Bayesian Non-linear System Identification and Frequency Response Analysis with Application to Soft Smart Actuators

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy

William Jacobs

Department of Automatic Control and Systems Engineering

July 2016

Acknowledgements

I would like to thank my supervisor Dr Sean Anderson for his support, guidance and all of our coffee fuelled discussions. Thanks also go to Professor Tony Dodd for his invaluable comments on my thesis.

Thanks to Dr Emma Wilson from the Sheffield Robotics group who collected the experimental data used throughout this thesis and Dr Tareq Assaf from the Bristol Robotics Laboratory for designing and building the experimental rig.

A special thank you to my family for supporting me all the way, and to all of my friends. I have enjoyed all of our adventures over the years and look forward to many more.

Finally, I would like to thank Tara for being a constant source of support in my life.

Abstract

Newly emerging classes of next generation soft-smart actuators are set to have a huge impact on the fields of robotics, orthotics and prosthetics due to their light-weight, high-strain and muscle-like properties. Like muscle, these actuators can be used in multiple roles, *e.g.* both as actuators and brakes, due their variable compliance. One important class of soft actuator is the dielectric elastomer actuator (DEA). However, DEAs are extremely difficult to control due to their non-linear and time varying dynamics. A crucial step in the advancement of this technology is the development of techniques for systems level modelling and analysis, which is the focus of this thesis.

In the first part of the thesis, a set of DEAs are identified and analysed using standard methods from the field of system identification, obtaining non-linear autoregressive with exogenous input (NARX) models. These provide a benchmark against which later methods are evaluated. The key novelty in this part is the development of NARX models of DEAs for use in non-linear frequency-domain analysis. This result provides insight for the first time into how a set of similarly fabricated DEAs vary in different ways.

A further aspect of DEA behaviour is their unexplained time varying behaviour. The system identification approach used to identify NARX models of DEAs is in a convenient form such that it can be easily extended to cater for this time varying behaviour. There are however very few available methods for the frequency domain analysis of time varying systems. A novel method for time varying frequency domain analysis of NARX systems is developed in this work and applied to the DEAs. The analysis procedure is used to provide insight on how the dynamic behaviour of DEAs change over time.

In the second part of the thesis a novel approach to the joint structure detection and parameter estimation of NARX models is developed using a sparse Bayesian method. The Bayesian framework allows for the estimation of posterior distributions over model parameters, characterising the model uncertainty. Analytic solutions are found that describe model uncertainty in the frequency-domain as confidence bounds on both linear and higher order frequency response functions.

The sparse Bayesian identification algorithm is applied to the DEA data sets and is used to give the first non-linear dynamic model of DEAs with uncertainty bounds plus the first description of DEA dynamics in the frequency-domain, again with uncertainty bounds.

Contents

List of Figures

List of Tables

Nomenclature

A list of the variables and notation used in this thesis is defined below. The definitions and conventions set here will be observed throughout unless otherwise stated.

$\boldsymbol{A}$	ARD value
$\boldsymbol{x}$	Vector of lagged input/outputs
Γ	Gamma function operator
$\hat{\boldsymbol{\theta}}$	Estimated parameter vector
$\mathcal{L}$	Variational lower bound
$\mathcal{M}_m$	m'th model structure
ϕ_m	m'th basis function
τ	Precision
Gam	Gamma distribution
Θ	Set of parameters and hyper-parameters
G_n	n'th order NOFRF
H_1	Linear FRF
H_n	n'th order GFRF
M	Total number of model terms
n_p	Maximum polynomial order
n_u	Maximum dynamic order of the input
n_y	Maximum dynamic order of the output

St	Student t distribution
U_n	n'th order input spectrum
Y	Total output spectrum
Y_n	n'th order output spectrum
ARD	Automatic relevance determination
ARMAX	Autoregressive Moving Average with Exogenous inputs
ARX	Autoregressive with exogenous inputs
DEA	Dielectric elastomer actuator
EAP	Electro active polymer
ERR	Error reduction ration
FFT	Fast Fourier transform
FRF	Frequency response function
GFRF	Generalised frequency response function
LASSO	Least absolute shrinkage and selection operator
LS	Least Squares
MC	Monte Carlo
mpo	Model predicted output
MSPE	Mean squared prediction error
MSSE	Mean squared simulation error
NARMAX	Non-linear autoregressive moving average with exogenous input
NARX	Non-linear autoregressive with exogenous inputs
NOFRF	Non-linear output frequency response function
osa	One-step-ahead prediction
RLS	Recursive least squares
SERR	Simulated error reduction ration
SVB-NARX	Sparse variational Bayes NARX

Chapter 1

Introduction

1.1 Background

There are currently 10.3 million people over the age of 65 in the UK alone, an 80% increase over the last 60 years. The growth of the 65+ population is projected to increase with an expected total of 16.9 million by 2035 [105]. The phenomenon of an ageing population is already imposing a large burden on health care institutions and the government's ability to finance them. Mechatronic devices, such as exoskeletons and orthotics, will help to provide a solution to these problems, keeping the elderly in the home for longer and providing greater independence.

These devices will also find an application in the treatment and assistance of people who have lost or damaged limbs [59]. Increase in technology will see exoskeletons that can greatly assist in the rehabilitation of damaged limbs, helping to affectively regain mobility. Where rehabilitation is not possible, dexterous, agile exoskeletons and prosthetics will replace original limb function.

Conventional robotic devices are predominantly built of rigid materials such as steel and stiff plastics, they are designed to be fast, powerful and precise in their actions, specialising to specific tasks that they can achieve with a high efficiency. Robots have revolutionised manufacturing, vastly outperforming their human counterparts at certain tasks. However, they are generally designed to be used in a controlled environment where they seldom come into contact with humans because of the risk of injury.

In order to introduce mechatronic devices, such as exoskeletons, into our everyday lives a new generation of robotics is required, referred to as soft robotics. The new devices must be able to function in man made, as well as natural environments, without posing a danger to those around it. They are also required to perform a diverse range of applications ranging from medical care [32] to human-

robot cooperation [126].

To achieve this the devices must possess an entirely new set of characteristics that allow them to operate effectively and safely within their new environments. New soft robots will be highly compliant and capable of amorphous kinematics and dynamics. Soft wearable robotic orthotics, prosthetics and exoskeletons will help with human motor assistance and must be comfortable, safe, light-weight and move in cooperation with the user. Similarly, medical robots must be able to interact with patients without causing discomfort, and not cause damage in the result of collision.

One of the challenges in the path to realising the potential of soft robots is the design and implementation of soft actuation technologies that allow the robots to effectively perform their assigned tasks [127]. These actuators must be lightweight and compliant while being sufficiently powerful. Actuators of this type are being increasingly referred to as variable impedance actuators, although the term soft actuator is used throughout this thesis. There is already an array of potential candidates, in the large part based on pneumatic actuation [62, 111]. However, pneumatic actuators require off board pneumatic pumps in order to operate. Electro Active Polymers (EAP) provide an alternative solution to the problem of actuation. They have been much studied in recent years because they display a promising collection of characteristics that have led them to be compared to biological muscles [10].

The control of EAP currently offers a significant challenge due to the inherently non-linear dynamics displayed in response to an applied electric field, as well as the time varying characteristics, inherent variability in the dynamic behaviour and the non-uniformity of current fabrication techniques. The accurate description of the non-linear dynamic of a system can be a challenging task and current methods largely consist of first principle models [108, 118, 128, 129]. In order to implement a control structure a data driven modelling approach is more appropriate.

1.2 Motivation

The lack of available techniques for systems level modelling and analysis of EAPs is the key motivation of this thesis. However, once those models have been identified more questions are raised: How can the actuators time varying behaviour be characterised? If the identified models of the system capture the true system dynamics, what can be learnt about the system from the models themselves? And how do we account for the inherent uncertainty in the behaviour of the actuators?

These questions are investigated using experimentally gathered data from a

subclass of EAPs, dielectric elastomer actuators (DEA)s. DEAs are of particular interest for the application of soft robotics because they possess a collection of promising characteristics; combining large actuation strains and a high response speed while being light weight and compliant. The experiment is designed in a manner such that the DEA dynamic behaviour is constrained to one dimension. This allows for the gathering of single input single output (SISO) data over which system identification can be performed. The model class is chosen such that it is easily extendible to a high dimensional problem.

The field of system identification provides us with a collection of tools for obtaining a mathematical description of a system through operations on the observed inputs and outputs of the system. The aim is to obtain a suitable model that accurately captures the dynamics of the system. There are many non-linear model structures available in the literature such as NARMAX, neural networks, fuzzy models, Box Jenkins and output-error [113]. The non-linear Autoregressive Moving Average with eXogenous input (NARMAX) framework provides a unified time domain modelling and analysis tool that has proven successful for many applications in the past [14]. By adopting a parametric modelling approach, tracking the system behaviour over time can be achieved using a number of available recursive parameter estimation techniques. An advantage of the NARMAX framework is the ability to map the time domain model of a non-linear system into the frequency domain as Generalised Frequency Response Functions (GFRF)s and non-linear output frequency response functions (NOFRF)s which facilitates the further analysis of the system in the frequency domain.

System identification methods can therefore be applied to identify data-driven models that are suitable for the application of control. The system identification procedure is performed by a number of steps. Once a suitable model structure has been chosen structure detection algorithms can be used to determine the significant model terms. The model parameters are then estimated using parameter estimation techniques. If a number of models are available to choose between, model selection is performed and the final model is tested for suitability using model validation techniques. In this thesis the models are identified using the NARMAX model structure. The structure detection procedure is initially performed using an algorithm in which term selection is driven by minimising the simulation error. Using this criteria for structure detection has advantages for the predictive ability of the models that are identified.

Identifying models using the NARMAX model class allows the use of advanced model-based frequency domain analysis techniques in which the parameters of the NARMAX model are mapped into the frequency domain via GFRFs.

Although these techniques are available in the time invariant case, little precedent exits for extending these methods to time varying systems. A time varying frequency domain description could however be advantageous in monitoring the changes in system dynamics over time in a transparent and intuitive way.

The variability in DEA behaviour cannot be accounted for using standard methods. One approach to characterising the variability is to characterise the uncertainty in the model. A natural way to incorporate the uncertainty into the model is provided by performing the identification within a Bayesian framework. This allows for prior distributions to be placed over the model parameters and by observing the data the posterior distribution over the parameters can be obtained via Bayes' rule. There is however very little available literature for identifying NARMAX models in this way. In order to address this problem a fully Bayesian structure detection and parameter estimation algorithm is developed based on the concept of sparse estimation.

The uncertainty associated with a NARMAX model is characterised by the variance-covariance of the model parameters. Given that methods exist for mapping the NARMAX model parameters into the frequency domain, a final question is asked in this thesis: How can the uncertainty in the model be manifested in the frequency domain? The problem is complicated by the fact that the frequency domain is commonly represented as a complex variable, *i.e.* it contains a real and imaginary part. Unlike the uncertainty associated with a real variable there is no standard method for uncertainty analysis of complex variables. In order to address this, multivariate uncertainty propagation is performed such that the variance-covariance in the real and imaginary part of the complex variable is approximately mapped from the uncertainty in the model parameters.

1.3 Aims and objectives

The overarching aim of this thesis is the systems level modelling and analysis of EAPs using advanced non-linear modelling and frequency domain analysis techniques.

This key objectives of this thesis are given below.

- The development of a novel framework for the data-driven control focussed identification and model based analysis of DEAs involving:
  - The identification of simple and accurate models using simulation based system identification techniques.
  - Advanced non-linear model-based analysis in the frequency domain.

  - The characterisation of time varying behaviour.

- The development of a novel method for extending the frequency domain analysis techniques to the time varying case and the extension of the above framework to include this.

- The application of the modelling framework to experimentally gathered input output data for six DEAs in order to provide new results into the dynamic behaviour of the actuators as well as analysing the consistency of the fabrication process.

- The development of an identification algorithm capable of performing joint structure detection and parameter estimation of polynomial NARX models within a Bayesian framework in order to evaluate the uncertainty associated with the model.

- Development of a novel method for propagating uncertainty into the frequency domain description of both linear and non-linear systems by considering the uncertainty in the model parameters.

1.4 Thesis overview

The thesis is structured into 8 chapters as follows.

- Chapter 2 introduces the research problems associated with soft actuators for the application of robotics. The state of the art in DEA technology is discussed with an emphasis on modelling techniques and identifies barriers to the implementation of these devises in real life applications. The experimental process used to gather input output data from a set of soft actuators is described and a preliminary analysis of the data is performed.

- Chapter 3 provides a review of system identification techniques and an introduction to the Bayesian inference problem. Linear and non-linear systems theory is introduced and a number of model structures for describing both linear and non-linear systems are discussed. The system identification process is outlined including a review of methods for performing parameter estimation, structure detection, model selection and model validation which represent the main steps in the identification process. An introduction to Bayesian inference for linear regression models is made and the advantages

of identification within a Bayesian context is discussed.

- Chapter 4 focusses on the development of a novel framework for the data driven modelling and frequency domain analysis of DEAs. Simulation based system identification techniques are used in order to identify simple and accurate NARX models of DEAs targeted at the application of control design. GFRF based analysis is used to map the NARX model parameters into the frequency domain. The framework is demonstrated by application to a set of six DEAs in order to provide new results into the dynamic behaviour of the actuators and the consistency of the fabrication process.

- Chapter 5 extends the DEA modelling and analysis framework to include a characterisation of the time varying dynamics of the DEA system. Parameters are tracked over time using Kalman filtering. A novel time varying frequency domain analysis method is introduced based on NOFRFs. The new method provides an intuitive and transparent method for analysing the underlying system behaviour of a nonlinear system as it changes across time. The time varying analysis method is applied to the set of six DEAs in order to provide a new insight into the time varying behaviour.

- Chapter 6 introduces a novel sparse Bayesian identification algorithm for the joint structure detection and parameter estimation of NARX models. Term selection is driven by a sparsity inducing hyper-prior which forces the weighting of irrelevant basis functions to zero, a process named automatic relevance determination. Approximate inference is performed in closed form using the variational approximation in which the posterior distribution is approximated as the product of simpler distributions. This leads to a set of closed form update equations. The algorithm is tested against a benchmark system and is shown to perform favourably in comparison to conventional identification approaches. Identification of the DEA actuators is performed with the new algorithm which capture the uncertainty inherent to the system.

- Chapter 7 presents a novel method for the propagation of uncertainty into the frequency domain description of both polynomial ARX and NARX models. A discussion of the uncertainty in complex valued data is made and

the multivariate uncertainty propagation law is introduced as a method for mapping the uncertainty in NARX model parameters into the frequency domain. The method is demonstrated to accurately capture the uncertainty associated with models of DEAs.

- Chapter 8 provides a conclusions to the work presented in this thesis. Suggestions to future research directions stemming from this work are discussed.

1.5 Summary of contributions

The novel contributions arising from this thesis to both the DEA and system identification communities are listed below:

1. Chapter 4: A novel framework for the data driven identification of control oriented models and model based frequency domain analysis of DEAs is developed. The framework performs simulation based structure detection to identify compact and accurate models of DEAs with a focus on long range predictive ability. Current models of DEAs mostly consist of first principles models that are unsuitable for the purpose of control-focused analysis and design because they tend to be overly complex descriptions, incorporating many terms and parameters. The development of modelling approaches for DEAs that are specifically control-focused hence provides a significant contribution to the field. Furthermore, the framework incorporates advanced non-linear frequency domain analysis techniques based on GFRFs and NOFRFs - for the first time allowing frequency response analysis of non-linear DEA dynamics. The advantage of using GFRFs stems from the fact that the identified model equations are difficult to interpret and directly use for analysis. However, the frequency response is a powerful, widely used method in control for interpreting system dynamics.

   The application of this framework to a set of film type DEAs represents a further contribution originating from this chapter. The modelling and analysis framework is successful in identifying accurate system descriptions of each actuator. The accompanying non-linear frequency domain analysis provides new results on the dynamic behaviour of DEAs and the consistency of the fabrication process.

2. Chapter 5: The framework introduced in Chapter 4 is extended to characterise the time variations inherent in DEA behaviour. Current DEA models provide a limited description of time varying behaviour and cannot explain the time variations displayed by the film-type DEAs under investigation. The data driven modelling and analysis framework does not require that the physical origin of the time variations be known and can characterise the time varying behaviour by recursive estimation of model parameters based only on the data. Current model based frequency analysis techniques are not sufficient to analyses the change in the frequency response due to time varying effects. A novel method for the time varying frequency domain analysis of non-linear systems is developed based on NOFRFs to address this gap in the literature. The new analysis technique provides a novel contribution for model based analysis in general as well as for the analysis of time varying DEAs. Application of the extended framework to the set of film type DEAs provides new results on the evolution of DEA behaviour over time showing that different actuators have significant differences in their time varying behaviour.

3. Chapter 6: A novel sparse Bayesian structure detection algorithm is introduced for the identification of dynamic systems of the NARX model class. The comparison of DEAs in Chapters 4 and 5 are based on models identified using a frequentist modelling approach. It is therefore unclear if the apparent differences in DEA behaviour can be attributed to true differences in the dynamic behaviour between actuators or if it is due to inaccuracies in the modelling process. This motivates a method for characterising the uncertainty in the identified models of DEA. By using a Bayesian modelling approach, parameter uncertainty can naturally be incorporated into the identified model. There is, however, a lack of available Bayesian methods for the identification of non-linear models of the NARX class. Existing methods are based on repeated sampling techniques, and are complex and inefficient. The novel algorithm introduced in this chapter is computationally efficient, consisting of a sequence of closed form updates. Structure detection is driven by the inclusion of a sparsity inducing hyper prior such that terms that are not relevant to the generation of the data can be iteratively pruned from the model. The final model structure is chosen automatically using Bayesian model selection.

   Non-linear models of DEAs identified using the new identification algorithm include a characterisation of the model uncertainty through the uncertainty

in the model parameters. Uncertainty in the frequency domain is evaluated by sampling from the uncertain model parameters to build up a distribution in the frequency domain. The uncertain frequency domain analysis provides an improved comparison between actuators, representing a further novel contribution resulting from this chapter.

4. Chapter 7: A novel method for the propagation of uncertainty from the posterior model parameter distribution into both linear and higher order frequency response functions (FRF)s is presented. Current methods for analysing frequency domain uncertainty are limited to propagating the uncertainty in experimentally gathered data into the (complex valued) FRF, or by sampling from the posterior distribution of the model parameters. The sampling approach is however very computationally inefficient, especially when considering higher orders. There is therefore a gap in the literature regarding methods for model based uncertainty analysis of frequency domain descriptions for both linear and non-linear systems. The new uncertainty propagation method therefore represents a significant novel contribution originating from this chapter.

1.6 Research outputs

The work presented in this thesis has been published in part in a journal paper and a peer reviewed conference paper. A further Journal paper is currently in preparation.

- W.R. Jacobs, E.D. Wilson, T. Assaf, J. Rossiter, T.J. Dodd, J. Porrill and S.R. Anderson. Control-focused, nonlinear and time-varying modelling of dielectric elastomer actuators with frequency response analysis. *Smart Materials and Structures* 24.5 (2015): 055002.
  The journal paper incorporates elements from Chapters 4 and 5.

- W.R. Jacobs, T. Baldacchino and S.R. Anderson. Sparse Bayesian Identification of Polynomial NARX Models. 17th IFAC Symposium on System Identification, *SYSID 2015*.
  The Conference paper incorporates elements from Chapter 6.

- W.R. Jacobs, E.D. Wilson, T.J. Dodd, Z-Q Lang and S.R. Anderson. Analysis of time-varying non-linear systems in the frequency domain using non-linear

output frequency response functions. (In preparation)
The prospective journal paper incorporates elements from Chapters 4 and 5.

- W.R. Jacobs, T. Baldacchino and S.R. Anderson. Sparse Bayesian system identification and frequency domain uncertainty analysis. (In preparation)
The prospective journal paper incorporates elements from Chapters 6 and 7.

Chapter 2

Dielectric Elastomer Actuators for Soft Robotics: Applications, Theory and Experimental Design.

Soft robots promise a range of advantages and opportunities that are not possible with traditional robots. Constructing robots with compliant, amorphous materials such as polymers, liquids and gels leads to the possibility of actuation in multiple degrees of freedom in a way that was previously not possible [124]. As well as applications in exoskeleton and orthotics devices the similarity to biological systems allows new research potential in biomimetics and the emulation of natural systems [61, 70, 82, 96] and facilitating new routes into the modelling and understanding of natural concepts.

There are many challenges to overcome before the potential of soft robotic devices can be truly realised. The compliant nature of soft robots as well as their methods of actuation, often through deformation in shape by stretching, contracting or varying stiffness, provides a difficulty when it comes to forming mathematical models of the system [73]. The system response is often inherently non-linear and follows different kinematic rules to traditional rigid robots. Current control strategies rely on precise knowledge of the robots location which may be difficult to achieve with soft actuators. Another challenge is the lack of soft actuator technologies and corresponding control structures [52].

One of the key problems preventing the progress of the soft robotics field is the lack of readiness and availability of soft actuator technologies [73]. Soft actuators are required to be light and compliant while also producing the required force for their application.

McKibben actuators (or pneumatic artificial muscles (PAMs)) were an early

candidate for use in soft robots and have mostly found applications in medical rehabilitation robots [26, 62]. The actuator is formed of a pneumatic bladder surrounded by woven fibres, actuation is induced by inflating the bladder with pressurised air, the surrounding woven fibres restrict the volume of the bladder causing the ends to be pulled together. The construction of the PAM allows the actuator to be very light as well as compliant. There are big disadvantages to the actuator however, heavy air pressurising equipment is required to be carried (or for the robot to be tethered) in order to operate the actuators.

More recently, longitudinal cables and transverse shape memory alloys have been used to actuate soft robotic arms for the OCTOPUS project [70]. The arms are capable of lifelike movement underwater in multiple degrees of freedom as well as grasping objects. A soft locomotive robot has been build, actuated by inflating air bladders built into the soft silicon structure[111]. The robot is capable of sophisticated locomotion with multiple gaits as well as navigating simple obstacles. The actuation method shares the disadvantage of needing to be tethered in order to operate.

Active polymers are another promising candidate for applications in soft robotics. Active Polymers are comprised of a broad family of dimension changing polymers that have great potential for actuation purposes as well as for sensing and energy harvesting [10, 66].

The remainder of this chapter is structured as follows: In Section 2.1 an overview of electro active polymer technology is provided highlighting different materials and their advantages for the purpose of actuation. Dielectric elastomers are then introduced in Section 2.2. The working principles of dielectric elastomer actuators are discussed, including how different configurations can produce different actuation modes. A discussion of the current state of the art in the modelling of actuators of this type is then made. In Section 2.3 the experimental procedure used to gather input output data from a simple dielectric elastomer actuator system is first described. Pre analysis of data gathered from the experiment is then performed and a number of features displayed by the data are discussed. In the final section, Section 2.4, a discussion of the findings in this chapter is made.

2.1 Electro active polymers as smart actuators

Active polymers are materials for which deformation is induced by a range of different stimuli. In this thesis the focus will be placed on a subset of active polymers (EAP)s named electro active polymers for which deformation is induced through stimulation by an applied voltage.

Increased interest in the field of EAPs has been facilitated by advances in polymer technology in the past decade [7, 9, 89]. EAPs have the ability to directly transform electrical energy into mechanical work by deforming in shape. Their potential to produce large actuation strains while remaining light weight, coupled with their compliance and resilience, has driven research in recent years.

There is a further variety of materials falling into this category which can be subdivided into two main types, namely, ionic EAPs and electronic EAPs

- **Ionic EAPs** Ionic EAP exploit the displacement of ions inside a polymer in order to induce deformation. Large deformations can be achieved at a low voltage (just a few volts), however, the operating principle is based on the diffusion of ions across an electrolyte which requires the system to be submerged in liquid in many cases, limiting the applicability [7, 10].

- **Electronic EAPs** For electronic EAP, deformation is achieved by the displacement of electric charge within the polymer material, inducing electrostatic forces that in turn lead to deformation of the material.

A summary of common types of both ionic and electronic EAP can be found in Table 2.1 as well as a discussion of the advantages/disadvantages of each material in Table 2.2.

Electronic EAP	**Ionic EAP**
Dielectric elastomer EAP	Carbon Nanotubes
Electrostrictive Graft Elastomers	Conductive Polymers
Electrostrictive Paper	ElectroRheological Fluids
Electro-Viscoelastic Elastomers	Ionic Polymer Gels
Ferroelectric Polymers	Ionic Polymer Metallic Composite
Liquid Crystal Elastomers	

Table 2.1: List of the main types of electronic and ionic EAP [10].

In the context of actuation in robotics, Ionic EAP are limited by low strain rates and slow response times [50]. Many electronic EAPs are also limited by their achievable strains, such as electrostrictive and pizoelectric polymers [50]. Dielectric elastomers however show great potential, displaying characteristics that are highly favourable for this application. Dielectric elastomers will be discussed in more detail in the following section.

	Electronic EAP	Ionic EAP
Actuation	Actuation is induced in the planer direction due to in-plane deformation. Actuation is independent of voltage polarity.	Actuation is induced by the contraction of one side of the material causing bending. Bending direction dependant on voltage polarity.
Required voltage	Large Voltages required for actuation (~kV).	Low voltages (~mV)
Response time	Rapid response (~ms)	Slow Response (~s)
Actuation strain	Displays relatively hight actuation strains (Compared to traditional actuators)	Displays relatively low actuation strains (Compared to traditional actuators)

Table 2.2: Summary of the properties of electronic and ionic EAPs [10].

2.2 Dielectric elastomer EAPs

The focus of this thesis is on the modelling and analysis of dielectric elastomer's, a type of electronic EAP. Dielectric elastomer's are of particular interest to the field of robotics due to a collection of attractive properties. Strains of over 100% have been achieved as well as response times in the micro-seconds, exceeding the properties of biological muscle [93]. This exceptional combination of properties is especially applicable for the development of biologically inspired mechanisms and has led to the comparison of dielectric elastomers to biological muscle and as such they have been dubbed 'Artificial muscles'.

2.2.1 Working principles

The basic structure of a dielectric elastomer actuator (DEA) consists of an insulating polymer sandwiched between two compliant electrodes. In essence, the actuator works as a compliant capacitor. An electrostatic attraction is formed by applying a potential difference across the two electrodes, squeezing the (incompressible) polymer. Simultaneously, the like charges in the (compliant) electrodes repel each other. These two factors lead to an expansion in the planer direction, see Figure 2.1.

The described principle of actuation allows the DEA to perform work against external loads in two main directions; In-plane due to planar expansion, and in the thickness direction due to contraction, see Figure 2.1. DEAs can be fabricated in many different configurations to produce different modes of actuation. Config-

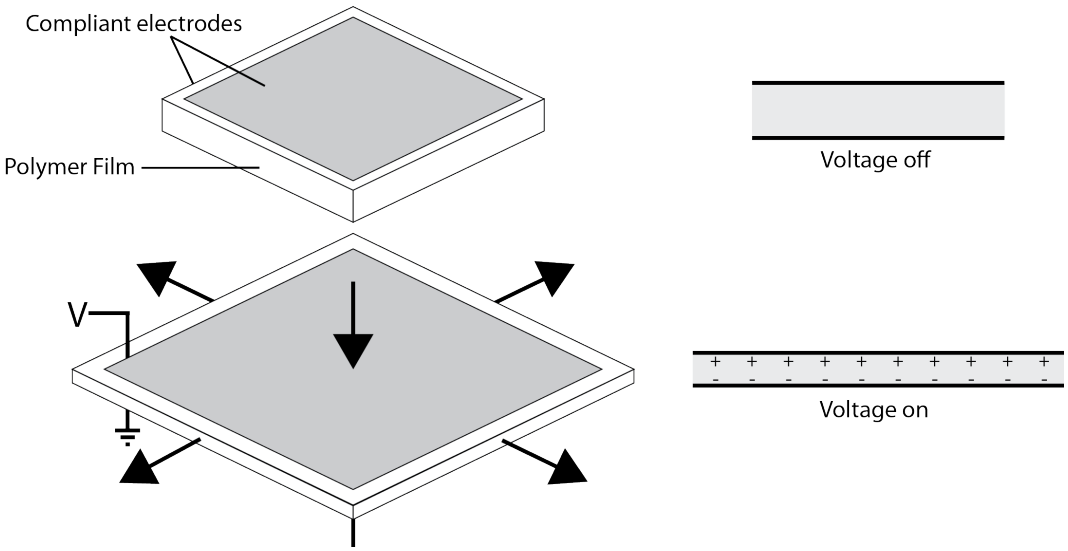

Figure 2.1: **Working Principles of DEAs.** When a voltage is applied across compliant electrodes placed on either side of a polymer film, electrostatic attraction causes the film to contract in the thickness direction and expand in the planer direction.

urations include diaphragms (films), unimorphes and biomorphes, bow-tie, cone and roll [89] some of which are discussed in more detail in Section 2.2.3.

The driving voltage required to induce actuation is typically very high (up to 150 $V/\mu m$) and no deformation in the material will be induced until a minimum voltage in this range is reached [89]. As the thickness of the DE film decreases there is a resultant increase in the electric field strength, if the electric field reached a critical point then it will break down, damaging the DE material. This phenomenon is referred to as the electromechanical instability [10]. Application of a voltage above a certain threshold can therefore damage the dielectric elastomer material because the induced deformation causes a decrease in the thickness direction and hence an increase in the electric field strength. Pre-straining of the DE material has been found to increase the actuators resistance to this phenomena.

2.2.2 Pre-straining of dielectric elastomer material

Pre-straining of DE films can have the effect of improving the magnitude of the induced deformation while reducing thickness as well as the required voltage [93]. The electrical breakdown strength may also be increased in this way, thus increasing the actuators electromechanical stability. However, pre-straining reduces the lifetime of the actuator due to relaxation and fatigue in the film [74].

2.2.3 Actuators and configurations of dielectric elastomers

The compliant nature of elastomer material allows for the constraining of the DE material in different ways, forming different configurations that can actuate in different dimensions. A large variety of actuators have been constructed that can provide actuation for a wide array of different applications, both current and potential. It is most common for the DEAs to harness the planar expansion of the material in order to perform actuation. Actuation in the out of plane direction has been achieved by stacking a number of actuators into one system.

Bow-tie actuators have already found uses in robotic devices [36]. They consist of a pre-strained film fixed to a rigid frame on two opposing sides and hinged spars on the remaining sides, bent in the shape of bow-tie. Actuation is induced in the planer direction either perpendicular too, or in the direction of, elongation. Multifunctional electro-elastomer roll (MER) actuators are a promising new technology that combines load bearing, actuation, and sensing functions. They are fabricated by rolling highly pre-strained DE films onto a central compression spring. The actuator has the ability to bend as well as extend providing multiple degrees of freedom of actuation [92].

Diaphragms can be formed by stretching a film of DE material over a frame and smearing both surfaces with a conductor such as copper grease to form the electrodes. A diverse range of applications have been found for DE diaphragms including loudspeakers [51], pumps [137], and refreshable Braille displays [101]. DE diaphragms are often pre-strained to alter or enhance certain properties of the actuator [31].

2.2.4 Theory and first principle model descriptions

Current models of the dynamic behaviour of EAP actuators are mostly constrained to first principle models. Early work in this field successfully predicted the linear response of DEAs that is displayed at low voltages, and hence low strains (<10%), by considering linear-elasticity and free boundary approximations [25]. However, DEA's have been shown to produce strains much greater than that predicted by the former theory [93]. At larger strains the response becomes nonlinear and further theory is required to explain it. As well as displaying nonlinear dynamics the DEs also show unexplained time varying characteristics.

The large strains exhibited by DE's can be primarily attributed to electrostriction and Maxwell stress effects. The induced strain arising from electrostriction effects is given by

$$S_{\text{electrostriction}} = -Q\epsilon_0^2(\epsilon_r)^2E^2 \tag{2.1}$$

where Q is the electrostrictive coefficient, ϵ_0 is the permittivity of free space, ϵ_r is the relative permittivity (dielectric constant) and E is the applied electric field.

Maxwell stress arises from the change in the distribution of the electric field inside the dielectric material and is given by

$$S_{\text{Maxwell}} = -\frac{1}{2} s \epsilon_0 \epsilon_r E^2 \tag{2.2}$$

where s is the elastic compliance. It can be noted that both effects exhibit a quadratic response similar to that displayed by the DE's in this investigation.

More sophisticated models have since been developed to account for the non-linearity observable at high strains based on various hyper-elasticity models such as Ogden [88], Mooney-Revlin [84] and Yeoh [140]. These are derived by considering the Maxwell stress and the elasticity of the situation in combination, and by assuming a constant electric permittivity. These models as well as variations of the same have been used extensively in the EAP litrature [25, 134].

It has been shown, however, that the permittivity (dielectric constant) of the DEs is not a constant at large deformations as assumed in previous models [135], varying by as much as a factor of 2, expressing a need for a model that does not rely on this assumption. A new type of model has been developed based on the theory of the thermodynamics of deformable dielectrics [118, 119], that accounts for the variable permittivity as well as showing that the Maxwell stress can only account for part of the voltage induced deformation.

2.3 On the behaviour of one dimensionally constrained film type DEAs

In this thesis the dynamic behaviour of DEAs is investigated using the experimental set-up described in the following section. The experiment was deliberately designed in order to constrain actuation to one dimension in order to provide a system that is as simple as possible whilst still displaying the systems non-linear and time varying behaviour. The simplicity of the system allows for the collection of single-input single-output data. The experimental design and data collection was performed by collaborators at the Bristol Robotics Laboratory for the purpose of system identification.

2.3.1 Experimental design

The DEAs under investigation comprise of a thin elastomer film, sandwiched between two compliant electrodes. The DEA used was formed of an acrylic elas-

tomer (3M VHB 4905), fixed to a rigid perspex frame, inner and outer diameters 80mm and 120mm respectively. The elastomer was biaxially pre-stretched and preloaded by 350%. Conductive carbon grease (MG chemicals) was smeared on both sides of the elastomer to form the electrodes, the electrodes were approximately circular, diameter $\approx$ 35mm. A spherical 3g mass was placed at the centre of the circular DEA allowing a vertical displacement to be measured with a laser displacement sensor (Keyence LK-G152), see Figures 2.2 and 2.3.

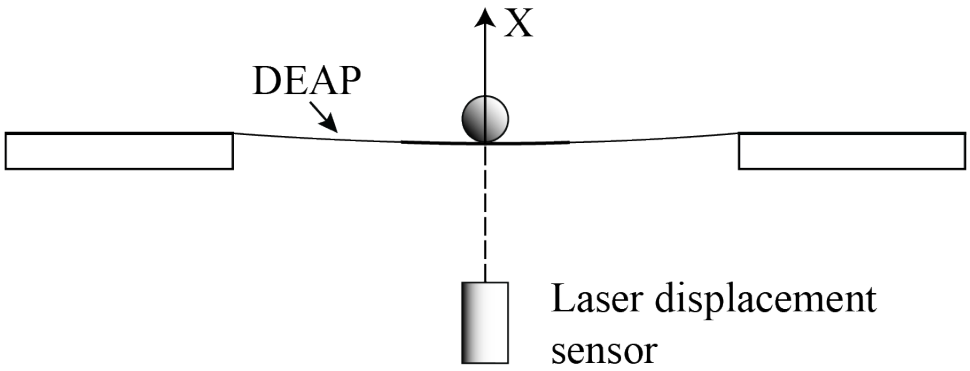

Figure 2.2: **DEA dynamic behaviour is constrained to one dimension by the experimental set-up.** A cross section of the DEA actuator. Application of an applied electric field causes the metal ball to move in the out of plane direction. The displacement of the ball (system output) is recorded by a laser displacement sensor.

The electrodes are squeezed together in response to an applied input voltage, causing planer expansion of the elastomer film. In turn this causes actuation in the out of plane direction and therefore the ball to move up and down. The system output is measured as the displacement of the ball relative to its location at zero input.

The excitation signal was formed of a zero mean normally distributed white noise signal with variance $\sigma_u^2 = 0.36$. The signal is then band limited to the range [0-1]Hz in order to excite the actuator in the frequency range appropriate to the future application.

A host Laptop combined with a CompactRio (CRIO-9014, National Instruments) platform, with input module NI-9144 (National Instruments) and output module NI-9264 (National Instruments) were used for the control hardware. Input and output signals were simultaneously sampled at 50Hz. The signal measured by the laser displacement sensor was supplied to the input module of the CRio.

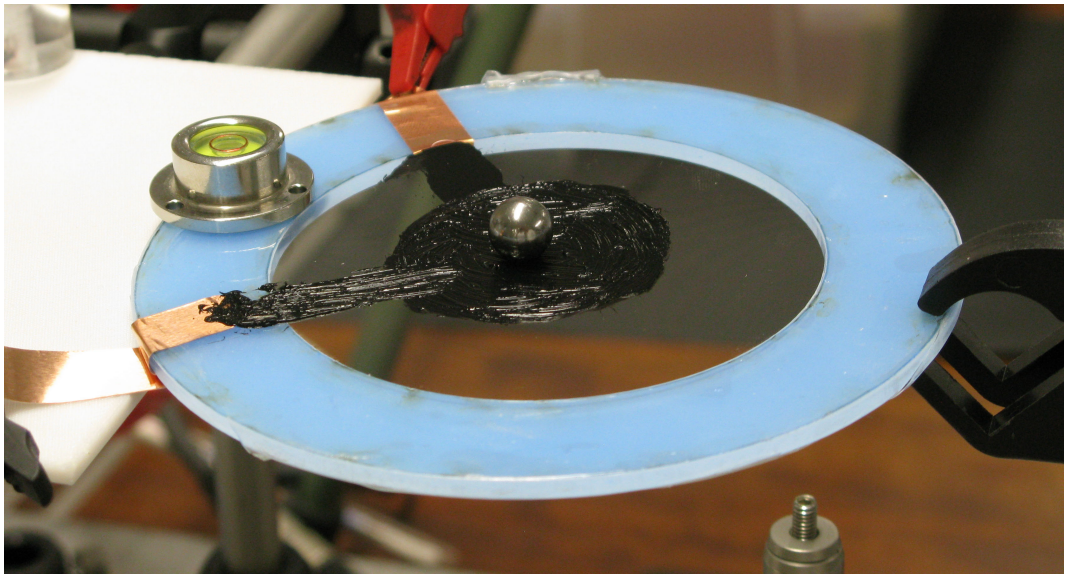

Figure 2.3: **A photograph of the DEA experimental set-up.** The photograph shows the dielectric elastomer film stretched over a perspex frame. Electrodes are painted onto each side of the elastomer film with carbon grease. A metal ball is placed in the center of the film to act as a weight.

Voltages in the range 1.1V-3.75V were then produced by the CRio output module by up-shifting the input signal by 2.4V and amplifying with the ratio 15V:12kV before being fed into the DEA, see Figure 2.4.

Input output data was collected for six different DEAs using the experimental set-up described, see Figure 2.5. All six DEA actuators were fabricated to the same specifications such that the dynamic behaviour is expected to be consistent across the six actuators. Any differences in there behaviour can hence be attributed to unknown degradation effects or problems with the fabrication process. An initial discussion of both the non-linear and time varying dynamics of the different actuators is presented in the remainder of this chapter.

2.3.2 Time varying dynamics of DEAs

Elastomer films, such as the VHB elastomer used in this study, are known to display time varying behaviour [104]. The relative permittivity of the elastomer, ϵ_r, has been shown to be time dependant in some cases due to relaxation in the elastomer material or by the gradual orientation of the polar parts of the elastomer chains in the presence of a high electric field [81]. Pre-strained DEAs also commonly display stress relaxation and creep [60].

By inspection of the output data against time for all six DEA actuators it is clear that there is some form of time varying dynamics taking place, see Figure

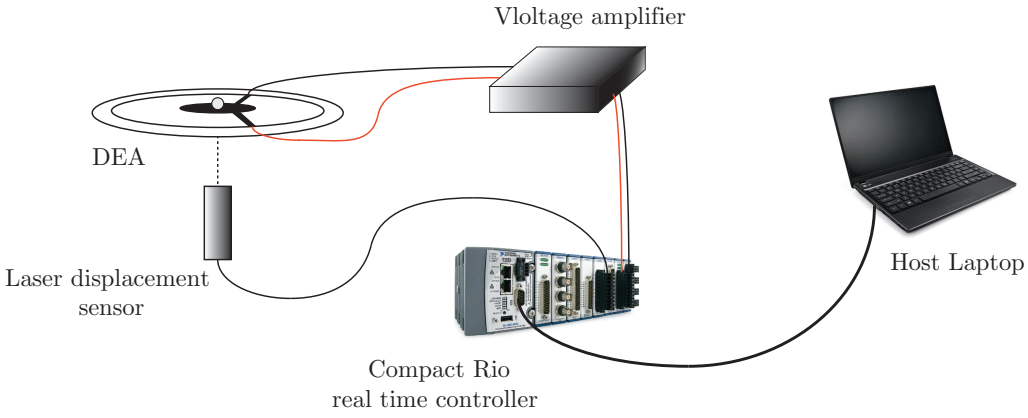

Figure 2.4: **Summary of the experimental set-up used to collect input-output data for DEAs.** The CompactRio simultaneously collects data from the laser displacement sensor and produces an output signal fed from the host laptop. The output signal is amplified before being fed into the DEA.

2.5. The mechanism behind this behaviour is not due to stress relaxation or creep because these phenomena are repeatable and have known characteristics [104]. It is interesting to note that the effect is manifested differently with each actuator. For example, DEA 1 exhibits a consistent decrease in the average displacement whereas DEAs 2 and 3 show a consistent increase in average displacement. Actuator 6 experiences a rapid shift in mean value followed be a period of apparent invariance. The effect may be due to a time dependant relative permittivity of the material, however it has not been established how to model this behaviour [81]. The origin of the time varying behaviour is therefore assumed unknown.

Observing the data over a short time scale it can be seen that DEA 1 exhibits behaviour not displayed by the other actuators Comparing the behaviour of DEA 1 at time $10-30$s with that for $1000-1020$s and the same time periods for DEA 5 the output is noted to be significantly less smooth in the period $10-30$s, particularly in the output range $\approx 0.4-0.5$mm, see Figure 2.6. Rather than removing this data set from the investigation it shall be included in order to see how the frequency domain analysis tools, introduced later in this thesis, interpret this behaviour.

2.3.3 Non-linear dynamics

The non-linear nature of the DEAs can be seen by plotting the input against the output for each DEA, see Figure 2.7. Firstly, by observing the trend of the data, it can be seen that the actuators display an approximately linear response at low voltages (below ≈ 2.5V), and at higher voltages (above ≈ 2.5V) non-linear effects

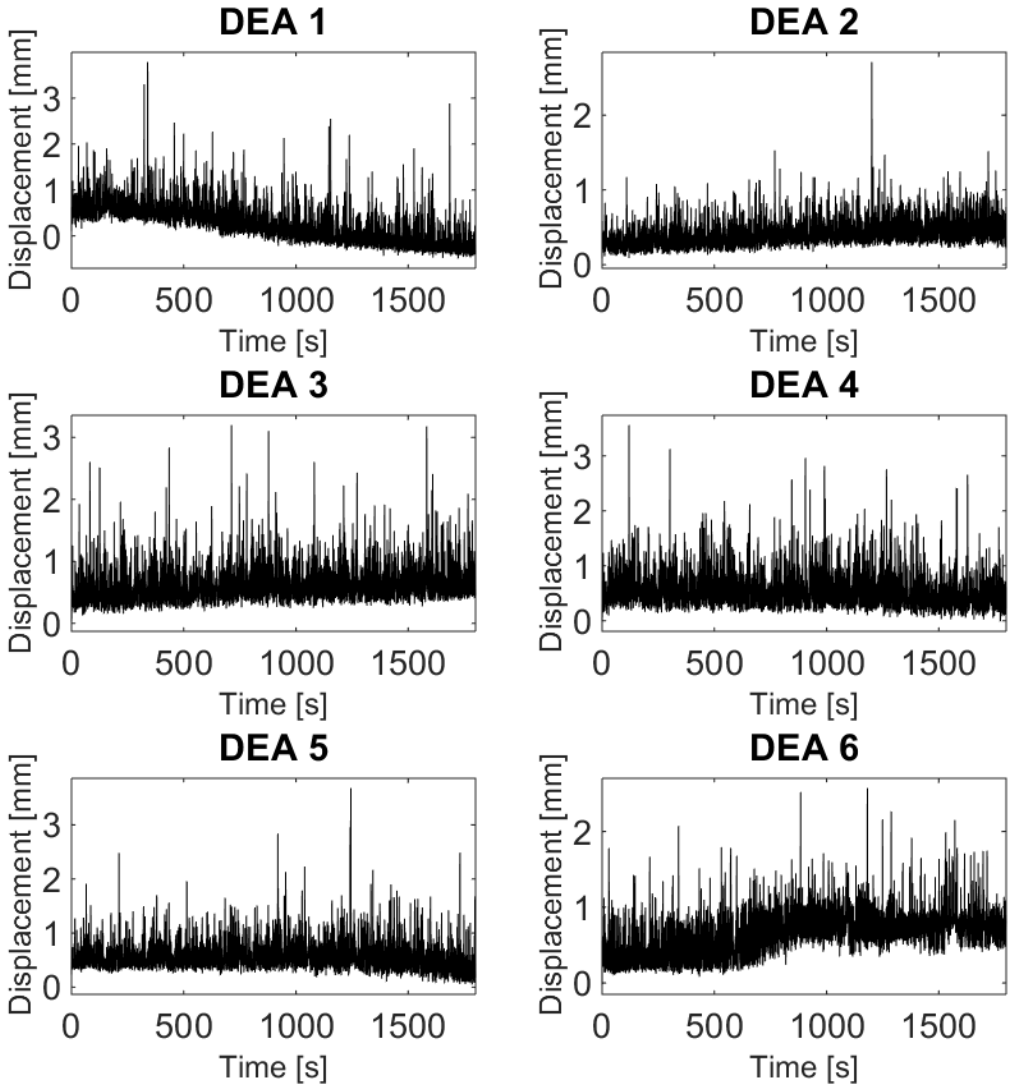

Figure 2.5: **DEAs exhibit significant, non-consistent, time varying behaviour.** Output displacement data for six different DEAs excited by a 0-1Hz band limited white noise signal. The DEAs show significant time variations in their output response.

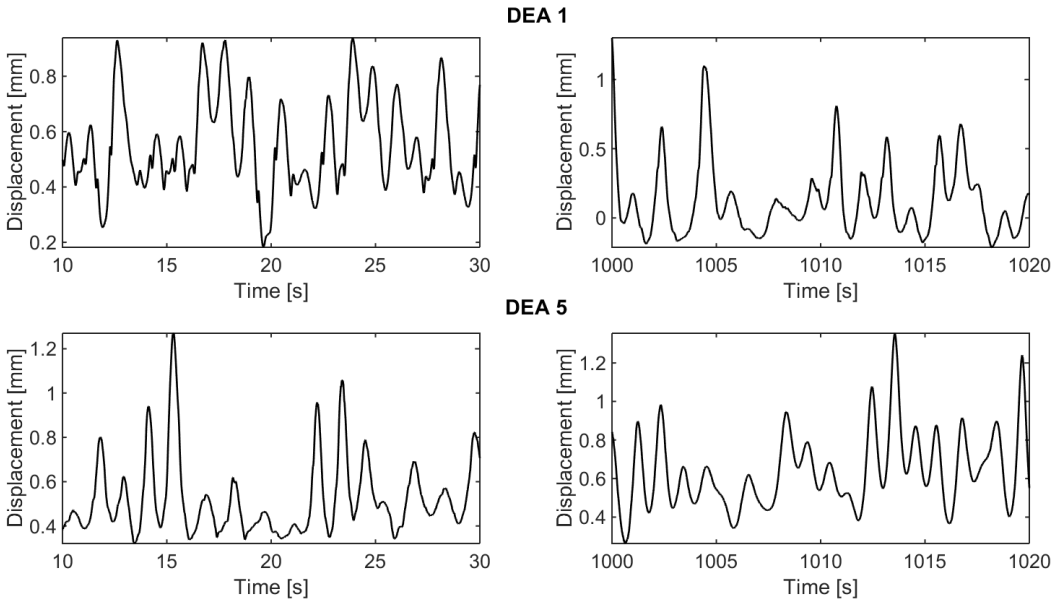

Figure 2.6: **DEA 1 Exhibits behaviour not displayed by any of the other actuators.** The output of DEA 1 and 5 plotted at two different time periods, $t = 10 - 30s$ and $t = 1000 - 1020s$.

are displayed. Secondly the system clearly exhibits hysteresis which can be seen by observing that the system output takes a different path in response to an increasing input than it does the the input it decreasing. Both of these observations are in line with the literature regarding DEAs as discussed in Section 2.2.4, see Figure 2.7

Further evidence of the inherent non-linearities of the system can be seen by observing the frequency domain response. The time domain input and output signals are mapped into the frequency domain by performing a fast Fourier transform (FFT). The spectrum of the input signal appears as expected with a frequency band at $0 - 1$Hz with a noisy but uniform amplitude and zero amplitude at frequencies greater than 1Hz, see Figure 2.8**A**. The spectrum of the output signal shows that some of the energy has been transferred from the $0 - 1$Hz frequency band to higher frequencies ,see Figure 2.8**B**. This behaviour is only displayed by non-linear systems [14].

2.4 Discussion

DEAs have great potential for use in robotic devices, however, there are a number of factors that pose a barrier to widespread implementation in real life devices. The current generation of models for DEAs are based on first principle descrip-

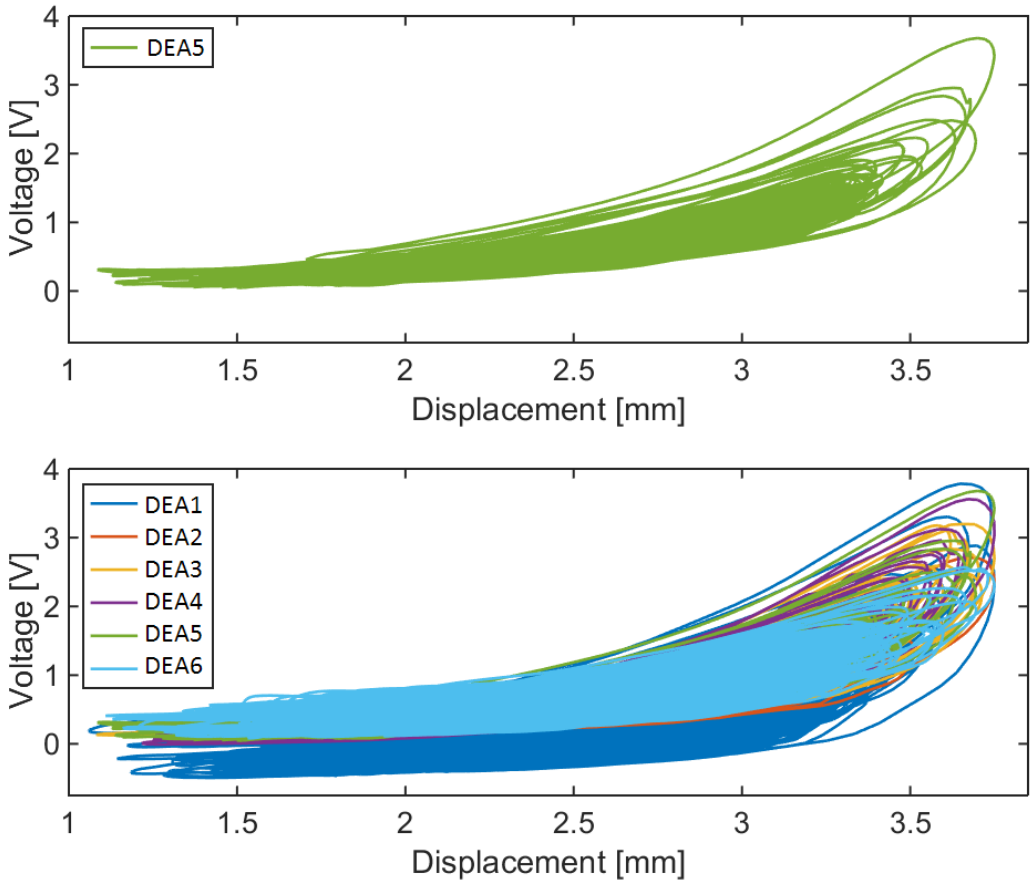

Figure 2.7: **DEAs display non-linear behaviour as well as Hysteresis.** Voltage vs. time for all six DEAs. The output response is seen to be approximately linear at low voltages but becomes non-linear at $\approx$ 2.5V. Hysteresis can also be observed.

tions. These models, which are discussed in Section 2.2.4, have been instrumental in driving forward the development of DEAs. However, they are generally not well suited to control-focused analysis and design because they tend to be overly complex descriptions, incorporating many terms and parameters. Furthermore, DEAs tend to exhibit time-varying characteristics which most current models do not describe, or are limited to a simplified assumption of time varying phenomena [100, 134]. There is, therefore, a significant gap in the current methodology for modelling DEAs in control-oriented tasks. Hence, it is now timely to develop modelling approaches for DEAs that are specifically control-focused due to the increased use of DEAs in various applications.

Section 2.3 details an experimental procedure specifically designed to constrain the dynamics to one dimension. Input-output data gathered from the experimental system therefore reflects the underlying system behaviour. A preliminary anal-

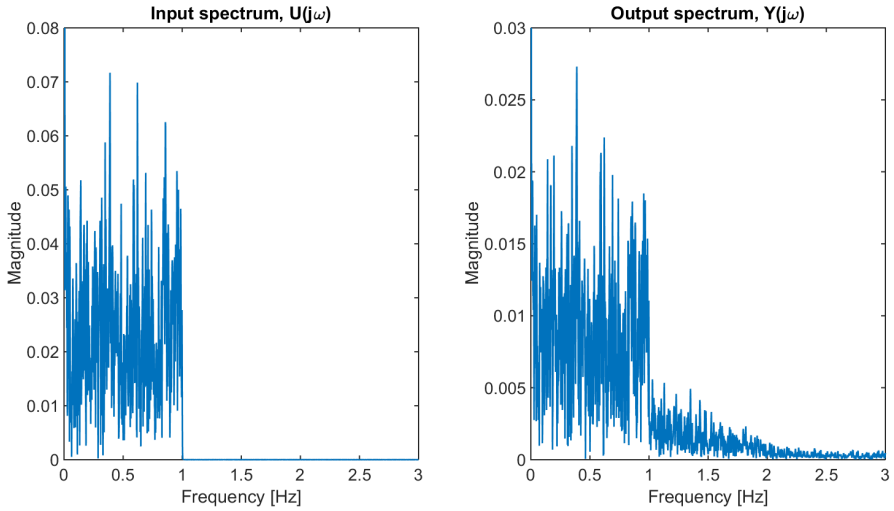

Figure 2.8: **The output spectrum of a DEA in response to a $0-1$Hz band limited excitation signal displays significant energy at frequencies not included in the input spectrum.** Left: The spectrum, $U(j\omega)$ of the excitation signal. Right: The spectrum of the output signal $Y(j\omega)$.

ysis of this data highlights some basic properties of the DEA. Firstly the dynamic response is clearly non-linear, this can be seen from the form of input vs output plot, Figure 2.7, as well as the energy transfer in the frequency domain input output data, see Figure 2.8. This non-linear behaviour is expected and is in line with the literature. However, the literature does not explain the time variations displayed by the DEAs. Interestingly the time variations are markedly different in each of the different actuators.

Given the non-linear nature of DEA behaviour it is clear that it is necessary to consider non-linear modelling techniques in order to accurately capture the behaviour of the system. System identification provides a series of tools for identifying models of a system from gathered input-output data. Model structures are available in which to identify model suitable for the application of control. The next chapter provides an overview of system identification and the identification procedure.

Chapter 3

Introduction to System Identification and Bayesian Inference

3.1 The behaviour of linear and non-linear systems.

The previous chapters highlight the need for control oriented models of DEAs in order to facilitate control design. In this thesis the modelling problem is addressed by treating the system as a black box, meaning that no knowledge of the physical properties of the system are observed or considered as part of the modelling process. Instead models of the system are identified from only experimentally collected input-output data. This approach to modelling broadly falls into the field of system identification. Treating the system as a black box does not mean that the underlying system behaviour is not of interest however. By identifying models that accurately reflect the true system, model based analysis techniques can be employed to describe behaviour that may not be able to be explained by first principle methods.

Linear systems theory is a thoroughly studied and well understood research area. A wealth of literature supports methods for both identification and analysis. However, in reality most real life systems display some level of non-linearity and the assumption of linearity may lead to incorrect system descriptions resulting in poor model performance in analysis, prediction or control. If this is the case then non-linear system identification is necessitated. Non-linear systems theory is commonly based on an extension of well known methods originally designed for linear systems. It is therefore helpful to introduce linear systems theory before non-linear systems are themselves discussed.

This chapter provides an introduction to the field of linear and non-linear system identification and describes the entire identification process. The literature provides a large variety of model types and algorithms for identification. The focus is placed on parametric models in line with the work presented in the rest of the thesis. A number of both linear and non-linear parametric model structures are reviewed. In black box models the parameters are not linked to the physical system and are instead estimated from the data. Two of the most common parameter estimation methods for linear regression problems, least squares and maximum likelihood, are introduced.

The structure detection problem is then introduced which represents possibly the most challenging part of the identification procedure. Structure detection involves identifying a set of basis functions, selected from a super set, that best represent the system with the objective of achieving a parsimonious representation. The difficulty arises in the size of the search space which in some cases can be extremely large. A number of common structure detection algorithms are reviewed as well as their criteria for term selection. The steps of model selection and model validation are also discussed.

Bayesian inference offers an alternative approach to the system identification task and provides a method for uncertainty to be naturally incorporated into the modelling process. Other significant advantages are found in the ability to include prior beliefs or assumptions and a natural penalisation for model complexity. A discussion of the inference procedure is made highlighting the advantages and disadvantages of taking this approach are discussed.

3.1.1 Linear systems

Traditional systems theory is based on linear time invariant (LTI) systems. The LTI property of a system is usually an approximation, however it is often the case that the approximation is justified. Consider a system, S with input u and output y. The system is LTI if it satisfies the properties of superposition and homogeneity defined below.

- **Principle of Superposition** For two different inputs, u_1 and u_2, that produce the system outputs y_1 and y_2 respectively then the the input $u_1 + u_2$ must produce the output $y_1 + y_2$.
- **Principle of Homogeneity** For the input u_1 that produces the output y_1 then the input αu_1 must produce the output αy_1, where alpha is a scalar.

Further to the above two properties linear systems have a unique equilibrium point, that is a point for which, once the system reaches that point, will stay there

for all time [114].

If the LTI system is also causal, the output at time t is dependent on the input up to that time only, then the system can be completely described by its impulse response $h(\tau)$ such that

$$y_t = h(t) * u_t \tag{3.1}$$

where $*$ is the convolution operator such that

$$y_t = \int_{\tau=0}^{\infty} h(\tau) u_{t-\tau} d\tau. \tag{3.2}$$

Given knowledge of $h(\tau)$ $\forall \tau \geq 0$ and u_s $\forall s \leq t$ the output y_t can be calculated for any input. The impulse response, $h(\tau)$, therefore completely characterises the system [75].

The system impulse response can then be used to find another important systems concept, the linear frequency response function (FRF), which in this thesis will be denoted $H_1(j\omega)$ where the subscript 1 indicates that it is the first order (linear) FRF in order to differentiate from the higher order FRFs that will be introduced later. The FRF is a frequency domain description of a system that is commonly used for systems analysis as well as control design. The tool used to convert from the time domain into the frequency domain is the Fourier transform (FT) such that the FT of the impulse resonance function $h(t)$

$$H_1(j\omega) = \boldsymbol{FT}[h(t)] = \int_{-\infty}^{\infty} h(t) e^{-j\omega t} \mathrm{d}t. \tag{3.3}$$

where $H(j\omega)$ is the frequency domain description of $h(t)$, where $j = \sqrt{-1}$ denotes the complex variable. The FRF is an invaluable tool in understanding system behaviour. It is common to display the FRF in gain and phase form.

The frequency domain description allows us to investigate another property of the LTI system. Consider the system S excited by two sinusoidal inputs with frequencies ω_1 and ω_2. The output of a linear system in response to this input will be the sum of two scaled sinusoid at the frequencies ω_1 and ω_2.

3.1.2 Non-linear systems

The behaviour of non-linear systems is much more complex and the properties of superposition and homogeneity no longer hold. The response of a non-linear system to an external input is also much more complex than in the LTI case.

It is perhaps easiest to demonstrate the complexity of the situation in the fre-

quency domain. Consider an input excitation consisting of two sinusoidal inputs ω_1 and ω_2 such as we had for the LTI case. Non-linear systems are able to transfer energy between frequencies. If the system contains a static quadratic non-linearity then harmonics of the system will cause energy transfer to $2\omega_1$ and $2\omega_2$. Intermodulation will cause energy to be transferred to the sum of the differences of the inputs, *i.e* at $\omega_1 + \omega_2$ and $\omega_2 - \omega_1$. Therefore even for an extremely simple case multiple output frequencies are generated. In this case it is possible to account for all of the output frequencies by inspection. In general the spectrum of the output will be too rich to consider this approach. Later in the thesis a number of approaches based on generalised frequency response functions (GFRFs) [20] will be introduced that can account for these non-linear effects.

Non-linear systems frequently exhibit multiple equilibrium points. This will be important to note later in Chapter 4 when it becomes necessary to shift the equilibrium of non-linear systems to the zero point.

3.2 The system identification procedure

System identification is concerned with the building of mathematical models of dynamic systems that achieve two main goals; Firstly to accurately map the input(s) to the output(s) of the given system and to provide the best possible predictions based on new unseen data. The second aim is to provide a description of the underlying system behaviour. This second goal allows for a better understanding of the system under examination model based analysis.

In this thesis the focus is laid on black box models, in which no prior knowledge of the underlying physics of the system is known but a model structure can be chosen from a flexible framework. The system identification literature provides us with a host of different modelling structures to choose from [113]. Some of the popular methods are NARMAX, neural networks, fuzzy models, Box Jenkins and output-error as well as others [113].

The system identification procedure can be loosely split in to four main problems [75, 116];

1. **Structure detection** Determination of the model structure that describes the mapping from the input to the output variables.

2. **Parameter estimation** The estimation of the parameters that weight each basis function/model term for a given model structure

3. **Model selection** selecting the best model from a set of competing models.

4. **Model validation** validation of the chosen model.

This is, of course, an over simplification of the true process in which the sub-problems are entwined, perhaps solved together or iteratively. For parametric, linear in the parameters models, such as the non-linear autoregressive with exogenous input (NARX) model, parameter estimation is a thoroughly researched topic and many estimators exist. These four steps are discussed in detail in the following four sections in order to give an overview of the entire identification process.

3.3 Model structure

The system identification literature provides a host of different Black-box modelling structures including Volterra series [27], non-linear autoregressive moving average with exogenous input (NARMAX) [28, 72], neural networks [29, 75] and fuzzy models [71, 117, 120] to name a few. All of the above structures can be unified by considering them as a basis function expansion of some function that maps the past data to the space of the output [113].

The system identification task is then to find some mapping from the past inputs, $\mathbf{u}_{k-1} = [u_1\ u_2\ \dots\ u_{k-1}]^T$ and outputs, $\boldsymbol{y}_{k-1} = [y_1\ y_2\ \dots\ y_{k-1}]^T$ to future outputs

$$y_k = f(\boldsymbol{u}_{k-1}, \boldsymbol{y}_{k-1}) + e_k \tag{3.4}$$

where u_k and y_k are the input and output at sample k. and e_k is the error between the predicted output $f(\mathbf{u}_{k-1}, \boldsymbol{y}_{k-1})$ and the observed output y_k at sample k.

This thesis is primarily concerned with the identification and analysis of non-linear models using the NARX/NARMAX modelling framework, however an aspect of the analysis of linear models is discussed in Chapter 7. The discussion is therefore focussed on models of the NARX/NARMAX form.

3.3.1 Linear black-box models

The traditional, and widely used, representation of linear Black-box models is given by [75]

$$y_k = \frac{B(q)}{F(q)A(q)}u_k + \frac{C(q)}{D(q)A(q)}e_k \tag{3.5}$$

where

$$\begin{aligned} A(q) &= 1 - a_1 q^{-1} - \ldots - a_{n_a} q^{-n_a} \\ B(q) &= b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \\ C(q) &= 1 + c_1 q^{-1} + \ldots + c_{n_c} q^{-n_c} \\ D(q) &= 1 + d_1 q^{-1} + \ldots + d_{n_d} q^{-n_d} \\ F(q) &= 1 + f_1 q^{-1} + \ldots + f_{n_f} q^{-n_f} \end{aligned} \tag{3.6}$$

where a, b, c, d and $f \in \mathbb{R}$ are the model parameters that weight the relevant lagged input, output or error term. e_k is assumed to be drawn from a zero mean independent identically distributed (*i.i.d.*) white noise sequence, q represents the forward shift operator such that $q^{-1}y_k = y_{k-1}$. The choice of A, B, C, D and F defines a number of different linear model structures of which some of the most common are; autoregressive (AR), autoregressive with exogenous input(ARX) and autoregressive moving average with exogenous input (ARMAX) which are given by the following equations

AR model: $B(q) = 0$, $C(q) = D(q) = 1$

$$y_k = a_1 y_{k-1} + a_2 y_{k-2} + \ldots + a_{n_a} y_{k-n_a} + e_k \tag{3.7}$$

ARX model: $C(q) = D(q) = F(q) = 1$

$$y_k = a_1 y_{k-1} + \ldots + a_{n_a} y_{k-n_a} + b_1 u_{k-1} + \ldots + b_{n_b} u_{k-n_b} \tag{3.8}$$

ARMAX model: $D(q) = F(q) = 1$

$$\begin{aligned} y_k = a_1 y_{k-1} + \ldots + a_{n_a} y_{k-n_a} + b_1 u_{k-1} + \ldots + b_{n_b} u_{k-n_b} \\ + e_k + c_1 e_{k-1} + \ldots + c_{n_c} e_{k-n_c} \end{aligned} \tag{3.9}$$

Note that the linear models defined by equations (3.7)-(3.9) are all simple basis function expansions in the form of Equation (3.4). The AR and ARX models defined above are *linear-in-the-parameters* allowing the use of simple optimal parameter estimation techniques such as least squares. In contrast the ARMAX model includes lagged error terms causing it to be non-linear in the parameters requiring a more complicated estimation procedure.

3.3.2 Non-linear black-box models

If the system to be modelled displayed non-linear behaviour, such as is discussed in Section 3.1.2, then in order to accurately model the system dynamics, the model used to represent it must also be non-linear. The NARX model [72] provides a simple non-linear extension to the ARX model given by Equation (3.8). The NARX model is given by

$$y_k = f(\boldsymbol{x}_k) + e_k \tag{3.10}$$

where $\boldsymbol{x}_k = (y_{k-1}, \ldots, y_{k-n_y}, u_{k-1}, \ldots, u_{k-n_u})$, and n_u, n_y represent the maximum lags in the input and output respectively. The non-linear function $f(\cdot)$ is decomposed into a sum of weighted basis functions such that

$$f(\boldsymbol{x}_k) = \sum_{m=1}^{M} \theta_m \phi_m(\boldsymbol{x}_k) \tag{3.11}$$

$$= \boldsymbol{\phi}_k \boldsymbol{\theta} \tag{3.12}$$

where

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1, & \theta_2, & \ldots, & \theta_M \end{bmatrix}^T \in \mathbb{R} \tag{3.13}$$

$$\boldsymbol{\phi}_k = \begin{bmatrix} \phi_1(\boldsymbol{x}_k), & \phi_2(\boldsymbol{x}_k), \ldots, & \phi_M(\boldsymbol{x}_k) \end{bmatrix} \tag{3.14}$$

where θ_m is the m'th model parameter, $\phi_m(\boldsymbol{x}_k)$ the m'th basis function and m is the total number of model terms. Substituting Equation (3.12) into Equation (3.10) and then considering the entire data record, the vector of system outputs is given in matrix form as

$$\boldsymbol{y} = \Phi \boldsymbol{\theta} + \boldsymbol{e} \tag{3.15}$$

where $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^T$, $\boldsymbol{e} = [e_1, e_2, \ldots, e_N]^T$ and Φ is known as the regression matrix and is given by

$$\Phi = \begin{bmatrix} \boldsymbol{\phi}_1^T, & \boldsymbol{\phi}_2^T, \ldots, & \boldsymbol{\phi}_N^T \end{bmatrix}^T. \tag{3.16}$$

The NARX model basis functions can take various forms including polynomial, wavelet, or radial functions [14]. Although $f(\cdot)$ is a non-linear function it still maintains the property of linear-in-the-parameters providing that the basis functions are linearly independent of each other.

