Blind Identification of the Electromechanical Modes of a Power System using a Wiener Model

by

Tasnim Hossain

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

CALGARY, ALBERTA

APRIL, 2016

© Tasnim Hossain 2016
Abstract

In this thesis, a blind system identification technique using a Wiener model is used to estimate the power system modes. A Wiener model, a universal approximator consisting of a dynamic linear system followed by a memoryless nonlinear element, is used to estimate the power system nonlinearities. It also constructs a set of intermediate data, which can be used by a linear estimation technique such as subspace identification for estimating the power system electromechanical modes. In this research, a blind variant of a subspace method, Numerical Algorithm for Subspace State Space System IDentification (N4SID), is used. The algorithm is tested with simulation data from the Kundur two-area network. The accuracy and reliability of these estimates are accessed by carrying out Monte Carlo simulations. The estimated results obtained from the simulated system using a Wiener model were very accurate with reduced prediction errors and was a good fit for the power system.
Acknowledgements

This thesis would not have been possible without the help and cooperation of several persons. My supervisor, Dr. David Westwick deserves a great appreciation for guiding me through the difficulties and challenges faced during the research. His vast knowledge of system identification was vital for this research. His patience, motivation and guidance helped me a lot all throughout the research. I am also very thankful to my co-supervisor, Dr. William Rosehart. He introduced me to the concept of power system oscillations. I appreciate his views and suggestions that improved the final draft of the thesis to a great extent. Both of them has been very invaluable to me.

I would like to thank my colleagues who provided their insight and expertise that greatly assisted the research. I am also thankful to them for providing a fun and stimulating work environment.

I am very thankful to Shah, my husband, best friend, and constant companion, for being so supportive and putting up with me the whole time. His love and support enhanced my life to a large extent especially towards the end. I will be forever grateful.

Last but not the least, I would thank my parents, Azmal and Farida, for their love, encouragement and support.
# Table of Contents

Abstract ............................................................. ii  
Acknowledgements .................................................. iii  
Table of Contents .................................................... iv  
List of Tables ......................................................... vi  
List of Figures ....................................................... vii  
List of Symbols ...................................................... viii  
1 INTRODUCTION ...................................................... 1  
  1.1 Overview ....................................................... 1  
  1.2 Literature Review ............................................... 5  
    1.2.1 Ringdown Signal Analysis Methods .............................. 5  
    1.2.2 Ambient Data Analysis Methods ................................ 8  
  1.3 Objectives and Contributions .................................... 11  
  1.4 Overview and Organization ...................................... 13  
2 BACKGROUND REVIEW ............................................... 15  
  2.1 Introduction ................................................... 15  
  2.2 System Identification ........................................... 15  
  2.3 Linear Time Invariant Systems ................................... 18  
    2.3.1 Blind Identification .......................................... 21  
    2.3.2 Multi Channel Systems ......................................... 22  
  2.4 Subspace identification .......................................... 22  
    2.4.1 State space models ............................................ 23  
    2.4.2 Stochastic Subspace Identification ............................ 24  
    2.4.3 Conversion of Discrete-time to Continuous-time Model ....... 31  
  2.5 Mode Estimation ................................................ 32  
    2.5.1 Power System Model .......................................... 34  
    2.5.2 Small Signal Analysis ........................................ 36  
  2.6 Summary ....................................................... 44  
3 NONLINEAR IDENTIFICATION ....................................... 45  
  3.1 Introduction ................................................... 45  
  3.2 Power System Nonlinearity ...................................... 46  
  3.3 Nonlinearity Detection .......................................... 48  
  3.4 Nonlinear Systems .............................................. 52  
    3.4.1 Volterra and Wiener Series .................................... 52  
    3.4.2 Wiener Model ................................................ 54  
    3.4.3 Blind Identification of Wiener Model .......................... 55  
  3.5 Summary ....................................................... 60  
4 BLIND IDENTIFICATION OF ELECTROMECHANICAL MODES USING A WIENER MODEL ........................................... 61  
  4.1 Introduction ................................................... 61  
  4.2 Proposed Algorithm ............................................. 61  
    4.2.1 Linearity Test ................................................. 62  
    4.2.2 Estimating Nonlinearity ....................................... 63
### List of Tables

1.1 Modal Estimation Techniques .................................................. 11

4.1 Dynamic states of a sub transient generator .................................. 67
4.2 Comparison of estimated results with the real modes of Kundur two area network .............................................................. 77
4.3 Estimated values of Mode shape1 normalized w.r.t Bus 1 ..................... 80
4.4 Estimated values of Mode shape 2 normalized w.r.t Bus 1 ..................... 80
4.5 Comparing model fits with linear and Wiener Model .......................... 81
List of Figures and Illustrations

