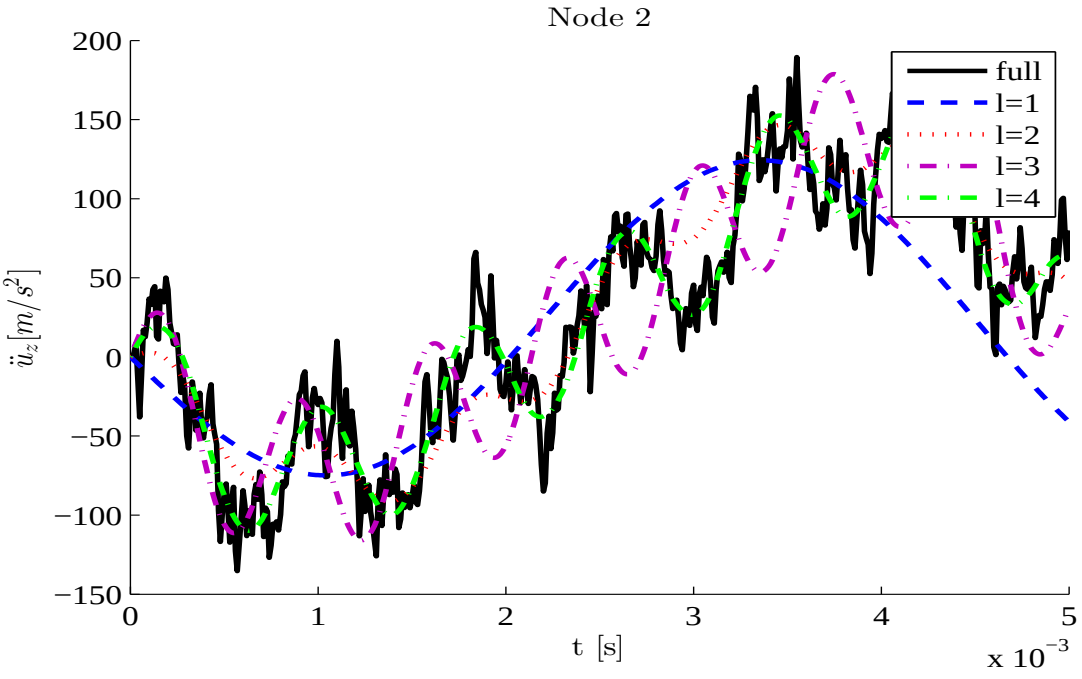
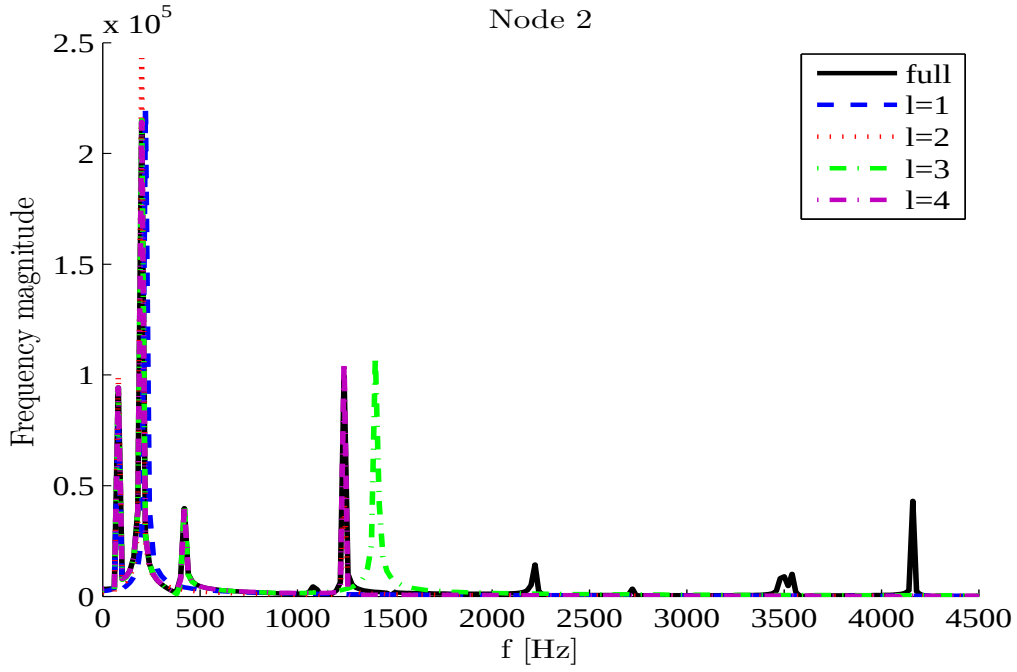


Figure 5.68: Damaged case - Point 2 - $\mathcal{F}[\ddot{u}_z]$



In order the reduced order model to be able to catch all the frequency peaks of the full order model a high number of POMs is required. The importance of the higher order POMs is basically related to the fact that the model is more flexible and hence more relevant mode configurations are allowed.

5.3.2 POD Kalman observer

Let us now consider the damage identification filter and its application to the 10×10 elements FE model. Since the effect of the measurement noise, the process noise and the initial conditions on the operation of the algorithm has been already shown, let us consider which is the parameter estimation, given a certain number of POMs retained. In all these simulations, the same settings chosen for the benchmark analysis have been used: $\sigma_{\mathbf{v}} = 10^{-5}$, $\sigma_{\mathbf{w}} = 10^{-5}$, $\mathbf{d}_0 = \mathbf{0}$. The observations are the same ones described in equation (5.10).
Tables 5.69 and 5.70 show the estimation of the damage parameters d_i when respectively 2 or 3 POMs are used. Unlike in the benchmark case (Table 5.39), the filter is not able to estimate any damage parameter: this is an expected result because we are now considering a model with a larger number of degrees of freedom and therefore the grade of reduction is much higher. Considering 3 POMs, the damage parameter estimation is guaranteed: this is a remarkable result, since we are able to estimate 4 parameters of a structure with 726 degrees of freedom, using a state vector of dimension 9.

Figure 5.69: Parameters d_i estimation - 2 POMs

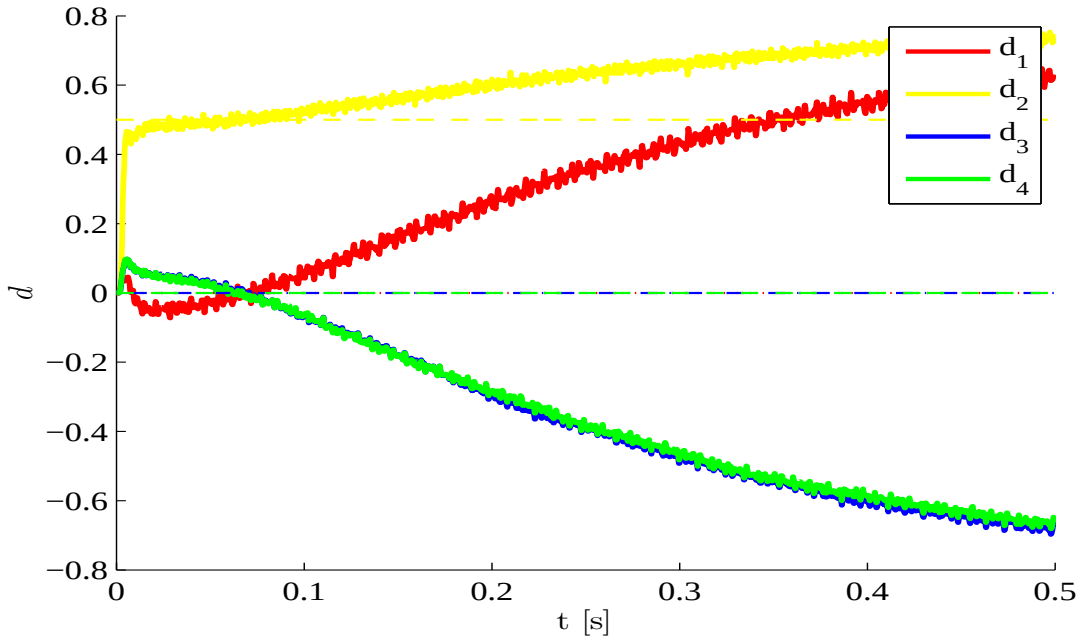
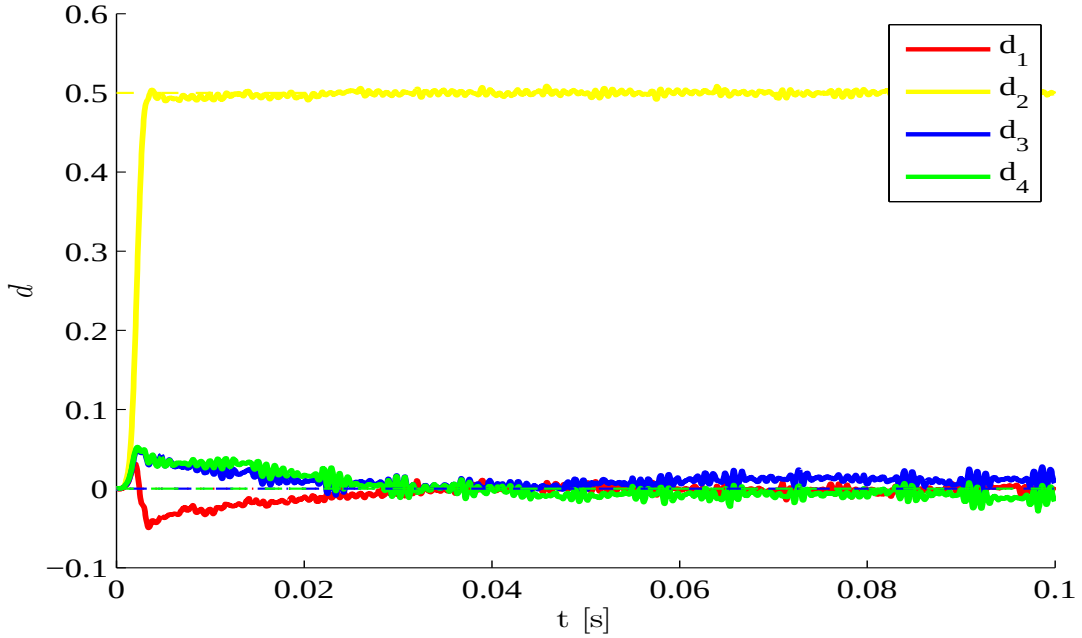


