IDENTIFICATION OF NONLINEARITIES IN STRUCTURAL DYNAMICS BY
USING ARTIFICIAL NEURAL NETWORKS AND OPTIMIZATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ANIL KOYUNCU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MECHANICAL ENGINEERING

SEPTEMBER 2013

Approval of the thesis:

IDENTIFICATION OF NONLINEARITIES IN STRUCTURAL DYNAMICS BY USING ARTIFICIAL NEURAL NETWORKS AND OPTIMIZATION

submitted by **ANIL KOYUNCU** in partial fulfillment of requirements for the degree of **Master of Science in Mechanical Engineering Department**, **Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Süha Oral
Head of Department, **Mechanical Engineering** _____

Prof. Dr. H. Nevzat Özgüven
Supervisor, **Mechanical Engineering Dept., METU** _____

Assist. Prof. Dr. Ender CİĞEROĞLU
Co-Supervisor, **Mechanical Engineering Dept., METU** _____

Examining Committee Members

Prof. Dr. Mehmet ÇALIŞKAN
Mechanical Engineering Dept., METU _____

Prof. Dr. H. Nevzat Özgüven
Mechanical Engineering Dept., METU _____

Assist. Prof. Dr. Ender CİĞEROĞLU
Mechanical Engineering Dept., METU _____

Assist. Prof. Dr. Gökhan O. ÖZGEN
Mechanical Engineering Dept., METU _____

Assist. Prof. Dr. M. Bülent ÖZER
Mechanical Engineering Dept., TOBB ETU _____

Date: 02.09.2013

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name :   Anıl KOYUNCU

Signature

ABSTRACT

IDENTIFICATION OF NONLINEARITIES IN STRUCTURAL DYNAMICS BY USING ARTIFICIAL NEURAL NETWORKS AND OPTIMIZATION

KOYUNCU, Anıl
M.S. Department of Mechanical Engineering
Supervisor: Prof. Dr. H. Nevzat ÖZGÜVEN
Co-Supervisor: Assist. Dr. Ender CİĞEROĞLU

September 2013, 92 pages

Many real life engineering structures exhibit nonlinear behavior in practice. Although there are sophisticated methods including the effect of nonlinearities on dynamic response of structures in literature, uncertainties about nonlinear elements make further investigation necessary for modeling nonlinearity, and this is usually achieved by using experimental data taken from real systems. Therefore, identification of nonlinearities -determining location, type and parameters of the nonlinear elements- is critical in dynamical structures. In this study, a new approach is proposed for identification of structural nonlinearities by employing neural networks. Linear finite element model of the system and frequency response functions measured at arbitrary locations of the system are used in this approach. Using the finite element model, a training data set is created, which appropriately spans the possible nonlinear configurations space of the system. A classification neural network trained on these data sets then localizes and determines the type of nonlinearity associated with the corresponding degree of freedom in the system. A new training data set spanning the parametric space associated with the determined nonlinearities is created to facilitate parametric identification. Utilizing this data set, a feed forward regression neural network is trained, which parametrically identifies the related nonlinearity. The proposed approach does not require data collection from the degrees of freedoms related with nonlinear elements, and furthermore, the proposed approach is sufficiently accurate even in the presence of measurement noise. Identified parameters are improved utilizing optimization. The application of the proposed approach is demonstrated on an example system with nonlinear elements and a real life experimental setup with a local nonlinearity.

Keywords: Neural networks, Nonlinearity identification, Nonlinearity classification, Nonlinear vibrations, Harmonic balance method

ÖZ

YAPISAL DİNAMİKLERDEKİ DOĞRUSAL OLMAYAN ÖZELLİKLERİN YAPAY SİNİR AĞLARI VE OPTİMİZASYON İLE BELİRLENMESİ

KOYUNCU, Anıl
Yüksek Lisans, Makina Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. H.Nevzat Özgüven
Tez Eş Danışmanı: Y. Doç. Dr. Ender CİĞEROĞLU

Eylül 2013, 92 Sayfa

Gerçek hayattaki birçok mekanik yapı uygulamada doğrusal olmayan özellikler göstermektedir. Literatürde doğrusal olmayan eleman etkilerinin sistemlere eklenmesi ile ilgili başarılı yöntemler bulunmasına rağmen, doğrusal olmayan özelliklerin modellenmesi için bu elemanlarla ilgili belirsizlikler ileri araştırmaları gerekli kılmaktadır ve bu da genellikle gerçek sistem üzerinden alınan ölçümlerle başarılmaktadır. Bu yüzden dinamik sistemlerdeki doğrusal olmayan elemanların belirlenmesi- yerlerinin, tiplerinin ve parametrelerinin - sistemdeki doğrusal olmayan özelliklerin tanımlanması açısından önemlidir. Bu çalışmada yapay sinir ağları kullanılarak sistemlerde doğrusal olmayan özelliklerin tanımlandığı yeni bir yaklaşım öne sürülmektedir. Bu yaklaşımda sistemlerin doğrusal sonlu eleman modelleri ve rastgele noktalardan ölçülen Frekans Tepkisi Fonksiyonları (FTF) kullanılmıştır. Doğrusal sonlu eleman modeli kullanılarak hesaplanan doğrusal olmayan muhtemel sistem veri kümeleri öğrenme işlemleri için elde edilir. Bu veri kümeleri sınıflandırma sinir ağlarına öğretildikten sonra sistemdeki doğrusal olmayan elemanları bölgeler ve tipleri ilgili serbestlik noktasıyla birlikte bulunmaktadır. Tespit edilen doğrusal olmayan elemanlara ait parametrelerin bulunması için yeni bir veri kümesi yaratılır. Yaratılan veri kümesiyle, ilgili doğrusal olmayan elemana ait parametrelerin bulmak için ileri besleme regresyon sinir ağları öğretilir. Öne sürülen yöntem doğrusal olmayan elemanların bulunduğu noktalardan ölçüm alınmasını gerektirmediği gibi ölçümden kaynaklı gürültülerin olması durumunda dahi yeterince doğru sonuçlar alınabilmektedir. Elde edilen parametreler optimizasyon kullanılarak iyileştirilebilir. Önerilen yöntemin uygulanması örnek sayısal ve deneysel sistemler üzerinde gösterilmiştir.

Anahtar Kelimeler: Yapay Sinir Ağları, Doğrusal Olmayan Eleman Tanımlama, Doğrusal Olmayan Eleman Sınıflandırma, Doğrusal Olmayan Titreşimler, Harmonic Denge Metodu

To My Family

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$\{b\}$ Bias vector

c Damping

$[C]$ Damping matrix

$[D]$ Dynamic stiffness matrix

$\{e\}$ Error vector

E Modulus of elasticity

$\{f\}$ Harmonic forcing vector

$\{F\}$ Amplitude vector of harmonic forcing

$\{f_{NL}\}$ Nonlinear restoring forcing vector

$\{F_{NL}\}$ Amplitude of nonlinear restoring forcing vector

$\{G\}$ Modified nonlinear residual vector

$[H]$ Structural damping matrix

$[J]$ Jacobian matrix

k Stiffness

$[K]$ Stiffness matrix

m Mass

$[M]$ Mass matrix

$\{n\}$ Net input vector

N_R Number of inputs

N_S Number of hidden neurons

$\{o\}$ Output vector

$\{q\}$	Unknown vector of arc-length equation
$\{R\}$	Nonlinear residual vector
s	Arc-length radius
t	Time
$\{t\}$	Target vector
$\{w\}$	Weight vector
$[W]$	Weight matrix
x	Response
X	Complex amplitude of response
$\{y\}$	Input vector
$[\alpha]$	Receptance matrix
ζ	Scale parameter of arc-length step size
η	Structural damping coefficient
λ	Relaxation parameter
μ	Coefficient of friction
ν	Poisson's ratio
ρ	Density
ψ	Angle of harmonic oscillation
ω	Frequency

Subscripts

h	Harmonic number
i	Iteration number
l	Linear coordinate
n	Nonlinear coordinate
NL	Nonlinear

Superscripts

Im	Imaginary part
k	Index of solution point
l	Linear part
n	Nonlinear part
Re	Real part

Abbreviations

DOF	Degree of freedom
FEM	Finite element model
FFT	Fast Fourier Transform
FRF	Frequency response function
HBM	Harmonic Balance Method
LM	Levenberg-Marquardt
MDOF	Multi degree of
MSE	Mean Squared Error
SCGB	Scaled Conjugate Gradient Backpropagation
SDOF	Single degree of

CHAPTER 1

INTRODUCTION

1.1 Nonlinear System Identifications

Determining dynamical characteristic of structures which are subjected to mechanical vibrations is demanded to improve design process for the industry. Mathematical modeling of dynamic structures to predict its performance increases in importance with the aid of developing technology. Despite today's computational power which can handle sophisticated mathematical models, uncertainties about structures make system identification necessary to obtain a correct model and correct prediction for dynamic response.

Linear system identification has been widely used in mechanical vibrations for many years. Identification techniques based on modal testing are quite simple and reliable tools for linear systems. However, linear system assumptions are not valid in the presence of nonlinearities in the structures.

Most of the engineering structures are exposes to nonlinear effects in real life. Typical nonlinear behaviors are mainly originated from geometric nonlinearities, material nonlinearities, inertial nonlinearities, damping nonlinearities and nonlinear boundary conditions [1, 2]. Effect of nonlinearities is modeled by coupling appropriate nonlinear elements to mathematical model of underlying linear system. Restoring force diagrams of commonly used nonlinear elements are given in Figure 1.1.

Nonlinear system identification becomes crucial for nonlinear system modeling since nonlinearities appear in structures unintentionally most of the time. Purpose of the nonlinear system identification is summarized in four main headings.

- Detection
- Localization
- Type Determination
- Parameter Estimation
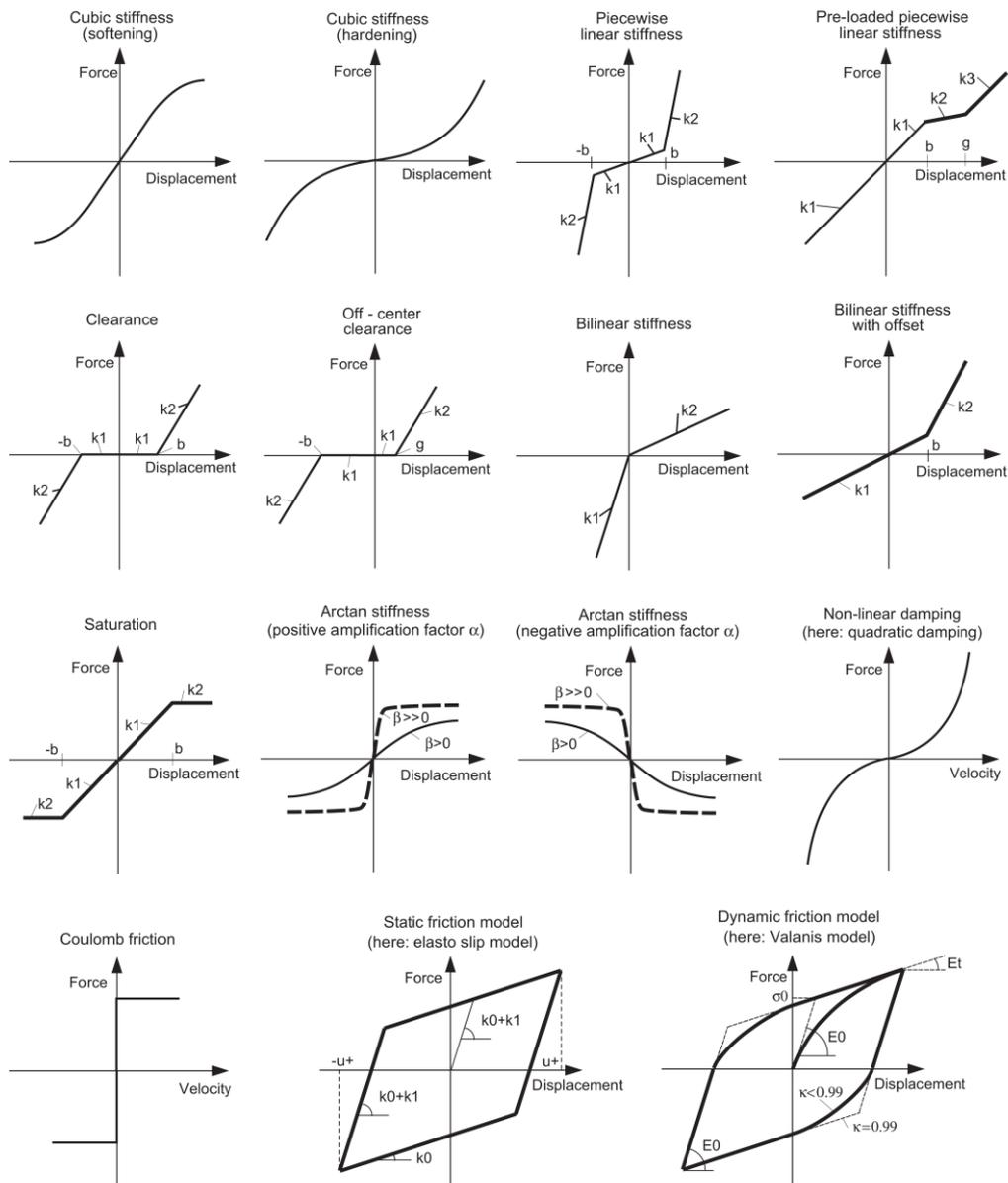
Figure 1.1.   Commonly used nonlinear elements [3]

1.2   Scope of Thesis

In this thesis study, it is aimed to develop a new method for nonlinearity identification in structural dynamics. The primary objective is to achieve complete nonlinear system identification by using measurements taken from arbitrary locations on a system.

Main topics are summarized as follows:

- to propose a method for nonlinearity localization and type identification using neural networks,
- to propose a method for parameter estimation of nonlinearities using neural networks,
- to improve the identification results using optimization algorithms.

The outlines of the chapters are given below.

1.3   Outline of Thesis

Chapter 2 gives comprehensive background information about nonlinear system identification and a brief literature review is presented in this chapter.

Chapter 3 reviews the nonlinear vibration analysis methods under harmonic excitation. Multi harmonic Harmonic Balance Method (HBM) is re-formulated in this chapter. Several numerical methods are introduced to solve nonlinear system equations. Fixed point iteration, Newton's method and Newton's method with arc length continuation are described in detail. A decoupling technique is represented in this chapter to reduce the number of nonlinear equations needed to be solved. Finally, performances of the numerical solution techniques are compared on simple case studies for single harmonic and multi harmonic solutions.

Chapter 4 presents theory of artificial neural networks. Basic elements of a neural network are introduced in this chapter. Training data generation and training operations are described in detail. Network configurations and training algorithms are suggested for classification and regression networks.

Chapter 5 proposes a new method for nonlinear system identification. General methodology of the method proposed is outlined in this chapter. Application of the method is demonstrated on a simple system. Identified system has unknown nonlinearities at unknown locations. Measurements are taken from an arbitrary location on the system. At the first state, locations and type of the nonlinear elements are determined by means of classification network. Then, parameters of the classified nonlinearities are identified using regression network. At the last step, identified parameters are optimized via using genetic algorithm. In this case study, identification is achieved successfully based on both noisy and noise free measurements.

Chapter 6 presents the application of the method proposed in Chapter 5 on a real life structure. Linear part of the system is modeled as a MDOF system. In this experimental case study, identification of the nonlinearity in the system and model

updating of the linear part are both performed. Accuracy of the proposed method is demonstrated on experimental measurements.

Chapter 7 discusses the results obtained from previous chapters. The conclusion of thesis is given in this chapter. Finally, contributions to nonlinear system identification are summarized.

Chapter 8 gives recommendations for future studies on this subject.

CHAPTER 2

LITERATURE REVIEW AND SUMMARY OF THESIS

Identification of structural nonlinearities in dynamic structures has become the interest of researchers in the past four decades [1]. Studies on this subject focused on two parts: localizing and characterizing the nonlinearity, and estimating the parameters of the nonlinearity based on experimentally measured data [2–6].

2.1 Nonlinearity Localization and Type Determination

The most challenging part of nonlinearity identification is determination of location and type of nonlinearity; in other words, classification of the nonlinear system from system responses. Classification of nonlinearities in a system is a phenomenon hard to achieve by observation most of the time.

Masri and Caughey [7] developed a method named as *Restoring Force Surface* (RFS) method to estimate nonlinear restoring forces using least squares Chebychev polynomial approach in time domain.

Göge *et al.* [3] demonstrated this behavior in their study. Nonlinear characteristics of an aircraft prototype are investigated in the study by observing change of resonant frequency and displacement amplitude with respect to excitation energy. Measured data is showed on the plots named as *"linearity plots"*. Analytical linearity plots of common nonlinear restoring forces are obtained for a SDOF system using frequency domain and time domain solutions. The main idea behind the proposed method is to detect and characterize nonlinearities in a system via looking at analytical linearity plots. However, it is revealed that some nonlinearities behave like a linear one and could not be observed in linearity plots. Although linearity plots give some idea about type information for specific nonlinearities, it could not help distinguishing some nonlinearities especially the ones exist in a system as a combination.

Nonlinearity classification methods are separated into two main groups. They are frequency domain methods and time domain methods. For nonlinearity classification problems, frequency domain methods are mostly preferred. He and Ewins [8] used frequency response functions (FRFs) obtained at different forcing levels in order to detect nonlinearities in a system. Hilbert transform is also a frequency domain method for nonlinearity detection well documented in the literature [9].

5

The method developed by Özer and Özgüven [6] determines possible locations of nonlinearities and identifies their types and parameters by using describing functions, which is as well a frequency domain method. This method requires measurements collected from all degrees of freedoms (DOFs) and complete FRFs of the linear part of the system for localization purposes. Linear system response is needed for localization and type identification. Authors suggest using nonlinear system responses at low forcing level. However, nonlinearities such as dry friction are dominant at low forcing level. Moreover, type identification works, provided that nonlinearities are localized between two coordinates in the system which can be thought as a major limitation.

Later, Arslan *et al.* [10] developed a method based on the method given in [6] for type and parametric identification of localized nonlinearities. The proposed method determines model parameters of the nonlinear system from displacement controlled FRF measurements. Describing functions of the localized nonlinearities are calculated by using classical FRF measurements and identified modal parameters.

In a later study, Aykan *et al.* [11] improved the method given in [6] by calculating complete FRFs of linear part of the system using linear modal identification and modal superposition. Aykan *et al.* also utilize Describing Function Inversion method to obtain describing function of the corresponding nonlinear element based on measured response from nonlinear DOFs. Type and parameter of the nonlinear element are determined by polynomial curve fitting to describing function data points.

A common feature of most of the methods in literature is the identification of types of nonlinearities by observing system responses or nonlinear restoring forces, which is a time consuming process. It is not suitable for identification of nonlinearity in a series of products due to manufacturing errors and assembly differences.

2.2   Parameter Estimation

Parameter estimation of nonlinearities in a classified system is an easily handled problem. There exist time domain and frequency domain methods targeting parameter identification in literature [1, 12, 13]. Although these methods have their own handicaps, they are promising in specific application areas.

Conditioned Reverse Path (CRP) proposed by Richards and Singh [14] uses spectral analysis in frequency domain to compute the coefficients of nonlinearity matrix applying Gaussian random excitations.

Another parameter identification method proposed by Atkins *et al.* [15]. Parameter identification of nonlinear system is performed via using modal restoring force surface method in the study. Underlying linear part of the nonlinear system is assumed to be known. The proposed method decouples nonlinear set of equations in modal domain to make the cross-modal displacements in the equations zero. Zero modal

displacements are achieved applying appropriate excitation forces determined by Volterra series approach. Considering MDOF system, satisfying zero modal displacements brings some practical limitation to application as well.

Neural network classifiers have been used for parameter identification of structural nonlinearities [16, 17]. Early study on this subject is performed by Masri *et. al* [16]. Artificial Neural Network is used for discrete-time system identification. The method proposed in the study targets estimating linear and/or nonlinear parameters of a system expressed in Chebychev basis functions mathematically. Chen and Billings [18] utilize Artificial Neural Network for discrete-time parameter identification of nonlinear restoring forces expressed in NARMAX model mathematically. Identification performances of different neural networks architectures are evaluated in the study.

Neural networks combined with fuzzy theory are utilized by Liang *et al.* [19] for parametric identification of nonlinear restoring forces. The proposed method requires to measure displacements, velocities and accelerations of all the degrees of freedoms at discrete times. Type and locations of linear and nonlinear elements are to be known *a priori*.

The method proposed in this study, identifies nonlinearities associated with the system utilizing neural networks. Determining type of a nonlinearity by using neural networks is a new application in structural dynamics. Constructed networks are to be trained with sample data sets which are frequency response functions of selected points on the system. Nonlinear frequency responses of the system are obtained by analyzing possible nonlinear system configurations. It should be noted that in order to perform these analyses, a mathematical model of the system is required. After training the networks, location and type of the nonlinearity is determined by running trained networks with measured systems responses as the input. Parameters of the classified nonlinearities are identified by means of a regression network utilizing the same input data used for classification. Identified parameters are further improved by optimizing mean squared error (MSE) between the nonlinear model and the measured system.

The method proposed does not require taking measurements from the nonlinear DOFs; hence, system responses measured from arbitrary locations can be used for identification purposes. Moreover, with the proposed method, it is as well possible to handle measurement noise by injecting noise into data during the training process. Once networks are trained, identification is very fast; hence, the method proposed is very suitable for identification of nonlinearities in a series of products due to manufacturing errors and assembly differences.

CHAPTER 3

NONLINEAR VIBRATION ANAYSIS UNDER HARMONIC EXCITATIONS

As outlined in the previous chapter, stepped sine testing method is to be used for identification purposes in this study. Therefore, tested nonlinear systems are excited harmonically. This chapter devoted to the analysis of harmonically excited nonlinear systems, and consists of three sections. Mathematical formulation of forced response of nonlinear systems is given and discussed in the first section. In the following section, performance of commonly used numerical methods is investigated. Receptance decoupling method is introduced to reduced number of coupled nonlinear equations in the last section.

3.1 Mathematical Formulation

General time domain representation of equation of motion of a linear discrete system excited by a harmonic external forcing is given by

$$[M]\{\ddot{x}\}+[C]\{\dot{x}\}+i[H]\{x\}+[K]\{x\}=\{f(t)\}, \tag{3.1}$$

where $[M]$, $[C]$, $[H]$ and $[K]$ are mass, viscous damping, structural damping and stiffness matrices of the linear system, respectively. In Eq. (3.1), $\{x(t)\}$ and $\{f(t)\}$ are response and excitation force vectors of the system, respectively.

In case any nonlinear effect exists within a linear system, a nonlinear restoring forcing term appears in Eq. (3.1). Equation of motion of the nonlinear system becomes

$$[M]\{\ddot{x}\}+[C]\{\dot{x}\}+i[H]\{x\}+[K]\{x\}+\{f_{NL}(\{x\},\{\dot{x}\},...)\}=\{f(t)\}, \tag{3.2}$$

where $\{f_{NL}(t)\}$ represents nonlinear restoring forces.

For a linear system under single harmonic forcing, external force and response of the system can be written in the following form

$$\{f(t)\}=\text{Im}(\{F\}e^{i\omega t}), \tag{3.3}$$

$$\{x(t)\} = \mathrm{Im}\left(\{X\}e^{i\omega t}\right). \tag{3.4}$$

Response of a linear system to a periodic forcing can be written as a superposition of responses corresponding to each harmonic. However, the same property is not valid for nonlinear systems. In addition to this, even a nonlinear system is excited by a single harmonic forcing, system response may contain multiple harmonics.

If a nonlinear system is excited by a single harmonic forcing as in Eq. (3.3), response of the nonlinear system is to be

$$\{x(t)\} = \mathrm{Im}\left(\sum_{h=1}^{\infty}\{X\}_h\, e^{ih\omega t}\right) + \{X\}_0, \tag{3.5}$$

where h represents the harmonic number. If the nonlinear system is excited by a periodic forcing, it can be written in terms of multiple harmonics by using Fourier series representation as follows

$$\{f(t)\} = \mathrm{Im}\left(\sum_{h=1}^{N}\{F\}_h\, e^{ih\omega t}\right) + \{F\}_0. \tag{3.6}$$

Response of the nonlinear system is still going to be periodic and Eq. (3.5) is valid as well. Although external forcing includes finite number of harmonics, nonlinear system responds in infinite number of frequencies, which are generated by the internal nonlinear forcing. Internal nonlinear forcing can be expressed in multiple harmonics as follows

$$\left\{f_{NL}\left(\{x(t)\},\{\dot{x}(t)\},\cdots,\right)\right\} = \{F_{NL}\}_0 + \mathrm{Im}\left(\sum_{h=1}^{\infty}\{F_{NL}\}_h\, e^{ih\omega t}\right), \tag{3.7}$$

where $\{F_{NL}\}_h$ is the complex amplitude vector of the hth harmonic, which is a function of system response and frequency, and can be defined as

$$\left\{F_{NL}\left(\{X\},\omega\right)\right\}_0 = \frac{1}{2\pi}\int_0^{2\pi}\left\{f_{NL}\left(\{x\},\{\dot{x}\},\cdots,\right)\right\}d\psi, \tag{3.8}$$

$$\left\{F_{NL}\left(\{X\},\omega\right)\right\}_h = \frac{i}{\pi}\int_0^{2\pi}\left\{f_{NL}\left(\{x\},\{\dot{x}\},\cdots,\right)\right\}e^{-ih\psi}d\psi, \qquad h = 1,\ 2,\ .... \tag{3.9}$$

where $\psi = \omega t$.

Substituting Eqs. (3.5)-(3.7) into Eq. (3.2) the following nonlinear equation of motion is obtained

$$\sum_{h=1}^{N}\{F\}_h\, e^{ih\omega t} + \{F\}_0 = [K]\{X\}_0 + \sum_{h=1}^{\infty}\left([K]+i[H]-(h\omega)^2[M]+i(h\omega)[C]\right)\{X\}_h\, e^{ih\omega t}$$
$$+\{F_{NL}\}_0 + \sum_{h=1}^{N}\{F_{NL}\}_h\, e^{ih\omega t} \qquad\qquad h=1,2,...,N \qquad\qquad, (3.10)$$

where $\{X\}_0$, $\{F_{NL}\}_0$, $\{F\}_0$ are non-oscillatory time independent bias terms. Balancing the terms, the following set of complex algebraic equations is obtained.

$$[K]\{X\}_0 + \{F_{NL}\}_0 = \{F\}_0 \qquad\qquad (3.11)$$

$$\left([K]+i[H]-(h\omega)^2[M]+i(h\omega)[C]\right)\{X\}_h + \{F_{NL}\}_h = \{F\}_h \qquad\qquad (3.12)$$

Above equations are converted into a set of real valued algebraic equations which can be represented in matrix form as follows

$$\begin{bmatrix} [K] & [0] & \cdots & [0] \\ [0] & [D(\omega)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ [0] & \cdots & \cdots & [D(h\omega)] \end{bmatrix} \begin{Bmatrix} \{X\}_0 \\ \{X\}_1^{Re} \\ \{X\}_1^{Im} \\ \vdots \\ \vdots \\ \{X\}_h^{Re} \\ \{X\}_h^{Im} \end{Bmatrix} + \begin{Bmatrix} \{F_{NL}\}_0 \\ \{F_{NL}\}_1^{Re} \\ \{F_{NL}\}_1^{Im} \\ \vdots \\ \vdots \\ \{F_{NL}\}_h^{Re} \\ \{F_{NL}\}_h^{Im} \end{Bmatrix} = \begin{Bmatrix} \{F\}_0 \\ \{F\}_1^{Re} \\ \{F\}_1^{Im} \\ \vdots \\ \vdots \\ \{F\}_h^{Re} \\ \{F\}_h^{Im} \end{Bmatrix}, \qquad (3.13)$$

where

$$D(h\omega) = \begin{bmatrix} [K]-(h\omega)^2[M] & -(h\omega)[C]-[H] \\ (h\omega)[C]+[H] & [K]-(h\omega)^2[M] \end{bmatrix} \qquad\qquad (3.14)$$

The number of harmonics that should be used in Eq. (3.13) depends on the excitation frequencies and the type of nonlinearities. There is a simple but iterative method named as harmonic balancing method commonly used to determine effective harmonics. The method is well-illustrated by Worden and Tomlinson [9] on a simple example. In this study it is assumed that nonlinear restoring forces excite periodically. In some cases nonlinear restoring forcing expressions have terms, such as $\cos(\sqrt{2}\omega)$,

$\sin(2.5\omega)$, that causes aperiodicity in the system. In such cases, multitone harmonic balancing should be used to balance the system of nonlinear equations [9, 20].