1.1 Initial modelling failure in the California Oregon Intertie for WECC outage [42] 2
2.1 Representation of a system ................................. 16
2.2 System Identification Procedure ............................ 17
2.3 SISO LTI System ............................................. 19
2.4 System with known outputs, \( y(t) \) and unknown inputs, \( e(t) \) ........ 21
2.5 An example of a local mode ................................ 33
2.6 An example of an inter-area mode ......................... 34
3.1 Block diagram of Wiener series approximator ............. 55
3.2 Wiener Model .................................................. 57
4.1 Kundur two-area Power System ............................ 68
4.2 Time series of Voltage Angle of Bus 3 .................... 69
4.3 Step.1: Passing data through first order derivative filter .... 70
4.4 Step.2: Passing through low pass filter .................... 70
4.5 Step.3: Downsampling ....................................... 71
4.6 Step.4: Passing through high pass filter ................. 71
4.7 Step.5: Getting rid of transients ......................... 72
4.8 Gaussianity test results of power system data .......... 74
4.9 Gaussianity test results of intermediate signal data from the Wiener model .... 75
4.10 Linearity test results of intermediate signal data from the Wiener model .... 75
4.11 Inverse nonlinear function using the Wiener model .... 76
4.12 Estimating the inverse nonlinear function using different methods and their corresponding fit errors. .................. 77
4.13 Comparison of Mode shape 1 ............................. 79
4.14 Comparison of Mode shape 2 ............................. 79
4.15 PDF of prediction errors using Wiener and linear models .... 82
4.16 Correlation plots of the residuals from Output \( y_5 \) using a Wiener and Linear Model Data .................. 83
4.17 Singular Values of the Prediction Error ................... 84
### List of Symbols, Abbreviations and Nomenclature

<table>
<thead>
<tr>
<th>NOTATION</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AR</strong></td>
<td>Auto Regressive</td>
</tr>
<tr>
<td><strong>ARMA</strong></td>
<td>Auto Regressive Moving Average</td>
</tr>
<tr>
<td><strong>CSD</strong></td>
<td>Cross Spectral Density</td>
</tr>
<tr>
<td><strong>ERA</strong></td>
<td>Eigensystem Realization Algorithm</td>
</tr>
<tr>
<td><strong>EMD</strong></td>
<td>Empirical Mode Decomposition</td>
</tr>
<tr>
<td><strong>FDD</strong></td>
<td>Frequency Domain Decomposition</td>
</tr>
<tr>
<td><strong>FFT</strong></td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td><strong>LS</strong></td>
<td>Least Squares</td>
</tr>
<tr>
<td><strong>LTI</strong></td>
<td>Linear Time Invariant</td>
</tr>
<tr>
<td><strong>MIMO</strong></td>
<td>Multiple Input Multiple Output</td>
</tr>
<tr>
<td><strong>MISO</strong></td>
<td>Multiple Input Single Output</td>
</tr>
<tr>
<td><strong>MEYW</strong></td>
<td>Modified Extended Yule-Walker</td>
</tr>
<tr>
<td><strong>N4SID</strong></td>
<td>Numerical Algorithm for Subspace State Space System</td>
</tr>
<tr>
<td><strong>PDF</strong></td>
<td>Probability Density Function</td>
</tr>
<tr>
<td><strong>PEM</strong></td>
<td>Prediction Error Method</td>
</tr>
<tr>
<td><strong>PMU</strong></td>
<td>Phasor Measurement Unit</td>
</tr>
<tr>
<td><strong>PSD</strong></td>
<td>Power Spectral Density</td>
</tr>
<tr>
<td><strong>R3LS</strong></td>
<td>Regularized Robust Recursive Least Squares</td>
</tr>
<tr>
<td><strong>RLS</strong></td>
<td>Recursive Least Squares</td>
</tr>
<tr>
<td><strong>RRLS</strong></td>
<td>Robust Recursive Least Squares</td>
</tr>
<tr>
<td><strong>SISO</strong></td>
<td>Single Input Single Output</td>
</tr>
</tbody>
</table>
**SIMO**  
Single Input Multiple Output

**SVD**  
Singular Value Decomposition

**WECC**  
Western Electricity Coordinating Council

**YW**  
Yule Walker

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>State Vector</td>
</tr>
<tr>
<td>$u$</td>
<td>Input Vector</td>
</tr>
<tr>
<td>$y$</td>
<td>Output Vector</td>
</tr>
<tr>
<td>$e$</td>
<td>Estimation Error Vector</td>
</tr>
<tr>
<td>$A$</td>
<td>System State Matrix</td>
</tr>
<tr>
<td>$B$</td>
<td>Input Matrix</td>
</tr>
<tr>
<td>$C$</td>
<td>Output Matrix</td>
</tr>
<tr>
<td>$D$</td>
<td>Control Output Matrix</td>
</tr>
<tr>
<td>$K$</td>
<td>Noise Gain Matrix</td>
</tr>
<tr>
<td>$I$</td>
<td>Unit Matrix</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Eigenvalue</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Damping Ratio</td>
</tr>
<tr>
<td>$f$</td>
<td>Frequency</td>
</tr>
<tr>
<td>$v$</td>
<td>Right Eigenvector</td>
</tr>
<tr>
<td>$u$</td>
<td>Left Eigenvector</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>Normal Standard Distribution</td>
</tr>
<tr>
<td>$\text{det}$</td>
<td>Determinant</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Overview

Electromechanical modes are natural oscillations that occur in interconnected power systems. They indicate the stability of the power system to sudden swings or small disturbances. The synchronous nature of power systems tends to keep the speed of all the generators at a constant nominal value, whereas natural electrical load variations cause fluctuations in the generator speeds. These exchanges of electromagnetic and mechanical energy result in oscillations dictated by the electromechanical modes.