Figure 5.70: Parameters d_i estimation - 3 POMs



5.4 Conclusions

First of all, a general description of the structural problem used to evaluate the performances of the procedure has been given in Section 5.1: a simply supported thin square plate subjected to a smooth load is considered, both in a undamaged and damaged case.
In Section 5.2, a benchmark analysis has been performed considering the most simple discretization of the plate, i.e. a 4-elements FE model. The model order reduction has been exploited (Section 5.2.1) in order to compare the results obtained using the full and the reduced models, both for the damaged and an undamaged case. It has been shown that even if the number of POMs retained is low, i.e. $l = 2$, the full model evolution is reproduced with a high accuracy. Considering higher order models, the cumulative error is basically constant over time. The comparison between the damaged and the undamaged case leads to conclude that a less symmetrical structure requires a higher number of POMs in order to reach the same level of accuracy. The convergence of the POMs has been shown, when a different number of snapshots is retained.
Afterwards, in Section 5.2.2, the procedure described in Figure 4.1 has been evaluated in terms of damage estimation and dynamic tracking. In order to summarize the results, we can lead to the following conclusions:

- A 1-POM reduced model is able to detect the damage but can not assess it; higher order models can estimate the damage parameters with increasing accuracy.

- A little variation of the initial conditions is not detrimental for the estimation: it has been shown that initial conditions that goes from 0 to 1.5 times the target values of the stiffness are acceptable. Higher initial values leads to divergence of the estimates.

- The lower is the measurement noise level, the more accurate and fast is the estimation of the parameters; on the other hand, a variation of the measurement noise $\mathbf{v}$ does not affect so much the dynamic tracking of the system because the applied load is a given datum for the filter. High levels of $\mathbf{v}$ (standard deviation of the measurement noise $\sigma_{\mathbf{v}}$ higher than 10% of the signal amplitude) prevents the filter from identifying the damage, because it conceals the real dynamics of the system.

- High levels of process noise induce high fluctuations of the estimated parameters; on the other hand, a quick convergence is reached. The estimation becomes slow but more accurate if the level of $\mathbf{w}$ is low.

- The number of measurements affects the estimation in terms of accuracy. It has been shown that using a 3-POMs reduced model, two measurement

are sufficient to estimate the damage parameter. Assuming only one measurement is available, some degrees of freedom can retain a higher level of information. For this reason, depending on the position and type of measurement used, we can either detect the damage or lead to divergence of the estimation. Nevertheless, one single measurement is never sufficient to quantify the damage.

- The convergence of the POMs over time is shown to be monotonic and the effect of the POM updating phase is underlined.

In Section 5.3 the model order reduction method and the damage identification procedure are applied to a 10×10 elements FE model of the damaged plate described in Section 5.1. Even if the number of degrees of freedom is much higher than in the benchmark case, it has been shown that the damage parameters can be still well estimated with a 3-POMs reduced model.

Chapter 6

Structural health monitoring system

In the previous Chapters, an identification damage procedure has been presented and tested through a simulated analysis. The first main goal of the method is of course the estimation of some damage indexes correlated to a certain number of regions in which the structure has been conceptually divided. Considering the wide range of system identification methods based on Bayesian recursive filters that are used nowadays, the most well-established and reliable systems are off-line procedures. Therefore, the data acquisition and computational steps are performed in different time intervals. This is due to the high computational cost of these recursive calculations. In this work, we have used an improved method proposed in [62], with the aim of dramatically reducing the computational demand and make it an on-line or nearly on-line identification method. As explained before, in order to reach these objective, both a reduced order model scheme and a particle filter using a re-sampling technique have been adopted.
In the present Chapter we will focus on the description of the acquisition system that could be used for a possible experimental implementation of the method, and could practically lead to its application on a real structure as well.
According to [13], the definition of a data acquisition system should always consider the following conceptual main features:

- the selection of the type of sensors;
- the choice of the optimal number and location of the sensors;
- the definition of an appropriate network that allow to control the sensors, collect and store the data.

Last but not least, a computational unit should be used to switch on and off the sensors, control the timing of the acquisition signal and perform the computation of the identification model.

In the following Sections we will focus on these topics: first of all, a brief review about the state of the art will be given; then, the proposed acquisition system will be described.

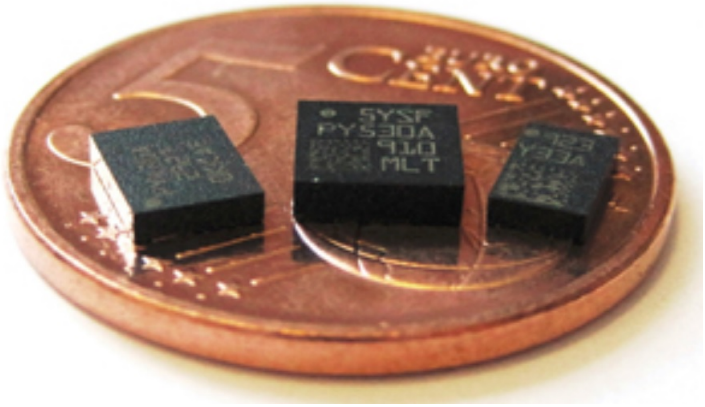
6.1 Sensors

In the last few years, as a consequence of the improvements in electronic engineering, a large number of different sensors has been developed: the choice of the type of sensors depends of course on the type of input data required by the identification model, and the type of structure to be assessed. Some remarkable technologies are: fiber optics sensors used to measure strains [115, 116] or detect defects [117] and delamination [118]; piezoelectric sensors for electrical impedance-based [119]; elastic wave-based SHM method [120]; magnetostrictive sensors [121]. The most used and affordable methods are based on acceleration or rotation signals. Different types of accelerometers are available nowadays: AC charge-mode piezoelectric, AC voltage mode piezoelectric, DC capacitive or DC piezoresistive. Despite the traditional accelerometer technologies are very accurate and reliable, due to their weight and sizes, they can be employed only in structures whose mechanical behavior is not affected by the sensors themselves. Therefore, they are usually used to assess civil structures, such as bridges [9], buildings [122] or infrastructures [123].
Consider now lightweight structures, used for example in aeronautical applications. The deployment of heavy sensors would affect the mechanical response of the system and hence could not be considered as an efficient method to acquire data. Because of this problem and due to the fact that traditional high accurate accelerometers could have very high costs per unit, in recent years a new type of accelerometeetrs have been exploited in SHM systems: the so-called micro electro-mechanical systems (MEMS). Comparing the MEMS accelerometers with the conventional ones, some important advantages can be underlined:

- very low price;
- small size (see Figure 6.1);
- very low weight;
- low power consumption;
- multi-chip mounted wafers.