The non-linear extension to the ARMAX model is the NARMAX model [72] given by Equation (3.10) where instead

$$\boldsymbol{x}_k = (y_{k-1}, \ldots, y_{k-n_y}, u_{k-1}, \ldots, u_{k-n_u}, e_{k-1}, \ldots, e_{k-n_e}) \tag{3.17}$$

As before with the NARX model, the NARMAX model is also represented by the linear sum of basis function given by Equation (3.12). Like the ARMAX model however, the NARMAX model is *not* linear-in-the-parameters because the error terms depend on the model output resulting in biased parameter estimates.

In this thesis modelling is performed using models of the NARMAX class. This choice of model class is made based on some of the advantages the NARMAX methodology possesses. One of the key advantages of NARMAX is the wealth of supporting literature available for both identification and analysis purposes. More specifically however, NARMAX models are capable of modelling a wide range of non-linear systems with a compact, parsimonious model description. Perhaps most importantly for this thesis, NARMAX models can be mapped directly into the frequency domain as GFRFs allowing non-linear frequency domain model based analysis, the GFRF concept will be explained in Chapter 4.

3.4 Parameter estimation

In general the parameters of a system will be unknown and as such techniques for finding (unbiased) estimates of the unknown parameters are required. There are a variety of different estimators available for this purpose. Two of the commonly used estimators are LS and Maximum likelihood (ML).

Least squares

LS was first proposed by Gauss as a solution to problems in astronomy and the technique has been used in many scientific disciplines [40].

To derive the LS estimator consider Equation (3.15) where $\boldsymbol{\theta}$ is the unknown parameter vector to be determined. The LS estimator can be obtained by minimizing the sum of squared errors, given by the objective function

$$J_{LS} = (\boldsymbol{y} - \Phi\boldsymbol{\theta})^T(\boldsymbol{y} - \Phi\boldsymbol{\theta}). \tag{3.18}$$

Equation (3.18) is a hyperbolic function such that the global minimum can be found by differentiating with respect to the parameters and equating to zero:

$$\frac{dJ_{LS}}{d\boldsymbol{\theta}} = -2\Phi^T(\boldsymbol{y} - \Phi\boldsymbol{\theta}) = 0 \tag{3.19}$$

Rearranging Equation (3.19), the result is the least squares estimate of $\boldsymbol{\theta}$ obtained as

$$\hat{\boldsymbol{\theta}}_{LS} = (\Phi^T\Phi)^{-1}\Phi^T\boldsymbol{y} \tag{3.20}$$

For a linear-in-the-parameters model the LS estimator is optimal under the condition of a normally distributed error signal. For non-linear regression the estimator will be biased and it is necessary to adopt variants of the LS estimator [75].

Maximum likelihood

The ML estimator was introduced by Aström and Bohlin for the estimation of ARMAX model parameters in an early milestone paper [2] and has become another common parameter estimation technique. Parameters are estimated by maximising the likelihood function, $L(\boldsymbol{\theta}|\boldsymbol{y})$, such that

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\theta} L(\boldsymbol{\theta}|\boldsymbol{y}) \tag{3.21}$$

The likelihood function, given by $L(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\theta}, \Phi)$, is the probability of observing the output data, $\boldsymbol{y}$, conditional on a parameter vector $\boldsymbol{\theta}$. The error is assumed to be a *i.i.d.* zero mean sequence drawn from the normal distribution, $\mathcal{N}(e_k|0, \sigma_e^2)$, where σ_e^2 is the error variance. Because each output data point is assumed to be drawn independently from a normal distribution the likelihood function is given by the product of the distributions for each data point such that

$$\begin{aligned} L(\boldsymbol{\theta}|\boldsymbol{y}) &= p(\boldsymbol{y}|\boldsymbol{\theta}, \Phi) && (3.22)\\ &= \prod_{k=1}^{N} \mathcal{N}(y_k|\boldsymbol{\phi}_k\boldsymbol{\theta}, \sigma_e^2) && (3.23)\\ &= \mathcal{N}(\boldsymbol{y}|\Phi\boldsymbol{\theta}, \sigma_e^2\boldsymbol{I}_N). && (3.24) \end{aligned}$$

The maximisation is simplified by working with the log likelihood, this is justified by noting that the logarithm function is a monotonically increasing function and so maximising the log likelihood is equivalent to maximising the likelihood. Applying the logarithm to the likelihood function

$$\begin{aligned} \ln L(\boldsymbol{\theta}|\boldsymbol{y}) &= \ln \mathcal{N}(\boldsymbol{y}|\boldsymbol{\theta}, \sigma_e^2\boldsymbol{I}_N) && (3.25)\\ &= \ln\left(\frac{1}{\sigma_e(2\pi)^{\frac{k}{2}}}\right) - \frac{1}{2\sigma_e^2}(\boldsymbol{y} - \Phi\boldsymbol{\theta})^T\mathbf{I}_N(\boldsymbol{y} - \Phi\boldsymbol{\theta}), && (3.26) \end{aligned}$$

and differentiating the log likelihood with respect to $\boldsymbol{\theta}$ and equating to zero

$$\frac{\mathrm{d}\ln L(\boldsymbol{\theta}|\boldsymbol{y})}{\mathrm{d}\boldsymbol{\theta}} = \frac{1}{2\sigma_e^2}\Phi^T(\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}}_{ML}) = 0. \tag{3.27}$$

Finally, rearranging Equation (3.27) provides the ML estimate of the parameter vector

$$\hat{\boldsymbol{\theta}}_{ML} = (\Phi^T\Phi)^{-1}\Phi^T\boldsymbol{y} \tag{3.28}$$

Comparing (3.28) and (3.20) we can see that the two estimators are equivalent for the case of Normally distributed additive noise.

Extended least squares

Parameters estimated for ARMAX or NARMAX models using the above parameter estimation schemes will result in biased parameters. This is because the noise model depends on the residuals, which are themselves dependent on the model parameters. A number of iterative methods are available to estimate unbiased parameters for ARMAX/NARMAX models such as generalised lease squares (GLS), instrumental variables (IV) and extended least squares (ELS) [75].

The ELS method represents a common solution to the parameter estimation problem. The concept is that the residuals and the model parameters are iteratively updated until convergence. The method is given in Algorithm 3.1 where δ is a termination threshold, $\hat{\boldsymbol{e}}_i$ denotes the residuals at the i'th iteration of the algorithm and $\Phi|_{\hat{e}_i}$ indicated the regression matrix is updated with the new residuals at each iteration, i. The algorithm is initialised by setting the residuals at $i = 0$ equal to zero, $\hat{\boldsymbol{e}}_0 = 0$. At the first algorithm iteration all the terms in the regression matrix that are dependant on the lagged error are therefore equal to zero. The parameter estimate made using this regression matrix is then used to calculate the updated residuals which are used in the regression matrix at the next iteration.

Algorithm 3.1 The extended least squares algorithm.

Initialise
$i = 0,\ \hat{\boldsymbol{e}}_0 = 0$
Procedure
Iteratively update parameter estimates
while $\sum_{k=1}^{N} |\hat{e}_i(k) - \hat{e}_{i-1}(k)|^2 > \delta$
$\Phi_i = \Phi|_{\hat{e}_i}$
Estimate parameters with least squares
$\hat{\boldsymbol{\theta}}_i = (\Phi_s^T\Phi_i)^{-1}\Phi_s^T\boldsymbol{y}$
$\hat{\boldsymbol{e}}_{i+1} = \boldsymbol{y} - \Phi_i\hat{\boldsymbol{\theta}}_i$
$i = i + 1$
end while
end Procedure

3.4.1 Regularised and sparse parameter estimation

A regularisation term can be included into the cost function in order to either control over fitting or encourage sparsity in the parameter estimates [24]. In general a regularised estimator is found by the introduction of a regularisation term $R(f)$ into a cost function such that

$$J = \sum_{k=1}^{N} V(f(\boldsymbol{x}_k), \hat{y}_k) + \lambda R(f) \tag{3.29}$$

where V is the cost of some term that we wish to minimise and $\lambda \geq 0$ is the weighting on the regularisation term. For the least squares estimator, where $V = J_{LS}$ given by Equation (3.18), a common choice of regularisation term is given by the p-norm

$$R(f) = \sum_{m=M}^{m} |\theta_m|^p \tag{3.30}$$

where different regularisers are defined by the value of p. The choice of $p = 1$ and $p = 2$ lead to LASSO and ridge regression respectively. To understand the effect that different regularisation terms have on the parameter estimates it can first be noted that minimising Equation (3.29) for $V = J_{LS}$ and regularisation term given by (3.30) is equivalent to minimising the regularised LS cost function subject to the constraint

$$\sum_{m=1}^{M} |\theta_m|^p \leq \eta \tag{3.31}$$

where η is a positive constant, see appendix A.1 for proof. By visualising the constraint in a 2-dimensional parameter space it can be seen how they act on the minimisation, see Figure 3.1.

Ridge regression

Ridge regression is the name given to the choice of regularisation term given by Equation (3.30) for $p = 2$ which is equivalent to the squared l_2 (or Euclidean) norm. The ridge regression cost function is given by

$$J_{RR} = (\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}})^T(\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}}) + \lambda\boldsymbol{\theta}'\boldsymbol{\theta}. \tag{3.32}$$

The derivation of the ridge regression estimator closely follows that for LS, see appendix A.2, resulting in the estimator

$$\hat{\boldsymbol{\theta}}_{RR} = (\Phi^T\Phi + \lambda\boldsymbol{I})^{-1}\Phi^T\boldsymbol{y}. \tag{3.33}$$

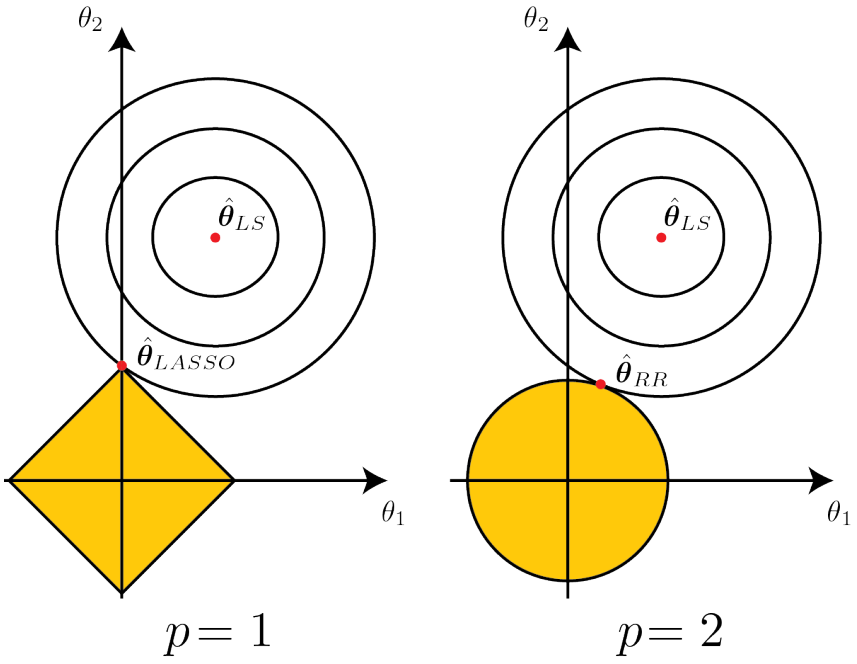

Figure 3.1: **The choice of regularisor enforces constraints on the minimisation.** Adapted from [24]. Contour plot of the cost function subject to the constraint (3.30) for: Left) $p = 1$ (LASSO) provides a sparse solution in which parameter $\theta_1 = 0$. Right) $p = 2$ (ridge regression) causes shrinkage in both parameters.

The constant,λ, therefore adds a positive value to the diagonal (ridge) of the sample covariance matrix $\Phi^T\Phi$. Ridge regression causes parameter shrinkage, *i.e.* the model parameters shrink towards zero, with the value of λ controlling the amount of shrinkage. Ridge regression does not force the parameters to be exactly zero, rather it forces them to be small. This can be seen pictorially in Figure 3.1 for $p = 2$.

LASSO

The least absolute selection and shrinkage operator (LASSO) is defined similarly to ridge regression but with $p = 1$, equivalent to the l_1 norm, as the regularisation term such that

$$J_{LASSO} = (\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}})^T(\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}}) + \lambda \sum_{m=1}^{M} |\theta_m|^1. \tag{3.34}$$

LASSO encourages sparsity in the parameter estimates. Sparsity refers to the estimation of parameters in which some of the parameters are forced to zero. Such a scheme can be useful in selection basis functions that are not relevant to the generation of the data. It can be seen from Figure 3.1 how the constraint on the parameter size for the LASSO estimator encourages sparsity. Unlike ridge regression

there is no closed form solution for LASSO estimates. LASSO therefore requires the use of convex optimisation methods to solve. LASSO has the advantage over ridge regression in that it can force parameter estimates to be exactly zero and is therefore more suitable to the task of detecting terms that are not relevant to the data.

3.4.2 Recursive parameter estimation

Many real-world systems display time-varying/non-stationary behaviour [37]. For linear systems, modelling of time-varying systems is a well studied area and a variety of adaptive recursive methods are available such as recursive least squares (RLS), least mean squares (LMS) and Kalman filtering [76].

Recursive least squares

RLS is a recursive algorithm that iteratively minimizes a weighted LS cost function in order to find the model parameters at each time step. The cost function is given by

$$J_{RLS} = \sum_{k=1}^{n} \lambda^{n-k}(y_k - \boldsymbol{\phi}_k \boldsymbol{\theta}_k)^2 \tag{3.35}$$

The parameter estimate is updated each time a new data point arrives. The cost function is weighted by a 'forgetting factor', $\lambda \in \mathbb{R}^+$, which gives exponentially less weight to older error samples.

The RLS equations are given by [76]:

$$e_k = y_k - \boldsymbol{\phi}_k \boldsymbol{\theta}_{k-1} \tag{3.36}$$

$$\boldsymbol{g}_k = \boldsymbol{P}_{k-1}\boldsymbol{\phi}_k'(\boldsymbol{\phi}_k \boldsymbol{P}_{k-1}\boldsymbol{\phi}_k' + \lambda)^{-1} \tag{3.37}$$

$$\boldsymbol{P}_k = \lambda^{-1}(\boldsymbol{I} - \boldsymbol{g}_k \boldsymbol{\phi}_k)\boldsymbol{P}_{k-1} \tag{3.38}$$

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_k + e_k \boldsymbol{g}_k \tag{3.39}$$

Kalman filtering

The Kalman filter was originally formulated for state estimation of the state space model which in discrete time is given by

$$\boldsymbol{x}_k = \boldsymbol{A}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1} \tag{3.40}$$

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k \tag{3.41}$$

where $\boldsymbol{x}$ is the state vector, $\boldsymbol{u}$ is the input vector, $\boldsymbol{y}$ is the output vector, $\boldsymbol{A}$ is the state transition matrix, $\boldsymbol{B}$ is the input matrix and $\boldsymbol{H}$ is the observation model. The

state space model separates the noise into the process noise $\boldsymbol{w}_k$ and observation noise $\boldsymbol{v}_k$ drawn from the normal distributions

$$\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q}_k), \quad \boldsymbol{v}_k \sim \mathcal{N}(0, \boldsymbol{R}_k).$$

The algorithm works recursively in two steps, prediction and correction. In the prediction step the the Kalman filter estimates the current states, $\boldsymbol{x}$, along with the uncertainty in the estimate $\boldsymbol{P}_k^-$. These estimates are then updated in the correction step by considering the error between the model prediction and the true system output.

The Kalman filter can conveniently be cast into a parameter estimation problem. For the model given by Equation (3.15), setting $\boldsymbol{x}_k = \boldsymbol{\theta}_k$, $\boldsymbol{A} = \boldsymbol{I}, \boldsymbol{B} = 0 \; \forall k$ and $\boldsymbol{H_k} = \boldsymbol{\phi}_k'$, then the Kalman filter equations are given by

Prediction

$$\boldsymbol{\theta}_k^- = \boldsymbol{A}_{k-1}\boldsymbol{\theta}_{k-1}^+ \tag{3.42}$$
$$\boldsymbol{P}_k^- = \boldsymbol{A}_{k-1}\boldsymbol{P}_{k-1}^+\boldsymbol{A}_{k-1}' + \boldsymbol{Q}_{k-1} \tag{3.43}$$

Correction

$$e_k = y_k - \boldsymbol{\phi}_k\boldsymbol{\theta}_k^- \tag{3.44}$$
$$\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{\phi}_k'(\boldsymbol{\phi}_k\boldsymbol{P}_k^-\boldsymbol{\phi}_k + \boldsymbol{R}_k)^{-1} \tag{3.45}$$
$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{\phi}_k')\boldsymbol{P}_k^- \tag{3.46}$$
$$\boldsymbol{\theta}_k^+ = \boldsymbol{\theta}_k^- + e_k\boldsymbol{K}_k. \tag{3.47}$$

There is a clear similarity between the RLS equations (3.36)-(3.39) and the Kalman filtering equations (3.42)-(3.47) for parameter estimation. The key difference is in the prediction step, Equation (3.43). The Kalman filter introduces a process noise term, $\boldsymbol{Q}$, which is added onto the propagation matrix, $\boldsymbol{P}$, at every iteration in order to inflate the covariance. $\boldsymbol{Q}$ introduces a fading memory, such that the large the propagation noise, the less weight is placed on the previous measurements of the system. By considering the measurement noise and the process noise separately the Kalman filer is often able to provide much better tracking of the system.

3.5 Structure detection

Structure detection is the task of finding the best subset of regressors and corresponding parameter estimates that describe the behaviour of the system. Of course, here, the term 'best' depends on many factors such as the models application and how the quality of the model is validated. This step is necessary in the identification of both linear and non-linear black box models, however, in the case of non-linear models the problem is significantly more difficult. For polynomial basis function type models, such as those dealt with in this thesis, the standard approach is to conduct a search amongst some predefined superset of possible model terms in order to arrive at a parsimonious description [16, 99, 113]. In the case of linear systems this super set is typically small allowing for a brute force approach in which each possible combination of basis functions can be checked against some criterion for measuring the model quality. By checking all possible combinations the 'best' structure is guaranteed to be found. For non-linear systems however, the search space rapidly becomes extremely large as the complexity of the system increases making this approach infeasible. In order to find a parsimonious system description it is therefore necessary to use structure detection methods.

To illustrate the difficulty of the task, consider the superset of all polynomial NARMAX model terms generated by a maximum polynomial order, $n_p = 3$, and dynamic order $n_y = n_u = n_e = 3$. The search space contains [14]

$$M = \frac{(n + n_p)!}{n! n_p!} \tag{3.48}$$

model terms, where $n = n_y + n_u + n_e = 9$ is the number of linear terms. For this relatively low order example the superset of basis functions contains $M = 220$ terms. Increasing the polynomial order to $n_p = 4$ increases the size of the set to $M = 715$ terms and $n_p = 5$ generates $M = 2002$. When dealing with large dynamic or polynomial orders the total number of regression terms to choose from rapidly become vast.

The structure detection problem falls into two main categories; linear and non-linear regression. Linear regression is justified when the model is in a linear-in-the-parameters form and the model terms are not a function of the residuals. If this is not the case non-linear regression is required.

3.5.1 Linear regression

A number of different approaches to structure detection for linear regression are available including:

Exhaustive Search

An exhaustive search as a method for structure detection has been mentioned previously. This involves comparing each possible combination of basis functions against some criterion to measure model quality. This approach is closely related to model selection, discussed in the following section, because it is essentially a large scale model comparison problem. Exhaustive methods are only available when the search space of the basis functions is very small, *i.e.* the dynamic and polynomial orders are small, because of the computational cost required in the model evaluations.

Forward selection

Forward selection involves adding basis functions to the model one at a time based on some criterion. The model is initialised as the empty set and each of the competing terms in the superset is tested to see how it increases with some measure of model quality. The term that increases the model quality the most is included in the model. The process is repeated until a stopping criterion is reached.

Backward elimination

For Backward selection the model is initialised as the superset of all possible model terms. Terms are iteratively pruned from the model based on some criteria. The iteration is terminated when a stopping criterion is reached. Commonly backward elimination is performed as a pruning step for models that are found using, for example, forward selection.

Stepwise Regression

Stepwise regression consists of an iterative combination of forward selection and backward elimination. The model is initialised as in the case for forward regression. After each forward regression step every term in the model set is tested and may be removed from the model via backward elimination. This is done because model terms that were added at a previous iteration may become redundant at a

future iteration. The process continues until a stopping criterion is reached. Stepwise regression can result in a better model structure than forward regression or backward elimination separately [22].

3.5.2 Structure detection algorithms and selection criteria

Structure detection algorithms are partially defined by the criteria used for including/removing model terms.

Orthogonal Least Squares

Orthogonal Least Squares (OLS) is a commonly used method for linear regression. Terms are selected based on their contribution to the maximisation of an error reduction ratio (ERR) [64]. An orthogonal decomposition of the regression matrix is made which allows for each term to be tested sequentially and independently. The ERR is a measure of the ability to explain the output variance exhibited by each model term given by

$$ERR_i = \frac{\sum_{i=1}^{N} g_i^2 w_i^2(t)}{\sum_{k=1}^{N} y_k^2 - \frac{1}{N}\{\sum_{k=1}^{N} y_k\}^2}. \tag{3.49}$$

where g_i is the parameter corresponding to the i'th auxilary orthogonal regressor, w_i.

Terms with a value below some predefined threshold can be considered as insignificant and pruned from the model. This method was extended to the forward regression orthogonalisation (FRO) algorithm in [17, 30]. FRO is a forward regression scheme in which the ERR criterion is used to select the model term to be added at each step. At each iteration the competing basis functions are re-orthogonalised with respect to the previously selected basis functions.

The ERR is dependent on the order that the terms are chosen and in practice the first few terms chosen often explain the majority of the output variance leading to a rapid drop off in ERR for further terms if the data is not well conditioned [99]. This can lead to important terms being overlooked or becoming sensitive to noise in the data. [6] provide an alternative method of dealing with fast sampled data by introducing the delta operator. Down sampling the data may also be advantageous for structure detection.

FRO has been used extensively in the literature for the system identification of many real world problems and is used frequently as a benchmark because of its widespread usage as well as the comprehensive body of literature developed for the identification of NARMAX models [14]. It will be used later as a benchmark

for the identification methods developed in this thesis.

Simulation error minimisation

An alternative approach is adopted by both Piroddi and Spinelli [99] and Billings and Mau [18]. In both the former references the authors consider the minimization of a simulated error. Piroddi and Spinelli name this the Simulation Error Reduction Ratio (SERR) and it is used as a criteria for term selection. Hence term selection is driven by the ability to predict future data, as given by

$$SERR_i = \frac{MSSE(\mathcal{M}_i) - MSSE(\mathcal{M}_{i+1})}{\frac{1}{N}\sum_{k=1}^{N} y_k^2} \tag{3.50}$$

where $\mathcal{M}_i$ is the model at the ith iteration and $\mathcal{M}_{i+1}$ is the model to be tested at the subsequent iteration and $MSSE$ is the mean squared error between the measured output and the simulated output. The method is designed to produce models that are more robust than those identified by prediction error criteria. It is reported to produce more accurate models in some cases for fast sampled data and supports a better long range prediction accuracy. The advantages posed comes at the price of the high computational cost inherent in repeated simulation and there is no efficient algorithm for their computation.

Sparse estimation with LASSO

A further alternative approach to the structure detection task is using sparse parameter estimation techniques to drive the parameters of irrelevant basis functions to zero. LASSO has been investigated as a method for direct structure detection [65] as well as a method reducing the search space [98]. LASSO is the addition of the l_1 norm as a regularisation term into the standard LS cost function, see Section 3.4.1. This has the property of forcing some of the parameters to zero. Structure detection can be performed by applying LASSO to a super-set of model terms and removing terms with a parameter weighting close to zero. It has been reported that LASSO alone results in an overly parametrised system description but can be used to significantly reduce the search space with little reduction in the prediction accuracy of identified models [98].

3.5.3 Non-linear regression

The linear regression structure detection methods mentioned above cannot be directly applied to the non-linear regression problem such as is required for the

structure detection of NARMAX models. This is because the regression matrix contains residual error terms which are themselves dependent on the model structure. A modification of the ERR based identification procedures must be used instead [64] This is achieved by the following

1. Perform the OLS estimation procedure for the process parameters and calculate the ERR assuming the residuals to be zero. Remove terms based on the ERR criterion with some user defined threshold. Re-estimate the parameters of the remaining model terms.

2. Calculate the residuals between the system output and the output of the model identified in step 1.

3. Estimate the parameters associated with the residual error, calculate the ERR and remove relevant terms.

4. repeat step 3 until convergence.

The noise model (basis functions containing the residuals) is therefore estimated after the process model has been identified.

3.6 Model validation

Once a model has been identified it is necessary to perform model validation tests in order to ensure the obtained model is adequate in its ability to describe the system behaviour. A number of validation test are available. For linear models there are two standard approached; Firstly to directly test the models fit to an independent validation data set. The second is to test the residuals for their correlation with the input. If the model is good then the residuals will form a white noise sequence. The test can be performed by computing the autocorrelation function and the cross-correlation function [75].

Most correlation based methods available for linear models are insufficient in the non-linear case. However, the concept remains the same, testing that the residuals are unpredictable and hence that all the predictable information in the data has been captured [14].

3.6.1 Model fit

Validation can be performed by testing the models ability to predict future outputs of a system given a known input and hence how well it represents the system dynamics. An estimation data set is used to identify the model and a validation

data set is used to test the models predictive ability by comparison to a simulated output.

The prediction error can be calculated as

$$\boldsymbol{e} = \boldsymbol{y} - \hat{\boldsymbol{y}} \tag{3.51}$$

where $\hat{\boldsymbol{y}}$ is the simulated output. The simulated output can be either one step ahead (OSA) or model predicted output (MPO). OSA predictions are made with knowledge of both the past inputs and past outputs up to the current time step. MPO is a model simulation using a known input and previous outputs generated by the model. Hence MPO gives a better indication of the models predictive ability.

The model performance can be assessed using various standard measures [75] such as the Mean Square Error (MSE)

$$MSE(\mathbf{e}) = \frac{1}{N}\sum_k e_k^2 \tag{3.52}$$

$$= \frac{1}{N}\boldsymbol{e}^T\boldsymbol{e} \tag{3.53}$$

or normalised mean squared error (NMSE)

$$NMSE(\boldsymbol{e}) = \frac{MSE(\boldsymbol{e})}{MSE(\boldsymbol{y})} \tag{3.54}$$

and variance accounted for (VAF)

$$VAF = \frac{1 - \mathrm{Var}(\boldsymbol{y} - \hat{\boldsymbol{y}})}{\mathrm{Var}(\boldsymbol{y})} \tag{3.55}$$

which quantifies how much of the variance in the data is explained by the model.

Depending on whether $\boldsymbol{e}$ is calculated from an OSA or MPO the MSE can be named mean square prediction error (MSPE) or the mean square simulation error (MSSE) respectively.

3.6.2 Correlation tests

A set of statistical correlation tests have been developed by Billings et al. in order to check that the residuals of the model are unpredicted from all linear and combinations of past inputs [13, 21]. Letting Φ_{ab} represent the correlation between a

and b, then the relevant correlation tests are:

$$\begin{aligned} \Phi_{ee} &= \delta_\tau \\ \Phi_{eu} &= 0, \quad \forall\tau \\ \Phi_{e(eu)} &= 0, \quad \tau >= 0 \\ \Phi_{eu^2} &= 0, \quad \forall\tau \\ \Phi_{e^2u^2} &= 0, \quad \forall\tau \end{aligned} \tag{3.56}$$

where $\boldsymbol{e}$ is the sequence of residuals and $\boldsymbol{u}$ is the input sequence. In practice the correlation functions will not equal zero because of a finite data length. The model is thus accepted if the correlation function lies within a 95% confidence interval ($\pm 1.96/\sqrt{N}$, where N is the data length). It should be noted that for fast sampled systems this method may lead to misleading results due to the correlation between samples.

3.7 Model selection

Model selection is the choice of a final model from a set of competing models. This is subtly different from structure detection in which the terms that constitute a model are chosen. Structure detection algorithms often provide a number of different competing models from which to choose. This might be the case from, for example, varying the ERR threshold in the FRO algorithm. The model selection task would then be to select the best model from those generated by varying the threshold. The objective of the system identification task as a whole is to identify a parsimonious description of a system that avoids over fitting to the data. This indicates that the criteria for assessing the model quality should somehow penalise complexity while promoting a good model fit. In Section 3.8.3 it will be shown how, by taking a Bayesian perspective, the penalisation of model complexity is naturally incorporated and hence problem of over fitting can be avoided. For now the trade-off between model complexity and accuracy will be discussed in another way, as the bias and variance trade-off.

3.7.1 The bias-variance trade-off.

Different sources of error in the modelling process lead to both bias and variance in the model predictions. The error due to bias is the difference between the expected model prediction and the true value. The error due to variance is the variability in the model prediction. In a perfect scenario both the bias and variance should be

minimised to zero, however, given a finite data limit it can be shown that the best choice of model will involve a trade-off between the two as reducing one leads to an increase in the other.

To define the concept mathematically consider the general system given by Equation (3.4) where the function $\hat{f}(\boldsymbol{x})$ is approximating the true function $f(\boldsymbol{x})$. The expected value of the squared error is given by

$$\boldsymbol{E}[(y - \hat{f}(\boldsymbol{x}))^2] \tag{3.57}$$

which can be decomposed into

$$\boldsymbol{E}[f(\boldsymbol{x}) - \hat{f}(\boldsymbol{x})] + Var[\hat{f}(\boldsymbol{x})] + Var[y] \tag{3.58}$$

$$\text{Bias}^2 + \text{Variance} + \text{Irreducible Error} \tag{3.59}$$

see appendix A.3 for derivation.

The squared error has been decomposed into three terms; a squared bias term, a variance term and an irreducible error term. This third term represents the noise in the true system which cannot be reduced. The task of minimising the squared error is therefore the task of minimising both the squared bias and the variance.

In the context of the basis function regression model with parameters estimated by least squares, the bias variance trade-off can be directly linked to the complexity of the model through the number of parameters. By increasing the number of model terms the flexibility of the model increases resulting in a better fit to the training data and hence a low bias. However this can result in over fitting (fitting to the noise) which increases the variance. Decreasing the amount of model terms leads to a decreased fit and hence a higher bias but may produce a smaller variance. The quality of the model as measured over a validation set will be poor in both the overly simplistic and overly complex case. The overly simple model cannot capture the true system behaviour and the overly-complex model is fitting to a random noise sequence. It is natural to conclude that the choice of the best model structure is at a trade-off between the bias and the variance which can be controlled through the final choice of model model structure.

Introducing a regularisation term into the estimator, such as in Section 3.4.1, represents another way to control the model complexity. Consider fitting the parameters of fixed model structure to different training data sets: Highly regularised parameters (a large value of regularisation constant) lead to a model with consistent outputs (low variance) over the data sets but a poor fit to the true system (high bias). A small regularisation constant leads to a model whose outputs provide a good fit to the training data and so perform well on average (low bias)

but are inconsistent (high variance) because of over fitting to the data. The bias-variance trade-off is therefore affected by both the number of terms in the model and by way of regularisation.

3.7.2 Metrics for model selection.

A number of metrics are available for comparing different models in order to select a final model structure. From the previous section it is clear that any measure of model quality must include some trade between model fit and model complexity. Metrics for performing this task are called information criterion. Popular information criteria are;
Akaike's Information Criterion (AIC) [5],

$$AIC = 2M - 2\ln(L) \tag{3.60}$$

Bayesian (Schwarz) Information Criterion (BIC) [110]

$$BIC = M\ln(N) - 2\ln(L) \tag{3.61}$$

and Final Prediction Error (FPE) [4]

$$FPE = L\frac{N+M}{N-M} \tag{3.62}$$

where L is the maximised value of the likelihood function, M is the number of model terms and N is the number of data points over which the model was estimated. The first two criteria are both similar in form, with the first term penalising the number of model terms and the second term decreasing with model complexity. The model with the smallest value is selected as the best.

3.8 Introduction to Bayesian inference

The remainder of this chapter provides an introduction to the concept of Bayesian inference and how it can be used to solve linear regression problems. Combining the product rule of probability, $p(A,B) = p(B|A)p(A)$ and with the symmetry property $p(A,B) = p(B,A)$ the following relationship is easily found

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}. \tag{3.63}$$

where A and B are two conditional random variables .

Equation (3.63) is known as Bayes' theorem and it is integral to many modelling techniques, particularly in the field of machine learning. It is simple to show that the denominator is given by

$$p(A) = \sum_{B} p(A|B)P(B) \tag{3.64}$$

by noting that the sum over B of the distribution $p(B|A)$ must equate to one as it represents all the possible outcomes. If the probability distributions in question are continuous the summation in (3.64) can be replaced with an integral with no loss of generality. The denominator of Bayes' theorem can therefore be expressed in terms of the quantities in the numerator. It acts as a normalisation constant, ensuring that the left hand side of Equation (3.63) is a proper probability distributional (that sums to one over all values of B).

Bayesian inference provides a different approach to the system identification methods discussed so far in this chapter. Up until this point in the chapter the uncertainty associated with the models that have been discussed has not been taken into account. Taking a Bayesian perspective provides a natural way for incorporating this uncertainty and leads to three main advantages in the context of system identification:

1. Uncertainty is incorporated naturally into the modelling process.
2. Prior beliefs and assumptions about the system can be included into the model via the prior distribution.
3. Model complexity is penalised automatically.

The remainder of this section gives an introduction to how these properties of Bayesian inference arise and how Bayesian inference can be be applied in the context of linear regression.

3.8.1 Bayesian linear regression

Lets start the discussion by applying Bayes' theorem to the inference of model parameters in a linear regression problem. Rewriting Equation (3.63) in terms of the parameters of the NARX model given by Equation (3.10) and the system output gives

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})}. \tag{3.65}$$

Looking at the terms on the right hand side of Equation (3.65), the first term in the numerator, $p(\boldsymbol{y}|\boldsymbol{\theta})$, is the *likelihood function* which was introduced in Section

3.4. The likelihood function encapsulates the model of the system and represents the probability that the data was generated by different choices of the parameter vector $\boldsymbol{\theta}$. The second term, $p(\boldsymbol{\theta})$, is the *prior* distribution of the model parameters. The prior represents any assumptions or prior knowledge of the parameters before observing the data.

The term on the left hand side, $p(\boldsymbol{\theta}|\boldsymbol{y})$, represents the *posterior* distribution of the parameters after observing the data. Noting that the denominator is constant, the posterior distribution is therefore proportional to the likelihood function multiplied by the prior distribution. If the both the posterior and prior distribution are members of the same family of distributions then the prior is *conjugate* to the likelihood function. The choice of conjugate distributions allows the posterior distribution to be calculated analytically, greatly simplifying the inference task. However in most problems of interest an analytical solution is not possible to achieve.

Defining the likelihood function, $p(\boldsymbol{y}|\boldsymbol{\theta})$, as before by Equation (3.24), a conjugate prior is given by the normal distribution

$$p(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_0, \boldsymbol{\Sigma}_0) \tag{3.66}$$

with mean $\boldsymbol{\theta_0}$ and covariance $\boldsymbol{\Sigma_0}$.

The posterior distribution can now be calculated and will be proportional to the product of the likelihood function and the prior distribution. Because of the choice of a conjugate normal prior distribution the posterior will also be a normal distribution. The posterior can therefore be evaluated exactly by the method of completing the square in the exponential from which the normalisation constant can be determined. The choice of a conjugate prior greatly simplifies the inference task by removing the need to perform the high dimensional integral required in calculating the denominator in Bayes' theorem.

The posterior distribution can be determined by the method of completing the square, see appendix A.4, such that

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \mathcal{N}(\boldsymbol{\theta}_N, \boldsymbol{\Sigma}_N) \tag{3.67}$$

where

$$\boldsymbol{\theta}_N = \boldsymbol{\Sigma}_N(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\theta_0} + \sigma_e^{-2}\Phi'\boldsymbol{y}) \tag{3.68}$$

$$\boldsymbol{\Sigma}_N^{-1} = \boldsymbol{\Sigma}_0^{-1} + \sigma_e^{-2}\Phi'\Phi \tag{3.69}$$

Instead of the point estimate determined by the frequentest approaches of ML

and LS the Bayesian approach results in a probability distribution over the parameters. A point estimate of the posterior distribution can be made by taking the maximum a posteriori (MAP) estimate by finding the mode of the posterior distribution. For the normally distributed posterior (3.67) the MAP estimate is simply $\boldsymbol{\theta}_{MAP=}\boldsymbol{\theta}_N$ given by Equation (3.68).

Choosing the infinitely broad prior distribution by setting $\Sigma_0 = \beta \boldsymbol{I}$ for $\beta \to \inf$ the MAP estimate is given by

$$\boldsymbol{\theta}_{MAP} = \boldsymbol{\Sigma}_N(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\theta}_0 + \sigma_e^{-2}\Phi'\boldsymbol{y}) \tag{3.70}$$

$$= (\cancelto{0}{\boldsymbol{\Sigma}_0^{-1}} + \sigma_e^{-2}\Phi'\Phi)^{-1}(\cancelto{0}{\boldsymbol{\Sigma}_0^{-1}}\boldsymbol{\theta}_0 + \sigma_e^{-2}\Phi'\boldsymbol{y}) \tag{3.71}$$

$$= (\Phi'\Phi)^{-1}\Phi'\boldsymbol{y} \tag{3.72}$$

$$= \boldsymbol{\theta}_{ML}. \tag{3.73}$$

such that the MAP estimate is equal to the ML estimate if there is no influence from the prior distribution. Conversely, setting $N = 0$ results in the posterior distribution being entirely defined by the prior distribution. Comparing equations (3.33) and (3.70) it can be seen that the MAP estimator for the Bayesian linear regression problem corresponds to ridge regression with the choice of hyper-parameters

$$\boldsymbol{\theta}_0 = 0 \tag{3.74}$$

$$\boldsymbol{\Sigma}_0 = \frac{1}{\lambda\sigma_e^2}\boldsymbol{I}. \tag{3.75}$$

If the full Bayesian treatment is considered prior distributions are placed over all of the hyper-parameters in the model, $\boldsymbol{\theta}_0$ and $\boldsymbol{\Sigma}_0$ as well as the model parameters $\boldsymbol{\theta}$. The marginalisation over all of these variables is intractable, *i.e.* no closed form solution exists.

3.8.2 The marginal likelihood

The complexity involved in the inference problem often arises in evaluating the denominator, which can be re-written as

$$p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \tag{3.76}$$

as before. $p(\boldsymbol{y})$ is commonly named the *marginal likelihood* because the parameters are marginalised (integrated or summed) out of the likelihood function. The integral over the entire parameter space in Equation (3.76) is often prohibitively difficult to compute. It is therefore necessary to approximate the integral. This is

commonly performed by Monte Carlo sampling methods methods such as Markov chain Monte Carlo [42]. More recently, approximate deterministic methods have gained in popularity, most notably expectation propagation [83] and variational Bayes' [11] which shall be discussed in detail in Chapter 7.

3.8.3 Bayesian model selection

The simple linear regression problem given by Equation (3.65) can be written more generally as

$$p(\Theta|y,\mathcal{M}) = \frac{p(\mathbf{y}|\Theta,\mathcal{M})p(\Theta|\mathcal{M})}{p(\mathbf{y}|\mathcal{M})}. \tag{3.77}$$

where all of the parameters associated with the model have been collected into the set Θ. The dependence on the model structure $\mathcal{M}$ has also been explicitly included.

The aim is to find the model structure that maximises the probability distribution over the different models for a given data set. The posterior distribution over the model structure conditional on the data can be found by a further application of Bayes' theorem such that

$$p(\mathcal{M}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathcal{M})p(\mathcal{M})}{p(\mathbf{y})}, \tag{3.78}$$

noting that the marginal likelihood in Equation (3.77) appears in the numerator on the right hand side. Assigning a non-informative prior in Equation (3.78) so that all models are considered equally likely a priori, and given that the denominator is constant for a given data set, the inference depends solely on the marginal likelihood and therefore

$$p(\mathcal{M}|\mathbf{y}) \propto p(\mathbf{y}|\mathcal{M}) \tag{3.79}$$

The most probable model is therefore the one that maximises the marginal likelihood, $p(\mathbf{y}|\mathcal{M})$. The marginal likelihood can be found by integrating over the parameters

$$p(\mathbf{y}|\mathcal{M}) = \int p(\mathbf{y}|\Theta,\mathcal{M})p(\Theta|\mathcal{M})\ \mathrm{d}\Theta. \tag{3.80}$$

Equation (3.80) gives insight into one of the key advantages of Bayesian methods - complex models are naturally penalised in the inference step [78]. This is because the integration over the parameter space penalises models with a greater complexity because complex models can a priori model a larger range of data sets [11, 78]. This property is often named Occam's razor referring to Occam's philosophy that simple descriptions should be chosen over complex ones.

To understand this concept consider three competing models, $\mathcal{M}_1$, $\mathcal{M}_2$ and

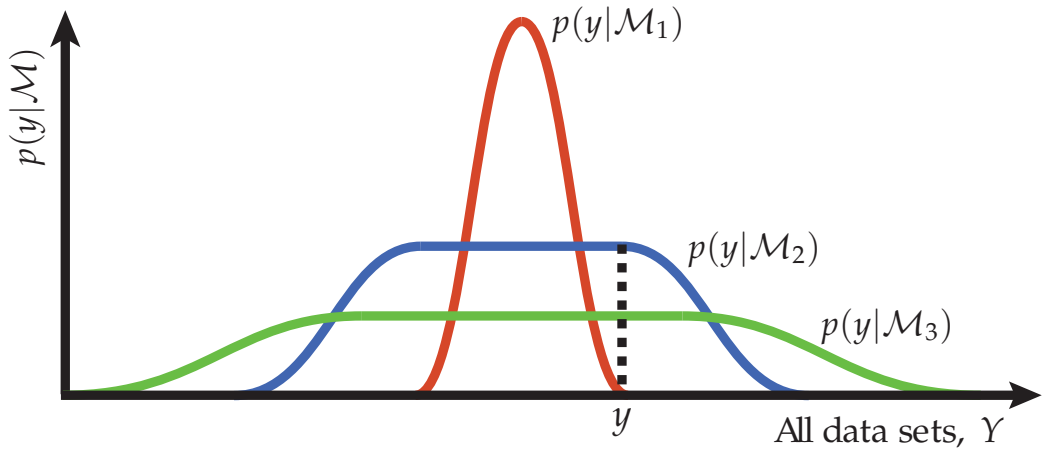

Figure 3.2: Occam's razor: Bayesian inference penalises complex models. Adapted from [11, 78]. The marginal likelihood $p(\mathbf{y}|\mathcal{M})$ is plotted against the space of all possible data sets Y. $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$ are models with increasing complexity giving rise to the probability distributions $p(\mathbf{y}|\mathcal{M}_1)$, $p(\mathbf{y}|\mathcal{M}_2)$ and $p(\mathbf{y}|\mathcal{M}_3)$ respectively. Because the marginal likelihood must integrate to 1 with respect to $\mathbf{y}$, the complex model $\mathcal{M}_3$, that models a larger range of data sets, must spread its probability distribution more. Model $\mathcal{M}_1$ can only model a small range of data sets so its probability distribution is highly peaked. The data set $\mathbf{y}$ falls within the probability distribution of $\mathcal{M}_2$ and $\mathcal{M}_3$ but not $\mathcal{M}_1$ and so the model $\mathcal{M}_2$ has the highest probability.

$\mathcal{M}_3$ with increasing complexity, giving rise to the probability distributions $p(\mathbf{y}|\mathcal{M}_1)$, $p(\mathbf{y}|\mathcal{M}_2)$ and $p(\mathbf{y}|\mathcal{M}_3)$ respectively. $\mathcal{M}_3$ is the most complex and as such can model a wide range of data sets. $\mathcal{M}_2$ and $\mathcal{M}_1$ are less complex and so can model a smaller range of data sets. The integral of the marginal likelihood over all possible data sets must integrate to one and so $\mathcal{M}_3$ will model a greater range of data sets with smaller probability and $\mathcal{M}_1$ will model a small range of data sets with greater probability. Thus if the data set $\mathbf{y}$ is outside of the probability range of $\mathcal{M}_1$ but inside that of $\mathcal{M}_2$ and $\mathcal{M}_3$ then $\mathcal{M}_2$ will be the most probable model [11]. It is therefore possible to reject both overly complex and overly simple models based on the marginal likelihood. A pictorial representation of Occam's razor is given in Figure 3.2.

Two competing models can be directly compared using the Bayes' factor which is the ratio of the marginal likelihoods [58]

$$B_{ij} = \frac{p(\boldsymbol{y}|\mathcal{M}_i)}{p(\boldsymbol{y}|\mathcal{M}_j)}. \tag{3.81}$$

A value of $B_{ij} > 1$ suggests that model $\mathcal{M}_i$ is more strongly supported by the

data than model $\mathcal{M}_j$ and its magnitude represents the strength of the evidence by the following vales

$2 \ln B_{ij}$	B_{ij}	Strength of evidence
0 to 2	1 to 3	Not worth a bare mention
2 to 6	3 to 20	Positive
6 to 10	20 to 150	Strong
> 10	> 150	Very strong

The Bayesian technique therefore encompass the bias-variance trade-off, discussed in Section 3.7, automatically via the marginal likelihood. It was shown in Section 3.8.1 how the inclusion of the prior distribution over the parameters leads to regularisation. In a fully Bayesian model in which prior distributions are placed over all of the hyper-parameters, the weight of the regularisation term can also be chosen automatically by the data. Complexity can therefore be penalised automatically through both the number of model terms and via regularisation.

3.8.4 Bayesian identification of NARMAX models

Despite the growing popularity of Bayesian inference and the widespread use of NARMAX models in system identification there is little precedent for NARMAX identification algorithms within a Bayesian framework. To the authors knowledge the only instance uses the reversible jump Markov chain Monte Carlo (RJMCMC) algorithm [44]

RJMCMC

The RJMCMC procedure has previously been used for the identification of linear autoregressive models [125] as well as for neural networks [33]. It was developed for the the identification of NARMAX models in [8]. The RJMCMC algorithm consists of moves between different model structures by either a birth or death move, such that terms are added or removed respectively.

The features of the RJMCMC approach to NARMAX identification are

1. *The birth move:* Sampling of competing model terms that are tested for inclusion in the model. This encourages a global search of the space of the model terms.

2. *The death move:* Sampling of selected model terms that are tested for exclusion from the model. A pruning step is thus incorporated into the procedure encouraging parsimony.

3. *Estimation of model and parameter distributions:* Distributions are naturally evaluated within a Bayesian framework.

4. *Bayesian model selection:* The Bayesian framework allows for complexity in the model structure to be automatically penalised.

The algorithm has been shown to compare favourably to standard approaches, such as FRO, in the identification of known benchmark problems [8] as well as for the identification of real world problems [121]

3.9 Summary

This chapter provides an overview of system identification and the system identification procedure. An introduction to the properties of both linear and non-linear systems is made. The entire identification procedure is discussed in detail and a number of different methods are introduced for the main steps: parameter estimation, structure detection, model selection and model validation. Following this, Bayesian inference is introduced providing a different perspective on the modelling task. By identifying models within a Bayesian framework uncertainty is naturally incorporated. In the next chapter non-linear system identification and frequency domain analysis methods are applied to the application of film-type DEAs.

Chapter 4

Nonlinear System Identification and Frequency Domain Analysis of Dielectric Elastomer Actuators

In the previous chapter DEAs were introduced as a new type of actuation device with huge potential for replacing traditional actuators such as motors for many applications, especially where robotic devices are interacting with humans because of the inherent compliance of the constituent materials. However, in order to realise the potential of these actuators the non-linearity and time-variation in the dynamic behaviour of DEAs necessitates the need for advanced control algorithms before these actuators can be successfully implemented [90, 133, 139]. A crucial step in achieving this is the development of techniques for control-oriented modelling and analysis.

In this chapter a novel control-focused analysis framework for DEAs is developed. The framework is based on the NARX model, which is discussed in detail in Section 3.3.2. In the context of control, the NARX model provides a number of advantages: it is able to produce parsimonious model descriptions with a high prediction accuracy that naturally incorporate a description of the plant non-linearities [91]. Many approaches are available for the structure detection of NARX models, a selection of which are discussed in Section 3.5. Here a simulation prediction error identification approach is taken in which term selection is driven by the models ability to predict unseen data, in-line with the objectives of this work.

A key attribute of the framework developed in this investigation, as well as throughout the thesis, is the use of non-linear frequency domain analysis techniques by both GFRF [20] and NOFRF [67] analysis. This allows, for the first time,

the frequency response analysis of non-linear DEA dynamics. Linear frequency domain methods are widely used in the field of control and provide a powerful tool for both implementing control architectures as well as analysis and interpretation of system dynamics [35]. The GFRF and NOFRF methodologies extends the use of frequency response analysis from linear to non-linear models.

The developed analysis framework is applied to the set of film-type DEA actuators presented in Section 2.3. This is done with an aim to demonstrate the applicability of the analysis framework as well as to investigate the insight that can be obtained from the non-linear frequency domain analysis. This analysis enables direct comparisons between the different actuators providing insight into their dynamic similarities and differences. This is especially interesting when DEAs are fabricated to the same specifications using identical fabrication methods.

The aims of this chapter are summarised below:

1 The identification of control focussed non-linear models of DEAs that are compact and accurate.

2 The use of identified non-linear models of DEAs in model based frequency domain analysis.

3 Demonstration of the analysis framework on a set of film-type DEAs, used to gain insight into the dynamic similarities and differences across this set of actuators.

Part of the work presented in this chapter has been published previously by the author in a peer reviewed journal [54].

The remainder of this chapter is structured as follows: In Section 4.1 the joint structure detection and parameter estimation used for identification of non-linear models for DEAs is introduced. If identified non-linear models contain a DC component it is necessary to remove this term from the model prior to performing frequency domain analysis, this is discussed in Section 4.2. In Section 4.3 advanced frequency domain analysis techniques are introduced. Section 4.4 collects the modelling and analysis methods into a framework which is then applied to a set of DEA actuators. The chapter is concluded with a discussion of the results and a summary in Sections 4.5 and 4.6 respectively.

4.1 Simulation based structure detection with the SEMP identification algorithm

The first step in the framework is the joint structure detection and parameter estimation of non-linear models of the DEA system of the NARX class. The model structure was identified in this work using a simulation based term selection algorithm named the *simulation error minimization with pruning* (SEMP) algorithm - which has been shown to provide greater discrimination between model terms than one-step-ahead predictive algorithms [99]. Term selection is driven by the reduction of the squared error over simulated model predictions using the MSSE encouraging the identification of models that perform well at long term prediction [99]. The criterion for assessing terms is the simulated error reduction ratio (SERR), given by Equation (3.50) and repeated here for clarity

$$SERR_i = \frac{MSSE(\mathcal{M}_i) - MSSE(\mathcal{M}_{i+1})}{\frac{1}{N}\sum_{k=1}^{N} y_k^2} \tag{4.1}$$

where $\mathcal{M}_i$ is the model at the ith iteration and $\mathcal{M}_{i+1}$ is the model to be tested at the subsequent iteration. $MSSE$ is the Mean Squared Simulation Error, given by

$$MSSE = \frac{1}{N}\sum_{k=1}^{N} \hat{e}_k^2 \tag{4.2}$$

where $\hat{e}_k = y_k - \hat{y}_k$, with $\hat{y}_k$ being the simulated system output at sample time k given by

$$\hat{y}_k = \sum_{m=1}^{M} \hat{\theta}_m \phi_m(\hat{\mathbf{x}}_k) \tag{4.3}$$

where

$$\hat{\mathbf{x}}_k = \left(\hat{y}_{k-1}, ..., \hat{y}_{k-n_y}, u_{k-1}, ..., u_{k-n_u}\right) \tag{4.4}$$

is a vector of lagged system input terms and simulated system output terms at sample time k.

The structure detection algorithm proceeds as follows: some initial model structure, $\mathcal{M}_0$, is chosen. Typically this is composed of the empty set or a linear basis of a predefined dynamic order. $SERR_0$ is initialised as

$$SERR_0 = \frac{MSSE(\mathcal{M}_0)}{\frac{1}{N}\sum_t y_k^2} \tag{4.5}$$

At the ith iteration a forward selection step is performed. The $MSSE_i$ is calculated

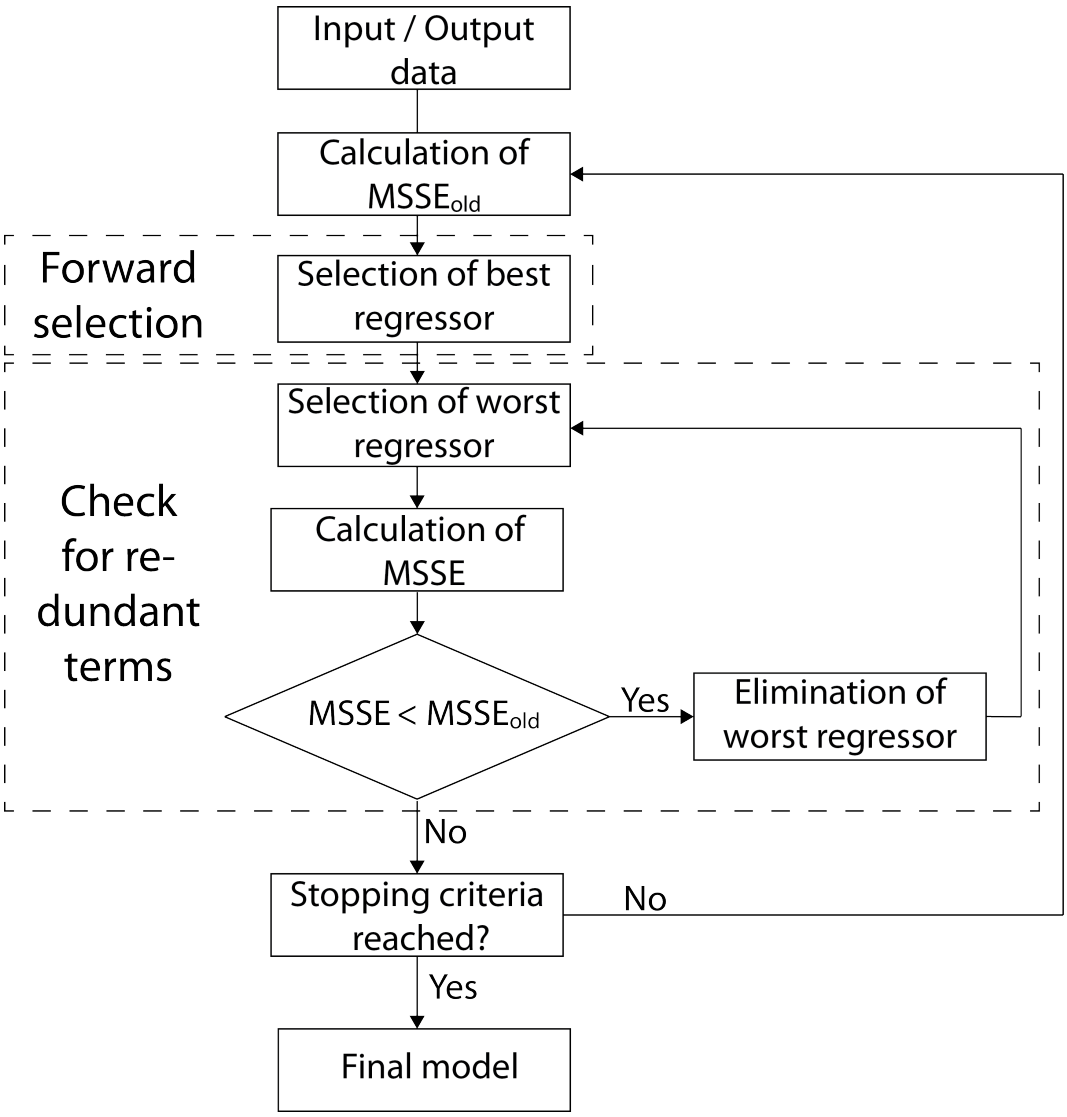

Figure 4.1: **The SEMP algorithm identifies models based on their prediction accuracy.** Schematic of the SEMP algorithm containing two distinct parts: Forward selection and removal of redundant terms at each iteration.

by Equation (4.1) for a proposed model $\mathcal{M}_{i+1}$. The proposed model $\mathcal{M}_{i+1}$ consists of the current model, $\mathcal{M}_i$, with the addition of a basis function from the set of all remaining basis functions, where each new basis function is tested one at a time. The model $\mathcal{M}_{i+1}$ that produces the minimum value of $MSSE_{i+1}$ is then chosen and the newly selected basis function is removed from the set of remaining basis functions. After a new term has been selected a check for redundant terms is performed. Each term in the proposed model $\mathcal{M}_{i+1}$ is removed and the MSSE is calculated. The worst basis function is selected as the one that produces the minimum of the MSSE. If the MSSE taken for the worst term is less than $MSSE_i$

then the term is actually removed from the proposed model and returned to the set of potential model terms and another check for redundant terms is performed. The model is updated as $\mathcal{M}_i = \mathcal{M}_{i+1}$ with $MSSE_i = MSSE_{i+1}$ where $MSSE_{i+1}$ is the MSSE of the newly selected model and the SERR is calculated by Equation (4.1). The algorithm can be terminated when $MSSE_{i+1} > MSSE_i$ indicating that the proposed model decreases the prediction accuracy. For each proposed model the parameters are necessarily re-estimated using LS and simulated over the data set in order to assess the MSSE. The algorithm is shown schematically in Figure 4.1.

Due to a finite data limit the algorithm may not terminate after the correct model structure is found and redundant terms may be selected that are over fitting to the noise. The SERR can be used to assess which terms should be included in the model by assessing the relative magnitudes of the SERR values.

If the mean level of the system output is non zero then a constant (or DC) term should be included into the structure detection as biased parameters may be estimated by LS if the DC component is neglected [14]. This is simply achieved by including a basis function into the superset of basis functions, with output unity for any excitation. The inclusion of a DC term in the final model structure affects subsequent analysis and so needs to be taken into account.

4.2 Treatment of mean levels in non-linear systems

Many linear and non-linear systems have a non zero mean output. If this is the case it may be necessary to take this mean level into account in any subsequent modelling and analysis of the system [95].

The treatment of mean levels for linear systems is well understood and it is common practice to remove mean levels from the data or to explicitly introduce a DC component into the model [1, 75]. The DC component might either represent an additive disturbance to the system, in which case it may be removed, or if it represents some aspect of the true system behaviour, it can be incorporated into the systems analysis. Either task is justified because linear systems do not shift energy to different frequencies so any DC (zero frequency) input only effects the zero mean output.

In the domain of non-linear systems simply removing mean levels from the data cannot be justified and the presence of a DC component in a model introduces further complexities into systems analysis. [95] show how the application of techniques that are common in the pretreatment of data for linear analysis, such as pre filtering or mean level removal, may lead to the identification of an incorrect

model structure or the estimation of biased parameters. In the paper the authors show that these problems can be addressed by explicitly including a constant term into the model structure.

The NARMAX model structure as well as the frequency domain analysis techniques introduced initially in Section 4.3 are all based on the Volterra model description. Volterra series models are valid around an equilibrium point y_0. For NARX models with no DC term the model will have at least one zero equilibrium point, $y_0 = 0$, so that zero input leads to zero output. However, the presence of a DC term in the model acts to propagate the equilibrium point through the system and this must either be taken into account in the modelling/analysis or removed from the model taking into account the way in which it interacts with the non-linear terms [95]. In the previous citation the authors present an 'unravelling' algorithm for removing the DC component from a model. The algorithm is not repeated here, however the concept is simple and is explained by way of an example.

Example 4.1 Removal of a DC term from NARX models.
Consider the following NARX model which explicitly includes a DC term, θ_1, in the parameter vector

$$y_k = \theta_1 + \theta_2 y_{k-1} + \theta_3 u_{k-1} + \theta_4 y_{k-1}^2 + \theta_5 y_{k-1} u_{k-1}. \tag{4.6}$$

The system is operating around some equilibrium position y_0, which can be equivalent to the steady state of the system but in general may take other values [114]. Shifting the system so it is operating around the zero equilibrium can be performed by substituting y_k with $y_k^* + y_0 \; \forall \; k$ such that

$$y_k^* + y_0 = \theta_1 + \theta_2(y_{k-1}^* + y_0) + \theta_3 u_{k-1} + \theta_4(y_{k-1}^* + y_0)^2 + \theta_5(y_{k-1}^* + y_0)u_{k-1}. \tag{4.7}$$

which can easily be rearranged so that it takes the same form as (4.6)

$$\begin{aligned} y_k^* \;\; = \;\; & \theta_4 y_0^2 + (\theta_2 - 1)y_0 + \theta_1 + y_{k-1}^*(\theta_2 + 2\theta_4 y_0) \\ & + u_{k-1}(\theta_3 + \theta_5 y_0) + \theta_4 y_{k-1}^* + \theta_5 y_{k-1}^* u_{k-1}. \end{aligned} \tag{4.8}$$

Because this new system $y_k^* = y_k + y_0$ is operating around the zero equilibrium position, y_k^* must be zero when $u_k = 0 \; \forall \; k$ which can only be true if there is no constant term present in the model. It hence follows that the appropriate value of y_0 can be determined by solving the quadratic equation formed by equating the constant terms, coloured red in Equation (4.8) to zero. This results in a NARX

model with same form as Equation (4.6) with an altered parameter vector

$$y_k^* = \theta_2^* y_{k-1}^* + \theta_3^* u_{k-1} + \theta_4^* y_{k-1}^{*2} + \theta_5^* y_{k-1}^* u_{k-1}. \tag{4.9}$$

where

$$\boldsymbol{\theta}^* = \left[0, \theta_2 + 2\theta_4 y_0, \theta_3 + \theta_5 y_0, \theta_4, \theta_5\right]^T \tag{4.10}$$

△

It should be noted that in the above example only the parameter vector was modified, but in general new terms may be introduced. The algorithm given in [95] requires knowledge of the value y_0 which can always be found directly by setting all $u(\cdot) = 0$ and $y(\cdot) = y_0$ and solving the resultant equation for y_0.

4.3 Analysis of non-linear systems in the frequency domain

The analysis of systems in the frequency domain is of fundamental importance across many disciplines of science, and particularly so in engineering. Frequency domain descriptions of linear systems are the basis for many methods for model analysis and control design [35, 75]. Many of these methods are based on the linear FRF $H_1(\omega)$ introduced in Section 3.1.1. The linear FRF provides a complete description of the behaviour of the linear system and is therefore invariant of the time domain model used to describe the system.

In the non-linear case the situation is vastly more complicated and no single FRF can be used to completely characterise a non-linear system. Despite this it is perhaps surprising that relatively little research exists for the frequency domain analysis of non-linear systems, given that this approach is so widely used in the linear case.

One approach is to consider the extension of the linear FRF to higher orders as the GFRF, $H_n(\omega_1, \ldots, \omega_n)$. In order to completely describe the behaviour of a non-linear system, GFRFs up to order N_m must be considered, where in some cases N_m may be infinite. One GFRF is used to describe each order of non-linearity displayed by the system. If $N_m = \infty$ the approximation $N_m \approx N_m^*$ is made, where N_m^* is the maximum order under consideration. As in the linear case, the GFRFs provide an invariant description of the underlying system behaviour, given that the model accurately described the system. More importantly, GFRFs can be used to describe complex non-linear behaviour in a much more transparent and physically meaningful way than is possible in the time domain [14, 20].

GFRF analysis is based on the Volterra series and are defined as the Fourier

transform of the Volterra series kernels [107]. The Volterra series model is given by

$$y(t) = \sum_{n=1}^{\infty} y_{n,t} \tag{4.11}$$

where

$$y_{n,t} = \int_{-\infty}^{\infty} \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^{n} u_{t-\tau_i} \mathrm{d}\tau_i \tag{4.12}$$

is the n'th order output of the system and $h_n(k_1, \dots, k_n)$ is the n'th order Volterra kernel, or equivalently the n'th order impulse response function, of the system. Taking the Fourier transform of $h_n(\tau_1, \dots, \tau_n)$ produces an expression for the n'th order GFRF:

$$H_n(j\omega_1, \dots, j\omega_n) = \int_{-\infty}^{\infty} \dots \int h_n(\tau_1, \dots, \tau_n) e^{-j(\omega_1\tau_1 + \dots + \omega_n\tau_n)} \mathrm{d}\tau_1 \dots \mathrm{d}\tau_n. \tag{4.13}$$

The function H_n provides the basis for all of the non-linear frequency domain analysis techniques that will be discussed in this thesis. In this thesis the notation H_n will be reserved exclusively for the linear FRF when $n = 1$ and the n'th order GFRF when $n > 1$.

Computation of GFRFs from identified NARX models: The probing method

Efficient computation of GFRFs can be performed by recursive algorithms, however, here an analytic approach based on probing [20] will be presented as it provides a more transparent insight into how the GFRFs are calculated and how they relate to the original NARMAX description.

The Volterra kernel and corresponding GFRF are not guaranteed to be symmetric i.e. the function is dependent on the order of its arguments. Both functions can be replaced by their symmetric counterpart such that [107]

$$h_n(\tau_1, \dots, \tau_n)_{symmetric} = \frac{1}{n!} \sum_{\substack{\text{all} \\ \text{permutations}}} h_n(\tau_1, \dots, \tau_n)_{asymmetric} \tag{4.14}$$

$$H_n(j\omega_1, \dots, j\omega_n)_{symmetric} = \frac{1}{n!} \sum_{\substack{\text{all} \\ \text{permutations}}} H_n(j\omega_1, \dots, j\omega_n)_{asymmetric} \tag{4.15}$$

which will be convenient for the later analysis. For notational simplicity the subscript will be dropped and both Volterra kernels and GFRFs will be assumed to be asymmetric unless otherwise stated.

GFRFs relating to the Volterra kernels in Equations (4.11-4.12) can be found using the probing method [12] named because the system is 'probed' by an ex-

perimental input. Consider the system given by Equation (4.12) excited by the input

$$u_t = \sum_{q=1}^{Q} \mathbf{A}_q \, \mathrm{e}^{j\omega_q t} \tag{4.16}$$

where $\mathbf{A}_q$ is the amplitude of the signal and $\omega_q \in \mathbb{R}$. Substituting into Equation (4.12)

$$\begin{aligned} y_t &= \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^{n} u_{t-\tau_i} \sum_{q=1}^{Q} \mathbf{A}_q \, \mathrm{e}^{j\omega_q (t-\tau_i)} \, \mathrm{d}\tau_i \qquad (4.17) \\ &= \sum_{n=1}^{\infty} \sum_{q=1}^{Q} \dots \sum_{q=1}^{Q} \prod_{i=1}^{n} \mathbf{A}_q \, \mathrm{e}^{j\omega_{q_i} t} \int_{-\infty}^{\infty} \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^{n} \mathrm{e}^{-j\omega_{q_i}\tau_i} . \qquad (4.18) \end{aligned}$$

Then noting that the latter half of the previous equation can be substituted for Equation (4.13) yields the expression

$$y_t = \sum_{n=1}^{\infty} \sum_{q_1=1}^{Q} \dots \sum_{q_n=1}^{Q} \left[\mathbf{A}_{q_1} \dots \mathbf{A}_{q_n} H_n(j\omega_{q_1}, \dots, j\omega_{q_n}) \right] \mathrm{e}^{j(\omega_{q_1} + \dots + \omega_{q_n})t} \tag{4.19}$$

which in the case that $Q = n$, $\mathbf{A}_Q = 1$ for all $q = 1, 2, \dots, n$ reduces to

$$y_t = \sum_{n=1}^{\infty} \sum_{q_1=1}^{n} \dots \sum_{q_n=1}^{n} H_n(j\omega_{q_1}, \dots, j\omega_{q_n}) \, \mathrm{e}^{j(\omega_{q_1} + \dots + \omega_{q_n})t} . \tag{4.20}$$

Note that the summation in (4.20) produces $n!$ terms with frequency $\omega_1 + \omega_2 + \dots + \omega_n$ corresponding to all the permutations over the possible frequencies. For a system that can be described by the Volterra series (4.11-4.12) a harmonic input must produce a harmonic output [14]. Therefore the system GFRFs, H_n, can be found by substituting (4.20) and (4.16) into the NARX model equation and comparing coefficients of $\mathrm{e}^{j(\omega_1 + \dots + \omega_n)t}$.