Low frequency electromechanical modes can be classified into two types: (i) local modes and (ii) inter-area modes. Local modes result from the swinging of one machine against a group of machines. They generally occur in the frequency range 1-2 Hz. Inter-area modes refer to the swinging of a group of machines in one area against a group of machines in another area. Their frequencies usually range from 0.1-1Hz [20], [41]. These modes are usually associated with weak transmission links, random load fluctuations and heavy power flows. Characterizing the electromechanical modes means measurement of the frequency, damping ratio, amplitude and phasing of the natural oscillations. It is shown in [20] that the eigenvalues and eigenvectors of a linearized state space model are the modes and mode shapes respectively. The mode is usually described by a complex value, whose real and imaginary parts determine the damping and frequency of the mode. Damping provides direct evaluation of the mode’s stability. The mode shape determines the amplitude and phasing of the mode. It helps in identifying the most powerful/energetic oscillations.

The potential for unstable electromechanical oscillations remains one of the major concerns for reliable operation of power systems. Nowadays with increased interconnectivity, random
load fluctuations and increased power transfers, electromechanical oscillations are inevitable. The ability to accurately estimate these modes is crucial for safe and reliable operation of the power system. If these oscillations persist for a certain period of time, not only will there be disruption in power transfer but they may also lead to a system breakdown. There were several outages in the past due to the inability to estimate, and hence control, unstable modes. For example, an unidentified mode which grew out of control led to system wide outage in 1996 in the Western Electricity Coordinating Council (WECC) [19]. The Northern and Southern parts of the WECC began to oscillate against each other. When these growing oscillations reached an amplitude of about 1000 MW, the WECC was broken into four separate islands. A total of 30,000 MW load and 27,000 MW generation was lost. 7.5 million customers across Western Canada, the Western U.S. and NW Mexico suffered lost connections. Figure 1.1 shows that the planning models predicted the mode as stable even when the oscillations broke the system apart. In order to prevent consequences from getting as severe as the above example, lightly damped low-frequency oscillations must be identified.

Figure 1.1: Initial modelling failure in the California Oregon Intertie for WECC outage [42]
using reliable and accurate models

Electromechanical modes cannot be measured directly because power systems have random components derived from random load variations which are constantly changing [9]. This is the reason that they must be estimated. The accuracy of the estimates is determined by many factors including the type of model used and the statistical properties of the estimation method. In 1996, the standard approach involved linearizing the power system model and estimating the system’s modes from the linearized model. Unfortunately the method was proven to be inaccurate and WECC went through a massive power outage [19].

Estimating the electromechanical modes of a power system can be challenging because the underlying system is both stochastic and nonlinear. The stochastic nature of the system is due to random changes in load etc. Power systems are highly nonlinear due to nonlinear elements such as generators, loads etc. The interconnection of these different individual elements causes a large variety of dynamic interaction within the whole system [25]. Power system nonlinearities can affect the accuracy and reliability of the estimators. In [54], different kinds of power system nonlinearities affecting modal estimation are acknowledged and modelled. Generator nonlinear effects were modelled as the swing equations, magnetic saturation and excitations. Load nonlinear effects were modelled as constant impedance, constant current and constant power loads. Power systems are continuously experiencing changes in the loading and pattern of generation [35]. Power systems are also prone to disturbances which tend to introduce nonlinearity into the power system. Examples of dramatic disturbances are: lightning causing a line to be tripped, boiler tube failure resulting in a generator going off-line and lines sagging on trees being tripped [28]. Therefore a method to accurately estimate power system modes considering the effect of nonlinearity is needed, that is a nonlinear estimator.

Wiener nonlinear model structures have been used in several applications such as the microwave and RF technology [5], chemical process industry [31], biology [14] and physiology
Wiener models have a very simple structure - a linear time invariant system followed by a static nonlinearity [23]. Modern power systems are highly nonlinear and as a result, data obtained from them cannot be accurately explained by a linear model. Wiener models can place all the nonlinearity present anywhere in the power system at the output of the model, which does not necessarily correspond to the layout of the physical system. However, Wiener systems with multiple connections between the linear and nonlinear elements have been shown to be universal approximator in [3]. In this thesis, Wiener models with multiple input multiple output (MIMO) linear parts and MIMO nonlinearities will be used to estimate the power system modes. Wiener models are further detailed in Chapter 3.

Modal estimators require power system data to estimate modes. Different estimation methods analyze different kinds of power system data. There are three different kinds of power system data:

1. Ambient: When the power system is in equilibrium and undergoes small disturbances such as small amplitude random load changes, the measurement data is known as ambient data [37].

2. Ring-down: Measurement data as a result of large disturbances such as faults, line tripping, loss in transmission are ring-down data. They usually result in observable oscillations [11], [8].

3. Probing: The data obtained by intentionally injecting random noises into the power system are probing responses. These type of data are used to test the performance of the power system [69], [38].

In this thesis, ambient data is used. Ambient data though considered to be linear might have certain nonlinearities due to the nonlinear nature of the power systems. These nonlinearities must be considered by the estimator for accurate estimation results with minimum errors.
The remainder of the chapter is organized with a brief review of previous research methods used to estimate electromechanical modal properties of a power system. Then, the original technique and the contributions made by the thesis are outlined. At last, a general overview of the thesis is provided.

1.2 Literature Review

There are several methods used to estimate the electromechanical modes of a power system. The methods can be classified according to the type of data used for analysis.

1.2.1 Ringdown Signal Analysis Methods

Prony analysis:

Prony analysis originated in an earlier century [39]. Prony’s methods along with their added extensions are designed to estimate the parameters represented in an exponential form by fitting functions to an observed amount of output data. That is; it is a polynomial method which includes the process of finding the roots of a polynomial.