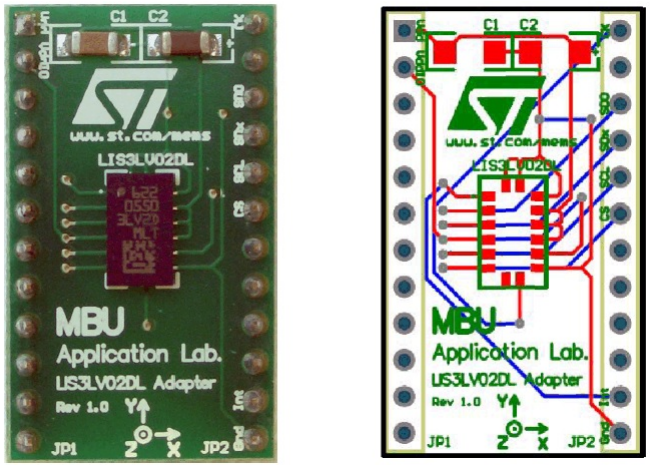
Several types of MEMS are available on the market, depending on the characteristics required by the applications. Despite the first devices were characterized by low performance and reliability compared to the conventional ones, nowadays

Figure 6.1: Size comparison of a MEMS ultra-compact LGA package



commercial off-the-shelf MEMS are approaching those standards. In the application presented here and in general in SHM systems, some MEMS that can be found on the consumer market have been used: the relatively low performances can be overcome by using a higher number of sensors.
In this particular application, a MEMS inertial 3-axis digital output low voltage linear accelerometer (LIS3LV02DL [5]) produced by STMicroelectronics could be used. The device can sense and measure the accelerations through the variation of capacity: a set of suspended micrometric silicon structures, which are connected to the die at some points, can move towards one of the directions of the accelerometer axes. Once the base is accelerated, the mass displacements induce a variation of capacity that is sensed by an appropriate circuit and converted to a digital signal.
The device is able to measure accelerations with output data rates (ODR) up to 2560 Hz, a device resolution of 15.6 mg and the package size is 4.4x7.5x1.0 mm. In order to hold also two filtering capacitors and provide the required connections to the network system through a standard socket, each device is attached to an adapter board (Figure 6.2 [4]). Considering the extremely small sizes of the adapter board (30x18 mm) and the maximum frequency bandwidth ($ODR/4$), that is the highest frequency signal that can be sampled by the aforementioned output data rate without the occurrence of aliasing effect, this type of accelerometer can be applied even for extremely stiff and small structures as the one used for the benchmark analysis. Comparing in fact the first natural frequencies of the undamaged structure (Table 5.8) with the accelerometer bandwidth, the system can sense up to the fourth mode of the structure itself.
As it will be explained later on, the communication with the computational unit (PC) occurs directly through the adapter board pins and a low level data interface. An alternative solution to the adapter board could be the use of an

Figure 6.2: Adapter board photograph and board layout [4]



evaluation board [124] that allows to obtain ready-to-use data through a USB communication standard port and a computer user interface. The major drawback of this solution is the motherboard and USB connector sizes: moreover, in some particular cases, the board stiffness could affect the mechanical response of the monitored structure. A second disadvantage is that the USB communication protocol is not a flexible solution and does not allow to connect all the devices through a smart networking scheme.
Once the basic characteristics of the sensor have been described, let us now explain the digital communication protocol. The device could work either with a I^2C (Inter-Integrated Circuit) or SPI (Serial Peripheral Interface) bus. Despite the fact SPI requires a larger number of communication lines, thanks to its easier implementation and faster operation, it is more widespread and hence it will be here explained.
The SPI is a bi-directional digital interface between the sensor (called slave) and the acquisition and computational unit (called master). The communication is based on four signals and thus four wires are required:

- Chip Select (CS) is a digital signal controlled by the SPI master and basically starts and closes the communication at the beginning and at the end of each byte;

- Serial Port Clock (SPC) is an output from the master and, using a digital sequence of evenly spaced time series of 0 and 1 bits, controls the timing of all the other signals;

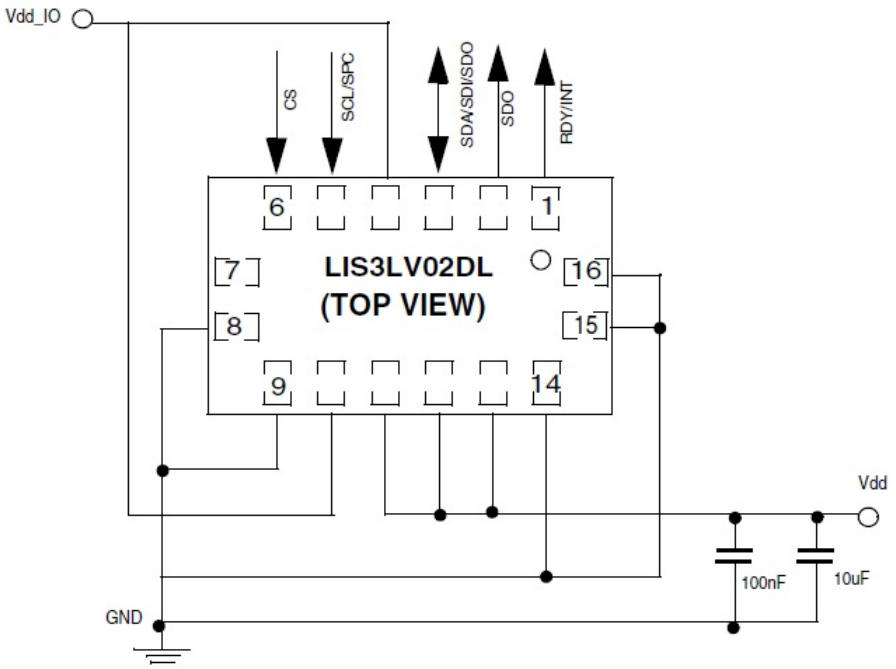
- Serial Data Output (SDO) is a master output and provides the register's code needed to control and query the device and hence its response;

- Serial Data Input (SDI) is the output from the slave to the master and follows the requests given by the SDO.

Further details on the communication sequence between the sensor and the computational unit will be given in Section 6.3.
In Figure 6.3, all the electrical connections the accelerometer requires are shown. In addition to the four SPI lines previously described, the other connections represents:

- Data-Ready (RDY/INT) is a signal that goes to 1 whenever all the data from the three axes accelerations are available;
- GND represents the reference voltage level for all the other pins (0 V power supply);
- Vdd and VddIO represent the power supplies respectively for the device and for the input/output communication; their voltage level are referred to the GND voltage.

Figure 6.3: Scheme of the electrical connections [5]



The two $100\mu F$ and $10\mu F$ capacitors indicated in Figure 6.3 are the same components shown on the adapter board and respectively named as $C1$ and $C2$ (Figure 6.2).

6.2 Optimal Placement

The optimal placement of the sensors is a critical issue in structural health monitoring, because it can lead to important cost saving. In other words, given a certain number of sensors, an optimal deployment of devices can provide more information and hence an increased accuracy of the overall identification procedure.
Several methods have been developed in order to achieve the goal, and they can be grouped into two basic methodologies:

- the sub-optimal methods based on iterative techniques; for example, one of these methods is the Effective Independence Method [125] which reduces iteratively the number of sensors by deleting the degrees-of-freedom of a particular information matrix, leading then to the selection of the only important sensors;

- classical optimization formulation [126], where an objective function is defined and through a variety of methods the function is optimized.