Solving the resulting set of nonlinear algebraic equations by a nonlinear equation solver given by Eq.(3.13), forced response of the nonlinear system can be obtained.

3.2 Numerical Solution Techniques

In this section, solution methods of Eq. (3.13) are investigated. Principal numerical methods utilized in nonlinear algebraic system of equation solution are introduced and discussed. Derived nonlinear algebraic equations are written in a form

$$\{R(\{X\},\omega)\} = \{0\},\tag{3.15}$$

where $\{R(\{X\},\omega)\}$ is the nonlinear residual vector. Further expanding the residual vector $\{R(\{X\},\omega)\}$, for our problem we have

$$\{R(\{X\},\omega)\} = \begin{bmatrix} [K] & [0] & \cdots & [0] \\ [0] & [D(\omega)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ [0] & \cdots & \cdots & [D(h\omega)] \end{bmatrix} \begin{Bmatrix} \{X\}_0 \\ \{X\}_1^{Re} \\ \{X\}_1^{Im} \\ \vdots \\ \vdots \\ \{X\}_h^{Re} \\ \{X\}_h^{Im} \end{Bmatrix} + \begin{Bmatrix} \{F_{NL}\}_0 \\ \{F_{NL}\}_1^{Re} \\ \{F_{NL}\}_1^{Im} \\ \vdots \\ \vdots \\ \{F_{NL}\}_h^{Re} \\ \{F_{NL}\}_h^{Im} \end{Bmatrix} - \begin{Bmatrix} \{F\}_0 \\ \{F\}_1^{Re} \\ \{F\}_1^{Im} \\ \vdots \\ \vdots \\ \{F\}_h^{Re} \\ \{F\}_h^{Im} \end{Bmatrix} = \{0\},$$

$$\tag{3.16}$$

3.2.1 Fixed Point Iteration

Iterative formula for fixed point iteration can be given as follows [21]

$$\{X\}_{i+1} = \{G(\{X\}_i,\omega)\}.\tag{3.17}$$

where, $\{G(\{X\}_i,\omega)\}$ is the function stationary point of which is sought. Iteration procedure continues until the relative error between $\{X\}_{i+1}$ and $\{X\}_i$ drops below a desired value. Selection of $\{G(\{X\}_i,\omega)\}$ is important for convergence; however, for

nonlinear set of equations, fixed point iteration rarely converges to the correct root; hence, as suggested by Ciğeroğlu and Özgüven [22], in order to improve its convergence, relaxation is used.

$$\{X\}_{i+1}^{new} = \lambda \{X\}_{i+1}^{old} + (1-\lambda)\{X\}_{i}^{old}$$ (3.18)

where λ is a weighting factor. If λ is between 0 and 1, the modification is called *underrelaxation* and it is preferred to make a nonconvergent systems convergent. If λ is between 1 and 2, the modification is called *overrelaxation* and it is used to increase the convergence rate [22].

Fixed point iteration converges if there exists a constant τ such that

$$\left\| \left[J\left(\{X\}_i, \omega \right) \right] \right\| < \tau < 1$$ (3.19)

where $\left[J\left(\{X\}_i, \omega \right) \right]$ is the Jacobian matrix of $\{G(\{X\}_i, \omega)\}$ at frequency ω and $\| \ \|$ denotes matrix norm operator [23].

Numerical root-finding algorithms usually need a good initial guess for the first iteration; therefore, linear system responses can be taken as the initial guess at the starting frequency assuming that the effect of nonlinearity is not significant at that frequency. Initial guesses for the subsequent frequencies can be taken from the solution at the previous frequency step.

3.2.2 Newton's Method

Newton's method is one of the popular root-finding numerical solution techniques based on the first order Taylor series expansion. The iterative formula for Newton's method is given as follows

$$\{X\}_{i+1} = \{X\}_i - \left[J\left(\{X\}_i, \omega \right) \right]^{-1} \left\{ R\left(\{X\}_i, \omega \right) \right\},$$ (3.20)

where, i is the iteration number. Jacobian matrix can be written in a more simplified form as

$$\left[J\left(\{X_i\}, \omega \right) \right] = \begin{bmatrix} [K] & [0] & \cdots & [0] \\ [0] & [D(\omega)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ [0] & \cdots & \cdots & [D(h\omega)] \end{bmatrix} + \left[\frac{\partial \{F_{NL}(\{X\}, \omega)\}}{\partial \{X\}} \right]$$ (3.21)

13

Newton's method converges quadratically to a unique solution point ensuring $\{R(\{X\},\omega)\}$ is three times differentiable [24].

The same initial guess determination procedure described for fixed point iteration method can be utilized for Newton's method as well.

Although convergence speed of Newton's method is quite satisfying, numerical solution process can be failed at the points where Jacobian is close to singular. Points at which Jacobian is singular are illustrated in Figure 3.1. Black and red dots on the curve locate singularity points for increasing and decreasing frequency directions, respectively.
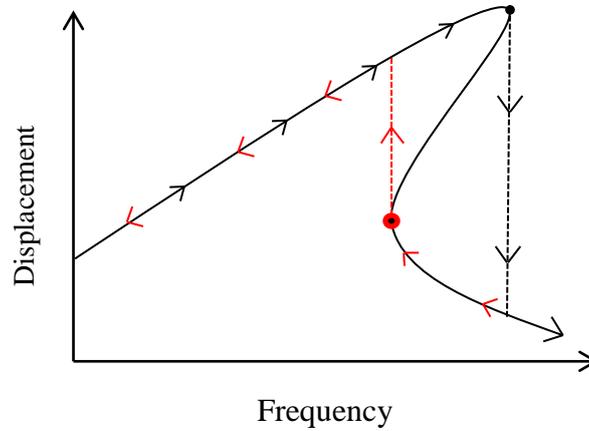
Figure 3.1   Singular points and jump phenomenon

The singularity points are also called as turning points. Turning point phenomenon is encountered in nonlinear frequency response functions frequently, especially if there exist softening or hardening type of nonlinearities in a system. In the following section, a modified version of Newton's method is utilized to handle singularity problems.

3.2.3   Arc-Length Continuation Method

Newton's method has two main handicaps. First, Jacobian of the residual vector is close to singular at the turning points, second it requires a good initial guess around turning points. In arc-length continuation method, an additional equation is introduced to the set of nonlinear algebraic equations given by Eq.(3.20) and frequency ω is set as an unknown. The added equation, called arc-length equation, forces the solution to trace the path by changing the frequency even if the path turns back [25, 26]. The arc-

length equation is an equation of an *n*-dimensional sphere whose center is located at the previous solution (Figure 3.2) and it is given as follows

$$\left(\left\{X^k\right\}-\left\{X^{k-1}\right\}\right)^2+\left(\omega^k-\omega^{k-1}\right)=s^2,\tag{3.22}$$

where *s* is the radius of multidimensional hypothetical sphere; in other words arc-length parameter. $\left\{X^k\right\}$ is the response of the nonlinear system at the kth frequency point, ω^k. Therefore, the iterative formula for Newton's method with arc-length continuation is [27, 28]

$$\left\{q_{i+1}^k\right\}=\left\{q_i^k\right\}-\left[\bar{J}\left(\left\{X_i^k\right\},\omega_i^k\right)\right]^{-1}\left\{\bar{R}\left(\left\{X_i^k\right\},\omega_i^k\right)\right\}\tag{3.23}$$
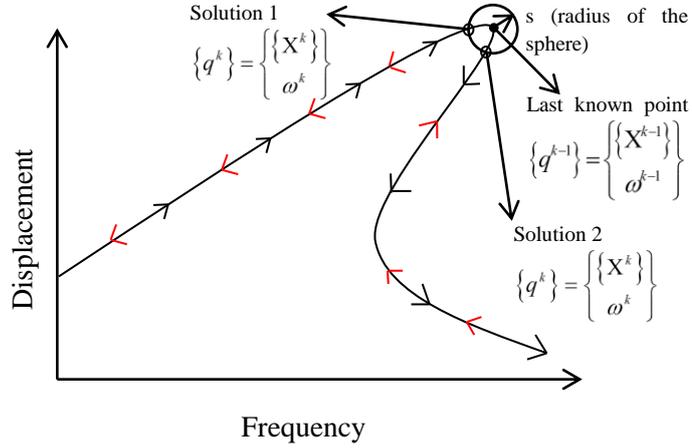
Figure 3.2   Solution Paths and Constraint Arc-Length Curve

where i is the iteration number and $\left[\bar{J}\left(\left\{X_i^k\right\},\omega_i^k\right)\right]$ is the new Jacobian matrix which is given below as

$$\left[\bar{J}\left(\left\{X_i^k\right\},\omega\right)\right]=\begin{bmatrix}\left[J\left(\left\{X_i^k\right\},\omega\right)\right] & \dfrac{\partial\left\{R\left(\left\{X_i^k\right\},\omega\right)\right\}}{\partial\omega_i^k}\\[4mm]\dfrac{\partial h\left(\left\{X_i^k\right\},\omega\right)}{\partial\left\{X_i^k\right\}} & \dfrac{\partial h\left(\left\{X_i^k\right\},\omega\right)}{\partial\omega_i^k}\end{bmatrix}\tag{3.24}$$

15

$$\{q\} = \left\{ \begin{array}{c} \{X\} \\ \omega \end{array} \right\}, \tag{3.25}$$

$$h\left(\{X^k\}, \omega\right) = \{\Delta q^k\}^T \{\Delta q^k\} - s^2 = 0, \tag{3.26}$$

$$\{\Delta q^k\} = \left\{ \begin{array}{c} \{\Delta X^k\} \\ \Delta \omega^k \end{array} \right\} = \left\{ \begin{array}{c} \{X^k\} - \{X^{k-1}\} \\ \omega^k - \omega^{k-1} \end{array} \right\}, \tag{3.27}$$

$$\{\bar{R}\left(\{X\}, \omega\right)\} = \left\{ \begin{array}{c} \{R\left(\{X\}, \omega\right)\} \\ h\left(\{X\}, \omega\right) \end{array} \right\} = \{0\}. \tag{3.28}$$

In order to increase the rate of convergence and decrease the computational time instead of using the previous solution point as the initial guess of the next step tangent predictor can be used. Initial guess for $\{X\}$ at the next solution point is given as [25, 29, 30]

$$\{X_0^k\} = \{X^{k-1}\} - \left(\frac{\partial \{R\left(\{X^{k-1}\}, \omega^{k-1}\right)\}}{\partial \{X^{k-1}\}} \right)^{-1} \frac{\partial \{R\left(\{X^{k-1}\}, \omega^{k-1}\right)\}}{\partial \omega^{k-1}}, \tag{3.29}$$

where

$$\frac{\partial \{R\left(\{X^{k-1}\}, \omega^{k-1}\right)\}}{\partial \{X^{k-1}\}} = \left[J\left(\{X^{k-1}\}, \omega^{k-1}\right) \right] \tag{3.30}$$

In Eq. (3.39), $\{X^{k-1}\}$ is the solution at the previous solution point and $\{X_0^k\}$ is the improved initial guess for the next, k^{th}, solution point. Initial guess the for frequency, ω is given as [31]

$$\omega_0^k = \omega^{k-1} \pm \frac{s}{\sqrt{\left(\left[J\left(\{X^{k-1}\}, \omega^{k-1}\right) \right]^{-1} \frac{\partial \{R\left(\{X^{k-1}\}, \omega^{k-1}\right)\}}{\partial \omega^{k-1}} \right)^2 + 1}}, \tag{3.31}$$

For frequency parameter, there are two solutions where sign of the second term at the right hand side of Eq. (3.30) is needed to be determined. Replacing $\pm$ sign in the Eq. (3.30) with sign of determinant of Jacobian matrix works quite well for most of the

cases [32]. Finally initial guess $\{q_0^k\} = \left(\{X_0^k\}^T \quad \omega_0^k\right)^T$ is obtained as illustrated on Figure 3.3.
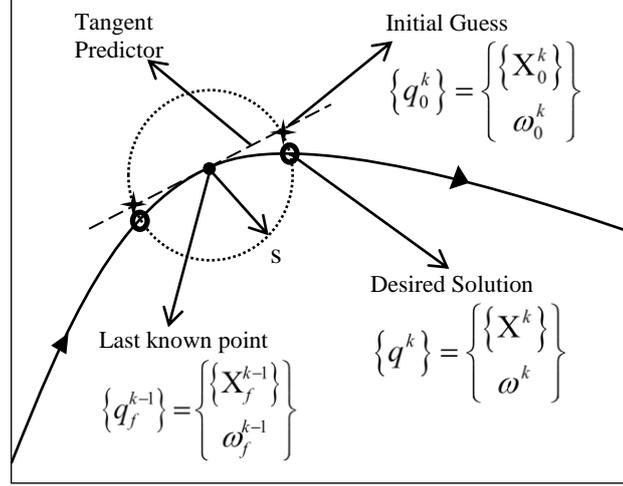
Figure 3.3   Constraint equations for initial guess

For sharp turning points arc-length parameter, *s* should be small enough to follow the path; however such a small radius length for entire solution path increases computational cost. There is a simple method which automatically updates the arc-length parameter as indicated below [33]

$$s_i = \xi s_{i-1} \tag{3.32}$$

where

$$\xi = \begin{cases} \dfrac{N_{opt}}{N_{i-1}} & if \quad 0.5 \le \dfrac{N_{opt}}{N_{i-1}} \le 2 \\[2ex] 0.5 & if \quad 0.5 \ge \dfrac{N_{opt}}{N_{i-1}} \\[2ex] 2 & if \quad \dfrac{N_{opt}}{N_{i-1}} \ge 2 \end{cases} \tag{3.33}$$

N_{opt} is the optimal number of iterations and N_{i-1} is the number of iterations in the previous solution step. It is observed in this study that N_{opt} should be around 3 to 5 for arc-length continuation method to obtain a smooth solution curve.

3.2.4 Reduction of the Number of Nonlinear Equations by Using Receptance Method

In order to determine the solution of an n-DOFs system, $(2N_h + 1)n$ nonlinear algebraic equations are required to be solved. This study basically aims classification and identification of localized nonlinearities in the systems where the number of nonlinear DOFs is significantly less than total number of DOFs. This is also the case in real life systems.

Eq. (3.10) can be divided into linear and nonlinear parts, which decreases the total number of nonlinear equations to be solved, by using the receptance method developed by Menq et al. [34] . Before dividing Eq. (3.10) coordinate numbers of the nonlinear system should be arranged such that linear and nonlinear DOFs are grouped:

$$\{X\} = \begin{Bmatrix} X^l \\ X^n \end{Bmatrix} \tag{3.34}$$

Multiplying Eq. (3.13) with multi-harmonic receptance matrix yields

$$
\begin{Bmatrix} \{X^l\}_0 \\ \{X^n\}_0 \\ \{X^l\}_1 \\ \{X^n\}_1 \\ \vdots \\ \{X^l\}_h \\ \{X^n\}_h \end{Bmatrix}
= -
\begin{bmatrix} [K]^{-1} & [0] & \cdots & [0] \\ [0] & [\alpha(\omega)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ [0] & \cdots & \cdots & [\alpha(h\omega)] \end{bmatrix}
\begin{Bmatrix} \{0\} \\ \{F_{NL}\}_0 \\ \{0\} \\ \{F_{NL}\}_1 \\ \vdots \\ \vdots \\ \{0\} \\ \{F_{NL}\}_h \end{Bmatrix} + \ldots
$$

$$
\begin{bmatrix} [K]^{-1} & [0] & \cdots & [0] \\ [0] & [\alpha(\omega)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ [0] & \cdots & \cdots & [\alpha(h\omega)] \end{bmatrix}
\begin{Bmatrix} \{F^l\}_0 \\ \{F^n\}_0 \\ \{F^l\}_1 \\ \{F^n\}_1 \\ \vdots \\ \vdots \\ \{F^l\}_h \\ \{F^n\}_h \end{Bmatrix}
\tag{3.35}
$$

Separating equations of linear and nonlinear coordinates in Eq.(3.35), the following set of nonlinear equations is obtained

$$
\begin{pmatrix} \{X^n\}_0 \\ \{X^n\}_1 \\ \vdots \\ \{X^n\}_h \end{pmatrix} = - \begin{pmatrix} [K_{nn}]^{-1}\{F_{NL}{}^n\}_0 \\ [\alpha_{nn}(\omega)]\left(\{F_{NL}{}^n\}_1 - \{F^n\}_1\right) - [\alpha_{nl}(\omega)]\{F^l\}_1 \\ \vdots \\ [\alpha_{nn}(h\omega)]\left(\{F_{NL}{}^n\}_h - \{F^n\}_h\right) - [\alpha_{nl}(h\omega)]\{F^l\}_h \end{pmatrix}. \tag{3.36}
$$

The number of nonlinear equations in Eq. (3.36) is significantly reduced for locally nonlinear systems. Nonlinear forcing vector is a function of nonlinear coordinates only; therefore, solving Eq. (3.36) response of the nonlinear degrees of freedoms is obtained.

Response of linear coordinates can be obtained from Eq. (3.37), since the nonlinear forcing vector is known

$$
\begin{pmatrix} \{X^l\}_0 \\ \{X^l\}_1 \\ \vdots \\ \{X^l\}_h \end{pmatrix} = - \begin{pmatrix} [K_{ln}]^{-1}\{F_{NL}{}^n\}_0 \\ [\alpha_{ln}(\omega)]\left(\{F_{NL}{}^n\}_1 - \{F^n\}_1\right) - [\alpha_{ll}(\omega)]\{F^l\}_1 \\ \vdots \\ [\alpha_{ln}(h\omega)]\left(\{F_{NL}{}^n\}_h - \{F^n\}_h\right) - [\alpha_{ll}(h\omega)]\{F^l\}_h \end{pmatrix}. \tag{3.37}
$$

3.3   Case Study: Response Calculation Using Harmonic Balance Method

A 2-DOF lumped parameter model given in Figure 3.4 is used as a case study. Parameters of the corresponding model are given in Table 3.1.

Table 3.1   Parameters of Linear Model

m_1 (kg)	m_2 (kg)	k_1 (N/m)	k_2 (N/m)	k_3 (N/m)	c_1 (Ns/m)	c_2 (Ns/m)	c_3 (Ns/m)
1	1	1000	1000	1000	1	1	1
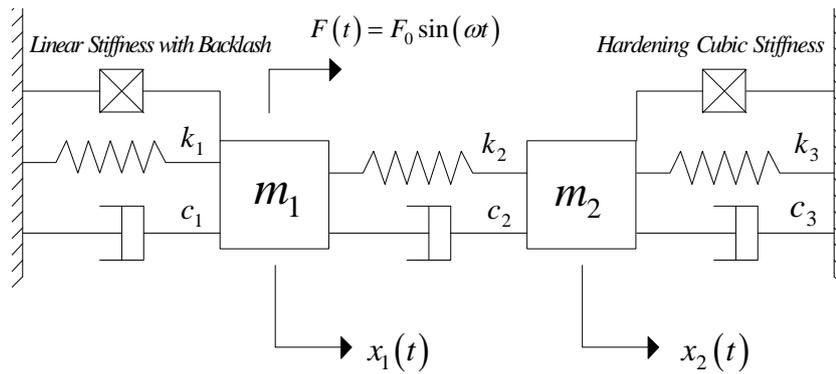
Figure 3.4   2 DOF lumped system with a linear stiffness with backlash and hardening cubic stiffness

A symmetric gap element with a stiffness of 2000 N/m and a backlash of 0.03 m is coupled to the linear system between the first DOF, and ground and a hardening cubic stiffness with a coefficient of 8×10^4 N/m^3 attached between the second DOF and ground. The system is excited from the first DOF with 50N harmonic forcing
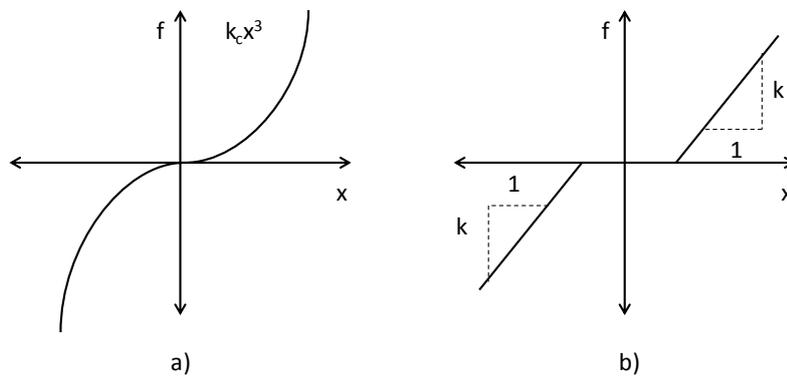
Figure 3.5   Restoring force diagram of the nonlinear elements a) cubic stiffness, b) symmetric gap nonlinearity

Frequency response of the system is, first, obtained by using ODE45 solver of MATLAB by performing time integration at each frequency from 0 to 12 Hz with 0.1 Hz increments. Time domain integration is carried out $128/f$ seconds, where f is the excitation frequency in Hz, to ensure that the system comes to steady state. After the system reaches to steady state, half of the peak-to-peak vibration amplitude is recorded as the maximum vibration amplitude.

In addition to time integration, harmonic analyses of the system are repeated by using single and multi-harmonic HBM methods. Resulted frequency responses obtained from time marching and HBM solutions are presented in Figure 3.6. Results show that using sufficient number of harmonics it is possible to approximate the response of the nonlinear system accurately. However, the frequency responses around the fundamental frequencies can be captured accurately by using only a single harmonic.
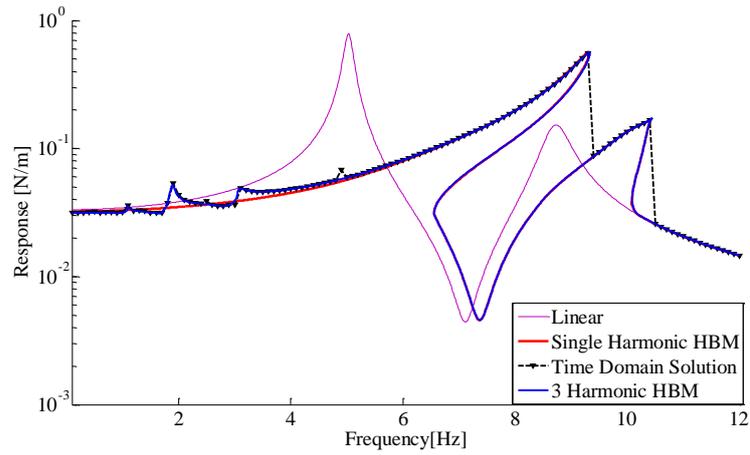
Figure 3.6   Comparison of nonlinear FRFs obtained from time integration and HBM

CHAPTER 4

THEORY OF ARTIFICIAL NEURAL NETWORKS

Learning abilities of human brain are more sophisticated compared with computers. Despite the high level of technology, artificial intelligence is still struggling on learning problems which is a simple task for a human brain. Although, operation speed of a computer processor is million times faster than that of human nerves, parallel processing of impulses through the billions of neurons makes the difference [35].

Artificial neural networks draw inspiration from biological neural networks. First artificial neuron model is proposed by McCulloch and Pitts [36]. An artificial neural network model is characterized by three items: number of neurons, number of layers and transfer functions.

4.1   Neurons

Fundamental element of a neural network is called as *neuron*. Mathematical model of a simple neuron is

$$o = f\left(\sum_{i=1}^{N_R} w_i y_i + b \right) = f\left(\{w\}^T \{y\} + b \right) = f(n), \tag{4.1}$$

$\{y\}$ is the input vector with number of elements, $\{w\}^T$ is the weight vector , b is the bias term, o is the output of the neuron, n is the net input, N_R is the number of elements in the input vector and f is the transfer function used (Figure 4.1). Inputs, weights and bias term are all real-valued numbers. $\{y\}$ is an $(N_R \times 1)$ dimensional vector.
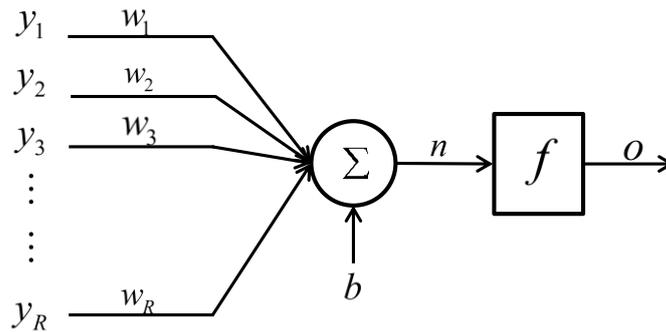
Figure 4.1   Simple Neuron Model

4.2   Transfer Functions

Transfer functions or activation functions play an important role in the structures of neural networks. Selection of a transfer function depends on input and output vectors and training algorithm. Commonly used transfer functions in neural network models are given in Table 4.1. In this section, three of the transfer functions in Table 4.1 are discussed. They are mostly preferred as transfer functions because of their differentiable characteristics.

Linear transfer function is suggested to use in output layer of the network performing curve fitting. Log-sigmoid transfer function is used in hidden layers. Hyperbolic-tangent sigmoid transfer function can be used in hidden and output layers. Studies on performance of transfer functions reveal that resulted error minimized when hyperbolic-tangent transfer function is used as both hidden and output layer transfer function for classification problems [37, 38].

4.3   Neural Network Architecture

4.3.1   Multi-Neuron Neural Network Layers

In Section (4.2), mathematical model of a single neuron is introduced. In neural networks, input vector $\{y\}$ can be connected to multiple neurons. The group of neurons is called *layer* (Figure 4.2). Layered type networks are used in *feedforward* neural networks. In feedforward neural networks, input of a layer is output of the previous layer. In contrast, a neuron in a layer can feed a neuron from the following or

previous layers in *recurrent* neural networks. In this study, main attention focuses on feedforward neural networks. Mathematical representation of a layer in a matrix form is

$$\{o\} = \{f([W]\{y\} + \{b\})\} \tag{4.2}$$

Table 4.1. Commonly used transfer functions [39]

Name	Input/Output Relation	Name	Input/Output Relation
Hard Limit	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$	Symmetrical Saturating Linear	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$
Symmetrical Hard Limit	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$	Log-Sigmoid	$a = \dfrac{1}{1 + e^n}$
Linear	$a = n$	Hyperbolic-Tangent Sigmoid	$a = \dfrac{e^n - e^{-n}}{e^n + e^{-n}}$
Saturating Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$	Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$
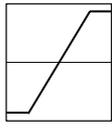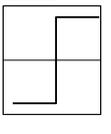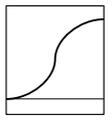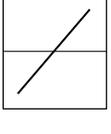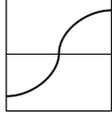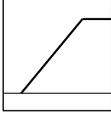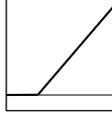
$[W], \{b\}, \{o\}$ are weight matrix, bias vector and output vector of the layer, respectively. Transfer function of each neuron can be different in a layer. However, the same transfer function is utilized throughout a layer in this study. $\{w^i\}$ is i^{th} row of the weight matrix. The single-layered network shown on the Figure 4.2 has N_s number of neurons. Therefore, the weight matrix $[W]$ is an $N_S \times N_R$ matrix.
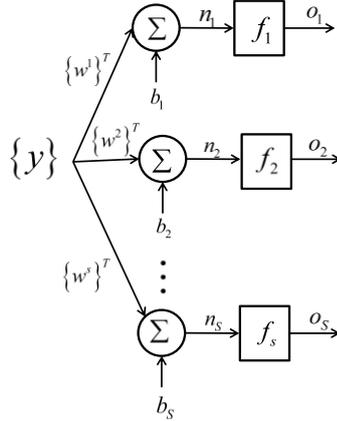
Figure 4.2   Single layer neural networks

4.3.2   Multilayer Neural Networks

Additional layer can be employed in neural network structures to increase capability of the network. In a network, the last layer which generates outputs is named as output layer and the remaining layers are named as hidden layers. The neurons in the hidden layers are called hidden neurons Schematic representation of a 2-layered neural network is illustrated in Figure 4.3.