Let the output data \( y(t) \) consist of \( N \) samples \( y(t_k) \) that are evenly spaced by \( \Delta t \). The notation in simplified exponential form will be:

\[
\hat{y}(t) = \sum_{i=1}^{n} A_i e^{\sigma_i t} \cos(\omega_i t + \theta_i)
\]

Using sample times \( t_k \) Equation 1.1 can be discretized to:

\[
\hat{y}(k) = \sum_{i=1}^{n} B_i z_i^k
\]

where \( z_i \) are the roots of the polynomial. The method can be summarized in the following steps:

1. The selected data samples are assembled into a Toeplitz data matrix,
2. The data is fitted into discrete linear prediction model for example least squares.
3. The roots of the characteristic polynomial Equation 1.2 associated with the model in step 1 are determined.

4. Using the roots, the modal frequencies, amplitude and phasing for each mode are determined.

Prony analysis was used in [11] and [8] where it was applied to nonlinear systems and shown to include the effects of nonlinearity. In [11], Prony analysis is used to estimate the modal properties. In [57], local and inter area modes were determined by Prony analysis for obtaining the transfer functions of Power System Stabilizer (PSS) using root-locus and sequential decentralized control techniques.

Eigensystem Realization Algorithm (ERA):
Eigensystem Realization Algorithm is a system identification algorithm introduced in [36]. The algorithm is used for modal identification and model reduction of linear systems. The ERA was introduced within the aerospace community and has been used extensively. ERA was used to estimate modes in [16]. ERA is based on the singular value decomposition (SVD) of the Hankel matrix of data with the linear ringdown of the system. The method is as follows:

1. The selected data samples are assembled into a Hankel data matrices.

2. SVD of the Hankel matrices is performed to estimate the system order, which is equal to the number of non-zero singular values.

3. The discrete system matrices are computed.

4. The discrete matrices are converted to continuous system matrices and the system response computed.

5. The eigenvalues of system state matrix $A$ are the poles/modes of the system.
Matrix Pencil Method:
The Matrix Pencil method was introduced by [13] for pole estimation. Initially it was used for extracting poles from a antennas’s electromagnetic transient responses. The matrix pencil method produces a matrix whose roots provide modes of the system. The steps are summarized as follows:

1. The selected data samples are assembled into Hankel data matrices.

2. The data is fitted into discrete linear prediction model, for example: least squares.

3. The $V$ matrices from SVD of the Hankel matrices are used to create $Y_1 = V_1^T V_1$ and $Y_2 = V_2^T V_2$ matrices.

4. The desired poles are found as the generalized eigenvalues of the matrix pair $[Y_2] - \lambda [Y_1]$. The eigenvalue set $\lambda(Y_1, Y_2)$ is contained in the square matrices $Y_1$ and $Y_2$, as the pencil values or roots of $Y_2$ relative to $Y_1$.

Modes were accurately estimated using matrix pencil in [21].

Hilbert Spectral Analysis:
A complex signal is constructed from the data series by adding an imaginary signal to the original function [27]. This is known as the Hilbert transform. The real part is an interpretation of the damping and the imaginary part is the frequency.

Empirical Mode Decomposition(EMD):
EMD provides an analytical basis for the nonlinear decomposition of any signal into a finite set of essentially band-limited components or basis functions called Intrinsic Mode Functions (IMFs). The IMF has two parts: a slowly varying residue and a fast component superimposed on the slow component. The IMFs are extracted by a iterative process called sifting from
the original signal. The IMFs are checked for maxima and minima and then the process is repeated until there are no longer any maxima or minima in the residual. Then the last stage involves forming a complex signal and then extracting modal frequency and damping associated with the oscillating components using Hilbert analysis. Details of the procedure can be found in [27].

Other algorithms applied to ringdown data signals include Wavelet Decomposition [17], Hankel total least squares methods [21] and Steiglitz-McBride [43]. For more information the referenced papers can be referred to.

1.2.2 Ambient Data Analysis Methods

Ambient data analysis has been done with frequency and time domain analysis.

Spectral Method:
The Spectral method uses the spectral density functions: cross spectral density (CSD) and power spectral density (PSD). CSD is a complex function of frequency of measured signals and PSD is a positive real function of the frequencies.

1. The measured signals are tested for the number of observable states.
2. If the measured signals have highly observable states, the peaks of the PSDs will occur near the frequencies of the modes.
3. The phasing of the mode is directly estimated by the angle of the CSD.
4. The mode shape is estimated from the ratio of CSD between two different frequencies and the PSD of one of the frequency.

Spectral method uses any of the algorithms in [46] to determine the spectral density functions. In [10] and [52] different methods to estimate the spectral density functions are shown.
Frequency Domain Decomposition (FDD):

FDD requires the singular value decomposition (SVD) of an estimated spectral density matrix for the measured signals.

1. A spectral density matrix is estimated from the measured data.

2. SVD of the spectral density matrix is carried out.

3. The mode in the vicinity of any given frequency will have the largest relative singular value.

4. The mode shapes are found by taking the ratio of the elements of the singular vectors corresponding to the largest singular value.

FDD for modal estimation in power systems was used in [22].