We recall here and briefly explain an optimal sensor placement method developed in [6] for the case of thin plates, monitored through a MEMS pattern. Let us first of all consider an undamaged structure and a damaged structure with the same geometry of the first one but with an unknown damage (in our case see Section 5.1). The structures are then discretized in N elements with area A_i. The basic idea of the optimization scheme is to look for that pattern of sensors for which the difference $||\boldsymbol{\varphi}_{ki} - \hat{\boldsymbol{\varphi}}_i||$ between the rotations φ_i, measured from the sensors, of the two structures is maximized. $\boldsymbol{\varphi}_{ki}$ is the rotations vector measured in the i-th element for a damage located in the k-th element. A discrete density field x_i is defined: the generic x_i value is associated to the i-th sensor that is theoretically deployed at the i-th element. x_i assumes value 1 if the i-th sensor should be placed according to the optimal solution, 0 if not. The definition of the aformentioned utility function can be therefore given:

$$f(x_i) = \sum_{k=1}^{s} \alpha_k \frac{\sum_{i=1}^{N} x_i^p ||\boldsymbol{\varphi}_{ki} - \hat{\boldsymbol{\varphi}}_i||}{\max_i \left[\sum_{i=1}^{N} x_i^p ||\boldsymbol{\varphi}_{ki} - \hat{\boldsymbol{\varphi}}_i||\right]} \tag{6.1}$$

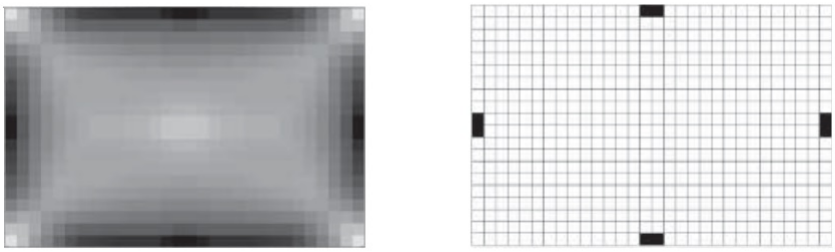
Then, the definition of the optimization problem is:

$$\arg\max[f(x_i)] \text{ subject to } \frac{1}{A}\sum_{i=1}^{N} x_i A_i \leq \bar{N} \tag{6.2}$$

The latter constraint is required in order to take into account the fact that the number $\bar{N}$ of sensors deployed can be lower than the number of elements N. The relevant weights α_k takes into account the relevance of each damage contribution

to the objective function. Practically $\alpha_k > 1$ when a particular zone of the structure should be monitored with more attention. Since different damage locations could lead to different amplitudes of the effect on the structural response, each term of the sum is weighted by the relevant maximum.
In Figure 6.4 the contour plot of the objective function using the formulation described in equation (6.1) is shown on the left; the figure refers to a flexible plate where 8 sensors are deployed and assuming the position of the damage is unknown. The figure on the right shows the discrete density field x_i: the black squares correspond to the positions for which the maximum variation of rotation can be sensed and hence the 8 sensors should be placed.

Figure 6.4: Optimal placement of 8 MEMS sensors on a flexible plate damaged in an unknown position [6]



6.3 Data acquisition and network system

Let us now describe the more widespread types of communication networks and the one we propose for this particular type of application. Considering the permanently installed networks, there are basically two types of systems: embedded sensors inside the materials, such as fiber Bragg grating sensors [127] or piezoelectric sensors, and surface mounted sensors. The first type of deployment is characterized by two major drawbacks:

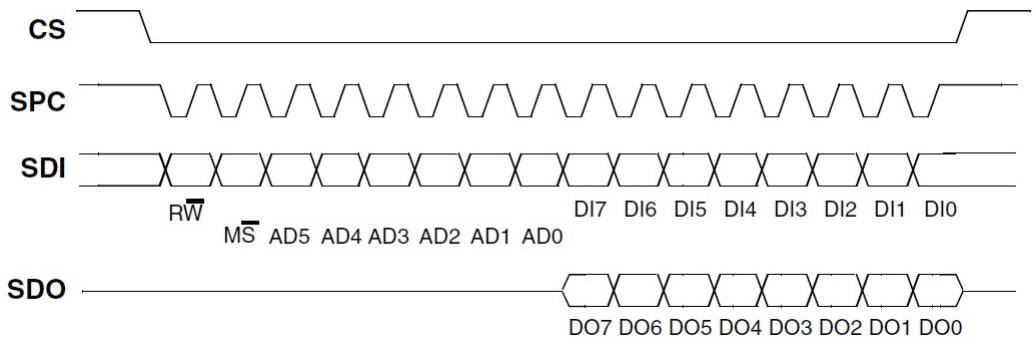
- in recent years, it has been shown in several research works [128, 129, 130, 131] that the embedded sensors can affect the behavior of the structure and, especially for composite materials, could induce unwanted cracking growth;
- the sensors can not be reached and hence replaced or checked.

Regarding the network, several wireless solutions have been presented in recent years [12]. Despite its obvious advantages in large structures due to the reducing of costs in the cable system, considering a possible application for lightweight structures, two disadvantages can be underlined: the battery life-time and the weight of the wireless receivers. For these reasons, a surface mounted wired

system is here proposed.
The sensors could be connected to the data acquisition system through either a one-to-one connection or with a unique sensor-to-sensor bus. This latter solution allows to reduce the length of cables and the complexity of the network. As previously explained, all the signals are controlled by a master, which includes both an acquisition system, that converts the digital *SPI* signal in the USB standard, and a computational unit, such as a consumer PC. We propose here the use of a commercial device:

- a 32-channel bidirectional digital input/output module (NI 9403 [132]);
- a National Instruments multipurpose chassis able to hold several modules and convert the signal to a USB standard (NIcDAQ-9178 [133]).

Figure 6.5: Read/write SPI protocol [5]



In Figure 6.5 an example of a read/write *SPI* protocol is shown. The CS signal opens and closes the byte communication while the SPC signal sets the timing at which each bit is created. The RW bit selects the read/write mode, DO(7:0) bits (corresponding to 1 byte) represents the acceleration data coming from the device towards the master, AD(5:0) and DI(7:0) bits are respectively the address of the registers and the register's values that are sent to the device in order to select the settings and query the wanted acceleration axis. Some examples of the purposes of the registers are:

- querying the identification number of the device;
- calibration of the zero acceleration level offset;
- calibration of the system sensitivity, i.e. the proportional coefficient that links the output values with the real acceleration value;
- on and off device powering;
- selection of the data rate at which the acceleration samples are created;

- selection of the actual axis x, y or z;
- selection of the acceleration range;
- choosing the bit encoding and representation settings.

In figure 6.6 the scheme of the hypothesized network for multiple accelerometers is shown.