It is important to note that this method is not valid for NARX models that contain a DC term in the model structure. Although there are methods for incorporating this term into the analysis [23, 56, 95], an equally valid and simpler method is to transform the model such that the DC term is removed, by for example using the method described in Section 4.2.

In order to apply the GFRF concept to the NARX model it is necessary for them to be described in discrete time. The relationship between the continuous and discrete time case is simply given by

$$H_n(j\omega_1, \dots, j\omega_n) = H_{n,discrete}(jT_s\omega_1, \dots, jT_s\omega_n) \tag{4.21}$$

where T_s is the sampling period.

Because of the simplicity of the transformation between the two cases the dependence on T_s will be assumed and dropped from subsequent equations such that functions in both continuous and discrete time are given by $H_n(j\omega_1, \ldots, j\omega_n)$.

Example 4.2 Calculation of GFRFs using the probing method.

Consider the NARX model given by the following equation

$$\begin{aligned} y_k &= \theta_1 + \theta_2 y_{k-1} + \theta_3 y_{k-2} \\ &\quad + \theta_4 u_{k-1} + \theta_5 y_{k-2}^2, \end{aligned} \tag{4.22}$$

where

$$\boldsymbol{\theta} = [-0.1, 1, -0.3, 0.1, 0.5]^T. \tag{4.23}$$

The system is used to generate noise free input-output data. The input, u_k, is a white noise sequence drawn from the zero mean normal distribution with variance $\sigma_u^2 = 0.25$.

By setting $u_k = 0$ and $y_k = y_0 \ \forall k$ and solving, the equilibrium point, y_0, can be found from the equation

$$0 = y_0^2\theta_5 + y_0(\theta_2 + \theta_3) + \theta_1, \tag{4.24}$$

which can be solved to give the two solutions: $y_{0,1} = -0.2385$, $y_{0,2} = 0.8385$. The solution $y_{0,1}$ is close to the mean value of the system output and is chosen for the value of y_0. Removing the DC component from the NARX model via the method described in Section 4.2 leads to the modified parameter vector

$$\boldsymbol{\theta}^* = \begin{bmatrix} 0 \\ \theta_2 \\ \theta_3 - 2y_0\theta_5 \\ \theta_4 \\ \theta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0.5385 \\ 0.1 \\ 0.5 \end{bmatrix} \tag{4.25}$$

$$\begin{aligned} \theta_1^* = 0,\ \theta_2^* = 1,\ \theta_3^* = 0.5385, \\ \theta_4^* = 0.1,\ \theta_5^* = 0.5. \end{aligned} \tag{4.26}$$

Applying the input given by equation (4.16) with $Q = 1$ and $\boldsymbol{A_k} = 1$ such that

$$u_k = e^{j\omega_1 k}, \tag{4.27}$$

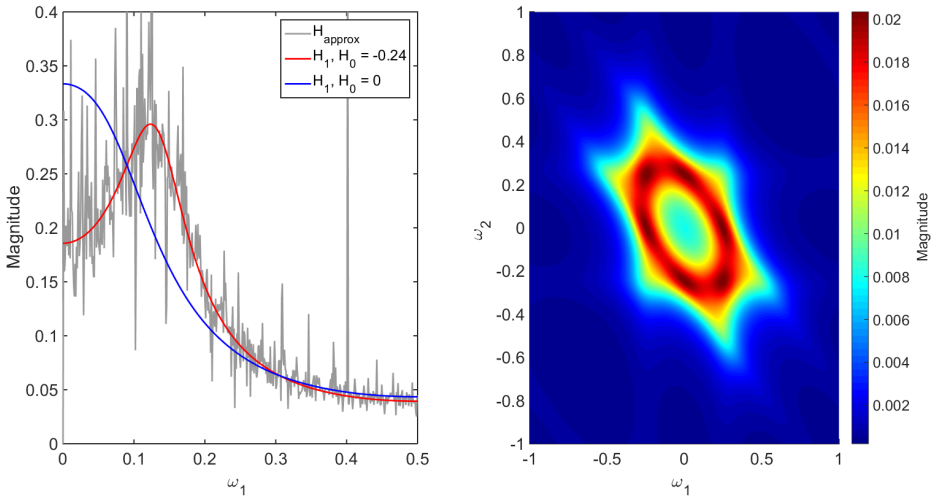

Figure 4.2: **GFRFs can be used to analyse the frequency domain characteristics of a non-linear system. It is necessary to consider any DC component included in the model. A**) The first order GFRF H_1, equivalent to the well known FRF, for the model given by Equation (4.22) is shown taking into account the DC component (Red) and ignoring the DC component (Blue). An approximation of the first order transfer function, assuming a linear system, is given for comparison (Grey). **B**) Second order GFRF, H_2, for the model given by Equation (4.22) the color map indicates the gain at combinations of frequencies ω_1 and ω_2.

it follows that the lagged input can be given by

$$u_{k-\tau} = e^{j\omega_1(k-\tau)}. \tag{4.28}$$

A harmonic input must lead to a harmonic output. The output response can therefore be found from Equation (4.20) where the infinite upper limit in the summation over the non-linear orders can be replaced with a summation up to the maximum order under consideration. This is because only these terms display a coefficient that can match that of the input in the next step. The output is hence given by

$$y_{k-\tau} = H_1(j\omega_1)e^{j\omega_1(k-\tau)}, \tag{4.29}$$

where the lagged output follows as with the lagged inputs. Substituting (4.28) and (4.29) into (4.22) and rearranging gives the first order GFRF

$$H_1^*(j\omega) = \frac{\theta_4^* e^{-j\omega}}{\theta_2^* e^{-j\omega} + \theta_3^* e^{-j\omega} - 1}. \tag{4.30}$$

Similarly, setting $Q = 2$ in Equation (4.16) gives the input

$$u_{k-\tau} = e^{j\omega_1(k-\tau)} + e^{j\omega_2(k-\tau)} \tag{4.31}$$

and the output response is again found from Equation (4.20) as

$$\begin{aligned} y_k = H_1(j\omega_1)e^{j\omega_1 k} + H_1(j\omega_2)e^{j\omega_2 k} + H_2(j\omega_1, j\omega_2)e^{j(\omega_1+\omega_2)(k-\tau)} \\ + H_2(j\omega_1, j\omega_1)e^{j2\omega_1(k-\tau)} + H_2(j\omega_2, j\omega_2)e^{j2\omega_2(k-\tau)}. \end{aligned} \tag{4.32}$$

Substituting (4.31) and its lagged forms as well as (4.32) into Equation (4.22), equating coefficients of $e^{(j(\omega_1+\omega_2)k)}$ and rearranging for H_2 leads to the asymmetric second order GFRF

$$H_2(j\omega_1, j\omega_2) = \frac{\theta_4 0.5 H_1(\omega_1)H_1(\omega_2)e^{-j(2\omega_1+2\omega_2)}}{\theta_2 e^{-j(\omega_1+\omega_2)} + \theta_3 e^{-j2(\omega_1+\omega_2)} - 1}. \tag{4.33}$$

The GFRFs given above are asymmetric. The symmetric GFRFs can be easily found from the asymmetric GFRFs via Equation (4.15). Further higher order GFRFs can be calculated by following the same procedure whilst increasing the value of K in the input Equation (4.16). Terms with coefficients matching that of the input will always be able to be found in equation (4.20) if there is either a pure non-linear output term or a cross term in the NARX model. In order to completely characterise the system it is therefore necessary to calculate infinite higher order GFRFs. In practice however the higher order GFRFs will rapidly become insignificant with increasing order. For example, the system given by Equation (4.22) has no third order terms so the contribution to the third order GFRF will come entirely from the second order term y_{k-2}^2. Considering the terms generated by Equation (4.20) with the summation from $n = 1, \ldots, 3$, substituting this into the model equation and comparing coefficients, a factor of $n!$ will divide the contribution of this term in the GFRF equation. With increasing n the contribution of this term will rapidly decrease to a negligible value.

The GFRFs are functions of n different frequency inputs, making them multi dimensional functions for $n > 1$. This has some important implications. Firstly, evaluating the GFRFs over a frequency range becomes increasingly computationally cumbersome as the order increases. Secondly, for $n > 2$ the GFRFs are hard to visualise making them hard to interpret. This has been resolved to some extent in [141] in which slices through the high dimensional space are investigated by setting some of the frequency inputs equal.

Note that if the constant term is not removed from the model, and a method for determining the GFRFs that took this term into account is used, the resulting

first order GFRF is given by

$$H_1(j\omega) = \frac{\theta_4 e^{-j\omega}}{\theta_2 e^{-j\omega} + (\theta_3 + 2H_0\theta_5)e^{-j\omega} - 1}, \tag{4.34}$$

which is equivalent to Equation (4.30) with $\theta_3^* = \theta_3 + 2H_0\theta_5$ which matches the parameter vector, Equation (4.25). Blindly removing the constant term from Equation (4.30) leads to the First order GFRF given by Equation (4.34) with $H_0 = 0$. Figure 4.2**A** shows the first order frequency response for both the case when the constant term has been correctly included in the analysis and when it has not been included . An approximate FRF, assuming a linear system, has also been calculated as $H_{approx} = Y/U = FFT(y)/FFT(u)$ from the simulated data and also plotted. The difference in the two responses emphasises the necessity of the inclusion of any additive constants in the frequency domain analysis of non-linear systems.

The magnitude of H_2 is much smaller than for H_1 at all frequencies, see Figure 4.2**B**. It can also be seen from Figure 4.2**A** that the first order frequency response H_1 very closely follows the approximation H_{approx}. Considering these two statements it could be naively concluded that the first order dynamics dominate system behaviour and thus neglecting the non-linearity would lead to a reasonable approximation. However, it will be shown later in this chapter that this conclusion is incorrect and by considering the output frequency response a much better understanding can be obtained. △

The example demonstrates how GFRFs of a non-linear system of the NARX class can be calculated.

The general n'th order GFRF is given by

$$H_n(\omega_1, \ldots, \omega_n) = \frac{H_{n_u}(\omega_1, \ldots, \omega_n) + H_{n_{uy}}(\omega_1, \ldots, \omega_n) + H_{n_y}(\omega_1, \ldots, \omega_n)}{1 - \sum_p \theta_p e^{-j(\omega_1, \ldots, \omega_n)k_p}}, \tag{4.35}$$

where the summation over p indicates the summation over all the linear output terms with corresponding lag k_p. The terms H_{n_u}, $H_{n_{uy}}$ and H_{n_y} represent contributions to the n'th order GFRF from pure input, cross and pure output terms respectively. A number of iterative algorithms exist for determining these terms in an efficient manner [23, 55, 56].

GFRF analysis of a system helps to interpret how input frequencies combine to create complex non-linear effects at the output. For linear FRFs analysis is often focused on identifying resonant peaks, which indicate areas of the output

spectrum which will be strong. GFRF analysis of non-linear systems is analogous to the linear case but analysis focuses on the identification of ridges as well as peaks in the GFRF. The location and direction of the ridges then indicate frequency combinations that will produce strong harmonic or intermodulation effects at the output [20].

GFRF analysis can therefore be extremely valuable in understanding non-linear behaviour. However, they are not able to provide a complete description of how the output spectrum of a non-linear system is composed. In order to achieve this, the input spectrum must be considered and built into the analysis framework, motivating the discussion in the following sections.

From the concepts introduced in this chapter and in light of the given example a few important remarks regarding GFRFs can be made

1 GFRFs provide a mapping from the time-domain to the frequency-domain for models of the NARX class and as such can be seen as an extension of the well known linear FRF to the non-linear case.

2 GFRFs allow analysis of non-linear systems not possible with traditional linear methods.

3 GFRFs are multi-dimensional functions, they are therefore hard to visualise for $n > 2$.

4 In order to provide the correct results, the DC component of a NARX model must be removed or taken into account in the GFRF analysis.

4.3.1 Output frequencies of non-linear systems

The output frequency response of a LTI system can be expressed as

$$Y(j\omega) = H(j\omega)U(j\omega), \tag{4.36}$$

where $U(j\omega)$ and $Y(j\omega)$ are the spectra of the input and output respectively. Given knowledge of the FRF $H(j\omega)$ the output response of a linear system is completely characterised by Equation (4.36) which remains valid for any input signal.

In the realm of non-linear systems, the situation is much more complicated and Equation (4.36) no longer holds. Each non-linear order of the system contributes to the output spectra and, unlike linear systems, these are not confined to the frequency range of the input signal, resulting in a much richer output frequency spectrum than that of the input.

The Volterra series model used as the basis for the GFRF analysis in the previous section will be employed again here. The model is redefined by combining the equation pair (4.11-4.12) such that

$$y_t = \sum_{n=1}^{N_m} \int_{-\infty}^{\infty} \dots \int h_n(\tau_1, \dots, \tau_n) \prod_{i=1}^{n} u(t-\tau_i) \mathrm{d}\tau_i \tag{4.37}$$

where the infinite upper bound of the summation over the non-linear order, n, has been replaced by some maximum order, N_m, to reflect a practical analysis scenario.

Now, taking the inverse Fourier transform of the n'th order GFRF, H_n, yields an expression for the n'th order Volterra kernal

$$h_n(\tau_1, \dots, \tau_n) = \int_{-\infty}^{\infty} \dots \int H_n(j\omega_1, \dots, j\omega_n) e^{j(\omega_1\tau_1 + \dots + \omega_n\tau_n)} \mathrm{d}\omega_1 \dots \mathrm{d}\omega_n. \tag{4.38}$$

then substituting this back into Equation (4.37)

$$y_t = \sum_{n=1}^{\infty} \frac{1}{(2\pi)^n} \int_{-\infty}^{\infty} \dots \int H_n(j\omega_1, \dots, j\omega_n) \prod_{i=1}^{n} U(j\omega_i)\, \mathrm{e}^{j(\omega_1 + \dots + \omega_n)t}\, \mathrm{d}\omega_i \tag{4.39}$$

which follows by decomposing the sum in the exponential into a product of exponentials and using the definition of the Fourier transform given by Equation (3.3).

From Equation (4.39) it can be shown that the output frequency response of a non-linear system that is stable at zero equilibrium and that can be described by the Volterra series model is given by [68]

$$\begin{aligned} Y(j\omega) &= \sum_{n=1}^{N_m} Y_n(j\omega) \\ Y_n(j\omega) &= \frac{n^{-1/2}}{(2\pi)^{n-1}} \int_{\omega} H_n(j\omega_1, \dots, j\omega_n) \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega} \end{aligned} \tag{4.40}$$

where H_n is the n'th order GFRF given by Equation (4.13), ω is the hyperplane defined by $\omega = \omega_1 + \dots + \omega_n$ and $\mathrm{d}\sigma_{n\omega}$ denotes the area of a minute element on the hyperplane ω.

Equation (4.40) shows that the output frequencies of a non-linear system can be represented as a sum over the output contributions from each non-linear order $Y_n(j\omega)$. The frequency range, f_Y of the output for nonlinear systems can be much greater than for linear systems and can be found by considering the output frequency range at each non-linear order, f_{Y_N}, such that

$$f_Y = \bigcup_{n=1}^{N} f_{Y_N} \tag{4.41}$$

A general expression for the output frequency range of the n'th order non-linearity, f_{Y_n}, due to an input frequency in the range $[a, b]$ was derived in [69] and can be determined by the following set of equations

$$\begin{aligned} f_{Y_n} &= \begin{cases} \bigcup_{k=0}^{i^*-1} I_k, & \text{when } \frac{nb}{(a+b)} - \left\lfloor \frac{na}{(a+b)} \right\rfloor < 1 \\ \bigcup_{k=0}^{i^*} I_k, & \text{when } \frac{nb}{(a+b)} - \left\lfloor \frac{na}{(a+b)} \right\rfloor \geq 1 \end{cases} \\ i^* &= \left\lfloor \frac{na}{(a+b)} \right\rfloor + 1 \\ I_k &= [na - k(a+b), nb - k(a+b)],\ k = 0, 1, \ldots, i^* - 1 \\ I_{i^*} &= [0, nb - i^*(a+b)]. \end{aligned} \tag{4.42}$$

along with Equation (4.41). Although this allows the calculation of the possible output frequency range, it provides no information regarding exactly what the output frequencies will be. In order to find out the exact output frequencies the concept of NOFRFs will be introduced.

4.3.2 Non-linear output frequency response functions (NOFRFs)

The output frequency response of the system to a general input can be expressed by equations (4.40). Given both a static linear and nonlinear system respectively given by

$$y_k = ku_k \tag{4.43}$$

$$y_k = ku_k^2, \tag{4.44}$$

then it follows that the frequency domain representations are given by

$$Y(j\omega) = kU(j\omega) \tag{4.45}$$

$$Y(j\omega) = kU_2(j\omega), \tag{4.46}$$

respectively, where U_2 is the Fourier transform of u_t^2 and can be expressed as

$$U_2(j\omega) = \frac{2^{-1/2}}{(2\pi)^{(2-1)}} \int\limits_{\omega_1+\omega_2=\omega} \prod_{i=1}^{2} U(j\omega_i) \mathrm{d}\sigma_{\omega 2} \tag{4.47}$$

which is given in terms of the input spectrum $U(j\omega)$. Equation (4.47) can hence be regarded as an extension of the input spectrum $U(j\omega)$ to the second order non-linear case. It then follows that the general expression for the n'th order non-linear case is given by

$$U_n(j\omega) = \frac{2^{-1/n}}{(2\pi)^{n-1}} \int\limits_{\omega_1+\ldots,\omega_n=\omega} \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{\omega n}. \tag{4.48}$$

Equation (4.40) can be then be rewritten in terms of Equation (4.48) as

$$\begin{aligned} Y_n(j\omega) &= \frac{\int\limits_{\omega_1+\ldots+\omega_n=\omega} H_n(j\omega_1,\ldots,j\omega_n) \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}}{\int\limits_{\omega_1+\ldots+\omega_n=\omega} \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}} \\ &\quad \frac{n^{-1/2}}{(2\pi)^{n-1}} \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega} \\ &= G_n(j\omega) U_n(j\omega), \end{aligned} \tag{4.49}$$

where

$$G_n(j\omega_n) = \frac{\int\limits_{\omega_1+\ldots+\omega_n=\omega} H_n(j\omega_1,\ldots,j\omega_n) \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}}{\int\limits_{\omega_1+\ldots+\omega_n=\omega} \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}}. \tag{4.50}$$

which is valid when

$$\int\limits_{\omega_1+\ldots+\omega_n=\omega} \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega} \neq 0. \tag{4.51}$$

Note that for the linear case ($n = 1$) Equation (4.49) is reduces to Equation (4.36) as expected, with $Y_1(j\omega) = Y(j\omega)$, $G_1(j\omega) = H(j\omega)$ and $U_1(j\omega) = U(j\omega)$.

The output frequency response of a nonlinear system given by Equation (4.40) can be now re-written as

$$Y(j\omega) = \sum_{n=1}^{N_m} Y_n(j\omega) = \sum_{n=1}^{N_m} G_n(j\omega) U_n(j\omega). \tag{4.52}$$

where the n'th order NOFRF $G_n(j\omega)$ is given by Equation (4.50). It can be seen

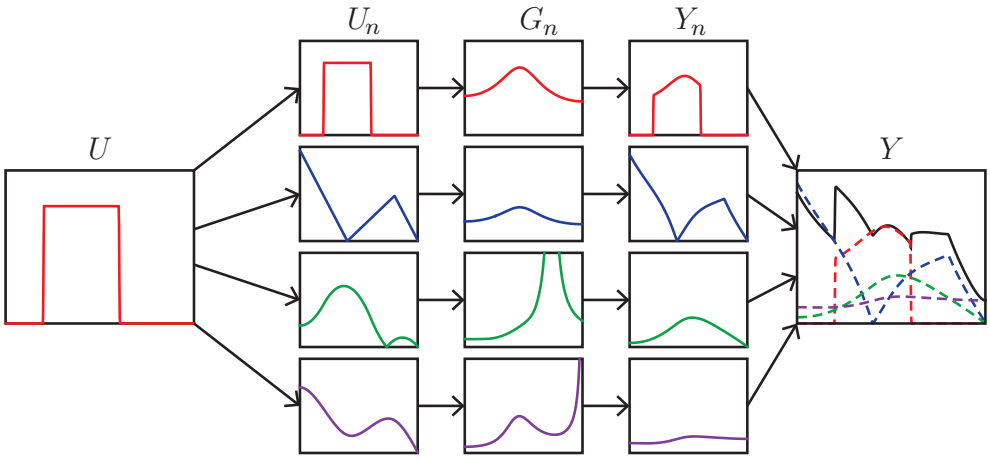

Figure 4.3: **A pictorial representation of the NOFRF analysis method.** The n'th order input spectra is calculated from the input spectra via Equation (4.48). NOFRFs are then found from the system GFRFs and the n'th input spectra using Equation (4.50). G_n acts as a filter on U_n to produce the nth order output spectra Y_n. The total output spectrum, Y is found from Equation (4.52).

from the form of (4.52) that this definition of the output frequency response for non-linear systems is a direct extension of the linear case given by Equation (4.36) and can be recovered by taking $N_m = 1$, $G_1(j\omega) = H(j\omega)$ and $U_1(j\omega) = U(j\omega)$.

Rearranging Equation (4.52) for the G_n

$$G_n(j\omega) = Y_n(j\omega)/U_n(j\omega) \tag{4.53}$$

and noting that, from the properties of non-linear systems, both Y_n and U_n must be equal to zero when ω is outside the frequency range f_{Y_n} defined by Equation (4.42). It then follows that G_n is only valid over the frequency range f_{Y_n}.

A further property of the NOFRFs is that they are insensitive to a change in the input spectrum by some constant gain which can be seen by considering

$$G_n(j\omega)\big|_{U(j\omega)=\alpha\bar{U}(j\omega)} = \frac{\alpha^n \int_{\omega_1+\ldots+\omega_n=\omega} H_n(j\omega_1,\ldots,j\omega_n)\prod_{i=1}^n \bar{U}(j\omega_i)\mathrm{d}\sigma_{n\omega}}{\alpha^n \int_{\omega_1+\ldots+\omega_n=\omega} \prod_{i=1}^n \bar{U}(j\omega_i)\mathrm{d}\sigma_{n\omega}}, \tag{4.54}$$

which implies that $G_n(j\omega)\big|_{U(j\omega)=\bar{U}(j\omega)}$ because the factor of α^n appears outside of the integral in both the numerator and the denominator. This property will be used to derive an efficient algorithm for the evaluation of NOFRFs in Section 4.3.3.

The process can be seen pictorially in Figure 4.3. From the figure it can be seen how G_n effectively acts as a filter on U_n within the frequency range for which it is defined, and can be calculated using equations (4.42), to produce the output Y_n, which may contain frequency components out of the range of the input U.

The NOFRF theory presented in this chapter provides a tool for calculating exactly how frequency generation effects occur in non-linear systems in order to transfer energy to frequencies not included in the input spectrum. The behaviour of non-linear systems is known to be input dependent and this can now be seen explicitly by noting that NOFRFs of the system are a function of the system input in Equation (4.50). They are also a function of the GFRFs $H_n(j\omega)$, which as discussed, may in general be a function of any lower order GFRF.

In order to see exactly how the non-linear behaviour of the system acts to transfer energy to new frequencies the combined effects of each order of non-linearity must be considered for a given input. However, standard input excitation signals used for system identification may not be ideal for performing NOFRF analysis. As such the system may instead be probed with a suitably designed inspection signal [67].

A summary of the NOFRF analysis procedure is given below

1. Identify the non-linear model structure and estimate model parameters.
2. Calculate GFRFs, H_n, for $n = 1, \ldots, N_m$ from the model equations using the methods described in Section 4.3
3. Calculate the n'th order input spectra using Equation (4.48) for $n = 1, \ldots, N_m$.
4. Calculate the n'th order NOFRF using Equation (4.50) for $n = 1, \ldots, N_m$.
5. Filter the n'th order input spectra with the n'th order NOFRF to find the n'th order output spectra and sum to construct the full output spectrum using Equation (4.52).

In the above procedure steps 2 and 4 can be both complex and computationally cumbersome to compute, largely because of the integration across a high dimensional space that is required in the evaluation of Equation (4.54). An efficient algorithm that is able to evaluate the NOFRFs directly without knowledge of the GFRFs is given in [67] and is presented in the following section.

4.3.3 Least squares based evaluation of NOFRFs

The output spectra of a nonlinear system is given by Equation (4.52). Decomposing both the NOFRFs and the input spectra into real and imaginary parts and dropping the notational dependence on $j\omega$ for clarity by defining

$$G_n^R = \operatorname{Re} G_n(j\omega),\ G_n^I = \operatorname{Im} G_n(j\omega),\ U_n^R = \operatorname{Re} U_n(j\omega),\ U_n^I = \operatorname{Im} U_n(j\omega)$$

the output spectra can then be written as

$$Y(j\omega) = \sum_{n=1}^{N_m}(G_n^R + jG_n^I)(U_n^R + jU_n^I) \tag{4.55}$$

$$= \sum_{n=1}^{N_m}(G_n^R U_n^R - G_n^I U_n^I) + j(G_n^R U_n^I - G_n^I U_n^R). \tag{4.56}$$

From Equation (4.56) it follows that the output spectra can be decomposed into a vector of real and imaginary components as

$$\begin{bmatrix} \mathrm{Re}\, Y(j\omega) \\ \mathrm{Im}\, Y(j\omega) \end{bmatrix} = \begin{bmatrix} U_1^R, \ldots, U_N^R, -U_1^I, \ldots, -U_N^I \\ U_1^I, \ldots, U_N^I, U_1^R, \ldots, U_N^R \end{bmatrix} \begin{bmatrix} G^R \\ G^I \end{bmatrix} \tag{4.57}$$

where Re and Im indicate the real and imaginary part of a complex number respectively.

Consider an input signal $u_k^\star$ scaled by a constant α such that

$$u_k = \alpha u_k^\star, \tag{4.58}$$

From Equation (4.48) the n'th order spectrum of the input u_k is given by

$$U_n(j\omega) = \frac{n^{-1/2}}{(2\pi)^{n-1}} \int_{\omega_1+\ldots,\omega_n=\omega} \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{\omega n} \tag{4.59}$$

$$= \alpha^n \frac{n^{-1/2}}{(2\pi)^{n-1}} \int_{\omega_1+\ldots,\omega_n=\omega} \prod_{i=1}^{n} U^\star(j\omega_i) \mathrm{d}\sigma_{\omega n} \tag{4.60}$$

$$= \alpha^n U_n^\star(j\omega). \tag{4.61}$$

In order to estimate the NOFRFs, G_n, for the input $U^\star$, Equation (4.57) is employed with a number of different scaled inputs and the problem is formulated such that it can be solved using least squares.

N_m^* different input signals, u_p are designed as

$$u_{p,k} = \alpha_p u_k^\star, \quad \text{for } p = 1, \ldots, N_m^* \tag{4.62}$$

where $N_m^* \geq N_m$ where N_m is the maximum non-linear order to be considered (Note that $N_m \geq n_p$ depending on the NARX model terms) and $\alpha_{N_m^*}, \alpha_{N_m^*-1}, \ldots, \alpha_1$ are constants satisfying the condition $\alpha_{N_m^*} > \alpha_{N_m^*-1} > \ldots, \alpha_1 > 0$. The system is simulated for each input giving rise to N_m^* system outputs with frequency response given by

$$Y^p(j\omega), \quad p = 1, \ldots, N_m^*. \tag{4.63}$$

Using the condition

$$G_n(j\omega)\Big|_{U(j\omega)=\alpha\bar{U}(j\omega)} = G_n(j\omega)\Big|_{U(j\omega)=\bar{U}(j\omega)} \tag{4.64}$$

enables us to write down the following equality

$$\mathbf{Y}^{1,\ldots,N_m^*}(j\omega) = \mathbf{X}(j\omega)\begin{bmatrix} G^{\star R} \\ G^{\star I} \end{bmatrix} \tag{4.65}$$

where

$$\mathbf{Y}^{1,\ldots,N_m^*}(j\omega) = \begin{bmatrix} \operatorname{Re} Y^1(j\omega) \\ \operatorname{Im} Y^1(j\omega) \\ \vdots \\ \operatorname{Re} Y^{N_m^*}(j\omega) \\ \operatorname{Im} Y^{N_m^*}(j\omega) \end{bmatrix}, \tag{4.66}$$

$$\mathbf{X}(j\omega) = \begin{bmatrix} \alpha_1 U_1^{\star R}, \ldots, \alpha_1^{N_m} U_{N_m}^{\star R}, -\alpha_1 U_1^{\star I}, \ldots, -\alpha_1^{N} U_{N_m}^{\star I} \\ \alpha_1 U_1^{\star I}, \ldots, \alpha_1^{N_m} U_{N_m}^{\star I}, \alpha_1 U_1^{\star R}, \ldots, \alpha_1^{N} U_{N_m}^{\star R} \\ \vdots \\ \alpha_{N_m^*} U_1^{\star R}, \ldots, \alpha_{N_m^*}^{N_m} U_{N_m}^{\star R}, -\alpha_{N_m^*} U_1^{\star I}, \ldots, -\alpha_{N_m^*}^{N_m} U_{N_m}^{\star I} \\ \alpha_{N_m^*} U_1^{\star I}, \ldots, \alpha_{N_m^*}^{N_m} U_{N_m}^{\star I}, \alpha_{N_m^*} U_1^{\star R}, \ldots, \alpha_{N_m^*} w^{N_m} U_{N_m}^{\star R}. \end{bmatrix} \tag{4.67}$$

The NOFRFs of the system in response to the input $u^\star(k)$ can now be estimated by formulating Equation (4.65)

$$\begin{aligned} \begin{bmatrix} G^{\star R} \\ G^{\star I} \end{bmatrix} &= [G_1^{\star R}(\omega), \ldots, G_{N_m}^{\star R}(\omega), G_1^{\star I}(\omega), \ldots, G_{N_m}^{\star I}(\omega)]^T \\ &= [\mathbf{X}(j\omega)^T \mathbf{X}(j\omega)]^{-1} \mathbf{X}(j\omega)^T \mathbf{Y}^{1,\ldots,N_m^*}(j\omega)], \end{aligned} \tag{4.68}$$

The Least squares estimate of $G^\star$ will hence be the one that best describes the output spectra Y for all of the scaled inputs u_p for $p = 1 : N_m^*$. See Algorithm 4.1 for a summary.

The NOFRF calculation procedure is made vastly simpler by performing Algorithm 4.1 and is no longer prohibitively computationally cumbersome at higher orders. It should be noted that the algorithm is not just applicable to models of the NARX class and can be used with any model for which input output data

Algorithm 4.1 An algorithm for determining NOFRFs of a generative system.

Initialise
Test input signal - u_k^*
Scaling constants - $\boldsymbol{\alpha}$
Number of simulations - N_m^*
} (see Equation (4.62))
Procedure
Transform the time domain input into the frequency domain.
for $n = 1, \ldots, N_m^*$
Calculate U_n via Equation (4.61).
end for

Simulate the system for N_m^* proportional inputs.
for $p = 1, \ldots, N_m^*$
Design $u_p(k)$ via Equation (4.62).
Simulate the system to produce $y_p(k)$.
Find $Y^p(j\omega)$ from $y_p(k)$ via Equation (3.3).
end for

Perform the LS step
Construct matrices $\boldsymbol{Y}^{1,\ldots,N_m^*}(j\omega)$ and $\boldsymbol{x}(j\omega)$ via equations (4.66) and (4.67).
Calculate G^* via Equation (4.68).
end Procedure

can be generated. It may also be used directly with experimental data if suitable input-output data can be gathered.

Example 4.3 Calculation of NOFRFs

In this example the analysis performed in example 4.2 will be extended to the output frequencies of the system given by Equation (4.22) using the NOFRF method by application of algorithm 4.1. The highest order term in Equation (4.22) is a pure output non-linearity of order 2. As discussed in Section 4.3 any pure output or cross-term in the time domain model equation gives rise to higher order contributions to the frequency response in the form of higher order GFRFs and hence give a contribution to the output spectrum. It is therefore necessary to consider these higher order terms as part of the analysis.

In example 4.2 the system is simulated in response to a white noise sequence which led to the system being approximately represented by just the first order frequency response. Consider instead the same system given by Equation (4.22) in response to a band limited white noise signal, which can be found by, for example, low pass filtering the original white input, see Figure 4.4. Considering the output spectrum of the system excited by the two different signals the output spectra for

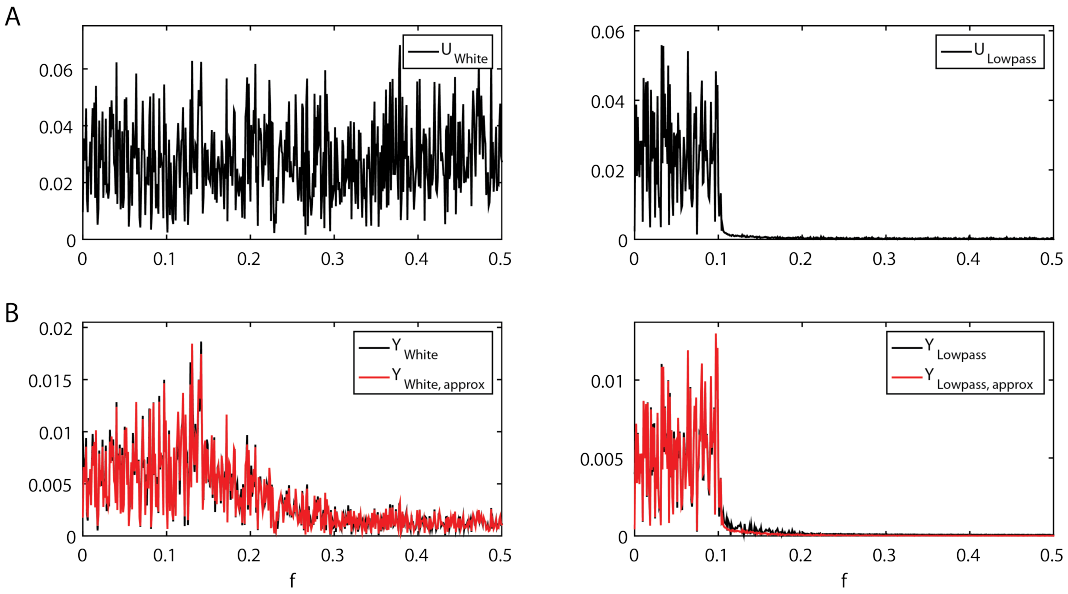

Figure 4.4: **Non-linear systems are input dependent. GFRFs are invariant to the input and it is necessary to calculate the NOFRFs of the system in order to fully understand the non-linear mechanisms that produce the output. They also transfer energy from the frequency range of the input to frequencies outside that range. A**) The spectrum of: Left, a white noise input. Right, a band limited white noise input. **B**) The spectrum of: Left, the output of the generative system given by Equation (4.22) in response to a white noise input. Right, the output of the same system in response to a band limited input.

both cases has been approximated by taking the convolution of the input spectra with the first order GFRF/NOFRF. It is not obvious from the white noise case that the approximation is a poor one. In the band limited case the approximation is good in the range $f = [0, 0.1]$ but there is clearly un-modelled behaviour resulting in the output in the region $f = [0.1, 0.2]$.

From linear systems theory it is known that the frequency range of the input is the same as that for the output of a linear system. From this we know that the linear contribution to the output spectrum is going to be in the range $[0, 0.1]$ to match the spectrum of the input. The range of available output frequencies for each non-linear order can be calculated from equation set (4.42). The result is simply $fy_n = [0, n * 0.1]$, i.e. multiples of the input frequency range.

In order to use Algorithm 4.1 an appropriate input signal must be designed. The input signal is designed to be a frequency band that is constant in the range

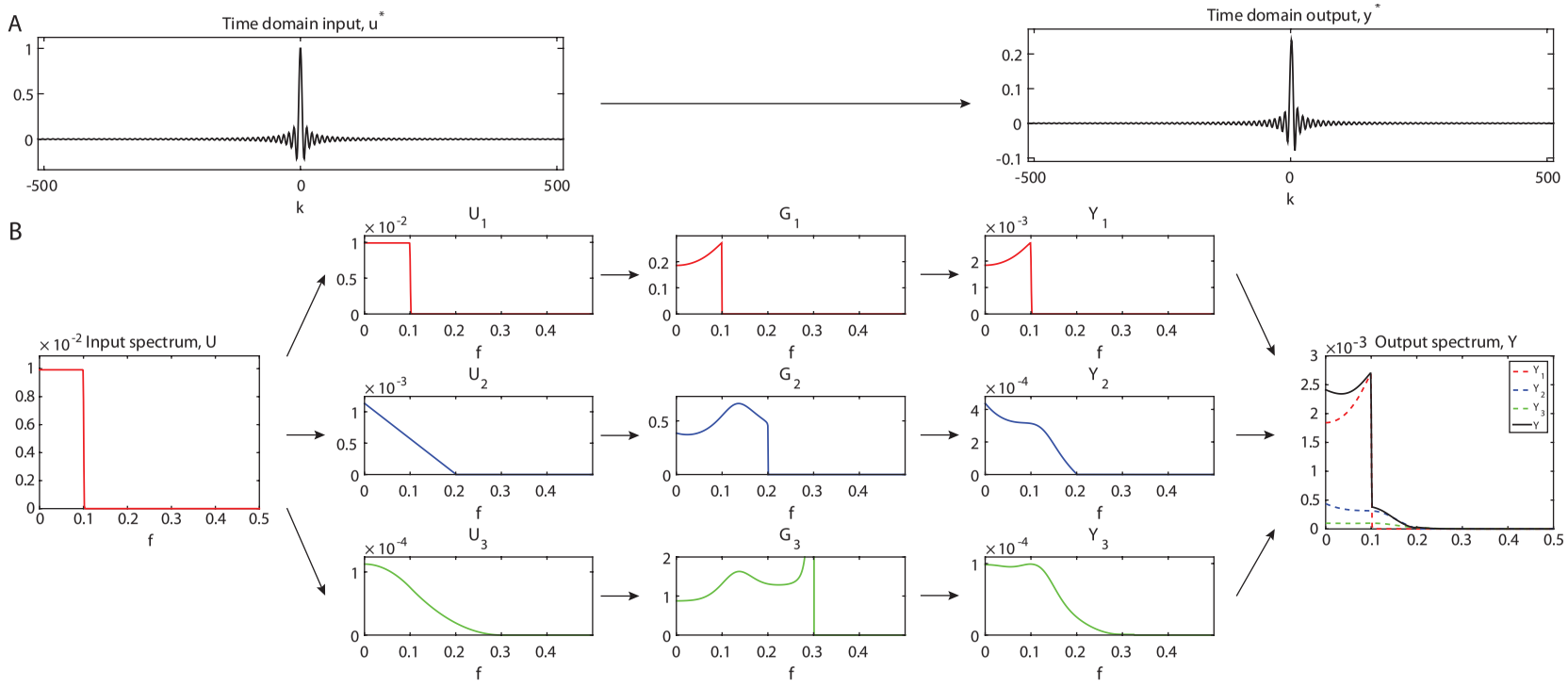

Figure 4.5: **The NOFRF concept allows the investigation of how each non-linear order of a system contributes to the output. The transformation from the input to the output can be described by a set of functions: the NOFRFs. A**) The input and output in the time domain. **B**) The input spectrum is decomposed into N_m n'th order input spectra. The NOFRFs act as a filter on the n'th order input spectra to produce the n'th order output spectra. These can be summed to form the output spectrum.

$[a, b]$ and 0 elsewhere. A signal of this form can be constructed as

$$\begin{aligned} u_k^* &= \frac{3}{(2\pi)} \frac{\sin(2\pi bk\Delta t) - \sin(2\pi ak\Delta t)}{k\Delta t} \\ k &= -511, \ldots, 512 \end{aligned} \tag{4.69}$$

as recommended in [67], where Δt is the sampling time. The spectra of the n'th order inputs U_n can then be found by Equation (4.48), which can be performed efficiently by computing

$$U_n(j\omega) = \frac{n^{-1/2}}{(2\pi)^{n-1}} FT(\mathbf{u^n}) \tag{4.70}$$

where $FT(\mathbf{u^n})$ is the Fourier Transform of the input to the n'th power. The time domain input and output signals, u_k and y_k, as well as the n'th order input spectra can be seen in Figure 4.5. It can be seen how the n'th order input spectra are limited to the frequency range defined by (4.42).

The algorithm estimates both the NOFRFs and the n'th order output spectra for the system, both of which are shown in Figure 4.5. The figure serves to demonstrate how the NOFRFs $G_1, \ldots, G_{N_m}$ act as filters on the n'th order input spectra in order to produce the n'th order output spectra that themselves sum to form the output spectrum. The NOFRFs are only shown over the frequency range for which they are defined. △

The example highlights how the concept of NOFRFs directly extends traditional frequency domain analysis using the FRF to the realm of non-linear functions. It also shows how the relationship between the input and output of a nonlinear system is significantly more complicated than in the linear case.

4.4 Results

The remainder of this chapter is concerned with the application of the methods introduced so far to the modelling and analysis of DEAs. The methods are first collected into a framework and then result are given that provide new insight into DEA behaviour.

4.4.1 A framework for the identification and analysis of DEAs.

The methods detailed in the previous sections of this chapter constitute the components of the data driven and control-oriented analysis framework and are sum-

marised below.

1 Identification of the non-linear model structure using the SEMP algorithm.

2 Removal of any DC component present in the model.

3 Calculation of system GFRFs.

4 Calculation of system NOFRFs.

In the next section this framework is demonstrated with application to six film-type DEAs. A preliminary analysis of these DEAs was performed in Chapter 2 and indicates the necessity for non-linear system identification techniques in order to achieve accurate models of the system. In the remainder of this chapter the non-linear system identification and frequency domain analysis techniques introduced so far in the chapter will be applied to the set of DEAs.

All of the DEA actuators considered in this study display unexplained TV dynamics. A characterisation of these time variations will be performed in Chapter 5, in this chapter identification and analysis will be performed on an approximately time invariant section of data. For this purpose 160 seconds (8000 data points prior to sub-sampling) of input-output data are chosen for each actuator and are considered to be non time-varying, see Appendix B.1 for the data segments chosen. All of the identification, validation and analysis in this chapter is performed over these data sets.

4.4.2 Data preprocessing

It is well known that the sampling time of the data under investigation affects both the dynamic behaviour of the final model estimate but also the performance of structure detection and parameter estimation algorithms [14]. If the system is over sampled then the data will be highly correlated such that $y_k \approx y_{k-1}$ leading to numerical problems. If it is under sampled then some of the dynamic behaviour of the system will be lost. Therefore it may be necessary to down-sample the data record in order to achieve satisfactory results. The sampling time can be chosen using a number of procedures [103]. A simple and common approach is to select the sampling time based on the correlation function [3]. This method has been expanded to also consider non-linear correlations [15].

Considering both the linear and non-linear [19] correlation functions

$$\Phi_{yy}(\tau_c) = \boldsymbol{E}[(y_k - \overline{y_k})(y_{k-\tau_c} - \overline{y_k})], \quad \tau_c = 0, 1, \ldots \tag{4.71}$$

$$\Phi_{y^{2'}y^{2'}}(\tau_c) = \boldsymbol{E}[(y_k^2 - \overline{y_k^2})(y_{k-\tau_c}^2 - \overline{y_k^2})], \quad \tau_c = 0, 1, \ldots. \tag{4.72}$$

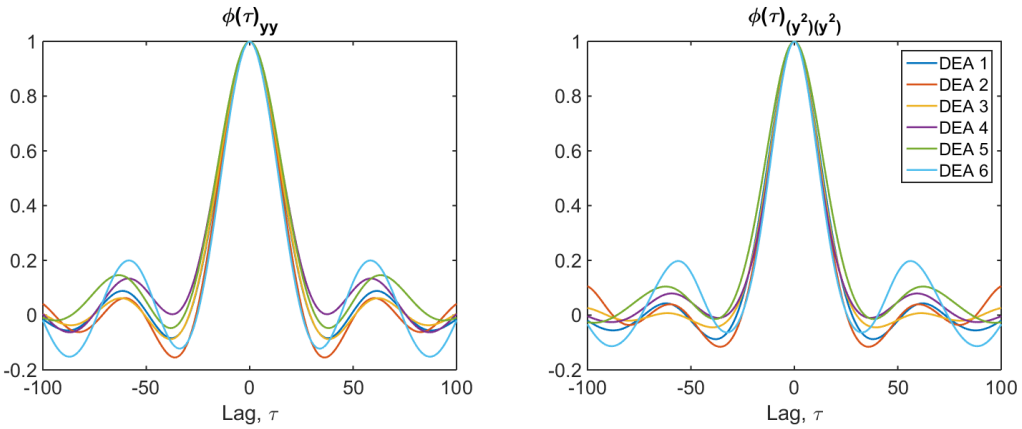

Figure 4.6: **An appropriate sampling time for performing system identification can be chosen using linear and non-linear correlation functions.** The autocorrelation functions Φ_{yy} and $\Phi_{y^{2'}y^{2'}}$ for all six DEAs. The appropriate sampling time can be chosen based on the first minima of the two functions

The sampling time can then be chosen in the range [3, 15]

$$\frac{\tau_m}{20} \leq T_s \leq \frac{\tau_m}{5} \tag{4.73}$$

where

$$\tau_m = \min(\tau_y, \tau_{y^{2'}}) \tag{4.74}$$

where τ_y and $\tau_{y^{2'}}$ are the first minimums of Φ_y and $\Phi_{y^{2'}y^{2'}}$ respectively.

The above method is applied to the training data for all six DEAs, see Figure 4.6. The first minima over all of the actuators is found as $\tau_m = 36$. From the rule given by Equation (4.73) the sampling time can thus be chosen in the range $[\tau_c/20,\ \tau_c/5] = [1.8, 7.2]$. The experimental sampling time is $T_s = 1/50$, the system can therefore be down sampled into the range $[0.036, 0.144]$. The sampling time $T_s = 0.08$ is chosen corresponding to $1/4$ of the original sampling rate.

4.4.3 Identification of the non-linear structure

In this section the simulation based structure detection and parameter estimation algorithm, the SEMP algorithm, introduced in Section 4.1 is applied to the six DEAs. The modelling procedure is first demonstrated on DEA 5 and then results are provided for all six actuators in order to provide a comparison over the data sets. DEA 5 is chosen as an example based on the modelling performance, as measured by the MSSE, being close to the average over all the actuators. Identification is performed over 2000 samples over the down sampled data equally divided into estimation and validation data sets of 1000 data samples each.

Selection of model orders.

In order to perform identification with the SEMP algorithm it is first necessary to generate a super-set of model terms in which to perform the search. This requires choosing the maximum dynamic order n_u and n_y as well as the maximum polynomial order n_p. A standard method is based on the identification of linear models of the system [130] and the approach used here is based on this method, however in this work the criterion for variable selection is based on the simulation error rather than a MSPE based error measure.

An appropriate dynamic order is determined by performing an exhaustive search for linear ARX models *i.e.* setting the polynomial order $n_p = 1$) over all the terms generated by an increasing dynamic order $n_y = n_u = 1, \ldots, 5$. The dynamic order $n_y = n_u$ is then chosen as the one that, on increasing it further, achieves a negligible decrease in the MSPE. The polynomial order is next estimated by repeating the process with the chosen dynamic order and increasing the polynomial order $n_p = 1, \ldots, 4$. For the polynomial order an exhaustive approach is infeasible due to the size of the search space. The SEMP algorithm is therefore applied at each polynomial order and selected based on the same criterion as before.

The procedure is applied to the DEA system for all six actuators. First the dynamic order is selected. Increasing the dynamic order from first to third significantly increases the model fit, averaged over the six actuators. After third order dynamics the model fit flattens off, see Figure 4.7**A**. This behaviour is expected when no further information can be obtained by including regressors with a higher dynamic order. Based on this the dynamic order is set as $n_u = n_y = 3$. Similarly the polynomial order is selected by identifying models with increasing polynomial order. An improvement in the MSSE is again seen up to $n_p = 3$ after which the MSSE increases greatly, it is suggested that fourth order basis functions introduce incorrect predictions over a long range. As such the dynamic order is chosen as $n_p = 3$, see Figure 4.7**B**.

Determining the non-linear model structure

Once the maximum dynamic and polynomial orders have been estimated the SEMP algorithm is applied to the data in order to identify the non-linear model structure. The SEMP algorithm requires an initial model structure $\mathcal{M}_0$. For the identification of the DEAs the initial model structure is chosen as

$$\mathcal{M}_0 = [y_{k-1}, y_{k-2}, y_{k-3}, u_{k-1}, u_{k-2}, u_{k-3}], \tag{4.75}$$

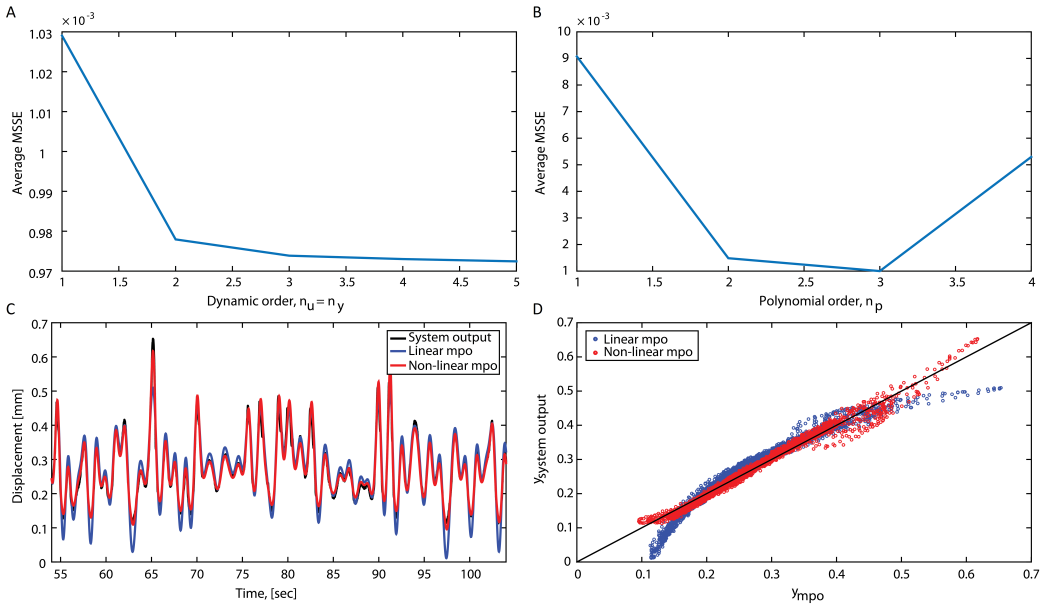

Figure 4.7: **Maximum dynamic and polynomial orders must be chosen before the identification algorithm can be performed. Non-linear models perform significantly better in prediction in comparison to linear models further indicating the necessity of non-linear modelling techniques.** **A**) The Maximum dynamic order is chosen based on the MSSE of identified ARX models with increasing order $n_u = n_y$ measured over the MPO. **B**) The Maximum polynomial order is chosen based on the MSSE of identified NARX models with increasing order n_p measured over the MPO. **C-D**) The MPO of the best Non-linear model for DEA 5 with $n_u = n_y = 3$ and $n_p = 3$ displays a significantly better fit than the best linear model with $n_u = n_y = 3$ and $n_p = 1$. Simulation is performed over a section of the validation data.

where the parameter vector $\boldsymbol{\theta}$ is estimated using LS. This initial model structure is chosen based on the knowledge that the behaviour of dielectric elastomer material can be well described by a linear model at low voltages. Note that the algorithm has the flexibility to remove these terms at any iteration.

The SEMP algorithm is used to find the model structure for DEA 5 with the variables selected in the previous section, namely $n_u = n_y = 3$, $n_p = 3$ and choosing an SERR threshold of 1×10^{-5}. The SERR falls below the threshold at iteration 12 of the algorithm at which point the MSSE flattens off, see Figure 4.8. This leads to a final model structure with 11 terms shown in Table 4.1.

Correlation analysis is performed on both the identified NARX model of DEA 5 using the correlation tests given by equations (3.56). The analysis is also performed on an ARX model containing all possible terms generated by a maximum dynamic order $np = 3$. The correlation function $\Phi_{(u^2)e}$ shows significant correlations in

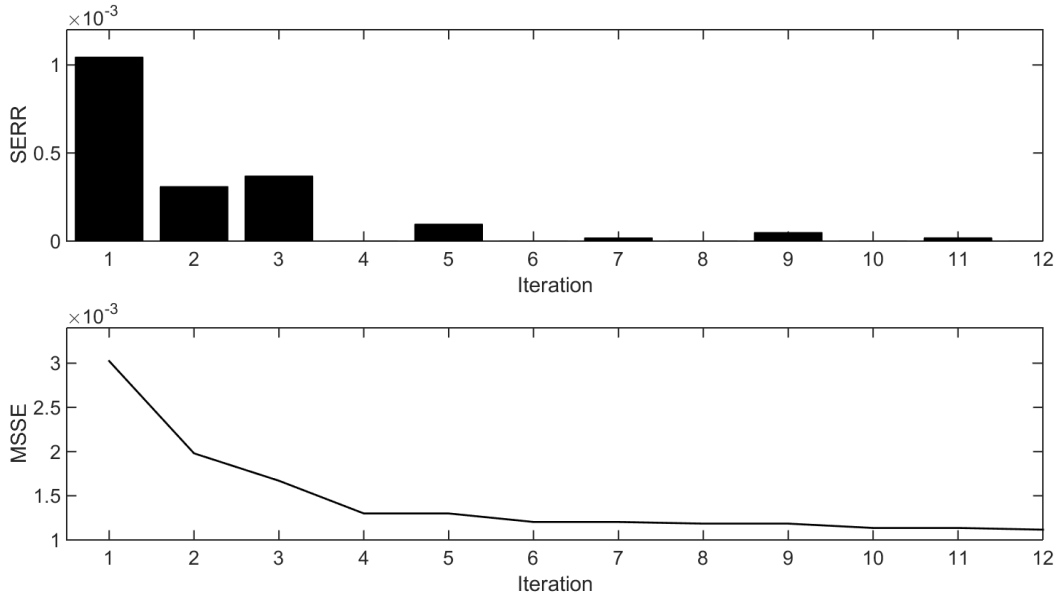

Figure 4.8: **In the forward term selection stage of the SEMP algorithm the final model is selected when the addition of a new term causes the reduction in the SERR to fall below some threshold.** SERR, calculated as part of the SEMP algorithm for the identification of DEA 5. The SERR falls below the threshold of 1×10^{-5} at the 12th algorithm iteration, leading to a model with 12 terms prior to pruning.

the linear model for $\tau < 6$ as well as large auto-correlations in residuals and in $\Phi_{e(ue)}$. The non-linear model shows significant improvement in $\Phi_{(u^2)e}$ in which the correlations are approximately within the 95% confidence intervals. Auto-correlations in the residuals are still present outside of the confidence intervals as well as in $\Phi_{\tilde{\xi}(ue)}$. The correlation function $\Phi_{(u^2)e}$ indicates significant missing second order terms. However, including further terms is not found to improve the results, nor including terms with a higher dynamic order. This is a feature across all six of the actuators see Appendix B.2-B.2.

The correlations in Φ_e and $\Phi_{e(ue)}$ indicate the presence of coloured noise which leads to the estimation of biased parameters. A noise model is fitted by selecting terms based on the correlation functions, see Table 4.2. Unbiased parameters are estimated using ELS via Algorithm (3.1). The correlation analysis on the identified NARMAX model shows that the correlations in Φ_e and $\Phi_{e(ue)}$ have been brought approximately within the 95% confidence intervals except for in $\Phi_{e(ue)}$ with $\tau = 0$. This is another feature of the DEAs. The correlation could suggest the inclusion of a feed-forward term (the inclusion of u_k into the model), however this makes no improvement in the correlation tests.

Repeating this process for the remaining five DEAs leads to a compact model

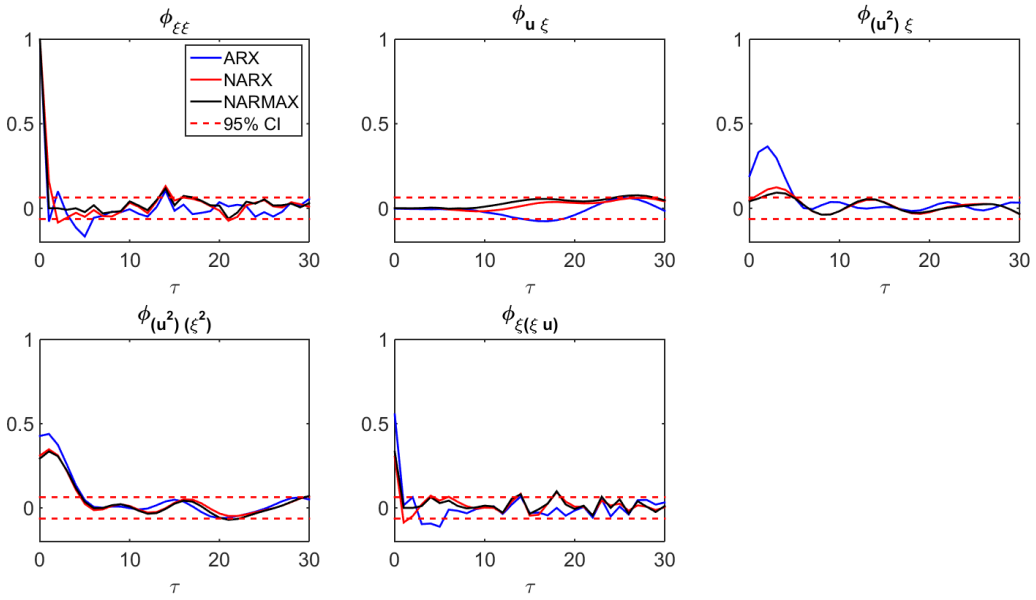

Figure 4.9: **Correlation tests can be used to validate that a sufficient model structure has been identified.** Correlation tests for linear (Blue), non-linear (Red) and non-linear with noise model (black) models of DEA 5.

description for each of the actuators, see Table 4.1 with associated noise models, see Table 4.2. The NARMAX models identified in this manner accurately capture the dynamics of the system, demonstrated by simulation comparison with the independent validation data, see Figure 4.7**B-C**. Correlation tests for the set of identified models are given in Appendix B. The identified NARX models show little consistency in selected model terms except in the selection of linear terms, see Table 4.1. The relatively poor fit of DEAs 1 and 3 and 6 might be explained by the presence of time varying dynamics within the data record over which they have been identified and validated. It is also interesting that the identified models contain a different amount of model terms. In particular, DEA 1 contains 15 terms in comparison to just 7 for DEA 6. In the previous chapter it was noted that DEA 1 displays behaviour that is inconsistent with the other DEAs which may explain this.

The identified NARX models of the DEAs all include a DC component which was explicitly included into the modelling procedure. In order to transform the models into a form that is appropriate for the frequency domain analysis tools introduced in this chapter the DC component must be removed from the model. This is done using the procedure outlined in Section 4.2. The resulting models now include a larger set of basis functions that are generated by the transformation, see Table 4.3. Interestingly the new set of models show more similarity than before

Terms	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6
DC term	0.113394	0.202312	0.179079	0.061684	0.047522	0.239233
y_{k-1}	-	-	0.541589	0.877008	0.895000	-
y_{k-2}	-	0.182184	-	-	-	-
y_{k-3}	0.732689	0.071688	-	-	0.041525	0.174577
u_{k-1}	0.207556	0.171188	0.132531	0.603688	0.411452	0.584029
u_{k-2}	-	-0.198380	-0.593643	-0.673238	-0.421602	-0.230338
u_{k-3}	-	-	0.197257	0.127794	0.067124	-
$y_{k-1}y_{k-2}$	-	-	-	-	-0.076994	-
$y_{k-1}u_{k-1}$	-	0.893292	1.148154	-	-	-
$y_{k-2}u_{k-2}$	1.397738	-	-	-	-	-
$y_{k-2}u_{k-3}$	-1.653767	-	-	-	-	-
u_{k-1}^2	1.496631	-	-	-	-	0.406254
$u_{k-1}u_{k-2}$	-	-0.058285	-	-	-	-
u_{k-2}^2	-0.821075	-	-0.167632	-	-	-
$u_{k-2}u_{k-3}$	-	-	0.052803	-	-	-
$y_{k-1}^2u_{k-1}$	-	-	-0.073564	-	-	0.039074
$y_{k-1}y_{k-3}u_{k-1}$	-	-	-	0.146836	-	-
$y_{k-1}u_{k-1}^2$	-	-	-	0.979018	0.656646	-
$y_{k-1}u_{k-1}u_{k-3}$	-	-	-0.679573	-0.229460	-	-
$y_{k-1}u_{k-2}^2$	-	-	-	-1.257500	-0.738311	-
$y_{k-2}^2y_{k-3}$	-	-	-	-0.034998	-	-
$y_{k-2}y_{k-3}u_{k-3}$	-0.526294	-	-	-	-	-
$y_{k-2}u_{k-1}^2$	-	0.941410	0.263654	-	-	-
$y_{k-2}u_{k-1}u_{k-2}$	-	-1.156016	-	-	-	-
$y_{k-2}u_{k-1}u_{k-3}$	-	-	-	-	0.278564	-
$y_{k-2}u_{k-2}^2$	-	-	-	0.689155	-	-
$y_{k-2}u_{k-2}u_{k-3}$	2.128494	-	-	-	-	-
$y_{k-3}^2u_{k-1}$	0.868314	-	-	-	-	-
$y_{k-3}u_{k-1}^2$	-1.370184	-	-	-0.195967	-0.036604	-
$y_{k-3}u_{k-1}u_{k-3}$	-1.703940	-	-	-	-	-
$u_{k-1}^2u_{k-2}$	-	-	-	-	-	0.076650
$u_{k-1}^2u_{k-3}$	1.969487	-	-	-	-	-
u_{k-2}^3	-1.409335	-	-	-	-	-
MSSE ($\times 10^{-3}$)	14.98	0.30	3.69	0.66	0.73	1.53
Term count	14	9	11	12	11	7

Table 4.1: **Using the NARX model structure allows for a parsimonious system description and good prediction accuracy.** NARX models for DEA 1-6 identified using the SEMP algorithm along with their corresponding parameter values. The model fit is assessed by the MSSE.

Terms	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6
e_{k-1}	0.7566	0.5941	0.3207	0.2326	0.3296	0.7429
e_{k-2}	0.2013	0.3253	0.2499	0.2326	0.0416	0.3380
e_{k-3}	-0.1020	-	0.0660	-	-	-0.0790
e_{k-4}	0.0486	-	-	-	0.0203	-0.0719
e_{k-5}	-	-	-	-	-	0.0517
e_{k-6}	-0.0581	-	-	-	-	-
e_{k-7}	-	-	0.0931	-	-	-0.0088
$u_{k-1}e_{k-1}$	0.4299	-	-	-	-0.3776	-
$u_{k-2}e_{k-2}$	-	-0.3703	-0.3701	-0.2936	-0.1689	-
$u_{k-3}e_{k-3}$	0.1014	-0.2879	-	-	-	-
$u_{k-4}e_{k-4}$	0.0711	-0.1258	-	-	-	-
MSSE $(\times 10^{-3})$	4.30	0.28	3.27	0.52	0.72	1.45

Table 4.2: **Identified NARX models for the DEA actuators result in biased residuals requiring the identification of a noise model such that the estimated parameters are unbiased.**

the DC term was removed with non-linear term u_{k-1}^2 being consistent across all of the different model structures.

4.4.4 Non-linear frequency domain analysis of DEAs

GFRF analysis

In order to provide further insight into the dynamic behaviour of the DEA system the GFRF analysis framework is applied. Using the probing method discussed in Section 4.3 the GFRFs can be found directly from the identified NARX model after the noise model has been discarded.

Continuing with DEA 5 as an example and starting from the DC removed NARX model equation from Table 4.3

$$\begin{aligned} y_k = \quad & \theta_1 y_{k-1} + \theta_2 y_{k-2} + \theta_3 y_{k-3} + \theta_4 u_{k-1} + \theta_5 u_{k-2} + \theta_6 u_{k-3} + \theta_7 y_{k-1} y_{k-1} \\ & + \theta_8 u_{k-1}^2 + \theta_9 u_{k-1} u_{k-3} + \theta_{10} u_{k-2}^2 + \theta_{11} y_{k-1} u_{k-1}^2 + \theta_{12} y_{k-1} u_{k-2}^2 \\ & + \theta_{13} y_{k-2} u_{k-1} u_{k-3} + \theta_{14} y_{k-3} u_{k-1}^2 \end{aligned} \tag{4.76}$$

the system GFRFs can be found using an efficient algorithm based on the probing method [56]. The following asymmetric GFRFs are produced for $n = 1, \ldots, 3$

$$H_1(\omega_1) = -\frac{\theta_4 e^{-j\omega_1 \omega} + \theta_5 e^{-j\omega_1 2\omega_1} + \theta_6 e^{-j\omega_1 3\omega_1}}{\theta_1 e^{-j\omega_1 1\omega_1} + \theta_2 e^{-j\omega_1 2\omega_1} + \theta_3 e^{-j\omega_1 3\omega_1} - 1} \tag{4.77}$$

Terms	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6
y_{k-1}	-	-	0.541589	0.877008	0.858429	-
y_{k-2}	-	0.182184	-	-0.015572	-0.036572	-
y_{k-3}	0.732689	0.071688	-	-0.007786	0.041525	0.174577
u_{k-1}	0.363808	0.413404	0.569834	0.636355	0.411452	0.587311
u_{k-2}	0.592926	-0.198380	-0.593643	-0.673238	-0.421602	-0.230338
u_{k-3}	-0.796240	-	0.197257	0.127794	0.067124	-
$y_{k-1}y_{k-2}$	-	-	-	-	-0.076994	-
$y_{k-1}u_{k-1}$	-	0.893292	1.090678	0.069258	-	0.022650
y_{k-2}^2	-	-	-	-0.016508	-	-
$y_{k-2}y_{k-3}$	-	-	-	-0.033015	-	-
$y_{k-2}u_{k-2}$	1.397738	-	-	-	-	-
$y_{k-2}u_{k-3}$	-1.877023	-	-	-	-	-
y_{k-3}^2	-	-	-	-0.016508	-	-
$y_{k-3}u_{k-1}$	0.736684	-	-	0.069258	-	-
$y_{k-3}u_{k-3}$	-0.223256	-	-	-	-	-
u_{k-1}^2	0.915394	0.255263	0.102997	0.369340	0.294517	0.406254
$u_{k-1}u_{k-2}$	-	-0.371738	-	-	-	-
$u_{k-1}u_{k-3}$	-0.722818	-	-0.265477	-0.108229	0.132316	-
u_{k-2}^2	-0.821075	-	-0.167632	-0.268070	-0.350694	-
$u_{k-2}u_{k-3}$	0.902915	-	0.052803	-	-	-
$y_{k-1}y_{k-1}u_{k-1}$	-	-	-0.073564	-	-	0.039074
$y_{k-1}y_{k-3}u_{k-1}$	-	-	-	0.146836	-	-
$y_{k-1}u_{k-1}^2$	-	-	-	0.979018	0.656646	-
$y_{k-1}u_{k-1}u_{k-3}$	-	-	-0.679573	-0.229460	-	-
$y_{k-1}u_{k-2}^2$	-	-	-	-1.257500	-0.738311	-
$y_{k-2}^2y_{k-3}$	-	-	-	-0.034998	-	-
$y_{k-2}u_{k-3}^2$	-0.526294	-	-	-	-	-
$y_{k-2}u_{k-1}^2$	-	0.941410	0.263654	-	-	-
$y_{k-2}u_{k-1}u_{k-2}$	-	-1.156016	-	-	-	-
$y_{k-2}u_{k-1}u_{k-3}$	-	-	-	-	0.278564	-
$y_{k-2}u_{k-2}^2$	-	-	-	0.689155	-	-
$y_{k-2}u_{k-2}u_{k-3}$	2.128494	-	-	-	-	-
$y_{k-3}^2u_{k-1}$	0.868314	-	-	-	-	-
$y_{k-3}u_{k-1}^2$	-1.370184	-	-	-0.195967	-0.036604	-
$y_{k-3}u_{k-1}u_{k-3}$	-1.703940	-	-	-	-	-
$u_{k-1}^2u_{k-2}$	-	-	-	-	-	0.076650
$u_{k-1}^2u_{k-3}$	1.969487	-	-	-	-	-
u_{k-2}^3	-1.409335	-	-	-	-	-
Term count	19	9	12	21	14	7

Table 4.3: **Removing the DC component of a NARX model gives rise to a modified parameter vector as well as the generation of new model terms.** NARX models for 6 DEAs reported in Table 4.1 after the DC component has been removed. New model terms are coloured red, terms with an altered parameter are coloured blue.

$$H_2(\omega_1,\omega_2) = -\frac{\left(\theta_7 H_1(\omega_1)H_1(\omega_2)\frac{1}{2!}e^{-j(\omega_1+2\omega_2)} + \theta_8\frac{1}{2!}e^{-j(\omega_1+\omega_2)} + \theta_9\frac{1}{2!}e^{-j(1\omega_1+3\omega_2)} + \theta_{10}\frac{1}{2!}e^{-j(2\omega_1+2\omega_2)}\right)}{\theta_1 e^{-j(\omega_1+\omega_2)} + \theta_2 e^{-j(2\omega_1+2\omega_2)} + \theta_3 e^{-j(3\omega_1+3\omega_2)} - 1} \tag{4.78}$$

$$H_3(\omega_1,\omega_2,\omega_3) = -\frac{\left(\begin{array}{c}\frac{2}{3!}\theta_7 H_1(\omega_1)H_2(\omega_2,\omega_3)e^{-j(1\omega_1+2\omega_2+2\omega_3)} \\ +\frac{1}{3!}\theta_{11}H_1(\omega_1)e^{-j(1\omega_1+\omega_2+\omega_3)} + \frac{1}{3!}\theta_{12}H_1(\omega_1)e^{-j(1\omega_1+2\omega_2+2\omega_3)} \\ +\frac{1}{3!}\theta_{13}H_1(\omega_1)e^{-j(2\omega_1+1\omega_2+3\omega_3)} + \frac{1}{3!}\theta_1 4H_1(\omega_1)e^{-j(2\omega_3+1\omega_2+1\omega_3)}\end{array}\right)}{\theta_1 e^{-j(\omega_1+\omega_2+\omega_3)} + \theta_2 e^{-j(2\omega_1+2\omega_2+2\omega_3)} + \theta_3 e^{-j(3\omega_1+3\omega_2+3\omega_3)} - 1}. \tag{4.79}$$

The symmetric GFRFs can be found from the asymmetric GFRFs via Equation (4.15). Note that because of the presence of cross terms in Equation (4.76) there will be infinitely many higher order GFRFs, however, the higher order GFRFs give a negligible contribution to the system output.