Neither the Spectral nor the FDD method provide any information about the damping as they do not assume anything about the model that is driving the system. Prony’s method applied to the inverse discrete Fourier Transform (IDFT) of the PSD resulted in an autocorrelation sequence. These estimated autocorrelation sequences from several data channels could be applied to the multichannel Prony’s method of [56]. Similarly for FDD, Prony’s method is applied to signal obtained from IDFT of the frequency domain sequence of the largest singular value [22]. These approaches succeeded in providing the estimates of the damping.

Yule Walker (YW) Method:

The Yule Walker (YW) was used for mode estimation in [37].

1. The ambient power system data are applied to autoregressive (AR) model.

2. The mode estimates are derived from the characteristic AR equations.
Power system data were fit into autoregressive moving average (ARMA) models using Modified Extended Yule Walker (MEYW) in [66]. The mode estimates derived by the ARMA equations.

Channel Matching Method:
A method of estimating the mode shape through a channel matching algorithm was proposed in [6]. The method can be summarized as:

1. A narrow-band filter is constructed to isolate a required mode.
2. The mode shape is estimated by constructing a first order filter between two different measured power system outputs, comparing their amplitude and phasing.
3. The response of the filter solved at the modal frequency is used as the mode shape. This provided an estimate of the magnitude and phasing of one signal with respect to another.

In [68] improved results were achieved by constructing a transfer function between the pair of outputs instead. The response of the transfer function at the complex mode was used as the mode shape estimate.

Recursive Least Squares (RLS), Robust Recursive Least Squares (RRLS) and Regularized Robust Recursive Least Squares (R3LS):
Online estimation of modes were also carried out. Least Mean Squares (LMS) algorithm was used in [67] to estimate power system modes. From [58], Recursive Least Squares (RLS) algorithms were applied for power system mode estimation. Robust Recursive Least Squares (RRLS) and Regularized Robust Recursive Least Squares (R3LS) algorithms were used in [70] and [72], respectively. All these methods use online parametric modal estimation methods.
Subspace Identification:
Subspace methods have been used to estimate modal properties in power systems. In [69] subspace identification using Canonical Variate Algorithm (CVA) and in [71] using Numerical Algorithm for Subspace State Space System Identification (N4SID) were used for power system mode estimation. Basic subspace approach to estimate modes was used in [15]. All these methods are same, just different weighted versions of the basic subspace method [60]. The modes are estimated by the dynamic state matrix of the linearized subspace model. This method is explained in details in Chapter 2. N4SID forms the linear dynamic part of the Wiener model used in this thesis.

All the discussed methods are summarized in Table 1.1

<table>
<thead>
<tr>
<th>Ambient Data Analysis</th>
<th>Ringdown Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectral Method</td>
<td>Prony analysis</td>
</tr>
<tr>
<td>FDD</td>
<td>Eigenvalue Realization Algorithm (ERA)</td>
</tr>
<tr>
<td>YW and Modified YW on AR and ARMA</td>
<td>Matrix Pencil Method</td>
</tr>
<tr>
<td>RLS, RRLS, R3LS</td>
<td>Empirical Mode Decomposition (EMD)</td>
</tr>
<tr>
<td>Channel Matching Method</td>
<td>Hilbert spectral analysis</td>
</tr>
<tr>
<td>TF Method</td>
<td>Hankel Total Least Squares</td>
</tr>
<tr>
<td>Subspace Identification: N4SID, CVA</td>
<td>Steiglitz-McBride</td>
</tr>
</tbody>
</table>

1.3 Objectives and Contributions

Different methods of estimating the electromechanical modes of a power system used previously were stated in the last section. Subspace Identification of power system modes was carried out successfully in [69],[71] and [15]. In [7], subspace identification using the N4SID algorithm was considered to be the best overall mode estimator when compared to other estimators. N4SID could provide very good results during ambient conditions, but when these conditions were interrupted with nonlinear disturbances, the estimation results were
compromised. However, power systems consist of many interconnected nonlinear elements so the presence of nonlinearity in power system data is inevitable. In the current form, subspace identification cannot be used to estimate modes from nonlinear data with high accuracy. Wiener models, dynamic linear systems followed by memoryless nonlinearities, are relatively simple nonlinear systems that nevertheless have universal approximation properties, provided that there can be an arbitrary number of connections between the linear and nonlinear elements. The chief objective of this thesis is to provide a unique system identification method for estimating the modes of a power system using the Wiener model to account for any nonlinearity present in the data, and hence reduce the estimation error and increase the prediction accuracy, as compared to a purely linear approach. The contributions of the thesis are:

1. Estimating the nonlinearity present in the power system data.
2. Accurately estimating the electromechanical modal frequency, damping ratio and mode shape from nonlinear power system data.
3. Achieving reduced prediction errors.

These contributions are significant because eliminating the nonlinearities from the data reduces the effective noise level, when considering a linear model, and hence reduces variance of the results and provides more reliable and accurate results. The model prediction errors are reduced getting an improved model for power systems. Using a Wiener model, the effects of nonlinearity present anywhere within the power system can be modelled and removed from the data. These data can then be used along with subspace Identification using N4SID to estimate modes from any kind of data. That is; the model is suitable for both simulated and real data as it can estimate power system modes from both ringdown and ambient data with minimum estimation errors.
1.4 Overview and Organization

This chapter provides a brief literature review involved in mode estimation of power systems. For more detailed information about these methods, the provided references can be read. The remainder of the thesis is organized as follows:

- Chapter 2 provides a background review of system identification theory. It consists of linear system identification of multiple input multiple output (MIMO) systems using subspace identification. Stochastic subspace identification using Numerical Algorithm for Subspace State Space System Identification (N4SID) is described. Usage of these system identification tools leads to MIMO blind identification of electromechanical modes of a power system. Modal estimation using eigenanalysis is also described. This chapter is an overview of system identification and mode estimation theory applied to this research.