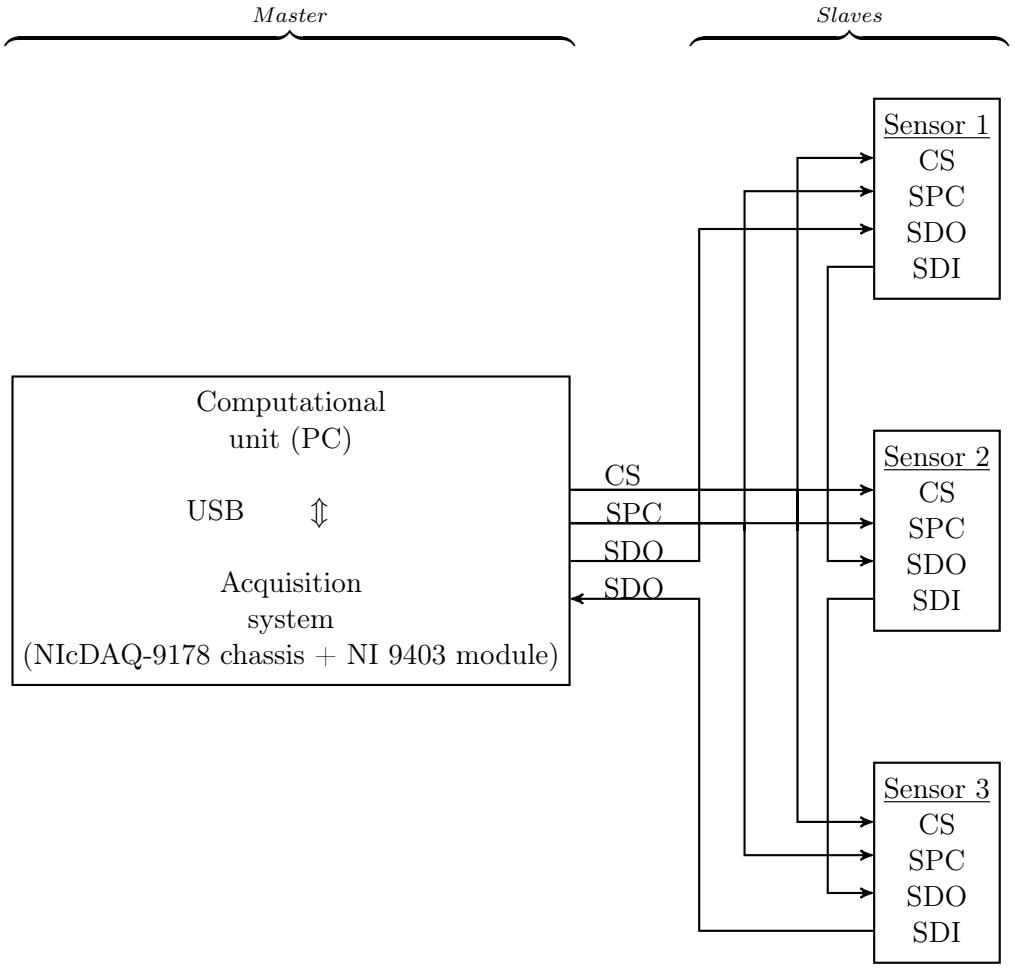


Figure 6.6: Block diagram representation of the master-slave communication scheme

Figure (6.7) shows the block diagram of the Labview program written to implement the SPI protocol and the power supply of the 3-axes accelerometer LIS3LV02DL. Each line in the program represents a wire: the lines on the left side of the image are the Vdd and VddIO power supplies lines, fed through a sub-function that creates a constant 1 value digital signal. The lines on the right represent the SPI communication: the green box at the beginning is a

sub-function that creates the programming variables, the blue boxes are the SPI read and write sub-functions, implemented following the scheme of Figure 6.5. The input given to the latter ones are basically the registers previously described: a sequence of bits are represented through an array of boolean values, where T (True) and F (False) stand respectively for 1 and 0. Each read function returns only an 8-bit byte, therefore since a single data is made of 16 bits, two read function are required for each value of acceleration. For instance, in order to read the acceleration A_x, two read functions are called: the first one is fed with its own register and give back the A_x^{LOW} byte, the second one is fed with an other different register and give back the A_x^{HIGH} byte. Finally A_x^{LOW} and A_x^{HIGH} are concatenated. The large grey box represents a loop: for each iteration, one set of data is acquired and stored in a given file, indicated by the sub-functions at the end of each line. The data acquisition and thus the iterations are stopped by a specific command, depicted as a red button.
Once the acceleration data at each step have been acquired in a binary format, the values are converted into numerical data. In order to calculate the rotations using the acceleration signals, two steps are required [134]: a calibration and a transformation.

1. Calibration: since natural misalignment arises when the sensor is not used, at the beginning of each measurement phase, a calibration procedure should be performed in order to compensate them. The misalignments are the angles between the sensor axes and the package axes and are usually related to the positioning of the silicon die with respect to package substrate. Despite the fact in ordinary applications the calibration is not necessary, in the present application, as we can see from Section 5.2.2, a high accuracy is required and the signals should be compensated. The sensor is calibrated positioning it for approximately 10 s in 6 known stationary positions and measuring the three accelerations. Using the least-square method, a calibration matrix $\mathbf{X} \in \mathbb{R}^{3\times3}$ is calculated. Therefore, supposing the accelerometers axes coincide with the structure axes, the raw non-calibrated accelerations $\bar{A}_x$, $\bar{A}_y$ and $\bar{A}_z$ are transformed in the calibrated components of the acceleration A_x, A_y and A_z according to:

$$\mathbf{A} = \mathbf{X}\bar{\mathbf{A}} \tag{6.3}$$

where

$$\mathbf{A} = \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} \text{ such that } \sqrt{A_x^2 + A_y^2 + A_z^2} = 1 \tag{6.4}$$

where the latter equation is required in order for the accelerations to be normalized.

2. Transformation: the normalized compensated accelerations $\mathbf{A}$ are used to calculate the rotations φ_x (roll) and φ_y (pitch). In order for the sensitivity

to be constant over the angles, the following equations should be used:

$$\varphi_x = \arctan\left(\frac{A_y}{\sqrt{A_x^2 + A_z^2}}\right) \tag{6.5}$$

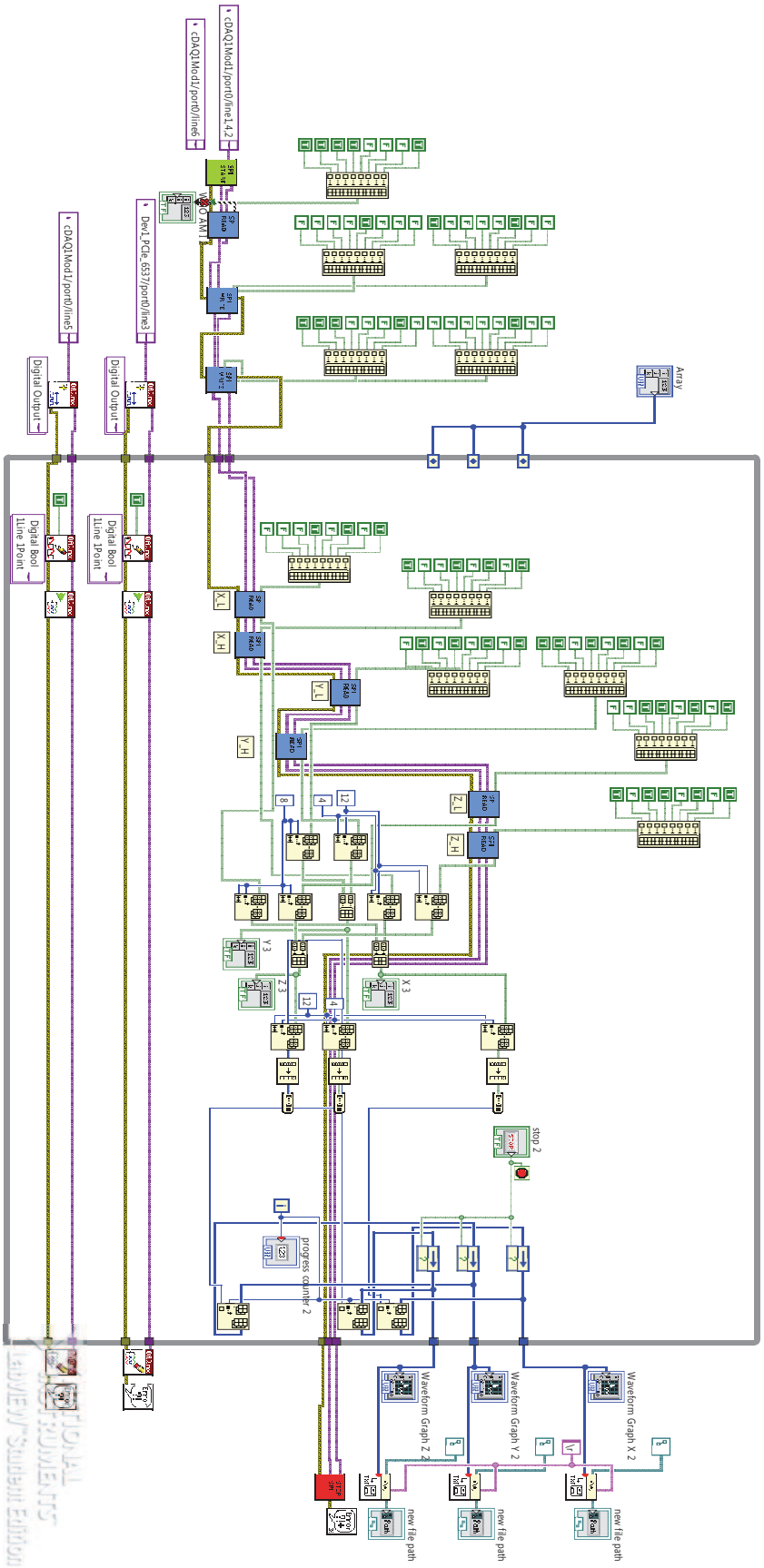
$$\varphi_y = \arctan\left(\frac{A_x}{\sqrt{A_y^2 + A_z^2}}\right) \tag{6.6}$$

6.4 Summary

In the present chapter a proposal for a structural health monitoring system has been described. This system could be used to measure the observations required by the recursive Bayesian filter described in Chapter 4.
The system is characterized by the following features:

- the use of MEMS accelerometers: thanks to their reduced weight and sizes, they can be deployed in lightweight structures without affecting the structural response; moreover, their low-cost allow to apply them even when a high number of sensors is required;
- the number of sensors deployed on the plate can be reduced by adopting the optimal placement strategy described in Section 6.2;
- the MEMS accelerometers are attached to their adapter board and connected through a wired network; a Labview programming interface is used to acquire the data through a SPI protocol.