The GFRFs are evaluated for the first and second order allowing comparisons between their frequency responses. Considering the whole spectrum, the first order GFRFs display varying frequency responses, however, because the system is only excited at low frequencies (0-1 Hz) it is known from the discussion of the output response of non-linear systems in Section 4.3.1 that output frequencies originating from the first order terms can only be in the range of the input spectrum. The focus of the analysis should hence be on the low frequency region, < 3 Hz. In this region all six actuators exhibit a similar response although the magnitude of the gain varies across the different actuators, see Figure 4.10**A**. Similarly, the second order GFRFs exhibit a range of behaviour at higher frequencies, at low frequencies < 2 they all exhibit a ridge at $\omega_1 + \omega_2 = 0$.

4.4.5 NOFRF analysis

GFRFs can provide a unique insight into the behaviour of non-linear systems, however they are not able to provide a complete description of the frequency response. In order to provide a fuller description, NOFRFs of the DEA system are calculated using the LS based algorithm given in Section 4.3.3.

Firstly, the frequency range of the output spectra can be calculated using equation set (4.42). Given the experimental band limited noise input spectrum in the region $[0, 1]$Hz then the n'th order output spectra can cover the range $[0, n]$Hz, although it is expected that the output of the system will be negligible from the higher order non-linearities.

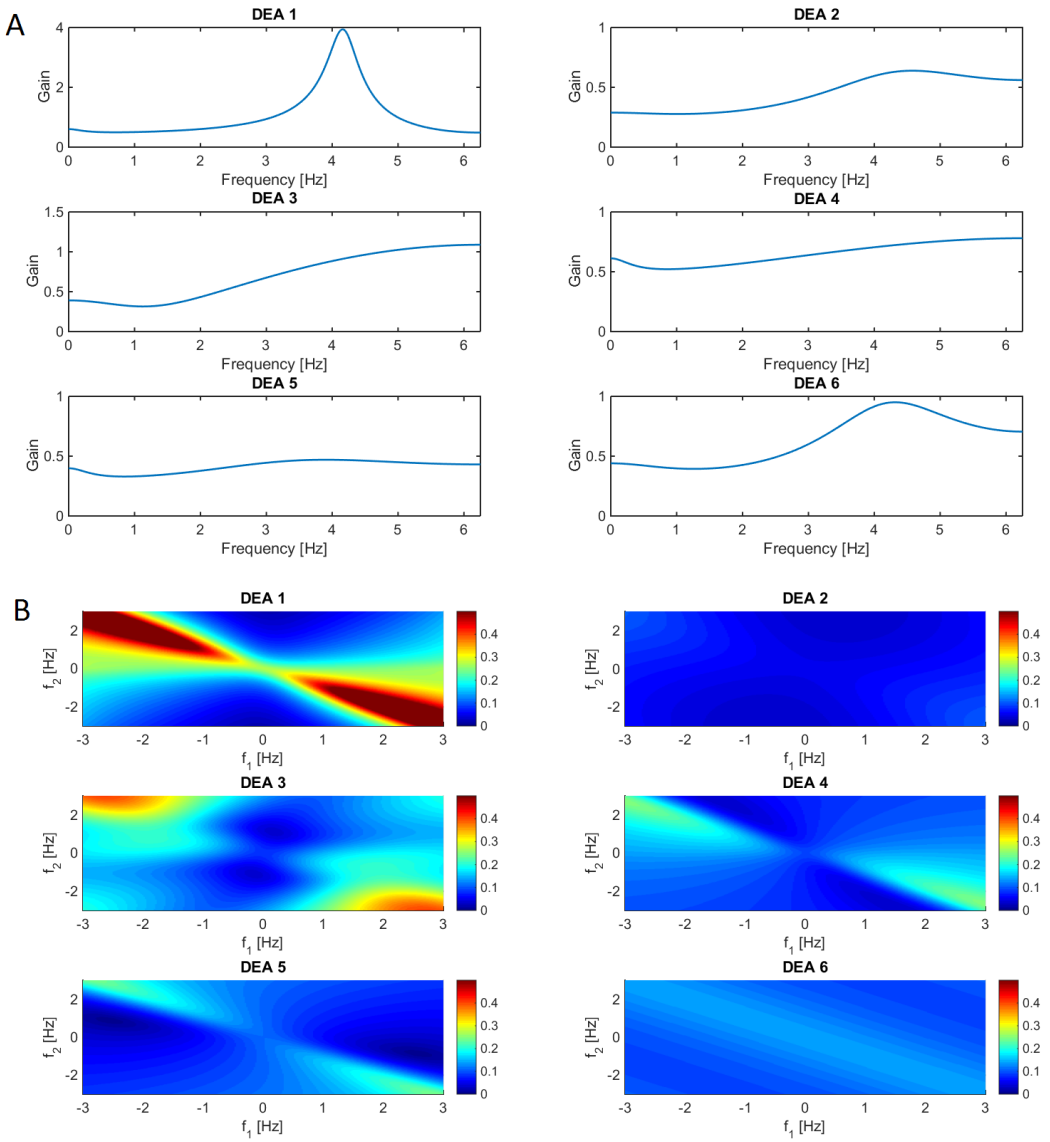

Figure 4.10: **GFRFs provide insight into the frequency domain characteristics of a non-linear system.** **A**) First order GFRFs for the set of DEAs, equivalent to the linear FRF. The first order GFRFs all exhibit similar characteristics at low frequencies, however the high frequency behaviour varies significantly. **B**) Second order GFRFs for the set of DEAs. Although there are broad differences between the second order GFRFs they all exhibit a ridge at $\omega_1 + \omega_2 = 0$ for small ω.

The experimental input is inappropriate for estimating the NOFRFs using Algorithm 4.1 and as such a test input, u_k^*, with similar characteristics to the true input is designed. The test input can be constructed using Equation (4.69) where the frequency band is in the region $[a, b] = [0, 1]$, $\Delta t = T_s = 1/50$ such that

$$\begin{aligned} u_k^* &= \frac{3}{(2\pi)} \frac{\sin(2\pi 0k) - \sin(2\pi 1k)}{k} \\ k &= -511 \times 0.02, \ldots, 512 \times 0.02 \end{aligned} \tag{4.80}$$

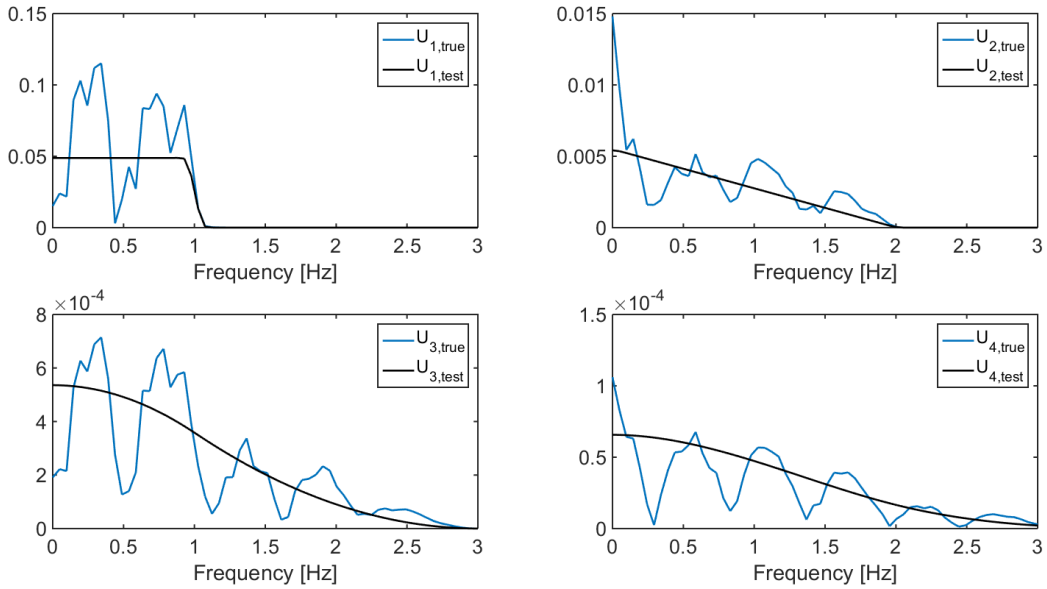

Figure 4.11: **Random Inputs that are appropriate for system identification are not appropriate for the estimation of NOFRFs, A test signal is used instead to excite the non-linear model.** The n'th order input spectra found from the time domain signal generated by Equation (4.80) (Black) match the input spectra of the true input signal (Blue).

The input spectrum produced is flat in the region [0,1] Hz and zero for all higher frequencies imitating the spectrum of the true input in a noise free case. The n'th order input spectra are found by Equation (4.70). The true and estimated n'th order inputs are then comparable, see Figure 4.11. This input spectra are used to estimate the system NOFRFs for all six DEAs using Algorithm 4.1.

The output spectra of all six DEAs exhibit a similar shape and composition at all orders except for in the case of DEA 1. However, as we have noted before the behaviour of DEA 1 is different from the other actuators. It is therefore interesting to note the difference in the output spectra which clearly indicate a divergence from the normal system behaviour.

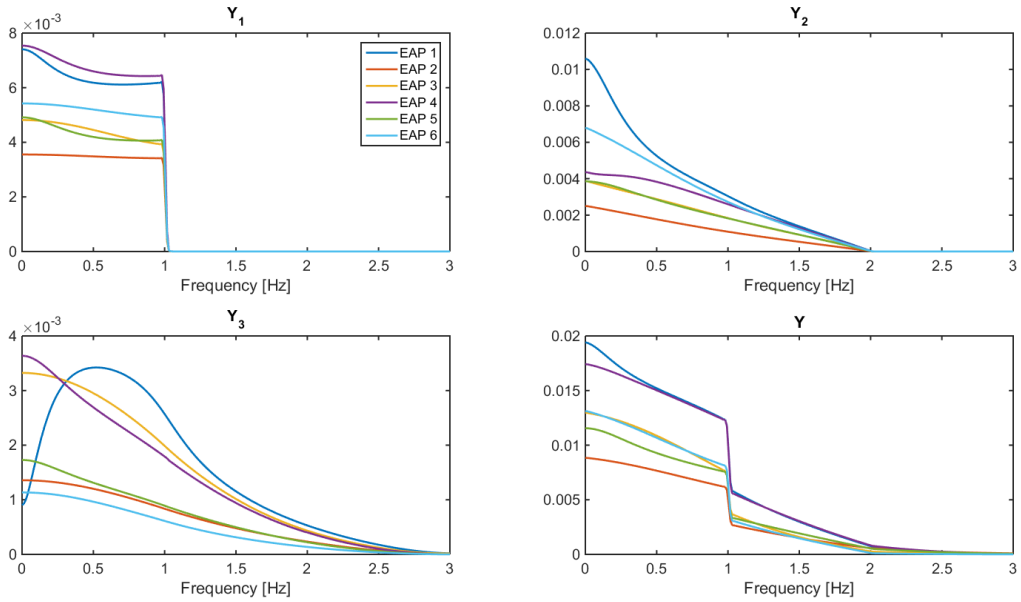

Figure 4.12: **NOFRFs give a unique insight into how the different non-linear orders of the system contribute to the output spectrum.** The n'th order output spectra at orders $n = 1, \ldots, 4$, of the DEA system for all six actuators are shown along with the total output spectra (Black)

4.5 Discussion

The aim of this chapter was to develop a framework for modelling and analysis of DEAs that would describe the non-linear dynamics of these actuators. In particular the models were identified with a focus on their predictive ability and in a form useful for the application of control. The framework was developed using the NARMAX methodology, including joint structure detection and parameter estimation, as well as frequency domain analysis using both GFRFs and NOFRFs. The framework that has been developed was then demonstrated using film-type DEAs. The identified models were found to accurately capture the behaviour of the DEAs over independent validation data. The accuracy of the predictions was shown to be vastly improved in comparison to identified linear models.

The necessity of considering the inherent non-linear behaviour when modelling DEAs has been emphasised in this work and is consistent with other DEA models. The need to use non-linear models to accurately describe DEA dynamics, found here, is consistent with current models of DEAs derived by considering the physical derived by considering the physical properties of the material [63, 108, 119, 134]. A distinct, novel contribution of the work presented in this chapter is that the NARX models identified are compact difference equation descriptions with a parsimonious model structure, which is in contrast with many

first-principle type models, making them particularly amenable to control design. Further to this, the NARMAX modelling framework is able to represent a broad range of non-linear systems because the model class can be used to describe general multiple-input multiple-output systems with a defined input and output [16]. Therefore, it is likely that the modelling and analysis framework could be applied to a wider range of actuator configurations in the future.

GFRF and NOFRF analysis has been performed on the DEA system leading to a frequency domain description of the system that is much more readily interpreted than a time domain description alone. The structure of the NARMAX models identified for the set of DEAs under investigation varies significantly between different actuators. However, this provides no evidence to the similarities/differences present in the underlying system. In contrast, the frequency domain description provided by the GFRFs and NOFRFs provides an invariant description that is directly comparable across different models. In particular the NOFRFs estimated for the models provide a convenient method for comparing the behaviour. This is because the NOFRFs can be displayed in two dimensions at all orders of non-linearity.

By comparing the NOFRFs for the set of DEAs it is seen that the dynamic behaviour of the actuators it inconsistent across the set despite all of the DEAs being fabricated to the same specifications. The difference in the behaviour of the DEAs thus represents inconsistency in the fabrication procedure of the actuators. In the absence of improved fabrication techniques the inconsistency in the DEA behaviour necessitates control design that is able adapt to the different properties or that can be tailored to each specific actuator.

It is known that the DEAs exhibit significant time varying characteristics which represents another modelling challenge. However, the analysis techniques used in this chapter are also only appropriate for time invariant systems. The investigation into extending the modelling and analysis framework presented in this chapter to the time varying case is the subject of the next chapter.

4.6 Summary

In this chapter a framework has been developed for the identification of control-focused models of a set of DEAs which accurately describes the non-linear behaviour of the system. The framework is then extended to include advanced frequency domain analysis techniques using both GFRFs and NOFRFs. This analysis allows direct comparisons to be made between different actuators, a feature that is not possible when considering the time domain alone. The framework

was demonstrated on six different DEAs, revealing differences in the dynamic behaviour despite being fabricated using the same method. These results demonstrate the successful application of the methodology as well as providing further insight into the behaviour of film type DEAs.

In the next chapter the modelling and analysis framework will be extended to cater for the inherent unexplained time-varying behaviour exhibited by film type DEAs.

Chapter 5

Frequency Response Analysis of DEAs with Time Dependent Characteristics

In this chapter three main novel contributions are made; The first contribution addresses the problem of modelling the time variations displayed by DEAs. Secondly, a novel method for time varying frequency domain analysis, based on time varying NOFRFs (TV-NOFRF)s, is proposed. Thirdly, the new TV-NOFRF method is applied to the set of six film-type DEAs in order to provide further insight into the dynamic behaviour of these actuators.

The set of DEA actuators that have been studied in the previous chapters are shown to display time variations in their dynamic behaviour. In Chapter 2 the current state of the art in the modelling of EAPs was discussed and it was noted that most current models do not describe, or are limited to, a simplified assumption of time varying phenomena [100, 118], however, the origin of time variations displayed by the actuators in this study are unknown. The lack of methods for accurately handling these time variations represents a significant gap in the current methodology for modelling DEAs, especially for control-oriented tasks.

The field of system identification provides a wealth of techniques for linear TV systems. Traditionally these are based on recursive methods such as RLS and Kalman filtering as well as others [75]. More advanced methods have gained popularity in recent years in which the TV parameter vector is expanded using a set of basis functions in order to make the TV problem into a time invariant problem that can be solved with, for example, least squares [86, 131, 142]. Incorporating the TV phenomena into the modelling framework developed in the previous chapter is greatly simplified by linear-in-the-parameters form of the NARX model used to

represent the system. The linear-in-the-parameters form allows that many of the algorithms available for linear time invariant problems are readily applicable to models of the NARX class.

Although the parameter estimation is simplified, the frequency domain analysis of TV non-linear systems represents a significant challenge. To the best of the author's knowledge, only one attempt to tackle this problem has been made in the literature. In that work the authors use GFRFs in order to perform the frequency domain mapping [49]. However, GFRFs are unable to provide a complete description of the output frequency response of the system.

The problem of characterising the dynamic behaviour of DEAs with respect to their time variations therefore motivates the development of a novel method for frequency domain analysis of TV non-linear systems. The method is based on NOFRFs which are detailed in Section 4.3.2. As discussed in that section, the advantages of using NOFRFs are significant in both implementation, analysis and simplicity of calculation.

The new analysis method allows for a greater insight into the time varying phenomena and provides a comparison across actuators that are fabricated to the same specifications providing a method for gauging the consistency of the fabrication process. It also provides much clearer insight into the time-varying dynamics of the DEAs than examining the time-variation of the model parameters alone.

The objectives of this chapter are summarised below

1. The extension of the identification framework developed in the previous chapter to accommodate time variations in DEA systems.

2. The development of a novel method for the time varying frequency domain analysis of non-linear systems.

3. The application of the modelling and analysis framework to the set of DEA actuators studied in previous chapters in order to analyse time varying behaviour in the frequency domain.

Part of the work presented in this chapter has been published previously by the author in a peer reviewed journal [54].

The remainder of this chapter is structured as follows: Firstly, in Section 5.1 a method for representing time varying non-linear systems is introduced in both the time domain as time varying NARX (TV-NARX) models and in the frequency domain as time varying GFRFs (TV-GFRF)s. This approach is then extended to the concept of TV-NOFRFs and is incorporated into the modelling and analysis framework introduced in the previous chapter. In Section 5.2 the time varying

frequency domain analysis procedure is applied to the set of six DEA actuators and results are presented. The chapter is concluded with a discussion of the results and a summary.

5.1 Frequency domain analysis of time-varying non-linear systems

To the best of the author's knowledge the work in [49] represents the only method for the frequency domain analysis of TV non-linear systems found in the literature. The work is based on extending the GFRF concept to the TV case by evaluating the system GFRFs for the model parameters at each sample time. In the previous chapter it was shown how NOFRFs can be seen as an extension to GFRFs. The TV-NOFRF method is similarly linked to that of TV-GFRFs. The TV-GFRF method is briefly discussed in the following section before being extended to the novel TV-NOFRF method.

5.1.1 Representation of time varying non-linear systems

Polynomial NARX/NARMAX models are discussed in detail in the previous chapters and are here extended in order that they can represent TV non-linear systems. The extension of the NARX/NARMAX model to the time varying case can simply be achieved by replacing the parameter vector $\boldsymbol{\theta}$ with a time varying parameter vector $\boldsymbol{\theta}(k)$ such that

$$y_k = \boldsymbol{\phi}_k \boldsymbol{\theta}(k) + e_k. \tag{5.1}$$

The estimation of the TV parameter vector, $\boldsymbol{\theta}(k)$ is greatly simplified by the linear in the parameters form of the model class. Parameter estimation for the TV-NARX model can be performed by a number of standard recursive parameter estimation techniques.

5.1.2 Time varying generalised frequency response functions

Investigating the frequency response of time varying non-linear systems is a difficult task. Methods that are available for time invariant systems mainly depend on taking transforms over a sufficient number of data points from a time domain data record. Capturing rapid changes in system dynamics may therefore not be possible. However, For a non-linear system that can be represented by a parametric model, many methods exist for tracking the parameters as they evolve over time, for example, Kalman filters [57], recursive least squares [75] and wavelets

[131]. If the use of a fixed model structure can be justified over the entire data record then the GFRF based frequency domain analysis methods discussed in the previous chapter can easily be extended to the TV case, greatly simplifying the analysis procedure [49].

The GFRFs obtained from NARX models are a function of the model structure and parameters. Therefore, assuming a fixed model structure leads to time variations in the parameter vector only. The n'th order GFRF, given by equation 4.35, can then be extended to the TV case by explicitly including a dependency on time t such that

$$H_n(\omega_1,\ldots,\omega_n,t) = \frac{H_{n_u}(\omega_1,\ldots,\omega_n,t) + H_{n_{uy}}(\omega_1,\ldots,\omega_n,t) + H_{n_y}(\omega_1,\ldots,\omega_n,t)}{1-\sum_p \theta_p(t) e^{-j(\omega_1,\ldots,\omega_n)k_p}}, \tag{5.2}$$

where again, the time dependency of the functions $H_{n_u}(\omega_1,\ldots,\omega_n,t)$, $H_{n_{uy}}(\omega_1,\ldots,\omega_n,t)$ and $H_{n_y}(\omega_1,\ldots,\omega_n,t)$ is solely due to the time varying parameters.

As discussed in Section 4.3, the GFRF based analysis of non-linear systems is largely based on the identification of peaks and ridges in the first and higher order GFRFs respectively. This is because the peaks and ridges in the frequency response indicate frequencies and combinations of frequencies of the input spectrum that produce strong non-linear effects in the output [20]. The relative directions and magnitudes of the peaks and ridges can provide a description of how the non-linear behaviour exhibited by the system.

[49] show that the direction of the ridges in the gain of n'th order GFRF are mainly dependent on the NARX model structure and that there is always a ridge in the $\omega_1+\omega_2+\ldots+\omega_n = C_i$ direction with the possible existence of extra ridges in other directions. The position C_i of the ridge is found at the minimum of the denominator in Equation (5.2) such that

$$C_i = \underset{\omega_1+\omega_2+\ldots+\omega_n}{\arg\min} \left| 1-\sum_p \theta_p(t) e^{-j(\omega_1+\omega_2+\ldots+\omega_n)k_p} \right|. \tag{5.3}$$

Visualisation of TV non-linear systems with TV-GFRFs for the n'th order GFRF requires a $(n+2)$ dimensional space. Therefore, for $n > 1$, so that the dimensionality of space required is > 3, visualisation of the TV-GFRF is difficult. However, it is often the case that only a small amount of ridge directions are observed, and the direction depends on the NARX model structure. By averaging in the ridge direction the dimensionality of the problem can be reduced to 2 dimensions (with the third being time). If there is only one ridge displayed by the system then no in-

formation is lost in the dimensionality reduction. The averaging can be performed via the equations

$$\left|H_n^{\omega_1+\omega_2+\ldots+\omega_j}(\omega,t)\right| = \int_{\omega_1+\omega_2+\ldots+\omega_j=\omega} \frac{1}{N_\omega} \left|H_n(\omega_1,\ldots,\omega_n,t)\right| \mathrm{d}\omega \tag{5.4}$$

and

$$\phi\left(H_n^{\omega_1+\omega_2+\ldots+\omega_j}(\omega,t)\right) = \int_{\omega_1+\omega_2+\ldots+\omega_j=\omega} \frac{1}{N_\omega} \phi\left(H_n(\omega_1,\ldots,\omega_n,t)\right) \mathrm{d}\omega, \tag{5.5}$$

for the gain and phases of the system respectively, where $\phi(\cdot)$ denotes the phase and N_f is the number of samples along the direction of the ridge at each frequency $\omega = \omega_1 + \omega_2 + \ldots + \omega_j$. The variable j depends on the models structure.

The TV-GFRF analysis scheme [49] for a TV nonlinear system that can be described by a fixed model structure with a TV parameter vector follows as

Step 1: Calculate the system TV-GFRFs, $H_n(\omega_1,\ldots,\omega_n,t)$, via Equation (5.2) up to the required order, N_m^*.

Step 2: Evaluate the TV-GFRFs at each sample time $t = kT_s$.

Step 3: Determine the GFRF ridge directions.

Step 4: Average the evaluated TV-GFRFs via equations (5.4) and (5.5).

This approach hence allows the higher order GFRFs of a TV non-linear system to be visualised in a lower dimensionality providing insight into how the input frequencies combine to generate complex non-linear effects. However, there are disadvantages to this scheme in terms of the computational cost involved in evaluating the GFRF due to their high dimensionality. This is further exacerbated in calculating the symmetric GFRF because the number of permutations across all the input frequencies also increases rapidly with the GFRF order. This calculation must then be performed at each time step.

More importantly, from the TV-GFRF alone it is not possible to find exactly how the system dynamics behave in generating the system output for a given input excitation. NOFRFs, an extension of the GFRF framework, were introduced in the previous chapter and are used in the next section to address the time varying case.

5.1.3 Time varying non-linear output frequency response functions

In this section a novel approach to the problem of frequency domain analysis of TV non-linear systems is introduced. NOFRFs that were introduced in the previous

section are extended to the TV case in a simple and efficient manner.

In the previous section the n'th order GFRF was extended to the time varying case given by Equation (5.2). Similarly, the TV-NOFRF can be constructed by introducing the dependency in time, t, into the n'th order GFRF - $H_n(j\omega_1, \ldots, j\omega_n)$ such that

$$G_n(j\omega_n, t) = \frac{\int\limits_{\omega_1+\ldots+\omega_n=\omega} H_n(j\omega_1, \ldots, j\omega_n, t) \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}}{\int\limits_{\omega_1+\ldots+\omega_n=\omega} \prod\limits_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega}}. \tag{5.6}$$

which is now dependent on time through the TV parameters of the NARX model.

Following the work by [49] a TV-NARX model with a fixed structure is considered such that the time dependency is entirely on the parameter vector. The evaluation of TV-NOFRFs can then be achieved by considering the TV-NARX model as N time invariant NARX models with parameter vectors $\boldsymbol{\theta} = \boldsymbol{\theta}(t)$ which can each be simulated separately for a given test input of interest. The TV-NOFRFs are then evaluated simply and efficiently by the iterative application of Algorithm 4.1 for each model. The novel TV-NOFRF analysis scheme is summarised in Algorithm 5.1.

Algorithm 5.1 An algorithm for determining TV-NOFRFs of a TV generative system.

Procedure
Convert the TV-NARX model into N NARX models.
for $k = 1, \ldots, N$
$\mathcal{M}_k^{TI} = \mathcal{M}^{TV}|_{\theta_k}$
end for

Generate TV-NOFRF
for $k = 1, \ldots, N$
Apply Algorithm 4.1 to model $\mathcal{M}_i^{TI}$ to produce $G_1^k, \ldots, G_N^k$
end for
end Procedure

The TV-NOFRF method provides a tool for determining how the non-linearities of the system transfer energy from the input spectrum to the output spectrum at each time point. This allows time varying frequency domain analysis that is not possible using other methods.

The new method has a significant advantage over the TV-GFRFs in that the system GFRF equations do not need to be evaluated, or even known because the NOFRFs are estimated using least squares over a set of simulated outputs trans-

formed into the frequency domain. Firstly, this makes it significantly simpler to implement. Secondly it is much less computationally expensive because the high dimensional function evaluations are not required. The necessity to reduce the dimensionality by averaging over the ridge directions is also removed because NOFRFs are 2-dimensional functions for all orders.

Note that although the input spectrum is independent of time in Equation (5.6) the extension to a time varying input signal is trivial if the required test input can be designed, and this incurs no extra computational cost.

5.1.4 TV-NOFRF analysis procedure

In order to apply the TV-NOFRF analysis procedure to a non-linear model/system it is necessary for the model/system to be operating around the zero equilibrium position. It may therefore be necessary to remove any DC component from the NARX model should it be present. This can be performed using the method discussed in Section 4.2.

Standard system identification techniques designed for time invariant systems may not be valid for TV systems. It is therefore necessary to validate that the identified model is able to describe the system dynamics across all time. In the case that such a model cannot be justified a number of TV identification techniques are discussed in [49].

The procedure is summarised as follows

Step 1: Identify a nonlinear TV model of the system using standard identification techniques.

Step 2: Estimate the TV-NARX model parameters via, for example, RLS, Kalman filtering, etc.

Step 4: Calculation of TV-NOFRFs for using Algorithm 5.1.

The procedure is demonstrated and validated by way of two examples. Firstly the case where the system is operating around the zero equilibrium and secondly where a time varying DC component is present.

Example 5.1 Time varying frequency domain analysis of a NARX system.

The following TV-NARX system is given as an example in [49] in which analysis is performed using GFRFs and it is repeated here to, firstly, provide a direct comparison and secondly to demonstrate the advantages in the method presented in this work.

$$y_k = \theta_1(t)y_{k-1} + \theta_2(k)y_{k-2} + \theta_3(k)u_{k-1} + \theta_4(t)u_{k-1}^2, \tag{5.7}$$

where the parameters dependence on the time, t, indicates that they vary with time such that

$$\begin{aligned}
\theta_1 &= \begin{cases} 1, & 0 \leq k\Delta t < 0.2s, \\ 0.4, & 0.2 \leq k\Delta t < 0.4s, \\ 0.8, & 0.4 \leq k\Delta t \leq 1s, \end{cases} \\
\theta_2 &= -0.3, \\
\theta_3 &= 0.1, \\
\theta_4 &= \begin{cases} 0.2, & 0 \leq k\Delta t < 0.6s \\ 0.5, & 0.6 \leq k\Delta t < 0.8s \\ 0.4, & 0.8 \leq k\Delta t \leq 1s \end{cases}
\end{aligned}$$

The TV system is used to generate input-output data. The input, u_k, is a white noise sequence drawn from the zero mean normal distribution with variance $\sigma_u^2 = 1$. The error signal e_k is i.i.d. noise drawn from the normal distribution with variance $\sigma_e^2 = 0.0004$. The sampling frequency is $f_s = 400$ Hz, $\Delta T = 1/400$ and the system is simulated for $1s$ leading to $N = 400$ input-output data samples.

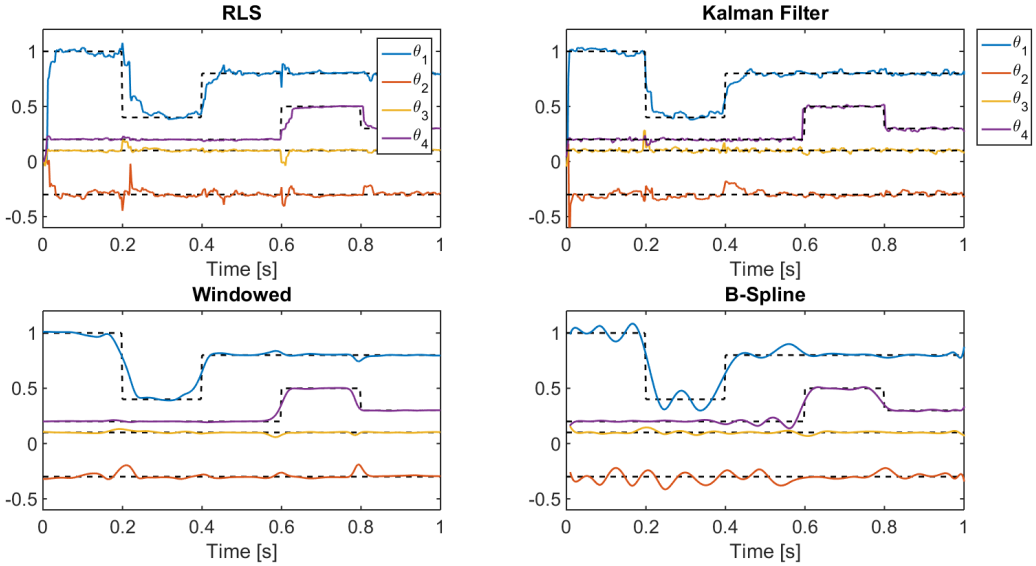

Figure 5.1: **Time varying parameter estimation techniques can be used to estimate parameters as the system changes over time.** Time varying parameter estimates for the system given by Equation (5.7) using: RLS, Kalman filtering, Windowed LS and B-Spline wavelets.

Assuming that the non-linear model structure is known, the TV parameters can be estimated from the system input-output data. TV parameter estimation is performed using both on-line and off line methods. The on-line methods of RLS and Kalman filtering are applied to the system using equations (3.36)-(3.39) and (3.42)-(3.47) respectively. The initialisation for the RLS algorithm was $\lambda = 0.83$ and $P_0 = 10\boldsymbol{I}$. The Kalman filter was initialised as $R = 0.1\boldsymbol{I}$, $Q = 0.1\boldsymbol{I}$ and $P_0 = 1^5\boldsymbol{I}$ where the initialisation of the parameters was chosen based on preliminary tests.

For both algorithms the initial parameter vector is assumed unknown and as such is initialised to zero. The initial confidence in the parameter estimate is therefore low leading to the selection of the values of P_0. Two further off-line methods were used. Firstly by windowing the data and calculating the parameters with Least squares. A window length of 50 was used incremented one sample at a time such that the dimensionality of the data is not reduced. A Hanning window is applied to the windowed data. A second off-line method using B-Splines is also applied following [49].

The parameters estimated with RLS have a slow transition when the system exhibit a rapid change in comparison to the parameters found using a Kalman filter which provides a significantly improved parameter estimate. The off-line parameter estimates can see future data allowing them to react to a change in the system dynamics before they occur in time. The parameters calculated using a moving window can be made to react faster to sudden parameters changes by reducing the window length at the expense of decreasing the accuracy of the estimate, see Figure 5.1.

After the time varying parameter estimates of the system have been made NOFRFs can be calculated by the application of Algorithm 5.1 and hence the relevant variables need to be initialised. The highest order non-linear term in the NARX model given by Equation (5.7) is a second order pure input term, u_{k-2}^2, and no other terms exist for $n > 1$. Therefore the highest order GFRF of the system will be second order and as such no higher orders will influence the output of the system and therefore do not need to be considered. The algorithm variables are hence set as $\hat{N} = 2$ and α takes 6 values linearly spaced in the range $\alpha = 0.1, \ldots, 1$. The system is analysed under a number of different excitation signals. The input test signals are generated using Equation (4.70) with $[a, b] = [0, 200]$, $[0, 80]$, $[60, 140]$, and $[120, 200]$Hz.

Both the TV-GFRFs and the TV-NOFRFs are calculated using the true system parameters. The TV-NOFRFs are found for all four of the test inputs. As should be expected in this case, the first order TV-GFRF and TV-NOFRFs are equal when the test input signal is uniform across all frequencies and the second order functions

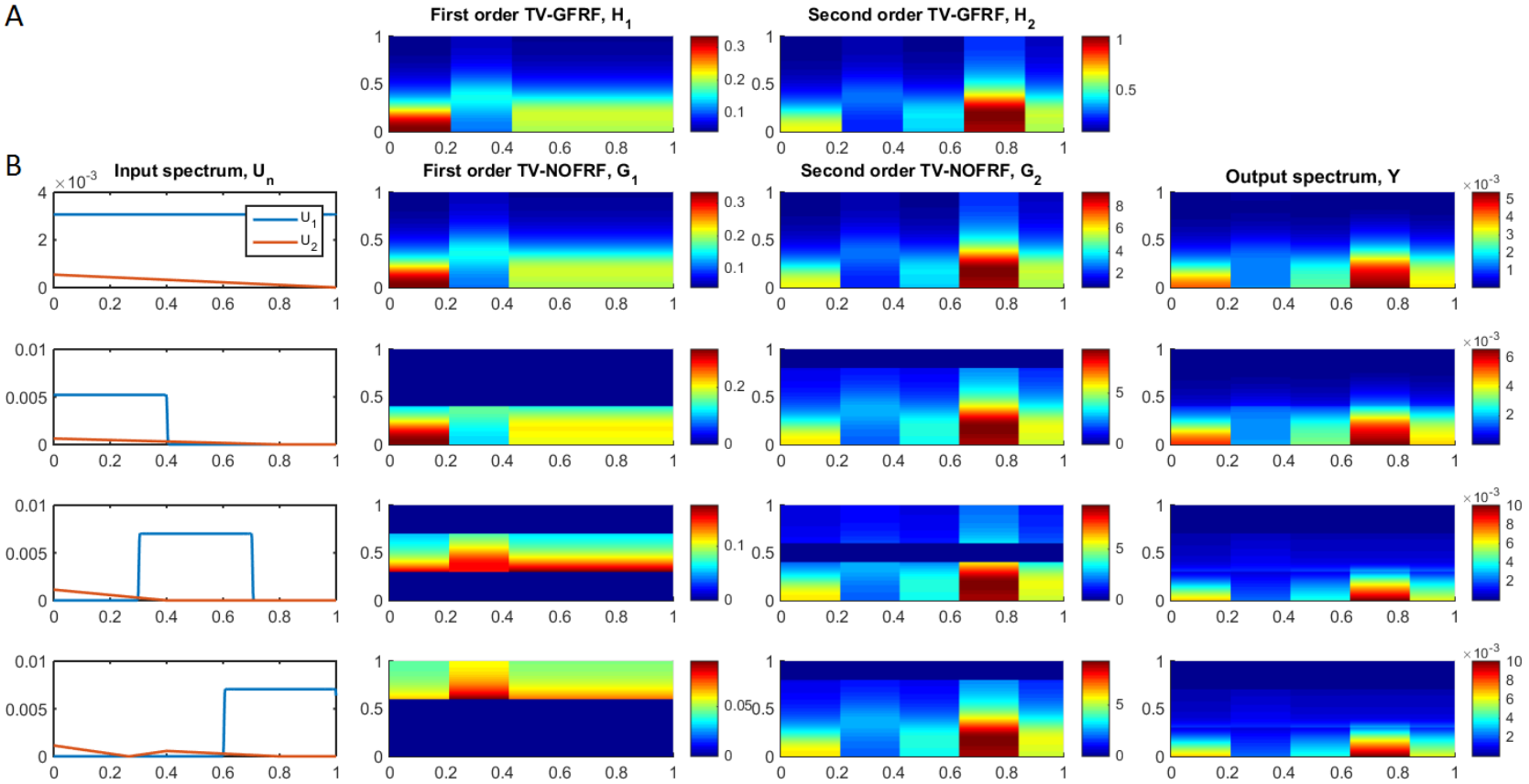

Figure 5.2: **Calculating the systems NOFRFs over time using Algorithm 5.1 provides an input dependent frequency domain description that cannot be found with GFRFs alone.** A) Center: The first order GFRF H_1 Right: the averaged second order GFRF H_2 B) Left: The n'th order spectrum of test input signals used to generate the NOFRFs. Center: The first order NOFRFs resulting from the test inputs in A. Right: The second order NOFRFs resulting from the test inputs in A. The NOFRF is response to a broadband excitation signal gives the same frequency response as the averaged GFRF in this case. For any non uniform input spectrum the NOFRFs give a frequency domain description not provided by the GFRFs.

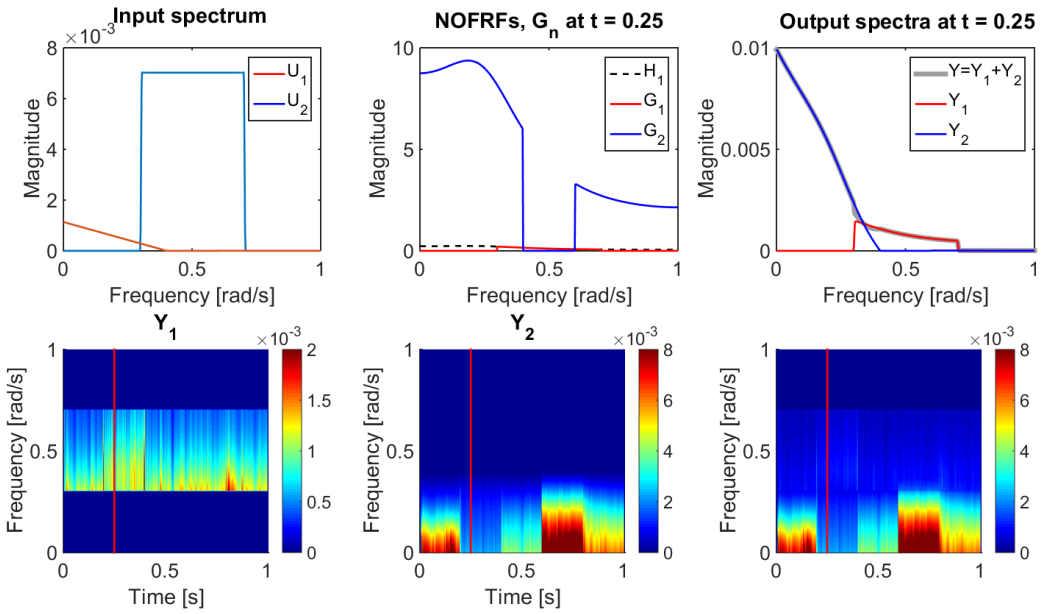

Figure 5.3: TV-NOFRFs evaluated with parameters estimated by Kalman filtering provide a good description of the system output Spectra. TV-NOFRFs excited by a frequency band input excitation in the region $[60, 140]$Hz. The NARX model parameters are estimated with a Kalman filter.

are proportional at all frequencies. This occurs because the set of basis functions that form the model contains only a pure input non-linearity. There are therefore only ridges in a single direction. The integration involved in the averaging in Equation (5.6) is equivalent to the integration for the calculation of the TV-NOFRF in Equation (5.4) except for a scaling factor, when the input spectrum is uniform, see Figure 5.2.

The TV-NOFRFs calculated for the three band excitation inputs clearly show that the system exhibits very different behaviour in response to different excitations. The dissection of the frequency domain discretion into their corresponding model orders in this way is not possible using other methods.

The TV-NOFRFs are also found for the parameters estimated with the Kalman filter with the excitation signal of frequency band $[60, 140]$Hz, see Figure 5.3. The total output spectra is observed to be very similar to that calculated with the true parameters indicating the usefulness of the method for real systems if the system can be well represented by a NARMAX model, see Figure 5.3.

△

Example 5.2 Time varying frequency domain analysis of a NARX system

with a time varying DC component.

In this example the effect of a TV DC shift in a TV non-linear system is demonstrated using the TV-NOFRF analysis procedure. The example is given in order to demonstrate how the changing mean level in a non-linear system needs to be taken into account in order to produce correct frequency domain analysis results. The change in the mean level, modelled by a time varying DC component, will change the equilibrium position over time. It is important to keep track of the equilibrium position in order to perform the frequency domain mapping. The example imitates the sort of behaviour expected to be displayed by the film type DEAs. Consider the generative system including a time dependent DC term given by the equation

$$y_k = \theta_1(t) + \theta_2 y_{k-1} + \theta_3 y_{k-2} + \theta_4 u_{k-1} + \theta_5 y_{k-2}^2, \tag{5.8}$$

where

$$\boldsymbol{\theta(t)} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \end{bmatrix} = \begin{bmatrix} 0.1t \\ 0.8 \\ -0.3 \\ 0.1 \\ 0.25 \end{bmatrix}, \tag{5.9}$$

such that the DC term takes linearly spaced values in the range $[0, 0.1]$ over time time interval $t = [0, 1]$, see Figure 5.4.

In order to apply the TV-NOFRF analysis method to the system it is first necessary to transform the system so that it is operating around the zero equilibrium position for all time. This is achieved with the method used in Section 4.2. The system equilibrium position is found by setting $y_k = y_0$ and $u_k = 0 \; \forall \; t = kT_s$ such that

$$y_0 = \theta_1 + \theta_2 y_0 + \theta_3 y_0 + \theta_5 y_0^2 \tag{5.10}$$

which rearanges to form a quadratic equation in y_0

$$0 = y_0^2\theta_5 + y_0(\theta_2 + \theta_3 - 1) + \theta_1 \tag{5.11}$$

the roots of which are simple to find. The DC component is now removed component from the system given by Equation (5.8). The parameter vector is time varying and so the equilibrium position is also time varying $y_0 = y_0(t)$, see Figure

5.4**B**. The new system is then given by

$$z_k = \theta_2^* z_{k-1} + \theta_3^* z_{k-2} + \theta_4^* u_{k-1} + \theta_5^* z_{k-2}^2, \tag{5.12}$$

where

$$\boldsymbol{\theta}^* = \begin{bmatrix} \theta_1^* \\ \theta_2^* \\ \theta_3^* \\ \theta_4^* \\ \theta_5^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0.8 \\ -0.3 + 2\theta_5 y_0 \\ 0.1 \\ 0.25 \end{bmatrix},$$

and the two systems are linked by the equation $y_k = z_k + y_0(t)$. The parameter vector of the new system has been modified such that θ_4^*, corresponding to the z_{k-2} term, is now dependent on time, see 5.4**C**.

The TV-NOFRF analysis method can be applied to the system given for z_k which has the same response as the system given for y_k at all frequencies except for the zero frequency component. The analysis reveals how the dynamics of the system change over time. The DC constant has propagated through all of the non-linear orders of the system and its effect on the n'th order output is clearly visible, see Figure 5.4**D-E**.

△

The two examples demonstrate how the TV-NOFRF analysis method can successfully be applied to the analysis of non-linear systems both with and without a DC term present.

5.2 Time varying frequency domain analysis of DEAs

The remainder of this chapter concerns the analysis of DEAs using the TV frequency domain analysis techniques for non-linear systems that have been introduced. The time varying behaviour of DEAs was first discussed in Chapter 2. It is clear from the observed data that the actuators are influenced by some form of time-varying dynamics, this can clearly be seen in the changes in DC level and gain for the DEAs. The physical interpretation of these dynamics are unknown, however, and it is necessary for the phenomena to be described for the development of a conventional control strategy. The advantage of the NARX model approach is that the underlying phenomena causing time-variation does not need to be known. Instead, the time-variation can be described in the model using the observed data.

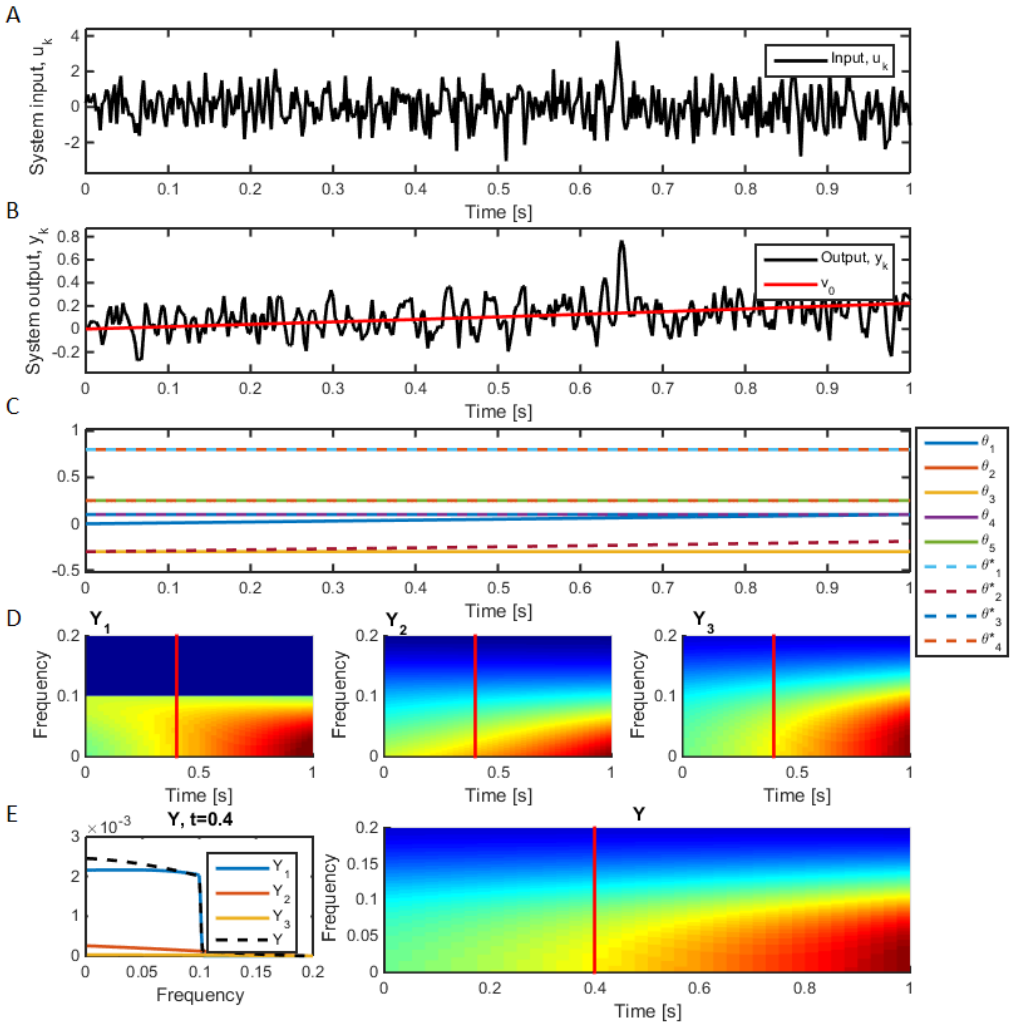

Figure 5.4: **The TV-NOFRF analysis reveals how the time varying DC component of a NARX model affects the energy transfer at all frequencies. A**) Time invariant input excitation signal u_k. **B**) TV system output y_k (Black) with the TV system equilibrium position $y_0(t)$ **C**) Parameters $\boldsymbol{\theta}$ of the system given by Equation (5.8) (Solid Lines) with parameters $\boldsymbol{\theta}^*$ of the DC removed system given by Equation (5.12) (Dashed Lines). **D**) n'th order output spectrum of the DC removed system given by Equation (5.12). **E**) Total output spectrum of the DC removed system given by Equation (5.12).

As was argued in the previous chapter, inspection of the time-domain DEA data in Figure 2.5 does not directly reveal whether the system dynamics are consistent across the different actuators. Similarly, calculating parameters using either on-line or off-line TV parameter estimation techniques provides a method for accurately predicting the system output, but provides no direct information as to how the dynamics of the system are changing. However by extending the fre-

quency domain analysis methods to the TV case, the frequency response of the system plotted as a function of time provides a common description that allows for a much better understanding.

As in the previous chapter, the TV-NOFRF method is first demonstrated on DEA 5 and then results for the entire set of actuators are given in order to provide a comparison.

5.2.1 Pre-analysis of the TV DEA data set

For clarity of reading, the raw data for DEA5 is re-plotted in Figure 5.5. The band limited excitation signal for all the DEAs is stationary, see Figure 5.5**A**, and so any TV in the output is solely due to time variations in the system dynamics. By only observing the time domain output it is unclear how the system dynamics are changing.

The system response is roughly linear at lower voltages with the non-linearities having a larger effect as the voltage increases. The general form of the non-linearity remains similar across time. This conclusion was reached by observing that the shape, or functional form, of the voltage displacement relationship was an approximately quadratic non-linearity that mainly varies in average displacement over time, see Figure 5.5**C**. Therefore it is assumed that the DEAs can be described by a fixed non-linear model structure with time-varying parameters, this is further justified in the next section.

Windowing is applied to the data using a Hanning window with a relativity long window length of 10000 samples and an increment of 300 samples. An estimate of the signal power of the output is taken over each data window with the DC component removed so that the shift in the mean level of the signal is not included in the analysis. The result is heavily influenced by the large peaks in the output data but it is clear that there is a general increase in the signal power in the region $500 < t < 1500$, see Figure 5.5**D**. This indicates that the system is not only changing in its average displacement but that the gain of the system is also changing in response to some unknown system dynamics.

5.2.2 Estimation of time varying parameters

The NARMAX models identified in Chapter 4 are used to describe the DEA actuators, see Table 4.1 for the process model and Table 4.2 for the noise model. It is assumed that the system dynamics can be accurately captured across all time by a time-invariant non-linear model structure. In order to justify this assumption the identified model structure, given in Table 4.1, are simulated over four different

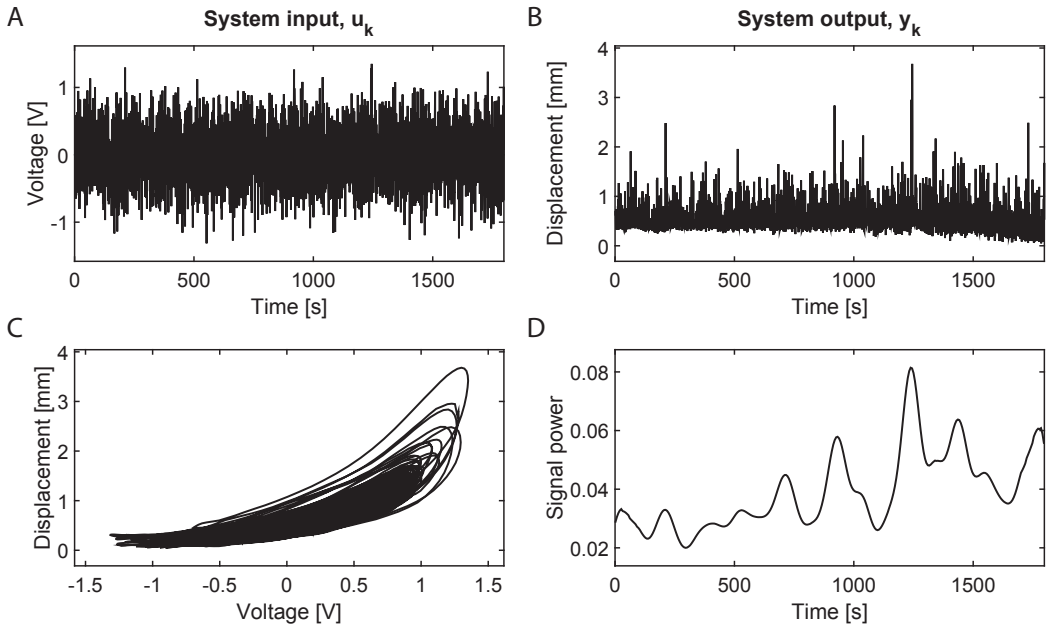

Figure 5.5: **The DEA system exhibits unknown time varying dynamics that is observable in an average change in displacement as well as changes to the gain.** **A**) 1Hz Bandpassed random input excitation signal. **B**) System Output exhibiting time varying effects. **C**) System input plotted as a function of the system output revealing the hysteresis. **D**) Approximate signal power for frequencies $\approx > 0Hz$ calculated using windowing.

time periods linearly spaced across the data set ($4-104$s, $304-404$s, $604-704$ s, $904-1004$s and $1204-2304$ s), with parameters re-estimated in each case using least-squares. Note that the first time segment is the same as the data over which the model structure was identified. The model gave accurate predictions in each case with model fit measured by the MSSE of $= 6.67,\ 8.37,\ 6.13$ and 31.74 $(\times 10^{-4})$ respectively. The model therefore provides a good fit to the data well over each of the different time segments. The relatively poor fit for the model over time $1204-2304$s can be explained by TV behaviour over the data. This conclusion can be drawn by noting that the model predictions are consistently larger than the system output, see Figure 5.6

Given that the model structure well represents the data with a time varying parameter vector, parameters are now estimated across time. Time varying parameters are calculated recursively using the Kalman filter equations (3.42)-(3.47). The Kalman filter can be used on-line and so is appropriate for the TV modelling for the application of control where future predictions need to be made. Off-line Kalman smoothing [38], using equations (C.4)-(C.12), is also performed in order to remove rapid changes in the TV parameter estimates, see Figure 5.6. This is

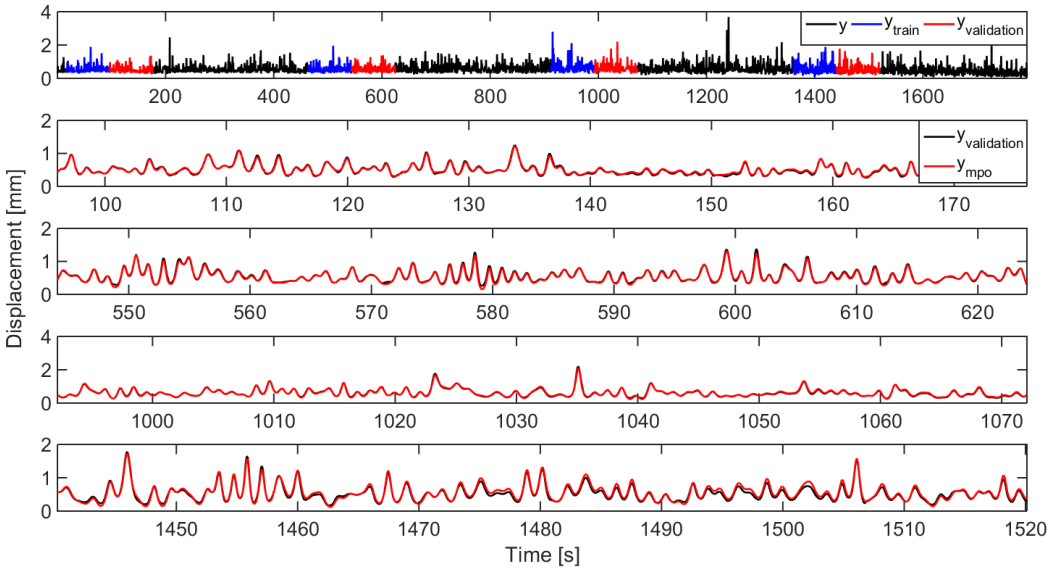

Figure 5.6: **The non-linear model structure of the DEA identified in Chapter 4 can accurately represent the system across time with a time varying parameter vector.** The model predicted output plotted with the true output for DEA5 at 5 different time segments. For each time segment parameters are estimated on an estimation data set and the fit is measured on a separate validation data set. The MSSE for each data segment is MSSE $=$ 6.67, 8.37, 6.13 and 31.74 ($\times 10^{-4}$) respectively.

done in order to obtain a clear view of the frequency domain behaviour. In order to perform subsequent NOFRF analysis it is necessary to remove the DC term from the model. After the parameters are estimated the noise model is discarded and the DC term of the TV-NARX model is removed using the method detailed in Section 4.2 at each time point. This recovers the time varying equilibrium point of the system $y_0(k)$, see Figure 5.7**B**. As it might be expected, the equilibrium point is approximately proportional to the average displacement of the data. The process is repeated for the entire set of actuators, time varying parameters and equilibrium position are shown in appendix C.2.

5.2.3 TV-NOFRF analysis of DEAs

TV-NOFRF analysis can now be performed on the DEA system using the DC removed TV-NARX model via Algorithm 5.1. The algorithm variables are chosen as they were previously in Section 4.4.4 and the same test input signal is used as before at each time point. Considering the results of the analysis for DEA 5 it can be seen that that the dynamic behaviour of the system changes significantly over the data record. It can be seen that the magnitude of the first order NOFRF

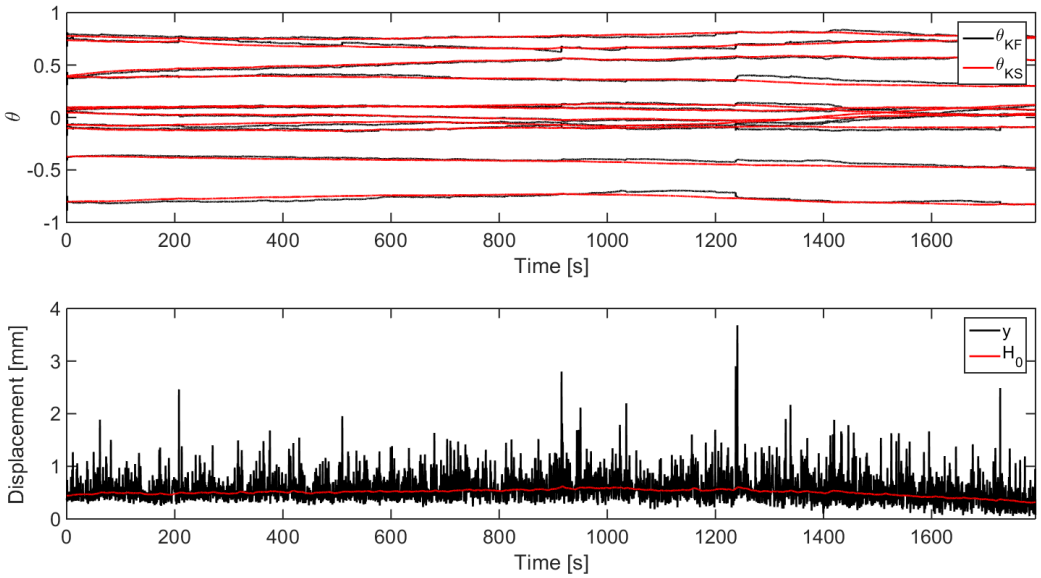

Figure 5.7: **Time varying parameters are estimated on-line using a Kalman filter and updated off-line using a Kalman smoother, the DC component is then removed from the model.** Top: Time varying parameters for DEA 5 calculated using a Kalman filter (Black) and updated using a Kalman smoother (Red). Bottom: The output $_k$ for DEA 5 (Black) with the TV equilibrium position, $y_0(k)$

and output spectra increase with time up to $t \approx 1250$ at which point it starts to decrease, see Figure 5.8. Interestingly this time point coincides to a large peak in the time domain output of the system, see Figure 5.7. The operating point also increases up to this time point after which it starts to decrease. In contrast the third order output spectra increases with time across the entire data record. The second order output spectra is comparatively uniform, although it also peaks at $t \approx 1250$. The Output spectra at time $t = 1100s$ is also plotted to shown the form of the spectra in a lower dimensional representation.

The TV-NOFRF procedure is repeated on the remainder of the set of DEAs. Time varying parameters and equilibrium points are plotted in Appendix 5.2 for each DEA, see Figures C.1-C.2. The results of the NOFRF analysis are plotted in Figures 5.9 and 5.10. It is observed that all of the actuators undergo changes in their dynamic behaviour over time. Firstly, it can be noted that the largest contribution to the output comes from the first order (linear) behaviour. For the majority of the actuators investigated here (DEAs 2,3,4 and 5) the first order output spectra increases in magnitude over time. This phenomena seems to be independent of how the operating point shifts. It can also be be noted that the magnitude of the third order spectra appears to increase as the operating point decreases (DEAs 4,5 and the latter section of the data record for DEA 1). The inverse of this is observed

in DEAs 2 and 3, such that the magnitude of the third order spectra decreases as the operating point increases. In order to form direct conclusions regarding how the frequency domain descriptions relate to true behaviour a much larger sample size would need to be investigated.

5.3 Discussion

The aim of this chapter was to perform TV non-linear analysis on the set of DEA actuators in order to gain further insight into the dynamic characteristics of these actuators, particularly in respect to the inherent TV effects displayed. Very few frequency domain analysis methods exist for non-linear systems [14]. In the case where the system in question exhibits TV properties, to the best of the author's knowledge, only one method is available in the literature [49]. The lack of available methods motivated the development of TV-NOFRF analysis in this work. The novel method has significant advantages for both analysis, simplicity in implementation and computational efficiency.

The concept of the TV-NOFRF method is to consider a TV model of a non-linear system as N time invariant models over which time invariant analysis techniques can be performed. The method is simplified by working with linear in the parameters models such as those of the NARX type. This allows recursive parameter estimation methods such as Kalman filtering and RLS to be applied to efficiently track parameter changes across time, assuming a fixed model structure.

It is demonstrated that the NARMAX model structure identified in the previous chapter provides a good description of the system across time. Time varying parameters are calculated using a Kalman filter and subsequently smoothed. Applying the TV-NOFRF analysis procedure to the set of DEAs reveals TV dynamic behaviour that cannot be seen with traditional methods.