It can be observed that output of each layer becomes the input of the next layer. Dimensions of weight matrices and bias vectors can be adjusted according to the problem requirements except R and S_2 values which are determined by the length of the input and the target vectors. In this study, a two-layer network is used for both classification and identification purposes. Any continuous function can be represented successfully by using two-layer neural networks [40].
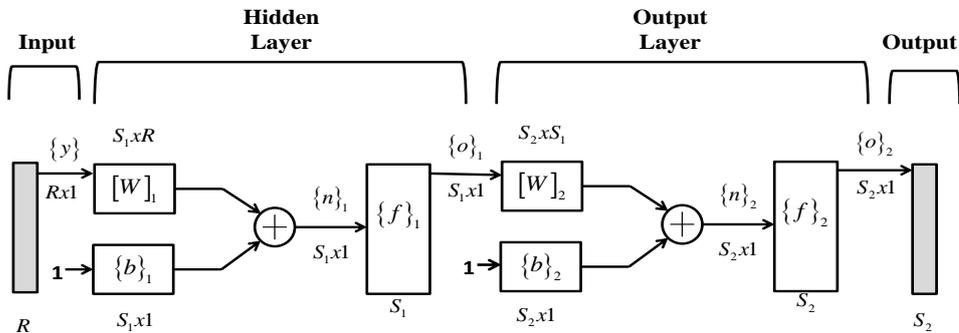
Figure 4.3   A two layer neural networks

26

4.3.3 Inputs and Targets of the Neural Networks

Inputs of neural networks are real valued vectors. Outputs of the physical system are taken as input vectors for system identification. Outputs of the system are needed to be processed before assigning as input vector due to network efficiency. The first operation is scaling values of input vector into a pre-defined interval. The Second operation is to remove the rows which are constant throughout the input vectors [41].

Purpose of neural network training is to minimize the error between target and output vectors. Target vectors simply represent the unknown parameters for parameter estimation problems; whereas, for pattern recognition problems, target vectors represent the classification of the input vectors under consideration. Target vectors can be assigned to each class in different ways; however, selection of target vectors which are not proper may cause information loss which yields misclassification. Creating an n-dimensional orthogonal target vector space for n-class systems minimizes misclassifications. For example, one can assign numbers to each target class such as 1, 0 and -1 for a three-class pattern recognition network. However, labeling three classes in a one dimensional target space might yield poor results. Therefore, the associated target vectors of each input should be orthogonal such as $target_1 = \{1 \quad 0 \quad 0\}$, $target_2 = \{0 \quad 1 \quad 0\}$, $target_3 = \{0 \quad 0 \quad 1\}$.

4.4 Neural Network Training

Performance of a network is quantified through mean squared error (MSE) between the network output vector, $\{o\}$, and target vector, $\{t\}$ as:

$$MSE = \frac{1}{N} \int_{i=1}^{N} e_i^{\ 2} = \frac{1}{N} \int_{i=1}^{N} (o_i - t_i)^2 . \tag{4.3}$$

Objective of a neural network problem is to minimize MSE via tuning the elements of weight matrices and bias vectors. The tuning process is called as *training*, where weight matrices and bias vectors are updated according to "the training algorithm". In this study, MATLAB Neural Network Toolbox is used for training operations. The toolbox utilizes Scaled Conjugate Gradient Backpropagation (SCGB) algorithm for nonlinearity localization and classification; whereas, Levenberg-Marquardt (LM) algorithm is used for identification of parameters of the nonlinear elements by default [42]. Although Levenberg-Marquardt algorithm is referred as the most efficient training algorithm for small and medium size networks regardless of the network type, computational cost coming from Hessian inversion causes drop in training speed such

that the method loses its advantage for large scale systems [42]. Because of that, SCGB algorithm is used in curve fitting networks in this study.

Before starting to the training process, collected data is required to be divided into three subsets as: training, validation and test data sets. Weight matrices and bias vectors are updated based on the training data set; whereas, network error based on the validation data set is used as the stopping criteria in order to prevent over fitting. Although the test data set is not used in training, errors based on the test data set are useful in post processing. All the data subsets should represent the entire data set. Otherwise there might be large discrepancies between targets and outputs. Optimal division of training data set is one of the main concerns in neural network training. There are studies in the literature investigating proportion and distribution of the training data [43–46]. Shahin *et al.* [44] observe that best network performance is obtained when the proportion of training, validation and test data sets are 80%, 10% and 10% respectively. In a later study, they observed that the optimal data proportions are 56%, 20% and 24% [46]. They conclude that there is no distinct relationship between data proportions and network performance. Crowther and Cox [45] demonstrated that performance of the network is higher when the training data set is divided into equal portions. They also emphasize that effect of data subsets becomes insignificant as the number of training samples is increased. Since large number of training samples is used in the case studies, proportions for training data are adapted from MATLAB defaults. In this study, 75% of samples are allocated as training data set, 15% of samples are allocated as the validation data set and the remaining 10% is used as the test data set.

CHAPTER 5

NONLINEAR SYSTEM IDENTIFICATION USING ARTIFICIAL NEURAL NETWORKS

Theory of artificial neural networks is given in the previous chapter. In this chapter, a new method is introduced to identify nonlinear systems using artificial neural networks. Proposed method is also illustrated on case studies using simulated data. Finally, an experimental case study is conducted on a test rig to demonstrate the performance of the method on a real life system.

5.1 Methodology

The proposed method is composed of two parts: in the first part, a classification network is used to identify the locations and the types of the nonlinear elements, whereas in the second part, a regression network is utilized to determine the parameters associated with the identified nonlinearities. Classification and regression networks are created according to the targeting system and then created networks are needed to be trained.

For training of the networks, forced response function of an arbitrary degree of freedom is used; therefore, a mathematical model of the system is required. There is no restriction on the mathematical model to be used; hence, finite element models (FEMs) can as well be used with the proposed method. Using the mathematical model together with the nonlinear elements and their possible configurations, simulations are performed to generate the training data by using harmonic balance method (HBM).

Training process is to be performed using the frequency response of one of the coordinates, which does not need to be measured from the nonlinear DOFs. Classical methods require response of the system taken from nonlinearity locations [3–6, 11]. However, the proposed method is capable of performing identification by measuring nonlinear system response from an arbitrary location. In order to include the effect of measurement noise, input vectors are polluted by addition of random numbers having Gaussian distribution with zero mean and realistic standard deviation. Input training data is generated from the frequency response of the nonlinear system as follows:

$$\{y\} = \left\{ \mathrm{Re}\{X(\omega_1) \quad X(\omega_2) \quad \cdots \quad X(\omega_n)\} \mid \mathrm{Im}\{X(\omega_1) \quad X(\omega_2) \quad \cdots \quad X(\omega_n)\}\right\}^T \quad (5.1)$$

Target vectors define the configurations of the nonlinear system and the parameter values of the nonlinear elements in type determination and parametric identification problems. For classification networks, length of the target vectors are determined from the total number of possible nonlinear system configurations; therefore, orthogonality of target vectors can be achieved easily.

After obtaining the input and target vectors, the classification and regression networks are ready to be trained. Created networks are trained with training data at the following step. Classification and parameter identification operations are performed by simulating measured frequency responses on the trained networks. Once the networks are trained, classification and parameter identification can be performed instantaneously. Identified nonlinear system response is further improved by means of an optimization tool as a final step as shown in the flowchart of the proposed method given in Figure 5.1.
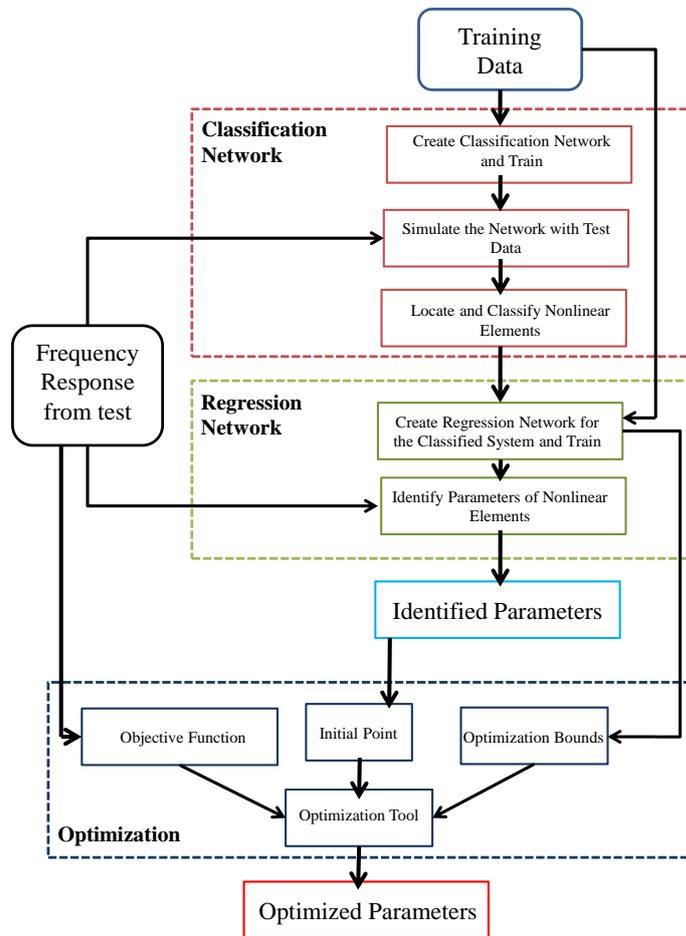
Figure 5.1. Flowcharts of nonlinearity classification and parametric identification

5.2 Case Studies

5.2.1 4-DOF Nonlinear System

In this section, application of the proposed approach is presented on a simple 4-DOF system with local nonlinearities shown in Figure 5.2.
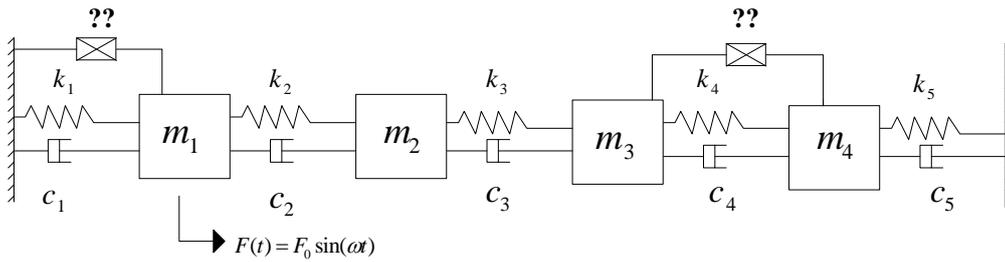
Figure 5.2. 4-DOF non-linear system schematic view

Parameters of elements of known linear system are given in Table 5.1.

Table 5.1. Parameters of elements of known linear system

k_1 (N/m)	k_2 (N/m)	k_3 (N/m)	k_4 (N/m)	k_4 (N/m)
500	500	500	500	500
c_1 (Ns/m)	c_2 (Ns/m)	c_3 (Ns/m)	c_4 (Ns/m)	c_5 (Ns/m)
5	5	5	5	5
m_1 (kg)	m_2 (kg)	m_3 (kg)	m_4 (kg)	
1	2	3	5	

From the physics of the problem or experiences it is possible to localize nonlinearities; hence, in this case study, the possible locations of the nonlinear elements are identified as between the ground and the first mass, and between the third and the fourth masses as shown in Figure 5.2. Moreover, from the physics of the problem or experience it is known that hardening type cubic stiffness nonlinearity and symmetric gap nonlinearity are the possible nonlinearities that can be encountered (Figure 5.3). This yields 8 possible nonlinear system configurations as indicated in Table 5.2 with their corresponding classification network targets.
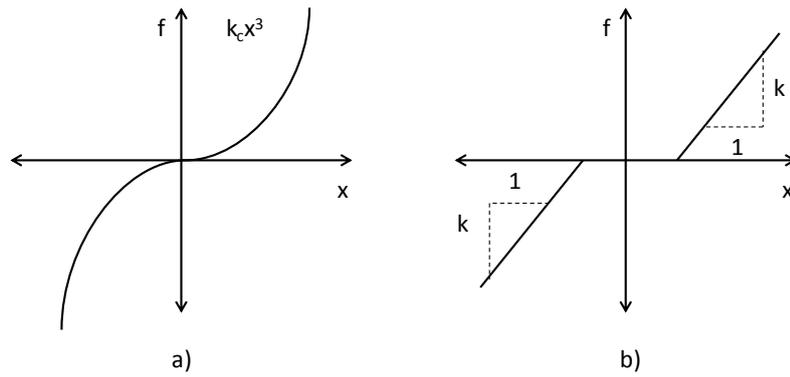
Figure 5.3.   Restoring force diagram of possible nonlinearities a) cubic stiffness, b) symmetric gap nonlinearity

Table 5.2.   Possible nonlinear system configurations and corresponding classification network targets

Conf. No.	DOF 1 and Ground	DOF 3 and DOF 4	Binary Network Targets
1	Hardening Cubic Stiffness	Linear Stiffness with Backlash	$\{1,0,0,0,0,0,0,0\}^T$
2	Linear Stiffness with Backlash	Hardening Cubic Stiffness	$\{0,1,0,0,0,0,0,0\}^T$
3	Hardening Cubic Stiffness	Hardening Cubic Stiffness	$\{0,0,1,0,0,0,0,0\}^T$
4	Linear Stiffness with Backlash	Linear Stiffness with Backlash	$\{0,0,0,1,0,0,0,0\}^T$
5	Hardening Cubic Stiffness	Linear	$\{0,0,0,0,1,0,0,0\}^T$
6	Linear	Hardening Cubic Stiffness	$\{0,0,0,0,0,1,0,0\}^T$
7	Linear Stiffness with Backlash	Linear	$\{0,0,0,0,0,0,1,0\}^T$
8	Linear	Linear Stiffness with Backlash	$\{0,0,0,0,0,0,0,1\}^T$

A two-layered classification network is created to determine the location(s) and type(s) of the nonlinearities. Numbers of neurons used in the hidden and the output layers are 50 and 8, respectively (Figure 5.4). At this stage input and target data sets are generated for training purposes. Nonlinear frequency response of the second DOF, which is not connected to any nonlinear DOF, is calculated under a 10 N harmonic excitation force applied at the first DOF in order to generate the input training data.

Frequency range of interest is taken as 0 to 8 Hz which covers all four modes of the system with increments of 0.01 Hz. For the simulations, the ranges of parameter values are selected as: 5×10^4 N/m^3 to 2×10^6 N/m^3 for cubic stiffness coefficient, and 2.5×10^{-4} m to 0.01 m gap for the symmetric gap nonlinearity with a known linear stiffness of 500 N/m. For each nonlinear system configuration at about 1600 frequency response vectors and a total of $12708 \simeq 1600 \times 8$ data sets are created. Training input vectors are polluted by normally distributed random numbers with a zero mean and 1 mm standard deviation, which corresponds to 2% of the maximum vibration amplitude, representing measurement noise. After dividing the samples into training, validation and test data; training operation is performed, which is completed after 198 epochs. A confusion matrix on which network outputs and network targets are compared is presented in Figure 5.5. It shows that 100% classification of nonlinearities in the system is achieved as indicated by the last column or row of the confusion matrix for training, validation and test sets.
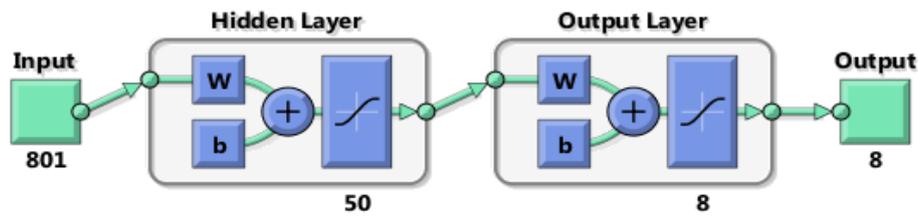
Figure 5.4.   Classification network

After determining location(s) and type(s) of the nonlinearities, parameters of the nonlinear elements are identified by means of a regression network. Eight regression networks are created for all nonlinear system configurations. For all nonlinear system configurations, 50 neurons are used in the hidden layer; whereas for the output layers 2 neurons and 1 neuron are used for configurations 1 to 4 and 5 to 8, respectively (Figure 5.6), since there is only one parameter identified in configurations 5 to 8 . Input data sets generated for the classification network can as well be used in the regression network. However, in this case, target vectors are the unknown parameters of the nonlinear elements corresponding to each input data vector. Using these data and the identified configuration, regression networks are trained similar to the classification network. Correlation coefficients of the resulted regression networks are presented in Table 5.3.
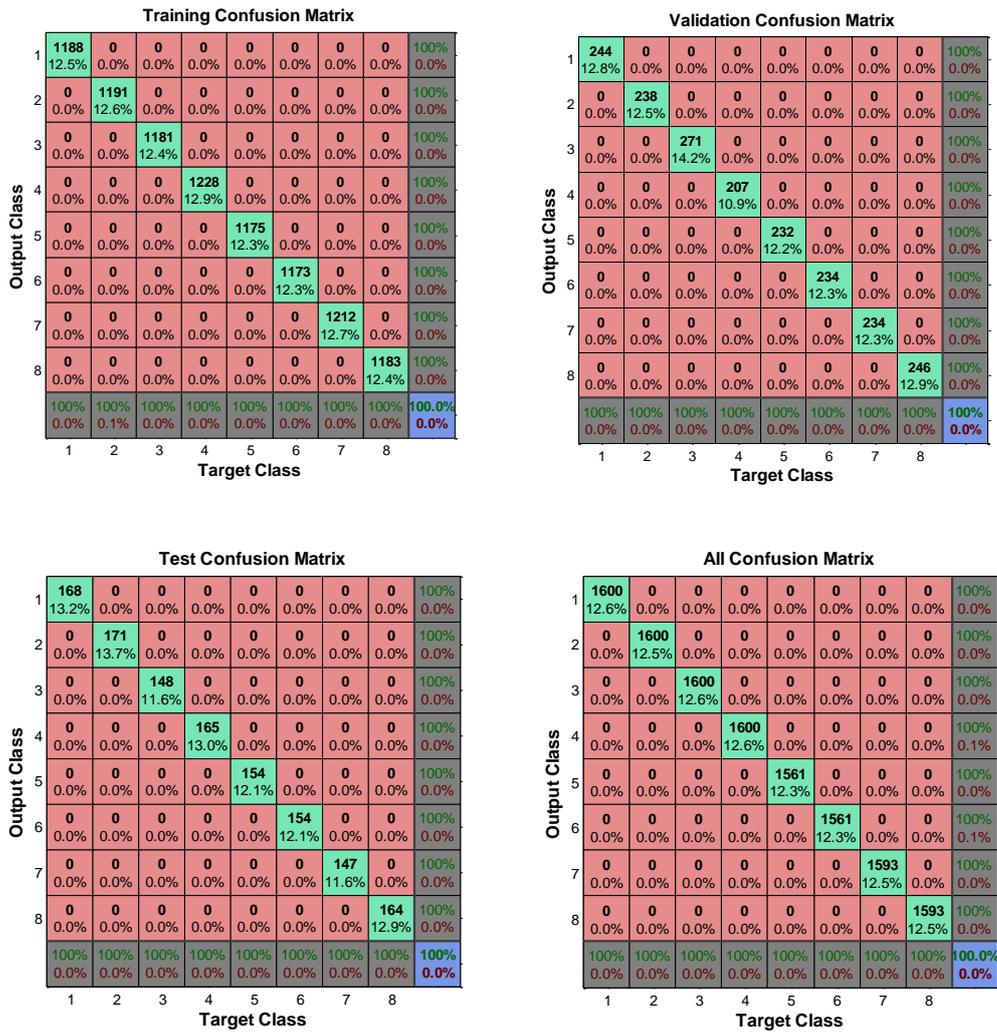
Figure 5.5.   Confusion matrices of classification network training
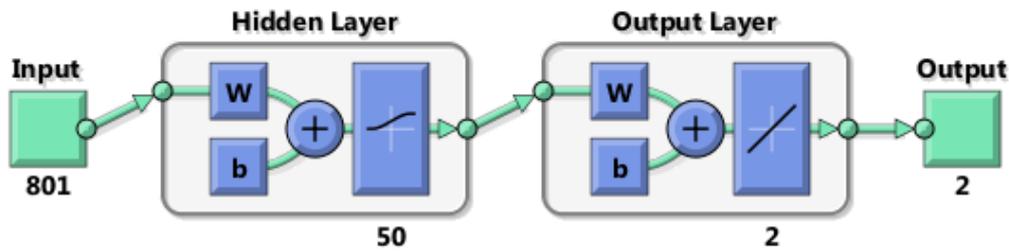
Figure 5.6.   Regression network

Table 5.3.   Correlation Coefficients for Each Regression Network

	Training Data Set	Validation Data Set	Test Data Set
Configuration 1	0.999	0.998	0.998
Configuration 2	0.998	0.997	0.997
Configuration 3	0.997	0.992	0.992
Configuration 4	0.997	0.994	0.993
Configuration 5	0.997	0.995	0.995
Configuration 6	0.990	0.992	0.991
Configuration 7	0.999	0.996	0.996
Configuration 8	0.996	0.996	0.996

In this study it is observed that MSE of noise free training inputs are less than that of noise injected ones. But, it should be remembered that error between network targets and corresponding outputs does not indicate actual performance of the network. Network training is performed with generated data while the network is going to be simulated with test data. In real life applications test data always has uncertainties and noise. If a regression network is trained with noise free generated data, the network may fail to identify the measured data due to overfitting. Therefore, it is preferred to use polluted training data.

Now, let us consider as an example the following nonlinear elements in the system given in Figure 5.2. A hardening cubic stiffness with a coefficient of 10^6 N/m attached between the first DOF and ground and a symmetric gap element with a stiffness of 500 N/m and a backlash of 0.005 m between the third and the fourth DOFs is considered. In order to simulate the experimentally measured frequency response from the second DOF, time domain solution of the nonlinear system, representing harmonic test result, is obtained by using ODE45 solver of MATLAB by increasing frequency from 0 to 8 Hz with 0.01Hz increments. Steady state oscillation amplitudes are used to obtain frequency response values at the corresponding frequency. In order to represent the effect of measurement noise, data obtained from time marching is polluted by adding random numbers having a normal distribution with zero mean and 0.5 mm standard deviation. The resulting frequency response function is given in Figure 5.7, which is used as the input to the trained networks.
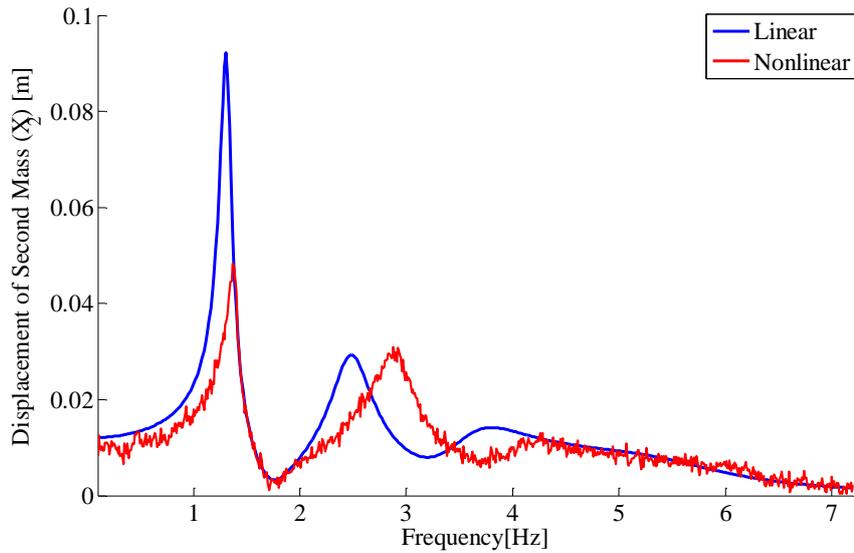
Figure 5.7.   Displacement of the second DOF

Classification and parameter identification process is also repeated by considering noise free response data. The results obtained for noisy and noise free cases are presented in Table 5.4 and 0, respectively. Comparison of identified and actual system responses are presented for noisy and noise free measurement cases in Figure 5.8 and Figure 5.9 respectively. It should be noted that, in the training process a single harmonic HBM is used to generate training data; whereas, in the simulated measurement data, time marching method is utilized. Even though identification is performed by using time domain simulation results of the system; it shows that the identified system response is very close to the actual system for both noisy and noise free measurement cases. However, it should be noted that, there is a slight increase in the error for the identified cubic stiffness parameter if noisy measurement is used. On the other hand, the error obtained for the parameters decreases if noise free measurements are used. However, such cases are not very realistic, since in real life applications it not possible to use noise free measurements.

Number of nonlinear system configurations and size of the training data generated depend on the number possible nonlinearities and their possible locations in the system. Number of nonlinear equations solved for generating training data also depends on the number of nonlinear elements used in the analysis. Therefore, for realistic finite element models, data generation process is a time consuming one, which can be overcome by employing reduction methods. In this study, the number of nonlinear equations is reduced by employing receptance method [34], which is a very effective method, especially if the nonlinearities are local.

36

Table 5.4. Classification and parameter identification results with noise free measurements (network trained with noisy data)

	Simulation with Noise Free Measurement		
	Target Vector	**Cubic Stiffness (N/m^3)**	**Backlash (m)**
Actual System	$\{1,0,0,0,0,0,0,0\}^T$	1000000	0.005
Identified System	$\{0.999,0,0,0.005,0,0,0,0\}^T$	1015500	0.004935
Error	$MSE= 3\times10^{-6}$	1.2%	1.3%

Table 5.5. Classification and parameter identification results with noisy measurements (network trained with noisy data)

	Simulation with Noisy Measurement		
	Target Vector	**Cubic Stiffness (N/m^3)**	**Backlash (m)**
Actual System	$\{1,0,0,0,0,0,0,0\}^T$	1000000	0.005
Identified System	$\{0.999,0,0,0.002,0.008,0,0,0\}^T$	1058882	0.005568
Error	$MSE= 8.4\times10^{-6}$	5.9%	11.4%
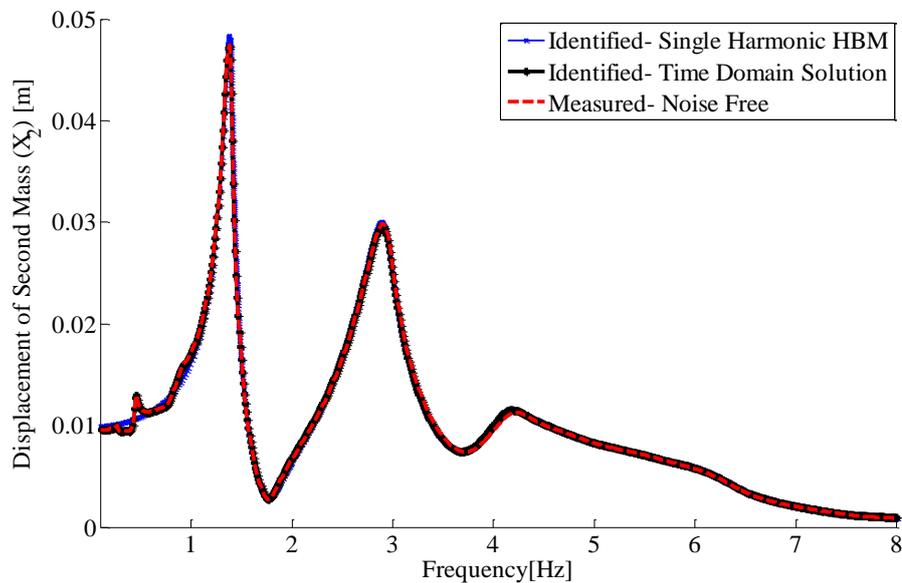
Figure 5.8. Comparisons of actual and identified system responses for noise free measurement
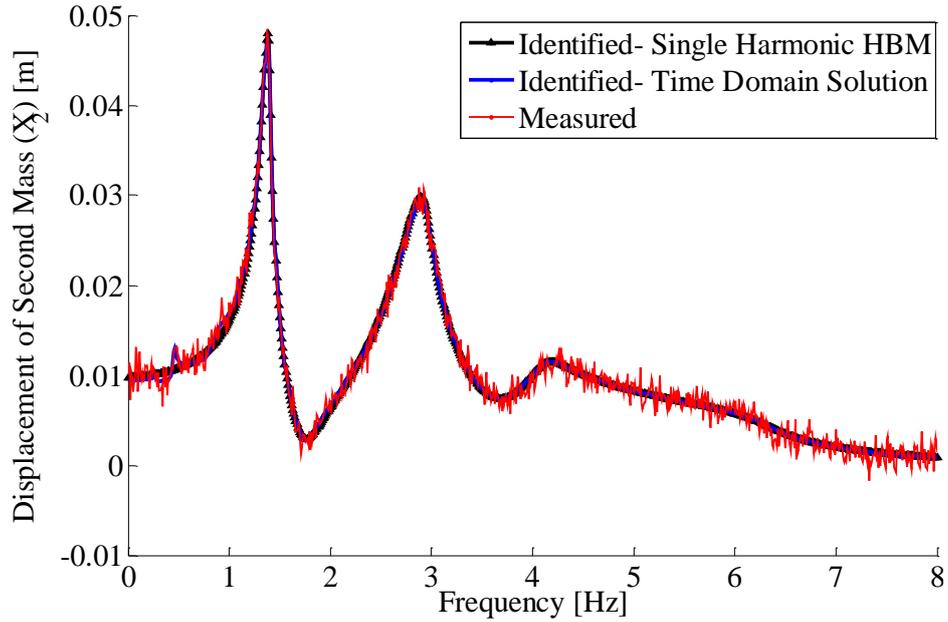
Figure 5.9. Comparisons of actual and identified system responses for noisy measurement case

In this study, two-layered networks give considerably accurate results; however, the number of neurons used in hidden layers should be optimized in order to achieve this accuracy. Therefore classification and identification trainings are repeated using different number of hidden neurons. Comparison of networks is performed by means of training results of test data sets. Classification results listed in Table 5.6 show that hidden neuron numbers between 20 and 50 improve performance of the classification network.