Figure 6.7: Labview block diagram of the *SPI* protocol for the 3-axis MEMS accelerometer LIS3LV02DL



Chapter 7

Conclusions

The main goal of the present thesis has been the evaluation and application of a particular recursive Bayesian filter, proposed in [62], to the estimation of the damage parameters of a certain structure, given some observations. The basic objectives the filter had to fulfilled were:

- estimate both the dynamical evolution of the system and the damage parameters, i.e. the local stiffness;
- guarantee an affordable computational cost, such that the system identification could be performed in a nearly real-time way;
- make the method independent with respect to the specific FE formulation used to discretize the structure.

The first target has been achieved by using a dual estimation framework, where the state vector that describes the system contains both the dynamic variables, and the damage parameters defined as a function of the ratio between the local stiffnesses of respectively the damaged and undamaged structures. The second goal has been fulfilled through the application of a model order reduction performed using the POD and a Galerkin projection of the full space; moreover, a hybrid extended Kalman particle filter and a resampling technique has been adopted. At last, a particular formulation of the process model has been used: the stiffness matrix has been expressed as a linear combination of some appropriate reference stiffness matrices, somehow normalized with respect to the local stiffness parameters.
A benchmark analysis has been presented, applying the algorithm to a simply supported plate, subjected to a smooth load. The following concluding remarks can be underlined:

- the proper orthogonal decomposition is able to successfully reduce the model of either a damaged and an undamaged structure, retaining most of the information in the first POMs;

- the higher the degree of symmetry, the less POMs are necessary to reach a certain accuracy;
- the convergence of the first POMs calculated through the SVD is guaranteed;
- the accuracy of the POD approximation increases in time;
- the choice of the number of POMs retained affects the accuracy of the damage estimation, and basically it is in accordance with the results obtained for the model order reduction; nevertheless, a reduced number of POMs is sufficient to efficiently estimate the parameters.
- if the initial guess is too far from the target values, the operation of the filter could be prevented; therefore, in some cases, a preliminary estimation can be necessary;
- a high measurement noise can prevent the ability of the filter to estimate the parameters;
- a high process noise introduce a large instability in the filter;
- unless the dimension of the reduced system is very low, i.e. one or two POMs retained, the number of sensors deployed does not play a crucial role on the behavior of the filter;
- the convergence of the POMs over time is monotonic, hence the accuracy increases over time;
- the operation of the filter is guaranteed even for a system with a high number of degrees of freedom.

To summarize, it can be remarked that the stability and accuracy of the estimation can be guaranteed by an appropriate choice of the filter settings previously described. For instance, both the number of sensors and above all the size of the reduced system can be estimated in advance by comparing the full and reduced order model through a simulation. Regarding the measurement noise, some filtering techniques can be applied to the data in order to reduce it. Hence the only parameter that can not be tuned a priori is the process model; nevertheless, if the model used is accurate, this term is very low.
Having said that, we can conclude that the dual estimation of the reduced state and the damage parameters of a given structure can successfully fulfill the objectives described at the beginning.

Nevertheless, further research and investigation are of course required. Some ideas for future investigations are:

- Despite the fact the method could be conceptually applied to any type of structure, some possible further research topics could regard the assessment of the algorithm when different characteristics of the structure are considered, such as:
  - three dimensional structures;
  - boundary conditions;
  - loading conditions (impacts);
  - non-linear material constitutive relations;
  - number of damage parameters to be estimated.
- In the present work an embedded data acquisition system based on MEMS accelerometers has been proposed. Therefore, this structural health monitoring system could be used to validate the algorithm, using real measurements in an experimental set.
- An enhanced ad-hoc structural health monitoring for specific real applications could be developed.

Bibliography

[1] D. Balageas, C. P. Fritzen, and A. Güemes. *Structural Health Monitoring*. Wiley-ISTE, 2006.

[2] M. Rathinam and L. R. Petzold. A new look at proper orthogonal decomposition. *SIAM Journal on Numerical Analysis*, 41:1893–1925, 2003.

[3] S. Haykin. *Kalman Filtering and Neutral Networks*. John Wiley & Sons, 2001.

[4] STMicroelectronics. LIS3LV02DL adapter board Datasheet. 2007.

[5] STMicroelectronics. LIS3LV02DL Datasheet. 2008.

[6] M. Bruggi and S. Mariani. Optimization of sensor placement to detect damage in flexible plates. *Engineering Optimization*, 45:659–676, 2012.

[7] C. R Farrar and K. Worden. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A*, 365:303–315, 2007.

[8] D. H. Allen. *Damage and Interfacial Debonding in Composites*. Elsevier Science, 1996.

[9] K. Y. Wong, C. K. Lau, and A. R. Flint. Planning and implementation of the structural health monitoring system for cable-supported bridges in hong kong. *Proceedings of SPIE*, 3995, 2000.

[10] Federal Highway Administration U.S. Department of Transportation. Highway bridge inspection: State-of-the-practice survey. 2001.

[11] National Concrete Pavement Technology Center. A feasibility study on embedded micro-electromechanical sensors and systems (mems) for monitoring highway structures. *Iowa State University Report*, 2011.

[12] S. Cho, C. B. Yun, J. P. Lynch, A. T. Zimmerman, B. F. Spencer, and T. Nagayama. Smart wireless sensor technology for structural health monitoring of civil structures. *Steel Structures*, 8:267–275, 2008.

[13] H. Sohn, C. R. Farrar, F. Hemez, and J. Czarnecki. A review of structural health monitoring literature 1996 Ű 2001. *Los Alamos National Laboratories*, 2004.

[14] O. S. Salawu. Detection of structural damage through changes in frequency: a review. *Engineering Structures*, 19:718–723, 1997.

[15] C. R. Farrar and S. W. Doebling. An overview of modal-based damage identification methods. *Structural Damage Assessment Using Advanced Signal Processing Procedures*, 1997.

[16] Z. Zhang and A. E. Aktan. Application of modal flexibility and its derivatives in structural identification. *Research in Nondestructive Evaluation*, 10:43–61, 1998.

[17] C. R. Farrar and D. A. Jauregui. Comparative study of damage identi?cation algorithms applied to a bridge: II. Numerical study. *Smart Materials and Structures*, 7:720–721, 1998.

[18] M. I. Friswell. Damage identication using inverse methods. *Philosophical Transactions of the Royal Society A*, 365:393–410, 2007.

[19] E. N. Chatzi, A. W. Smytha, and S. F. Masri. Experimental application of on-line parametric identication for nonlinear hysteretic systems with model uncertainty. *Structural Safety*, 32:326–337, 2010.

[20] J. N. Yang, S. Lin, H. Huang, and L. Zhou. An adaptive extended kalman filter for structural damage identification. *Structural Control Health Monitoring*, 13:849–867, 2006.

[21] J. L. Beck and S. K. Au. Bayesian updating of structural models and reliability using markov chain monte carlo simulation. *Journal of Engineering Mechanics*, 128:380–391, 2002.

[22] J. L. Beck and K. V. Yuen. Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, 130:192–203, 2004.

[23] H. Sohn and K. H. Law. A bayesian probabilistic approach for structure damage. *Earthquake Engineering Structural Dynamics*, 26:1259–1281, 1998.

[24] M. W. Vanik, J. L. Beck, and S. K. Au. Bayesian probabilistic approach to structural health monitoring. *Journal of Engineering Mechanics*, 126:738–745, 2000.

[25] S. Eftekhar Azam. *Online Damage Detection in Structural Systems*. Springer, 2014.

[26] A. N. Krylov. On the numerical solution of equation by which are determined in technical problems the frequencies of small vibrations of material systems. *News of Academy of Sciences of the USSR*, VII, Nr.4:491–539, 1931.