It was noted in Chapter 2.3.2 that DEA 1 shows behaviour at low time that is significantly different to the other actuators. The identification of the actuator was performed over a different region of the input output data. Observing the TV output spectra for DEA 1 for time $t \approx 0 - 300s$ the third order output is attenuated at low frequencies. The model has a fixed structure and so will behave poorly in the presence of un-modelled system dynamics. It is therefore likely that the frequency domain description is not an accurate representation of the true system in this region. This demonstrates a potential application of the TV-NOFRF method for fault detection in time series data. NOFRFs have already been proposed as a method for performing fault detection [94, 138]. Using the TV-NOFRF method it is suggested that faults could be detected in time series data by tracking the change

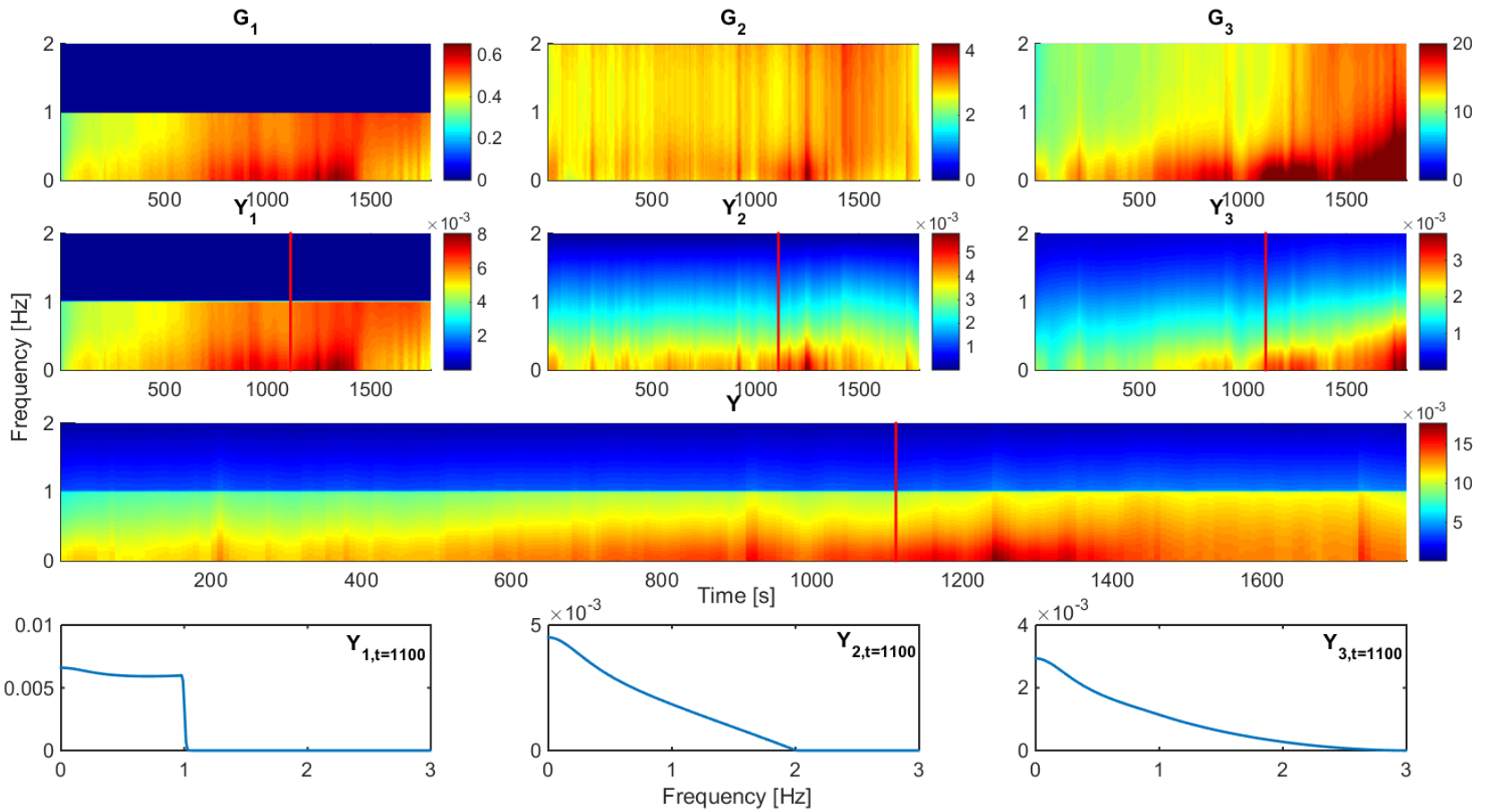

Figure 5.8: **TV-NOFRF analysis of DEA 5 reveals how the non-linear behaviour of the actuator changes significantly over time.** Top: n'th order NOFRFs, G_n and output spectra, Y_n for $n = 1,\ldots,3$ across all time in response to a 1Hz band limited input spectrum. Bottom: The Output spectrum across all time and the n'th order output spectra Y_n for $n = 1,\ldots,3$ at time $t = 1100s$.

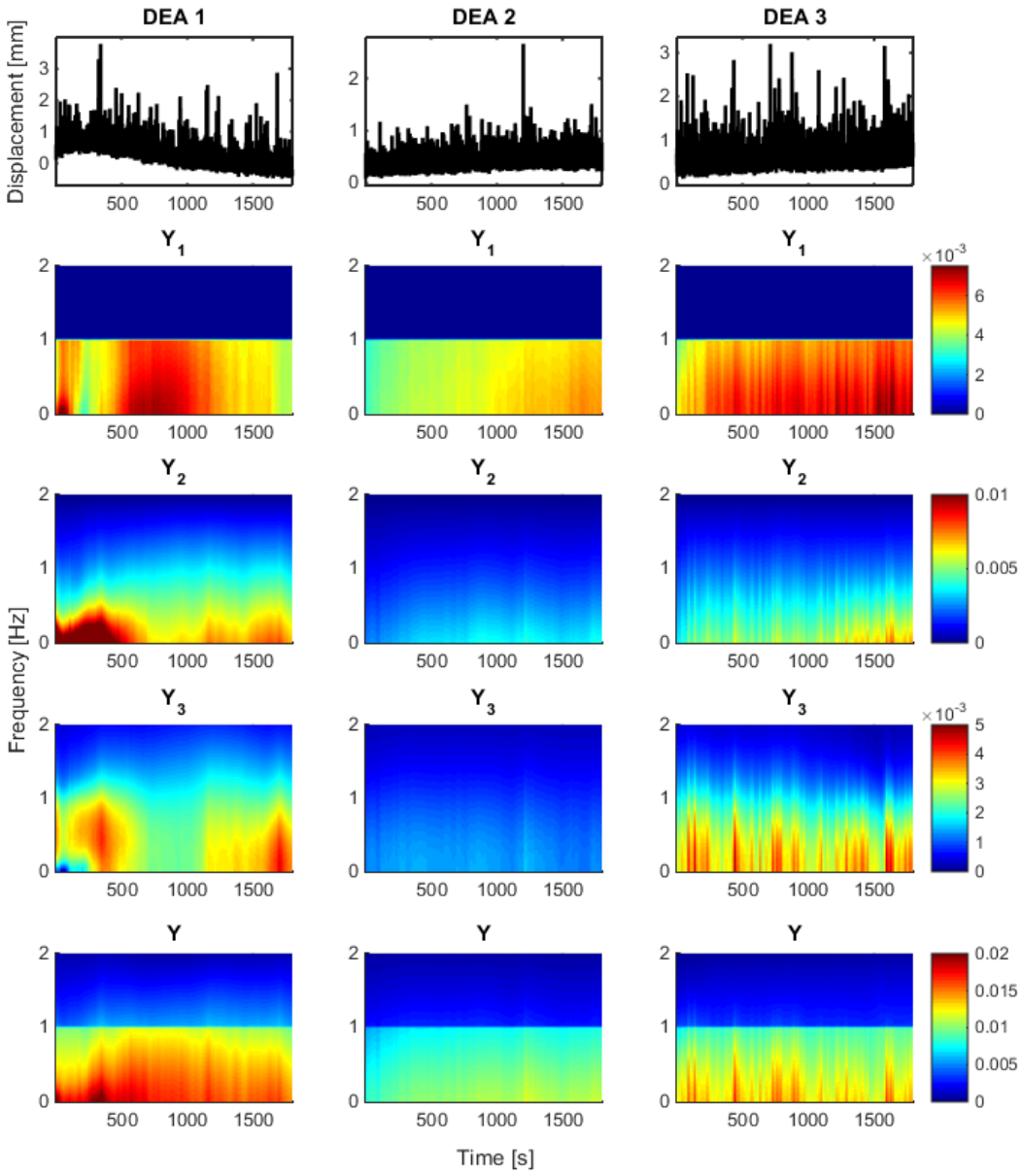

Figure 5.9: **TV-NOFRF analysis of DEAs show both significant and inconsistent time varying dynamics that cannot be observed in the time domain.** n'th order output spectra, Y_n for $n = 1, \ldots, 3$ across all time in response to a 1Hz band limited input spectrum for DEA 1-3.

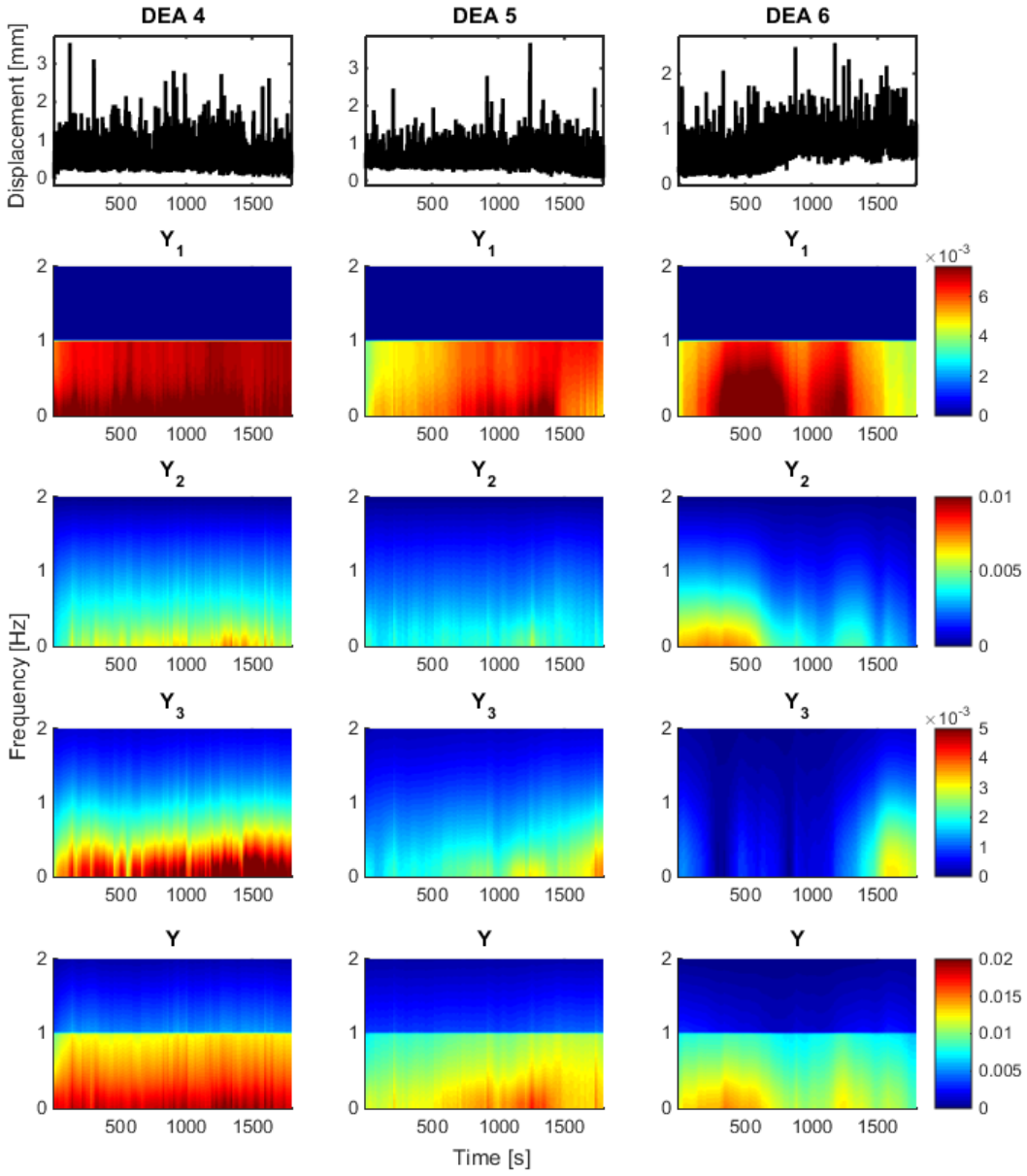

Figure 5.10: **TV-NOFRF analysis of DEAs show both significant and inconsistent time varying dynamics that cannot be observed in the time domain.** n'th order output spectra, Y_n for $n = 1, \ldots, 3$ across all time in response to a 1Hz band limited input spectrum for DEA 4-6.

in the NOFRFs for significant divergence from the model estimated for the healthy system.

The analysis over the set of DEAs shows that the TV effects significantly change the dynamic characteristic of the actuators across the data record and that the effects are manifested differently for each actuator, see Figures 5.9 and 5.10. This change in the underlying system behaviour represents a barrier to the implementation of traditional control architectures. Some connections between the frequency domain descriptions and the behaviour of the system in the time domain are made but it is suggested that a larger set of actuators need to be investigated before conclusions can be drawn.

The frequency domain analysis methods introduced so far in this thesis suggest that significant variation is shown in the underlying system behaviour of DEAs both across time as well as across different actuators. However, without quantifying the uncertainty in the identified models it is not possible to show if the exhibited behaviour is truly different or if it is a result of the system behaviour not being accurately captured by the model. This motivates the investigation into model uncertainty that is presented in the next chapter.

5.4 Summary

In this chapter a novel method for TV frequency domain analysis is introduced. The representation of TV systems within the NARMAX modelling framework is first discussed. TV-GFRFs are then introduced and methods for estimating them are discussed. The TV-GFRF method is extended to the novel TV-NOFRF analysis method which have significant advantages for the purpose of TV analysis, particularly in the simplicity of calculation.

The TV-NOFRF analysis method is then applied to the set of DEAs in order to investigate how underlying system behaviour changes over time. The results demonstrate that the actuators display inconsistencies in their behaviour which vary greatly over the data record and between actuators. In order to quantify how accurately the model describes the system, and hence how meaningful the comparisons between actuators is, it is suggested that a characterisation of the uncertainty in the model is required. In the following chapter, uncertainty is incorporated into the modelling process by taking a Bayesian approach.

Chapter 6

NARX Modelling Within a Bayesian Framework

The comparison of the behaviour of DEAs in Chapters 5 and 4 highlights the need to assess the confidence that is placed on the system model. To perform a meaningful comparison it is necessary to demonstrate that the observed difference in actuator behaviour is due to a true difference in the underlying dynamics rather than resulting from a poor system model. In order to take this approach, the uncertainty in the identified model needs to be characterised. This objective can be achieved by adopting the Bayesian philosophy in which model uncertainty can be naturally incorporated into the modelling process.

In this chapter two main contributions are made. Firstly the development of a novel Bayesian system identification algorithm for NARX models. Secondly, the application of this algorithm to the set of DEA actuators in order to investigate the uncertainty associated with the identified models. One of the key aims of this thesis is to develop methods that can characterise the uncertainty in systems, both in the modelling process and in the identified model. The identification methods that have been discussed so far in this thesis are all based on the frequentist approach; they provide the 'best' model of the system given some modelling parameters based on the minimisation of a prediction error. Although they can be successful in accurately modelling system behaviour they also possess major disadvantages in their inability to characterise the uncertainty associated with the system and that the model selection is performed by a subjective choice of ERR/SRR threshold.

The focus of this chapter is therefore the joint structure detection and parameter estimation of non-linear time series systems within a Bayesian framework. Bayesian methods offer a different perspective on the system identification prob-

lem. The key advantages of Bayesian modelling were discussed in Section 3.8. Two of the main points that are most relevant are emphasised here for completeness. Firstly, the Bayesian paradigm naturally incorporates uncertainty into the modelling procedure. Secondly, Bayesian model selection embodies Occam's razor, model complexity is inherently penalised and there is therefore a preference for simple models [78].

A problem with Bayesian modelling is the complexity involved in the inference task for all but simple problems. It is often the case that the integrals involved in the inference step are intractable. For this reason there has been much emphasis placed on approximate inference techniques, commonly based on MCMC sampling [42]. More recently Variational Bayesian (VB) inference has gained in popularity in the machine learning community [11].

Despite the recent popularity of Bayesian methods in both the machine learning and system identification communities there is little precedent for the identification of NARX models within a Bayesian framework. To the author's knowledge, the only instance of a Bayesian approach to parametric modelling of NARX models appears in [8]. The authors of that paper use a Markov chain Monte Carlo (MCMC) sampling method in order to numerically obtain posterior distributions of both the model structure and parameters. However, sampling methods are often computationally intensive to implement because they tend to rely on large numbers of samples to accurately estimate distributions [87].

There is therefore a need for computationally efficient Bayesian system identification techniques capable of producing parsimonious models of non-linear systems. In this chapter a novel Bayesian structure detection algorithm, named Sparse Variational Bayesian (SVB)-NARX, is presented which is capable of identifying accurate and compact NARX models, which is both simple to implement and relatively computationally efficient. The SVB-NARX identification algorithm uses variational Bayesian inference to perform parameter estimation, resulting in a sequence of closed-form equations in an iterative algorithm [11]. Structure detection is driven by the inclusion of a sparsity inducing hyper prior (i.e. one that encourages parameter values to go to zero), referred to as automatic relevance determination (ARD), which is used as a metric to prune redundant terms from the model [85]. To the author's knowledge ARD has not been used in the context of NARX models before and so represents a further novel contribution of this work.

The remainder of the chapter is structured as follows: Firstly in Section 6.1 the concept of ARD is introduced as a method for sparse parameter estimation. In Section 6.2, a Bayesian approach to NARX modelling is introduced and the model is defined. A discussion of Variational Bayesian inference follows in Section 6.3

and the update equations for Bayesian NARX models are derived. In Section 6.4 the SVB-NARX algorithm algorithm is introduced. Finally, results for the identification of both a synthetic and DEA system are presented in Sections 6.5 and 6.6.

6.1 Automatic relevance determination

ARD was formulated by [85] for detecting irrelevant inputs when there is a large number of inputs available for the training of neural networks [78]. In this work ARD is used as a method to automatically detect which model terms are not relevant to the generation of the system output data in order to prune these terms from the model. To the best of the author's knowledge ARD has not been applied in the context of NARX models and so its application in this work represents a novel contribution.

Considering a regression model defined by a super-set of basis functions, many of which are not relevant to the prediction of the output variable, standard parameter estimation techniques are not able to drive the parameters of these basis functions to zero given a finite data limit. The model is hence over-fitting to the irrelevant basis functions and will negatively influence model behaviour when predicting for new, unseen, data.

The concept of ARD is to introduce a prior distribution over the regression parameters which is able to infer from the data which model parameters are irrelevant to the data and to weight the parameters appropriately. This can be achieved by introducing separate variables for each basis function, that act as an independent regularisor to each term. The regularisors are estimated from the data as part of the inference. Model terms that are not relevant to the generation of the output are assigned a large regularisation variable indicating a small weight for the term [78].

It is common to perform ARD by optimisation of the hyper-parameters, by for example type-II maximum likelihood [24, 85, 123]. In this work the full Bayesian treatment is considered where posteriors over the hyper-parameter distributions are inferred.

6.2 System identification within a Bayesian framework

Bayes' rule allows us to infer the posterior distribution of the set of model parameters and hyper-parameters, denoted Θ, given the observed data set $\boldsymbol{y} = [y_1, \ldots, y_N]^T$ such that

$$p(\Theta|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\Theta)p(\Theta)}{p(\boldsymbol{y})}. \tag{6.1}$$

In the above equation the term $p(\boldsymbol{y}|\Theta)$ is referred to as the likelihood function and $p(\Theta)$ is the prior distribution of the parameters before observing the data. The denominator, $p(\boldsymbol{y})$ is named the marginal likelihood and is given as

$$p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\Theta)p(\Theta)d\Theta. \tag{6.2}$$

in order to normalise the posterior distribution.

As discussed in detail in Section 3.8, the evaluation of Equation (6.2) can be extremely challenging and necessitates the need for approximation methods in all but the most simple of cases. In this chapter, Equation (6.2) will be approximated by the use of variational Bayesian techniques.

6.2.1 Full Bayesian modelling of polynomial NARX models

In this section Bayesian inference of linear-in-the-parameters regression models is introduced. It is considered in the context of the polynomial NARX model given by Equation (3.10), which is repeated here for clarity of reading. As before we consider a single-input single-output dynamic system as some non-linear function, $f(.)$, of lagged system inputs, u_k, and outputs, y_k,

$$y_k = f(\mathbf{x}_k) + e_k \tag{6.3}$$

where $\mathbf{x}_k = \left(y_{k-1}, ..., y_{k-n_y}, u_{k-1}, ..., u_{k-n_u}\right)$ and e_k is a zero-mean normally distributed white noise process. n_u and n_y are the maximum lags, or dynamic orders, of the input and output respectively. The non-linear function $f(.)$ can be decomposed into a sum of weighted basis functions, which is a linear-in-the-parameters model,

$$f(\mathbf{x}_k) = \sum_{m=1}^{M} \theta_j \phi_m(\mathbf{x}_k) \tag{6.4}$$

$$= \Phi\boldsymbol{\theta} \tag{6.5}$$

where

$$\Phi = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_M], \tag{6.6}$$

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_M]^T. \tag{6.7}$$

and the composition of Φ will be determined by the algorithm presented later in

this chapter.

Likelihood function

Considering Equation (6.1) with $\Theta = \{\boldsymbol{\theta}, \tau\}$, where τ is named the precision and is the inverse of the variance of the error τ^{-1}, the likelihood function for the data, $\boldsymbol{y}$, can be written

$$\begin{aligned} p(y_k|\boldsymbol{\phi}_k\,\boldsymbol{\theta}, \tau) &= \mathcal{N}(y_k|\boldsymbol{\phi}_k\boldsymbol{\theta}, \tau^{-1}) && (6.8)\\ &= \left(\frac{\tau}{2\pi}\right)^{1/2} \exp(-\frac{\tau}{2}(y_k - \boldsymbol{\phi}_k\boldsymbol{\theta})^2), && (6.9)\end{aligned}$$

where $\mathcal{N}$ denotes the normal distribution with mean, $\boldsymbol{\phi}_k\boldsymbol{\theta}$, and variance, τ^{-1}. Under the assumption of a Normal *i.i.d.* noise sequence the data likelihood is hence given as

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{\phi}, \boldsymbol{\theta}, \tau) &= \prod_k p(y_k|\boldsymbol{\phi}_k, \boldsymbol{\theta}, \tau) && (6.10)\\ &= \prod_k \mathcal{N}(y_k|\boldsymbol{\phi}_k\boldsymbol{\theta}, \tau^{-1}) && (6.11)\\ &= \left(\frac{\tau}{2\pi}\right)^{N/2} \exp\left(-\frac{\tau}{2}\sum_k (y_k - \boldsymbol{\phi}_k\boldsymbol{\theta})^2\right) && (6.12)\end{aligned}$$

Priors

The use of conjugate priors in Bayesian modelling was discussed in Section 3.8.1. The likelihood function given by Equation (6.12) is a member of the exponential family and so the choice of an exponential prior is required for conjugacy (that the posterior distribution is in the same form as the prior).

The mean, $\boldsymbol{\phi}_k\boldsymbol{\theta}$, and precision, τ, of the likelihood are unknown parameters to be inferred and as such the choice of Normal-Gamma prior distribution is made [45] such that

$$\begin{aligned} p(\boldsymbol{\theta}, \tau|\boldsymbol{\alpha}) &= p(\boldsymbol{\theta}|\tau, \boldsymbol{\alpha})p(\tau) && (6.13)\\ &= \mathcal{N}(\boldsymbol{\theta}|0, (\tau\boldsymbol{A})^{-1})\mathrm{Gam}(\tau|a_0, b_0) && (6.14)\\ &= \left(\frac{|\boldsymbol{A}|}{2\pi}\right)^{M/2} \frac{b_0^{a_0}}{\Gamma(a_0)}\tau^{M/2+a_0-1}\exp(-\frac{\tau}{2}(\boldsymbol{\theta}\boldsymbol{A}\boldsymbol{\theta}^T + 2b_0)) && (6.15)\end{aligned}$$

where the Normal distribution has been further parametrised by $\boldsymbol{\alpha}$ where $\mathrm{diag}(\boldsymbol{A}) = \boldsymbol{\alpha}$ and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_M)^T$. Note that each α_i is independent. The introduction of $\boldsymbol{\alpha}$ into the model naturally incorporates ARD, this is the basis of the sparse estimation framework that will be discussed later. The variable $\boldsymbol{\alpha}$ is an unknown

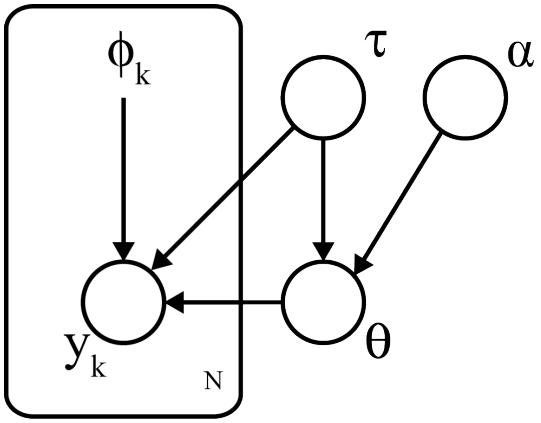

Figure 6.1: **Probabilistic graphical model of the hierarchical model represented in Equation (6.18)** The plate (box), denoted by the number of data samples N, indicates N *i.i.d* observations. The arrows indicate the direction of conditional dependence.

model parameter and will also be inferred and so requires the introduction of a hyper-prior (prior of a prior), $p(\boldsymbol{\alpha})$. The choice of a conjugate prior is again chosen in order to simplify the later analysis. The hyper-prior is therefore assigned a Gamma distribution,

$$p(\boldsymbol{\alpha}) = \prod_m \text{Gam}(\alpha_m|c_0, d_0) \tag{6.16}$$

$$= \prod_m \frac{d_0^{c_0}}{\Gamma(c_0)} \alpha_m^{c_0-1} \exp(-d_0\alpha_m) \tag{6.17}$$

Collecting all the model parameters produces the set $\Theta = \{\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}\}$ which have hyper-parameters a_0, b_0, c_0, d_0 which can be initialised so as to have broad/uninformative prior distributions so that the inference process is dominated by the influence of the data.

The joint distribution over all of the random variables can now be expressed hierarchically as

$$p(\boldsymbol{y}, \boldsymbol{\phi}, \boldsymbol{\theta}, \tau, \boldsymbol{\alpha}) = p(\boldsymbol{y}|\boldsymbol{\phi}, \boldsymbol{\theta}, \tau)p(\boldsymbol{\theta}|\tau, \boldsymbol{\alpha})p(\tau)p(\boldsymbol{\alpha}), \tag{6.18}$$

assuming τ and $\boldsymbol{\alpha}$ independent and noting that $\boldsymbol{\phi}$ is a function of the observed data and not a random variable. The decomposition can by made more transparent by considering the directed graphical model shown in Figure 6.1.

ARD Prior

ARD has been incorporated into the Bayesian model via the introduction of the hyper-parameter $\boldsymbol{\alpha}$ in Equation (6.13), α_m corresponds to the precision (inverse of the variance) of θ_m. α_m therefore controls the magnitude of θ_m, if $\alpha_m^{-1} = 0$ then the precision of θ_m is infinite and in order to maintain a high likelihood $\theta_m = 0$, indicating that the m'th model term is not relevant to the generation of the data. α_m is hence acting as a sparse regularisation term independently on each model weight $\boldsymbol{\theta}_m$. The values of α_i can then be used as a basis for pruning irrelevant basis functions from the model.

In the remainder of this chapter the ARD value corresponding to the m'th model term will be defined as α_m^{-1} such that small values indicate terms that are not relevant to the generation of the output. This value will be used to drive the structure detection in the SVB-NARX system identification algorithm introduced in Section 6.4.

Posterior distribution

The joint posterior distribution over the model parameters $\boldsymbol{\theta}$ and τ can be found by considering Equation (6.1), with $\Theta = \{\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}\}$, such that

$$p(\boldsymbol{\theta}, \tau|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\Phi, \boldsymbol{\theta}, \tau)p(\boldsymbol{\theta}|\tau, \boldsymbol{\alpha})p(\tau)p(\boldsymbol{\alpha})}{p(\boldsymbol{y})}. \tag{6.19}$$

The inclusion of the hyper-parameter, $\boldsymbol{\alpha}$, into the model causes the marginal likelihood in the denominator of Equation (6.19) to become intractable, *i.e.* no direct analytical solution is possible. Many methods exist for approximating the marginal likelihood, commonly these techniques are based on random sampling. Here, variational Bayesian inference will be used because the posterior distribution can be approximated in a series of closed form update equations, avoiding the use of computationally expensive sampling methods.

6.3 Variational Bayesian inference

In many non-trivial cases the evaluation of posterior distributions is infeasible as is the case with the linear regression with ARD introduced in the previous section. In fact, the full Bayesian treatment in closed form is only feasible for a limited class of models. Numerical integration techniques can always be used, however, the computational expense is often prohibitive. Variational Bayes' provides a method for approximating the posterior distribution. In this section some general results

of variational Bayesian inference are discussed followed by its application to the linear regression model given by Equation (6.19).

6.3.1 Variational optimisation of the Bayesian model

In the context of inference problems variational calculus can be used as a method for approximating posterior distributions. Consider a fully Bayesian model such that all the model parameters are assigned prior distributions as described in Section 6.2.1. By Bayes' theorem

$$p(\boldsymbol{\Theta}|\boldsymbol{y},\boldsymbol{x}) = \frac{P(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{p(\boldsymbol{y})}. \tag{6.20}$$

The aim of the inference problem is to find an approximation to the posterior distribution $p(\boldsymbol{\Theta}|\boldsymbol{y},\boldsymbol{x})$ assuming that the marginal distribution, $p(\boldsymbol{y})$, is intractable. For brevity of notation the dependency on the past data, $\boldsymbol{x}$, will be assumed implicit and dropped from the following discussion. As before, the marginal distribution is defined as

$$p(\boldsymbol{y}) = \int_{\boldsymbol{\Theta}} p(\boldsymbol{y}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})\, \mathrm{d}\boldsymbol{\Theta}. \tag{6.21}$$

Introducing any variational distribution, $Q(\boldsymbol{\Theta})$, to approximate $p(\boldsymbol{\Theta}|\boldsymbol{y})$ produces a lower bound on $p(\boldsymbol{y})$. In the case that this distribution is equal to the posterior, $p(\boldsymbol{\Theta}|\boldsymbol{y})$, then the bound is exact, however it is here assumed that working with the true posterior is intractable and as such a simpler form for the variational distribution must be used. This is achieved by first taking the natural logarithm of (6.21) and introducing $Q(\boldsymbol{\Theta})$ such that

$$\begin{aligned}
\ln p(\boldsymbol{y}) &= \ln \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \frac{p(\boldsymbol{y}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{Q(\boldsymbol{\Theta})}\, \mathrm{d}\boldsymbol{\Theta}. && (6.22)\\
&\geq \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{y}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{Q(\boldsymbol{\Theta})}\, \mathrm{d}\boldsymbol{\Theta}. && (6.23)\\
&= \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln p(\boldsymbol{y}|\boldsymbol{\Theta}) + Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{\Theta})}{Q(\boldsymbol{\Theta})}\, \mathrm{d}\boldsymbol{\Theta}. && (6.24)\\
&= \mathcal{L}[Q(\boldsymbol{\Theta})] && (6.25)
\end{aligned}$$

where (6.23) follows from Jenson's inequality which is applicable because the natural logarithm is a concave monotonically increasing function. The lower bound on the log marginal distribution $\ln p(\boldsymbol{y})$ is hence given by $\mathcal{L}[Q(\boldsymbol{\Theta})]$. By maximising this quantity an approximation of $\ln p(\boldsymbol{y})$ is acquired. Variational optimisation is employed to perform the variational maximization of $\mathcal{L}[Q(\boldsymbol{\Theta})]$ with

respect to the free distribution $Q(\boldsymbol{\Theta})$.

By noting that $Q(\boldsymbol{\Theta})$ is a proper probability distribution and therefore its integral with respect to $\boldsymbol{\Theta}$ is equal to unity and rearranging (6.23) - (6.25),

$$\begin{aligned}
\mathcal{L}[Q(\boldsymbol{\Theta})] &= \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{y}|\boldsymbol{\Theta})p(\boldsymbol{\Theta})}{Q(\boldsymbol{\Theta})} \, \mathrm{d}\boldsymbol{\Theta}. && (6.26)\\
&= \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{\Theta}|\boldsymbol{y})}{Q(\boldsymbol{\Theta})} + Q(\boldsymbol{\Theta}) \ln p(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{\Theta}. && (6.27)\\
\ln p(\boldsymbol{y}) &= \mathcal{L}[Q(\boldsymbol{\Theta})] - \int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{\Theta}|\boldsymbol{y})}{Q(\boldsymbol{\Theta})} \, \mathrm{d}\boldsymbol{\Theta} && (6.28)\\
&= \mathcal{L}[Q(\boldsymbol{\Theta})] + KL(Q(\boldsymbol{\Theta})||p(\boldsymbol{\Theta}|\boldsymbol{y})) && (6.29)
\end{aligned}$$

where

$$KL(Q(\boldsymbol{\Theta})||p(\boldsymbol{\Theta}|\boldsymbol{y})) = -\int_{\boldsymbol{\Theta}} Q(\boldsymbol{\Theta}) \ln \frac{p(\boldsymbol{\Theta}|\boldsymbol{y})}{Q(\boldsymbol{\Theta})} \, \mathrm{d}\boldsymbol{\Theta} \tag{6.30}$$

is the Kullback-Leibler (KL) divergence from Q to p. A few things of interest can be noted from Equation (6.28). First, the log marginal distribution, $p(\boldsymbol{y})$, can be de-constructed into a lower bound on the distribution and a measure of how the difference between the true posterior distribution, $p(\boldsymbol{\Theta}|\boldsymbol{y})$, and its approximating distribution, $Q(\boldsymbol{\Theta})$, in the form of the KL divergence. It is then evident that the optimal approximation to the marginal distribution is found by maximising the lower bound, or equivalently minimising the KL divergence. It is also clear that the choice of $Q(\boldsymbol{\Theta}) = p(\boldsymbol{\Theta}|\boldsymbol{y})$ would result in an exact match between the bound and the marginal distribution.

6.3.2 Factorised distributions

In order to make the variational optimisation of $\mathcal{L}[Q(\boldsymbol{\Theta})]$ feasible, the family of possible distributions, $Q(\boldsymbol{\Theta})$, over which the optimisation is performed, must be restricted. The assumption is made that the variational distribution can be factorised such that

$$Q(\boldsymbol{\Theta}) = \prod_j q_j(\Theta_j) \tag{6.31}$$

where each $q_j(\Theta_j)$ are independent. This approximation is known as the mean field theory in physics. The function can now be maximised with respect to each distribution $q_j(\Theta_j)$ separately while holding all others fixed.

Substituting the factorised distribution (6.31) into Equation (6.24) and then

separating out a singular distribution over which to perform the optimisation

$$\begin{aligned}\mathcal{L}[Q(\boldsymbol{\Theta}))] &= \int \prod_j q_j(\Theta_j) \left(\ln p(\boldsymbol{y}, \boldsymbol{\Theta}) - \sum_i \ln q_i(\Theta_i) \right) \mathrm{d}\boldsymbol{\Theta} \qquad (6.32)\\ &= \int q_i(\Theta_i) \left(\int \ln p(\boldsymbol{y}, \boldsymbol{\Theta}) \prod_{i \neq j} q_j(\Theta_j) \mathrm{d}\Theta_j \right) \mathrm{d}\Theta_j \\ &\quad - \int q_i(\Theta_i) \ln q_i(\Theta_i) \mathrm{d}\Theta_j + \text{const} = 0 \qquad (6.33)\end{aligned}$$

where the terms $\prod_{i \neq j} q_i(\Theta_i) \ln q_i(\Theta_i)$ have been absorbed into a constant term.

In order to find the distribution $q_i(\Theta_i)$ that maximises $\mathcal{L}[Q(\boldsymbol{\Theta})]$ a variational optimisation is performed with respect to $q_i(\Theta_i)$. The optimisation is performed using variational calculus by noting that Equation (6.33) is in the appropriate form to directly apply the fundamental lemma of the calculus of variations [41] such that

$$\begin{aligned}\frac{\delta \mathcal{L}[q_i(\Theta_i)]}{\delta q_i(\Theta_i))} &= \frac{\partial}{\partial q_i(\Theta_i)} \left(q_i(\Theta_i) \left(\int \ln p(\boldsymbol{y}, \boldsymbol{\Theta}) \prod_{i \neq j} q_j(\Theta_j) \mathrm{d}\Theta_j \right) \right. \\ &\quad \left. - q_i(\Theta_i) \ln q_i(\Theta_i) + \text{const} \right) \\ &= \int \ln p(\boldsymbol{y}, \boldsymbol{\Theta}) \prod_{i \neq j} q_j(\Theta_j) \mathrm{d}\Theta_j - \ln q_i(\Theta_i) + \text{const} \qquad (6.34)\end{aligned}$$

where the second term on the right hand side of Equation (6.34) has been found via the variational product rule. The integral that forms the first term on the right hand side of Equation (6.34) is the expectation of the log joint distribution where the expectation is taken with respect to all of the distributions $q(\Theta)$ for which $i \neq j$, such that

$$\mathbb{E}_{i \neq j}[\ln(\boldsymbol{y}, \boldsymbol{\Theta})] = \int \ln p(\boldsymbol{y}, \boldsymbol{\Theta}) \prod_{i \neq j} q_j(\Theta_j) \mathrm{d}\Theta_j. \qquad (6.35)$$

where $\mathbb{E}_{i \neq j}$ denotes the expectation with respect to the distributions q over all the variables in the set Θ for which $i \neq j$.

Substituting (6.35) into (6.34), equating to zero and rearranging for the variational distribution, a general expression for the optimal solution and therefore the update of the jth factor of the variational distribution $q_j^{(t+1)}(\Theta_j)$ is then given by

$$\ln q_j^{(t+1)}(\Theta_j) = \mathbb{E}_{i \neq j}[\ln(\boldsymbol{y}, \boldsymbol{\Theta})] + \text{const.} \qquad (6.36)$$

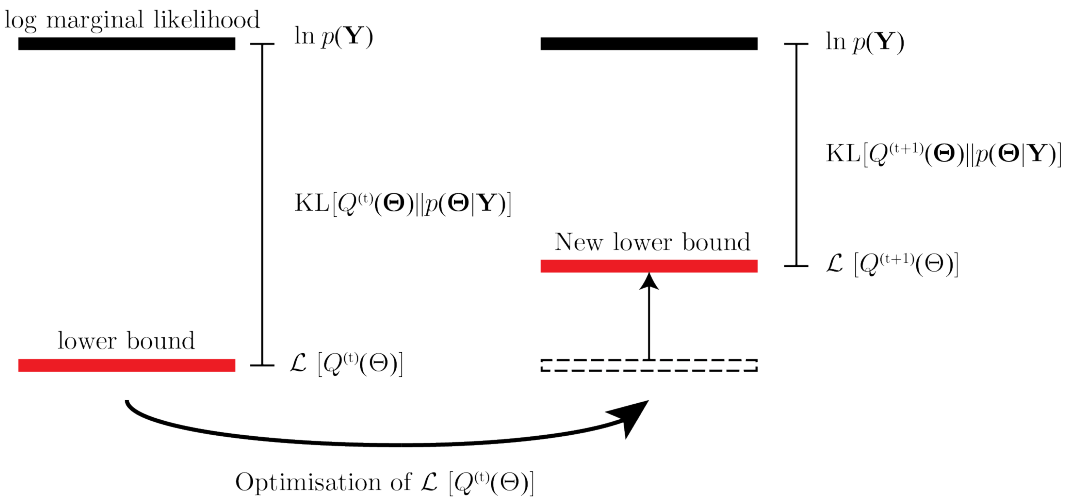

Figure 6.2: Variational inference is performed by iteratively updating the variational lower bound via an optimisation step. The variational Bayesian update according to Equation (6.36)

Performing the update (6.36) for each factor $q_j(\Theta_j)$ of the variational distribution completes one optimisation step of $\mathcal{L}[Q(\Theta)]$. The optimisation is performed iteratively where t indicates the current iteration such that $q_j^{(t+1)}$ is the update using the statistics of the distributions $q_j^{(t)}$ at the previous iteration, see Figure 6.2.

Equation (6.36) states that the log of the optimal solution for the ith factor of the variational distribution, $q_i^{(t+1)}(\Theta_i)$, is given by the expectation of the log of the joint distribution over the data and model parameters, where the expectation is taken with respect to all of the other factors q_i for $i \neq j$. Therefore the maximum of the lower bound, given the factorisation approximation (6.31), can be found by cycling through the set of equations defined by Equation (6.36) for $i = 1, \ldots, N_i$ where N_i is the total number of factors of $Q(\Theta)$. After initialising $q_i^{(t+1)}(\Theta_i)$ for all i, each computation of (6.36) is taken with respect to the revised parameter estimates from the previous calculations. The lower bound is guaranteed to converge because it is convex with respect to each of the factors of $q_i^{(t+1)}(\Theta_i)$.

In order to facilitate computation the variational lower bound can be written more explicitly as

$$\begin{aligned}
\mathcal{L}[Q(\mathbf{\Theta})] &= \int_{\mathbf{\Theta}} Q(\mathbf{\Theta}) \ln \frac{p(\boldsymbol{y}, \mathbf{\Theta})}{Q(\mathbf{\Theta})} \mathrm{d}\mathbf{\Theta} && (6.37)\\
&= \int Q(\mathbf{\Theta}) \ln p(\boldsymbol{y}, \mathbf{\Theta}) \mathrm{d}\mathbf{\Theta} - \int Q(\mathbf{\Theta}) \ln Q(\mathbf{\Theta}) \mathrm{d}\mathbf{\Theta} && (6.38)\\
&= \mathbb{E}[\ln p(\boldsymbol{y}, \mathbf{\Theta})] - \mathbb{E}[\ln Q(\mathbf{\Theta})] && (6.39)
\end{aligned}$$

where the last step follows from the definition of expectations [24]. Although the calculation of the variational lower bound is not required for the inference problem it provides a check that the algorithm, as well as the theory behind it, is functioning correctly as the bound is guaranteed to increase with each update of the variational distribution. Further than this it will be used later as a method for model comparison when the variational inference method is employed for the purpose of structure detection.

6.3.3 Variational inference of linear-in-the-parameters regression models

The variational Bayesian inference procedure can now be applied to the Bayesian linear regression model defined by Equation (6.4) following [34]. Applying the approximation defined by Equation (6.31), the assumption is made that the posterior distribution $p(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})$ can be approximated by a factorisation of the variational distribution, $Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})$, such that

$$Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}) = q(\boldsymbol{\theta}, \tau) q(\boldsymbol{\alpha}). \tag{6.40}$$

This factorisation allows us to perform a maximisation of the variational lower bound with respect to each factor. From Equation (6.23), the lower bound is given by

$$\mathcal{L}[Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})] = \iiint Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}) ln \frac{p(\boldsymbol{y}|\Phi, \boldsymbol{\theta}, \tau, \alpha) P(\boldsymbol{\theta}, \tau|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}))}{Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})} \mathrm{d}\boldsymbol{\theta} \mathrm{d}\tau \mathrm{d}\boldsymbol{\alpha}. \tag{6.41}$$

Using the results of the variational Bayesian inference derived above the optimisation of the bound is performed via Equation (6.36) for the distributions $q(\boldsymbol{\theta}, \tau)$ and $q(\boldsymbol{\alpha})$ resulting in a set of closed form update equations.

Update for $q(\boldsymbol{\theta}, \tau)$

The variational posterior $q(\boldsymbol{\theta}, \tau)$ is found by maximising the variational lower bound, $\mathcal{L}(Q)$, with fixed $q(\alpha)$. Using the general result for the update of the j'th distribution given by Equation (6.36) where $\Theta = \{\boldsymbol{\phi}, \boldsymbol{\theta}, \tau, \boldsymbol{\alpha}\}$, such that $p(\boldsymbol{y}, \Theta)$

is given by Equation (6.18), gives

$$\begin{aligned}\ln q(\boldsymbol{\theta},\tau) &= \ln p(\boldsymbol{y}|\Phi,\boldsymbol{\theta},\tau) + \mathbb{E}_{\boldsymbol{\alpha}}[\ln p(\boldsymbol{\theta},\tau|\boldsymbol{\alpha})] + \text{const} \qquad (6.42)\\ &= \left(\frac{M}{2} + a_0 - 1 + \frac{N}{2}\right)\ln(\tau)\\ &- \frac{\tau}{2}\left(\boldsymbol{\theta}^T\left(\mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}] + \sum_k \boldsymbol{\phi}_k^T\boldsymbol{\phi}\right)\boldsymbol{\theta}\right.\\ &+ \left.\sum_k y_k^2 - 2\sum_k \Phi_k y_k\boldsymbol{\theta}^T + 2b_0\right) + \text{const} \qquad (6.43)\end{aligned}$$

where $p(\boldsymbol{y}|\Phi,\boldsymbol{\theta},\tau)$ and $p(\boldsymbol{\theta},\tau|\boldsymbol{\alpha})$ are given by Equations (6.12) and (6.15) respectively and all the terms not dependent on $\boldsymbol{\theta}$ and τ have been absorbed into the constant term.

Given that the likelihood function, Equation (6.8), is in the form of a normal distribution, the conjugate normal-gamma prior, Equation (6.15), was chosen. The posterior distribution can hence be found analytically in the form

$$q_K(\boldsymbol{\theta},\tau) = \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_K,\tau^{-1}\boldsymbol{V}_N)\text{Gam}(\tau|a_K,b_K). \qquad (6.44)$$

The method of completing the square can be used to find the exponent of the normal distribution in the posterior. First, separating out all the coefficients of $-\frac{\tau}{2}\boldsymbol{\theta}^T\boldsymbol{\theta}$,

$$\begin{aligned}-\frac{\tau}{2}\boldsymbol{\theta}^T\boldsymbol{V}_K^{-1}\boldsymbol{\theta} &= -\frac{\tau}{2}\boldsymbol{\theta}^T\left(\sum_k \Phi_k^T\Phi_k + \mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}]\right)\boldsymbol{\theta} \qquad (6.45)\\ \boldsymbol{V}_K^{-1} &= \sum_k \Phi_k^T\Phi_k + \mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}]. \qquad (6.46)\end{aligned}$$

It follows that

$$\begin{aligned}-2\boldsymbol{\theta}^T\boldsymbol{V}_K^{-1}\boldsymbol{\theta}_K &= -2\boldsymbol{\theta}^T\sum_k \Phi_k^T y_k \qquad (6.47)\\ \boldsymbol{\theta}_K &= \boldsymbol{V}_K\sum_k \Phi_k^T y_k. \qquad (6.48)\end{aligned}$$

This leaves the remaining terms in Equation (6.43), as well as introducing the extra term $-\frac{\tau}{2}\boldsymbol{\theta}_k^T\boldsymbol{V}_K\boldsymbol{\theta}_k$ by completing the square, that have not been incorporated into

the normal such that

$$\begin{aligned}\ln q_K(\boldsymbol{\theta},\tau) &= \ln \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_K,\tau^{-1}\boldsymbol{V}_N)+\left(a_0-1+\frac{N}{2}\right)\ln(\tau)\\ &- \frac{\tau}{2}\left(\sum_k y_k^2+2b_0-\boldsymbol{\theta}_k^T\boldsymbol{V}_K^{-1}\boldsymbol{\theta}_k\right)+\text{const.}\end{aligned} \tag{6.49}$$

The remaining terms are then absorbed into the gamma distribution. Again, taking the exponent and equating terms

$$-\tau b_K = -\frac{\tau}{2}\left(\sum_k y_k^2+2b_0\right) \tag{6.50}$$

$$b_K = b_0+\frac{1}{2}\left(\sum_k y_k^2-\boldsymbol{\theta}_k^T\boldsymbol{V}_K^{-1}\boldsymbol{\theta}_k\right) \tag{6.51}$$

Similarly for a_K, equating powers of τ

$$(a_K-1)\ln\tau = \left(a_0-1+\frac{N}{2}\right)\ln\tau \tag{6.52}$$

$$a_K = a_0+\frac{N}{2} \tag{6.53}$$

The update of $\ln q(\boldsymbol{\theta},\boldsymbol{\tau})$ is then performed by the computation of equations (6.46), (6.48), (6.51) and (6.53).

Update for $q(\boldsymbol{\alpha})$

Similarly, the variational posterior $q(\boldsymbol{\alpha})$ is found by maximising the variational lower bound, $\mathcal{L}(Q)$, with fixed $q(\boldsymbol{\theta},\tau)$. Again, by Equation (6.36), and following the same process as for the maximisation of $q(\boldsymbol{\theta},\tau)$,

$$\begin{aligned}\ln q_K(\boldsymbol{\alpha}) &= \mathbb{E}_{\boldsymbol{\theta},\tau}(\ln p(\boldsymbol{\theta},\tau|\boldsymbol{\alpha}))+\ln(\boldsymbol{\alpha})+\text{const}\\ &= \prod_m \ln \text{Gam}(\alpha_m|c_K,d_{K_m}).\end{aligned} \tag{6.54}$$

where

$$c_K = c_0+\frac{1}{2} \tag{6.55}$$

$$d_{K_m} = d_0+\frac{1}{2}\mathbb{E}_{\boldsymbol{\theta},\tau}[\tau\theta_m^2]. \tag{6.56}$$

The update of $q(\boldsymbol{\alpha})$ is hence performed by the computation of (6.55) - (6.56).

The required expectations are found by considering the standard moments of the relevant distributions [24] such that

$$\mathbb{E}_{\boldsymbol{\theta},\tau}[\tau\theta_m^2] = \theta_{K_m}^2 \frac{a_K}{b_K} + \boldsymbol{V}_{K_{mm}} \tag{6.57}$$

and the required expectation for the update of $q(\boldsymbol{\theta}, \tau)$ is given by

$$\mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}] = \boldsymbol{A}_K \tag{6.58}$$

where $\boldsymbol{A}_K$ is a diagonal matrix with elements

$$\mathbb{E}_{\boldsymbol{\alpha}}[\alpha_m] = \frac{c_K}{d_{K_m}}. \tag{6.59}$$

After initialisation of the relevant variables, a_0, b_0, c_0 and d_0, the variational approximation, $Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}) = q(\boldsymbol{\theta}, \tau)q(\boldsymbol{\alpha})$, to the true posterior $p(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})$ can be found by iteratively updating $q(\boldsymbol{\theta}, \tau)$ and $q(\boldsymbol{\alpha})$ by the update equations defined by (6.46),(6.48),(6.51) and (6.53), and (6.55-6.56) respectively. These updates then form an iterative algorithm that is summarised in Algorithm 6.1.

Variational lower bound $\mathcal{L}$(Q)

The variational lower bound is found by considering Equation (6.39). The dependencies between the parameters in the first term in Equation (6.39) are easily determined by considering the joint distribution given by Equation (6.18) and the probabilistic graphical model in Figure 6.1. Expanding the second term follows from Equation (6.40) along with (6.44) and (6.54). The lower bound is then given by

$$\begin{aligned} \mathcal{L}[Q(\boldsymbol{\Theta})] &= \mathbb{E}[\ln p(\boldsymbol{y}, \boldsymbol{\Theta})] - \mathbb{E}[\ln Q(\boldsymbol{\Theta})] && (6.60) \\ &= \mathbb{E}_{\boldsymbol{\theta},\tau}[\ln(p(\boldsymbol{y}|\Phi, \boldsymbol{\theta}, \tau))] + \mathbb{E}_{\boldsymbol{\theta},\tau,\boldsymbol{\alpha}}[\ln p(\boldsymbol{\theta}, \tau|\boldsymbol{\alpha})] \\ &+ \mathbb{E}_{\boldsymbol{\alpha}}[\ln p(\boldsymbol{\alpha})] - \mathbb{E}_{\boldsymbol{\theta},\tau}[\ln q(\boldsymbol{\theta}, \tau)] - \mathbb{E}_{\boldsymbol{\alpha}}[\ln q(\boldsymbol{\alpha})] && (6.61) \end{aligned}$$

Taking the expectations of (6.12), (6.15), (6.17), (6.44) and (6.54) by again considering the moments of the relevant distributions;

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\theta},\tau} \ln p(\boldsymbol{y}|\Phi,\boldsymbol{\theta},\tau) &= \frac{N}{2}(\psi(a_K) - \ln b_K - \ln 2\pi) \\
&\quad -\frac{1}{2}\sum_k \left(\frac{a_K}{b_K}(y_k - \boldsymbol{\theta}^T\phi_k^T)^2 + \phi_k \boldsymbol{V}_K \phi_k^T\right), \qquad (6.62) \\
\mathbb{E}_{\boldsymbol{\theta},\tau,\boldsymbol{\alpha}} \ln p(\boldsymbol{\theta},\tau|\alpha) &= \frac{M}{2}(\psi(a_K) - \ln b_k + \psi(c_K) - \ln 2\pi) \\
&\quad -\frac{1}{2}\sum_m \left(\ln d_{K_m} + \frac{c_K}{d_{K_m}}\left(\frac{a_K}{b_K}\boldsymbol{\theta}^T\boldsymbol{\theta} + \boldsymbol{V}_{K_{mm}}\right)\right) \\
&\quad -\ln\Gamma(a_0) + a_0 \ln(b_0) \\
&\quad +(a_0 - 1)(\psi(a_K) - \ln b_K) - b_0\frac{a_K}{b_K}, \qquad (6.63) \\
\mathbb{E}_{\boldsymbol{\alpha}} \ln p(\boldsymbol{\alpha}) &= -M(\ln\Gamma(c_0) + c_0 \ln d_0) \\
&\quad +\sum_m \left((c_0 - 1)(\psi(c_K) - \ln d_{K_m}) - d_0\frac{c_K}{d_{K_m}}\right), \qquad (6.64) \\
\mathbb{E}_{\boldsymbol{\theta},\tau} \ln q_K(\boldsymbol{\theta},\tau) &= \frac{M}{2}(\psi(a_K) - \ln b_K - \ln 2\pi - 1) - \frac{1}{2}\ln|\boldsymbol{V}_K| \\
&\quad -\ln\Gamma(a_K) + a_K \ln b_K \\
&\quad +(a_K - 1)(\psi(a_K) - \ln b_K) - a_K \qquad (6.65) \\
\mathbb{E}_{\boldsymbol{\alpha}} \ln q_K(\boldsymbol{\alpha}) &= \sum_m((c_K - 1)(\psi(c_K) + \ln d_{K_m})) \\
&\quad -M(\ln\Gamma(c_K) - c_K). \qquad (6.66)
\end{aligned}$$

Substituting (6.62) - (6.66) into (6.61) provides an expression for the variational bound as

$$\begin{aligned}
\mathcal{L}(Q) &= -\frac{N}{2}\ln 2 - \frac{1}{2}\sum_k \left(\frac{a_K}{b_K}(y_k - \boldsymbol{\theta}_K^T\Phi_k^T)^2 + \Phi^T\boldsymbol{V}_K\Phi_K\right) + \frac{1}{2}\ln|\boldsymbol{V}_K| \\
&\quad -\ln\Gamma(a_0) + a_0\ln b_0 - b_0\frac{a_K}{b_K} + \ln\Gamma(a_K) - a_K\ln b_K + a_K \\
&\quad +M\left(\frac{1}{2} - \ln\Gamma(c_0) + c_0\ln d_0 + \ln\Gamma(c_K)\right) - \sum_m(c_K \ln d_{K_m}) \qquad (6.67)
\end{aligned}$$

The variational posterior distribution $Q(\boldsymbol{\theta},\tau,\boldsymbol{\alpha})$ can now be calculated by iteratively computing $q(\boldsymbol{\theta},\tau)$ and $q(\boldsymbol{\alpha})$ using the expectations of one distribution to calculate the other. At each iteration the variational lower bound, $\mathcal{L}(Q)$ can be computed via equation 6.67. The best approximation is found when $\mathcal{L}(Q)$ plateaus. The process is summarised in Algorithm 6.1.

Algorithm 6.1 Computation of variational posteriors

Initialise

$T_{\mathcal{L}_{(Q)}}, a_0, b_0, c_0, d_0$

Procedure

$t = 0$

while $\mathcal{L}(Q)_t - \mathcal{L}(Q)_{t-1} \leq T_{\mathcal{L}(Q)}$

Compute expectations $\mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}]$ *and* $\mathbb{E}_{\boldsymbol{\theta},\tau}[\tau\theta_m^2]$

$\mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}] = \boldsymbol{A}_K$

where $\boldsymbol{A}_K$ is a diagonal matrix with elements

$\mathbb{E}_{\boldsymbol{\alpha}}[\alpha_m] = \frac{c_K}{d_{K_m}}$

$\mathbb{E}_{\boldsymbol{\theta},\tau}[\tau\theta_m^2] = \theta_{K_m}^2 \frac{a_K}{b_K} + \boldsymbol{V}_{K_{mm}}$

Compute update equations for $q(\boldsymbol{\theta}, \tau)$

$\boldsymbol{V}_K = \sum_k \Phi_k^T \Phi_k + \mathbb{E}_{\boldsymbol{\alpha}}[\boldsymbol{A}]$

$\boldsymbol{\theta}_K = \boldsymbol{V}_K \sum_k \Phi_k^T y_k$

$a_K = a_0 + \frac{N}{2}$

$b_k = b_0 + \frac{1}{2}\left(\sum_k y_k^2 - \boldsymbol{\theta}_K^T \boldsymbol{V}_K^{-1} \boldsymbol{\theta}_K\right)$

Compute update equations for $q(\boldsymbol{\alpha})$

$c_K = c_0 + \frac{1}{2}$

$d_{K_m} = d_0 + \frac{1}{2}\mathbb{E}_{\boldsymbol{\theta},\tau}[\tau\theta_m^2]$

Compute variational lover bound $\mathcal{L}(Q)_t$ *via Equation (6.67)*

$t = t + 1$

end while

end Procedure

6.3.4 Predictive distribution

Predictions of a new, unseen, data point can be made by calculating a predictive distribution for the model at sample k. Given the training data, such that we can form Φ, the task is hence the evaluation of the distribution $p(y_{k+1}|\Phi)$ which can be achieved by again approximating the posterior $p(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})$ by the variational distribution $Q(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha}) = q(\boldsymbol{\theta}, \tau)q(\boldsymbol{\alpha})$. The predictive distribution can then be found by marginalising over the parameters such that

$$\begin{aligned} p(y_{k+1}|\Phi) &= \iiint p(\boldsymbol{y}|\Phi, \boldsymbol{\theta}, \tau)p(\boldsymbol{\theta}, \tau, \boldsymbol{\alpha})\mathrm{d}\boldsymbol{\theta}\mathrm{d}\tau\mathrm{d}\boldsymbol{\alpha} \\ &\approx \iiint p(\boldsymbol{y}|\Phi, \boldsymbol{\theta}, \tau)q(\boldsymbol{\theta}, \tau)q(\boldsymbol{\alpha})\mathrm{d}\boldsymbol{\theta}\mathrm{d}\tau\mathrm{d}\boldsymbol{\alpha} \\ &= \iiint \mathcal{N}(y_k|\boldsymbol{\phi}\boldsymbol{\theta}_K, \tau^{-1})\mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_K, \tau^{-1}\boldsymbol{V}_N)\mathrm{Gam}(\tau|a_K, b_K)\mathrm{d}\boldsymbol{\theta}\mathrm{d}\tau\mathrm{d}\boldsymbol{\alpha} \\ &= \mathrm{St}\left(y_k|\boldsymbol{\phi}\boldsymbol{\theta}_K, (1+\boldsymbol{\phi}^T\boldsymbol{V}_K\boldsymbol{\phi})^{-1}\frac{a_K}{b_K}, 2a_K\right) \end{aligned} \tag{6.68}$$

where the resulting distribution, denoted St, is a Student's t distribution. The distribution over $\boldsymbol{\alpha}$ does not appear in the third line of the above derivation because it is independent from the other distributions and hence it integrates to unity. In the final step, standard result from convolving conjugate distributions have been used [24]. The mean and variance of the distribution are given by

$$\mathbb{E}[y_k] = \boldsymbol{\phi}\boldsymbol{\theta}_K, \tag{6.69}$$

$$\mathrm{Var}[y_k] = (1+\boldsymbol{\phi}^T\mathbf{V}_K\boldsymbol{\phi})\frac{b_K}{(a_K-1)} \tag{6.70}$$

Example 6.1 ARD for sparse parameter estimation estimation

In this example, application of the variational Bayesian estimation framework with ARD to a NARX model is demonstrated. ARD is employed as a method for detecting redundant basis functions in a parametric NARX model. The following example demonstrates the influence of the sparsity inducing prior on the estimation of model coefficients as well as the calculated ARD values and their ability to detect the underlying model structure. Parameters are estimated by LS and regularised LS for comparison.

The following generative polynomial NARX model is simulated for $N = 1000$ samples

$$y_k = 0.3y_{k-1} + 0.1u_{k-1} + 0.4y_{k-1}y_{k-2} + e_k \tag{6.71}$$

where e_k is a white noise sequence drawn from the normal distribution $\mathcal{N}(e_k|0, \sigma^2)$,

where $\sigma^2 = 0.02$. Assuming no knowledge of the true model structure a superset of model terms is generated of maximum dynamic order $n_u = n_y = 2$ and maximum polynomial order $n_p = 2$ leading to a set containing 14 terms.

Parameters are estimated using the variational Bayesian algorithm with ARD by iteratively updating the variational posteriors using Algorithm 6.2 with initialisation $a_0 = c_0 = 1 \times 10^{-2}$ and $b_0 = d_0 = 1 \times 10^{-4}$. The parameters are also estimated with least squares and ridge regression as a comparison.

The true model terms are shown at index $m = 1, 3, 9$, see Figure 6.3**A**. The least squares estimate gives a high weighting to non model terms as might be expected, in particular large parameter weights are assigned to terms with index 5, 6 and 11. Ridge regression performs better causing some shrinkage of the redundant model parameters. However, weight is still assigned to incorrect terms. In comparison, parameters estimated within the variational framework with ARD very closely follow the true parameters, see Figure 6.3**A**. The three true model terms fall within 95% confidence intervals of the true value and all but one term (Term index 11) of the non model terms are within a 95% confidence interval of zero, see Figure 6.3**C**. The ARD values are inferred as part of the algorithm, these values indicate which terms are relevant to generating the data. The calculated ARD values are displayed by their natural log value, $\ln(ARD)$ to ease comparison as they can have a large range of values. The three correct model terms are assigned large ARD values by the algorithm. The model term at index 11 has a significant value in comparison to those assigned to the true model terms, see Figure 6.3**B**. Although the weight assigned to the parameter is very low, the ARD value is significant, incorrectly suggesting that the term at index 11 might be part of the true model. The example also serves to demonstrate the convergence of the variational lower bound, see Figure 6.3**D**.

The correct model structure could not be obtained by performing the inference with ARD alone. This then motivates the development of a Bayesian system identification algorithm which is the subject of the remainder of this chapter.

6.4 Sparse Bayesian identification of polynomial NARX models

The variational Bayesian inference procedure has provided a method for estimating the posterior distributions of parameters for models of the NARX form. Through the incorporation of ARD into the procedure a measure of how relevant each basis function is to the prediction of the data is also acquired. Importantly the

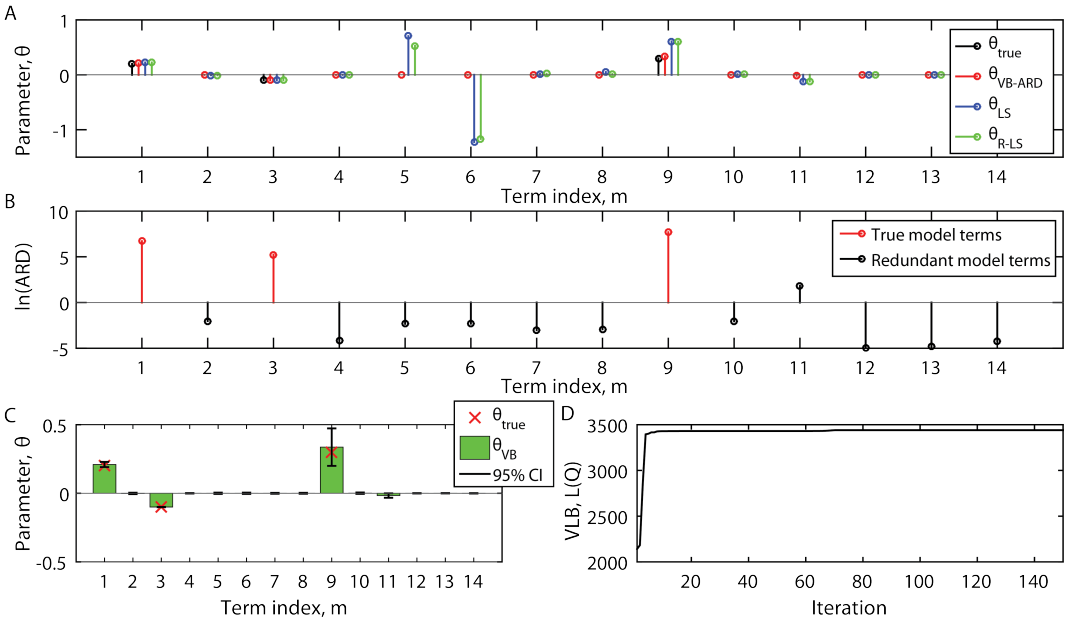

Figure 6.3: **Variational inference with ARD highly regularises non model terms to encourage sparsity in the prameter estimates. It is not able to reduce all non model parameters to zero.A**) Comparison of true and estimated model coefficients for the generative system given by Equation (6.71); True (Black), Variational inferance with ARD (Red), Least squares (Blue) and Regularised least squares (Green). **B**) ARD values corresponding to the parameter estimates. **C**) Variational Bayes' with ARD parameter estimates (Green bar) with 95% confidence intervals (Black) and true parameters (Red cross). **D**) Progression of the variational lower bound.

variational lower bound provides a measure of how good the approximate posterior distribution of the parameters is to the true posterior distribution and hence a method for model selection. In this section we will take advantage of these features of the variational Bayesian inference in order to develop an algorithm for the parsimonious structure detection of polynomial NARX models.

6.4.1 The SVB-NARX algorithm

The algorithm is initialised by generating a NARX model containing the set of all possible polynomial basis functions, denoted $\mathcal{M}_0$ such that

$$\mathcal{M}_0 = \{\phi^m\}_{m=1}^{M}, \tag{6.72}$$

where ϕ^j is the jth polynomial basis function and M is the total number of basis functions in the current model structure.

At each iteration of the algorithm, denoted by i, the variational Bayesian inference procedure is applied to the previous model structure, $\mathcal{M}_{i-1}$, using Algorithm

6.1. The variational Bayesian inference requires a further iteration denoted by t. The iteration is performed until the increase in the variational lower bound, $\mathcal{L}(Q)$, is less than some pre-defined threshold, $T_{\mathcal{L}(Q)}$, such that

$$\mathcal{L}(Q)_t - \mathcal{L}(Q)_{t-1} \leq T_{\mathcal{L}(Q)}. \tag{6.73}$$

When the lower bound has reached convergence the final value is saved as

$$\mathcal{L}(Q)_t = \mathcal{L}(Q)^s, \tag{6.74}$$

so as to record the lower bound value for each model $\mathcal{M}_s$.

ARD values associated with each model term of the previous structure, $\mathcal{M}_{s-1}$ are calculated as

$$ARD^s = \left\{ \left(\frac{c_K}{d_{K_m}} \right)^{-1} \right\}_{m=1}^{M}. \tag{6.75}$$

Terms that correspond to ARD values falling below some threshold T^s_{ARD} are pruned from the model in order to form the new model structure, $\mathcal{M}_i$. The threshold is updated at each algorithm iteration as

$$\ln T^s_{ARD} = \min(\ln ARD^s) + \frac{(\max(\ln ARD^s) - \min(\ln ARD^s)}{r}). \tag{6.76}$$

with the resolution, r, being a tuning parameter of the algorithm set by the modeller. Consequences of the choice of r are discussed in the following section.

The threshold is therefore dependent on the range of the $\ln(ARD)$ values and removes terms at the lower fraction of this range depending on the value of r. This choice of threshold has the advantage of removing increasingly less terms at each algorithm iteration and hence discriminating more in the pruning as the correct model structure is approached. $\ln(ARD)$ values are used to calculate the threshold because the ARD values associated with highly relevant model terms can be very high in comparison to less relevant (but still correct) model terms. $\ln(ARD)$ values will provide greater discrimination between less relevant terms in this case.

The algorithm is terminated when $M = 1$ (all but one term have been pruned from the model). The optimal model choice, $\mathcal{M}^*$, is now selected as the model, $\mathcal{M}_s$, corresponding to the maximum recorded lower bound such that

$$\mathcal{M}^* = \mathcal{M}^{s^*}, \quad \text{where } s^* = \arg\max_s \mathcal{L}(Q)^s. \tag{6.77}$$

The justification for the optimal model being selected as the one that max-

imises the lower bound is given in Section 6.4.2. The algorithm is summarised in Algorithm 6.2.

6.4.2 Algorithm properties

In this section descriptions and examples are given in order to explain some of the properties of the algorithm. The examples used in this section refer to the identification of the system generated by the polynomial NARX model

$$y_k = 0.3y_{k-1} + 0.1u_{k-1} + 0.4y_{k-1}y_{k-2} + e_k \tag{6.78}$$

where e_k is a white noise sequence drawn from the normal distribution $\mathcal{N}(e_k|0, \sigma^2)$, where $\sigma^2 = 0.01$. In this case the model has been chosen for its structural simplicity.