Table 5.6. Classification network performance for different hidden neurons

Number of Hidden Neurons	MSE			Classification Performance	Number of Epoch
	Training	Validation	Test		
10	8.356×10^{-8}	9.146×10^{-8}	9.692×10^{-8}	100%	687
20	8.078×10^{-8}	8.374×10^{-8}	6.833×10^{-8}	100%	366
30	6.912×10^{-8}	6.912×10^{-8}	4.806×10^{-8}	100%	284
40	4.264×10^{-8}	4.185×10^{-8}	4.659×10^{-8}	100%	336
50	5.383×10^{-8}	6.789×10^{-8}	6.251×10^{-8}	100%	301
60	1.018×10^{-5}	2.923×10^{-5}	9.532×10^{-6}	99.9%	234
70	1.300×10^{-7}	9.561×10^{-7}	1.181×10^{-7}	100%	246

Results of identification training for Configuration 1 are tabulated for different hidden neuron numbers in Table 5.7. It can be seen that, based on test data results, best identification results are obtained for 50 hidden neurons. However, increasing hidden neuron numbers from 10 to 50 also gives accurate results for regression network.

Table 5.7.   Regression network performance for different hidden neurons

Number of Hidden Neurons	Correlation Coefficient *R*			Performance (*MSE*)	Number of Epoch
	Training	Validation	Test		
10	0.9989	0.9966	0.9964	1.96×10^8	95
20	0.9994	0.9972	0.9967	1.53×10^8	142
30	0.9987	0.9973	0.9974	2.60×10^8	115
40	0.9998	0.9970	0.9967	9.12×10^7	312
50	0.9979	0.9965	0.9963	3.13×10^7	170
60	0.9997	0.9938	0.9941	5.91×10^8	292
70	0.9985	0.9976	0.9977	1.01×10^9	237

Since regression networks used in the case study are medium sized, LM algorithm can be used as training algorithm. For the regression network of Configuration 1, performance of LM training algorithm is compared with that of SCGB algorithm. Comparison is made by using correlation coefficients of test data sets and total training time. Training results are tabulated in Table 5.8. Although convergence rate of the LM algorithm is better than that of SCGB algorithm, training performance and training speed of SCGB algorithm are clearly better.

Table 5.8.   Comparison of SCGB and LM algorithms for the regression network

	Number of Hidden Neurons	Correlation Coefficient *R*			Training Time (s)	Number of Epoch
		Training	Validation	Test		
SCGB	10	0.9989	0.9966	0.9964	1	95
	20	0.9994	0.9972	0.9967	1	142
	30	0.9987	0.9973	0.9974	1	115
LM	10	0.9993	0.9770	0.9751	183	11
	20	1	0.9946	0.9948	1220	12
	30	0.9999	0.9964	0.9965	3396	10

Classification and regression networks are defined as two-layered networks with one hidden layer. Effect of number of hidden layers on the network performances is investigated by introducing an additional hidden layer to the networks. Results are given in 0 and Table 5.9. Comparison based on performance of test data reveals that performances of the single hidden layered networks are superior to 2 hidden layered networks.

Using the measured system responses as the inputs of the trained networks, nonlinearities in a 4-DOF system are classified and the parameters associated with them are identified. The results obtained show that the proposed method is capable of identifying the locations and types of the nonlinearities and the parameters associated with them even in the presence of measurement noise.

Table 5.9.   Performance of one hidden layered and two hidden layered regression network with different hidden neurons

Number of 1st Layer Hidden Neurons	Number of 2nd Layer Hidden Neurons	Correlation Coefficient R			Training Time (s)	Number of Epoch
		Training	Validation	Test		
0	10	0.9989	0.9966	0.9964	1	95
	20	0.9994	0.9972	0.9967	1	142
	30	0.9987	0.9973	0.9974	1	115
10	10	0.9992	0.9949	0.9936	1	98
	20	0.9949	0.9914	0.9892	0.8	53
	30	0.9990	0.9908	0.9917	1	117
20	10	0.9990	0.9950	0.9955	1	87
	20	0.9993	0.9960	0.9949	1	85
	30	0.9987	0.9955	0.9951	1	96
30	10	0.9994	0.9969	0.9968	2	92
	20	0.9961	0.9959	0.9961	2	80
	30	0.9990	0.9969	0.9965	2	92

Table 5.10.   Performance of one hidden layered and two hidden layered classification network with different hidden neurons

Number of 1st Layer Hidden Neurons	Number of 2nd Layer Hidden Neurons	MSE			Training Performance (s)	Number of Epoch
		Training	Validation	Test		
0	10	8.356×10^{-8}	9.146×10^{-8}	9.692×10^{-8}	%100	687
	20	8.078×10^{-8}	8.374×10^{-8}	6.833×10^{-8}	%100	366
	30	6.912×10^{-8}	6.912×10^{-8}	4.806×10^{-8}	%100	284
10	10	5.983×10^{-8}	6.293×10^{-8}	5.938×10^{-8}	%100	680
	20	3.703×10^{-8}	4.353×10^{-8}	6.668×10^{-8}	%100	402
	30	1.372×10^{-6}	1.395×10^{-6}	1.434×10^{-6}	%100	639
20	10	1.113×10^{-7}	1.224×10^{-7}	1.163×10^{-7}	%100	436
	20	1.195×10^{-7}	1.217×10^{-7}	1.256×10^{-7}	%100	414
	30	5.392×10^{-8}	5.536×10^{-8}	6.379×10^{-8}	%100	336
30	10	8.131×10^{-8}	7.884×10^{-8}	8.942×10^{-8}	%100	485
	20	1.094×10^{-7}	1.106×10^{-7}	1.235×10^{-7}	%100	310
	30	4.316×10^{-8}	4.898×10^{-8}	3.805×10^{-8}	%100	304

5.3   Optimization of Identified Parameters using Genetic Algorithm

Identified parameters are further improved by utilizing an optimization based identification procedure. Optimization is performed by using MATLAB Global Optimization Toolbox. Genetic algorithm is used as an optimization method. Inputs and outputs of the regression network are used to define the optimization problem. Mean squared error between the measured frequency response and calculated frequency response based on identified parameters is taken as objective function. Optimization routine which attempts to find nonlinear parameter values minimizes the objective function give in Eq. (5.2).

$$Objective\ Function = \frac{1}{N}\sum_{i=1}^{N}\left(\left\{X\left(\omega_i\right)\right\}^{Measured} - \left\{X\left(\omega_i, k_c, \delta\right)\right\}^{Calculated}\right)^2 \qquad (5.2)$$

Each k_c and δ pairs are the individuals in genetic algorithm, where a population is a set of individuals. Number of parameter pairs in a population is named as population size. Genetic algorithm calculates objective functions for each individual in a population at each step. Depending on scores of the objective function, the population is modified for the next iteration. The individuals which have the lowest scores are kept and directly transferred to the next population. The individuals which have the highest scores are modified with random numbers sticking to optimization problem constraints. The rest of the individuals are modified by interchanging parameters between themselves. At each iteration, the individuals get closer and the optimization problem converges to an optimum point. If the selected individuals are near the optimum point, convergence speed of the genetic algorithm is increased.

By default, genetic algorithm selects random parameter values within the optimization bounds as an initial population. Parameters obtained from regression network simulation are taken as starting values to reduce total number of iterations for optimization. For instance for the example problem considered, coefficient of cubic stiffness and backlash values obtained from the trained network are 1015500 N/m^3 and 0.00475 m, respectively. Difference between target parameters and regression network outputs named as absolute identification error is used to set the bounds of the parameter space. Afterwards, fractional distribution of the absolute identification errors among the training samples is calculated. Identification errors of the case study parameters with respect to cumulative probability distributions are given in Figure 5.10 and Figure 5.11. Parameter space is chosen such that absolute identification error of the training outputs should be within the limits obtained from absolute error figures. For instance, for parameter optimization of the identified system, upper and lower bounds of the parameters are chosen to fall in the error bound of the 90% of the network outputs. Error of the cubic stiffness and backlash are determined as ±75000 N/m^3 and ±6x10^{-4} m from Figure 5.10 and Figure 5.11, respectively Parameters of the

42

optimization problem are given in Table 5.11. Optimization is performed according to the parameters given in Table 5.11. Optimized parameters are given in Table 5.12.
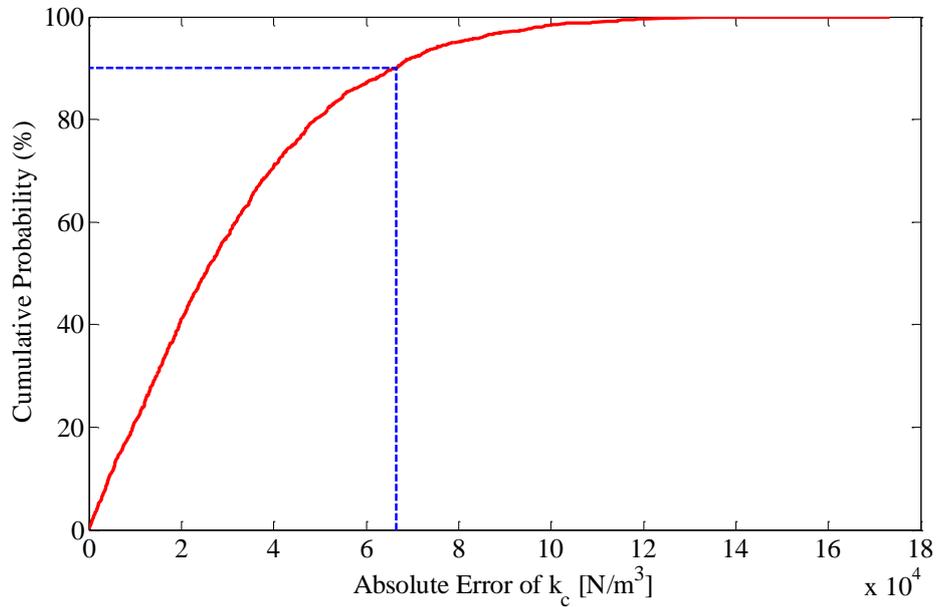
Figure 5.10. Absolute error of cubic stiffness coefficients with respect to cumulative probability
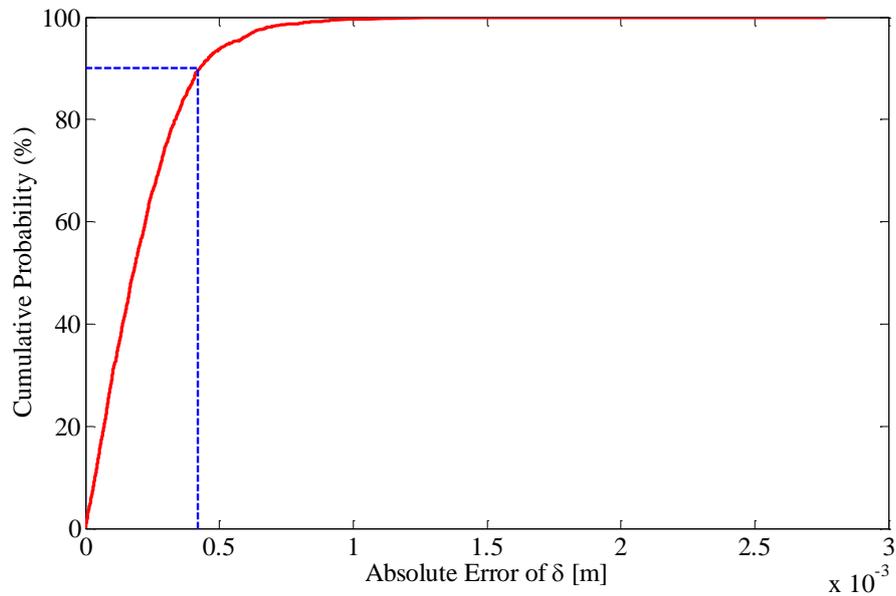
Figure 5.11. Absolute error of backlash with respect to cumulative probability

Table 5.11.   Optimization problem setup

Global Optimization Toolbox Parameters	Values
Solver	*ga*
Fitness Function	@*optim_fun*
Number Variables	2
Lower Bound	[983800 0.0049]
Upper Bound	[1133800 0.0062]
Population Size	100
Initial Population	[1058882 0.00556]
Fitness Scaling Function	*Rank*
Selection Function	*Stochastic Uniform*
Elite Count	10
Crossover Fraction	0.8
Mutation	*Constraint Dependent*
Crossover Function	*Scattered*
Migration Direction	*Forward*
Migration Fraction	0.2
Migration Interval	20
Stopping Criteria	
Generations	100
Time Limit	*Inf*
Fitness Limit	*-Inf*
Stall Generations	50
Stall Time Limit	*Inf*
Function Tolerance	1e-14

Table 5.12.   Optimization results

	Simulation with Noise Free Measurement		Simulation with Noisy Measurement	
	Cubic Stiffness (N/m^3)	**Backlash (m)**	**Cubic Stiffness (N/m^3)**	**Backlash (m)**
Actual System	1000000	0.005	1000000	0.005
Network Identified System Parameters	1015500	0.004935	1058882	0.005568
Optimization Identified System Parameters	996997	0.00502	990997	0.005101
Error in Network Identification	**1.2%**	**1.3%**	**5.9%**	**11.4%**
Error in Optimization based Identification	**0.3%**	**0.4%**	**0.9%**	**2.0%**

Frequency response of second DOF is compared with optimized results in Figure 5.12 and Figure 5.14. Figure 5.13 and Figure 5.15 show first resonance region of the frequency responses.

Figure 5.12. Comparisons of actual and optimized system responses for noise free measurement case
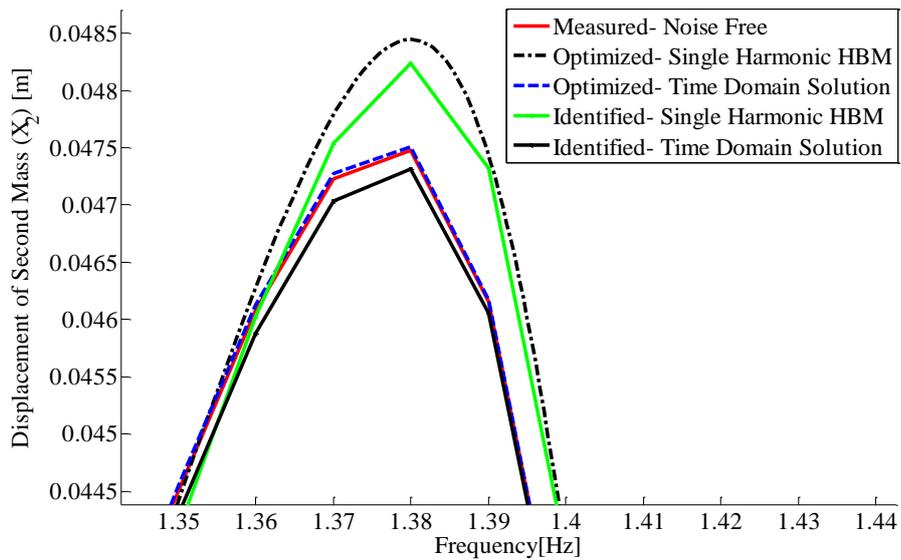
Figure 5.13. Comparisons of actual and optimized system responses around the first resonance for noise free measurement case
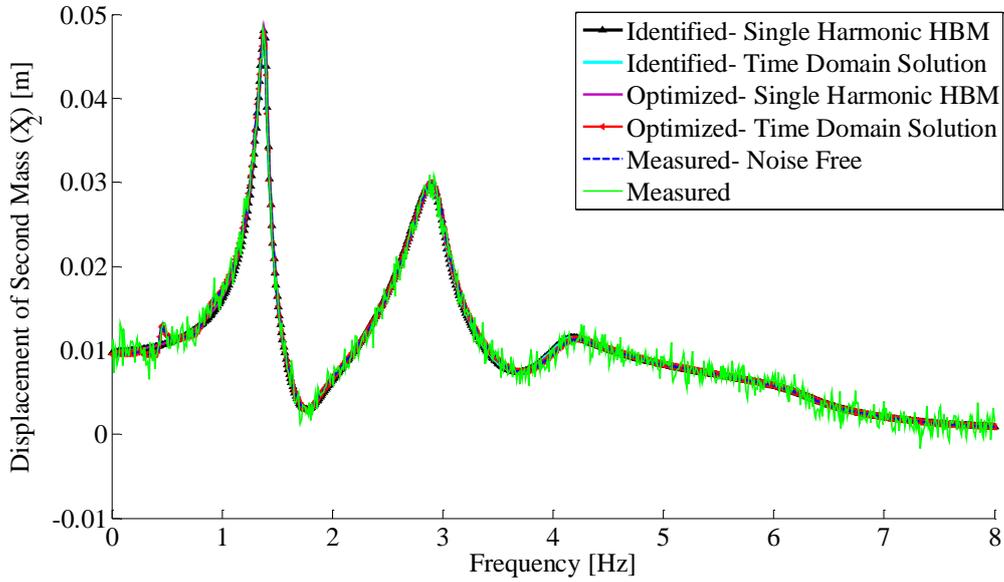
Figure 5.14. Comparisons of actual and optimized system responses for noisy measurement case
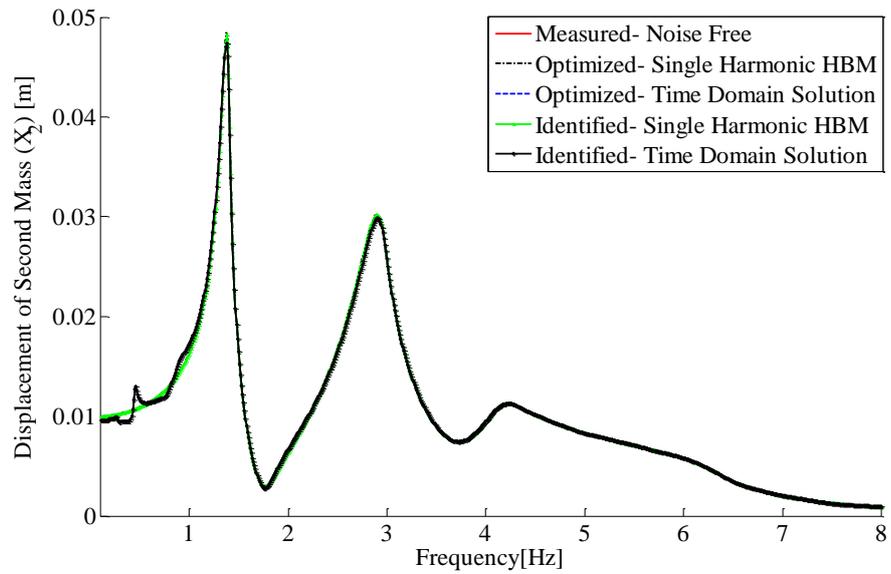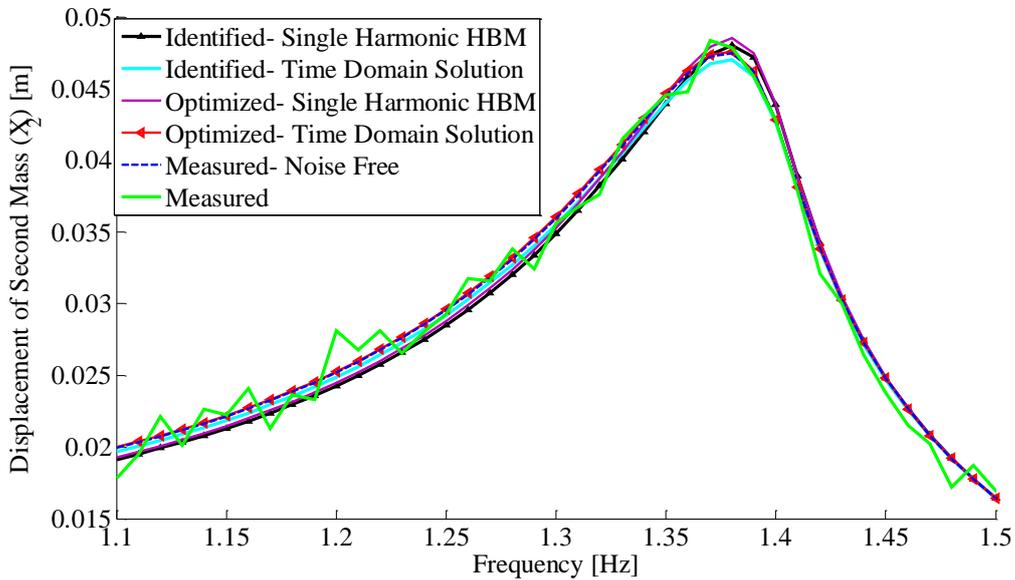

Figure 5.15. Comparisons of actual and optimized system responses around the resonance for noisy measurement case

5.4 Effect of Training Data Space on Identification

One of the important issues about neural networks is to generate training data space to map the actual system. Identification results can be inaccurate when an identified system is outside the bounds of the training data space. Three additional frequency responses are generated for nonlinear system Configuration 1. These frequency responses are on the boundary of the space and outside the space (Figure 5.16). Classification and parameter estimation results of the test data are given in Table 5.13 and Table 5.14. For test data B, C and D classification error is increased comparing with classification error of test data A. Parameter estimation of test data B and C is achieved in an acceptable error bound. However, identified parameters of test data D are not accurate.
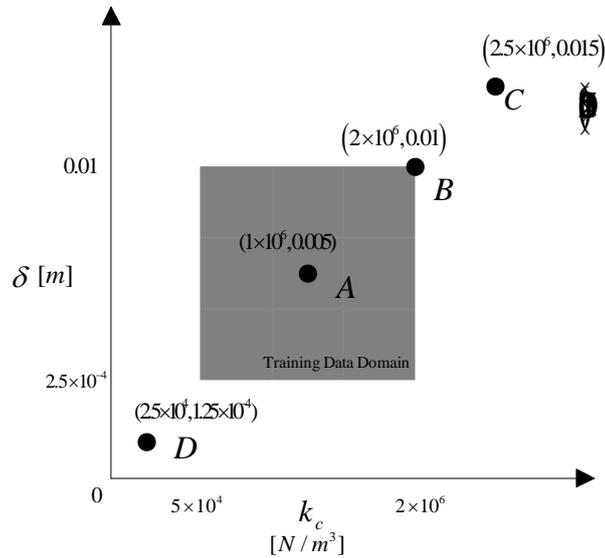
Figure 5.16.   Parameter domains for the regression network

Table 5.13.   Classification results of test data B, C and D

	Test Data B	Test Data C	Test Data D
	Target Vector	**Target Vector**	**Target Vector**
Actual System	$\{1,0,0,0,0,0,0,0\}^{\mathrm{T}}$	$\{1,0,0,0,0,0,0,0\}^{\mathrm{T}}$	$\{1,0,0,0,0,0,0,0\}^{\mathrm{T}}$
Identified System	$\{1,0,0,0.880,0,0,0,0\}^{\mathrm{T}}$	$\{1,0,0,0.897,0,0,0,0.106\}^{\mathrm{T}}$	$\{0.839,0,0,0,0.4,0,0,0\}^{\mathrm{T}}$
Error	$MSE= 0.096$	$MSE= 0.102$	$MSE= 0.0235$

47

Table 5.14.   Regression results of test data B, C and D

	Test Data B		Test Data C		Test Data D	
	Cubic Stiffness (N/m^3)	Backlash (m)	Cubic Stiffness (N/m^3)	Backlash (m)	Cubic Stiffness (N/m^3)	Backlash (m)
Actual System	2000000	0.01	2500000	0.015	25000	1.25×10^{-4}
Identified System	1.97×10^6	0.00986	2.39×10^6	0.013	1.15×10^4	1.04×10^{-5}
Relative Error	0.15%	1.4%	4.4%	16%	54%	91%

In order to demonstrate effect of optimization based identification on nonlinear system identification, only the identified parameters of test data D, which have the largest error, are improved by using optimization. Optimized parameters are given in Table 5.15. Although identified parameters are erroneous, they represent a narrow space and good starting direction for the optimization based identification. However, the results obtained by the optimization method contains more error compared to point D which suggest to increase the population size or run the optimization more than once in order to converge to the actual solution.

Table 5.15.   Optimization results of test data D

	Test Data D	
	Cubic Stiffness (N/m^3)	Backlash (m)
Actual System	25000	1.25×10^{-4}
Network Identified System Parameters	1150.0	1.04×10^{-5}
Optimization Identified System Parameters	23327	1.13×10^{-4}
Error in Network Identification	54%	91%
Error in Optimization based Identification	6.7%	9.6%

CHAPTER 6

EXPERIMENTAL VERIFICATION OF NONLINEAR SYSTEM IDENTIFICATION USING ARTIFICIAL NEURAL NETWORKS

6.1 Experimental Study

In this section, application of the proposed identification method on a real-life system is considered. An unknown nonlinearity on the test set-up is to be identified using neural networks and optimization based identification (Figure 6.1). The same test set-up was first used by Ferreira [47] for parameter identification of a known, cubic, nonlinearity existing in the system. Then, Siller [48] performed nonlinear modal tests of the same set-up in order to show the validity of the explicit formulation method developed in his study. A recent study is performed by Aykan [2] in order to locate and identify the type and parameters of the nonlinearity in the test rig.