[27] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139:153–176, 2006.

[28] A. Ammar. The proper generalized decomposition: a powerful tool for model reduction. *International Journal of Material Forming*, 3:89–102, 2010.

[29] K. F. Alvin, A. N. Robertson, G. W. Reich, and K. C. Park. Structural system identification: from reality to models. *Computer and Structures*, 81:1149–1176, 2003.

[30] A. Megretski. Lecture notes of the Model Order Reduction course. Massachussets Institute of Technology, Fall 2004.

[31] D. D. Kosambi. Statistics in function space. *Journal of the Indian Mathemathical Society*, 7:76–88, 1948.

[32] M. Loève. *Asymptotical study of dependent random variables*. PhD thesis, Université de Paris, 1941.

[33] K. Karhunen. Über lineare methoden in der wahrscheinlichkeitsrechnung. *Annales Academiae Scientiarum Fennicae Mathematica*, 37:1–79, 1947.

[34] V. S. Pougachev. General theory of the correlations of random functions. *Izvestiya Akademii Nauk USSR*, 17:1401–1402, 1953.

[35] M. A. Obukhov. Statistical desctiption of continuous fields. *Transactions of the Geophysical International Academy Nauk USSR*, 24:3–42, 1954.

[36] I. T. Jolliffe. *Principal Component Analysis*. Springer, 1986.

[37] A. I. Mees, P. E. Rapp, and L. S. Jennings. Singular-value decomposition and embedding dimension. *Physical Review A*, 1;36(1):340–346, 1987.

[38] M. Rathinam and L. R. Petzold. A new look at proper orthogonal decomposition. *SIAM Journal on Numerical Analysis*, 41(5):1893–1925, 2003.

[39] M. Rathinam and L. R. Petzold. Dynamic iteration usine reduced order models: a method for simulation of large scale modular systems. *SIAM Journal on Numerical Analysis*, 40(4):1446–1474, 2002.

[40] V. R. Algazi and D. J. Sakrison. On the optimality of the karhunen-lo'eve expansion. *IEEE Transactions on Information Theory*, IT-15:319–320, 1969.

[41] D. Amsallem and C. Farhat. Lecture notes of the Model Reduction course (CME 345). Stanford University, Spring 2011.

[42] W. H. Schilders, H. A. Van der Vorst, and J. Rommes. *Model Order Reduction. Theory, Research Aspects and Applications*. Springer, 2008.

[43] L. Sirovich. Turbulence and the dynamics of coherent structures. I - coherent structures. II - symmetries and transformations. III - dynamics and scaling. *Quarterly of Applied Mathematics*, 45:561–590, 1987.

[44] Y. C. Liang, H. P. Lee, S. P. Lim, W. Z. Lin, K. H. Lee, and C. G. Wu. Proper orthogonal decomposition and its applications, part i: Theory. *Jounal of Sound and Vibration*, 252:527–544, 2002.

[45] C. G. Wu, Y. C. Liang, W. Z. Lin, H. P. Lee, and S. P. Lim. A note on equivalence of proper orthogonal decomposition methods. *Journal of Sound and Vibration*, 265:1103–1110, 2003.

[46] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

[47] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6):417–441, 1933.

[48] G. Kerschen, J. C. Golinval, A. F. Vakakis, and L. A. Bergman. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: An overview. *Nonlinear Dynamics*, 41:147–169, 2005.

[49] D. Otte. *Development and Evaluation of Singular Value Analysis Methodologies for Studying Multivariate Noise and Vibration Problems*. PhD thesis, Katholieke Universiteit Leuven, 1994.

[50] K. Afanasiev and M. Hinze. Adaptive control of a wake flow using proper orthogonal decomposition. *Lecture Notes in Pure and Applied Mathemathics*, 216:317–332, 2001.

[51] B. F. Feeny and R. Kappagantu. On the physical interpretation of proper orthogonal modes in vibrations. *Journal of Sound and Vibration*, 211:607–616, 1998.

[52] G. Kerschen and J. C. Golinval. Physical interpretation of the proper orthogonal modes using the singular value decomposition. *Journal of Sound and Vibration*, 249:849–865, 2002.

[53] M. Geradin and D. Rixen. *Mechanical Vibrations, Theory and Application to Structural Dynamics*. Masson, 1994.

[54] B. F. Feeny and Y. Liang. Interpreting proper orthogonal modes of randomly excited vibration systems. *Journal of Sound and Vibration*, 265:953–966, 2003.

[55] S. Kotz and S. Nadarajah. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, London, 2000.

[56] R. Smith. *Extremes and Integrated Risk Management*, chapter Bayesian Risk Analysis, pages 235–252. Risk Publications, 2000.

[57] G. Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42:393–405, 1990.

[58] Z. Yang and B. Rannala. Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method. *Molecular Biology and Evolution*, 14:717–724, 1997.

[59] Deborah Ashby. Bayesian statistics in medicine: a 25 year review. *Statistics in Medicine*, 25:3589–3631, 2006.

[60] M. Sahami, S. Dumais, D. Heckerman, and E. Horvitz. A bayesian approach to filtering junk e-mailaaai'98 workshop on learning for text categorization. *AAAI Workshop on Learning for Text Categorization*, 1998.

[61] J. O. Berger. *Statistical decision theory and Bayesian Analysis*. Springer-Verlag, 1985.

[62] S. Eftekhar Azam. *Dual estimation and reduced order modelling of damaging structures*. PhD thesis, Politecnico di Milano, 2012.

[63] S. Mariani and A. Corigliano. Impact induced composite delamination: State and parameter identification via joint and dual extended kalman filters. *Computer Methods in Applied Mechanics and Engineering*, 194:5242–5272, 2005.

[64] A. Corigliano and S. Mariani. Parameter identification in explicit structural dynamics: performance of the extended kalman filter. *Computer Methods in Applied Mechanics and Engineering*, 193:3807–3835, 2004.

[65] S. Eftekhar Azam, M. Bagherinia, and S. Mariani. Stochastic system identification via particle and sigma-point kalman filtering. *Scientia Iranica*, 19:982–991, 2012.

[66] R. Pettai. *Noise in Receiving Systems*. Wiley-Interscience, 1984.

[67] H. W. Ott. *Noise Reduction Techniques in Electronic Systems*. Wiley-Interscience, 1976.

[68] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society*, 53:370–418, 1763.

[69] A. A. Markov. *Theory of Algorithms*. Imprint Moscow, Academy of Sciences of the USSR, 1954.

[70] S. Chapman. On the Brownian displacements and thermal diffusion of grains suspended in non-uniform fluid. *Proceedings of the Royal Society A*, 119:34–54, 1928.

[71] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50:174–188, 2002.

[72] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, pages 35–45, 1960.

[73] R. E. Kalman. Contributions to the theory of optimal control. *Boletin Sociedad Matematica Mexicana*, pages 102–119, 1960.

[74] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, pages 95–108, 1961.

[75] Y. C. Ho and R. C. K. Lee. A bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, AC-9:333–339, 1964.

[76] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. John Wiley & Sons, 2001.

[77] F. H. Lewis. *Optical Estimation with an Introduction to Stochastic Control Theory*. John Wiley & Sons, 1986.

[78] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice*. Prentice-Hall, 1993.

[79] H. L. Van Tress. *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 1968.

[80] G. L. Smith, S. F. Schmidt, and L. A. McGee. Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle. *National Aeronautics and Space Administration*, 1962.

[81] R. van der Merwe. *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models.* PhD thesis, Oregon Health & Science University, 2004.

[82] E. Wan, R. van der Merwe, and A. Nelson. Dual estimation and the unscented transformation. In *Neural Information Processing Systems 12*, 2000.

[83] S. Mariani. *Failure of layered composites subject to impacts: constitutive modeling and parameter identification issues*, chapter 3, pages 97–131. Nova Science Publishers, 2009.

[84] T. Lefebvre, H. Bruyninckx, and J. De Schutter. Kalman filters for nonlinear systems: a comparison of performance. *IEEE Transactions on Automatic Control*, 2001.

[85] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45:477–482, 2000.

[86] E. A. Wan and R. van der Merwe. The unscented kalman filter for nonlinear estimation. *Proceedings of IEEE Symposium on Adaptive Systems for Signal Processing Communications and Control (AS-SPCC)*, pages 153–158, 2000.