- Chapter 3 recognizes the different nonlinearities present in a power system. How these nonlinearities can alter the estimation results are also stated here. The presence of nonlinearity of data can be detected using linearity tests. The linearity test used in this thesis is described here. Modal identification using nonlinear power system data can be identified more accurately using a nonlinear model, Wiener models are introduced. Wiener models used to determine the system nonlinearities are discussed.

- Chapter 4 presents the original work of the thesis. The method first of all, determines the presence of nonlinearity in power system data, and uses Wiener model to estimate the nonlinearity. The data is then used to identify electromechanical power system modes using subspace identification with N4SID. Simulation results from a test bus system, Kundur two area network, are provided. For validation, the results of the Wiener model are also compared with
a linear model. The performance of the introduced algorithm is compared using Monte Carlo simulations.

- The thesis is concluded in Chapter 5. It also reviews the overall contributions made in the thesis as well as a discussion of the future work that should be performed in this area of research.
Chapter 2

BACKGROUND REVIEW

2.1 Introduction

This thesis presents a method of estimating the electromechanical modes of power systems using system identification. This chapter provides a background information for the analytical techniques used in this thesis. It starts with a brief explanation of system identification theory which extends to multiple input multiple output (MIMO) and blind identification problems. Subspace system identification methods that perform blind identification on MIMO power systems are reviewed. These methods identify models of power systems that have been linearized to form a set of state space equations whose modes can be estimated using eigenanalysis.

2.2 System Identification

System Identification aims to provide a systematic process for quantifying the relations that regulate the behaviour of the system. A system is an object with different variables acting as the excitation and response of the system. The variables responsible for exciting the system that can be influenced by the user are known as the inputs. Other signals that affects the system but are not controlled by the user are disturbances. Outputs are referred to the signals that are produced by the system in response to the inputs and disturbances [23]. Figure 2.1 depicts a simple representation of a system.

System Identification is the modelling of any dynamic physical system from input/output measurements. The system identification procedure has a natural flow and is shown in Figure 2.2 and summarized below:
The first step of the process involves collection of input-output data records. The input-output data are recorded by conducting identification experiments. Usually a subset of the most relevant variables subject to the constraints introduced to the experiment is used. Therefore the user determines which signals to measure and when to measure or a user-defined set of inputs can be used to generate the outputs.

The measured data are preprocessed. There can be many different steps for preprocessing the data. Examples are averaging over all data sets, filtering disturbances, removing linear trends or detrending to remove slow drifts.

After the data have been recorded and preprocessed, different models are tried to fit the measured data in order to find the most suitable model. This assessment is based on how well the model reproduces the measured data. Examples of basic approaches would be using the least squares (LS) method. This is the most difficult and important step of system identification process.

Choosing of fit criterion determines how the different candidate models should be accessed for the most suitable model. Usually the model’s ability to reproduce the system’s measured output is the main criterion.

The next step determines the goodness of fit of the model candidates are done using appropriate tools. This determines the most suitable model required for
Finally, the model validation compares the predicted and measured signals to determine whether the model is good enough or valid for the purpose. These involves various procedures to determine how the chosen model relates to observed data and its intended use. Every model can eventually reproduce the estimation data but the challenge is to determine whether it can also describe fresh data sets from the system. A way to compare two different models would be to evaluate their performance on validation data. The model with the better performance is chosen. This is known as cross-validation.
A system model represents the relationship between observed signals and the system variables. A mathematical model uses mathematical expressions and equations to represent this relationship such as difference or differential equations. Models can be further classified continuous and discrete time, linear or nonlinear, etc. An example of a discrete linear time invariant (LTI) model:

\[
x(t + 1) = Ax(t) + Bu(t)
\]
\[
y(t) = Cx(t) + Du(t)
\]
where \(A, B, C\) and \(D\) becomes \(A(t), B(t), C(t)\) and \(D(t)\) respectively to form a linear time varying system.

An example of a nonlinear system model:

\[
x(t + 1) = F(x(t), u(t))
\]
\[
y(t) = G(x(t), u(t))
\]
where, \(F\) and \(G\) are nonlinear functions of their arguments. In general, a model structure is a parameterized function that generates the predicted outputs from measured data. That is

\[
\hat{y}(\theta) = M(\theta, u(1)...u(N), y(1)...y(N))
\]
where, \(\theta\) is a vector containing the parameters that describe the model.

### 2.3 Linear Time Invariant Systems

A single input single output (SISO) LTI system can be described using Equation(2.4) and is depicted by Figure 2.3.

\[
y(t) = G(q, \theta)u(t) + H(q, \theta)e(t)
\]
where, \(y(t)\) is the the system output, \(u(t)\) is the system input, \(e(t)\) is the unknown random system input, \(q\) is a shift operator, \(\theta\) is the parameter vector and \(G\) and \(H\) are linear functions of \(q\) and \(\theta\). \(\theta\) usually includes information of \(G\) and \(H\) and also the parameters
of the probability density function (PDF) of $e(t)$. In case of identifying modes, the system inputs and outputs, $u(t)$ and $y(t)$, are assumed to be measurable, and $e(t)$ is assumed to be zero mean Gaussian White Noise. The PDF of $e(t)$ can be described by the first and second moments.