Model selection by the variational lower bound

In the previous section it was stated that the optimal model choice is taken to be the model that maximises the variational lower bound after it has reached convergence. This is justified by considering the Bayesian model selection procedure discussed in Section 3.8.3.

In Equation (6.79) the conditional dependencies on the model $\mathcal{M}_s$ were neglected. Explicitly including the conditional dependencies, Equation (6.79) can be written

$$p(\boldsymbol{\Theta}|\boldsymbol{x}, \mathcal{M}_s) = \frac{P(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\Theta}, \mathcal{M}_s)p(\boldsymbol{\Theta}|\mathcal{M}_s)}{p(\boldsymbol{y}|\mathcal{M}_s)}, \tag{6.79}$$

which is in the same form as Equation (3.78). From the discussion of model selection in Section 3.8.3, considering the posterior distribution over the models $\mathcal{M}_s$ conditional on the data and applying Bayes' theorem the posterior distribution over the model is given by

$$p(\mathcal{M}_s|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\mathcal{M}_s)p(\mathcal{M}_s)}{p(\boldsymbol{y})}. \tag{6.80}$$

The first term in the numerator on the right hand side of Equation (6.80) is the same as the marginal likelihood in Equation (6.79). Setting equal prior distributions $p(\mathcal{M}_s)$ for each model and noting that the denominator is constant for a given data set the posterior is proportional to the marginal likelihood in Equation (6.79)

$$p(\mathcal{M}_s|\boldsymbol{y}) \propto p(\boldsymbol{y}|\mathcal{M}_s). \tag{6.81}$$

Algorithm 6.2 The SVB-NARX Algorithm

Initialise
$T_{\mathcal{L}(Q)}, T_{\text{ARD}}, a_0, b_0, c_0, d_0,$
Initialise model structure to all basis functions $\mathcal{M}_0 = \{\Phi^m\}_{m=1}^{M}$
$s = 0$
start Procedure
while $M > 1$
$s = s + 1$
$t = 1$

Calculate variational posteriors
while $\mathcal{L}(Q)_t - \mathcal{L}(Q)_{t-1} \leq T_{\mathcal{L}(Q)}$
t = t+1
update parameter estimates for model $\mathcal{M}_{s-1}$ using Algorithm 6.1,
calculate $\mathcal{L}(Q)_t$ via Equation (6.67)
end while
Set $\mathcal{L}(Q)^s = \mathcal{L}(Q)_t$
Calculate $\{ARD^s\}_{j=1}^{M}$ via Equation (6.75)
Calculate T^s_{ARD} via Equation (6.76)

Perform pruning step
Initialise pruning terms set, $\mathcal{M}^- = \emptyset$,
for $m = 1 : |\mathcal{M}_{i-1}|$
if $ARD^s_j \leq T^s_{\text{ARD}}$
collect terms to prune, $\mathcal{M}^- = \mathcal{M}^- \cup \Phi^j$,
end if
end for

Update model structure
Set current model structure to $\mathcal{M}_s = \mathcal{M}_{s-1} \setminus \mathcal{M}^-$
Set $M = |\mathcal{M}_s|$
end while

Select final model structure
Set optimal model $\mathcal{M}^* = \mathcal{M}^{s^*}$ where $s^* = \underset{s}{\text{argmax}} \mathcal{L}(Q)^s$
end Procedure

The VLB, $\mathcal{L}(Q)^s$ calculated for each model is an approximation of the marginal likelihood, $p(\boldsymbol{y}|\mathcal{M}_s)$. Equation (6.81) therefore provides the justification for using the VLB as a criterion for selecting final model structure.

Tuning parameters

The single tuning parameter of the algorithm is the resolution, r, whose value is set in advance by the modeller. It is named resolution because it defines the region of ARD values that are pruned from the model via Equation (6.76). Increasing the value of r leads to a higher resolution, resulting in less terms being pruned at each iteration, s, because a smaller portion of the range of ARD values is selected for pruning. Consequently, computation time will increase. Conversely, reducing the value of r increases the number of terms pruned at each iteration because a larger portion of the range of ARD values is selected for pruning.

It is to be noted that if r is chosen too small then correct model terms may be incorrectly pruned from the model. The effect of the tuning parameter is demonstrated through example via the structure detection of the system generated by Equation (6.82). Structure detection is performed on the test system using the SVB-NARX algorithm initialised with $a_0 = c_0 = 1^{-2}$ and $b_0 = d_0 = 1^{-4}$ the resolution is set as $r = 25, 50, 75, 100$. The correct model structure is identified for $r = 50, 75, 100$ but the algorithm fails for $r = 25$ as extra terms were included in the final model, see Figure 6.4. Note that the peak variational lower bound value is lowest at $r = 25$ when the incorrect model structure is identified and is constant for the other values or r.

For a given model and data set the variational lower bound is independent of the resolution that produced it. This allows for multiple algorithm runs with varying values of r that produce models which are directly comparable.

6.5 Results

In this section the SVB-ARD algorithm is demonstrated in order to assess the performance for the purpose of joint structure detection and parameter estimation. The algorithm is then applied to the identification of a DEA system in order to validate the performance of the algorithm on real data. For both the synthetic and the real case the algorithm is benchmarked against the FRO and SEMP algorithms. The benchmark example is also compared against LASSO.

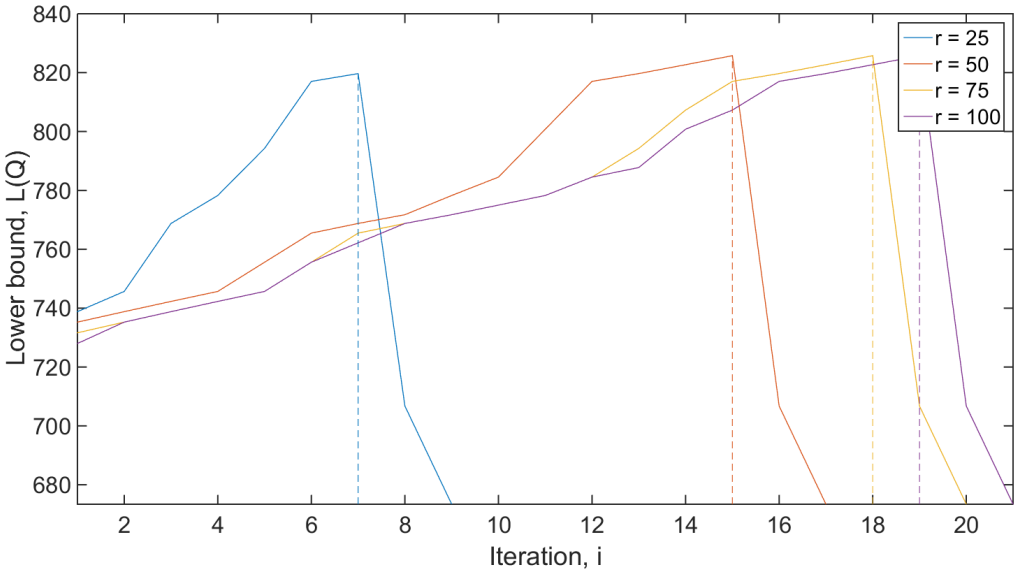

Figure 6.4: **The resolution r of the algorithm controls how many terms are removed at each iteration and therefore the total number of iterations performed.** Variational lower bound against iteration number for $r = 25, 50, 75, 100$. The dotted line indicates the maximum value of the variational lower bound for each value of r.

6.5.1 Numerical example 1: A non-linear benchmark

The SVB-NARX algorithm is demonstrated by application to the structure detection of the generative system given by

$$\begin{aligned} y_k &= \theta_1 y_{k-2} + \theta_2 y_{k-1} u_{k-1} + \theta_3 u_{k-2}^2 + \theta_4 y_{k-1}^3 \\ &\quad + \theta_5 y_{k-2} u_{k-2}^2 + e_k \\ \boldsymbol{\theta} &= [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T \\ &= [-0.5, 0.7, 0.6, 0.2, -0.7]^T \end{aligned} \tag{6.82}$$

where e_k is a normally distributed white noise sequence drawn from the distribution $\mathcal{N}(e_k|0, \sigma^2)$. The system is simulated for $N = 1000$ data samples with $\sigma^2 = 0.004$ corresponding to a SNR of $\approx 20dB$. The input, u_k, is drawn from a uniform distribution in the range $[-1, 1]$. This system was initially used in [80] and again as a benchmark in [99] and [8]. The system demonstrates a situation in which the FRO algorithm fails to select the correct terms, necessitating further identification steps.

Assuming no prior knowledge of the model structure the polynomial model order is conservatively set to $n_p = 4$ and the dynamic order in both the input

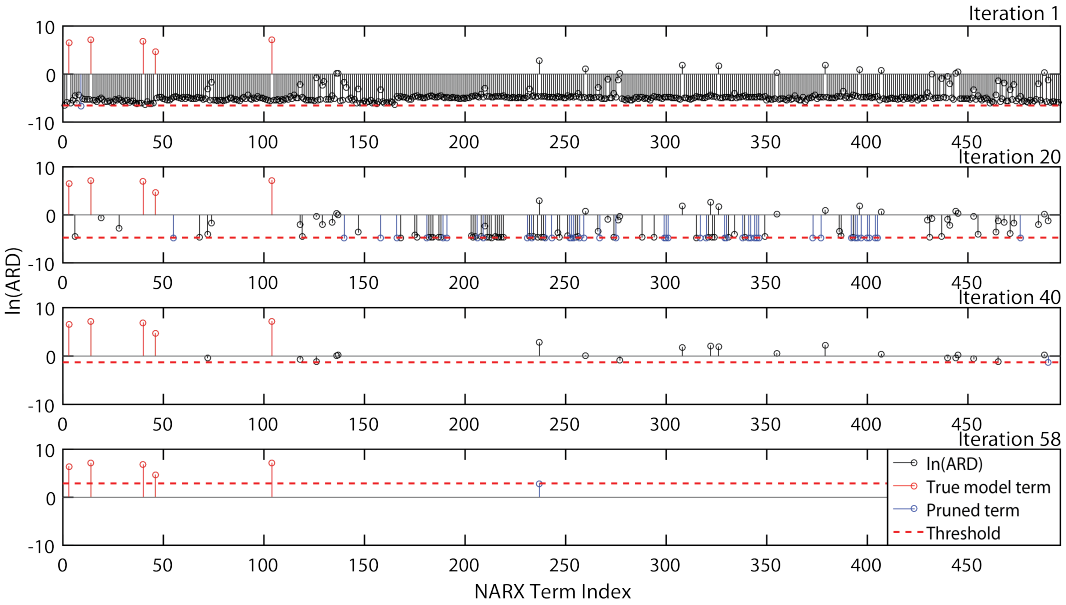

Figure 6.5: **The SVB-NARX algorithm iteratively prunes terms that fall below a threshold.** Model term index plotted against ARD value. The log(ARD) value of each model term is given by the black stems, true model terms are painted red. The dashed red line indicates the threshold, r, at each iteration if the algorithm terms with an ARD value below the threshold are pruned, these are painted blue. The correct model is found at iteration 58 although the algorithm continues untill only one term remains in the model.

and the output is set to $n_u = n_y = 4$. The assumption on the model order leads to a superset of $\mathcal{M} = 495$ model terms in which to search (including the DC term). The SVB-NARX algorithm requires initialisation of the hyper-parameters associated with the prior distributions, namely a_0, b_0, c_0 and d_0 as well as the resolution variable, r. Hyper-parameters were chosen as $a_0 = c_0 = 1 \times 10^{-2}$ and $b_0 = d_0 = 1 \times 10^{-4}$, so as to produce uninformative prior distributions. The mean of the Gamma distribution on τ^{-1} at these values is undefined but it has mode $b_0/(a_0 + 1) \approx 1 \times 10^{-4}$. This implies that the most likely variance on $\boldsymbol{\theta}$ will be small a priori. The a priori variance on τ^{-1} is also undefined at these values, however the variance on τ can be computed as $a_0/b_0^2 = 1 \times 10^6$. It can hence be concluded that although the prior distribution indicates a preference for $\boldsymbol{\theta}$ to take small values, this effect will be minimal on the inference because of the broad distribution. The same reasoning can be applied to the prior distribution on $\boldsymbol{\alpha}$.

The FRO, LASSO and SEMP algorithms are applied to the same data set to provide a benchmark. LASSO is applied in order to show a comparison to available sparse methods and to demonstrate the resultant over parametrisation. The FRO and SEMP algorithms provide benchmarks to standard identification meth-

Table 6.1: **The SVB-NARX algorithm selects the correct model structure.** Terms selected the SVB-NARX, LASSO and FRO algorithms for the system given by Equation (6.82).

SVB-NARX				
-	Basis function	ARD ($\times 10^3$)	Parameter estimate	Correct term?
	$y_{k-2}u_{k-2}^2$	1.1821	-0.7053	✓
	$y_{k-1}u_{k-1}$	1.2037	0.6990	✓
	u_{k-2}^2	0.8708	0.5999	✓
	y_{k-2}	0.6064	-0.5006	✓
	y_{k-1}^3	0.1047	0.2079	✓
FRO				
Iteration	Basis function	ERR	Parameter estimate	Correct term?
1	$y_{k-4}u_{k-2}^2$	0.3792	-0.0035	✗
2	u_{k-2}^2	0.1576	0.6006	✓
3	y_{k-2}	0.2681	-0.5006	✓
4	$y_{k-1}u_{k-1}$	0.1600	0.6990	✓
5	$y_{k-2}u_{k-2}^2$	0.0236	-0.7077	✓
6	y_{k-1}^3	0.0070	0.2028	✓
LASSO				
-	Basis function	-	Parameter estimate	Correct term?
	y_{k-1}		-0.4669	✓
	$y_{k-1}u_{k-1}$		0.6250	✓
	u_{k-1}^2		0.5610	✓
	y_{k-1}^3		0.1316	✓
	$y_{k-1}u_{k-1}^2$		-0.6137	✓
	$y_{k-1}u_{k-1}^2$		-0.0371	✗
	$y_{k-1}u_{k-1}^2$		-0.0160	✗
	$y_{k-1}u_{k-1}^2$		-0.0036	✗
	$y_{k-1}u_{k-1}^2$		0.0236	✗
	$u_{k-1}u_{k-1}^2$		0.0025	✗
	$y_{k-1}^2u_{k-1}^2$		0.0547	✗
	$y_{k-1}u_{k-1}u_{k-1}^2$		0.0742	✗
SEMP				
Iteration	Basis function	SRR	Parameter estimate	Correct term?
1	u_{k-1}^2	0.3226	0.5904	✓
2	y_{k-1}	0.2001	-0.5164	✓
3	$y_{k-1}u_{k-1}$	0.2344	0.6970	✓
4	$y_{k-1}u_{k-1}^2$	0.1525	-0.6556	✓
5	y_{k-1}^3	0.0888	0.2021	✓

ods. Both algorithms are terminated when the ERR/SERR value falls below a threshold, in this case the threshold was set to 0.01 in both cases. The LASSO parameter estimates are found by minimising the cost function, given by Equation (3.34), using a gradient descent algorithm [39]. The parameters are estimated with values of λ in the range $[0, 1]$. 10-fold cross validation is then used to estimate the MSPE for the parameter set at each value of λ. The final model is selected as the model with the greatest MSPE within one standard deviation of the minimum found MSPE, see Figure 6.6**C**.

The SVB-NARX algorithm correctly identifies the model structure at iteration 58, indicated by the maximum of the variational lower bound recorded at each iteration. 63 iteration were required before all but one term is left in the model terminating the algorithm, see Figure 6.5. The number of iterations can be reduced drastically by selecting a smaller value for r while still obtaining the correct model structure, however in a different scenario relevant terms could be incorrectly pruned. The correct model is selected at the maximum of the variational lower bound, see Figure 6.6**A**. The Bayesian framework allows for a natural calculation of the probability distributions over the model parameters, see Figure 6.7. The algorithm has estimated all the parameters within a 95% confidence interval, see Table 6.2 and Figure 6.7.

The FRO algorithm selects an incorrect term at the first iteration, see Table 6.1. The incorrect term selection is suggested to be a result of the local nature of the search performed by the algorithm [112]. It should be noted that this problem can be solved by the inclusion of an extension to the FRO algorithm that adds a pruning step at every iteration of the algorithm [99]. The SEMP algorithm identifies the correct model structure in 5 model iterations.

The LASSO parameter estimate resulting in the minimum MSPE is found at $\lambda = 0.0043$ with a sample standard deviation of 3.0443×10^{-04}. The model selected is then taken as the one with the maximum MSPE within one standard deviation and is found at $\lambda = 0.0090$. The resulting model has 12 non-zero parameter estimates, incorrectly including 7 extra terms to the true model structure and therefore being greatly over parametrised. The true model structure is not recovered at any value of λ.

△

6.5.2 Numerical example 2: Effect of noise on algorithm performance.

The generative system given by Equation (6.82) is used again to investigate the effect of increased noise on the effectiveness of the SVB-NARX algorithm. The

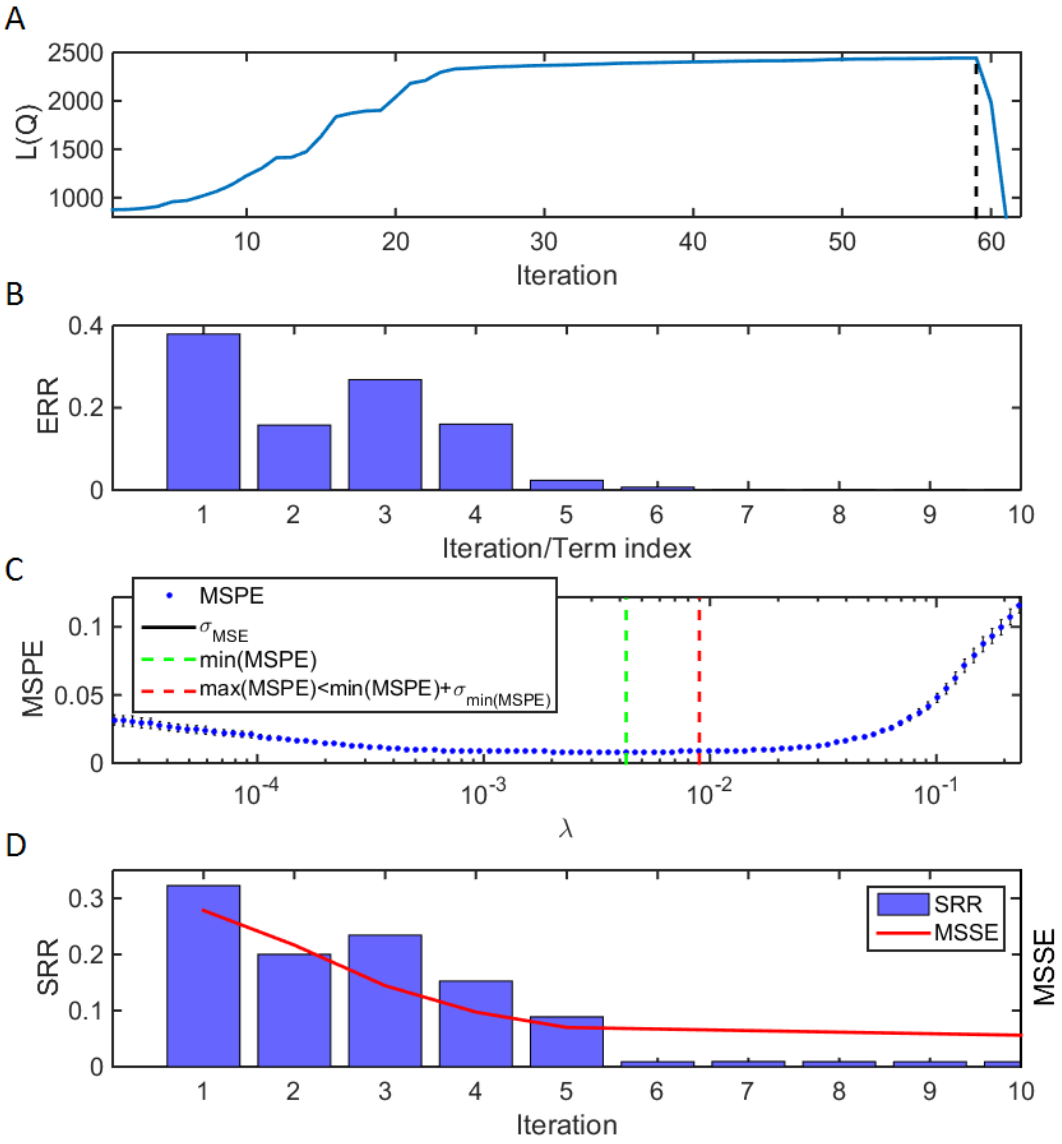

Figure 6.6: **Model selection is performed automatically for the SVB-NARX algorithm based on the variational lower bound. Model selection for FRO, LASSO and SEMP are based on a user defined threshold.** **A**) SVB-NARX model selection: The variational lower bound plotted against iteration number. The maximum value is indicated by the black dotted line. **B**) FRO model selection: Error reduction ratio plotted against iteration number. The final model structure is chosen when the ERR falls below the threshold 0.01, in this case an incorrect model structure is identified. **C**) LASSO model selection: The MSPE associated with the sparse parameter estimates found at values of λ in the range $[0, 1]$, the final model is chosen as the one with the greatest MSPE within one standard deviation of the minimum MSPE. **D**) SEMP model selection: Simulated error reduction ratio and MSSE plotted against iteration number. The final model structure is chosen when the SERR falls below 0.01.

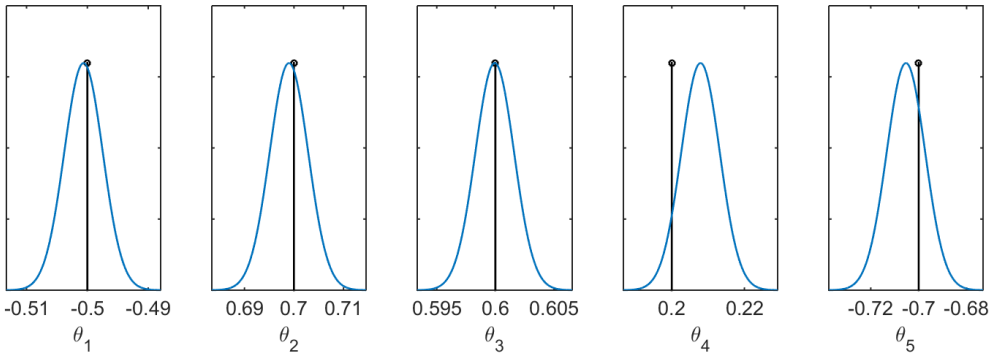

Figure 6.7: **Bayesian methods such as the SVB-NARX algorithm naturally produce distributions over the model parameters.** Distributions over the model parameters of Equation (6.82). The stem plots indicate the true values.

Table 6.2: **Parameter estimates are all within a 95% confidence interval.** Comparison of true and estimated parameters along with the associated 95% confidence interval for the system given by Equation (6.82)

Parameter	True parameter	Parameter estimate
θ_1	-0.5	-0.5006 $\pm$ 0.0062
θ_2	0.7	0.6990 $\pm$ 0.0076
θ_3	0.6	0.5999 $\pm$ 0.0032
θ_4	0.2	0.2079 $\pm$ 0.0104
θ_5	-0.7	-0.7053 $\pm$ 0.0158

system is simulated again under the same conditions as the previous example but for varying noise levels. The following four noise levels were used to generate the data: $\sigma = 0.004, 0.001, 0.025, 0.05$, generating signals with a signal to noise ratio of $SNR \approx 15, 10, 5$ and 2 dB respectively. The results from the previous example will also be included in the discussion for comparison.

The SVB-NARX algorithm is initialised as before with the same assumptions on the polynomial and dynamic order, namely $np = n_u = n_y = 4$. The algorithm selects the correct model structure at each noise level, this can be seen in Figure 6.8 where the dashed black line indicates the correct model structure and corresponds in all instances to the maximum of the variational bound. The estimated parameters are within the 95% confidence intervals at all noise levels, see Table 6.3. As should be expected with increasing noise the variance of the parameter distributions also increases, see Figure 6.9.

△

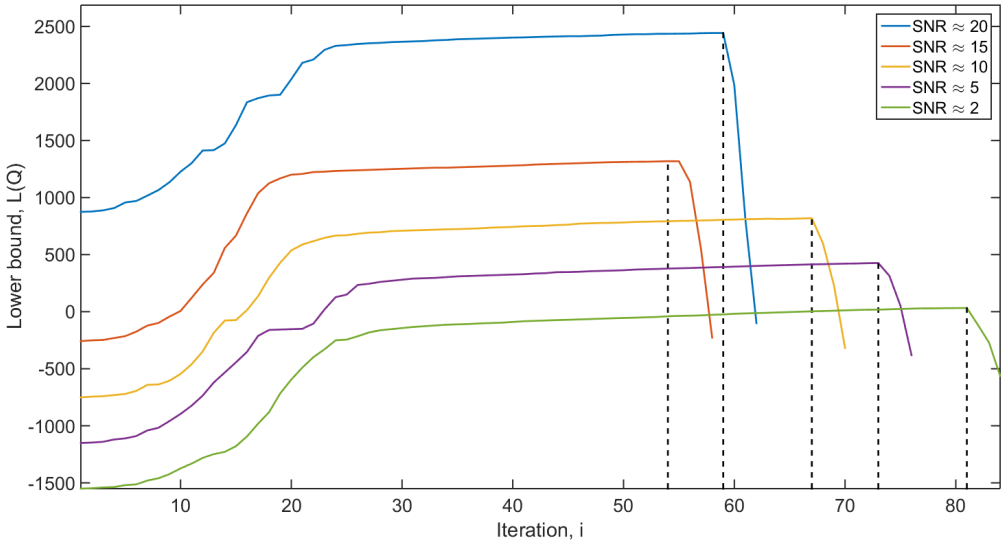

Figure 6.8: **The SVB-NARX algorithm selects the correct model structure at each noise level.** The variational lower bound plotted against iteration number fr each noise level. The correct model structure is indicated by the black dotted line and corresponds with the maximum of the variaitinal lower bound for all noise levels. The bound converges to a smaller value with increasing noise.

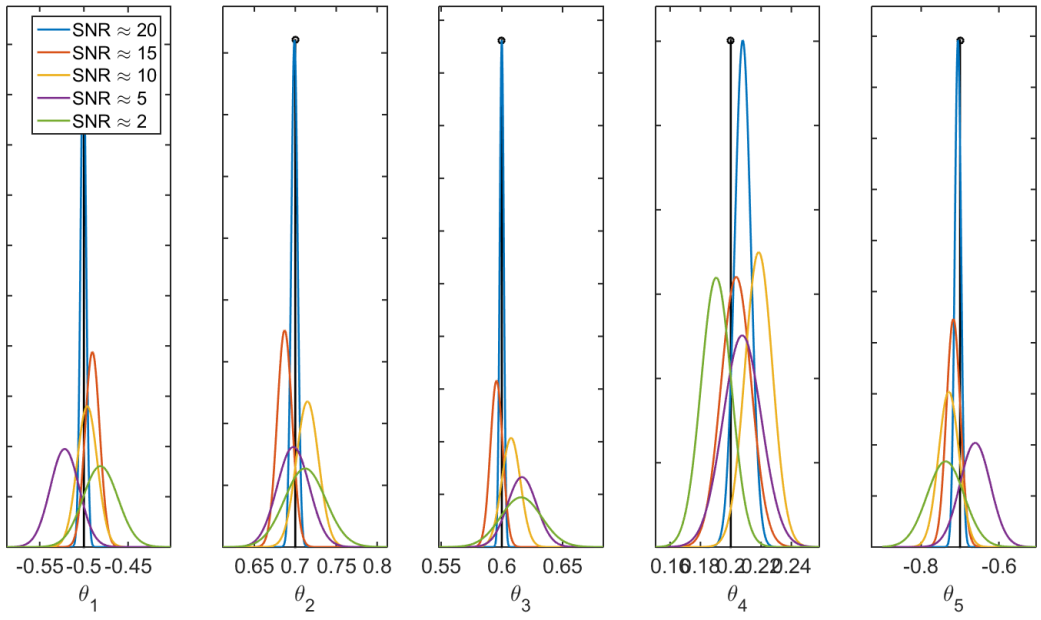

Figure 6.9: **Parameter distributions calculated at different noise levels.** The true parameter is given by the stem plot. As expected the parameter estimates are less certain at higher noise levels.

6.5.3 Numerical example 3: Assessing uncertainty in the frequency domain description of NARX models.

The uncertainty in the time domain description of the system is characterised by the predictive distribution but so far no discussion has been made of how the uncertainty in the frequency domain can be quantified. A previous study quantifies the uncertainty in GFRFs by taking MC samples from uniformly distributed model parameters and mapping them into GFRFs [136]. In this example a similar approach is taken but the samples are drawn from the model parameter covariance matrix. The sampled parameter vectors are then mapped into the systems NOFRFs.

In the previous example the effects of noise on the SVB-NARX algorithm were investigated resulting in parameter estimates for the system at different levels of noise. The algorithm infers a posterior distribution over the model parameters as a multivariate normal distribution defined by its mean and covariance matrix. 200 samples are drawn randomly from the parameter distributions at noise levels $SNR \approx 2, 5, 10$ and 20. Each sampled parameter vector is used to generate NOFRFs via Algorithm 4.1.

No first order FRF exists for the system given by Equation (6.82) so that the output spectra is approximately formed by the second and third order outputs, higher order outputs are assumed to be negligible. The second and third order output spectra are calculated for each sampled parameter vector for the four noise levels, see Figure 6.13. Note that sampled output spectra appear to be evenly distributed around the output spectra calculated at the MAP parameter estimate. In each case the output spectra of the true system is enclosed within the sampled outputs.

△

The three examples show the effectiveness of the SVB-NARX algorithm for the

Table 6.3: **Parameters estimated by the SVB-NARX algorithm are within the 95% confidence interval for all noise levels.** Estimated parameters values for the system given by Equation (6.82) at different noise levels.

θ_{True}	$\theta_{SNR\approx20}$	$\theta_{SNR\approx15}$	$\theta_{SNR\approx10}$	$\theta_{SNR\approx5}$	$\theta_{SNR\approx2}$
-0.50	-0.501 ± 0.006	-0.490 ± 0.016	-0.496 ± 0.023	-0.522 ± 0.033	-0.481 ± 0.040
0.70	0.699 ± 0.008	0.687 ± 0.018	0.715 ± 0.027	0.698 ± 0.039	0.712 ± 0.050
0.60	0.600 ± 0.003	0.596 ± 0.010	0.608 ± 0.015	0.617 ± 0.024	0.616 ± 0.034
0.20	0.208 ± 0.011	0.204 ± 0.020	0.218 ± 0.018	0.208 ± 0.025	0.190 ± 0.020
-0.70	-0.705 ± 0.016	-0.717 ± 0.036	-0.729 ± 0.053	-0.661 ± 0.078	-0.737 ± 0.095

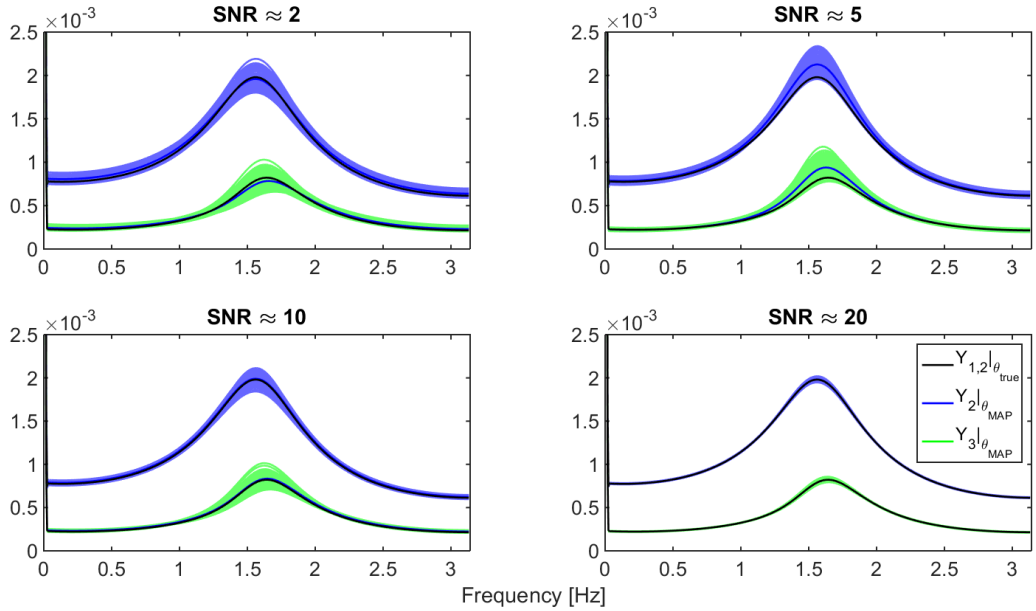

Figure 6.10: **The noise in the identified system is reflected in the uncertainty of n'th order system outputs calculated via its NOFRFs.** First (pale blue) and second (pale green) order output spectra of the system given by Equation (6.82) with parameters sampled from the covariance matrix with first (blue) and second (green) order output spectra taken at the MAP parameter estimate. True system outputs are given in black.

structure detection of known NARX systems. Of course, this represents a much simpler task than the identification of real systems. To demonstrate the use of the SVB-NARX algorithm on a real data set it is applied to the structure detection of the set of six film type DEAs in the following section.

6.6 Bayesian modelling and uncertainty analysis of DEAs

The SVB-NARX is applied to the identification of the set of DEA actuators introduced in Chapter 2. As discussed previously, all six DEA data sets display time varying behaviour. The algorithm is not appropriate for modelling time-varying systems. As such, and to provide a direct comparison to identification with the SEMP algorithm performed earlier in this thesis, the identification is performed using the same data used previously in Chapter 4. Following the examples in previous chapters, identification is demonstrated for DEA 5 and then results are given for the remainder of the set of actuators. The FRO algorithm is also applied to DEA 5 in order to benchmark the SVB-NARX algorithm against a further contemporary system identification method.

The SVB-NARX algorithm is initialised with $r = 1000$, and hyper-parameters are initialised so as to produce broad prior distributions: $a_0 = c_0 = 1 \times 10^{-2}$ and $b_0 = d_0 = 1 \times 10^{-4}$. The super set of model terms is generated by choosing maximum dynamic order $n_y = n_y = 3$ and maximum polynomial order n_p in line with the identification performed in Chapter 4. Application of the SVB-NARX algorithm to the DEA 5 data set results in an identified model containing 13 model terms in 28 iterations of the algorithm. The chosen model is found at iteration 17, indicated by the dashed line in Figure 6.12**A**. Selected model terms and estimated parameters are given with there 95% confidence intervals, see Table 6.4.

The model identified by the FRO algorithm contains 12 terms. The algorithm was terminated when the ERR value fell below 1×10^{-5}, see Figure 6.12**B**. The FRO algorithm assigns a large value of ERR (0.99) to the first model term it selects $y(k-1)$, after which the ERR rapidly falls to values of magnitude less that 1×10^{-3}. This makes selection of the final model structure difficult because there is no obvious point at which the algorithm should be terminated. After the chosen ERR threshold the values fluctuate around a value $\approx 1^{-6}$. The identification with the SEMP algorithm is performed in Chapter 4.

The SEMP and FRO algorithms both identify models with fewer terms, 11 and 12 respectively, see Table 6.4. This is significantly fewer than for the model identified with SVB-NARX by Algorithm 6.2. This does not necessarily indicate that the model identified by SVB-NARX is any more complex. The parameters estimated by SVB-NARX are regularised due to the prior distribution over the parameters. Model complexity/over fitting has been penalised automatically as part of the inference. This can be seen by considering the correlation tests performed on the model, see Figure 6.11. The correlation functions display similar features to those performed before the for models identified with SEMP, see Figures B.2-B.3. However in the case of the autocorrelation of the residuals and the function $\Phi_{\xi(\xi u)}$, both tests perform significantly better.

The SVB-NARX performs similarly to the FRO as measured by the MSSE metric calculated over the validation data set. The MSSE for the identified models are calculated as $MSSE = 7.26 \times 10^{-4}$ and $MSSE = 7.14 \times 10^{-4}$ for SVB-NARX and FRO respectively. The SEMP algorithm performs better for DEA 5 with $MSSE = 6.13 \times 10^{-4}$. The Bayesian framework allows for the predictive distribution to be calculated by Equation (6.70). Confidence intervals of 95% are shown along with the model predicted output and the true system output over small sections of data (200 data samples) for both the training and validation data. As should be expected, the model predicted output is confined to the bounds of the confidence intervals for the majority of the data record. The model has difficulty

Terms	$\boldsymbol{\theta}_{SVB-NARX}$	$\boldsymbol{\theta}_{SEMP}$	$\boldsymbol{\theta}_{FRO}$
DC	-	0.0475	0.0518
y_{k-1}	0.9056 $\pm$ 0.0226	0.8950	1.0024
y_{k-2}	-	-	-0.2471
y_{k-3}	0.1977 $\pm$ 0.0293	0.0415	0.1379
u_{k-1}	-	0.4115	-
u_{k-2}	-	-0.4216	-
u_{k-3}	-	0.0671	-
$y_{k-1}y_{k-2}$	-	-0.0770	-
$y_{k-1}u_{k-1}$	-	-	3.0050
$y_{k-1}u_{k-2}$	-0.5504 $\pm$ 0.0411	-	-5.1093
$y_{k-1}u_{k-3}$	-	-	1.7671
y_{k-2}^2	-0.2144 $\pm$ 0.0419	-	-
$y_{k-2}u_{k-1}$	0.5725 $\pm$ 0.0426	-	-2.2368
$y_{k-2}u_{k-2}$	-	-	4.4082
$y_{k-2}u_{k-3}$	-	-	-1.7435
u_{k-1}^2	0.8962 $\pm$ 0.1620	-	-
$u_{k-1}u_{k-2}$	-1.3596 $\pm$ 0.3081	-	-
$u_{k-2}u_{k-3}$	1.1632 $\pm$ 0.3193	-	-
u_{k-3}^2	-0.6669 $\pm$ 0.1730	-	-
$y_{k-1}^2u_{k-1}$	0.3082 $\pm$ 0.0497	-	-
$y_{k-1}y_{k-2}u_{k-3}$	-0.5610 $\pm$ 0.14721	-	-
$y_{k-1}u_{k-1}^2$	-	0.6566	- 0.2914
$y_{k-1}u_{k-2}^2$	-	-0.7383	-
$y_{k-1}u_{k-3}^2$	-	-	-0.2780
$y_{k-2}^2u_{k-3}$	0.3906 $\pm$ 0.1061	-	-
$y_{k-2}u_{k-1}u_{k-3}$	-	0.2786	-
$y_{k-3}u_{k-1}^2$	-	-0.0366	-
u_{k-1}^3	0.2016 $\pm$ 0.0386	-	-
$u_{k-1}u_{k-2}u_{k-3}$	-0.3676 $\pm$ 0.1317	-	-
$u_{k-2}^2u_{k-3}$	0.1755 $\pm$ 0.1030	-	-
MSSE	7.26×10^{-3}	6.13×10^{-3}	7.14×10^{-3}

Table 6.4: **The SVB-NARX, SEMP and FRO algorithms select different terms to represent the DEA 5 data set. SVB-NARX provides a natural method for calculating uncertainty bounds on parameter estimates.**, NARX model terms for DEA 5 identified using SVB-NARX, SEMP and FRO along with their corresponding parameter values. Parameter values for the SVB-NARX identified model are given with their 95% confidence intervals.

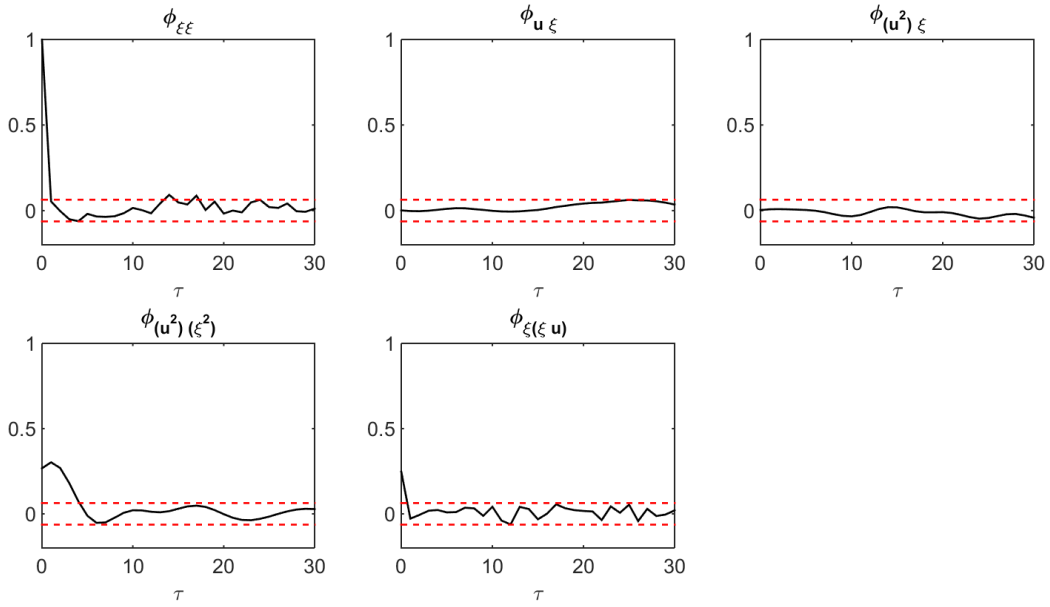

Figure 6.11: **Correlation tests for NARX models of DEAs identified with the SVB-NARX algorithm perform better than for the SEMP algorithm.** Linear and non-linear correlation tests for the NARX model of DEA 5 identified using the VB-NARX algorithm.

modelling the larger peaks in the output and model predictions are poor in these areas, see Figure 6.12**C-D**.

The SVB-NARX algorithm is performed on the rest of the set of DEA actuators. The identified models along with their parameters are given in Appendix D.1. Correlation tests for the model of each actuator are given in Appendix D.2. It is interesting to note that very different model terms are selected by the three different algorithms when modeling all six DEAs, see Table 6.4 and Appendix D.1. This mirrors the motivation for Chapters 4 and 5 in which emphasis was placed on comparisons between different models. It is argued that in order to check for consistency between models a frequency domain description is required. As was the case in Example 6.5.3, the Bayesian inference performed as part of the SVB-NARX algorithm provides the posterior distribution over the model parameters. Sampling from this distribution and mapping the sampled parameters into the frequency domain as NOFRFs results in a sampled distribution over the n'th order NOFRFs/output spectra.

In order to apply Algorithm 4.1 to the models in order to generate the system NOFRFs it is necessary to remove the DC component from the models. However, it is unclear how this affects the posterior distribution of the model parameters. To avoid this process sampling is performed from the original models and then

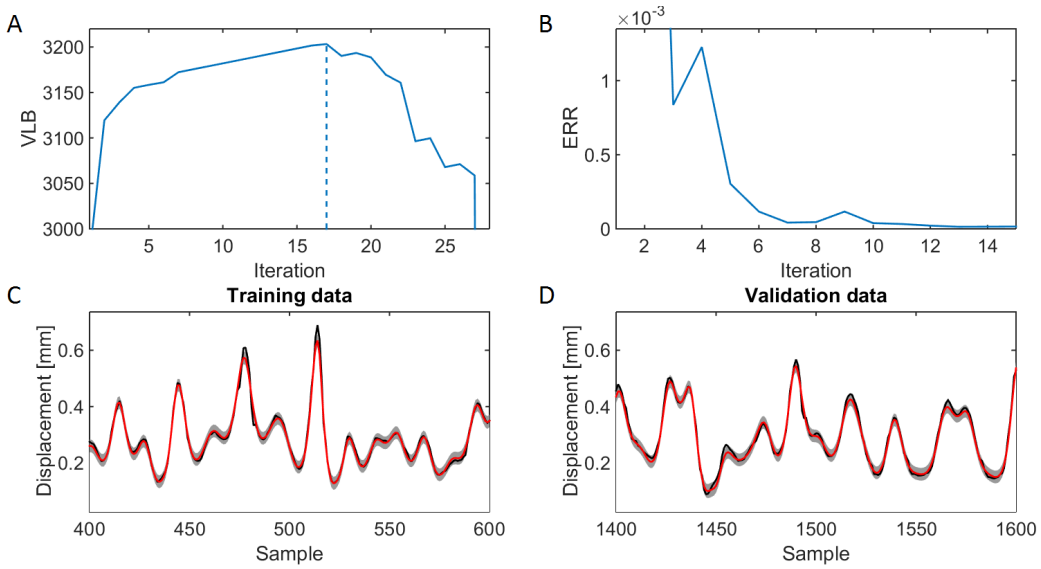

Figure 6.12: **The 'best' model is obtained automatically by the SVB-NARX algorithm as the model that maximises the variational lower bound, for both FRO and SEMP a subjective choice of threshold for the ERR is necessary.** **A**) The SVB-NARX algorithm selects the 'best' model structure at iteration 17 as marked by the dashed line. **B**) The ERR threshold for FRO is chosen as 1×10^{-5}, ending the algorithm at the 10'th iteration. **C-D** The model predicted output (Red) for the model identified by the SVB-NARX algorithm is shown with the true system output (Black) and 99% confidence intervals (Gray shaded area). A small section (200 samples) of both training **C**) and validation **D**) data are shown for clarity.

each sampled parameter vector is mapped to the zero equilibrium point.

100 MC samples are drawn from the posterior parameter distribution of the models of each DEA and the DC component is removed, see Appendix D.2. The frequency domain mapping is performed for each of the sampled parameter vectors via Algorithm 4.1. The algorithm is initialised as in Chapter 4 for the models identified with the SEMP algorithm and the n'th order output spectra are calculated for $n = 1, \ldots, 3$, see Figure 6.13. For EAP 1, 2, 4, 5 and 6 the output spectra calculated from the SVB-NARX and SEMP identified models are broadly similar, falling within the range of the spectra calculated by sampling from the posterior parameter distribution. In the case of DEA 3 the two models give quite different results. Although the 1st and second order output spectra are similar, the third order output spectra behaves very differently for the two models.

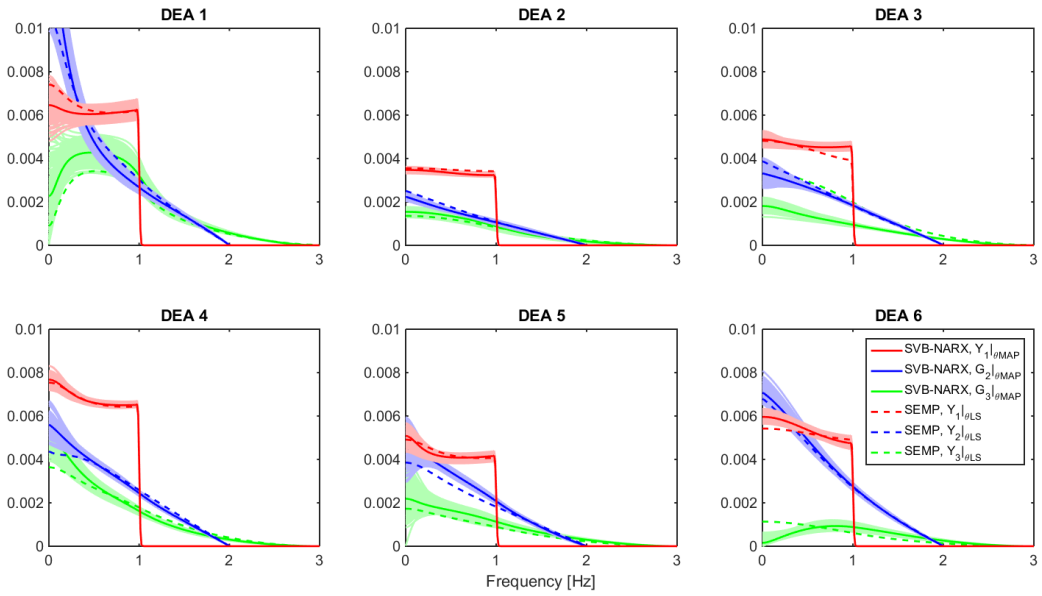

Figure 6.13: **The frequency domain provides an invariant description of a nonlinear system. Uncertainty can be incorporated into the frequency domain description by mapping parameters sampled from the posterior distribution.** n'th order output spectra for DEAs 1:6 calculated from models identified using **Solid lines**: SVB-NARX and **Dashed lines**: SEMP given in Table 6.4. **Pale colours**: A distribution over each output spectrum is made by sampling from the posterior parameter distribution.

6.7 Discussion

The aim of this chapter was the development of a novel method for the joint structure detection and parameter estimation of polynomial NARX models within a Bayesian framework and the application of this method to the Bayesian system identification of DEAs. This is achieved by the development of the SVB-NARX algorithm. The use of Bayesian analysis in system identification is appealing because model complexity is automatically penalised in model selection and uncertainty is naturally incorporated into the modelling process. However, to the best of the author's knowledge, no Bayesian algorithms for the identification of NARX models that can be performed in closed form exist.

Term selection in the SVB-NARX algorithm is driven by a sparsity inducing prior, referred to as ARD. ARD allows the data to automatically remove irrelevant basis functions from the model by regularising each term independently. However the inclusion of the ARD prior causes the inference problem to be intractable, necessitating the use of approximate inference methods. VB is used in which the posterior distribution is approximated by the product of simpler distributions. The

advantage of the VB approach is that the inference problem can be performed efficiently in a series of closed form update equations, negating the need for sampling methods.

One of the key advantages of Bayesian inference is that model complexity is automatically penalised such that model selection can be performed automatically. In the SVB-NARX algorithm the variational lower bound is conveniently approximated at every iteration of the algorithm. The VLB approximates the marginal likelihood or model evidence and so can be used as a measure of model quality. The SVB-NARX algorithm essentially provides a series of models, one identified at each iteration of the algorithm, on which model selection can be performed based on the value of the VLB.

A benchmark study is performed using a generative system that is commonly used in the literature. The algorithm performs favourably in comparison to other standard identification algorithms such as FRO and SEMP. However, in comparison to these alternative contemporary algorithms, the SVB-NARX algorithm developed here has the ability to identify a full uncertainty model and hence quantify the uncertainty associated with the identified model. For the purpose of identifying NARX models within a Bayesian framework only one other algorithm is available in the literature [8]. The SVB-NARX algorithm has significant advantages over the RJMCMC algorithm in the use of variational inference which allows the posterior distribution over the model parameters to be approximated by way of a series of closed form update equations. The inference step can hence be performed efficiently without the need for computationally burdensome sampling methods and is very simple to implement.

Applying the algorithm to the DEA data sets investigated throughout this thesis results in models of comparable accuracy to those identified in Chapter 4. By mapping the model parameters into the frequency domain as NOFRFs it can be seen how, despite the fact that the time domain models posses very different model terms, the frequency domain description is easily comparable. The result highlights the differences in the underlying behaviour of the different DEA actuators. By quantifying the uncertainty in the models it is now possible to show that the variations in the DEA behaviour cannot be explained by inaccuracy in the modelling but that the data indicates true differences. Further evidence for this hypothesis is shown in the similarity in the frequency domain behaviour of models separately identified using SEMP and SVB-NARX, see Figure 6.13.

The estimation of an uncertainty model for the DEAs raises an interesting question: how can the uncertainty be quantified in the frequency domain. This question is partially answered by sampling from the posterior distribution of the

model parameters and mapping the sampled parameters into the frequency domain as NOFRFs. This results in a distribution over the systems output spectra that give some insight into the characteristics of the frequency domain uncertainty. This motivates a more rigorous investigation into uncertain behaviour in the frequency domain which will be the subject of the next chapter.

6.8 Summary

In this chapter a novel approach to non-linear system identification of NARX models within a Bayesian framework is introduced: The SVB-NARX algorithm. The use of ARD for the pruning of irrelevant basis functions from NARX models has been demonstrated. Variational Bayesian inference is used in which the posterior distribution over the parameters is approximated by a factorisable variational distribution which can be lower-bounded. The Bayesian modelling framework with ARD then represents a method for sparse estimation of the model parameters. However, although this is shown to be successful for identifying terms that are not relevant to the generation of the output, it is not able to successfully identify the non-linear model structure in one step.

This motivates the development of the SVB-NARX algorithm. The algorithm uses ARD as a method for driving structure detection. At each iteration of the algorithm the ARD inference procedure is performed and irrelevant terms are pruned from the model. Bayesian model selection is then performed on the set of NARX models that are produced by the iterative pruning.

A series of results are presented to demonstrate the successful application of the new algorithm. Firstly a benchmark example is performed using data generated by a synthetic NARX system. The algorithm is then applied to the real world application of the identification of DEA systems. In all cases the algorithm is shown to perform well, identifying the correct model structure in the case of the synthetic examples. It is concluded that the SVB-NARX algorithm developed in this chapter provides a novel method for the identification of NARX models within a Bayesian framework.

In the final section the SVB-NARX algorithm is applied to the identification of DEA actuators. The identified models are shown to be broadly consistent with those identified in previous chapters by mapping them into the frequency domain as NOFRFs. MC samples are taken from the posterior parameter distribution and mapped into the frequency domain to provide a characterisation of the uncertainty in the system output spectra.

In the next chapter model uncertainty in the frequency domain is investigated

further for both linear and non-linear systems.

Chapter 7

Uncertainy Analysis in the Frequency Domain

This thesis is motivated by the modelling and frequency domain analysis of nonlinear systems. In the previous chapter, uncertainty was incorporated into the modelling process by the use of Bayesian inference. This allows for the natural description of the model uncertainty via the uncertainty in the model parameters. However, as has been argued throughout, the parameters of a time domain model provide limited insight into the underlying system behaviour that the model describes. Mapping the parameters into the frequency domain as GFRFs and NOFRFs provides a much more transparent insight into the system dynamics. The current frequency domain description does not take into account the uncertainty in the parameters, providing the motivation for this chapter. The aim of the work presented here is then: How can uncertainty be included into the frequency domain analysis?

The uncertainty associated with the FRF has been investigated in other disciplines of engineering. It is of particular interest in structural dynamics where the uncertainty is propagated through finite element models [122] as well as quantifying errors in experimentally gathered FRF measurements [109]. Both of these methods use the concept of uncertainty propagation, calculating the uncertainty in the output of some function by considering the uncertainty in the functions input variables. Classical uncertainty propagation is performed via a Taylor expansion of the function in question and then using the property of the sum of correlated variables as described in [53]. The former reference provides the international standard for calculating measurement uncertainty. This method allows for estimating the uncertainty on real valued data that is the output of some function.

Frequency domain descriptions of a system such as the FRF, however, are com-

plex, *i.e.* they consist of a real and imaginary part. For complex valued data there is no standard method for expressing the uncertainty. The complex statistics approach estimates the uncertainty in a complex variable as a symmetrical normal distribution in the real-imaginary space, such that its variance can be quantified by a single number [97]. Alternatively, the multivariate uncertainty method assigns a bivariate normal distribution to the complex variable such that a covariance matrix describes the correlation of the real and imaginary parts [46, 102]. The multivariate approach has been used to as the basis for propagating uncertainty into complex variables without the restrictive assumptions of the complex statistics approach [46, 47].

The multivariate uncertainty propagation approach has previously been used for quantifying the measurement uncertainty in experimentally gathered data [47, 109]. However, for the application of control, it is common that FRFs are calculated from data driven models. And, as discussed previously, the frequency domain description of a system provides the basis for many analysis as well as control methods. It would therefore be extremely advantageous to be able to estimate the uncertainty in the FRF directly from the model uncertainty. A previous work more in line with this goal is presented in [136] in which model parameter values are sampled from a uniform parameter distribution in some range before being mapped into the frequency domain using the probing method. Although this method is readily applicable it has significant disadvantages in the computational cost of mapping all of the sampled parameters. Furthermore, the assumption of a uniform distribution on the parameters may not reflect a realistic situation. A method for estimating the uncertainty in the linear as well as higher order FRFs is presented based on multivariate uncertainty propagation. The FRF's considered are calculated from both ARX and NARX models that are identified from system input output data. To the best of the authors knowledge this approach has not been proposed before and as such represents a novel contribution with potential applications in frequency domain analysis as well as control.

The remainder of the chapter is structured as follows. Firstly, in Section 7.1 the concept of complex uncertainty is introduced and a description of the complex propagation law is given. Then in Section 7.2 the complex propagation law is applied to first order FRFs as a function of the uncertain NARX model parameters. In Section 7.3 the uncertainty is propagated into NOFRFs via the GFRF. The uncertainty propagation method is then applied to the analysis of frequency domain uncertainty in DEAs in Section 7.4. Finally, a discussion of the results of this chapter is made in Section 7.5 followed by a summary in Section 7.6.

7.1 Complex uncertainty analysis

In this section a discussion is made of how uncertainty can be approximately propagated through some non-linear function into a complex output variable. Firstly the concept of uncertainty in complex variables is introduced followed by the available methods for propagating this uncertainty.

7.1.1 Uncertainty in complex valued quantities

When considering the uncertainty associated with a real valued measurement of a system it is very common to assume that the variable is drawn from a normal probability distribution [77]. The assumption of normality allows the distribution to be defined by the statistics of the normal distribution, the mean μ and the variance σ^2. The uncertainty can then be displayed by a percentage confidence interval which defines the interval in which the measurement falls within a percentage probability, see Figure 7.1**A**.

When the measurement is drawn from a bivariate normal distribution the statistics are defined by the vector mean $\boldsymbol{\mu}$ and covariance given by

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2} \end{bmatrix} \tag{7.1}$$

where $\sigma_{1,1}$, $\sigma_{2,2}$ are the variance in the 1'st and 2'nd variate and $\sigma_{1,2} = \sigma_{2,1}$ is the covariance between the two. The mean and covariance matrix define a probability distribution in the space to of two variates that characterise the uncertainty. Analogous to the univariate measurement where the uncertainty can be displayed as a confidence interval, for the bivariate measurement a percentage confidence area is defined by the uncertainty in each variate and the correlation between them [48]. In the space of the two variates the confidence area is an ellipse.

A complex valued variable is often represented in two parts, real and imaginary (commonly plotted on an argand diagram), where

$$x = a + jb = \mathrm{Re}(x) + j\,\mathrm{Im}(x). \tag{7.2}$$

Complex variables can hence be though of as a bivariate measurement:

$$X = [\mathrm{Re}(x), \mathrm{Im}(x)]. \tag{7.3}$$

In general the variance in the real and imaginary parts of the measurement will not be independent and can therefore be assumed to be drawn from the bivariate

normal distribution

$$X \sim \mathcal{N}\left(X| \begin{bmatrix} \mu_{\mathrm{Re}(x)} \\ \mu_{\mathrm{Im}(x)} \end{bmatrix}, \Sigma_{\mathrm{Re}(x),\mathrm{Im}(x)}\right). \tag{7.4}$$

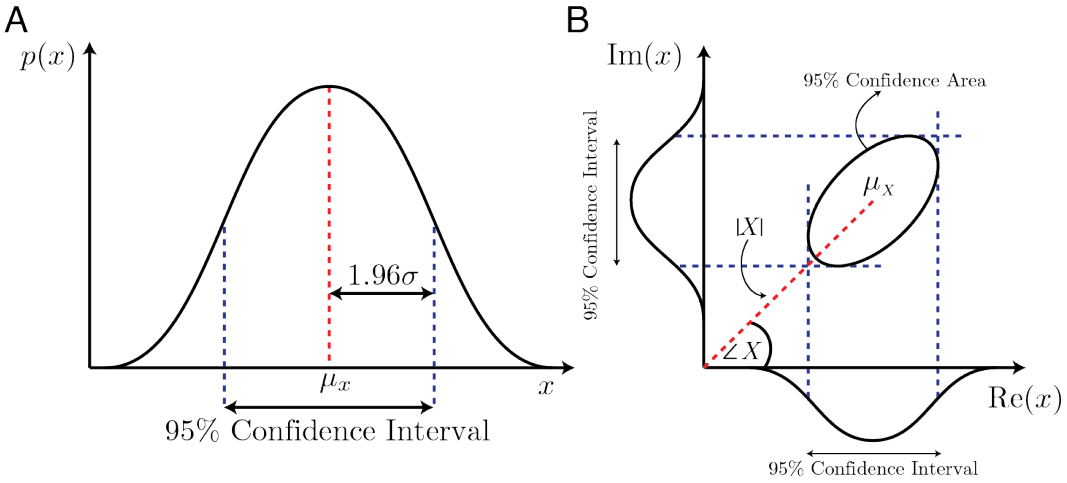

Figure 7.1: **The uncertainty in a complex variable can be represented by a bivariate normal distribution creating an elliptical uncertainty area in the real imaginary space. A**) The probability distribution of a real valued univariate variable with its 95% confidence intervals. **B**) The elliptical confidence area of a complex variable represented by a bivariate normal distribution.

The uncertainty in a complex variable is then displayed as an elliptical confidence area in the real-imaginary space, see Figure 7.1**B**. Another alternative representation for complex variables is in gain-phase form, this is very common when considering systems in the frequency domain and has been used throughout this thesis. The gain phase representation of a complex variable can also be considered as bivariate and so can be treated similarly. Care should be taken when using the gain phase representation however because the uncertainty predictions at near zero gain can be inaccurate [102].

In Chapters 3 and 4, different frequency domain descriptions, based on GFRFs and NOFRFs were introduced, that extend the linear FRF to higher order non-linear systems. The FRFs of all orders are complex valued and can be assumed to be drawn from a bivariate normal distribution. The following example demonstrates how the uncertainty is manifested in the FRF of a linear system, however the concept can easily be extended to the non-linear case.

Example 7.1 Complex uncertainty in first order FRFs.

Consider the following generative ARX model

$$y_k = \theta_1 y_{k-1} + \theta_2 y_{k-2} + \theta_3 u_{k-1} + \theta_4 u_{k-2} + e_k \tag{7.5}$$

where

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.1 \\ 0.1 \\ 0.05 \end{bmatrix}. \tag{7.6}$$

where e_k is an i.i.d white noise sequence drawn from the normal distribution $e_k \sim \mathcal{N}(e_k|0, \sigma_2^2)$ where $\sigma_e^2 = 0.0005$. The system is simulated for $N = 1000$ samples in response to the input excitation signal u_k drawn from a uniform distribution in the range $[-0.5, 0.5]$.

Parameters are estimated using VB inference by Algorithm 6.2, initialised with $a_0 = c_0 = 1 \times 10^{-2}$ and $b_0 = d_0 = 1 \times 10^{-4}$. The resulting posterior distribution on the parameters is normally distributed with mean and covariance given by

$$\theta_\mu = \begin{bmatrix} 0.1731 \\ 0.1285 \\ 0.0989 \\ 0.0520 \end{bmatrix}, \quad \Sigma_{\boldsymbol{\theta}} = \begin{bmatrix} 0.7909 & -0.3635 & -0.0010 & -0.0773 \\ -0.3635 & 0.4432 & -0.0016 & 0.0347 \\ -0.0010 & -0.0016 & 0.0063 & 0.0002 \\ -0.0773 & 0.0347 & 0.0002 & 0.0138 \end{bmatrix} \times 10^{-4}. \tag{7.7}$$

The frequency domain description of the system given by Equation (7.5) is its first order FRF,

$$H_1(\omega, \boldsymbol{\theta}) = \frac{\theta_3 e^{-i\omega} + \theta_4 e^{-2i\omega}}{1 - \theta_1 e^{-i\omega} - \theta_2 e^{-2i\omega}}. \tag{7.8}$$

The FRF, $H_1(\omega, \boldsymbol{\theta})$ is therefore a function of an uncertain variable, namely the vector $\boldsymbol{\theta}$, and therefore it is itself uncertain. Furthermore it is also known to be complex valued necessitating the use of complex uncertainty analysis.

A Monte Carlo simulation is performed on $H_1(\omega, \boldsymbol{\theta})$ by drawing $N_{MC} = 1000$ parameter values randomly from the multivariate normal distribution $\mathcal{N}(\boldsymbol{\theta}|\mu_{\boldsymbol{\theta}}, \Sigma_{\boldsymbol{\theta}})$. The real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ are plotted against each other for each Monte Carlo sample at three different frequency values, $\omega = 5.52, 5.64$ and 5.77, see Figure 7.2. The covariance matrices between the real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ are estimated from the Monte Carlo samples, assuming a bivariate normal distribution, and are given by

$$\Sigma_{H_1(\boldsymbol{\theta},\omega=5.52)} = \begin{bmatrix} 0.1193 & 0.0157 \\ 0.0157 & 0.1399 \end{bmatrix} \times 10^{-4}, \quad \Sigma_{H_1(\boldsymbol{\theta},\omega=5.52)} = \begin{bmatrix} 0.1295 & 0.0017 \\ 0.0017 & 0.1698 \end{bmatrix} \times 10^{-4},$$

$$\Sigma_{H_1(\boldsymbol{\theta},\omega=5.52)} = \begin{bmatrix} 0.1169 & 0.0100 \\ 0.0100 & 0.2287 \end{bmatrix} \times 10^{-4}. \tag{7.9}$$

The covariance matrices are used to calculate 90%, 95% and 99% confidence intervals which are overlaid on the sampled FRF. The real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ are also plotted against each other at all frequencies for the true parameter vector.

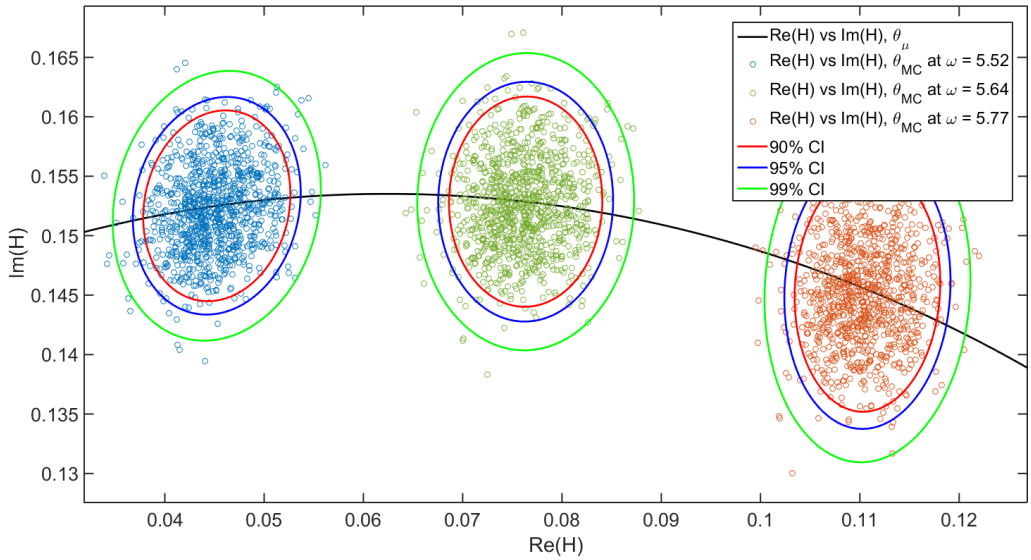

Figure 7.2: **The uncertainty associated with the Frequency response function can be described by the covariance between its real and imaginary parts.** Scatter plots of the real vs imaginary parts of H_1 calculated from Monte Carlo samples of the parameter distribution at $\omega = 5.52, 5.64$ and 5.77 plotted with confidence bounds and the real vs imaginary parts of H_1 at the true parameter vector.

The Figure shows that in this case the approximation that $H_1(\omega, \boldsymbol{\theta})$ can be described by a normal distribution is a good one. The distributions are also skewed, indicating that it is necessary to consider covariances between the real and imaginary parts. It is also clear that the uncertainty in the FRF is a function of the frequency.

△

The above example provides a motivation for the uncertainty analysis presented in the remainder of this chapter. It shows that there is covariance between the real and imaginary parts in the FRF and that these depend on the uncertainty in the model parameters as well as on the frequency.

7.1.2 Classical uncertainty propagation

Propagation of uncertainty is the calculation of the uncertainty associated with the output of some function by considering the uncertainty in its input variables and how these *propagate* through the equation. A discussion of the classical treatment of uncertainty propagation follows.