[87] S. J. Julier and J. K. Uhlmann. New extension of the kalman filter to nonlinear systems. In *Signal Processing, Sensor Fusion and Target Recognition VI*, 1997.

[88] S. Eftekhar Azam, A. Ghisi, and S. Mariani. Parallelized sigma-point kalman filtering for structural dynamics. *Computers and Structures*, 92:193–205, 2011.

[89] S. Mariani and A. Ghisi. Unscented kalman filtering for nonlinear structural dynamics. *Nonlinear Dynamics*, 49:131–150, 2007.

[90] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings - Part F: Radar and Signal Processing*, 140:107–113, 1993.

[91] J. Carpenter, P. Clifford, and P. Fearnhead. Improved particle filter for nonlinear problems. *Proceedings of the IEEE - Radar, Sonar, Navigation*, 1999.

[92] J. MacCormick and A. Blake. A probabilistic exclusion principle for tracking multiple objects. *Proceedings of the International Conference on Computer Vision*, pages 572–578, 1999.

[93] P. Del Moral. Non-linear filtering: Interacting particle solution. *Markov Processes Related Fields*, 2:555–580.

[94] K. Kanazawa, D. Koller, and S. J. Russell. Stochastic simulation algorithms for dynamic probabilistic networks. *Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence*, pages 346–351, 1995.

[95] D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50:736–746, 2002.

[96] A. Doucet. On sequential monte carlo methods for bayesian filtering. *Statistics and Computing*, 10:197–208, 2000.

[97] M. H. Kalos and P. A. Whitlock. *Monte Carlo methods. Volume I: Basics.* John Wiley & Sons, 1986.

[98] N. Bergman. *Recursive Bayesian estimation: Navigation and tracking applications.* PhD thesis, Linköpings Universitet, 1999.

[99] A. Doucet, J. F. G de Freitas, and N. J. Gordon. *Sequential Monte Carlo in practice*, chapter An introduction to sequential Monte Carlo methods. Springer-Verlag, 2001.

[100] A. Kong, J. S. Liu, and W. H. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89:278–288, 1994.

[101] J. S. Liu and R. Chen. Sequential monte carlo methods for dynamical systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.

[102] F. Cadini, E. Zio, and D. Avram. Monte carlo-based filtering for fatigue crack growth estimation. *Probabilistic Engineering Mechanics*, 24:367–373, 2009.

[103] G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5:1–25, 1996.

[104] S. Park, J. P. Hwang, E. Kim, and H. J. Kang. A new evolutionary particle filter for the prevention of sample impoverishment. *IEEE Transactions on Evolutionary Computation*, 13:801–809, 2009.

[105] K. Uosaki, Y. Kimura, and T. Hatanaka. Nonlinear state estimation by evolution strategies based particle filters. *IEEE Congress on Evolutionary Computation*, 1:884–890, 2004.

[106] S. Eftekhar Azam and S. Mariani. Dual estimation of partially observed nonlinear structural systems: A particle filter approach. *Mechanics Research Communications*, 46:54–61, 2012.

[107] N. M. Newmark. A method of computation for structural dynamics. *Journal of Engineering Mechanics, ASCE*, 85:67–94, 1959.

[108] Inc. The Aluminum Association. Aluminum standards and data. 2000.

[109] Dessault Systemes. *Abaqus/CAE 6-10.1 User's Manual.* 2010.

[110] S. Mariani, M. Bruggi, F. Caimmi, P. Bendiscioli, and M. De Fazio. Sensor deployment over damage-containing plates: a topology optimization approach. *Journal of Intelligent Material Systems and Structures*, 24:1105–1122, 2013.

[111] K. J. Bathe. *Finite Element Procedures.* Prentice Hall, 1996.

[112] S. Eftekhar Azam and S. Mariani. Investigation of computational and accuracy issues in pod-based reduced order modeling of dynamic structural systems. *Engineering Structures*, 54:150–167, 2013.

[113] S. Mariani, M. Bruggi, F. Caimmi, and P. Bendiscioli. Optimal placement of mems sensors for damage detection in flexible plates. *Structural Longevity.*

[114] R. D. Cook, D. S. Malkus, M. E. Plesha, and R. J. Witt. *Concepts and Applications of Finite Element Analysis.* John Wiley & Sons, 2001.

[115] S. Delepine-Lesoille, E. Merliot, C. Boulay, L. Quetel, M. Delaveau, and A. Courteville. Quasi-distributed optical fibre extensometers for continuous embedding into concrete: design and realization. *Smart Materials and Structures*, 15:931Ű938, 2006.

[116] B. Benmokrane, H. Rahman, and P. Mukhopadhyaya. Use of fibre reinforced polymer reinforcement integrated with fibre optic sensors for concrete bridge deck slab construction. *Canadian Journal of Civil Engineering*, 27:928–940, 2000.

[117] P. Rossi and F. Le Maou. New method for detecting cracks in concrete using fibre optics. *Materials and Structures*, 22:437–442, 1989.

[118] N. Elvin, C. K. Y. Leung, V. S. Sudarshanam, and S. Ezekiel. Novel fiber optic delamination detection scheme: theoretical and experimental feasibility studies. *Journal of Intelligent Material Systems and Structures*, 10:314–321, 2000.

[119] J. W. Ayres, F. Lalande, Z. Chaudhry, and C. A. Rogers. Qualitative impedance-based health monitoring of civil infrastructures. *Smart Materials and Structures*, 7:599–605, 1998.

[120] F. Wu and F. K. Chang. A built-in active sensing diagnostic system for civil infrastructure systems. *Smart Systems for Bridges, Structures, and Highways-Smart Structures and Materials*, 4330:27–35, 2001.

[121] H. Kwun and K. A. Bartels. Magnetostrictive sensor technology and its applications. *Ultrasonics*, 36:171–178, 1998.

[122] M. Celebi and A. Sanli. Real-time seismic monitoring needs of a building owner - and the solution: A cooperative effort. *Earthquake Spectra*, 20:333–346, 2004.

[123] B. Peeters and H. Van Der Auweraer. Operational modal analysis for estimating the dynamic properties of a stadium structure during a football game. *Shock and Vibration*, 14:283–303, 2007.

[124] STMicroelectronics. EK3LV02DL Evaluation Kit Datasheet. 2006.

[125] D. C. Kammer. Sensor placement for on-orbit modal identification and correlation of large space structures. *Journal of Guidance, Control and Dynamics*, 14:251–259, 1991.

[126] L. Canxing, L. Xingshan, Z. Ping, and D. Jing. A review on optimal sensor placement for health monitoring. *Electronic Measurement and Instruments*, 4:170–173, 2007.

[127] G. Zhou and L. M. Sim. Damage detection and assessment in fiber-reinforced composite structures with embedded fiber optic sensors-review. *Smart Materials and Structures*, 11:925–939, 2002.

[128] S. Butler, M. Gurvich, A. Ghoshal, G. Welsh, P. Attridge, H. Winston, M. Urban, and N. Bordick. Effect of embedded sensors on interlaminar damage in composite structures. *Journal of Intelligent Material Systems and Structures*, 22:1857–1868, 2011.

[129] V. Giurgioutiu. Tuned lamb wave excitation and detection with piezoelectric wafer active sensors for structural health monitoring. *Journal of Intelligent Material Systems and Structures*, 16:291–305, 2005.

[130] F. Ghezzo and S. Nemat-Nasser. Effects of embedded shm sensors on the structural integrity of glass fiber/epoxy laminates under in-plane loads. *Proceedings of SPIE*, 6530, 2007.

[131] S. Nemat-Nasser Y. Huang. Structural integrity of composite laminates with embedded micro-sensors. *Proceedings of SPIE*, 6530, 2007.

[132] National Instruments. NI 9403 Datasheet. 2008.

[133] National Instruments. NI cDAQ-9178 Datasheet. 2013.

[134] STMicroelectronics. Tilt measurement using a low-g 3-axis accelerometer. 2010.

[135] W. Fan and P. Qiao. Vibration-based damage identification methods: A review and comparative study. *Structural Health Monitoring*, 10:83–110, 2011.

[136] P. Del Moral. Measure valued processes and interacting particle systems. pplication to nonlinear filtering problems. *Annals of Applied Probability*, 8:438–495, 1998.