![Figure 2.3: SISO LTI System](image)

\[
\text{mean of } e = E[e(t)] = 0 \quad (2.5)
\]

\[
\text{variance of } e = E[e^2(t)] = \sigma_e^2 \quad (2.6)
\]

The model depends on the unknown parameters of $G$, $H$ and $e(t)$. In order to reduce the number of unknowns, the one step ahead predictor is introduced:

\[
\hat{y}(t|\theta) = y(t) - e(t) \quad (2.7)
\]

Substituting Eq.(2.4) in Eq.(2.7)

\[
\hat{y}(t|\theta) = H^{-1}(q, \theta)G(q, \theta)u(t) + [1 - H^{-1}(q, \theta)]y(t) \quad (2.8)
\]

19
This predictor model only depends on the measured data and the parameters of \( G \) and \( H \). Linear functions \( G \) and \( H \) can be represented using black-box models. This involves parametrizing \( G \) and \( H \) as rational functions and letting the parameters be the numerator and denominator coefficients. Examples would be output error and Box-Jenkins model. Detailed descriptions of these model structures are given in [23], [64].

There is already a set of collected data from the system. The problem of deciding how to use this data to select an appropriate parameter vector \( \hat{\theta} \); that is the proper model selection from the set of candidate models is known as parameter estimation method. The prediction error of a certain model is as follows:

\[
\epsilon(t, \theta) = y(t) - \hat{y}(t|\theta) \tag{2.9}
\]

A model is considered good if it is good at predicting. So smaller the prediction errors, the better the model. The step to find the appropriate parameter set \( \hat{\theta} \) is then defined by the minimization of the prediction error. The cost function is introduced:

\[
V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} l(\epsilon(\theta)) \tag{2.10}
\]

where, \( V_N(\theta) \) is a scalar function of the parameters, \( l \) is a scalar positive function for \( V_N \geq 0 \) and is equal to zero only when the error is zero for all times.

This approach is the prediction error method (PEM). The next step would be to find the set of parameters which is derived from the minimum of the cost function:

\[
\hat{\theta} = \arg \min_{\theta} V_N(\theta) \tag{2.11}
\]

\( \arg \min \) gives the value of \( \hat{\theta} \) at which the function \( V_N() \) is minimized.

The PEM is a general framework for the system identification procedure. Equation 2.11 requires solving an optimization problem. One method for solving this optimization problem is the LS method. In this case the cost function is chosen as:

\[
[e(t)]^2 = e^2(t) \tag{2.12}
\]
If the model parameters appear linearly in the predictor (2.9), then Eq.(2.11) may be solved in closed form. If the predictor is nonlinear in the parameters, then (2.11) will be a nonlinear least squares problem, and its solution will require iterative methods.

2.3.1 Blind Identification

Blind system identification identifies the model parameters of a system using only the outputs. 'Blind' refers to the system’s input not being available for the identification process. This kind of identification is suitable for applications where the data is generated from an unknown system by unknown inputs. Comprehensive reviews can be found in [49] and [1]. Different applications of blind system identification are: speech recognition [48], communications [29], [50], seismic signal processing [30], [24] etc.

Usually in blind system identification the purpose is to identify the inputs and/or the system function from the the outputs only. Figure 2.4 depicts the blind system Identification problem. When the system is linear and time invariant, the system output $y(t)$ can be represented by the following model [1]:

$$y(t) = e(t) * h(.)$$

(2.13)

where, * indicates convolution and $h(.)$ is the impulse response of the system.

![Figure 2.4: System with known outputs, y(t) and unknown inputs, e(t)](image)

It should be noted that if the system function or inputs are known, the identification becomes a much easier problem.
2.3.2 Multi Channel Systems

Most system identification techniques deal with single input single output (SISO) systems. For multiple input multiple output (MIMO) systems, basic identification methods usually do not work efficiently. Transfer function based models, such as those used in PEM methods, require one transfer function from each input to each output. This causes a dramatic increase in the number of parameters, as the models are not able to account for parameters that are common to multiple pathways. Thus, conventional methods either perform poorly, or fail. The state space model is most commonly used for MIMO systems because of its simple mathematical equations. Equations 2.1 and 2.2 represent state space models for linear and nonlinear systems respectively. Usage of state space models in subspace identification are explained in the following sections.

2.4 Subspace identification

Most classical system identification methods are transfer function-based, and hence use a SISO structure. Subspace System Identification overcomes this problem as they work with state-space models, and hence it can be applied to MIMO systems. Subspace Identification methods were used successfully in identifying electromechanical modes of oscillation in [69],[53] and [15]. They have several advantages:

1. System identified using subspace is very convenient for estimation, prediction and control [60]. It has numerical simplicity and robustness and are suitable for large MIMO systems such as power systems.

2. Subspace methods estimate modal properties such as frequencies, damping and mode shapes accurately from ambient and ring down data.

In the recent years, subspace identification has gained much attention. Subspace identification can be used to identify reduced order linear models of power systems. In [69] and
[53], ambient raw data are used to identify the power system modes. In [7], subspace identification of electromechanical modes using Numerical algorithm for Subspace State Space System IDentification (N4SID)\(^1\) was shown to be the most effective estimator.