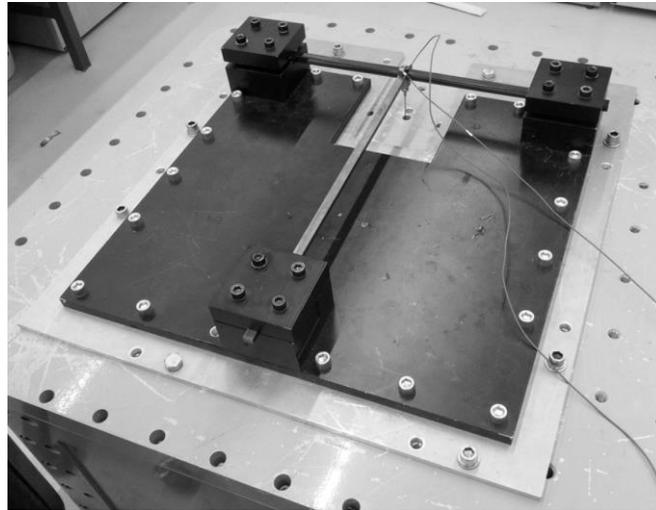**Figure 6.1**   The test rig

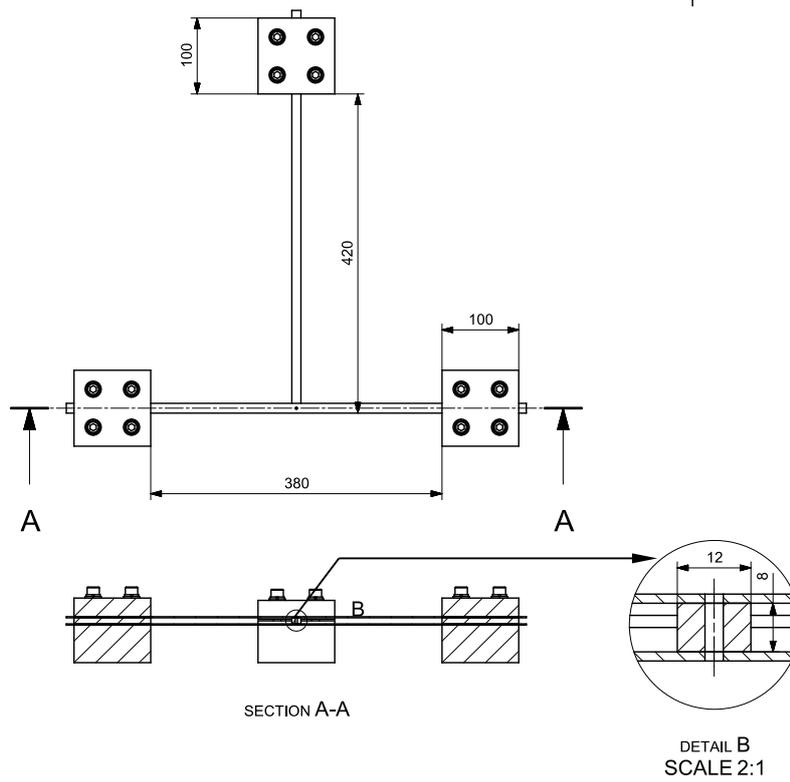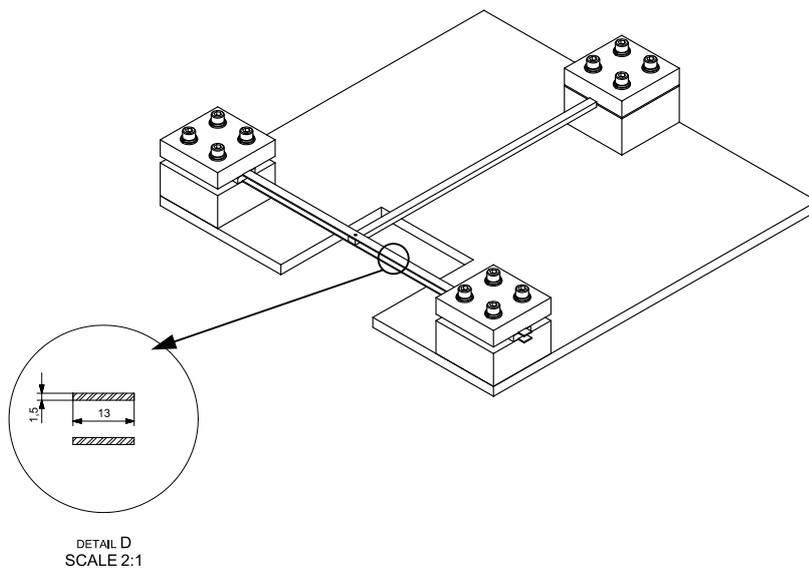Detailed geometric properties and the material specifications of the test-rig components are given in Figure 6.2.

DETAIL D
SCALE 2:1

100

420

100

380

A   A

B

12

8

SECTION A-A

DETAIL B
SCALE 2:1

Figure 6.2   Detailed geometric properties of the test rig (dimensions in mm)

50

The test set-up basically consists of rectangular cross sectioned cantilever beam and two identical thin strips. The thin strips are clamped at both ends lying in the transverse direction to the beam. The strips are attached to the free end of the cantilever beam from the midpoint. Fixed boundary conditions of the cantilever beam and the strips are obtained by clamping them to two bulky blocks with four M10 bolts. Each block is fixed to the ground plate by means of four M10 bolts. The beam has a length of 420 mm and a cross section of 8x12 mm. The strips have a length of 380 mm and a cross section of 1.5x13 mm. Material properties of the beam and the strips are given in Table 6.1.

Table 6.1   Physical properties of the components used in experiment

	Cantilever Beam	**Thin Strips**
Young Modulus, $E \left(N / m^2 \right)$	200×10^9	200×10^9
Density, $\rho \left(kg / m^3 \right)$	7850	7850
Poisons Ratio, ν	0.3	0.3
Hysteretic Damping, $\eta \left(\% \right)$	0.3	0.3

The cantilever beam corresponds to the linear part of the test rig; whereas, nonlinear behavior is imposed on the system due to the thin strips attached at the free end at of the cantilever beam. Thin strips act as a hardening cubic stiffness due to the geometric nonlinearity, which becomes significant after a certain displacement.

The structure is excited vertically at the end of the cantilever beam by using a shaker. Excitation force is transmitted from the shaker to the structure by means of a steel rod; where, a force transducer is placed between the rod and the beam. Two accelerometers are attached at the end and in the middle of the cantilever beam in order to measure the response (Figure 6.3).

Nonlinear testing is the most challenging part of the experiment, since it is required to collect frequency responses at constant forcing levels. Sinusoidal excitation is preferred as the excitation for two reasons. First, the frequency of the sinusoidal excitation dominates the response frequency of the structure and it yields good signal-to-noise ratio. Second, it is easier to control sinusoidal excitation, which is necessary to keep the applied forcing at a constant level.
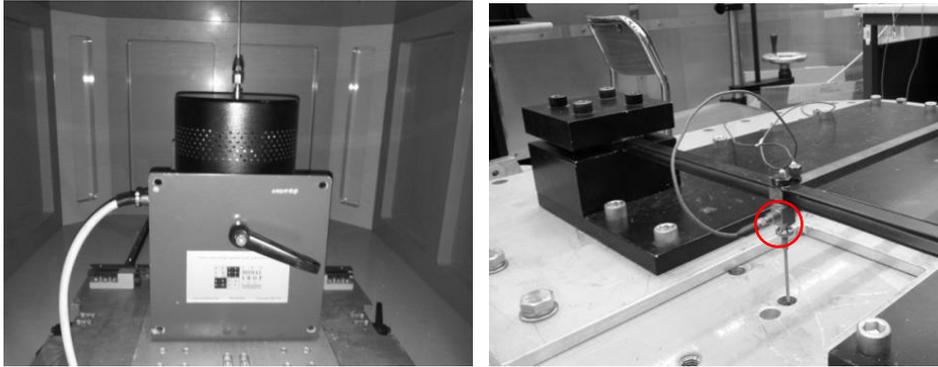
a)   b)
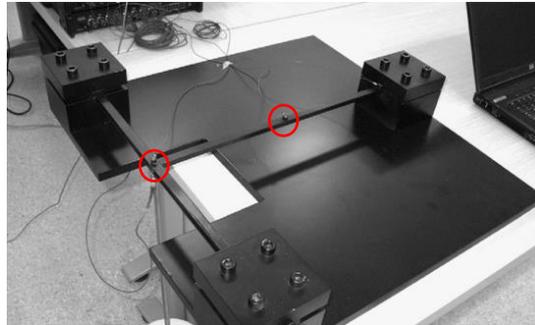
c)

Figure 6.3   a) The shaker used in the experiment b) Force transducer located at
the tip of the beam c) Accelerometers at the tip point and the midpoint

During the test, the test rig is excited at a frequency until the system reaches to the
steady-state, then the excitation frequency is increased or decreased to the next
frequency point. This is called as step-sine testing. Step-sine testing takes long testing
time, which can be thought as a main disadvantage. Test frequency range is limited to
the frequencies around the first resonance in order to shorten the testing time.
Frequency limits is obtained by using random profile excitation in a quick way. In the
experiment, the step sine testing is performed from 30 Hz up to 70 Hz with 0.1 Hz
increments. Magnitude and frequency of the excitation forcing is controlled via using
a closed-loop controller. Flowchart of the closed-looped feedback control of the test
set-up is given in Figure 6.4.
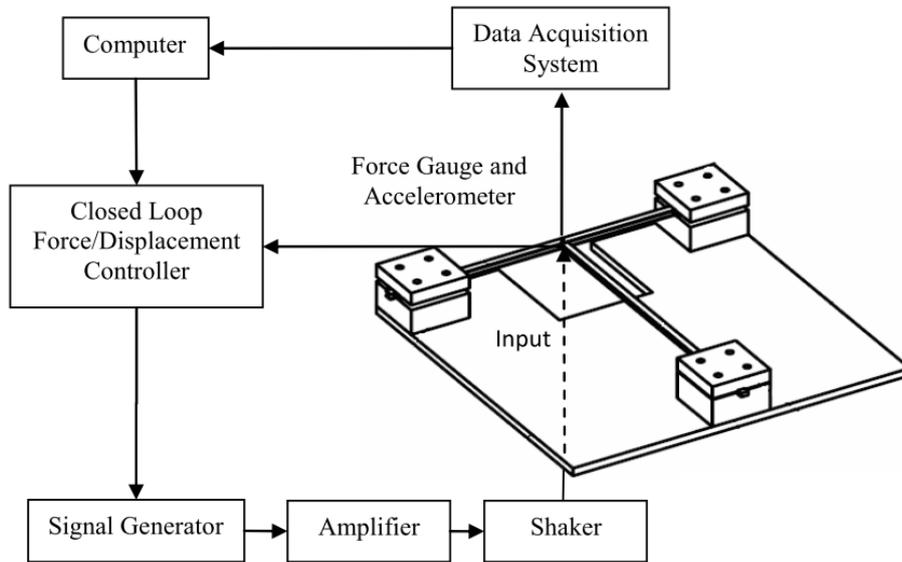
Figure 6.4   Flow chart of the testing process [49]

The time histories of the accelerations are measured by the accelerometers located on the structure. Frequency responses of the measured coordinates are obtained by means of harmonic estimator method [50]. At each frequency step, this method estimates amplitude and phase angle of a sine wave which oscillates in the excitation frequency. Parameters of the sine wave calculated for the best least squared error between the measured response and the sine wave. It means that higher harmonic effects are disregarded.

One of the main characteristic of nonlinear structures with cubic stiffness is jump phenomenon and it causes different FRFs for increasing and decreasing frequency directions. It should be noted that nonlinear systems can have multiple solutions at a single frequency. In the numerical analysis in frequency domain, stable and unstable solutions can be obtained if a sufficiently close initial guess is given. However, this is not the case in experimentation; therefore, during the test, it is observed that stepped sine testing is failed to track the response path in the increasing frequency direction. Eventually, response of the system jumps to the decreasing frequency response path at particular frequencies where multiple solutions exist. In order to handle these situations, step-sine tests are performed on systems in backward sweep directions.

The step sine-testing is repeated for different forcing levels. Frequency response of the end point and midpoint of the beam are given in Figure 6.5 and Figure 6.6, respectively.
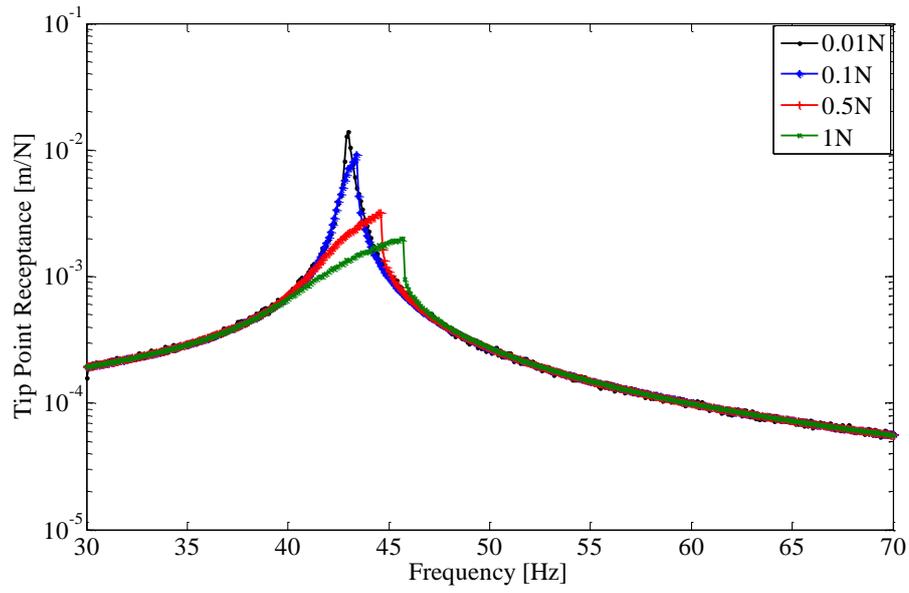
Figure 6.5   Frequency responses measured at tip of the beam
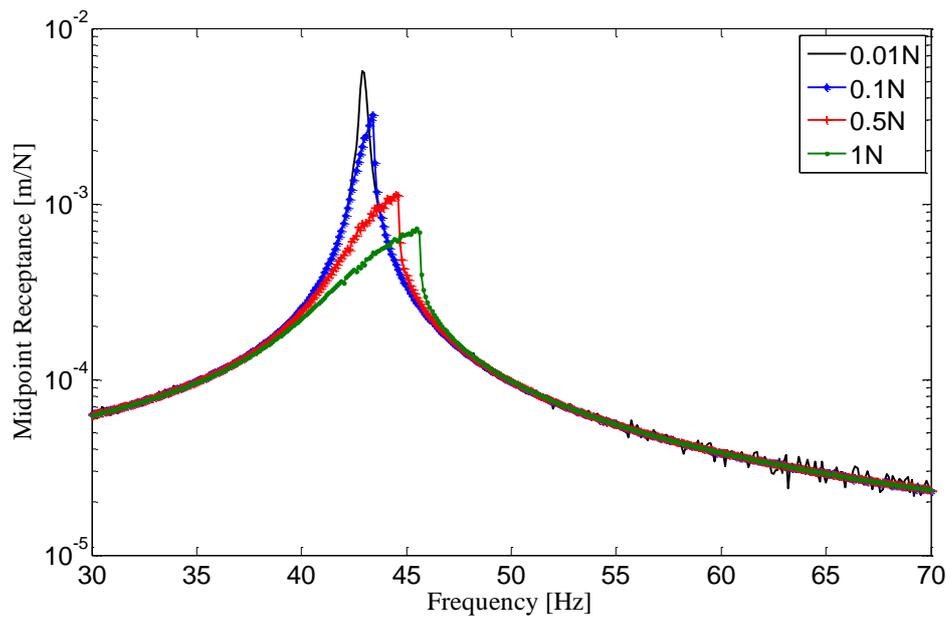
Figure 6.6   Frequency responses measured at the midpoint of the beam

54

6.2   Finite Element Modeling of the Test Rig

Finite element model of the cantilever beam and the thin strips are created in ANSYS Mechanical APDL. The cantilever beam is modeled by using BEAM188 element with 10 nodes and 6 DOFs at each node And the thin strips are modeled as a shell structure by using SHELL181 element with 10 shell elements and 6 DOFs at each node. A point mass is attached to the tip of the beam by using MASS21 element in order to represent the effect of the force transducer and the accelerometer. The cantilever beam and the strips are rigidly connected to each other by using four MPC184 elements at the tip. The nodes at the clamping regions are fully fixed for 6 DOFs. Finally 1 N load is applied to the tip of the beam in +Y direction (Figure 6.7).
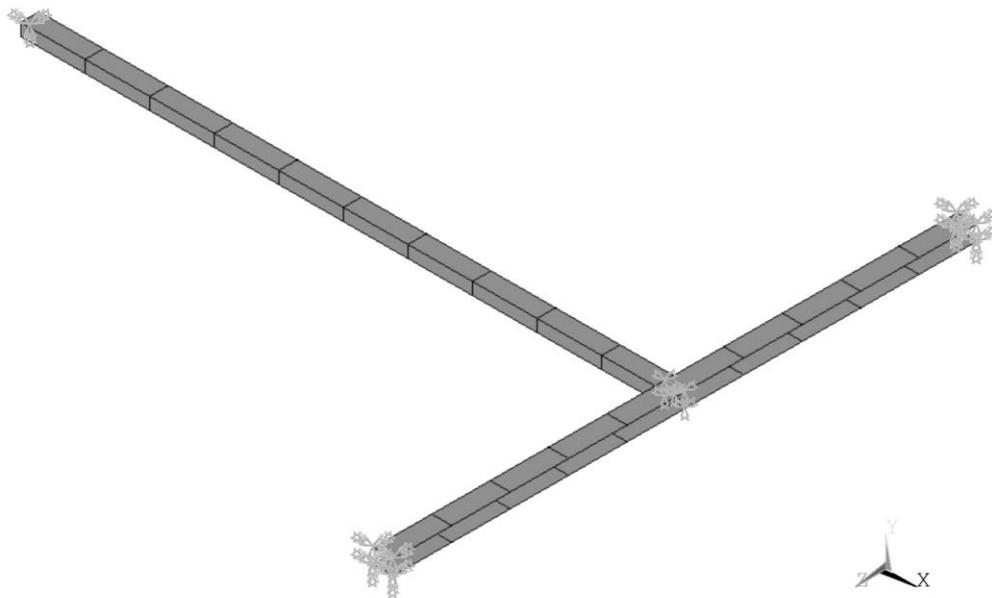
Figure 6.7   Finite element model of the test rig

After the finite element model is established, harmonic analysis is performed from 30 Hz to 70 Hz with 0.1 Hz increments. Displacements are calculated for the tip point (NODE 57) and midpoint (NODE 62) at each frequency point. Receptance of the tip point and midpoint are obtained from the FEM analysis and the tests are shown in Figure 6.8 and Figure 6.9, respectively. Results show that linear assumption is not valid for the tested structure. Increasing the level of applied force, deviates the response of the beam from the linear one. Although the response of the beam under 0.001N excitation is close to the one obtained from the FEM analysis, there is a slight shift in the results. It is considered that the difference is caused by the mass and

stiffness effect of the shaker. These disturbing effects that may come from the shaker are discussed by Siller [48].
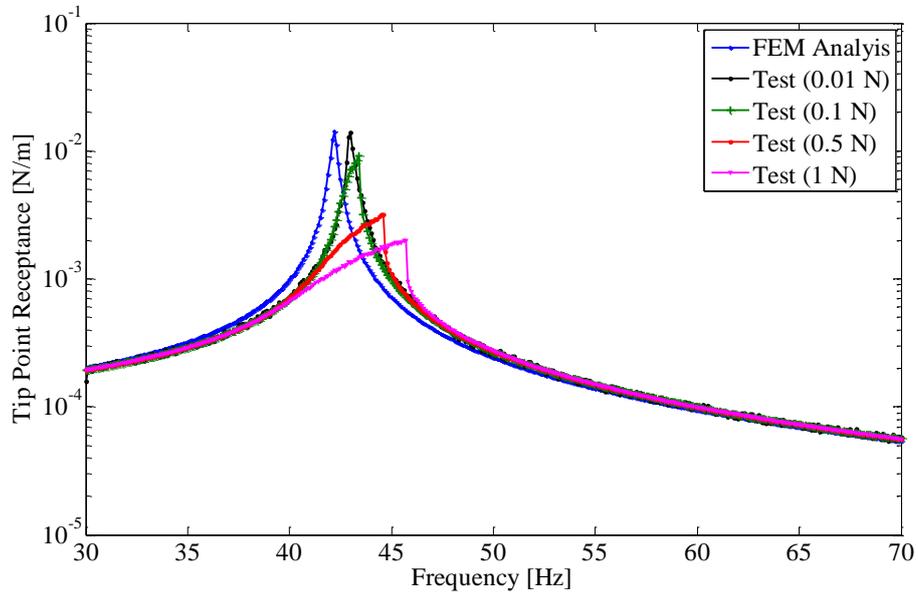
Figure 6.8   Comparison of linear and measured frequency responses at the tip point
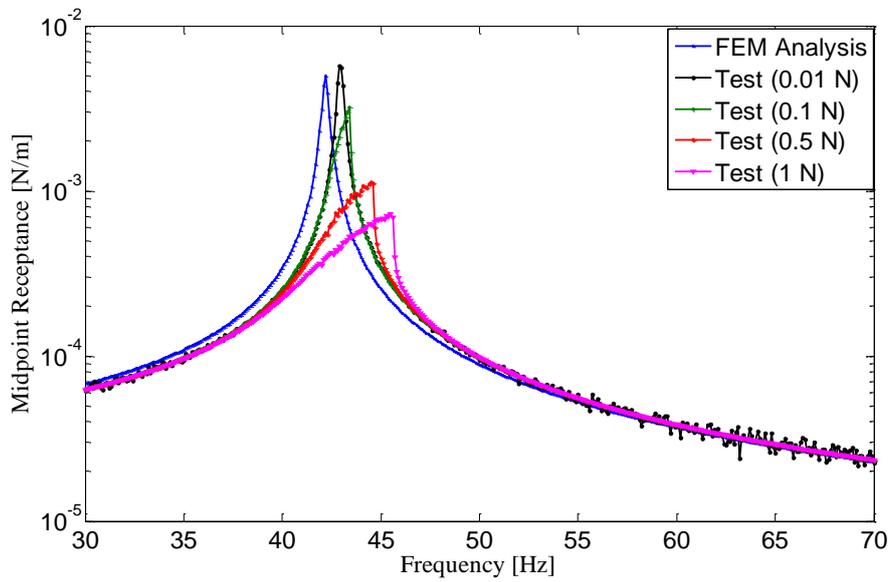
Figure 6.9   Comparison of linear and measured frequency responses at the midpoint

56

6.3   Identification of the Nonlinearity Using Neural Networks

In this section, the nonlinear behavior caused by the strips is identified using neural networks. In addition to nonlinearities, also uncertainties caused by the shaker effect are going to be identified. Identification process for type determination and parametric identification consist of three main steps: training data generation, network training and test data simulation.

Before generating training data, the nonlinear structure is dived into three substructures which are the cantilever beam (known linear part), unknown nonlinear element and unknown linear elements caused by the thin strips. Mass and stiffness matrices of the cantilever beam are generated from the finite element model; where, inactive DOFs are reduced from the mass and stiffness matrices. Since the cantilever beam is excited in Y direction, only the DOFs associated with displacement in Y direction and rotation about Z axis are considered. Displacements of the rest of the DOFs are zero for each node. 60 DOFs linear system is reduced to a 20 DOFs lumped system by using Guyan reduction technique (Figure 6.10). Reduced order model decreases matrix dimensions of linear equations in Eq. (3.37), whereas, the number of nonlinear equations are not affected since receptance method is utilized.
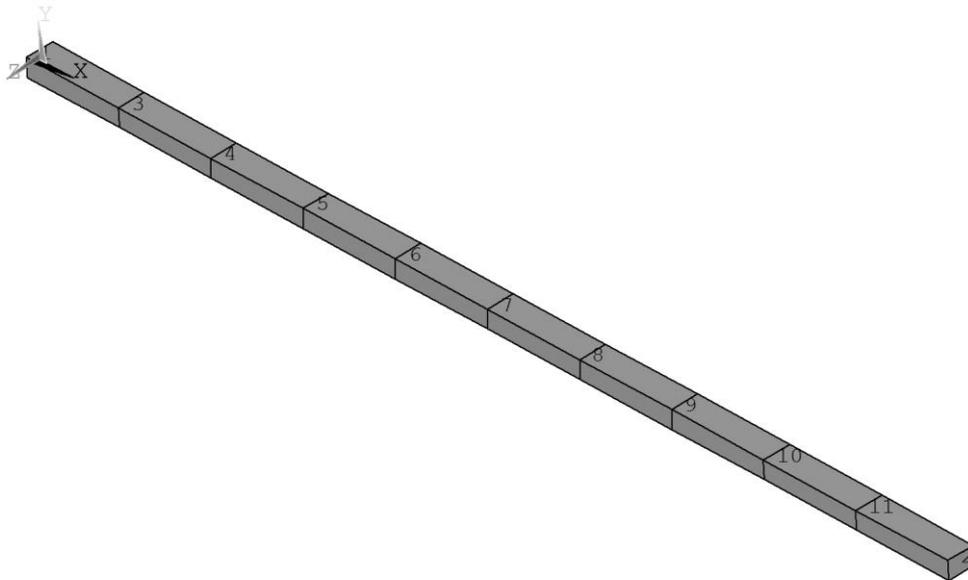
Figure 6.10   Finite element model of the cantilever beam used in the experiment

Analytical model of the test rig is constructed by coupling a nonlinear element, a concentrated mass and a linear stiffness at NODE 2. The accelerometer at the midpoint is modeled as a concentrated mass coupled to NODE 7 (Figure 6.11).

Nonlinear frequency response of the measurement points of the analytical model are calculated using single harmonic HBM. Frequency responses are generated at 0.1 N excitation force from 30 Hz to 70 Hz. It should be noted that type of the nonlinear element is unknown; therefore, possible nonlinear system configurations are generated first. Test results show that while excitation force increases, FRF curves both shift and decrease in amplitude. Looking at the test results, cubic stiffness and dry friction are selected as possible nonlinearities for this case (Figure 6.12). Possible nonlinear system configurations are listed in Table 6.2.
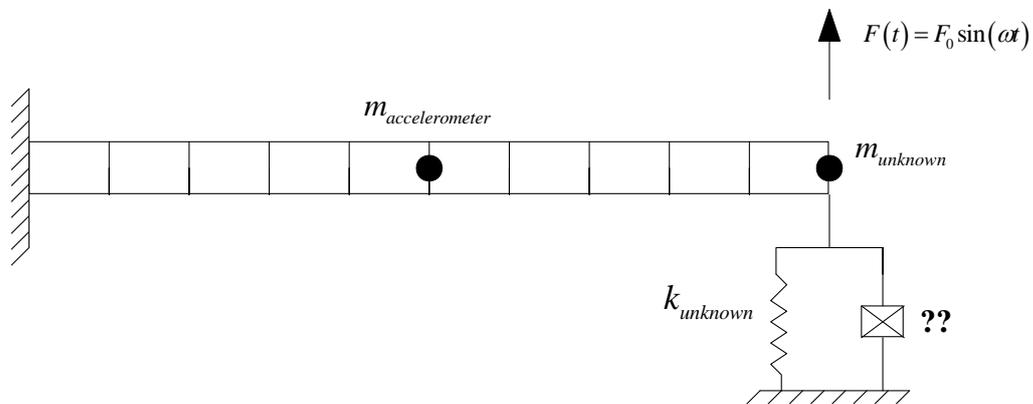
Figure 6.11   Schematic representation of the analytical nonlinear model of the test rig
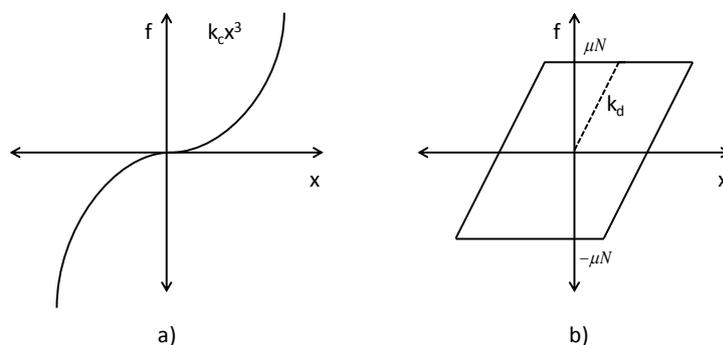
Figure 6.12   a) Restoring force diagram of hardening cubic stiffness b) Restoring force diagram of single slope macro-slip element

Table 6.2   Possible nonlinear system configurations

Conf No.	NODE 2 and Ground	Binary Network Targets
1	Hardening Cubic Stiffness	$\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$
2	Dry Friction with Single Slope	$\begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$

A parameter space is created for each configuration. Parameters are selected such that resonances of the frequency responses fall within the working frequency range. Lower and upper bound of the parameters are given in Table 6.3. 160000 different frequency responses are generated for each configuration for training operations. Before starting the training, generated data sets are polluted with noise to improve generalization. Test results obtained from the experiment show that noise level of measurements are quite low. Therefore, generated data sets are polluted with random distributed numbers having 5×10^{-7} mm standard deviation and zero mean.