Firstly consider a function f which is a linear combination of p variables $x_1, x_2, \ldots, x_p$ with coefficients $a_1, a_2, \ldots, a_p$, such that

$$y = f(x_1, x_2, \ldots, x_p) = \sum_{i=1}^{p} a_i x_i = \boldsymbol{ax} \tag{7.10}$$

where the uncertainty associated with the input variables is described by the covariance matrix

$$\Sigma^{\mathbf{x}} = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_{1,2}} & \cdots & \sigma_{x_{1,p}} \\ \sigma_{x_{2,1}} & \sigma_{x_2}^2 & \cdots & \sigma_{x_{2,p}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_{p,}} & \sigma_{x_{p,2}} & \cdots & \sigma_{x_p}^2 \end{pmatrix}. \tag{7.11}$$

The variance of the output variable, y, is then given by [53]

$$\sigma_y^2 = \sum_{i}^{p} \sum_{j}^{p} a_i \Sigma_{ij}^{\mathbf{x}} a_j \tag{7.12}$$

In the general case, f is allowed to take the form of some non-linear combination of $x_1, x_2, \ldots, x_p$ and a linearisation of f has to be performed (except in some special cases where the variance can be calculated exactly, see for example [43]). The classical propagation law is found by approximating f by a first order Taylor expansion

$$y \approx f_0 + \sum_{i=1}^{p} \frac{\partial f}{\partial x_i} x_i \tag{7.13}$$

which is valid only when the uncertainties associated with the input variables are small enough so that they satisfy the local linear approximation.

The variance of the non-linear function $f(x_1, x_2, \ldots, x_p)$ can be found from Equation (7.12) with

$$a_i = \frac{\partial f}{\partial x_i} \tag{7.14}$$

producing the classical law of uncertainty propagation

$$\sigma_y^2 = \sum_{i=1}^{n} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Sigma_{i,j}^{\mathbf{x}} \frac{\partial f}{\partial x_j}. \tag{7.15}$$

This method is used extensively for calculating errors in scientific measurements as recommended in *the Guide to the Expression of Uncertainty in Measurement* [53].

7.1.3 Multivariate uncertainty propagation

As shall be seen in Section 7.1.4, when the output variable is complex valued it is necessary to consider a multivariate form of the propagation law. The multivariate method allows for the estimation of uncertainty for multiple output variables simultaneously as well as the correlations between them.

The function f considered in the previous section is modified so that it is now a vector function denoted $\boldsymbol{f}$ of length q with a real valued vector output, $\boldsymbol{y}$, such that

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \langle f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_q(\boldsymbol{x}) \rangle \tag{7.16}$$

The multivariate uncertainty propagation equation is then given by [47, 102, 132]

$$\Sigma^y = \boldsymbol{J}\Sigma^x \boldsymbol{J}^T \tag{7.17}$$

where Σ^y is the covariance matrix for the output vector $\boldsymbol{y}$ and $\boldsymbol{J}$ is the Jacobian matrix given by

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_p} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_q}{\partial x_1} & \frac{\partial f_q}{\partial x_2} & \cdots & \frac{\partial f_q}{\partial x_p} \end{bmatrix} \tag{7.18}$$

Note that for a scalar output, Equation (7.17) reduces to the classical case given by (7.15). The multivariate propagation law is based on the same approximations of local linearity as in the univariate case and therefore shares the same limitations.

7.1.4 Propagation of uncertainty in complex valued variables

In order to incorporate the uncertainty analysis of complex valued variables a bivariate form of the propagation law is considered. The propagation of Uncertainty in complex valued variables is developed in [47, 102, 132]. In these works both the input and output variables of the function are complex valued. Here the input variables will be the real valued parameters of the NARX model and only the output variable will be complex valued. The discussion presented here is based on the referenced work but considering a complex valued output only.

The complex valued output to some vector function f can be represented as the vector $\boldsymbol{y} = [y_1, y_2]$, where y_1 and y_2 represent the real and imaginary components of the output, such that

$$\boldsymbol{y} = \langle f_1(\boldsymbol{x}), f_2(\boldsymbol{x}) \rangle \tag{7.19}$$

where the functions $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ map the input vector $\boldsymbol{x}$ into the real and imaginary part of the output respectively.

The variance propagation equation is hence given by (7.17) where $\boldsymbol{J}$ is a $[2 \times p]$ Jacobian matrix given by

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_p} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_p} \end{bmatrix} \tag{7.20}$$

The covariance matrix representing the uncertainty associated with the input variables remains unchanged (*i.e.* is given by equation (7.22)).

This method therefore allows uncertainty to be approximately propagated through a non-linear function of multiple input variables into a complex output. It is therefore in the correct form to approximate the uncertainty in the complex valued FRF as a function of uncertain model parameters. The approach is adopted in the following section.

7.2 Application of complex uncertainty propagation to model based frequency domain analysis.

In this section, variance propagation into a complex valued output variable, outlined in the previous section, is applied to the FRFs calculated from polynomial NARX models. The uncertainty in the model is assumed to be completely characterised by the uncertainty in the model parameters represented by a multivariate normal distribution. The model based frequency domain analysis techniques introduced in Chapters 3 and 4 make use of both linear and higher order frequency response functions which are also functions of the uncertain ARX/NARX model parameters. Uncertainty can therefore be propagated into the complex valued FRFs by considering the uncertainty associated with the model parameters.

7.2.1 Uncertainty propagation from linear ARX models into the real and imaginary parts of FRFs

The ARX model is defined by Equation (3.8) and is repeated here for clarity of

$$y_k = \sum_{r=1}^{n_y} \theta_r y_{k-r} + \sum_{s=1}^{n_u} \theta_{n_y+s} u_{k-s} \tag{7.21}$$

where the model uncertainty is characterised by the uncertainty in the parameter vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_M]^T$, drawn from the $M = n_u + n_y$ dimensional multivariate normal distribution $\boldsymbol{\theta} \sim \mathcal{N}(\mu_{\boldsymbol{\theta}}, \Sigma_{\boldsymbol{\theta}})$ with mean $\mu_{\boldsymbol{\theta}}$ and covariance

$$\Sigma^{\boldsymbol{\theta}} = \begin{pmatrix} \sigma^2_{\theta_1} & \sigma_{\theta_{1,2}} & \cdots & \sigma_{\theta_{1,M}} \\ \sigma_{\theta_{2,1}} & \sigma^2_{\theta_2} & \cdots & \sigma_{\theta_{2,M}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\theta_{M,1}} & \sigma_{\theta_{M,2}} & \cdots & \sigma^2_{\theta_M} \end{pmatrix}. \tag{7.22}$$

The Frequency response function for the ARX model is given by

$$H_1(\omega, \boldsymbol{\theta}) = \frac{\sum_{s=1}^{n_u} \theta_{n_y+s} e^{-j\omega s}}{1 - \sum_{r=1}^{n_y} \theta_r e^{-j\omega r}}. \tag{7.23}$$

where the dependence on theta has been shown explicitly. Using Equation (7.19) and noting that $H_1(\omega, \boldsymbol{\theta})$ is complex valued, Equation (7.23) can be written as

$$H_1(\omega, \boldsymbol{\theta}) = \boldsymbol{H}_1(\omega, \boldsymbol{\theta}) \begin{bmatrix} 1 \\ j \end{bmatrix} \tag{7.24}$$

where

$$\boldsymbol{H}_1(\omega, \boldsymbol{\theta}) = \langle \mathrm{Re}(H_1(\omega, \boldsymbol{\theta})), \mathrm{Im}(H_1(\omega, \boldsymbol{\theta})) \rangle \tag{7.25}$$

Equation (7.25) is now in the form of which we can apply the multivariate uncertainty propagation law given by Equation (7.17) such that

$$\Sigma_{\mathrm{Re}(H_1),\mathrm{Im}(H_1)} = J(\omega, \boldsymbol{\theta}) \Sigma_{\boldsymbol{\theta}} J(\omega, \boldsymbol{\theta})^T, \tag{7.26}$$

where

$$\boldsymbol{J}(\omega, \boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial \mathrm{Re}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_1} & \frac{\partial \mathrm{Re}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_2} & \cdots & \frac{\partial \mathrm{Re}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_M} \\ \frac{\partial \mathrm{Im}(H_1(\omega,\boldsymbol{\theta}))_2}{\partial \theta_1} & \frac{\partial \mathrm{Im}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_2} & \cdots & \frac{\partial \mathrm{Im}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_M} \end{bmatrix}. \tag{7.27}$$

Differentiating Equation (7.25) with respect to the m'th model parameter

$$\frac{\partial H_1(\omega, \boldsymbol{\theta})}{\partial \theta_m} = \frac{\partial \mathrm{Re}(H_1(\omega, \boldsymbol{\theta}))}{\partial \theta_m} + i \frac{\partial \mathrm{Im}(H_1(\omega, \boldsymbol{\theta}))}{\partial \theta_m} \tag{7.28}$$

it can be inferred that

$$\mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_m}\right) = \frac{\partial\,\mathrm{Re}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_m}, \tag{7.29}$$

$$\mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_m}\right) = \frac{\partial\,\mathrm{Im}(H_1(\omega,\boldsymbol{\theta}))}{\partial \theta_m}. \tag{7.30}$$

Substituting equations (7.30) and (7.29) into Equation (7.27) produces the simpler form of the Jacobian matrix

$$\boldsymbol{J}(\omega,\boldsymbol{\theta}) = \begin{bmatrix} \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_1}\right) & \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_2}\right) & \cdots & \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_M}\right) \\ \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})_2}{\partial \theta_1}\right) & \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_2}\right) & \cdots & \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_M}\right) \end{bmatrix}. \tag{7.31}$$

It is simple to construct a general formula for the partial derivatives required in the evaluation of the Jacobian matrix because of the simple form of the first order FRF.

$$\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_m} = \begin{cases} \frac{e^{-j\omega m}\sum_{s=1}^{nu}\theta_{ny+s}e^{-j\omega s}}{\left(1-\sum_{r=1}^{ny}\theta_r e^{-j\omega r}\right)^2}, & \text{if } m = 1,2,\ldots,n_y, \\ \frac{e^{-j\omega m}}{1-\sum_{r=1}^{ny}\theta_r e^{-j\omega r}}, & \text{if } m = ny+1, ny+2\ldots,n_y+n_u. \end{cases} \tag{7.32}$$

Example 7.2 Uncertainty propagation from a linear ARX model into the real and imaginary parts of the first order FRF

Consider again the generative system described in example 7.1 given by Equation (7.8) with first order FRF given by Equation (7.5). The variance is propagated into the FRF using equation (7.48) where $\Sigma_{\boldsymbol{\theta}}$ is given by Equation (7.7) and

$$\boldsymbol{J}(\boldsymbol{\omega},\boldsymbol{\theta}) = \begin{bmatrix} \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_1}\right) & \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_2}\right) & \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_3}\right) & \mathrm{Re}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_4}\right) \\ \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_1}\right) & \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_2}\right) & \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_3}\right) & \mathrm{Im}\left(\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_4}\right) \end{bmatrix} \tag{7.33}$$

where

$$\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_1} = \frac{e^{-j\omega}(\theta_3 e^{-1j\omega} + \theta_4 e^{-2j\omega})}{(1-\theta_1 e^{-j\omega} - \theta_2 e^{-2j\omega})^2}, \quad \frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_2} = \frac{e^{-2j\omega}(\theta_3 e^{-1j\omega} + \theta_4 e^{-2j\omega})}{(1-\theta_1 e^{-j\omega} - \theta_2 e^{-2j\omega})^2}$$

$$\frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_3} = \frac{e^{-j\omega}}{1-\theta_3 e^{-j\omega} - \theta_2 e^{-2j\omega}}, \quad \frac{\partial H_1(\omega,\boldsymbol{\theta})}{\partial \theta_4} = \frac{e^{-2j\omega}}{1-\theta_1 e^{-j\omega} - \theta_2 e^{-2j\omega}}$$

The covariance of the real and imaginary parts of $H_1(\omega,\boldsymbol{\theta})$ can then be calculated at each value of ω. The sampled covariance matrix of H_1 was calculated

previously in example 7.1 and can be used to provide a comparison. The variance on the real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ show a very good fit to the sampled variance. The Monte Carlo samples on the real and imaginary parts are shown in grey with the sampled variance shown by the blue dotted line. The propagated variance, shown by the red dotted line, overlays the sampled variance for both the real and imaginary parts, see Figure 7.3**A-B**.

A Nyquist plot can be constructed by plotting the real versus the imaginary part of $H_1(\omega, \boldsymbol{\theta})$ from $\omega = 0$ to 2π. Considering mean and covariance of $H_1(\omega, \boldsymbol{\theta})$ at each value of ω, the area of confidence in which the true value of $\partial H_1(\omega, \boldsymbol{\theta})$ lies within some probability bound can be plotted where the 99% confidence area is shown in grey, see Figure 7.3**C**. The sampled distribution of $H_1(\omega, \boldsymbol{\theta})$ at each value of ω is well approximated by the propagated variance for the sampled and estimated distributions at two values of ω, see Figure 7.3**D**. △

7.2.2 Propagation of uncertainty into the gain phase frequency description

The real-imaginary frequency description in the previous section is commonly used for systems stability analysis, however, the gain-phase description is widely used in many other systems analysis methods. To propagate uncertainty into the gain and phase of $H_1(\omega, \boldsymbol{\theta})$ two approaches can be taken. The first method is to first propagate the uncertainty into the real and imaginary parts for all ω following the process in Section 7.2.1 and then to reapply the multivariate propagation law, Equation (7.17), such that [109]

$$\Sigma_{|H_1|,\angle H_1} = \boldsymbol{J}(\omega, \boldsymbol{\theta}) \Sigma_{\mathrm{Re}(H_1),\mathrm{Im}(H_1)} \boldsymbol{J}(\omega, \boldsymbol{\theta})^T \tag{7.34}$$

where the explicit dependence of H_1 on ω and $\boldsymbol{\theta}$ has been dropped for clarity of reading. The Jacobian is given by

$$\boldsymbol{J}(\omega, \boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial |H_1|}{\partial \mathrm{Re}\, H_1} & \frac{\partial |H_1|}{\partial \mathrm{Im}\, H_1} \\ \frac{\partial \angle H_1}{\partial \mathrm{Re}\, H_1} & \frac{\partial \angle H_1}{\partial \mathrm{Im}\, H_1} \end{bmatrix} = \begin{bmatrix} \frac{\mathrm{Re}(H_1)}{\sqrt{\mathrm{Re}(H_1)^2 + \mathrm{Im}(H_1)^2}} & \frac{\mathrm{Im}(H_1)}{\sqrt{\mathrm{Re}(H_1)^2 + \mathrm{Im}(H_1)^2}} \\ -\frac{\mathrm{Im}(H_1)}{\mathrm{Re}(H_1)^2 + \mathrm{Im}(H_1)^2} & \frac{\mathrm{Re}(H_1)}{\mathrm{Re}(H_1)^2 + \mathrm{Im}(H_1)^2} \end{bmatrix}. \tag{7.35}$$

The second method allows the uncertainty to be propagated directly into the gain and phase without the intermediate step. This is achieved by again using the multivariate propagation law with

$$\Sigma_{|H_1|,\angle H_1} = \boldsymbol{J}(\omega, \boldsymbol{\theta}) \Sigma_{\boldsymbol{\theta}} \boldsymbol{J}(\omega, \boldsymbol{\theta})^T \tag{7.36}$$

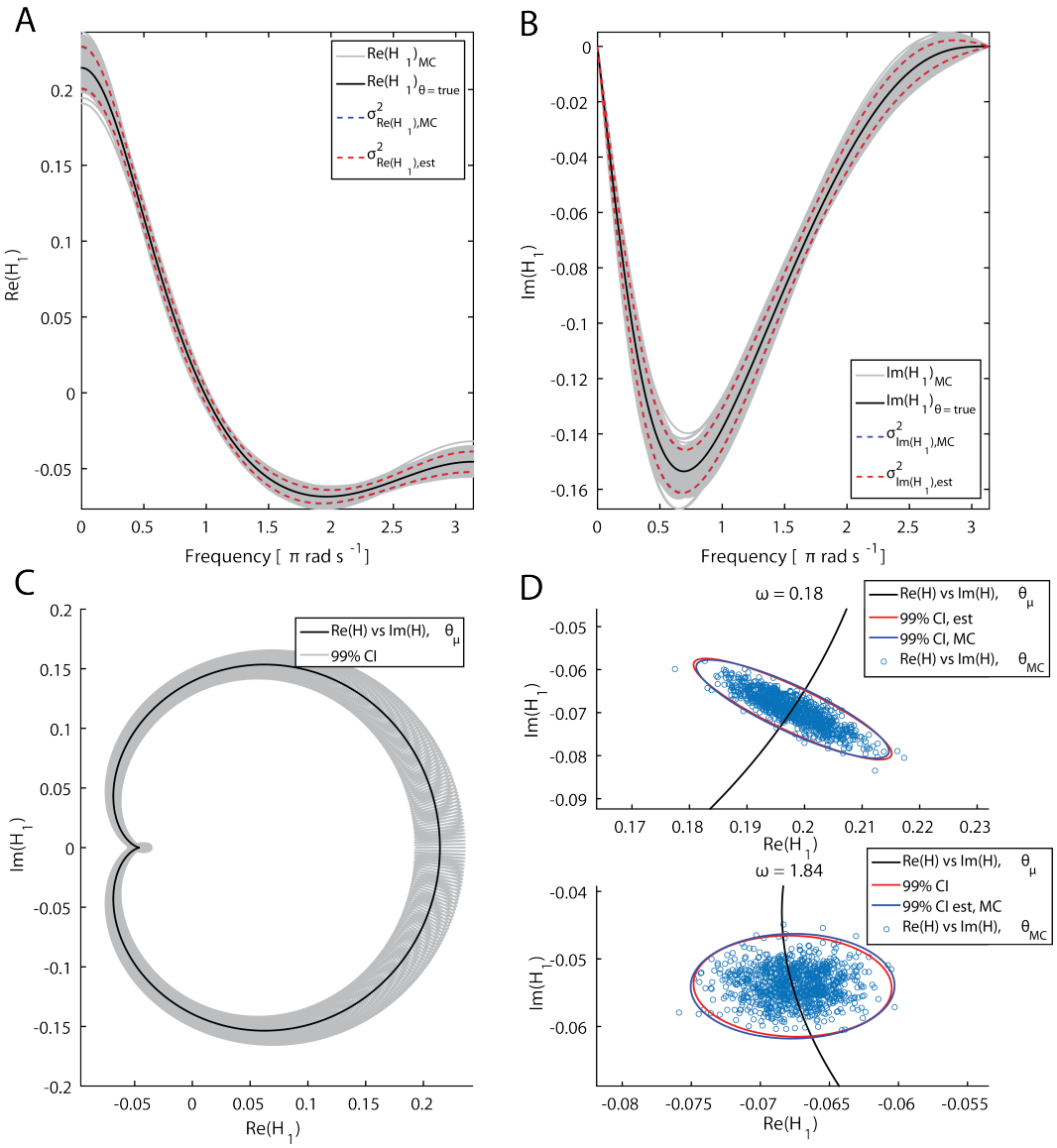

Figure 7.3: **Propagating uncertainty into the frequency domain as a FRF by considering the covariance in the ARX model parameters provides an accurate estimate in comparison to the sampled covariance.** The uncertainty in the real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$, given by Equation (7.5), with the ARX model parameter covariance given by Equation (7.7) **A**,**B**) The real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ plotted against frequency for the true parameter values (Black) with the Monte Carlo samples of $H_1(\omega, \boldsymbol{\theta})$ (Grey) with sampled variance (Blue dashed) and propagated variance (Red dashed). **C**) Nyquist plot for $H_1(\omega, \boldsymbol{\theta})$ (Black) with 99% confidence area (Grey). **D**) Monte Carlo sampled distribution of $H_1(\omega, \boldsymbol{\theta})$ at $\omega = 0.18$ and 1.84 with the sampled (Blue) and propagated (Red) confidence areas.

where

$$\boldsymbol{J}(\omega,\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial|H_1|}{\partial\theta_1} & \frac{\partial|H_1|}{\partial\theta_2} & \cdots & \frac{\partial|H_1|}{\partial\theta_p} \\ \frac{\partial\angle H_1}{\partial\theta_1} & \frac{\partial\angle H_1}{\partial\theta_2} & \cdots & \frac{\partial\angle H_1}{\partial\theta_p} \end{bmatrix}. \tag{7.37}$$

Generic expressions for the partial differentials in Equation (7.37) can be found in terms of the partial derivatives of H_1 with respect to the relevant parameter θ_i as follows:

Partial derivative of the gain of H_1 with respect to the m'th model parameter

Noting that

$$\mathrm{Re}(H_1) = \frac{1}{2}(H_1 + \overline{H_1}), \quad \mathrm{Im}(H_1) = \frac{1}{2j}(H_1 - \overline{H_1}) \tag{7.38}$$

where $\overline{H_1}$ represents the complex conjugate of H_1, and

$$\frac{\partial\,\mathrm{Re}(H_1)}{\partial\theta_m} = \frac{1}{2}\left(\frac{\partial\,\mathrm{Re}(H_1)}{\partial\theta_m} + \frac{\partial\,\mathrm{Re}(\overline{H_1})}{\partial\theta_m}\right), \quad \frac{\partial\,\mathrm{Im}(H_1)}{\partial\theta_m} = \frac{1}{2j}\left(\frac{\partial\,\mathrm{Im}(H_1)}{\partial\theta_m} - \frac{\partial\,\mathrm{Re}(\overline{H_1})}{\partial\theta_m}\right) \tag{7.39}$$

the partial derivative can be found by

$$\begin{aligned} \frac{\partial|H_1|}{\partial\theta_m} &= \frac{\partial}{\partial\theta_m}(H_1\overline{H_1})^{\frac{1}{2}} \\ &= \frac{\partial}{\partial H_1}(H_1\overline{H_1})^{\frac{1}{2}}\frac{\partial H_1}{\partial\theta_m} + \frac{\partial}{\partial\overline{H_1}}(H_1\overline{H_1})^{\frac{1}{2}}\frac{\partial\overline{H_1}}{\partial\theta_m} \\ &= -\frac{1}{2}\overline{H_1}(H_1\overline{H_1})^{\frac{1}{2}}\frac{\partial H_1}{\partial\theta_m} - \frac{1}{2}H_1(H_1\overline{H_1})^{\frac{1}{2}}\frac{\partial\overline{H_1}}{\partial\theta_m} \\ &= \frac{1}{2|H_1|}\left(\overline{H_1}\frac{\partial H_1}{\partial\theta_m} + H_1\frac{\partial\overline{H_1}}{\partial\theta_m}\right) \\ &= \frac{1}{|H_1|}\,\mathrm{Re}\left(\overline{H_1}\frac{\partial H_1}{\partial\theta_m}\right). \end{aligned} \tag{7.40}$$

Partial derivative of the phase of H_1 with respect to the i'th model parameter

Taking the derivative of the angle of H_1

$$\frac{\partial\angle H_1}{\partial\theta_m} = \frac{\partial}{\partial X}\arctan(X)\frac{\partial X}{\partial\theta_m}, \quad \text{where} \quad X = \frac{\mathrm{Im}(H_1)}{\mathrm{Re}(H_1)} \tag{7.41}$$

The derivative of X is given by

$$\begin{aligned}\frac{\partial X}{\partial \theta_m} &= \frac{\partial}{\partial \operatorname{Im}(H_1)}\left(\frac{\operatorname{Im}(H_1)}{\operatorname{Re}(H_1)}\right)\frac{\partial \operatorname{Im}(H_1)}{\partial \theta_m} + \frac{\partial}{\partial \operatorname{Re}(H_1)}\left(\frac{\operatorname{Im}(H_1)}{\operatorname{Re}(H_1)}\right)\frac{\partial \operatorname{Re}(H_1)}{\partial \theta_m} \\ &= \frac{1}{\operatorname{Re}(H_1)}\frac{\partial \operatorname{Im}(H_1)}{\partial \theta_m} - \frac{\operatorname{Im}(H_1)}{\operatorname{Re}(H_1)^2}\frac{\partial \operatorname{Re}(H_1)}{\partial \theta_m}. \end{aligned} \tag{7.42}$$

Employing equations (7.38) and (7.39) in Equation (7.42)

$$\frac{\partial X}{\partial \theta_m} = \frac{1}{j}\left(\frac{\partial H_1}{\partial \theta_m} - \frac{\partial \overline{H_1}}{\partial \theta_m}\right)\frac{H_1 + \overline{H_1}}{(H_1 + \overline{H_1})^2} - \frac{1}{j}\left(\frac{\partial H_1}{\partial \theta_m} + \frac{\partial \overline{H_1}}{\partial \theta_m}\right)\frac{H_1 - \overline{H_1}}{(H_1 + \overline{H_1})^2} \tag{7.43}$$

where the term on the left hand side has been multiplied by a factor of $H_1 + \overline{H_1}$ in the numerator and denominator. Collecting terms

$$\begin{aligned}\frac{\partial X}{\partial \theta_m} &= \frac{2}{j(H_1 + \overline{H_1})^2}\left(\overline{H_1}\frac{\partial H_1}{\partial \theta_m} - H_1\frac{\partial \overline{H_1}}{\partial \theta_m}\right) \\ &= \frac{4}{(H_1 + \overline{H_1})^2}\operatorname{Im}\left(\overline{H}\frac{\partial H_1}{\partial \theta_m}\right) \\ &= \frac{1}{\operatorname{Re}(H_2)^2}\operatorname{Im}\left(\overline{H}\frac{\partial H_1}{\partial \theta_m}\right). \end{aligned} \tag{7.44}$$

The derivative of $\arctan(X)$ is given by the identity

$$\begin{aligned}\frac{\partial}{\partial X}\arctan(X) &= \frac{1}{1 + X^2} \\ &= \frac{1}{1 + \left(\frac{\operatorname{Im}(H_1)}{\operatorname{Re}(H_1)}\right)^2} \\ &= \frac{\operatorname{Re}(H_1)^2}{|H_1|^2}. \end{aligned} \tag{7.45}$$

The general solution is then found by substituting equations (7.44) and (7.45) into Equation (7.42) such that.

$$\frac{\partial \angle H_1}{\partial \theta_m} = \frac{1}{|H_1|^2}\operatorname{Im}\left(\overline{H_1}\frac{\partial H_1}{\partial \theta_m}\right). \tag{7.46}$$

The partial differentials of the gain and phase of $H_1(\omega, \boldsymbol{\theta})$ can therefore be expressed compactly by the equations (7.40) and (7.46) which depend solely on the partial derivative of $H_1(\omega, \boldsymbol{\theta})$ with respect to the parameters and on $H_1(\omega, \boldsymbol{\theta})$ itself. The task of propagating the uncertainty into the gain and phase of the FRF therefore requires the calculation of the same differentials as in the real-imaginary

case. Both methods are able to provide good estimates of the variance in the gain and phase of the first order FRF

A further example is given to demonstrate the propagation of uncertainty into the gain and phase.

Example 7.3 Uncertainty propagation from a linear ARX model into the gain and phase of the first order FRF

In this example the use of the two methods for propagating the uncertainty in ARX model parameters into the gain and phase of the first order FRF are illustrated. The example system described in example 7.1 is again used to demonstrate the methods.

The first indirect method propagates the models parameter uncertainty first through the real and imaginary parts and then into the gain and phase. The covariance, $\Sigma_{\mathrm{Re}(H_1),\mathrm{Im}(H_1)}$, of the real and imaginary parts of $H_1(\omega, \boldsymbol{\theta})$ was found in example 7.2 for each value of ω. The covariance of the gain and phase can then be calculated by a further use of the multivariate propagation law given by Equation (7.17) where the Jacobian matrix is given by Equation (7.35) and $\Sigma_x = \Sigma_{\mathrm{Re}(H_1),\mathrm{Im}(H_1)}$.

The second method propagates the variance directly and can be performed using equations (7.34), (7.37, (7.40) and (7.46) where the covariance of the parameters are given in example 7.7 and the partial derivatives of $H_1(\omega, \boldsymbol{\theta})$ are found in example 7.2.

The covariance estimates found using the two methods are identical to working precision and so the results are shown together in Figure 7.4**A**-**D**. The sampled variance of the gain and phase of the system are calculated from the MC samples of $H_1(\omega, \boldsymbol{\theta})$ in order to provide a comparison. The propagated variance is a very close match to the sampled variance for both the gain and phase. The propagated variance (Red dashed line) overlays the sample variance (Blue dashed line) for both cases, see Figure 7.4**A**,**B**. The gain vs phase is plotted for all frequencies along with the 99% confidence area representing the system uncertainty, see Figure 7.4**C**. The distribution of the sampled $H_1(\omega, \boldsymbol{\theta})$ at $\omega = 0.18$ and 1.84 show the close fit between the covariance calculated from the sampled data to the propagated variance, see Figure 7.4**D**.

△

The example shows how either of the proposed methods can be used to propagate the variance into the gain and phase of the first order FRF.

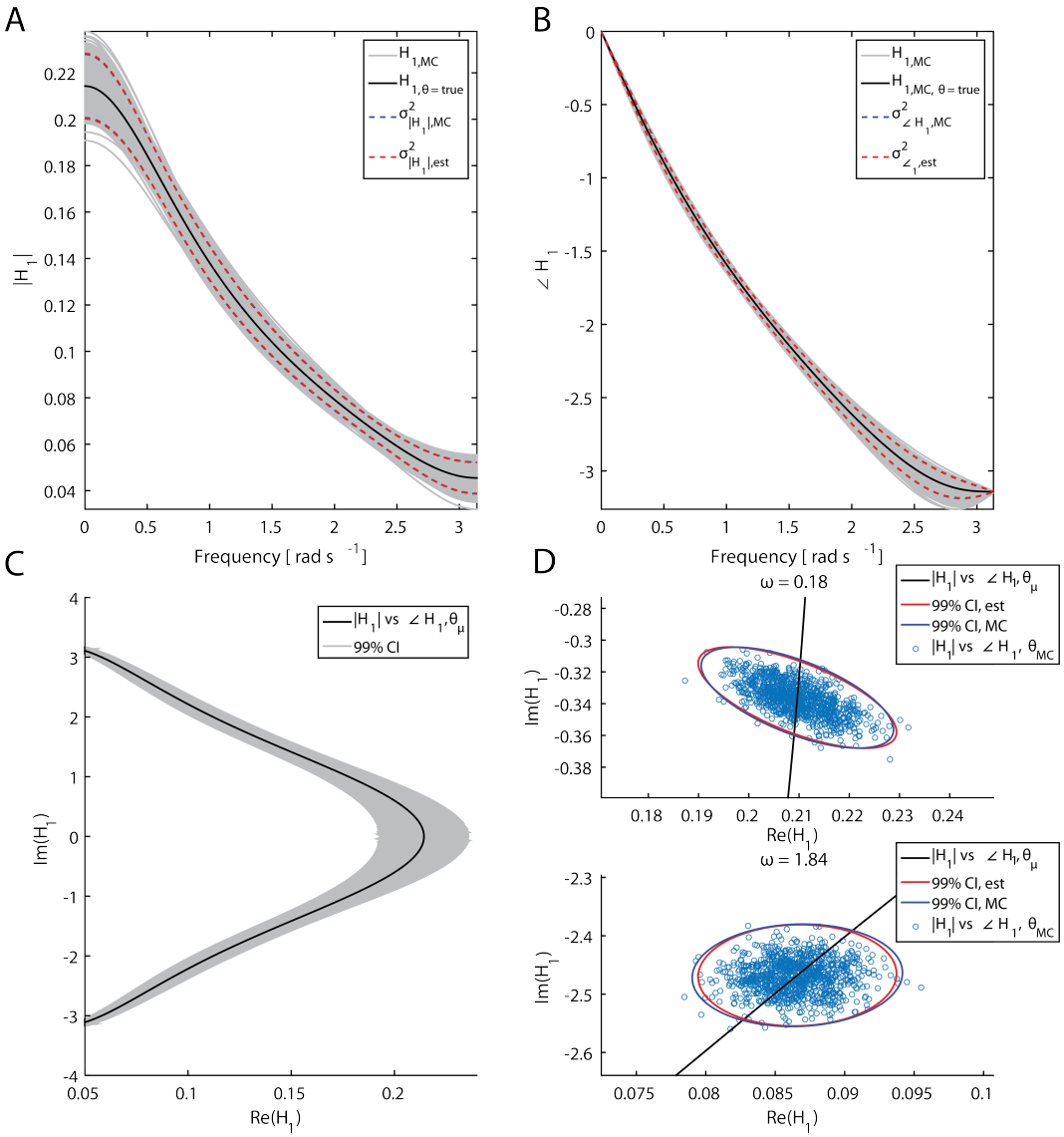

Figure 7.4: **Propagating uncertainty into the frequency domain as the gain and phases of the first order FRF by considering the covariance in the ARX model parameters provides an accurate estimate in comparison to the sampled covariance.** The uncertainty in the gain and phase of $H_1(\omega, \boldsymbol{\theta})$, given by Equation (7.5), with the ARX model parameter covariance given by Equation (7.7) **A**,**B**) The gain and phase of $H_1(\omega, \boldsymbol{\theta})$ plotted against frequency for the true parameter values (Black) with the Monte Carlo samples of $H_1(\omega, \boldsymbol{\theta})$ (Grey) with sampled variance (Blue dashed) and propagated variance (Red dashed). **C**) Gain vs phase plot for $H_1(\omega, \boldsymbol{\theta})$ (Black) with 99% confidence area (Grey). **D**) Monte Carlo sampled distribution of $H_1(\omega, \boldsymbol{\theta})$ at $\omega = 0.18$ and 1.84 with the sampled (Blue) and propagated (Red) confidence areas.

7.3 Uncertainty propagation into higher order frequency response functions of non-linear systems.

The uncertainty propagation method can readily be applied to higher order frequency response functions in the form of GFRFs. The propagation law remains unchanged and can be applied directly to the GFRFs. However, the dimensionality of the GFRFs makes visualising the associated uncertainty difficult even at the second order FRF because there is a covariance matrix associated with each input frequency combination. NOFRFs have the same dimensionality as the linear FRF and so provide a more transparent frequency domain description for expressing the uncertainty. In this section uncertainty propagation will be used to estimate the uncertainty in the n'th order output spectra of a NARX model where, as with the linear case, the uncertainty originates from the model parameters.

The output spectra of a non-linear system is given by Equation (4.40) and is repeated here for clarity of reading

$$\begin{aligned} Y(j\omega) &= \sum_{n=1}^{N_m} Y_n(j\omega) \\ Y_n(j\omega) &= \frac{n^{-1/2}}{(2\pi)^{n-1}} \int_{\omega} H_n(j\omega_1, \ldots, j\omega_n) \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega} \end{aligned} \tag{7.47}$$

Employing the multivariate propagation law, given by Equation (7.17), once again, the covariance in the gain and phase of the n'th order output spectra is given by

$$\Sigma_{|Y_n|,\angle Y_n} = J(\omega, \boldsymbol{\theta}) \Sigma_{\boldsymbol{\theta}} J(\omega, \boldsymbol{\theta})^T, \tag{7.48}$$

where

$$\boldsymbol{J}(\omega, \boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial |Y_n|}{\partial \theta_1} & \frac{\partial |Y_n|}{\partial \theta_2} & \cdots & \frac{\partial |Y_1|}{\partial \theta_M} \\ \frac{\partial \angle Y_n}{\partial \theta_1} & \frac{\partial \angle Y_n}{\partial \theta_2} & \cdots & \frac{\partial \angle Y_n}{\partial \theta_M} \end{bmatrix}. \tag{7.49}$$

From equations (7.40) and (7.46) the necessary differentials in (7.49) are given by

$$\frac{\partial |Y_n|}{\partial \theta_m} = \frac{1}{|Y_n|} \mathrm{Re}\left(\overline{Y_n} \frac{\partial Y_n}{\partial \theta_m} \right), \quad \frac{\partial \angle Y_n}{\partial \theta_m} = \frac{1}{|Y_n|^2} \mathrm{Im}\left(\overline{Y_n} \frac{\partial Y_n}{\partial \theta_m} \right). \tag{7.50}$$

The partial differential of the n'th order output spectrum with respect to the parameters that appear in equations (7.50) are simply computed from Equation (7.47) by noting that only the n'th order FRF is dependent on the parameters so that

$$\frac{\partial Y_n}{\partial \theta_m} = \frac{n^{-1/2}}{(2\pi)^{n-1}} \int_{\omega} \frac{\partial H_n(j\omega_1, \ldots, j\omega_n)}{\partial \theta_m} \prod_{i=1}^{n} U(j\omega_i) \mathrm{d}\sigma_{n\omega} \tag{7.51}$$

Note from Equation (7.51) that in order to perform the uncertainty propagation, partial differentials of the higher order frequency response functions with respect to the model parameters need to be taken. Conveniently, in general the higher order FRFs are a function of the model parameters and the FRFs of order $< n$. Therefore when differentiating with respect to the m'th parameter, any lower order FRF multiplying that parameter will not be a function of it as demonstrated by the following example.

Example 7.4 Differentials of higher order FRFs
Consider the following NARX model

$$y_k = \theta_1 y_{k-1} + \theta_2 y_{k-2} + \theta_3 u_{k-1} + \theta_4 u_{k-2} + \theta_5 u_{k-1}^2 + \theta_6 y_{k-1}^2 + e_k \tag{7.52}$$

The system is equivalent to the linear system given in example 7.1 with the addition of both a pure second order input and output non-linearity. The first order FRF is unchanged by the non-linear terms and is given by Equation (7.8). The 2'nd order FRF is given by

$$H_2(\omega_1, \omega_2) = \frac{\theta_5 e^{-i\pi(\omega_1+\omega_2)} + \theta_6 H_1(\omega_1) H_1(\omega_2) e^{-i\pi(\omega_1+\omega_2)}}{1 - \theta_1 e^{-i(\omega_1+\omega_2)} - \theta_2 e^{-2i(\omega_1+\omega_2)}}. \tag{7.53}$$

Note that in general $H_1(\omega)$ cannot be a function of the parameters associated with higher order terms, in this case θ_5 and θ_6. However the denominator of (7.8) is dependent on the linear terms.

Differentiating $H_2(\omega_1, \omega_2)$ with respect to θ_1

$$\frac{\partial H_2(\omega_1, \omega_2)}{\partial \theta_1} = \frac{\theta_6 e^{-i\pi(\omega_1+\omega_2)} \left(\frac{\partial H_1(\omega_1)}{\partial \theta_1} H_1(\omega_2) + \frac{\partial H_2(\omega_1)}{\partial \theta_1} H_1(\omega_2) \right)}{1 - \theta_1 e^{-i(\omega_1+\omega_2)} - \theta_2 e^{-2i(\omega_1+\omega_2)}} \tag{7.54}$$

$$+ \frac{e^{-i\pi(\omega_1+\omega_2)} \left(\theta_5 e^{-i\pi(\omega_1+\omega_2)} + \theta_6 H_1(\omega_1) H_1(\omega_2) e^{-i\pi(\omega_1+\omega_2)} \right)}{\left(1 - \theta_1 e^{-i(\omega_1+\omega_2)} - \theta_2 e^{-2i(\omega_1+\omega_2)}\right)^2} \tag{7.55}$$

and similarly for the differential with respect to θ_2. The partial differentials of the 1'st order FRF are given by Equation (7.32). Although the differentiation is simple to perform it indicates that, in general, it is necessary to evaluate the differentials of all lower order FRFs. △

Example 7.5 Differentiation of higher order FRFs with respect to the NARX model parameters.

Consider the generative NARX model given by Equation (7.52) where

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \\ \theta_5 \\ \theta_6 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.1 \\ 0.1 \\ 0.05 \\ 0.2 \\ 0.5 \end{bmatrix}, \tag{7.56}$$

and e_k is an *i.i.d* white noise sequence drawn from the normal distribution $e_k \sim \mathcal{N}(e_k|0, \sigma_2^2)$ where $\sigma_e^2 = 0.0005$. The system is simulated for $N = 1000$ samples in response to the input excitation signal u_k drawn from a uniform distribution in the range $[-0.5, 0.5]$.

Following from example 7.1, model parameters are again estimated using VB inference by Algorithm 6.2, initialised with $a_0 = c_0 = 1 \times 10^{-2}$ and $b_0 = d_0 = 1 \times 10^{-4}$, as before. The resulting posterior distribution on the parameters is normally distributed with mean and covariance given by

$$\mu_{\boldsymbol{\theta}} = \begin{bmatrix} 0.1887 & 0.1081 & 0.0991 & 0.0489 & 0.1978 & 0.6268 \end{bmatrix}^T$$
$$\Sigma_{\boldsymbol{\theta}} = \begin{bmatrix} 0.0246 & -0.0083 & 0.0000 & -0.0016 & -0.0004 & -0.1219 \\ -0.0083 & 0.0097 & -0.0000 & 0.0008 & -0.0004 & 0.0011 \\ 0.0000 & -0.0000 & 0.0002 & -0.0000 & -0.0000 & 0.0002 \\ -0.0016 & 0.0008 & -0.0000 & 0.0003 & 0.0001 & 0.0001 \\ -0.0004 & -0.0004 & -0.0000 & 0.0001 & 0.0013 & -0.0094 \\ -0.1219 & 0.0011 & 0.0002 & 0.0001 & -0.0094 & 1.5860 \end{bmatrix} \times 10^{-2}, \tag{7.57}$$

The covariance in the NARX model parameters is propagated into the n'th order output spectrum using equations (7.48)-(7.51). The evaluation of the integral in Equation (7.51) has to be performed approximately. Two different input signals are used, firstly a signal that is uniform across all frequencies and secondly a signal with uniform frequency in the range $[0, 0.4\pi]$ and zero elsewhere. Both signals are generated using Equation (4.69). The output spectra relating to the MC sampled parameters are also calculated using Algorithm 4.1 and the variance in the gain is found for comparison to the propagated uncertainty.

The propagated variance in the gain of the n'th order output spectra is a good fit to the sampled variance for both input excitations, see Figure 7.5**A**-**B**,**D**-**E**. The output spectrum is approximated as $Y \approx Y_1 + Y_2$, assuming higher orders have a negligible contribution to the output. The variance in the output spectrum is

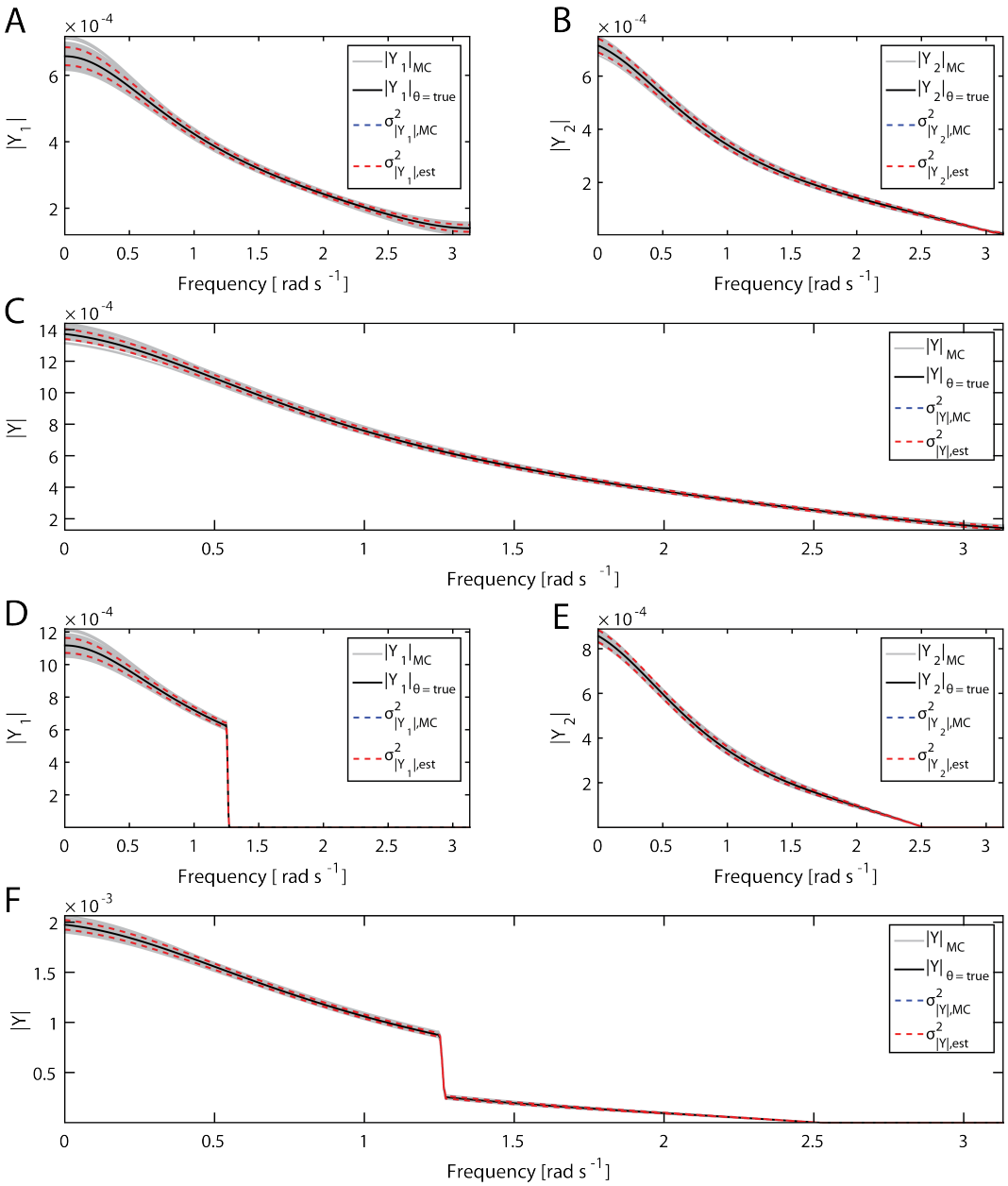

Figure 7.5: **Accurate estimates of the uncertainty in the n'th order output spectrum of non-linear systems can be made by the uncertainty propagation method.** The uncertainty in Y_n where the uncertainty originates from the covariance in the parameters given by Equation (7.57). **A**,**B**) The gain of Y_1 and Y_2 in response to a uniform input response plotted against frequency for the true parameter values (Black) with the Monte Carlo samples of Y_1 (Grey) with sampled variance (Blue dashed) and propagated variance (Red dashed). **C**) Gain of $Y \approx Y_1 + Y_2$ in response to a uniform input response **D,E,F**) As above in response to an input spectrum of uniform frequency in the range $[0, 0.4\pi]$ and zero elsewhere.

calculated as $\sigma_Y^2 = \sigma_{Y_1}^2 + \sigma_{Y_2}^2$ and similarly shows a good fit to the sampled variance, see Figure 7.5**C**,**F**. △

7.4 Frequency domain uncertainty analysis of DEAs using the propagation method.

The uncertainty propagation method developed in this chapter can be applied to DEAs in order to estimate confidence intervals for the frequency domain description. Following previous chapters the analysis is demonstrated using EAP 5. In Chapter 6 it was shown how the uncertainty in the frequency domain description of the DEAs can be found by sampling from the posterior distribution of the model parameters and mapping the sampled parameters into the frequency domain as NOFRFs. The use of the variance propagation method removes the need for sampling and the variance can be estimated directly from the model.

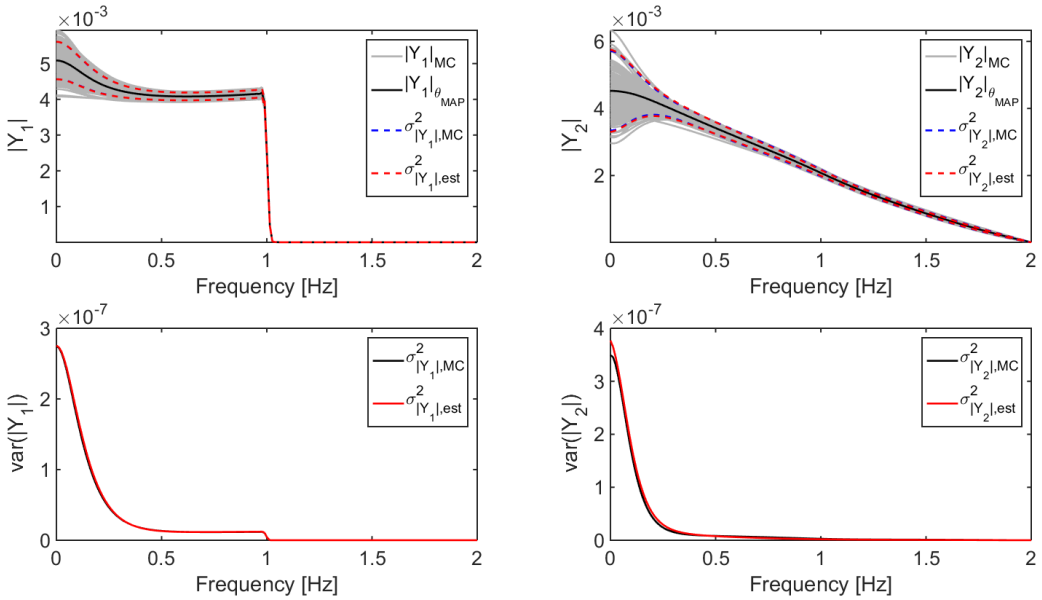

Figure 7.6: **Frequency domain confidence intervals for DEAs can be estimated accurately using the uncertainty propagation method.** First and second order output spectra for DEA 5 calculated using from the MAP parameter estimates (Black) and for MC samples of the posterior parameter distribution (Grey). 95% confidence intervals are shown calculated over the sampled outputs (Blue) and using the propagation method (Red).

First and second order GFRFs are calculated using the probing method for the NARX model of DEA 5 identified using the SVB-NARX algorithm in Chapter 6.

The variance in the first and second order NOFRFs of the system is estimated using the propagation method outlined in Section 7.3. NOFRFs for DEA 5 corresponding to the MAP estimate and for parameters sampled from the posterior parameter distribution are calculated in Chapter 6. The variance in the gain of the sampled first and second order output spectra is calculated. The propagated variance in the output spectra closely matches the variance in the gain of the sampled output spectra, see Figure 7.6. Confidence intervals are calculated for both and are seen to accurately contain the sampled output spectra.

The complex uncertainty propagation method has therefore allowed for confidence bounds to be estimated for both linear and higher order frequency domain descriptions to be calculated to a high degree of accuracy (compared to the variance of the distribution found via sampling from the posterior parameter distribution and mapping those into the frequency domain). It shows that the methodology works well in the real data case and that for the DEAs the assumption of Normality on the FRF distribution appears to be a good one. It is also further demonstrates the advantages of the NOFRF as the frequency domain description on which to display the uncertainty bounds. This is because of the low dimensionality of this representation.

7.5 Discussion

The aim of this chapter was the investigation of uncertainty in the frequency domain description of linear and non-linear systems and the development of methods for the propagation of uncertainty in both ARX and NARX model parameters into the frequency domain as uncertain FRFs. The frequency domain is commonly represented in two ways: in Cartesian co-ordinates, which can be considered as bivariate with real and imaginary parts, and in gain-phase form. Uncertainty can be considered in either form as an elliptical confidence area. In order to propagate parameter uncertainty in space of the FRF a bivariate approach is therefore required.

Uncertainty propagation is achieved using the bivariate uncertainty propagation law, mapping the variance in the model parameters into either the real and imaginary parts, or the gain and phase of the FRF. The propagation is based on the assumption that the uncertainty in the output is well approximated by a bivariate normal distribution and this is shown by example to be a good approximation in some cases. The uncertainty in the FRF of example systems is calculated using the propagation method and is shown to provide an accurate description by comparison to uncertainty estimates made by sampling from the parameter distribution.

The advantage of the uncertainty analysis methodology introduced in this section is that uncertainty in the FRFs of both linear and non-linear systems can be estimated accurately and efficiently. The uncertainty can be displayed in forms that are useful for both analysis and control design *i.e.* real-imaginary and gain-phase. In particular this could be beneficial for the purpose of robust control design. Robust control caters for plant-model mismatch by taking into account the uncertainty in the modelling process. This uncertainty is commonly represented as a percentage error in the physical parameters of some elements of the system [35]. IT is common, however, that models of the plant are identified using data driven system identification techniques. It is suggested that by following the uncertainty propagation method developed here. New methods for robust control based on the frequency domain uncertainty may be available where stability and performance can be informed by rigorous data-driven characterisation of the uncertainty in system dynamics.

In the previous chapter the uncertainty in the frequency domain descriptions of DEAs was estimated by sampling from the posterior distribution of the NARX model parameters and mapping each sampled parameter vector into the frequency domain to build up a distribution. The method developed in this chapter allows the frequency domain uncertainty to be propagated directly from the uncertainty in the uncertainty in the NARX model and therefore represents a significantly more efficient method. The uncertainty estimated by both methods is compared and the uncertainty propagation method is shown to provide an accurate estimate to the sampled distribution, see Figure 7.6.

7.6 Summary

In this chapter a novel approach to FRF based uncertainty analysis of linear and non-linear systems was developed. The method uses complex uncertainty propagation to propagate the uncertainty in the parameters of a NARX model into the linear FRF as well as higher order GFRFs or NFRFs.

The uncertainty associated with the FRF of both linear and non-linear systems is investigated. It is shown by example how the probability distribution of the FRF is well approximated by the bivariate normal distribution, where the two variates consist of the real and imaginary part of the signal.

The concept of uncertainty propagation is introduced in both the univariate and bivariate case and is then applied to the application of complex uncertainty propagation. General formula for the differentials of the gain and phase of a FRF with respect to the NARX model parameters are given. This allows for the

parameter uncertainty to be directly propagated into the uncertainty of the gain and phase.

The complex uncertainty propagation method is then applied to higher order FRFs in the form of NOFRFs by considering the differentials of the GFRFs with respect to the NARX model parameters. This allows the uncertainty in the higher order FRFs and NOFRFs to be estimated.

Chapter 8

Conclusions

8.1 Conclusions and Summary

Previous investigations into the dynamic behaviour of DEAs have focused on the use of physical first principles models. The potential widespread application of these actuators as soft actuation devices motivates the development of systems level modelling and analysis framework presented in this thesis. The framework allows for the data driven identification of compact and accurate models of DEAs in a form that is advantageous for control design. The NARMAX model structure in which the identification is performed allows for the mapping of the NARMAX model parameters into the frequency domain as GFRFs. GFRF and subsequent NOFRF analysis provides a frequency domain description of the DEA system that allows for a clearer understanding of the underlying system behaviour than analysing the model equations. The frequency domain also allows for a description of the system that is unique, unlike the time domain model, which may be dependent on the method used for identification.

Higher order frequency response based analysis can therefore be used to directly compare different DEAs, a comparison that is not possible by considering the time domain model equations alone. NOFRF analysis is of particular importance for performing these comparisons because the system NOFRFs are a function of only one frequency variable and can hence be displayed in two dimensions. The results of the NOFRF analysis clearly demonstrate that the actuators display different underlying system behaviours despite being fabricated to the same specifications.

As well as displaying variations in behaviour across actuators, time varying behaviour is also observed in each of the DEAs. The observed time variations cannot be explained by current models of DEAs. However, the characterisation of

this behaviour is essential in attempting to understand DEA dynamics, improve fabrication techniques, perform control design and perform fault diagnosis. It is shown by example that the DEA system can be well described by a single model structure with time varying parameters. The form of the model structure, *i.e.* that it is linear-in-the-parameters allows for recursive parameter estimation techniques to be used, vastly simplifying the estimation of the time varying parameters and providing a characterisation of the time varying behaviour.

In order to analyse the time varying system a novel time varying frequency domain analysis method is proposed and successfully demonstrated. The new method, TV-NOFRF, decomposes the time domain model into separate time invariant models at each sample time. NOFRFs can then be estimated efficiently from each time invariant model using a least squares based algorithm to build up a frequency domain description of the time varying phenomena. Application of the TV-NOFRF method to a set of DEAs shows that the time varying behaviour is manifested differently in different actuators, for the first time demonstrating the fact that current fabrication techniques may need improvement if they are to provide consistent dynamic behaviour in this emerging area.

The comparison of the different DEA actuators does not take into account the accuracy of the estimated model. There is therefore a question remaining as to the accurate of the frequency domain descriptions on which the comparisons are based. This motivates an investigation into characterisation of the uncertainty in these descriptions. In order to achieve this, the first step is to characterise the uncertainty in the model on which the frequency domain descriptions are based. There are, however, very few methods available for quantifying uncertainty in NARX model descriptions motivating the development of a new algorithm for the joint structure detection and parameter estimation of NARX models. A Bayesian approach is taken, in which parameter uncertainty is naturally incorporated into the modelling process. The novel algorithm, the SVB-NARX algorithm, uses ARD in order encourage sparsity in the parameter estimates and is used as a basis for driving term selection. The algorithm is benchmarked against other contemporary algorithms and shown to provide comparable performance on a synthetic example.

The algorithm has significant advantages originating from the Bayesian framework. Firstly, model selection is performed automatically such that no subjective choice of threshold for the choice of the final model is necessary, as is commonly the case with contemporary algorithms. Secondly, model uncertainty is incorporated naturally and a full probability model is estimated. The system identification task is repeated for the DEAs using the SVB-NARX algorithm. The identified

models have a prediction accuracy that is comparable to those identified with traditional techniques but now incorporates a posterior distribution over the model parameters that represents the uncertainty in the model.

Initially, the frequency domain uncertainty is estimated by sampling from the posterior distribution of the model parameters and mapping each parameter vector into the frequency domain as NOFRFs. Repeated random sampling builds up a distribution that is observed to be approximately normal for both synthetic examples and for the DEAs such that the variance can be calculated and used to provide confidence bounds on the uncertain NOFRFs. The NOFRF descriptions calculated using both traditional methods and the new SVB-NARX algorithm are compared and shown to be similar. The results add weight to the argument that the frequency domain provides an invariant system description.

The final novel contribution of this thesis is to provide an analytic approach to approximation of the uncertainty bounds in the GFRF and NOFRF by considering the uncertainty in the NARX model parameters, identified with the SVB-NARX algorithm. The sampling approach to estimating the distribution of system NOFRFs is computationally burdensome, particularly at higher non-linear orders, motivating an investigation into analytical methods presented here. Classical methods for estimating the uncertainty of the output of some function by considering the uncertainty in the input is performed by uncertainty propagation. The frequency domain is commonly represented as a complex number in Cartesian space or in gain-phase form in a polar space. Application of the propagation law to cater for a function with a complex output (such as FRFs) requires a multivariate form of the propagation law.

Previously, complex uncertainty propagation has only been considered for FRFs calculated directly from experimental data. These approaches differs from the novel method proposed in this work where uncertainty is propagated into the frequency domain from the uncertainty in the model parameters. General equations for the propagation of parameter uncertainty in linear ARX models are derived. The method is successfully demonstrated by propagating uncertainty into both the complex and gain-phase frequency descriptions. It is also successfully demonstrated that higher order frequency domain descriptions can be accurately estimated using the propagation method. The method is further demonstrated on DEAs using parameter uncertainty estimated with the SVB-NARX algorithm. It is shown that in both the synthetic and the real data case confidence intervals calculated from the estimated distributions are a very close to those found by the sampling method.

8.2 Future work

This section addresses both limitations and potential improvements to the methods introduced in this thesis. A number of different research directions of interest resulting from the work are also discussed.

- *Development of variational Bayesian recursive parameter estimation techniques.* In order to extend the uncertainty analysis of models identified using the SVB-NARX algorithm to the time varying case, methods for recursive parameter estimation within the variational Bayesian framework are required. On-line variational Bayes was proposed in [106] and the approach has been adapted to further applications [79, 115]. The former citation may be appropriate for data sets with rapidly changing parameters. Statistics are built up recursively as more observations are observed, when the system behaviour changes so that these statistics no longer represent the data they are reset and the inference is dominated by the prior. The latter citation proposes a variational Bayesian method for both filtering and smoothing within state-space models with point-process observations. However, for the DEAs analysed in this thesis an approach that is able to estimate slowly varying parameters is necessary.

  The implementation of on-line variational Bayesian inference for the recursive estimation of NARX model parameters requires solving the on-line inference problem for the Bayesian NARX model. Should this be achieved there is potential to extend the uncertain frequency domain analysis method to the time varying case. A potential application of this is in health monitoring for non-linear systems using the n'th order output spectra as signatures. A deviation of these signatures from their original (healthy) state greater than some level of confidence would indicate a system fault.

- *Investigation into the dynamic resolution of the SVB-NARX algorithm.* The resolution parameter used in the SVB-NARX algorithm directly affects how many model terms are removed at each algorithm iteration. Selecting a resolution value that is too small may result in the pruning of terms that may otherwise have been selected by the algorithm. However, a small resolution value also decreases the amount of iterations before termination. In order to achieve maximum algorithm efficiency while maintaining identification accuracy a dynamic resolution is therefore proposed. Two approaches are suggested:

  1. The resolution is set as a function of the variational lower bound such

that the resolution increases with the value of the lower bound. Higher discrimination will hence be achieved between more probably model structures.

2. A two stem approach could be investigated whereby the algorithm is reset with the superset of model terms recorded at the iteration before any maxima in the variational lower bound is observed.

By taking one of the two suggested approaches better discrimination between model terms may be achieved while also remaining efficient to compute.

- *Development of a toolbox for the DEA identification and analysis framework.* The system identification and model based analysis techniques that constitute the developed framework for the modelling and analysis for DEAs are difficult to implement for the non-expert. In order for the framework to be widely applied in the wider soft robotics/EAP communities a user-friendly toolbox would be advantageous. A toolbox that performs each step in the framework, including documentation, is currently under development in the MATLAB environment. The toolbox will perform the following steps;

  1. Joint structure detection and parameter estimation with SEMP.
  2. GFRF and NOFRF analysis.
  3. TV parameter estimation and equilibrium position.
  4. TV-NOFRF analysis

  The toolbox ideally would be extended in future with uncertainty propagation into higher orders. The current model based complex uncertainty propagation software only supports estimation of NOFRF uncertainty at low orders. This is primarily due to the complexity of the high dimensional integration over the space of the input frequencies that is required at high polynomial orders. It would therefore be advantageous to investigate methods for either approximating this integral or implementing parallel processing techniques to ease the computation burden. Such techniques are readily available with the recent upgrades to the MATLAB parallel processing toolbox.

  The toolbox will be available open-source so as to have maximum impact to the scientific community.

- *Development of a novel model-based smart manufacturing optimisation method for soft-smart actuators that will link fabrication parameters to actuator dynamics via*

identified non-linear models and the frequency-domain characteristics. An area of future work that would be interesting to pursue would be to develop a new type of model-based design method for smart manufacturing of soft-smart robot actuators. The key new attribute of this method would be that fabrication design parameters are included within a non-linear dynamic model of the actuator. This would allow a new type of model-based optimisation of the fabrication process, tuned to the specific role or task of the actuator. Due to the application of this technology to autonomous systems/robotics and particularly in the interaction with healthcare in the form of assistive robotic devices, it has the potential to make a useful contribution to this domain.

This approach would make use of outputs from this thesis. The key idea is in two parts: *i*) To include within the dynamic model the fabrication design parameters and then identify the actuator dynamics as a function of these fabrication parameters. *ii*) To perform model-based optimisation of the fabrication parameters that would give rise to desired dynamic performance of the actuator. The choice of dynamic characteristics to optimise against will include both the time and frequency domain, such as, *i*) time domain: rise-time, settling-time, overshoot; *ii*) frequency domain: gain and phase profiles, and features such as open loop bandwidth. The method could be tested initially for proof of concept using a single configuration of DEA.

- *Development of damage detection/fault diagnosis procedures based on frequency domain uncertainty analysis.* It is well known that the dynamic properties of a system can be changed by many forms of damage. The conventional approach to damage detection is to detect changes in the dynamic system behaviour from the behaviour observed from the healthy system. NOFRF based damage detection has been proposed and demonstrated in the literature [14, 94]. The NOFRF provides a convenient approach because it is a function of only one frequency variable and so has low dimensionality. The uncertainty associated with the NOFRF system representation can be found using the frequency domain uncertainty propagation method. It is suggested that the confidence bounds calculated from the NOFRF uncertainty will provide a better indication of whether the differences in the NOFRFs are due to damage or poorly modelled system behaviour.

- *Investigation into robust control design based on complex uncertainty propagation* The final suggestion for work following from this thesis is the investigation into new methods for linear and non-linear robust control. Robust control methods commonly consider the worst case scenarios on the physical mea-

surements (estimated from the known accuracy of the sensors) in order to form uncertainty bounds. Here, the bound on the frequency domain description can be found by the rigorous data-driven characterisation of the uncertainty in the system dynamics. Combining the sparse identification algorithm, SVB-NARX, with the new methods for complex uncertainty propagation provides a starting point for model based robust control design.

Appendix A

Derivations and proofs

A.1 Equivalence for norm regularisation and optimisation with respect to a norm constraint.

Consider the following two problems:

> **Problem one:** Find the minimum of the regularised cost function given by equation (3.29) with a regularisation term given by equation (3.30), such that the optimal value of the parameter vector for a given regularisation constant, λ, is given by
>
> $$\min_{\boldsymbol{\theta}} J_\lambda(\boldsymbol{\theta}) := V(\boldsymbol{\theta}) + \alpha|\boldsymbol{\theta}|^q \tag{A.1}$$
>
> **Problem two:** Find the minimum of the regularised cost function subject to a squared norm constraint such that optimal value of the parameter vector is given by
>
> $$\min_{\boldsymbol{\theta}} \frac{1}{2}|V(\boldsymbol{\theta})| \tag{A.2}$$
>
> $$s.t.\ |\boldsymbol{\theta}|^q - c \leq 0 \tag{A.3}$$

The condition for optimality of problem one is given by

$$\nabla_{\boldsymbol{\theta}} f_\lambda(\boldsymbol{\theta}^*(\lambda)) = 0 \tag{A.4}$$

where the optimal solution, $\boldsymbol{\theta}^*(\lambda)$ is given in terms of λ explicitly to show its dependence.

Problem two can be solved with Lagrangian multipliers, the Lagrangian is given by

$$\mathcal{L}(\boldsymbol{\theta}, \alpha) = V(\boldsymbol{\theta}) + \alpha(|\boldsymbol{\theta}|^q - c). \tag{A.5}$$

The optimal solution is found where the Lagrangian satisfies the Karush-Kuhn-Tucker conditions given by

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}^*, \alpha^*) = \nabla_{\boldsymbol{\theta}} J_\alpha(\boldsymbol{\theta}^*) = 0 \tag{A.6}$$

$$\alpha^*(\|\boldsymbol{\theta}^*\|^2 - c) = 0 \tag{A.7}$$

$$\alpha^* \geq 0. \tag{A.8}$$

Note that in the first condition, equation (A.6), the gradient of the Lagrangian is equivalent to the gradient of the cost function in problem one at $\lambda = \alpha*$ because the constant term, αc, in equation (A.5) differentiates to zero. The second condition, equation (A.7), implies that an optimal solution is found at $c = |\boldsymbol{\theta}^*|^q$. The optimal parameter vector for problem two is therefore found at $\boldsymbol{\theta}^* = \boldsymbol{\theta}(\boldsymbol{\lambda}^*)$, *i.e.* the same as for problem one at the point $c = |\boldsymbol{\theta}^*|^q$. The third condition simply states that α^* must be positive, which is the same condition given for the original regularisation problem. □

A.2 Least squares with ridge regression

The ridge regression cost function is given by

$$J_{RR} = (\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}})^T(\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}}) + \lambda\boldsymbol{\theta}'\boldsymbol{\theta}. \tag{A.9}$$

Differentiating with respect to the parameter vector and setting the result equal to zero

$$\frac{dJ_{LS}}{d\hat{\boldsymbol{\theta}}} = -2\Phi^T(\boldsymbol{y} - \Phi\hat{\boldsymbol{\theta}}_{RR}) + 2\lambda\hat{\boldsymbol{\theta}}_{RR} = 0 \tag{A.10}$$

$$\hat{\boldsymbol{\theta}}_{RR}(\Phi'\Phi + \lambda) = \Phi'\boldsymbol{y} \tag{A.11}$$

$$\hat{\boldsymbol{\theta}}_{RR} = (\Phi'\Phi + \lambda)^{-1}\Phi'\boldsymbol{y} \tag{A.12}$$

□

A.3 The bias variance trade-off decomposition

Consider the general system given by Equation (3.4) where the function $\hat{f}(\mathbf{x})$ is approximating the true function $f(\mathbf{x})$. The expected value of the squared error is

given by

$$\begin{aligned} \boldsymbol{E}[(y-\hat{f}(\boldsymbol{x}))^2] &= \boldsymbol{E}[y^2+\hat{f}^2-2y\hat{f}] && \text{(A.13)}\\ &= \boldsymbol{E}[y^2]+\boldsymbol{E}[\hat{f}^2]-\boldsymbol{E}[2y\hat{f}] && \text{(A.14)}\\ &= Var[y]+\boldsymbol{E}[y]^2+Var[\hat{f}]+\boldsymbol{E}[\hat{f}]^2-2f\boldsymbol{E}[\hat{f}] && \text{(A.15)}\\ &= Var[y]+Var[\hat{f}]+(f-\boldsymbol{E}[\hat{f}])^2 && \text{(A.16)}\\ &= Var[y]+Var[\hat{f}]+\boldsymbol{E}[f-\hat{f}]^2 && \text{(A.17)}\\ &= \sigma_y^2+Var[\hat{f}]+Bias[\hat{f}] && \text{(A.18)} \end{aligned}$$

where $\sigma_y^2 = Var[y]$ is an irreducible error and the following identity has been used.

$$Var[X] = \boldsymbol{E}[X^2]-\boldsymbol{E}[X]^2. \tag{A.19}$$

Equation (A.15) follows by noting that

$$y = f+e \tag{A.20}$$

and therefore

$$\begin{aligned} \boldsymbol{E}[y] &= \boldsymbol{E}[f+e] && \text{(A.21)}\\ &= \boldsymbol{E}[f] && \text{(A.22)}\\ &= f && \text{(A.23)} \end{aligned}$$

Since f is deterministic.