Table 6.3   Lower and upper bounds of the unknown parameters

Parameter	Lower Bound	Upper Bound
k_c	5×10^6 N/m^3	1×10^9 N/m^3
k_d	5×10^4 N/m	5×10^8 N/m
μN	0.015 N	0.075 N
$k_{unknown}$	1×10^3 N/m	1×10^4 N/m
$m_{unknown}$	0.01 kg	0.07 kg

Two classification networks are constructed with 2 layers, 50 hidden layer neurons and 2 output layer neurons for midpoint and tip point measurements (Figure 6.13). Generated data sets are dived into three groups for training (75%), validation (15%) and testing (10%) operations. Created networks are trained with input training data and corresponding target configuration vectors.
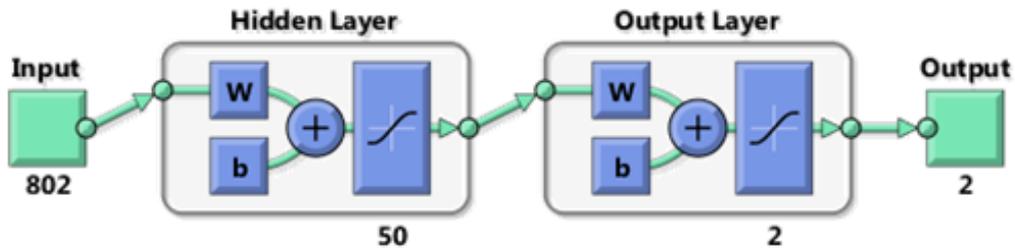
Figure 6.13   Classification network of the test rig

Training stage is completed after 352 and 504 epochs for midpoint and tip point classifications. 100% classification of training data sets is achieved for both classification networks as represented by the confusion matrices in Figure 6.14 and Figure 6.15. Configuration of the nonlinearity in the test rig is classified by simulating measured FRFs on midpoint classification network and tip point classification network separately. Simulation results are given in Figure 6.14.

Figure 6.14   Confusion matrices obtained from tip point classification process
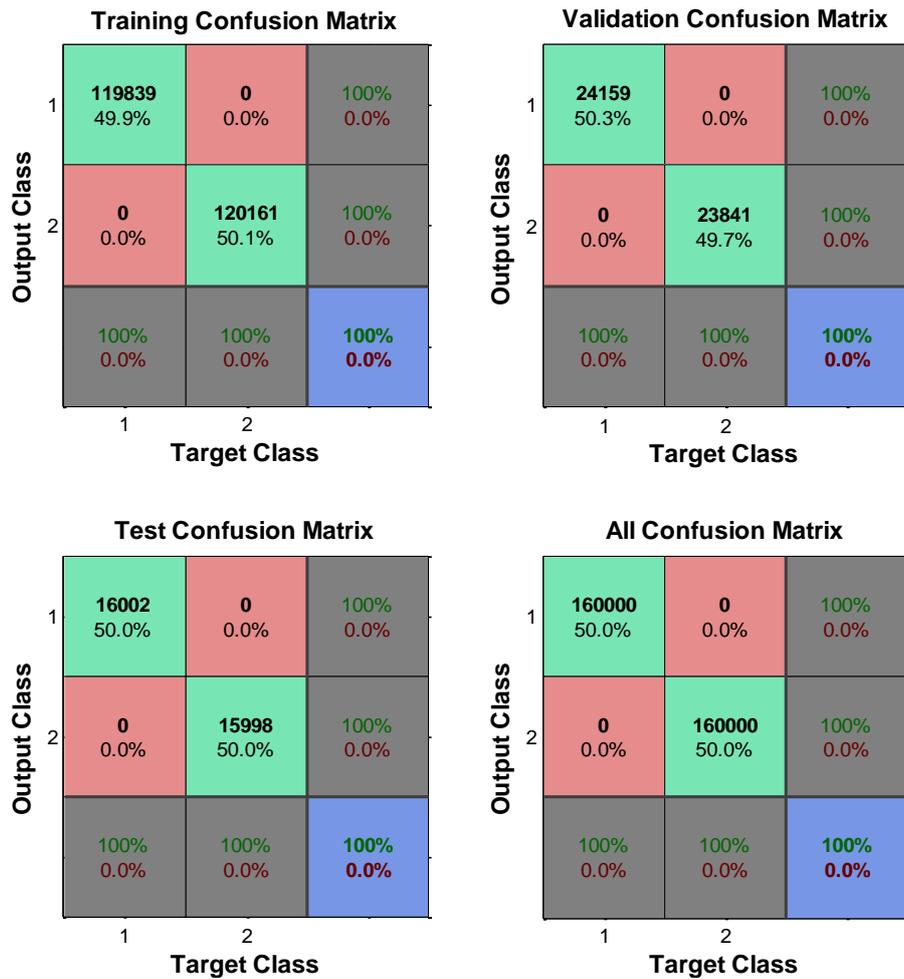
Figure 6.15   Confusion matrices obtained from midpoint classification process

Table 6.4   Classification results

	Actual System	Classified System	Error
Tip point Classification	$\{1 \quad 0\}^T$	$\{0.9991 \quad 0.0009\}^T$	$MSE = 7.7204 \times 10^{-7}$
Midpoint Classification	$\{1 \quad 0\}^T$	$\{0.9988 \quad 0.0012\}^T$	$MSE = 1.4753 \times 10^{-6}$

Classification results show that nonlinear behaviors of the strips are best explained with cubic stiffness nonlinearity. After determining the type of nonlinearity, regression networks are created for parameter identification of the cubic stiffness and the other linear elements (Figure 6.16).
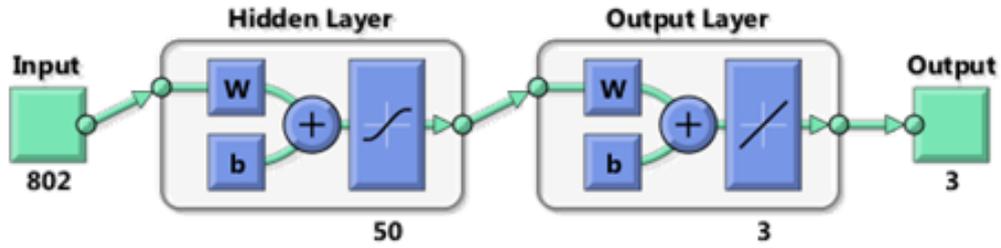
Figure 6.16   Regression network of the test rig

2 layered regression networks with 50 hidden layer neurons and 3 output layer neurons are trained with previously created input training data sets and corresponding target parameter vectors for coefficient of cubic stiffness, coefficient of unknown linear stiffness and unknown mass (Figure 6.16). Training operations are completed at the end of 788 and 1000 epochs for tip point and midpoint regressions. Regression results with correlation coefficients are given in Figure 6.17 and Figure 6.19. Trained regression networks are simulated with the frequency responses measured from the midpoint and the tip point. Simulation results give the parameters of identified elements in Table 6.5.

Table 6.5   Identification results

Parameter	Tip Point Identification	Midpoint Identification
$k_c \ \left(\mathrm{N/m}^3 \right)$	6.346×10^8	5.860×10^8
$k_{unknown} \ \left(\mathrm{N/m} \right)$	5.989×10^3	5.984×10^3
$m_{unknown} \ \left(\mathrm{kg} \right)$	0.0608	0.0612

Measured frequency response curves of tip point and midpoint of the cantilever beam are compared with the ones which are calculated based on the identified parameters in Figure 6.18 and Figure 6.20 respectively.
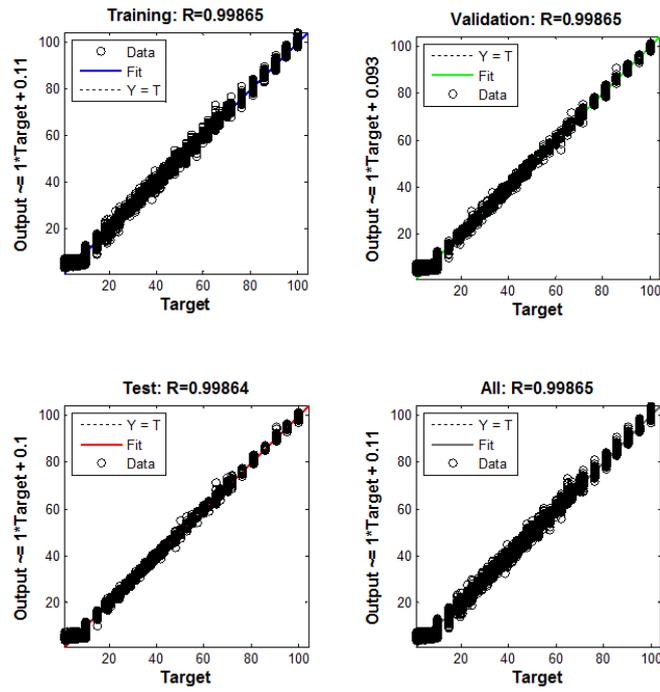
Figure 6.17   Regression plots of the tip point identification
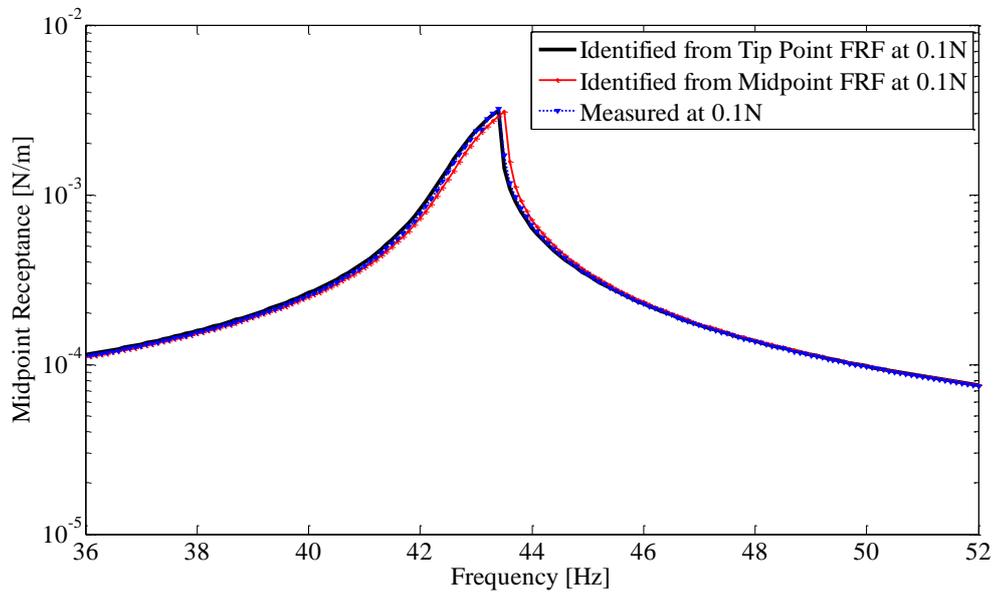
Figure 6.18   Comparison of calculated tip point frequency responses based on tip point and midpoint identified parameters and measured data under 0.1N harmonic excitation
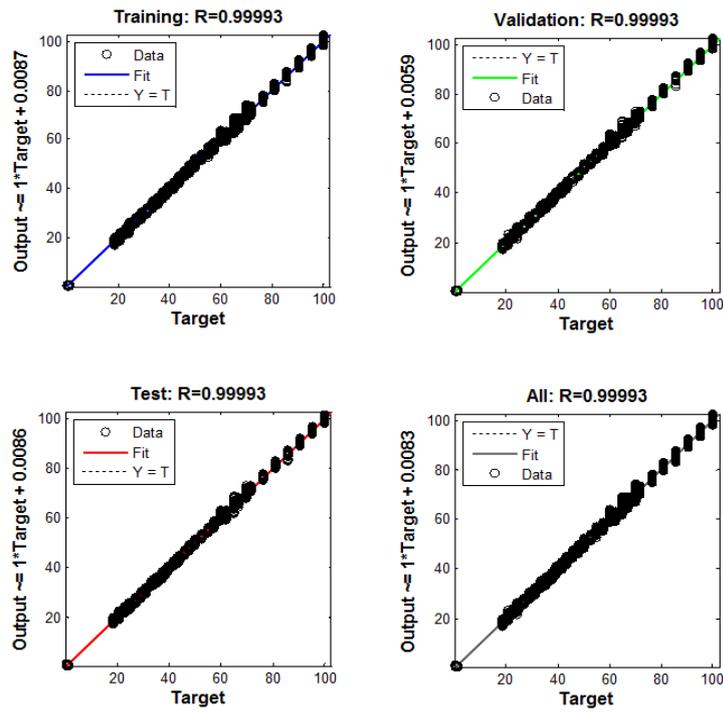
Figure 6.19   Regression plots of the midpoint identification
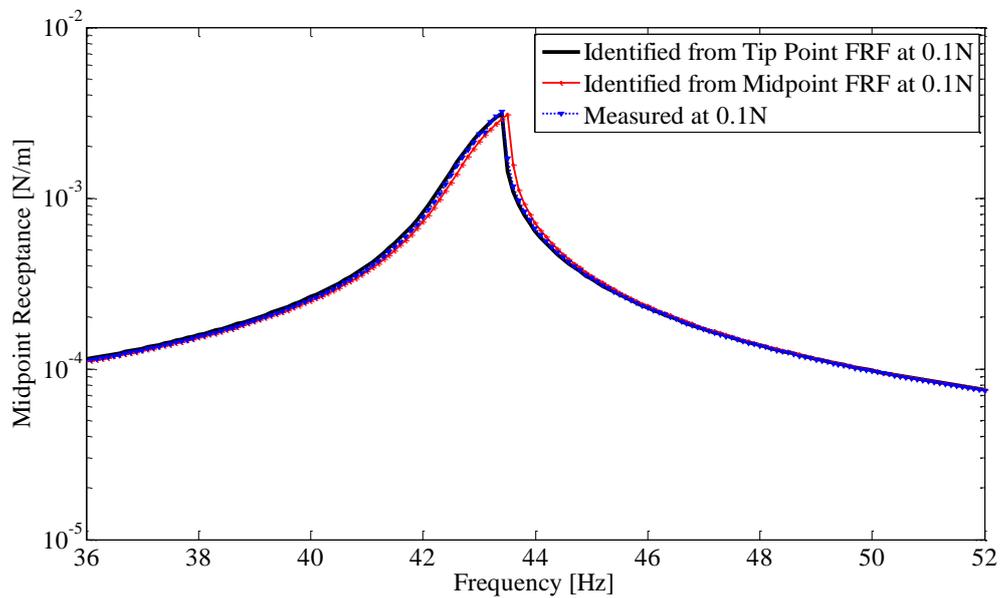
Figure 6.20   Comparison of calculated midpoint frequency responses based on tip point and midpoint identified parameters and measured data under 0.1N harmonic excitation

64

Comparing the parameters which are obtained midpoint and tip point identification processes, the unknown mass values and the unknown linear stiffness coefficients are close to each other. However there is still a slight difference between the coefficients of cubic stiffness. Final update is made on the parameters by performing an optimization based identification procedure. Starting point for the optimization is obtained from network identification results. Optimization bounds are determined from regression network errors given in Figure 6.21, Figure 6.22 and Figure 6.23. Parameters of the optimization problems are given in Table 6.6.

Table 6.6   Optimization problem setup

Global Optimization Toolbox Parameters	Values of Tip Point Optimization Based Identification	Values of Midpoint Optimization Based Identification
Solver	*ga*	*ga*
Fitness Function	*@optim_fun*	*@optim_fun*
Number Variables	3	3
Lower Bound	[5500 0.055 5.5e8]	[5800 0.057 4.5e8]
Upper Bound	[6500 0.065 7.1e8]	[6200 0.063 6.5e8]
Population Size	100	100
Initial Population	[5898 0.0608 6.34e8]	[5984 0.0611 5.47e8]
Fitness Scaling Function	*Rank*	*Rank*
Selection Function	*Stochastic Uniform*	*Stochastic Uniform*
Elite Count	10	10
Crossover Fraction	0.8	0.8
Mutation	*Constraint Dependent*	*Constraint Dependent*
Crossover Function	*Scattered*	*Scattered*
Migration Direction	*Forward*	*Forward*
Migration Fraction	0.2	0.2
Migration Interval	20	20
Stopping Criteria		
Generations	100	100
Time Limit	*Inf*	*Inf*
Fitness Limit	*-Inf*	*-Inf*
Stall Generations	50	50
Stall Time Limit	*Inf*	*Inf*
Function Tolerance	1e-14	1e-14
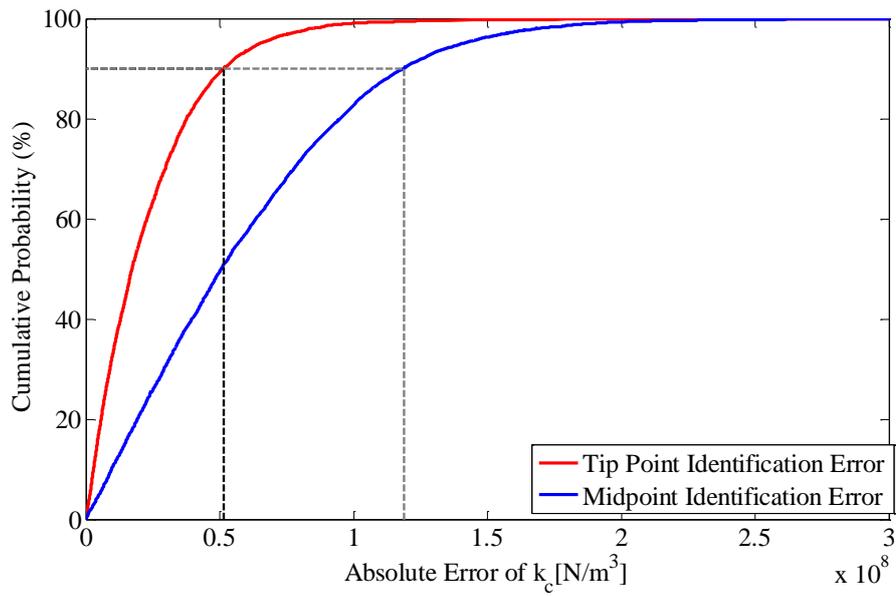
Figure 6.21   Absolute error of cubic stiffness coefficient with respect to cumulative probability for tip point and midpoint identification
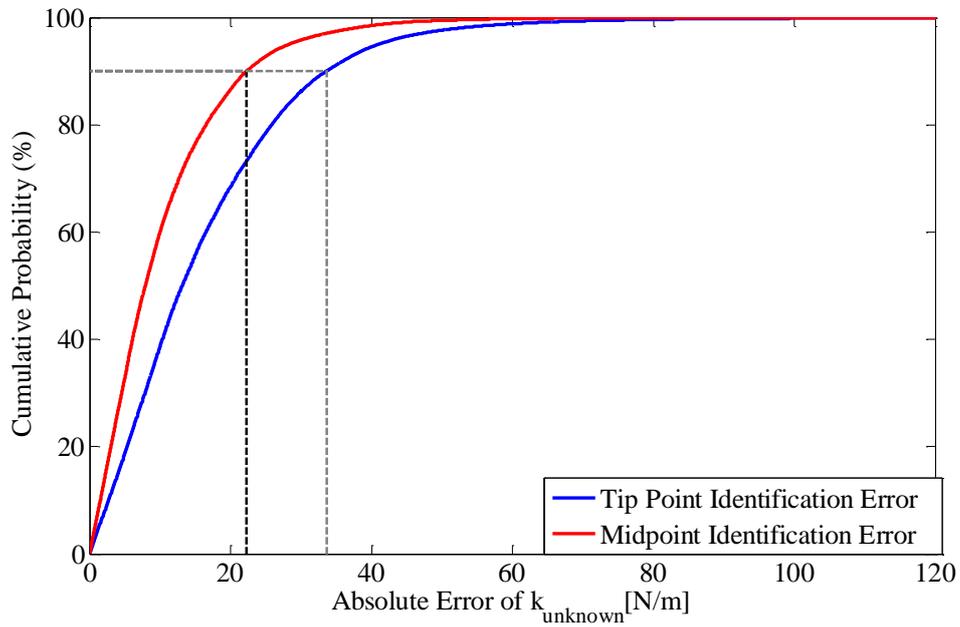
Figure 6.22   Absolute error of unknown stiffness coefficient with respect to cumulative probability for tip point and midpoint identification
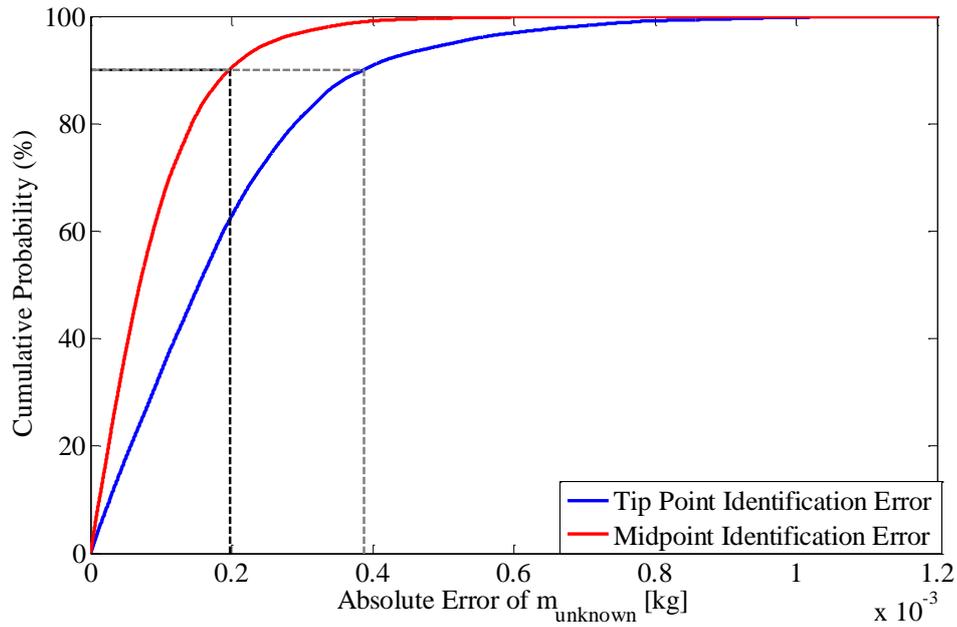
Figure 6.23   Absolute error of unknown mass with respect to cumulative probability for tip point and midpoint identification

Optimization based identification processes are successfully completed after satisfying desired function tolerance values given in Table 6.6. Optimized parameters are given in Table 6.7. Calculated frequency responses based on the optimized parameters and measured frequency responses under 0.1N harmonic excitation are shown in Figure 6.24 and Figure 6.25.

Table 6.7   Optimized parameters

Parameter	Tip Point Identification	Tip Point Optimization Based Identification	Midpoint Identification Based Identification	Midpoint Optimization Based Identification
$k_c \left(\text{N/m}^3 \right)$	6.346×10^8	5.518×10^8	5.860×10^8	5.476×10^8
$k_{unknown} \left(\text{N/m} \right)$	5.989×10^3	5.981×10^3	5.984×10^3	5.838×10^3
$m_{unknown} \left(\text{kg} \right)$	0.0608	0.0611	0.0612	0.0592
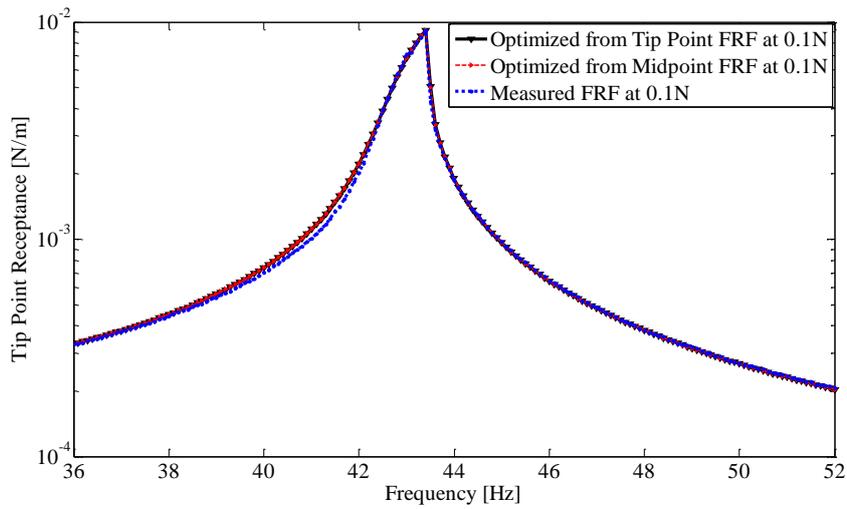
Figure 6.24   Comparison of calculated tip point frequency responses based on tip point and midpoint optimized parameters and measured data under 0.1N harmonic excitation

Using optimized parameters frequency responses of tip point and midpoint are calculated for 0.5N and 1N harmonic excitation. Results are compared with the measured frequency responses in Figure 6.26, Figure 6.27, Figure 6.28 and Figure 6.29.
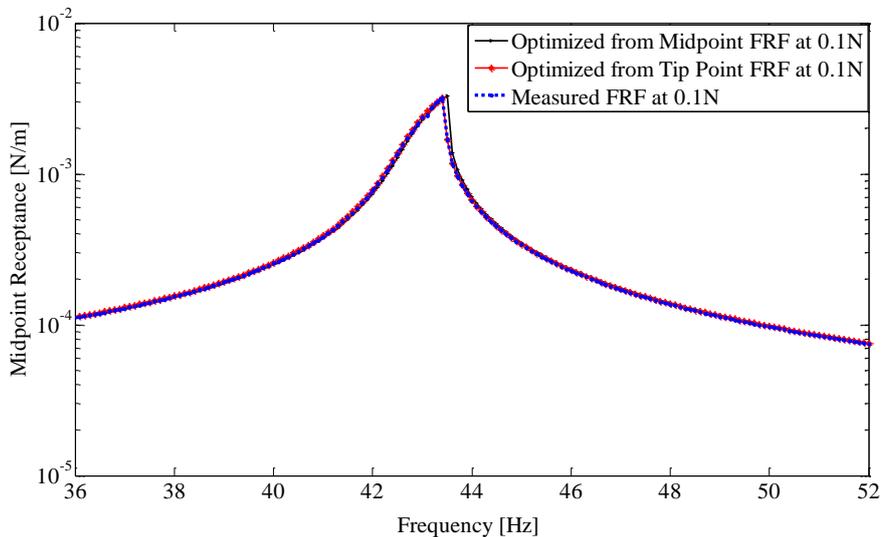
Figure 6.25   Comparison of calculated midpoint frequency responses based on tip point and midpoint optimized parameters and measured data under 0.1N harmonic excitation
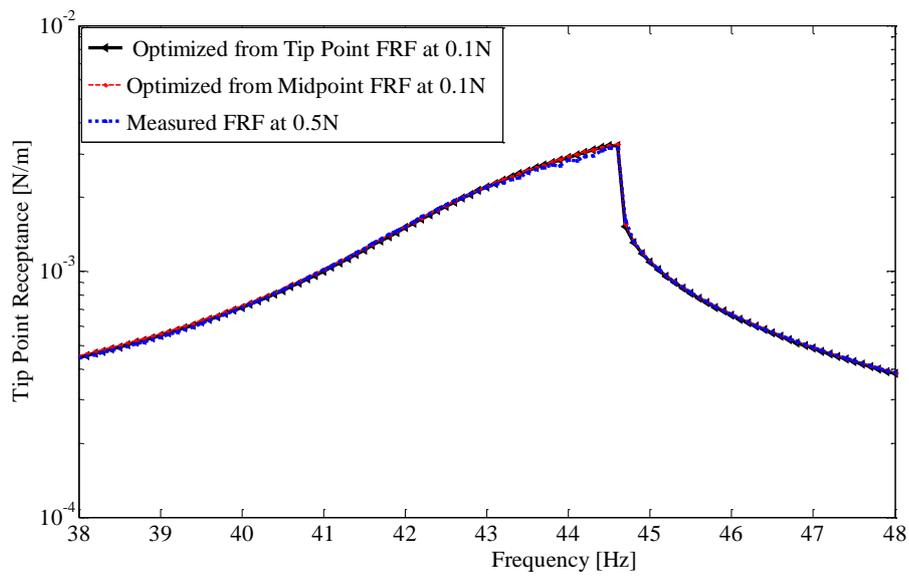
Figure 6.26   Comparison of calculated tip point frequency responses based on tip point and midpoint optimized parameters and measured data under 0.5N harmonic excitation
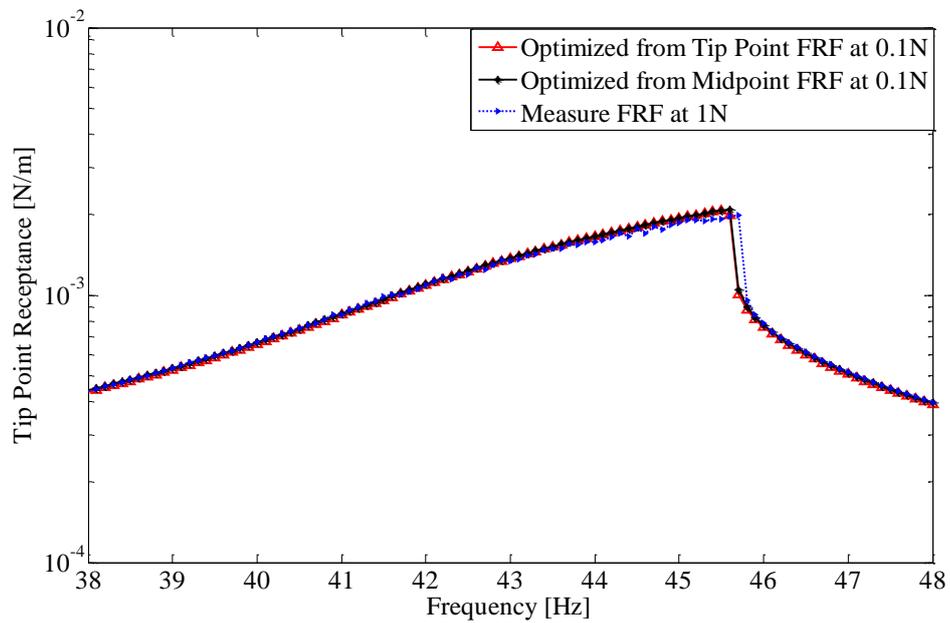
Figure 6.27   Comparison of calculated tip point frequency responses based on tip point and midpoint optimized parameters and measured data under 1N harmonic excitation
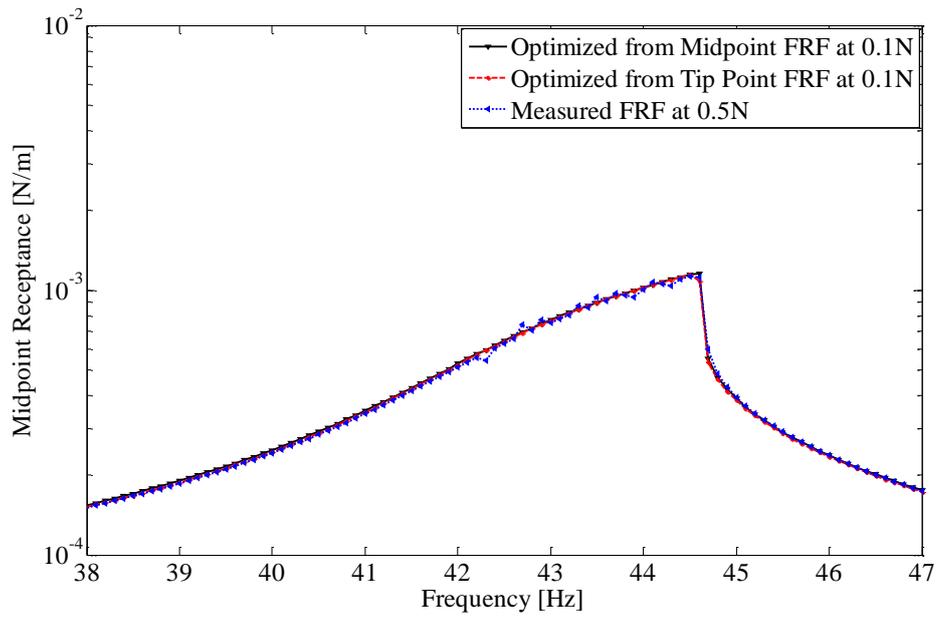
Figure 6.28   Comparison of calculated midpoint frequency responses based on tip point and midpoint optimized parameters and measured data under 0.5N harmonic excitation
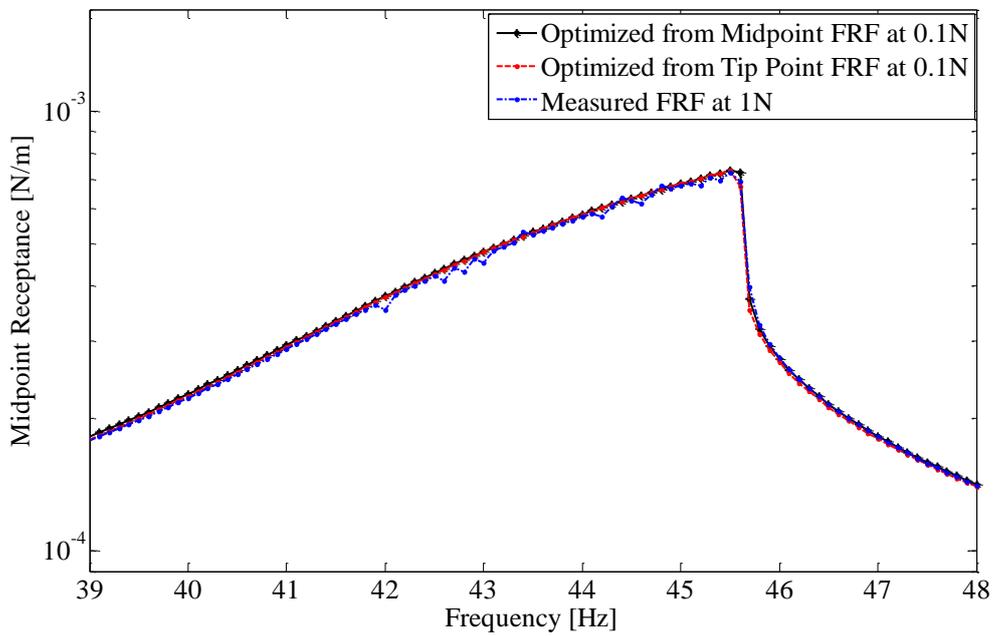
Figure 6.29   Comparison of calculated midpoint frequency responses based on tip point and midpoint optimized parameters and measured data under 1N harmonic excitation

CHAPTER 7

DISCUSSION AND CONCLUSIONS

In this study, artificial neural networks are used for nonlinearity classification and parameter identification. Proposed method has four main steps: generation of training data, training of the neural network, and identification by using the trained networks and optimization based identification.

Most of the effort is required in order to generate training data sets, which are frequency response sets of the system for all possible nonlinearity configurations and possible ranges of parameter values. Therefore, application of the proposed method to realistic finite element models requires reduction methods in order to decrease the computational time required for data generation process .A nonlinear solution method utilizing harmonic balance method with a single harmonic is used to obtain the required training data, which are employed in the training of the classification and regression networks. Numerical solution techniques are introduced to solve nonlinear algebraic equations obtained by using harmonic balance method.

Theory behind the regression networks are presented in this study. Basic elements form the network architectures are introduced and discussed. Appropriate layer numbers, transfer functions and training algorithm are proposed to attain desired identification performance.

Application of the method is demonstrated on a simple case study. A 4 DOF model with two different nonlinear elements at two different locations is identified successfully based on noisy measurements taken from a coordinate where none of the nonlinearities is connected. Identified parameters are improved by utilizing optimization based identification. Performance of the network architecture proposed in the theory part is compared with other alternative networks and the results show that the proposed one is superior to the others.

Finally, an experimental study is conducted to observe actual performance of the method on a real life system. Type determination and parameter estimation of the single nonlinear element are achieved on a test rig where location of the nonlinearity is known priori. Response measurements are taken from a point far from the nonlinear coordinate. Frequency responses of identified system obtained for single forcing level agree well with the actual system responses measured at different forcing levels.

Proposed method is capable of performing identification (classification and parameter identification) even the measurements are taken from DOFs where nonlinear elements are not connected. This is especially important, since before the classification process, locations of the nonlinearities are unknown. It should be noted that in order to increase the applicability of the proposed method, the generated data sets are polluted with random data in order to have a method that is prone to measurement noise.

Even though training data generation step is computationally demanding, once the networks are trained identification is very fast. Moreover, it is not required to re-train the networks for identification of different counterparts of the nonlinear system Therefore; the developed method is very a good candidate for online identification purposes.

In this study, complete system identification is achieved successfully using the proposed method. However, there are some aspects needed to be studied.

First of all, training data generation is one of the most crucial parts of the method. As the number of possible nonlinear systems increases, generation training data may exceed practical limits. A prior method should be utilized to reduce the possible nonlinear system configurations.

Moreover, measuring response of nonlinear systems by means of stepped sine testing is a time consuming process. The method should be improved such that faster testing techniques can be used to measure system response.

REFERENCES

[1] G. Kerschen, K. Worden, A. F. Vakakis, and J.-C. Golinval, "Past, present and future of nonlinear system identification in structural dynamics," *Mechanical Systems and Signal Processing*, vol. 20, no. 3, pp. 505–592, Apr. 2006.

[2] M. Aykan and H. Nevzat Özgüven, "Parametric identification of nonlinearity in structural systems using describing function inversion," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 356–376, Oct. 2013.

[3] D. Göge, M. Sinapius, U. Füllekrug, and M. Link, "Detection and description of non-linear phenomena in experimental modal analysis via linearity plots," *International Journal of Non-Linear Mechanics*, vol. 40, no. 1, pp. 27–48, Jan. 2005.

[4] D. E. Adams and R. J. Allemang, "Survey of Nonlinear Detection and Identification Techniques for Experimental Vibrations," in *International Seminar on Modal Analysis (ISMA 23)*, 1988.

[5] G. Kerschen, K. Worden, A. F. Vakakis, and J. C. Golinval, "Nonlinear System Identification in Structural Dynamics : Current Status and Future Directions," in *4th International Modal Analysis Conference*, 1986, no. 1, pp. 712–719.

[6] M. B. Özer, H. N. Özgüven, and T. J. Royston, "Identification of structural non-linearities using describing functions and the Sherman–Morrison method," *Mechanical Systems and Signal Processing*, vol. 23, no. 1, pp. 30–44, Jan. 2009.

[7] S. F. Masri and T. K. Caughey, "A Nonparametric Identification Technique for Nonlinear Dynamic Problems," *Journal of Applied Mechanics*, vol. 46, no. 2, p. 433, 1979.

[8] J. He and D. J. Ewins, "A Simple Method of Interpretation for the Modal Analysis of Non-linear Systems," in *5th International Modal Analysis Conference*, 1987, pp. 626–634.

[9] K. Worden and G. R. Tomlinson, *Nonlinearity in Structural Dynamics*. London: IOP Publishing Ltd, 2001.

[10] Ö. Arslan, M. Aykan, H. Nevzat Özgüven, and H. N. Özgüven, "Parametric identification of structural nonlinearities from measured frequency response data," *Mechanical Systems and Signal Processing*, vol. 25, no. 4, pp. 1112–1125, May 2010.

[11] M. Aykan and H. N. Özgüven, "Parametric Identification of Nonlinearity from Incomplete FRF Data Using Describing Function Inversion," in *Proceedings of SEM IMAC XXX Conference*, 2012, pp. 323–334.

[12] D. E. Adams and R. J. Allemang, "a Frequency Domain Method for Estimating the Parameters of a Non-Linear Structural Dynamic Model Through Feedback," *Mechanical Systems and Signal Processing*, vol. 14, no. 4, pp. 637–656, Jul. 2000.

[13] A. Chatterjee and N. S. Vyas, "Non-linear parameter estimation in multi-degree-of-freedom systems using multi-input Volterra series," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 457–489, May 2004.

[14] C. M. Richards and R. Singh, "Identification of Multi-Degree-of-Freedom Non-Linear Systems Under Random Excitations By the 'Reverse Path' Spectral Method," *Journal of Sound and Vibration*, vol. 213, no. 4, pp. 673–708, Jun. 1998.

[15] P. A. Atkins, J. R. Wright, and K. Worden, "An Extension of Force Appropriation To the Identification of Non-Linear Multi-Degree of Freedom Systems," *Journal of Sound and Vibration*, vol. 237, no. 1, pp. 23–43, Oct. 2000.

[16] S. F. Masri, A. G. Chassiakosl, and T. K. Caugheys, "Structure-unknown non-linear dynamic systems : identification through neural networks," vol. 1, pp. 45–56, 1992.

[17] S. L. Xie, Y. H. Zhang, C. H. Chen, and X. N. Zhang, "Identification of nonlinear hysteretic systems by artificial neural network," *Mechanical Systems and Signal Processing*, vol. 34, no. 1–2, pp. 76–87, Aug. 2013.

[18] S. Chen and S. a. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *International Journal of Control*, vol. 56, no. 2, pp. 319–346, Aug. 1992.

[19] Y. C. Liang, D. P. Feng, and J. E. Cooper, "Identification of Restoring Forces in Non-Linear Vibration Systems Using Fuzzy Adaptive Neural Networks," *Journal of Sound and Vibration*, vol. 242, no. 1, pp. 47–58, Apr. 2001.

[20] S. Lu, A. Narayan, and A. Mehrotra, "Continuation Method in Multitone Harmonic Balance," in *ISCAS '04*, 2004, pp. 520–523.

[21] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*. Society for Industrial and Applied Mathematics, 1995, pp. 65–68.

[22]   E. Ciğeroğlu and H. N. Özgüven, "Nonlinear vibration analysis of bladed disks with dry friction dampers," *Journal of Sound and Vibration*, vol. 295, no. 3–5, pp. 1028–1043, Aug. 2006.

[23]   R. L. Burden, J. D. Faires, and A. C. Reynolds, *Numerical Analysis*, 8th ed., vol. 87, no. 3. Thomson Learning Inc., 2005, pp. 598–606.

[24]   J. H. Mathews and K. D. Fink, *Numerical Methods using Matlab*, 4th ed. Pearson Prentice Hall, 2004, pp. 70–90.

[25]   E. Cigeroglu and H. Samandari, "Nonlinear free vibration of double walled carbon nanotubes by using describing function method with multiple trial functions," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 46, pp. 160–173, Sep. 2012.

[26]   C. Boral, E. Ciğeroğlu, and İ. Korkmaz, "Effect of Intentional Dry Friction Damping On the Performance of an Elastomeric Engine Mount," in *Proceedings of the ASME 2010 10th Biennial Conference on Engineering Systems Design and Analysis*, 2010, p. 59.

[27]   M. E. Yümer, "On The Nonlinear Vibration and Mistuning Identification of Bladed Disks," Middle East Technical University, 2010.

[28]   G. Orbay, "Nonlinear Vibrations of Mistuned Bladed Asseblies," Middle East Technical University, 2008.

[29]   G. Orbay and H. N. Özgüven, "Non-linear periodic response analysis of mistuned bladed disk assemblies in modal domain," in *9th International Conference on Vibrations in Rotating Machinery*, 2008, pp. 159–170.

[30]   M. A. Crisfield, *Non-linear Finite Element Analysis of Solids and Structures, Volume 1: Essentials*. John Wiley & Sons Ltd., 1991.

[31]   M. A. Crisfield, "A Fast Incremntal/Iterative Solution Procedure That Handles 'Snap-Through'," *Computers & Structures*, vol. 12, pp. 55–62, 1980.

[32]   P. G. Bergan and T. Soreide, "Solution of large displacement and instability problems using the current stiffness parameter," in *Finite Elements in Nonlinear Mechanics*, vol. 3, 1978, pp. 647–669.

[33]   A. De Niet, "Step-size control and corrector methods in numerical continuation of ocean circulation and fill-reducing orderings in multilevel ILU methods," University of Groningen, 2002.

[34]   C.-H. Menq, J. H. Griffin, and J. Bielak, "The influence of microslip on vibratory response, Part II: A comparison with experimental results," *Journal of Sound and Vibration*, vol. 107, no. 2, pp. 295–307, 1986.

[35] M. A. Fischler and O. Firschein, *Intelligence: The Eye, the Brain, and the Computer*. 1987, pp. 24–55.

[36] W. S. McCulloch and W. H. Pitts, "A Logical Calculus of the Ideas Immanent In Nervous Actıvıty," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943.

[37] H. Yonaba, F. Anctil, and V. Fortin, "Comparing Sigmoid Transfer Functions for Neural Network," no. April, pp. 275–283, 2010.

[38] I. S. Isa, Z. Saad, S. Omar, M. K. Osman, K. a. Ahmad, and H. a. M. Sakim, "Suitable MLP Network Activation Functions for Breast Cancer and Thyroid Disease Detection," in *Second International Conference on Computational Intelligence, Modelling and Simulation*, 2010, pp. 39–44.

[39] H. Demuth and M. B. M. Hagan, *Neural Network Toolbox ^{TM} 6 User' s Guide*. The Mathworks Inc., 2009, pp. 15–19.

[40] G. Hiiang and H. A. Babri, "General Approximation Theorem on Feedforward Networks," in *International Conferance on Information, Communications and Signal Processing*, 1997, no. September, pp. 698–702.

[41] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox ^{TM} User ' s Guide*. The MathWorks, Inc., 2013, pp. 1–14.

[42] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox^{TM} User' s Guide*. 2013, pp. 2–31, 8–17.

[43] G. J. Bowden, H. R. Maier, and G. C. Dandy, "Optimal division of data for neural network models in water resources applications," *Water Resources Research*, vol. 38, no. 2, pp. 2.1–2.11, 2002.

[44] M. A. Shahin, H. R. Maier, and M. B. Jaksa, "Engineering, Evolutionary Data Division Methods for Developing Artificial Neural Network Models in Geotechnical," 2000.

[45] P. S. Crowther and R. J. Cox, "A Method for Optimal Division of Data Sets for Use in Neural Networks," *KES 2005*, vol. LNAI 3684, pp. 1–7, 2005.

[46] M. A. Shahin, H. R. Maier, and M. B. Jaksa, "Data Division for Developing Neural Networks Applied to Geotechnical Engineering," *Journal of Computing In Civil Engineering*, vol. 18, no. 2, pp. 105–114, 2004.

[47] J. V. Ferreira, "Dynamic Response Analysis of Structures with Nonlinear Components," mperial College of Science, Technology and Medicine, 1998.

[48]   H. R. E. Siller, "Non-Linear Modal Analysis Methods for Engineering," Imperial College London / University of London, 2004.

[49]   M. Aykan, "Identification of Localized Nonlinearity for Dynamic Analysis of Structures," Middle East Technical University, 2013.

[50]   S. Orlando, B. Peeters, and G. Coppotelli, "Improved FRF estimators for MIMO Sine Sweep data," in *ISMA 2008*, 2008, no. September 2008.

PUBLISHED PAPERS

LOCALIZATION AND IDENTIFICATION OF STRUCTURAL NONLINEARITIES USING NEURAL NETWORKS[*]

A. Koyuncu[1, 2], E. Cigeroglu[1], M. E. Yumer[1, 3], H. N. Özgüven[1]

[1]*Middle East Technical University, 06800 Ankara, TURKEY*
[2]*Defense Systems Technologies Division, ASELSAN A.Ş., Ankara 06172, TURKEY*
[3]*Carnegie Mellon University, Pittsburgh, PA, USA*

ABSTRACT

In this study, a new approach is proposed for identification of structural nonlinearities by employing neural networks. Linear finite element model of the system and frequency response functions measured at arbitrary locations of the system are used in this approach. Using the finite element model, a training data set is created, which appropriately spans the possible nonlinear configurations space of the system. A classification neural network trained on these data sets then localizes and determines the type of nonlinearity associated with the corresponding degree of freedom in the system. A new training data set spanning the parametric space associated with the determined nonlinearities is created to facilitate parametric identification. Utilizing this data set, a feed forward regression neural network is

---

trained, which parametrically identifies the related nonlinearity. The application of the proposed approach is demonstrated on an example system with nonlinear elements. The proposed approach does not require data collection from the degrees of freedoms related with nonlinear elements, and furthermore, the proposed approach is sufficiently accurate even in the presence of measurement noise.

1. INTRODUCTION

Identification of structural nonlinearities in dynamic structures has become the interest of researchers in the past four decades [1]. Studies on this subject focused on two parts: localizing and characterizing the nonlinearity, and estimating the parameters of the nonlinearity based on experimentally measured data [2, 6].

One of the main problems of nonlinearity identification is the determination of the location and the type of nonlinearity; in other words, classification of the nonlinear system from system responses. For this problem, frequency domain methods are mostly preferred He and Ewins [7] used frequency response functions (FRFs) obtained at different forcing levels in order to detect nonlinearities in a system. Similarly, the method developed by Özer and Özgüven [4] determines possible locations of nonlinearities and identifies their types and parameters by using describing functions, which is as well a frequency domain method. Major restrictions of this method are the requirements of measurements collected from all degrees of freedoms (DOFs) and complete FRFs of the system for localization purposes. However, in a later study, Aykan *et al.* [5] improved the method using incomplete FRFs for localization purposes. A common feature of most of the methods in literature is the identification of types of nonlinearities by observing system responses or nonlinear restoring forces, which is a time consuming process and it is not suitable for identification of nonlinearity in a series of products due to manufacturing errors and assembly differences.

Parameter estimation of nonlinearities in a classified system is an easily handled problem. There exist time domain and frequency domain methods targeting parameter identification in literature [1, 8, 9]. Although these methods have their own handicaps, they are promising in specific application areas. Masri and Caughey [10] developed a method named as *Restoring Force Surface* (RFS) method to estimate parameters of nonlinearities using least squares approach in time domain. *Conditioned Reverse Path* (CRP) proposed by Richards and Singh [11] uses spectral analysis in frequency domain to compute the coefficients of nonlinearity matrix.

The method proposed in this study, classifies and identifies nonlinearities associated with the system utilizing neural networks. Neural network classifiers have been used for parameter identification of structural nonlinearities [12, 13]; however, nonlinearity classification using neural networks is a new application in structural dynamics. Constructed networks are to be trained with sample data sets which are frequency response functions of selected points on the system. Nonlinear frequency responses of the system are obtained by analyzing possible nonlinear system configurations. It should be noted that in order to perform these analyses, a mathematical model of the system is required. After training the networks, location and type of the nonlinearity is determined by running trained networks with measured systems responses as the input. Parameters of the classified nonlinearities are identified by means of a regression network utilizing the same input data used for classification. The proposed method does not require taking measurements from the nonlinear DOFs; hence, system responses measured

from arbitrary locations can be used for identification purposes. Moreover, with the proposed method, it is as well possible to handle measurement noise by injecting noise into data during the training process. Once networks are trained, identification is very fast; hence, the method proposed is very suitable for identification of nonlinearities in a series of products due to manufacturing errors and assembly differences.

2. THEORY

The proposed method is composed of two parts: in the first part, a classification network is used to identify the locations and the types of the nonlinear elements, whereas in the second part, a regression network is utilized to determine the parameters associated with the identified nonlinearities. For training of the networks, forced response function of an arbitrary degree of freedom is used; therefore, a mathematical model of the system is required. There is no restriction on the mathematical model; hence finite element models (FEMs) can as well be used with the proposed method. Using the mathematical model together with the nonlinear elements and their possible configurations, simulations are performed to generate the training data by using harmonic balance method (HBM), which is a frequency domain method. Details of the neural network models and the nonlinear mathematical modeling are explained in the following sections.

2.1 Neural Network Models

A typical neural network model is characterized by four items: number of neurons, number of layers, transfer functions and training algorithm. The fundamental element of a neural network is *a single neuron* mathematically represented below:

$$o_j = f_j \left(\sum_{i=1}^{R} w_{ji} y_i + b_j \right) = f_j \left(\{w_j\}^T \{y\} + b_j \right) = f_j \left(n_j \right) \tag{1}$$

$\{y\}$ is the input vector with R number of elements, $\{w_j\}^T$ is the jth row of the weight matrix $[W]$, b_j is the bias term, o_j is the output of the jth neuron, n_j is the net input and $f_j\left(n_j\right)$ is the transfer function used. Combination of multiple neurons is named as *a layer* and it can be written in matrix form as:

$$\{o\} = \left\{ f\left([W]\{y\} + \{b\} \right) \right\} = \left\{ f\left(\{n\} \right) \right\}. \tag{2}$$

Weight matrix [W] is a $S \times R$ matrix where S is number of neurons in a layer. As a result, dimension of the output vector $\{o\}$ is $S \times 1$. A network can be extended by increasing the number of hidden layers used, which are referred as multilayer networks. A two-layer network is given as an example in Figure 1. It can be observed that output of each layer becomes the input of the next layer. Dimensions of weight matrices and bias vectors can be adjusted according to the problem requirements except R and S_2 values which are determined by

81

length of input and target vectors. In this study, a two-layer network is used for both classification and identification purposes.
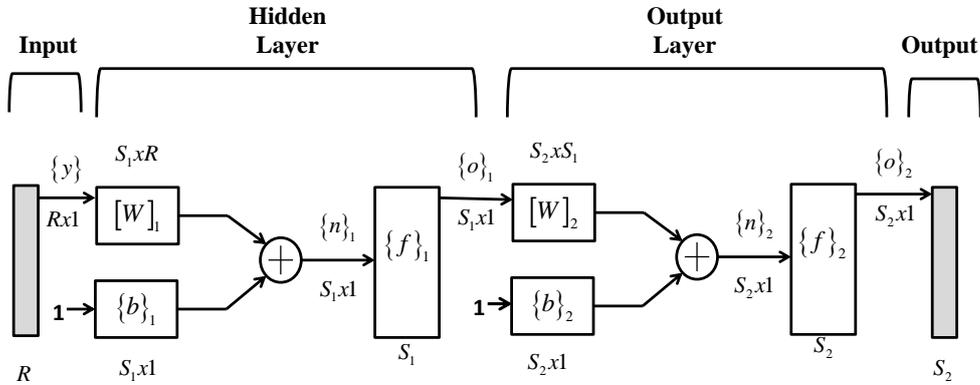
Figure 1 A two-layer network

Transfer functions play an important role in the structures of neural networks. Selection of a transfer function depends on input and output vectors and training algorithm. Commonly used transfer functions in neural network models are given in Table 1. In the classification network used in this study, transfer functions of Layer 1 and Layer 2 are both Hyperbolic-Tangent Sigmoid function; whereas, Hyperbolic-Tangent Sigmoid and Linear transfer functions are used in parameter identification networks for Layer 1 and Layer 2, respectively.

Table 1 Commonly Used Transfer Functions

Name	Input/Output Relation	Name	Input/Output Relation
Hard Limit	$a = 0 \quad n < 0$ $a = 1 \quad n \geq 0$	Symmetrical Saturated Linear	$a = -1 \quad n < -1$ $a = n \quad -1 \leq n \leq 1$ $a = 1 \quad n > 1$
Symmetrical Hard Limit	$a = -1 \quad n < 0$ $a = +1 \quad n \geq 0$	Log-Sigmoid	$a = \dfrac{1}{1 + e^n}$
Linear	$a = n$	Hyperbolic-Tangent Sigmoid	$a = \dfrac{e^n - e^{-n}}{e^n + e^{-n}}$
Saturated Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n \leq 1$ $a = 1 \quad n > 1$	Positive Linear	$a = 0 \quad n < 0$ $a = n \quad 0 \leq n$

Performance of a network is quantified through mean squared error (MSE) between the network output vector, $\{o\}$, and target vector, $\{t\}$ as:

$$MSE = \frac{1}{N} \int_{i=1}^{N} (e_i)^2 = \frac{1}{N} \int_{i=1}^{N} (o_i - t_i)^2 \ . \tag{3}$$

Objective of a neural network problem is to minimize MSE via tuning the elements of weight matrices and bias vectors. The tuning process is called as training, where weight matrices and bias vectors are updated according to *"the training algorithm"*. In this study, Scaled Conjugate Gradient Backpropagation algorithm is used for nonlinearity localization and classification; whereas, Levenberg-Marquardt algorithm is used for identification of parameters of the nonlinear elements [14]. Before starting to the training process, collected data is required to be divided into three subsets as: *training*, *validation* and *test* data sets. Weight matrices and bias vectors are updated based on the training data set; whereas, network error based on the validation data set is used as the stopping criteria in order to prevent over fitting. Although the test data set is not used in training, errors based on the test data set are useful in post processing. In this study, 75% of samples are allocated as training data set, 15% of samples are allocated as the validation data set and the rest is used as the test data set.