A.4 Bayesian inference: completing the square

The posterior distribution is proportional to the product of the likelihood function multiplied by the prior distribution. Taking the log of the posterior

$$\begin{aligned} \ln p(\boldsymbol{\theta}|\boldsymbol{y}) &= \ln p(\boldsymbol{y}|\boldsymbol{\theta})+\ln p(\boldsymbol{\theta})+C && \text{(A.24)}\\ &= -\frac{1}{2\sigma^2}(\boldsymbol{y}-\Phi\boldsymbol{\theta})'\mathbf{I}_N(\boldsymbol{y}-\Phi\boldsymbol{\theta})-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)+C && \text{(A.25)} \end{aligned}$$

where all other terms have beein incorparated into C. Colleting terms in $\boldsymbol{\theta}$

$$= -\frac{1}{2}\boldsymbol{\theta}'(\frac{1}{\sigma^2}\Phi'\Phi+\boldsymbol{\Sigma}_0^{-1})\boldsymbol{\theta}+\boldsymbol{\theta}'(\frac{1}{\sigma^2}\Phi'\boldsymbol{y}+\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\theta}_0)+C. \tag{A.26}$$

The likelihood and prior are normal and therefore the posterior is also normal by conjugacy. The parameters of the posterior distribution can therefore be read directly from equation (A.26) as

$$\boldsymbol{\theta}_N = \boldsymbol{\Sigma}_N(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\theta_0} + \sigma_e^{-2}\Phi'\boldsymbol{y}) \tag{A.27}$$

$$\boldsymbol{\Sigma}_N^{-1} = \boldsymbol{\Sigma}_0^{-1} + \sigma_e^{-2}\Phi'\Phi \tag{A.28}$$

□

Appendix B

Time invariant modeling of DEAs

B.1 Training and validation data for identification of DEAs

The training and validation data is chosen over an approximately time invariant section of the data record for each DEA. The data sections are given in table B.1 and shown in figure B.1.

DEA	Training data [s]	Validation data [s]
1	6001-7000	7001-8000
2	401-1400	1401-2400
3	101-1100	1101-2100
4	2501-3500	3501-4500
5	201-1200	1201-2200
6	201-1200	1201-2200

Table B.1: **Data sections over which training and validation data is performed.**

B.2 Correlation tests for models of DEAs identified with the SEMP algorithm.

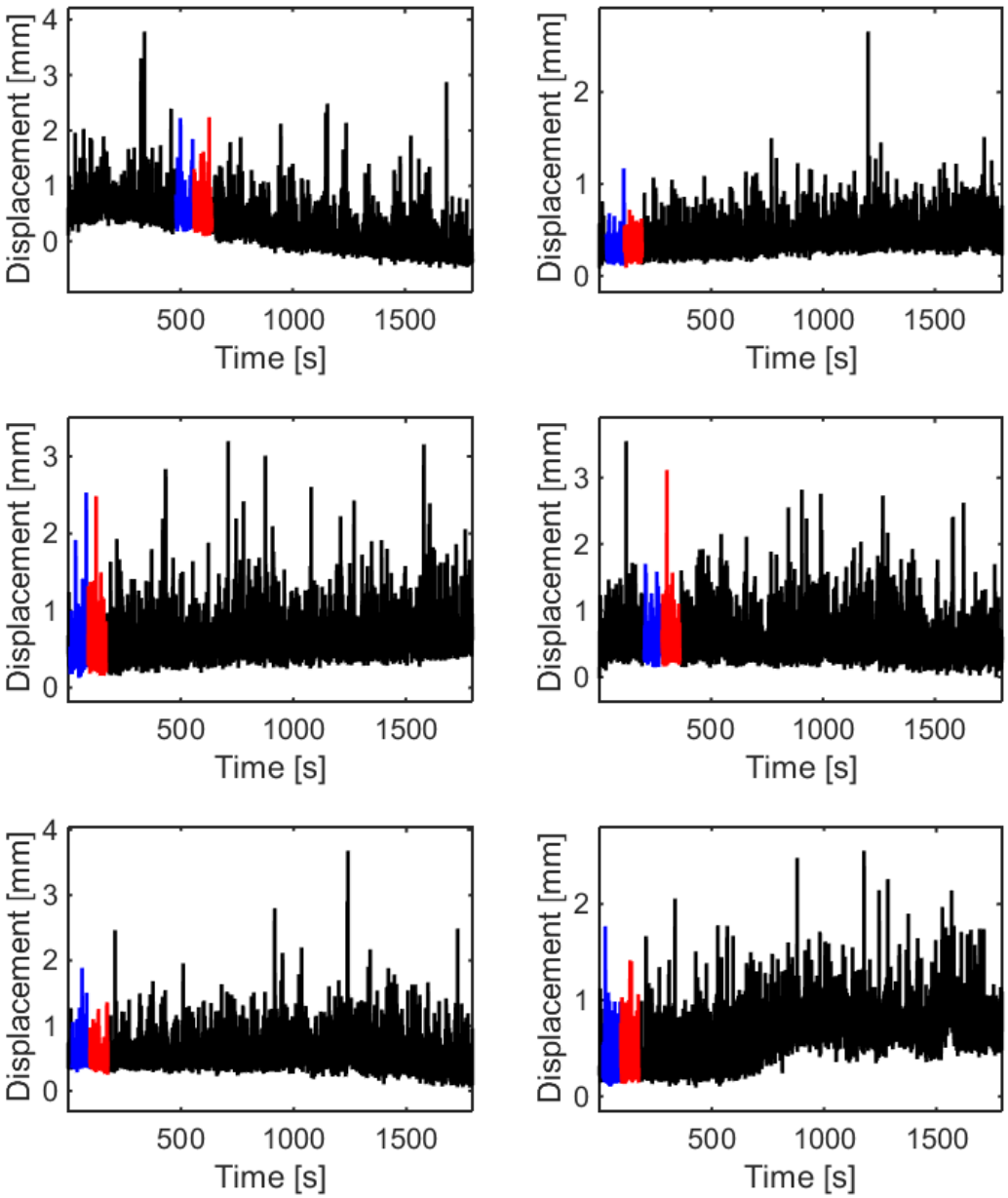

Figure B.1: **Training and validation data used for system identification is chosen over an approximately time invariant section of the data.** The system output for DEAs 1:6 (Black) showing the training data (blue) and the validation data (red).

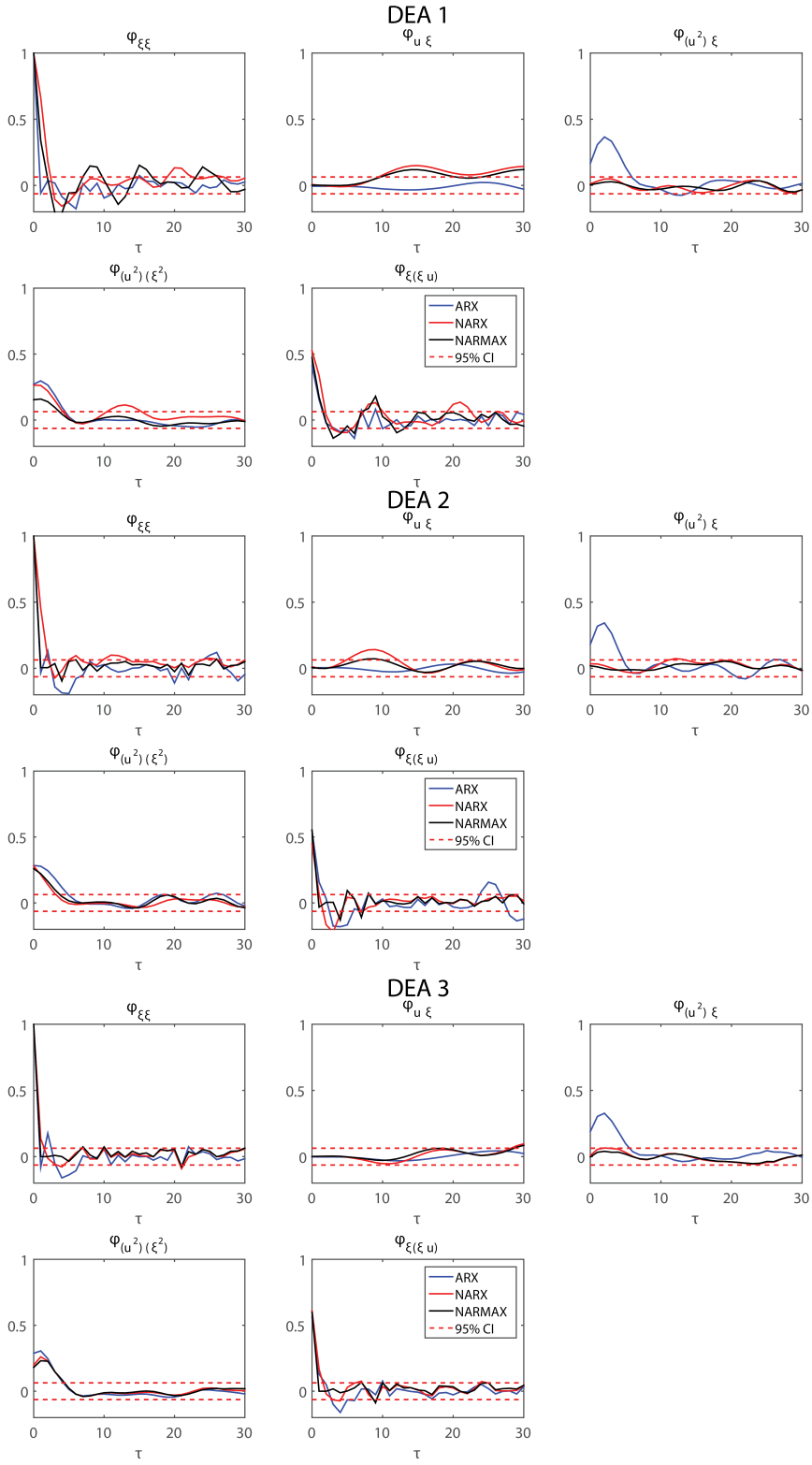

Figure B.2: **Linear and Non-linear correlation tests.** Linear and Non-linear correlation tests for ARX, NARX and NARMAX models of DEAs 1-3 identified with the SEMP algorithm.

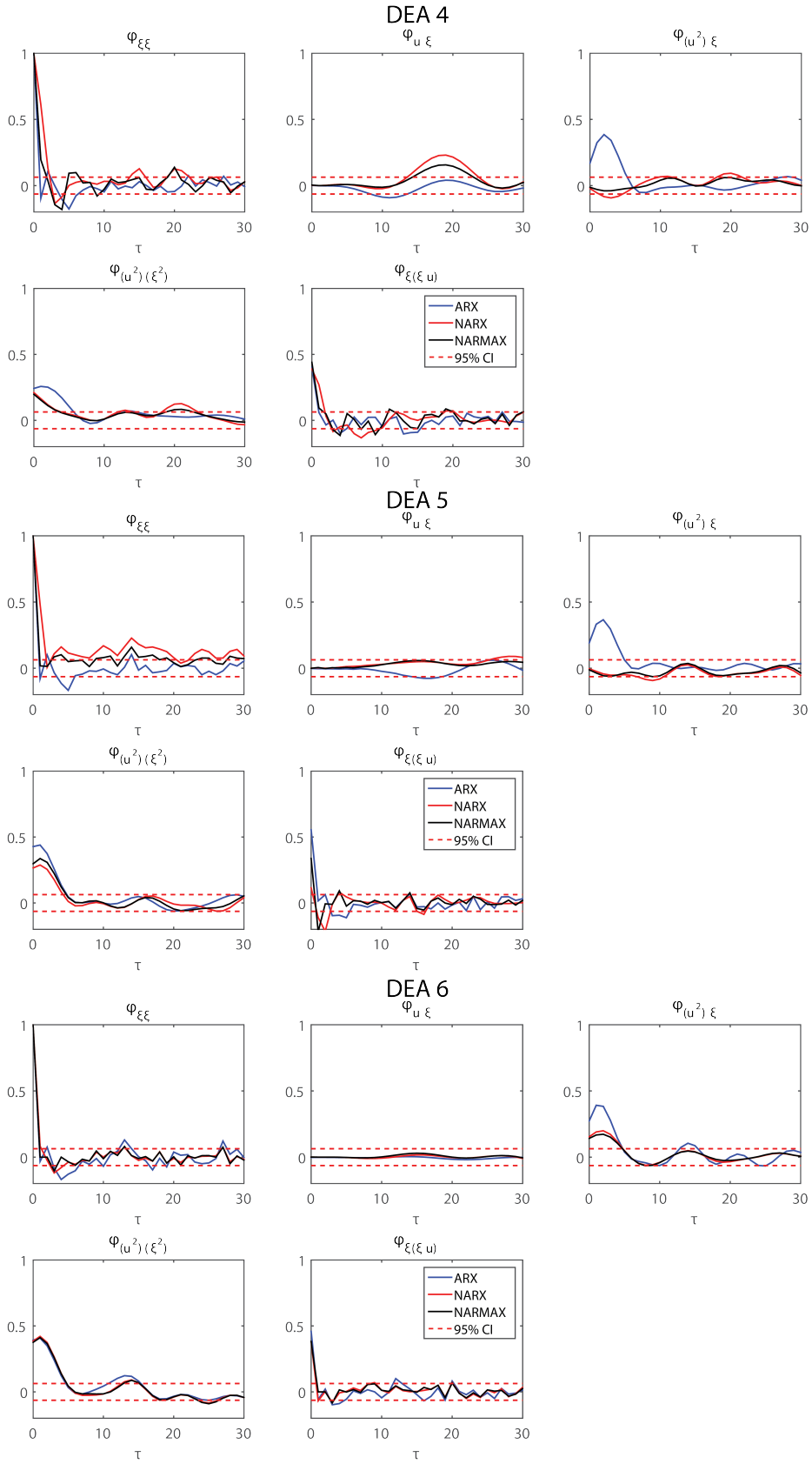

Figure B.3: **Linear and Non-linear correlation tests.** Linear and Non-linear correlation tests for ARX, NARX and NARMAX models of DEAs 4-6 identified with the SEMP algorithm.

Appendix C

Time varying parameters estimated by Kalman filtering with corresponding time varying equilibrium point

C.1 Kalman smoothing equations

Optimum linear smoothing smoothing for the state space model given by equation (3.41) with $\boldsymbol{B}_k = 0 \; \forall \; k$ can be performed by the following sets of equations:

Initialisation

$$\boldsymbol{y}_{bk}^{-} = 0 \tag{C.1}$$

$$\boldsymbol{\Upsilon}_{bk}^{-} = 0 \tag{C.2}$$

$$\tag{C.3}$$

Backwards Update

$$\boldsymbol{y}_{bk}^{+} = \boldsymbol{y}_{bk}^{-} + \boldsymbol{\phi}_k \boldsymbol{R}_k^{-1} z_k \tag{C.4}$$

$$\boldsymbol{\Upsilon}_{bk}^{+} = \boldsymbol{\Upsilon}_{bk}^{-} + \boldsymbol{\phi}_k \boldsymbol{R}_k^{-1} \boldsymbol{\phi}_k^{T} \tag{C.5}$$

Backwards Propagation

$$\boldsymbol{K}_{bk} = \boldsymbol{Y}^{+}_{bk+1}(\boldsymbol{Y}^{+}_{bk+1}\boldsymbol{Q}_k^{-1})^{-1} \tag{C.6}$$
$$\boldsymbol{y}^{-}_{bk} = \boldsymbol{A}_k^T(\boldsymbol{I} - \boldsymbol{K}_{bk})\boldsymbol{y}^{+}_{bk} \tag{C.7}$$
$$\boldsymbol{Y}^{-}_{bk} = \boldsymbol{A}_k^T(\boldsymbol{I} - \boldsymbol{K}_{bk})\boldsymbol{Y}^{+}_{bk+1}\boldsymbol{A}_k \tag{C.8}$$
$$\tag{C.9}$$

$$\boldsymbol{\theta}^{s}_k = (\boldsymbol{I} - \boldsymbol{K}^{s}_k)\boldsymbol{\theta}^{+}_k + \boldsymbol{P}^{s}_k\boldsymbol{y}^{-}_{bk} \tag{C.10}$$
$$\boldsymbol{K}^{s}_k = \boldsymbol{P}^{+}_k\boldsymbol{Y}^{-}_{bk}(\boldsymbol{I} + \boldsymbol{P}^{+}_k\boldsymbol{Y}^{-}_k)^{-1} \tag{C.11}$$
$$P^{s}_k = (\boldsymbol{I} - \boldsymbol{K}^{s}_k)P^{+}_k \tag{C.12}$$

C.2 Time varying parameter estimates and equilibrium position

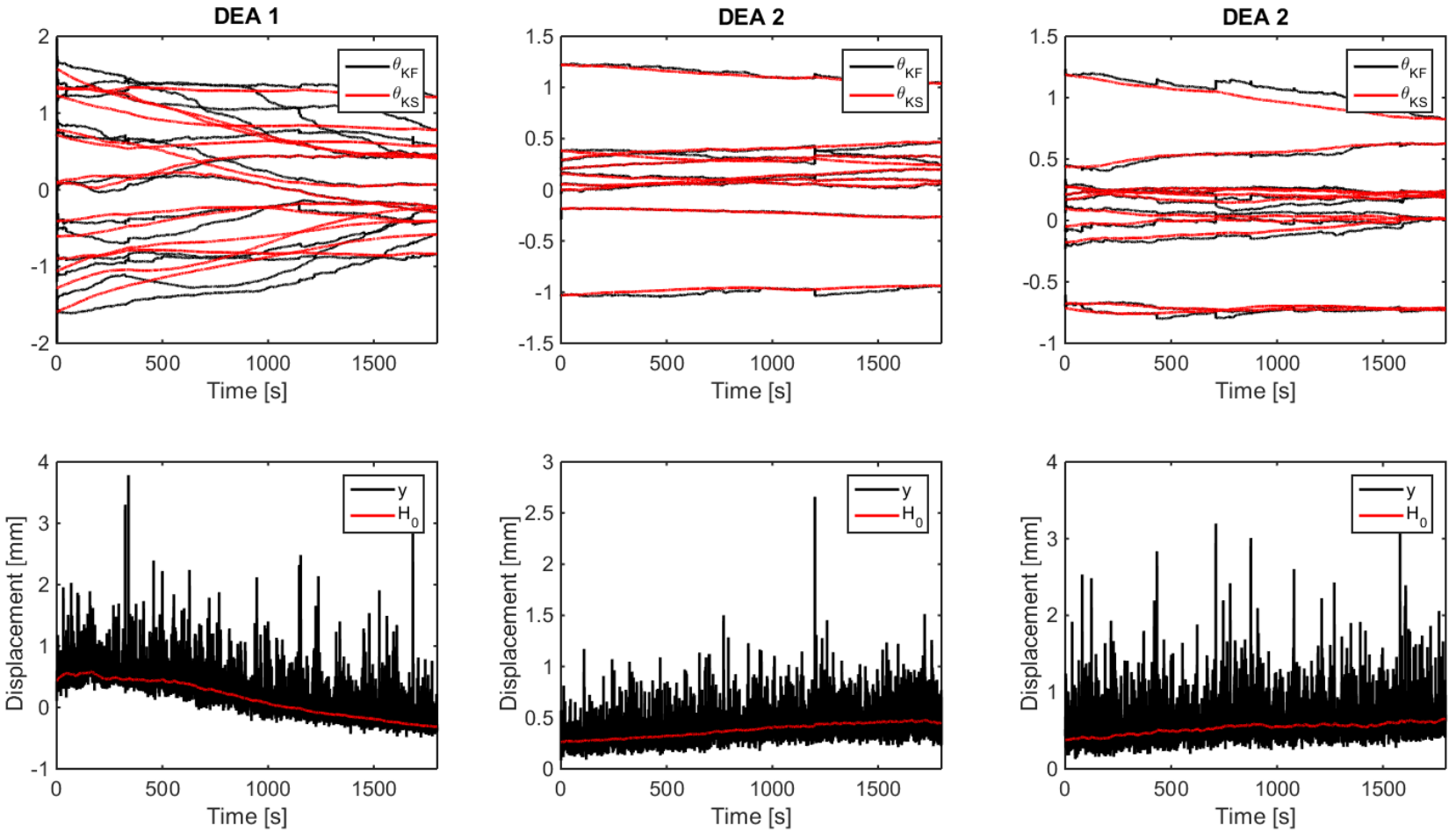

Figure C.1: **Time varying parameters are estimated on-line using a Kalman filter and updated off-line using a Kalman smoother, the DC component is then removed from the model.** Top: Time varying parameters for DEA 1-3 calculated using a Kalman filter (Black) and updated using a Kalman smoother (Red). Bottom: The output $_k$ for DEA 5 (Black) with the TV equilibrium position, $y_0(k)$.

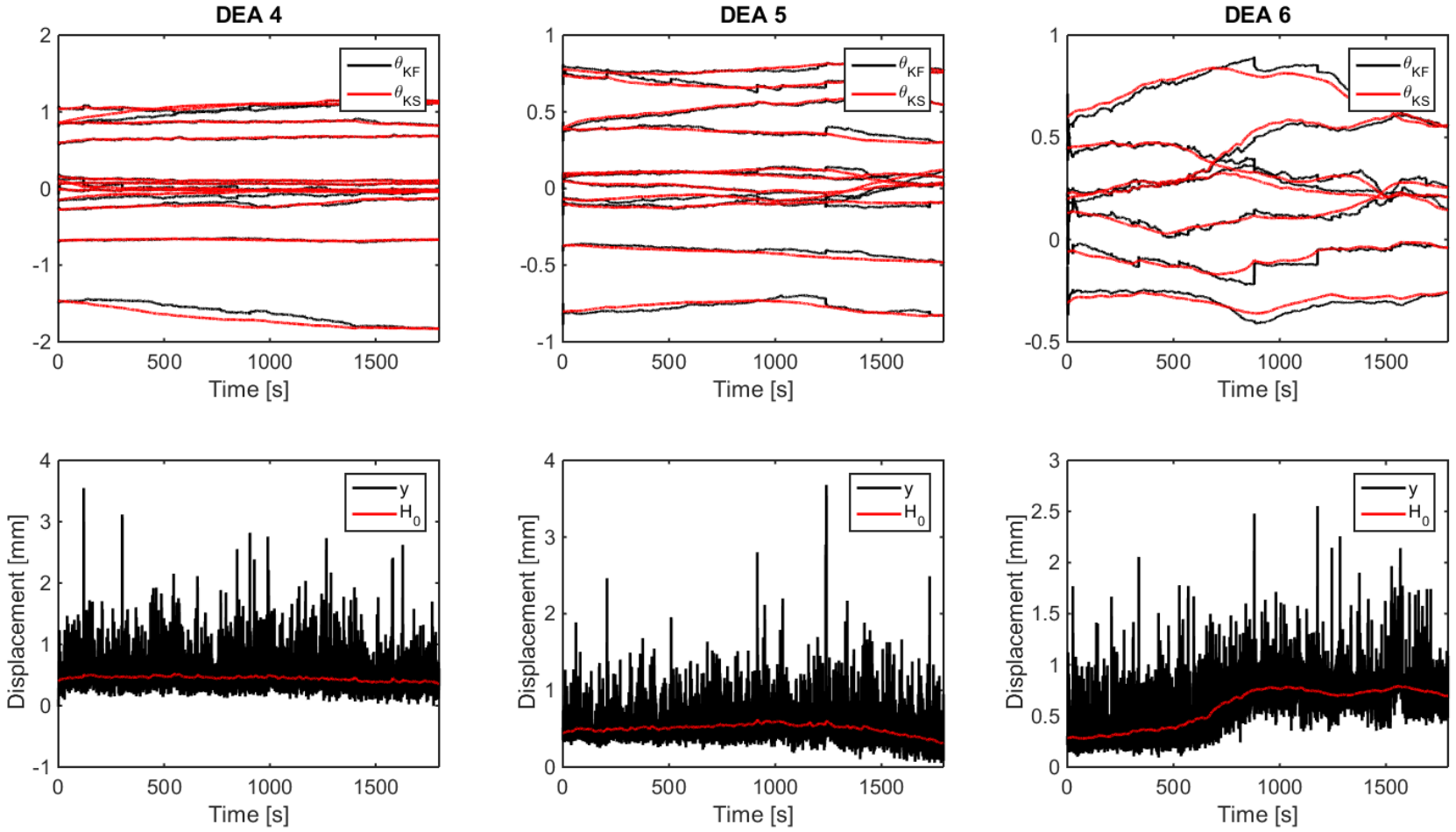

Figure C.2: **Time varying parameters are estimated on-line using a Kalman filter and updated off-line using a Kalman smoother, the DC component is then removed from the model.** Top: Time varying parameters for DEA 4-6 calculated using a Kalman filter (Black) and updated using a Kalman smoother (Red). Bottom: The output $_k$ for DEA 5 (Black) with the TV equilibrium position, $y_0(k)$.

Appendix D

Modeling of DEAs with the SVB-NARX algorithm

D.1 NARX models of DEAs identified with the SVB-NARX algorithm.

D.2 NARX models of DEAs identified with the SVB-NARX algorithm with DC component removed.

D.3 Correlation tests for NARX models of DEAs identified with the SVB-NARX algorithm

Terms	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6
DC term	-	0.110909	0.100324	-	-	0.106316
y_{k-1}	1.126108	0.635597	0.889559	0.595594	0.905585	0.730234
y_{k-2}	-0.173834	-	-0.223545	0.724595	-	-0.205051
y_{k-3}	0.128883	-	-	-0.180362	0.197705	0.101358
u_{k-1}	-	-	-	0.291770	-	0.531001
u_{k-2}	-	-	-	-0.612404	-	-0.448562
u_{k-3}	-	0.182089	0.075806	0.234188	-	0.153306
$y_{k-1}y_{k-3}$	-	-0.139791	-	0.422403	-	-
$y_{k-1}u_{k-1}$	1.590593	1.887260	1.063856	0.885335	-	-
$y_{k-1}u_{k-2}$	-1.391008	-1.555698	-1.015277	-	-0.550350	-0.214585
$y_{k-1}u_{k-3}$	-	-	-	-0.676577	-	-
y_{k-2}^2	-0.192015	-	-	-0.720368	-0.214408	-
$y_{k-2}y_{k-3}$	-	-	0.205926	-	-	-
$y_{k-2}u_{k-1}$	-0.433729	-0.357084	-	-	0.572534	-
$y_{k-2}u_{k-3}$	-	-	-	0.346834	-	-
$y_{k-3}u_{k-2}$	0.241243	-	-	-	-	-
$y_{k-3}u_{k-3}$	-	-0.176562	-	-0.186588	-	-
u_{k-1}^2	0.158786	-	-	0.261109	0.896250	0.782601
$u_{k-1}u_{k-2}$	-	-	-	-0.287463	-1.359614	-0.791393
$u_{k-1}u_{k-3}$	-	-	-	-	-	0.350320
u_{k-2}^2	-	-	-	-0.280386	-	-
$u_{k-2}u_{k-3}$	-	-	-	0.132932	1.163239	-
u_{k-3}^2	-	-	-	-	-0.666935	-
y_{k-1}^3	-	-	0.156436	-	-	-
y^2k-1y_{k-2}	-	-0.201361	-0.217866	-	-	-
y^2k-1u_{k-1}	-	-	-	-0.324119	0.308154	0.073970
y^2k-1u_{k-3}	-	-	-0.151059	-	-	-
$y_{k-1}y_{k-2}u_{k-3}$	0.221407	-	-	-	-0.561115	-
$y_{k-1}u_{k-1}^2$	-	-	0.184908	0.717225	-	-
$y_{k-1}u_{k-1}u_{k-2}$	-0.256614	-	-	-	-	-
$y_{k-1}u_{k-3}^2$	-	0.528004	-	-	-	-
y_{k-2}^3	-	-	0.073089	-	-	-
$y_{k-2}^2u_{k-3}$	-	-	-	0.325876	0.390603	-
$y_{k-2}y_{k-3}^2$	-	0.162389	-	-	-	-
$y_{k-2}u_{k-1}^2$	-	-	-	-0.439606	-	-
y_{k-3}^3	-	-	-0.029821	-	-	-
$y_{k-3}u_{k-1}^2$	-	-	-	-	-	-0.104582
u_{k-1}^3	0.154324	0.464028	-	-	0.201649	0.116520
$u_{k-1}^2u_{k-2}$	-	-0.752792	-	-	-	-
$u_{k-1}u_{k-2}^2$	-	0.259679	-	-	-	-
$u_{k-1}u_{k-2}u_{k-3}$	-	-	-	-	-0.367595	-
u_{k-2}^3	-	-	-	-0.093188	-	-
$u_{k-2}^2u_{k-3}$	-	-	-	-	0.175457	-
MSSE ($\times 10^{-3}$)	3.21	0.40	2.28	0.82	6.13	1.39
Term count	12	14	13	21	15	14

Table D.1: NARX models for DEAs 1-6 and their corresponding parameter values identified using the SVB-NARX algorithm.

Terms	DEA1	DEA2	DEA3	DEA4	DEA5	DEA6
y_{k-1}	1.071010	0.605321	0.966907	0.799192	0.906256	0.744502
y_{k-2}	1.144899	0.576953	0.877699	0.743252	0.925898	0.776536
y_{k-3}	1.143713	0.569749	0.947516	0.773512	0.903929	0.761207
u_{k-1}	1.078508	0.563436	0.963874	0.780212	0.900850	0.730774
u_{k-2}	1.095771	0.543638	0.904805	0.779177	0.905919	0.722372
u_{k-3}	1.168279	0.588031	0.906421	0.826058	0.901240	0.716879
y_{k-1}^2	-	0.598615	0.926397	-	-	-
$y_{k-1}y_{k-2}$	-	0.574787	0.893914	-	-	-
$y_{k-1}y_{k-3}$	-	0.523254	-	0.804111	-	-
$y_{k-1}u_{k-1}$	1.118131	0.582871	0.920236	0.769886	0.907134	0.697342
$y_{k-1}u_{k-2}$	1.078778	0.591956	0.895429	-	0.919934	0.745773
$y_{k-1}u_{k-3}$	1.147176	-	0.876061	0.809024	0.912086	-
y_{k-2}^2	1.184365	0.569525	0.908888	0.827937	0.889808	-
$y_{k-2}y_{k-3}$	-	0.597191	0.916219	-	-	-
$y_{k-2}u_{k-1}$	1.122950	0.562417	-	-	0.899310	-
$y_{k-2}u_{k-3}$	1.150481	-	-	0.863697	0.896809	-
y_{k-3}^2	-	0.530313	0.845378	-	-	-
$y_{k-3}u_{k-2}$	1.125967	-	-	-	-	-
$y_{k-3}u_{k-3}$	-	0.564670	-	0.862327	-	-
u_{k-1}^2	1.164117	-	0.856653	0.780608	0.913121	0.740314
$u_{k-1}u_{k-2}$	1.162109	-	-	0.810132	0.910705	0.726304
$u_{k-1}u_{k-3}$	-	-	-	-	-	0.682808
u_{k-2}^2	-	-	-	0.798357	-	-
$u_{k-2}u_{k-3}$	-	-	-	0.725628	0.903071	-
u_{k-3}^2	-	0.576991	-	-	0.922408	-
y_{k-1}^3	-	-	0.873323	-	-	-
$y_{k-1}^2y_{k-2}$	-	0.590116	0.910684	-	-	-
$y_{k-1}^2u_{k-1}$	-	-	-	0.831102	0.915645	0.744768
$y_{k-1}^2u_{k-3}$	-	-	0.905762	-	-	-
$y_{k-1}y_{k-2}u_{k-3}$	1.136026	-	-	-	0.916711	-
$y_{k-1}u_{k-1}^2$	-	-	0.828034	0.779376	-	-
$y_{k-1}u_{k-1}u_{k-2}$	1.144712	-	-	-	-	-
$y_{k-1}u_{k-3}^2$	-	0.549178	-	-	-	-
y_{k-2}^3	-	-	0.897069	-	-	-
$y_{k-2}^2u_{k-3}$	-	-	-	0.839682	0.918523	-
$y_{k-2}y_{k-3}^2$	-	0.565586	-	-	-	-
$y_{k-2}u_{k-1}^2$	-	-	-	0.822640	-	-
y_{k-3}^3	-	-	0.949978	-	-	-
$y_{k-3}u_{k-1}^2$	-	-	-	-	-	0.743875
u_{k-1}^3	1.127164	0.571477	-	-	0.900561	0.739936
$u_{k-1}^2u_{k-2}$	-	0.599565	-	-	-	-
$u_{k-1}u_{k-2}^2$	-	0.585009	-	-	-	-
$u_{k-1}u_{k-2}u_{k-3}$	-	-	-	-	0.901037	-
u_{k-2}^3	-	-	-	0.819124	-	-
$u_{k-2}^2u_{k-3}$	-	-	-	-	0.915745	-
Term count	18	23	21	21	22	14

Table D.2: DC removed NARX models for DEAs 1-6 and their corresponding parameter values identified using the SVB-NARX algorithm.

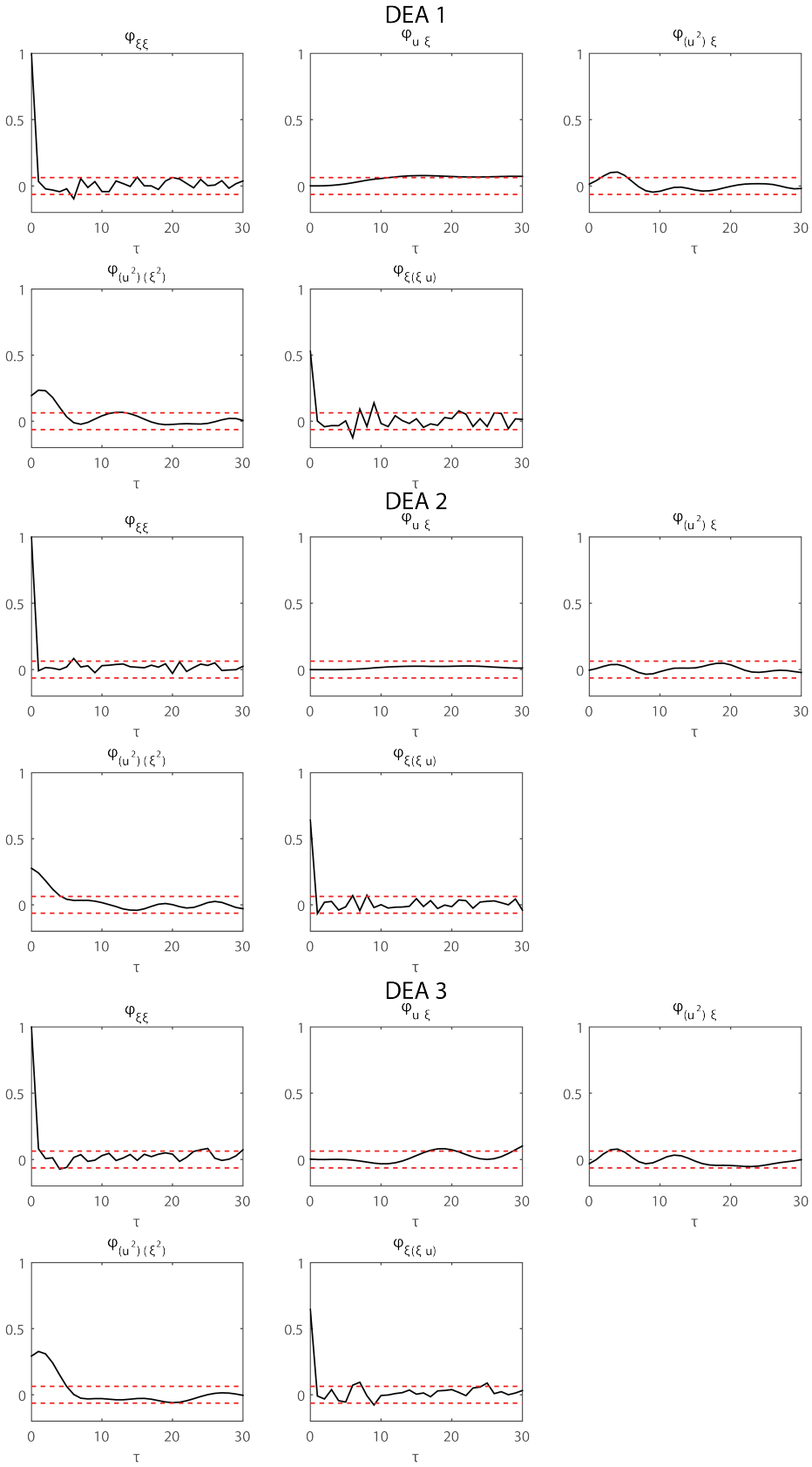

Figure D.1: **Linear and Non-linear correlation tests.** Linear and Non-linear correlation tests for NARX models of DEAs 1-3 identified with the SVB-NARX algorithm.

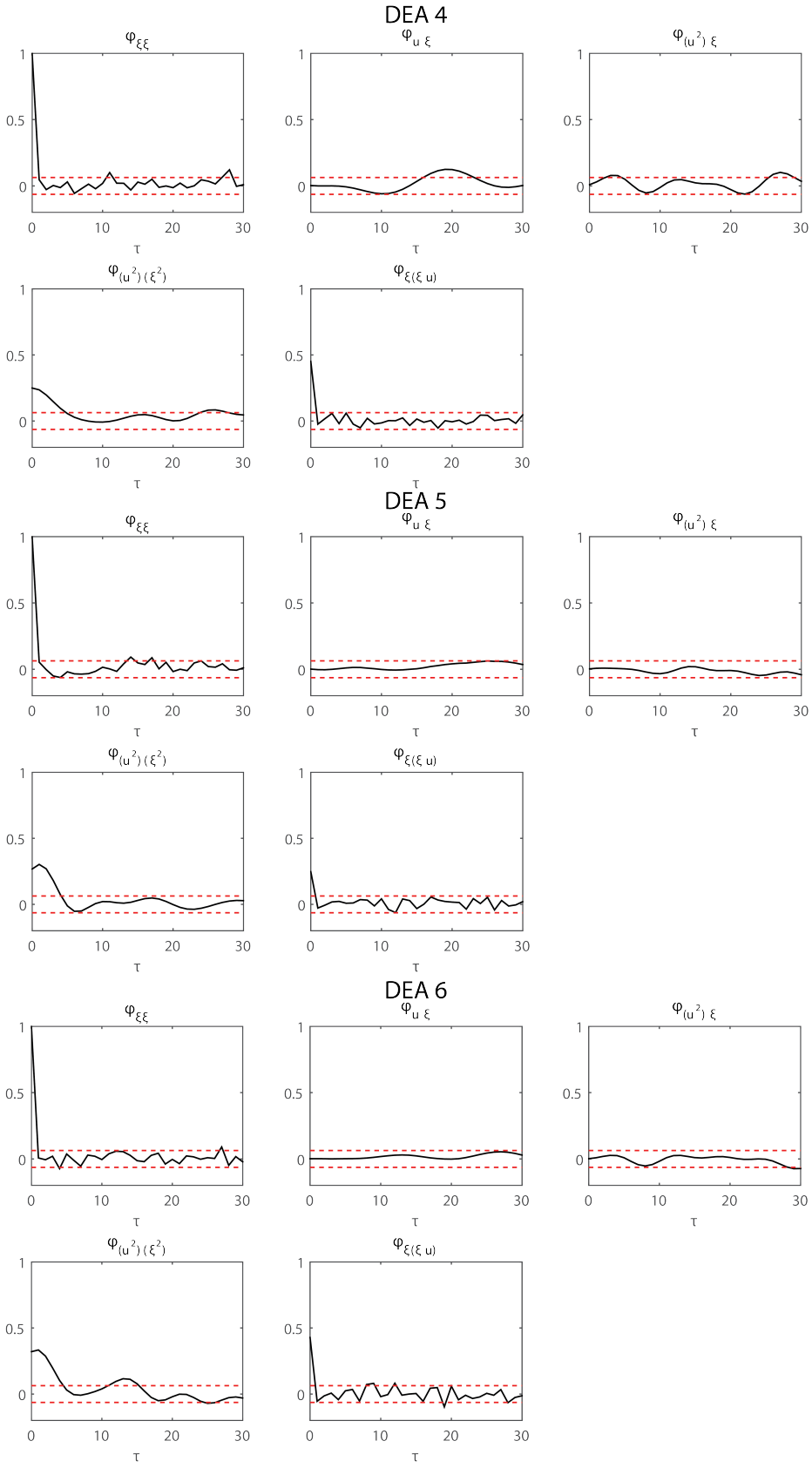

Figure D.2: Linear and Non-linear correlation tests. Linear and Non-linear correlation tests for NARX models of DEAs 4-6 identified with the SVB-NARX algorithm.

Bibliography

[1] Karl J. Åström. *Feedback systems: an introduction for scientists and engineers*. Princeton University Press, Princeton, 2008.

[2] Karl-Johan Åström and Torsten Bohlin. Numerical Identification of Linear Dynamic Systems from Normal Operating Records. In P. H. Hammond, editor, *Theory of Self-Adaptive Control Systems*, pages 96–111. Springer US, Boston, MA, 1966.

[3] Henry D. I Abarbanel, Reggie Brown, and James B. Kadtke. Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra. *Physical Review. A*, 41(4):1782–1807, February 1990.

[4] Hirotugu Akaike. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21(1):243–247, December 1969.

[5] Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

[6] S.R. Anderson and V. Kadirkamanathan. Modelling and identification of non-linear deterministic systems in the delta-domain. *Automatica*, 43(11):1859–1868, November 2007.

[7] Yousef Bahramzadeh and Mohsen Shahinpoor. A Review of Ionic Polymeric Soft Actuators and Sensors. *Soft Robotics*, 1(P):38–52, July 2013.

[8] Tara Baldacchino, Sean R. Anderson, and Visakan Kadirkamanathan. Computational system identification for Bayesian NARMAX modelling. *Automatica*, 49(9):2641–2651, September 2013.

[9] Yoseph Bar-Cohen. Transition of EAP material from novelty to practical applications: are we there yet? In *SPIE's 8th Annual International Symposium on Smart Structures and Materials*, pages 1–6, 2001.

[10] Yoseph Bar-Cohen. Artificial muscles using electroactive polymers (EAP): capabilities, challenges and potential. 2005.

[11] Matthew James Beal. *Variational algorithms for approximate Bayesian inference*. PhD thesis, University of London, 2003.

[12] Edward Bedrosian and Stephen O. Rice. The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs. *Proceedings of the IEEE*, 59(12):1688–1707, 1971.

[13] S. A. Billings. Structure detection and model validity tests in the identification of nonlinear systems. *Control Theory and Applications, IEE Proceedings D*, 130(4):193 – 199, 1983.

[14] S. A. Billings. *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. Wiley, 2013.

[15] S. A. Billings and Luis Antonio Aguirre. Effects of the sampling time on the dynamics and identification of nonlinear models. *International Journal of Bifurcation and Chaos*, 05(06):1541–1556, December 1995.

[16] S. A. Billings, S. Chen, and M. J. Korenberg. Identification of MIMO non-linear systems using a forward-regression orthogonal estimator. *International Journal of Control*, 49(6):2157–2189, June 1989.

[17] S. A. Billings, M. J. Korenberg, and S. Chen. Identification of non-linear output-affine systems using an orthogonal least-squares algorithm. *International Journal of Systems Science*, 19(8):1559–1568, January 1988.

[18] S. A. Billings and K.Z. Mao. Model identification and assessment based on model predicted output - Google Scholar. Technical Report 714, ACSE, University of Sheffield, May 1998.

[19] S. A. Billings and Q. H. Tao. Model validity tests for non-linear signal processing applications. *International Journal of Control*, 54(1):157–194, July 1991.

[20] S. A. Billings and K. M. Tsang. Spectral analysis for non-linear systems, Part I: Parametric non-linear spectral analysis. *Mechanical Systems and Signal Processing*, 3(4):319–339, 1989.

[21] S. A. Billings and W. S. F. Voon. Correlation based model validity tests for non-linear models. *International Journal of Control*, 44(1):235–244, July 1986.

[22] S. A. Billings and W. S. F. Voon. A prediction-error and stepwise-regression estimation algorithm for non-linear systems. *International Journal of Control*, 44(3):803–822, September 1986.

[23] S. A. Billings and H. Zhang. Computation of non-linear transfer functions when constant terms are present. *Mechanical Systems and Signal Processing*, 9(5):537–553, 1995.

[24] Christopher M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, New York, 2006.

[25] Paul Brochu and Qibing Pei. Advances in Dielectric Elastomers for Actuators and Artificial Muscles. *Macromolecular Rapid Communications*, 31(1):10–36, January 2010.

[26] Darwin G. Caldwell, N. G. Tsagarakis, Sophia Kousidou, Nelson Costa, and Ioannis Sarakoglou. "Soft" exoskeletons for upper and lower body rehabilitation: design, control and testing. *International Journal of Humanoid Robotics*, 4(03):549–573, 2007.

[27] Luigi Carassale and Ahsan Kareem. Modeling nonlinear systems by Volterra series. *Journal of Engineering Mechanics*, 136(6):801–818, 2009.

[28] S. Chen and S. A. Billings. Representations of non-linear systems: the NARMAX model. *International Journal of Control*, 49(3):1013–1032, March 1989.

[29] S. Chen, S. A. Billings, and P. M. Grant. Non-linear system identification using neural networks. *International Journal of Control*, 51(6):1191–1214, January 1990.

[30] S. Chen, S. A. Billings, and W. Luo. Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, 50(5):1873–1896, November 1989.

[31] Hyouk Ryeol Choi, Kwangmok Jung, Nguyen Huu Chuc, Minyoung Jung, Igmo Koo, Jachoon Koo, Joonho Lee, Jonghoon Lee, Jaedo Nam, Misuk Cho, and Youngkwan Lee. Effects of prestrain on behavior of dielectric elastomer actuator. In Yoseph Bar-Cohen, editor, *Proc. SPIE 5759, Smart Structures and Materials 2005: Electroactive Polymer Actuators and Devices*, pages 283–291, May 2005.

[32] Torbjørn Dahl and Maged Boulos. Robots in Health and Social Care: A Complementary Technology to Home Care and Telehealthcare? *Robotics*, 3(1):1–21, December 2013.

[33] João Ferdinando Gomes de Freitas. *Bayesian methods for neural networks*. PhD thesis, Citeseer, 2000.

[34] Jan Drugowitsch. Variational Bayesian inference for linear and logistic regression. *arXiv preprint*, arXiv:1310.5438, 2013.

[35] Ken Dutton, S. Thompson, and Bill Barraclough. *The art of control engineering*. Addison Wesley, Harlow ; Reading, Mass, 1997.

[36] Joseph Eckerle, Scott Stanford, John Marlow, Roger Schmidt, Seajin Oh, Thomas Low, and Subramanian V. Shastri. Biologically inspired hexapedal robot using field-effect electroactive elastomer artificial muscles. pages 269–280, June 2001.

[37] William J. Fitzgerald, editor. *Nonlinear and nonstationary signal processing*. Cambridge University Press, Cambridge ; New York, 2000.

[38] D. Fraser and J. Potter. The optimum linear smoother as a combination of two optimum linear filters. *IEEE Transactions on Automatic Control*, 14(4):387–390, August 1969.

[39] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.

[40] Carl Friedrich Gauss. *Theory of the motion of the heavenly bodies moving about the sun in conic sections*. Dover, Mineola, N.Y., 2004.

[41] I. M. Gelfand, S. V. Fomin, and Richard A. Silverman. *Calculus of variations*. Dover Publications, Mineola, N.Y, 2000.

[42] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov chain Monte Carlo in practice*. Chapman & Hall, Boca Raton, Fla, 1998.

[43] Leo A. Goodman. On the Exact Variance of Products. *Journal of the American Statistical Association*, 55(292):708, December 1960.

[44] Peter J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[45] Jim E. Griffin and Philip J. Brown. Inference with Normal-Gamma prior distributions in regression problems. *Bayesian Analysis*, 5(1):171–188, March 2010.

[46] B. D. Hall. Calculating measurement uncertainty for complex-valued quantities. *Measurement Science and Technology*, 14(3):368, 2003.

[47] B D Hall. On the propagation of uncertainty in complex-valued quantities. *Metrologia*, 41(3):173–177, June 2004.

[48] Wolfgang Karl Härdle and Léopold Simar. *Applied Multivariate Statistical Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[49] Fei He, Hua-Liang Wei, and Stephen A. Billings. Identification and frequency domain analysis of non-stationary and nonlinear systems using time-varying NARMAX models. *International Journal of Systems Science*, pages 1–14, November 2013.

[50] Hugh M. Herr and Roy D. Kornbluh. New horizons for orthotic and prosthetic technology: artificial muscle for ambulation. In Yoseph Bar-Cohen, editor, *SPIE Proceedings Vol. 5385: Smart Structures and Materials 2004: Electroactive Polymer Actuators and Devices (EAPAD)*, pages 1–9, July 2004.

[51] Richard Heydt, Roy Kornbluh, Joseph Eckerle, and Ron Pelrine. Sound radiation properties of dielectric elastomer electroactive polymer loudspeakers. In Yoseph Bar-Cohen, editor, *Proc. SPIE 6168*, pages 61681–61688, March 2006.

[52] Fumiya Iida and Cecilia Laschi. Soft Robotics: Challenges and Perspectives. *Procedia Computer Science*, 7:99–102, January 2011.

[53] Bureau International des Poids et Mesures internationale de normalisation , Commission electrotechnique internationale, and Organisation. *Guide to the Expression of Uncertainty in Measurement*. International Organization for Standardization, 1995.

[54] William R Jacobs, Emma D Wilson, Tareq Assaf, Jonathan Rossiter, Tony J Dodd, John Porrill, and Sean R Anderson. Control-focused, nonlinear and time-varying modelling of dielectric elastomer actuators with frequency response analysis. *Smart Materials and Structures*, 24(5):055002, May 2015.

[55] J. C. Peyton Jones and S. A. Billings. Recursive algorithm for computing the frequency response of a class of non-linear difference equation models. *International Journal of Control*, 50(5):1925–1940, November 1989.

[56] J.C. Peyton Jones and K. Choudhary. Efficient computation of higher order frequency response functions for nonlinear systems with, and without, a constant term. *International Journal of Control*, 85(5):578–593, May 2012.

[57] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, 1960.

[58] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, June 1995.

[59] Kazuo Kiguchi, Mohammad Habibur Rahman, Makoto Sasaki, and Kenbu Teramoto. Development of a 3DOF mobile exoskeleton robot for human upper-limb motion assist. *Robotics and Autonomous Systems*, 56(8):678–691, August 2008.

[60] Kwang Jin Kim and Satoshi Tadokoro. *Electroactive polymers for robotic applications artificial muscles and sensors*. Springer, London, 2007.

[61] Sangbae Kim, Matthew Spenko, Salomon Trujillo, Barrett Heyneman, Virgilio Mattoli, and Mark R. Cutkosky. Whole body adhesion: hierarchical, directional and distributed control of adhesive forces for a climbing robot. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 1268–1273. IEEE, 2007.

[62] H. Kobayashi, H. Suzuki, and M. Iba. Development of a Muscle Suit for the Upper Limb Motion Support A New Shoulder Mechanism and Posture Measurement. In *Automation Congress, 2006. WAC'06. World*, pages 1–6, 2006.

[63] Guggi Kofod, Peter Sommer-Larsen, Roy Kornbluh, and Ron Pelrine. Actuation Response of Polyacrylate Dielectric Elastomers. *Journal of Intelligent Materials Systems and Structures*, 14(12):787–793, December 2003.

[64] M. Korenberg, S. A. Billings, Y. P. Liu, and P. J. McIlroy. Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, 48(1):193–210, July 1988.

[65] Sunil L. Kukreja, Johan Lofberg, and Martin J. Brenner. A least absolute shrinkage and selection operator (LASSO) for nonlinear system identification. 2006.

[66] MickaeÏ Lallart, Pierre-Jean Cottinet, Laurent Lebrun, Benoit Guiffard, and Daniel Guyomar. Evaluation of energy harvesting performance of electrostrictive polymer and carbon-filled terpolymer composites. *Journal of Applied Physics*, 108(3):034901, 2010.

[67] Z. Q. Lang and S. A. Billings. Energy transfer properties of non-linear systems in the frequency domain. *International Journal of Control*, 78(5):345–362, March 2005.

[68] Zi-Qiang Lang and S. A. Billings. Output frequency characteristics of non-linear systems. *International Journal of Control*, 64(6):1049–1067, August 1996.

[69] Zi-Qiang Lang and S. A. Billings. Output frequencies of nonlinear systems. *International Journal of Control*, 67(5):713–730, January 1997.

[70] Cecilia Laschi, Matteo Cianchetti, Barbara Mazzolai, Laura Margheri, Maurizio Follador, and Paolo Dario. Soft Robot Arm Inspired by the Octopus. *Advanced Robotics*, 26(7):709–727, January 2012.

[71] C.C. Lee. Fuzzy logic in control systems: fuzzy logic controller. I. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2):404–418, March-April/1990.

[72] I. J. Leontaritis and S. A. Billings. Input-output parametric models for non-linear systems Part I: deterministic non-linear systems. *International Journal of Control*, 41(2):303–328, February 1985.

[73] Hod Lipson. Challenges and Opportunities for Design, Simulation, and Fabrication of Soft Robots. *Soft Robotics*, 1(P):21–27, July 2013.

[74] Liwu Liu, Yanju Liu, and Jinsong Leng. Theory progress and applications of dielectric elastomers. *International Journal of Smart and Nano Materials*, pages 1–11, October 2013.

[75] Lennart Ljung. *System identification: theory for the user*. Prentice Hall information and system sciences series. Prentice Hall PTR, Upper Saddle River, NJ, 2nd ed edition, 1999.

[76] Lennart Ljung and Svante Gunnarsson. Adaptation and tracking in system identification—A survey. *Automatica*, 26(1):7–21, January 1990.

[77] Aidan Lyon. Why are normal distributions normal? *The British Journal for the Philosophy of Science*, 65(3):621–649, 2014.

[78] David Mackay. Probable networks and plausible predictions — a review of practical Bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, 6(3):469–505, August 1995.

[79] Andrew Zammit Mangion, Ke Yuan, Visakan Kadirkamanathan, Mahesan Niranjan, and Guido Sanguinetti. Online variational inference for state-space models with point-process observations. *Neural computation*, 23(8):1967–1999, 2011.

[80] K. Z. Mao and S. A. Billings. Algorithms for minimal model structure detection in nonlinear dynamic system identification. *International Journal of Control*, 68(2):311–330, January 1997.

[81] Thomas G. McKay, Emilio Calius, and Iain A. Anderson. The dielectric constant of 3M VHB: a parameter in dispute. In Yoseph Bar-Cohen and Thomas Wallmersperger, editors, *Proc. SPIE 7287, Electroactive Polymer Actuators and Devices (EAPAD) 2009*, volume 72870P, March 2009.

[82] Arianna Menciassi, Dino Accoto, Samuele Gorini, and Paolo Dario. Development of a biomimetic miniature robotic crawler. *Autonomous Robots*, 21(2):155–163, August 2006.

[83] Thomas P. Minka. Expectation propagation for approximate Bayesian inference. In *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*, pages 362–369. Morgan Kaufmann Publishers Inc., 2001.

[84] M. Mooney. A Theory of Large Elastic Deformation. *Journal of Applied Physics*, 11(9):582, 1940.

[85] Radford M. Neal. *Bayesian learning for neural networks*. Number 118 in Lecture notes in statistics. Springer, New York, 1996.

[86] M. Niedzwiecki and T. Klaput. Fast recursive basis functions estimators for identification of time-varying processes. In *2000 IEEE International Conference on*, volume 1, pages 1–4. IEEE, 2000.

[87] Brett Ninness and Soren Henriksen. Bayesian system identification via Markov chain Monte Carlo techniques. *Automatica*, 46(1):40–51, January 2010.

[88] R. W. Ogden. Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 326(1567):565–584, February 1972.

[89] Ailish O'Halloran, Fergal O'Malley, and Peter McHugh. A review on dielectric elastomer actuators, technology, applications, and challenges. *Journal of Applied Physics*, 104(7):071101, 2008.

[90] Muzaffer Y. Ozsecen and Constantinos Mavroidis. Nonlinear force control of dielectric electroactive polymer actuators. March 2010.

[91] Ronald K. Pearson. *Discrete-time dynamic models*. Topics in chemical engineering. Oxford University Press, New York, 1999.

[92] Qibing Pei, Marcus A. Rosenthal, Ron Pelrine, Scott Stanford, and Roy D. Kornbluh. Multifunctional electroelastomer roll actuators and their application for biomimetic walking robots. In *Smart Structures and Materials*, volume 5051, pages 281–290, 2003.

[93] R. Pelrine. High-Speed Electrically Actuated Elastomers with Strain Greater Than 100%. *Science*, 287(5454):836–839, February 2000.

[94] Z.K. Peng, Z.Q. Lang, C. Wolters, S.A Billings, and K. Worden. Feasibility study of structural damage detection using NARMAX modelling and Nonlinear Output Frequency Response Function based analysis. *Mechanical Systems and Signal Processing*, 25(3):1045–1061, April 2011.

[95] J.C. Peyton Jones. Mean levels in nonlinear analysis and identification. *International Journal of Control*, 58(5):1033–1052, 1993.

[96] R. Pfeifer, M. Lungarella, and F. Iida. Self-Organization, Embodiment, and Biologically Inspired Robotics. *Science*, 318(5853):1088–1093, November 2007.

[97] Rik Pintelon and Johan Schoukens. Measurement of frequency response functions using periodic excitations, corrupted by correlated input/output errors. *Instrumentation and Measurement, IEEE Transactions on*, 50(6):1753–1760, 2001.

[98] L. Piroddi, V. Seghezza, and M. Bonin. NARX model selection based on simulation error minimisation and LASSO. *IET Control Theory & Applications*, 4(7):1157–1168, July 2010.

[99] L. Piroddi and W. Spinelli. An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781, November 2003.

[100] Jean-Sébastien Plante and Steven Dubowsky. On the performance mechanisms of Dielectric Elastomer Actuators. *Sensors and Actuators A: Physical*, 137(1):96–109, June 2007.

[101] Kailiang Ren, Sheng Liu, Minren Lin, Yong Wang, and Q.M. Zhang. A compact electroactive polymer actuator suitable for refreshable Braille display. *Sensors and Actuators A: Physical*, 143(2):335–342, May 2008.

[102] Nick M. Ridler and Martin J. Salter. An approach to the treatment of uncertainty in complex S-parameter measurements. *Metrologia*, 39(3):295, 2002.

[103] Michael T. Rosenstein, James J. Collins, and Carlo J. De Luca. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D: Nonlinear Phenomena*, 73(1-2):82–98, May 1994.

[104] Samuel Rosset, Benjamin M O'Brien, Todd Gisby, Daniel Xu, Herbert R Shea, and Iain A Anderson. Self-sensing dielectric elastomer actuators in closed-loop operation. *Smart Materials and Structures*, 22(10):104018, October 2013.

[105] Tom Rutherford. Population ageing: statistics - Commons Library Standard Note. `http://www.parliament.uk/Templates/BriefingPapers/Pages/BPPdfDownload.aspx?bp-id=sn03228`, February 2012.

[106] Masa-Aki Sato. Online model selection based on the variational Bayes. *Neural Computation*, 13(7):1649–1681, 2001.

[107] Martin Schetzen. *The Volterra and Wiener theories of nonlinear systems*. Wiley, New York, 1980.

[108] A. Schmidt, A. E. Bergamini, G. Kovacs, and E. Mazza. Experimental characterization and modeling of circular actuators made of interpenetrating polymer network-reinforced acrylic elastomer. *Journal of Intelligent Material Systems and Structures*, 24(10):1257–1265, January 2013.

[109] Todd Schultz, Mark Sheplak, and Louis N. Cattafesta. Application of multivariate uncertainty analysis to frequency response function estimates. *Journal of Sound and Vibration*, 305(1-2):116–133, August 2007.

[110] Gideon Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, March 1978.

[111] Robert F. Shepherd, Filip Ilievski, Wonjae Choi, Stephen A. Morin, Adam A. Stokes, Aaron D. Mazzeo, Xin Chen, Michael Wang, and George M. Whitesides. Multigait soft robot. *Proceedings of the National Academy of Sciences*, 108(51):20400–20403, 2011.

[112] Alex Sherstinsky and Rosalind W. Picard. On the efficiency of the orthogonal least squares training method for radial basis function networks. *Neural Networks, IEEE Transactions on*, 7(1):195–200, 1996.

[113] Jonas Sjöberg, Qinghua Zhang, Lennart Ljung, Albert Benveniste, Bernard Delyon, Pierre-Yves Glorennec, H\aakan Hjalmarsson, and Anatoli Juditsky.

Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

[114] J.J. E. Slotine and Weiping Li. *Applied nonlinear control*. Prentice Hall, Englewood Cliffs, N.J, 1991.

[115] Václav Šmídl. *The Variational Bayes Approach in Signal Processing*. PhD thesis, Trinity College, 2004.

[116] Torsten Söderström. *System identification*. Prentice Hall International series in systems and control engineering. Prentice Hall, New York, 1989.

[117] Michio Sugeno and Takahiro Yasukawa. A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on fuzzy systems*, 1(1):7–31, 1993.

[118] Z Suo, X Zhao, and W Greene. A nonlinear field theory of deformable dielectrics. *Journal of the Mechanics and Physics of Solids*, 56(2):467–486, February 2008.

[119] Zhigang Suo. Theory of dielectric elastomers. *Acta Mechanica Solida Sinica*, 23(6):549–578, December 2010.

[120] Tomohiro Takagi and Michio Sugeno. Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, (1):116–132, 1985.

[121] Tara Baldacchino. *Statistical Estimation and Identification of Nonlinear Dynamic Systems*. PhD thesis, The University of Sheffield, 2011.

[122] Thomas E. Fricker, Jeremy E. Oakley, Neil D. Sims, and Keith Worden. Probabilistic uncertainty analysis of an FRF of a structure using a Gaussian process emulator. *Mechanical Systems and Signal Processing*, 25(8):2962–2975, 2011.

[123] M. E. Tipping. Sparse bayesian learning and the relevance vector machine. *The Journal of Machine Learning Research*, 1:211–244, September 2001.

[124] Deepak Trivedi, Christopher D. Rahn, William M. Kier, and Ian D. Walker. Soft robotics: Biological inspiration, state of the art, and future research. *Applied Bionics and Biomechanics*, 5(3):99–117, December 2008.

[125] Paul T. Troughton and Simon J. Godsill. Bayesian model selection for time series using Markov chain Monte Carlo. In *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, volume 5, pages 3733–3736. IEEE, 1997.

[126] Nikolaos G. Tsagarakis and Darwin G. Caldwell. Development and control of a 'soft-actuated' exoskeleton for use in physiotherapy and training. *Autonomous Robots*, 15(1):21–33, 2003.

[127] B. Vanderborght, A. Albu-Schaeffer, A. Bicchi, E. Burdet, D.G. Caldwell, R. Carloni, M. Catalano, O. Eiberger, W. Friedl, G. Ganesh, M. Garabini, M. Grebenstein, G. Grioli, S. Haddadin, H. Hoppner, A. Jafari, M. Laffranchi, D. Lefeber, F. Petit, S. Stramigioli, N. Tsagarakis, M. Van Damme, R. Van Ham, L.C. Visser, and S. Wolf. Variable impedance actuators: A review. *Robotics and Autonomous Systems*, 61(12):1601–1614, December 2013.

[128] Rocco Vertechy, Antonio Frisoli, Massimo Bergamasco, Federico Carpi, Gabriele Frediani, and Danilo De Rossi. Modeling and experimental validation of buckling dielectric elastomer actuators. *Smart Materials and Structures*, 21(9), September 2012.

[129] Yongquan Wang, Huanhuan Xue, Hualing Chen, and Junhua Qiang. A dynamic visco-hyperelastic model of dielectric elastomers and their energy dissipation characteristics. *Applied Physics A*, May 2013.

[130] H. L. Wei, S. A. Billings, and J. Liu. Term and variable selection for non-linear system identification. *International Journal of Control*, 77(1):86–110, January 2004.

[131] Hua-Liang Wei and S. A. Billings. Identification of time-varying systems using multiresolution wavelet models. *International Journal of Systems Science*, 33(15):1217–1228, 2002.

[132] R. Willink and B. D. Hall. A classical method for uncertainty analysis with multidimensional data. *Metrologia*, 39(4):361, 2002.

[133] Emma D. Wilson, Tareq Assaf, Martin J. Pearson, Jonathan M. Rossiter, Sean R. Anderson, and John Porrill. Bioinspired Adaptive Control for Artificial Muscles. In *Biomimetic and Biohybrid Systems*, volume 8064, pages 311–322. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[134] Michael Wissler and Edoardo Mazza. Modeling and simulation of dielectric elastomer actuators. *Smart Materials and Structures*, 14(6):1396–1402, December 2005.

[135] Michael Wissler and Edoardo Mazza. Electromechanical coupling in dielectric elastomer actuators. *Sensors and Actuators A: Physical*, 138(2):384–393, August 2007.

[136] K. Worden. Confidence bounds for frequency response functions from time series models. *Mechanical Systems and Signal Processing*, 12(4):559–569, July 1998.

[137] Feng Xia, Srinivas Tadigadapa, and Q.M. Zhang. Electroactive polymer based microfluidic pump. *Sensors and Actuators A: Physical*, 125(2):346–352, January 2006.

[138] Xin Xia, Jianzhong Zhou, Chaoshun Li, and Wenlong Zhu. A novel method for fault diagnosis of hydro generator based on NOFRFs. *International Journal of Electrical Power & Energy Systems*, 71:60–67, October 2015.

[139] Sheng Q. Xie, P. F. Ramson, D. D. Graaf, E. P. Calius, and Iain A. Anderson. An adaptive control system for dielectric elastomers. In *Industrial Technology, 2005. ICIT 2005. IEEE International Conference on*, pages 335–340. IEEE, 2005.

[140] O. H. Yeoh. Characterization of Elastic Properties of Carbon-Black-Filled Rubber Vulcanizates. *Rubber Chemistry and Technology*, 63(5):792–805, November 1990.

[141] R. Yue, S. A. Billings, and Z.-Q. Lang. An investigation into the characteristics of non-linear frequency response functions. Part 1: Understanding the higher dimensional frequency spaces. *International Journal of Control*, 78(13):1031–1044, September 2005.

[142] Y. Zheng, Z. Lin, and D. B. H. Tay. Time-varying parametric system multiresolution identification by wavelets. *International Journal of Systems Science*, 32(6):775–793, January 2001.