2.2 Mathematical Modeling

Before starting the training process, training data should be generated first. Training data consists of input vectors and corresponding target vectors. Input vectors are outputs of the physical system in both classification and identification problems; whereas, target vectors are the location and type of the nonlinearity in the classification problem, and parameters of the nonlinearity in the identification problem. A simulated data set which covers full range of possible input space is required as input data. In order to obtain this, nonlinear frequency responses of all selected system configurations are required to be determined. Equation of motion of a nonlinear system can be written as follows

$$[M]\{\ddot{x}\} + [C]\{\dot{x}\} + i[H]\{x\} + [K]\{x\} + \{f_{NL}(\{x\},\{\dot{x}\},...)\} = \{f(t)\}, \tag{4}$$

where $[M]$, $[C]$, $[H]$ and $[K]$ are mass, viscous damping, structural damping and stiffness matrices of linear system, respectively. In this equation, $\{x(t)\}$, $\{f_{NL}(t)\}$ and $\{f(t)\}$ are response, nonlinear force and excitation force vectors of the system, respectively. For harmonic forcing, excitation force vector can be written as:

$$\{f(t)\} = \{F\}e^{i\omega t} \ . \tag{5}$$

For harmonic input, it can be assumed that system response is also harmonic, which can be written as follows

$$\{x(t)\} = \{X\}e^{i\omega t} \ , \tag{6}$$

using a single harmonic term. Substituting Eq. (5) and Eq. (6) into Eq. (4), the following equation is obtained:

$$\left([K]-\omega^2[M]+i\omega[C]+i[H]\right)\{X\}+\{F_{NL}(\{X\})\}=\{F\} \,. \tag{7}$$

The nonlinear forcing can be written using a single term Fourier series as follows

$$\{F_{NL}(\omega)\}=\frac{i}{\pi}\int_0^{2\pi}f_{NL}(\psi)e^{i\psi}\,d\psi \,, \tag{8}$$

where $\psi=\omega t$.

Equation (7) can be divided into linear and nonlinear parts, which decreases the total number of nonlinear equations to be solved, by using the receptance method developed by Menq et al. [15]. Receptance method employed in this study decreases the computational time required for the solution of nonlinear equations obtained, which makes the proposed method suitable for realistic finite element models. The resulting nonlinear equations can be solved by a nonlinear equation solver. In this study, Newton's method with Arc-Length Continuation [16, 17] is used to solve the resulting system of nonlinear equations.

Training process is to be performed using the frequency response of one of the coordinates, which does not need to be measured from the nonlinear DOFs. In order to include the effect of measurement noise, input vectors are polluted by addition of random numbers having Gaussian distribution with zero mean and 0.05mm standard deviation. Input training data is generated from frequency response of the nonlinear system as follows:

$$\{y\}=\left\{\mathrm{Re}\{X(\omega_1)\quad X(\omega_2)\quad\cdots\quad X(\omega_n)\}\ \vdots\ \mathrm{Im}\{X(\omega_1)\quad X(\omega_2)\quad\cdots\quad X(\omega_n)\}\right\}^T \tag{9}$$

Target vectors define the configurations of the nonlinear system and the parameter values of the nonlinear elements in classification and identification problems, respectively. For classification networks, length of the target vectors are determined from the total number of possible nonlinear system configurations; therefore, orthogonality of target vectors can be achieved easily.

After obtaining the input and target vectors, the classification and identification networks are ready trained. In Figure 2, flowchart of the proposed method is explained schematically.
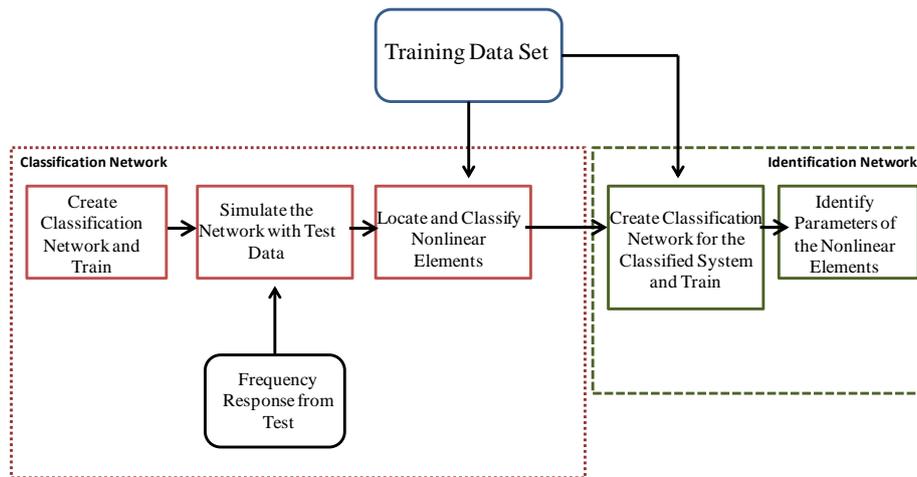
Figure 2 Flowcharts of Nonlinearity Classification and Parametric Identification

3. CASE STUDY: 4-DOF NONLINEAR SYSTEM

In this section, application of the proposed approach is presented on a simple 4-DOF system with local nonlinearities. From the physics of the problem or experiences it is possible to localize nonlinearities; hence, in this case study, the possible locations of the nonlinear elements are identified as between the ground and the first mass and between the third and the fourth masses as shown in Figure 3. Moreover, from the physics of the problem or experience it is known that hardening type cubic stiffness nonlinearity and symmetric gap nonlinearity are the possible nonlinearities that can be encountered (Figure 4). This yields 8 possible nonlinear system configurations as indicated in Table 2 with their corresponding classification network targets.
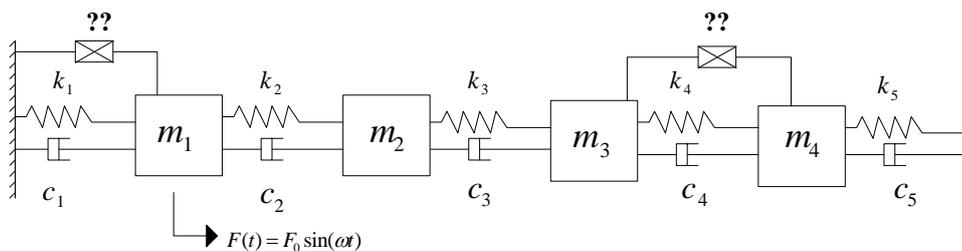

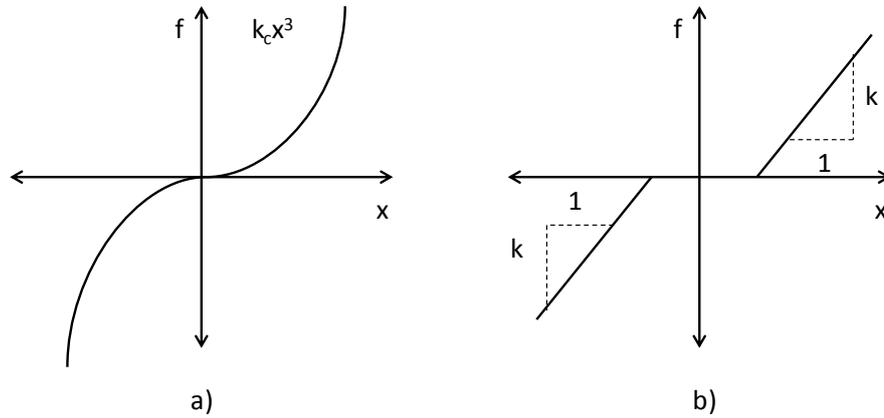**Figure 3** 4-DOF Non-Linear System Schematic View

Figure 4 Restoring Force Diagram of Possible Nonlinearities a) Cubic Stiffness, b) Symmetric Gap Nonlinearity

Table 2 Possible Nonlinear System Configurations and Corresponding Classification Network Targets

Conf No.	DOF 1 and Ground	DOF 3 and DOF 4	Binary Network Targets
1	Hardening Cubic Stiffness	Linear Stiffness with Backlash	$\{1,0,0,0,0,0,0,0\}^T$
2	Linear Stiffness with Backlash	Hardening Cubic Stiffness	$\{0,1,0,0,0,0,0,0\}^T$
3	Hardening Cubic Stiffness	Hardening Cubic Stiffness	$\{0,0,1,0,0,0,0,0\}^T$
4	Linear Stiffness with Backlash	Linear Stiffness with Backlash	$\{0,0,0,1,0,0,0,0\}^T$
5	Hardening Cubic Stiffness	Linear	$\{0,0,0,0,1,0,0,0\}^T$
6	Linear	Hardening Cubic Stiffness	$\{0,0,0,0,0,1,0,0\}^T$
7	Linear Stiffness with Backlash	Linear	$\{0,0,0,0,0,0,1,0\}^T$
8	Linear	Linear Stiffness with Backlash	$\{0,0,0,0,0,0,0,1\}^T$

A two-layered classification network is created to determine the location(s) and type(s) of the nonlinearities. Numbers of neurons used in the hidden and the output layers are 50 and 8 respectively. At this stage input and target data sets are generated for training purposes. Nonlinear frequency response of the second DOF, which is not connected to any nonlinear DOF, is calculated under a 10 N harmonic excitation force applied at the first DOF in order to generate the input training data. Frequency range of interest is considered as 0 to 8 Hz with increments of 0.01 Hz. For the simulations, the ranges of parameter values are selected as: 5×10^4 N/m^3 to 2×10^6 N/m^3 for cubic stiffness coefficient, and 2.5×10^{-4} m to 0.01 m gap for the symmetric gap nonlinearity with a known linear stiffness of 500 N/m. For each nonlinear

system configuration about 1600 samples a total of 12708 data sets are created. Training input vectors are polluted by normally distributed random numbers with zero mean and 0.5 mm standard deviation, representing measurement noise. After dividing the samples into training, validation and test data; training operation is performed, which is completed after 198 epochs. Confusion matrix presented in Figure 5 shows that 100% classification of nonlinearities in the system is achieved as indicated by the last column or row of the confusion matrix for training, validation and test sets.

After determining location(s) and type(s) of the nonlinearities, parameters of the nonlinear elements are identified by means of a regression network. Eight regression networks are created for all nonlinear system configurations. For all nonlinear system configurations, 50 neurons are used in the hidden layer; whereas for the output layers 2 neurons and 1 neuron are used for configurations 1 to 4 and 5 to 8, respectively, since there is only one parameter identified in configurations 5 to 8 . Input data sets generated for the classification network can as well be used in the regression network. However, in this case, target vectors are the unknown parameters of the nonlinear elements corresponding to each input data vector. Using these data and the identified configuration, regression networks are trained similar to the classification network. Correlation coefficients of the resulted regression networks are presented in Table 3.

As an example, a system with a hardening cubic stiffness with a coefficient of 10^6 N/m attached between the first DOF and ground and a symmetric gap element with a stiffness of 500 N/m and a backlash of 0.005 m between the third and the fourth DOFs is considered. In order to simulate the experimentally measured frequency response from the second DOF, time domain solution of the nonlinear system is obtained by using ODE45 solver of MATLAB increasing frequency from 0 to 8 Hz with 0.01Hz increments. Steady state oscillation amplitudes are collected as the frequency response values at the corresponding frequency. In order to represent the effect of measurement noise, data obtained from time marching is polluted by adding random numbers having a normal distribution with zero mean and 0.5 mm standard deviation. The resulting frequency response is given in Figure 6, which is used as the input to the trained networks. Classification and identification process is also repeated by considering noise free response data and. The results obtained for noisy and noise free cases are presented in Table 4. Comparison of identified and actual system responses are presented for noisy and noise free measurement cases in Figure 7 and Figure 8, respectively. It should be noted that, in the identification process a single harmonic HBM is used; whereas, in the simulated measurement data, time marching method is utilized. Therefore, even though identification is done by using a single harmonic, identified time domain simulation of the system shows that the identified system response is very close to the actual system for both noisy and noise free measurement cases. However, it should be noted that since networks are trained with noisy data, there is a slight increase in the error for parameter identification of coefficient of cubic stiffness if noisy free measurement is used. On the other hand, the error obtained for gap parameter decreases if noise free measurement is used.

4. CONCLUSION

In this study, artificial neural networks are used for nonlinearity classification and parameter identification. Proposed method has three main steps: generation of training data, training of the neural network and identification by using the trained networks. Application of the method is demonstrated on a simple case study. Most of the effort is required in order to generate training data sets, which are frequency response sets of the system for all possible nonlinearity configurations and possible ranges of parameter values. Therefore, application of the proposed method to realistic finite element models requires reduction methods in order to decrease the computational time required for data generation process. A nonlinear solution method utilizing harmonic balance method with a single harmonic is used to obtain the required training data, which are employed in the training of the classification and identification networks. It should be noted that in order to increase the applicability of the proposed method, the generated data sets are polluted with random data in order to have a method that is prone to measurement noise.
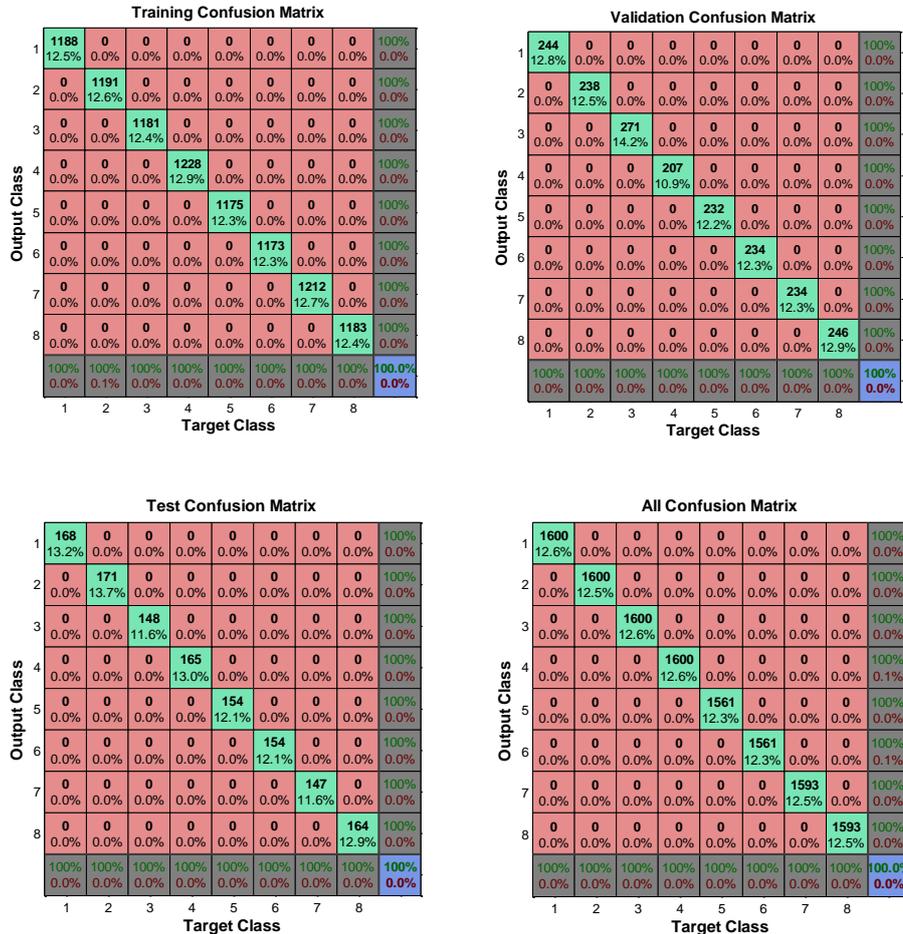

Figure 5 Confusion Matrices of Classification Network

Table 3 Correlation Coefficients for Each Regression Network

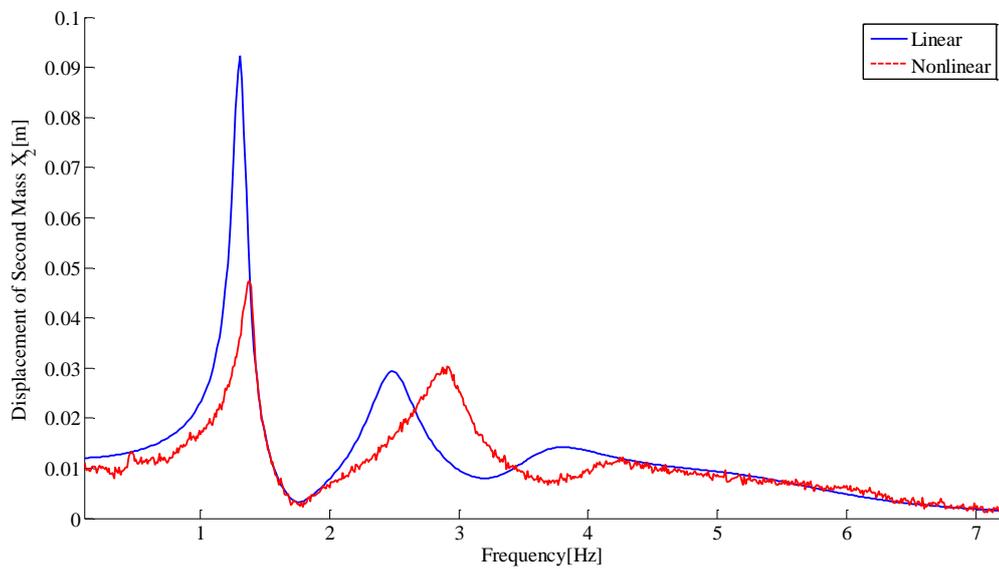|	Training Data Set	Validation Data Set	Test Data Set
Configuration 1	0.999	0.998	0.998
Configuration 2	0.998	0.997	0.997
Configuration 3	0.997	0.992	0.992
Configuration 4	0.997	0.994	0.993
Configuration 5	0.997	0.995	0.995
Configuration 6	0.990	0.992	0.991
Configuration 7	0.999	0.996	0.996
Configuration 8	0.996	0.996	0.996

Figure 6 Displacement of the Second DOF

Table 4 Classification and Parameter Identification Results

	Simulation with Noise Free Measurement			Simulation with Noisy Measurement		
	Target Vector	**Cubic Stiffness N/m³**	**Backlash (m)**	**Target Vector**	**Cubic Stiffness N/m³**	**Backlash (m)**
Actual System	$\{1,0,0,0,0,0,0,0\}^T$	1000000	0.005	$\{1,0,0,0,0,0,0,0\}^T$	1000000	0.005
Identified System	$\{0.999,0,0,0.005,0,0,0,0\}^T$	1015500	0.004935	$\{1,0,0,0.013,0,0,0,0\}^T$	998930	0.0047514
Error	*MSE*= 3x10⁻⁶	1.2%	1.3%	*MSE*= 1.7x10⁻⁴	0.1%	4.9%

89

In this study, two-layered networks give considerably accurate results; however, the number of neurons used in hidden layers should be optimized in order to achieve this accuracy. Using the measured system responses as the inputs of the trained networks, nonlinearities in a 4-DOF system are classified and the parameters associated with them are identified. The results obtained show that the proposed method is capable of identifying the locations and types of the nonlinearities and the parameters associated with them even in the presence of measurement noise.

Number of nonlinear system configurations and size of the training data generated depend on the number possible nonlinearities in the system. Number of nonlinear equations solved for generating training data also depends on the number of nonlinear elements used in the analysis. Therefore, for realistic finite element models, data generation process is a time consuming one, which can be overcome by employing reduction methods. In this study, the number of nonlinear equations is reduced by employing receptance method, which is a very effective method, especially if the nonlinearities are local.

Proposed method is capable to perform identification even the measurements are taken from linear DOFs. This is especially important, since before the classification process, locations of the nonlinearities are unknown. Moreover, by employing a noise injection process to the training data, proposed method can handle measurement noise successfully. Even though training data generation step is computationally demanding, once the networks are trained identification is very fast. Therefore, it is very good candidate for online identification purposes.
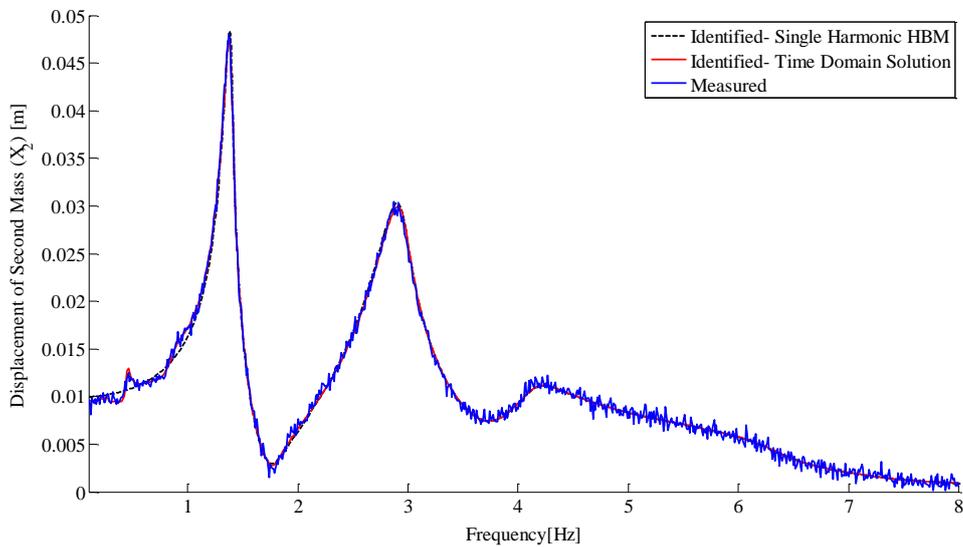

Figure 7 Comparisons of Actual and Identified System Responses for Noisy Measurement Case
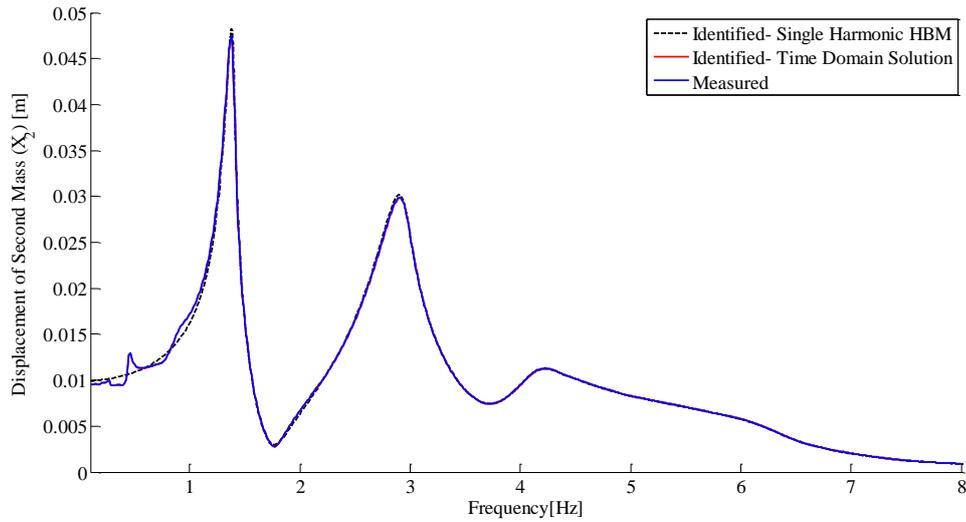
Figure 8 Comparisons of Actual and Identified System Responses for Noise Free Measurement Case

5. REFERENCES

[1]   G. Kerschen, K. Worden, A. F. Vakakis, and J. C. Golinval, "Past, present and future of nonlinear system identification in structural dynamics," *Mechanical Systems and Signal Processing*, vol. 20, pp. 505–592, 2006.

[2]   D. E. Adams and R. J. Allemang, "Survey of Nonlinear Detection and Identification Techniques for Experimental Vibrations," in *International Seminar on Modal Analysis (ISMA 23)*, 1988.

[3]   G. Kerschen, K. Worden, A. F. Vakakis, and J. C. Golinval, "Nonlinear System Identification in Structural Dynamics : Current Status and Future Directions," in *4th International Modal Analysis Conference*, 1986, no. 1, pp. 712–719.

[4]   M. B. Özer, H. N. Özgüven, and T. J. Royston, "Identification of structural non-linearities using describing functions and the Sherman–Morrison method," Mechanical Systems and Signal Processing, vol. 23, no. 1, pp. 30–44, Jan. 2009.

[5]   M. Aykan and H. N. Özgüven, "Parametric Identification of Nonlinearity from Incomplete FRF Data Using Describing Function Inversion," in Proceedings of SEM IMAC XXX Conference, 2012, pp. 323–334.

[6]   D. Göge, M. Sinapius, U. Füllekrug, and M. Link, "Detection and description of non-linear phenomena in experimental modal analysis via linearity plots," International Journal of Non-Linear Mechanics, vol. 40, no. 1, pp. 27–48, Jan. 2005.

[7]   J. He and D. J. Ewins, "A Simple Method of Interpretation for the Modal Analysis of Non-linear Systems," in 5 th International Modal Analysis Conference, 1987, pp. 626–634.

[8]   D. E. Adams and R. J. Allemang, "a Frequency Domain Method for Estimating the Parameters of a Non-Linear Structural Dynamic Model Through Feedback," Mechanical Systems and Signal Processing, vol. 14, no. 4, pp. 637–656, Jul. 2000.

[9]   A. Chatterjee and N. S. Vyas, "Non-linear parameter estimation in multi-degree-of-freedom systems using multi-input Volterra series," Mechanical Systems and Signal Processing, vol. 18, no. 3, pp. 457–489, May 2004.

[10]   S.F. Masri, T.K. Caughey, A nonparametric identification technique for nonlinear dynamic problems, Journal of Applied Mechanics 46 (1979) 433–447 (Introduction Section 3; Sections 3.2, 5.1, 6.1, 6.2).

[11]   C.M. Richards, R. Singh, Identification of multi-degree-of-freedom non-linear systems under random excitations by the reverse-path spectral method, Journal of Sound and Vibration 213 (1998) 673–708 (Sections 3.3, 6.6, 7.1).

[12]   S. F. Masri, A. G. Chassiakos, and T. K. Caughey, "Structure-unknown non-linear dynamic systems : identification through neural networks," vol. 1, pp. 45–56, 1992.

[13]   S. L. Xie, Y. H. Zhang, C. H. Chen, and X. N. Zhang, "Identification of nonlinear hysteretic systems by artificial neural network," Mechanical Systems and Signal Processing, pp. 1–12, Aug. 2012.

[14]   M. H. Beale, M. T. Hagan, and H. B. Demuth, Neural Network Toolbox TM User ' s Guide. 2012, pp. 18–84.

[15]   C.-H. Menq, J. H. Griffin, and J. Bielak, "The influence of microslip on vibratory response, Part II: A comparison with experimental results," Journal of Sound and Vibration, vol. 107, no. 2, pp. 295–307, 1986.

[16]   E. Cigeroglu and H. Samandari, "Nonlinear free vibration of double walled carbon nanotubes by using describing function method with multiple trial functions," Physica E: Low-dimensional Systems and Nanostructures, vol. 46, pp. 160–173, Sep. 2012.

[17]   M. E. Yümer, E. Ciğeroğlu, and H. N. Özgüven, "Non-linear forced response analysis of mistuned bladed disk assemblies," ASME Turbo Expo 2010: Power for Land, Sea, and Air (GT2010), 2010, pp. 757–766.