### 2.4.1 State space models

A linear time-invariant (LTI) state space model can be described by the following set of difference equations [60].

\[
x(t + 1) = Ax(t) + Bu(t) + w(t) \\
y(t) = Cx(t) + Du(t) + v(t)
\]  

(2.14)

with

\[
E\left[ \begin{pmatrix} w_r \\ v_r \\ w_s \\ v_s \\ T \end{pmatrix} \begin{pmatrix} w_s \\ S \\ T \\ T \\ R \end{pmatrix} \right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{rs} \geq 0
\]  

(2.15)

where, \( x(t) \in \mathbb{R}^n \) is a state vector; \( y(t) \in \mathbb{R}^l \) is the output vector consisting \( l \) outputs of the system; \( u(t) \in \mathbb{R}^m \) is the input vector consisting of \( m \) inputs; \( w(t) \in \mathbb{R}^n \) is the random disturbance vector; \( v(t) \in \mathbb{R}^l \) is the output random measurement error. \( v(t) \) and \( w(t) \) are unmeasurable signals assumed to be zero mean, white noise vectors.

\( A \in \mathbb{R}^{n \times n} \) is the system state matrix. It is also known as the dynamical system matrix as it can be used to describe the dynamics of a system.

\( B \in \mathbb{R}^{n \times m} \) is the control input matrix. This matrix represents the linear transformation by which the inputs affect the next state.

\( C \in \mathbb{R}^{l \times n} \) is the output matrix. This matrix describes how the internal state is transferred to the outputs.

\( D \in \mathbb{R}^{l \times m} \) is the control output matrix.

\( Q \in \mathbb{R}^{n \times n}, S \in \mathbb{R}^{n \times l} \) and \( R \in \mathbb{R}^{l \times l} \) are the covariance matrices of noise vectors \( v(t) \) and \( w(t) \).

\(^1\)N4SID is pronounced "enforce id"
2.4.2 Stochastic Subspace Identification

Subspace Identification methods are used to identify the parameters of a state space model from input-output data. The model parameters are obtained from the row or column subspace of the system and system related matrices formed from the input-output data. Usually the row space is used to determine the states sequence of the model and the column space is used to determine information about the model.

Some of the benefits of Stochastic Subspace Identification are small computational time, no disturbance is required to extract information from the measured data, and capability of dealing with signals containing noise.

Since the identification process is blind, that means the system has no external inputs or the inputs are unknown, \( u(t) = 0 \). In this case, the subspace identification is treated as a purely stochastic system. From [59], [60] and [51], the stochastic subspace algorithm is described as follows:

Stochastic subspace algorithms are used to compute state space models from output data \( y(t) \) only. So the equations become:

\[
\begin{align*}
x(t + 1) &= Ax(t) + w(t) \\
y(t) &= Cx(t) + v(t)
\end{align*}
\]

with white noise vectors \( w \) and \( v \) with covariance matrix:

\[
E\left[ \begin{pmatrix} w_r \\ v_r \\ w_s T \\ v_s T \end{pmatrix} \right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{rs}
\]

The stochastic state space representation is non-unique and there are infinitely many representations of the model. All of the model representations are equivalent since the second order statistics of the output generated are the same. Two different systems which are related by any similarity transform are equivalent. For example: if \( \tilde{x} = Tx(t) \), where \( T \) is any
invertible transformation. The state space equations become:

\[ \tilde{x}(t + 1) \rightarrow TAT^{-1}\tilde{x}(t) + \tilde{w}(t) \]
\[ y(t) \rightarrow CT^{-1}\tilde{x}(t) + v(t) \]  

(2.18)

where \( \tilde{w}(t) = Tw(t) \). The covariance matrices become:

\[ Q = TQT^T \]
\[ S = TS \]  

(2.19)

Examples of model representations are the Forward model, Forward Innovation model, Backward model, Backward Innovation model. First of all, some of the structural relations of the linear time invariant stochastic process in Equation 2.16 are determined. The state covariance matrix \( \Sigma^s \) is given as follows:

\[ E[x(t)x(t)^T] = \Sigma^s = A\Sigma^s A^T + Q \]  

(2.20)

The output covariance matrices are:

\[ E[y(t)y(t)^T] = \Lambda_0 = C\Sigma^s C^T + R \]
\[ E[x(t + 1)y(t)^T] = G = A\Sigma^s C^T + S \]  

(2.21)

The state space model in Equations 2.16 and 2.17 can be written by applying a Kalman filter to the system:

\[ x(t + 1) = Ax(t) + Ke(t) \]
\[ y(t) = Cx(t) + e(t) \]  

(2.22)

where \( K \) is the Kalman gain. To derive Equation 2.22, Forward innovation form is used because its noise model has fewer degrees of freedom than the other model structures, and it can therefore be identified appropriately from the output data. While using this form, the all the parameters are a part of the model. Therefore there is no need to estimate \( Q, R \) and \( S \) separately.

\[ E(y(t)y^T(t)) = E(Cx(t)x^T(t)C^T) + E(e(t)e^T(t)) \]  

(2.23)
Since, $x(t)$ and $e(t)$ are independent events, $E[e(t)e(t)^T] = (\Lambda_0 - CPC^T)$, where $P$ is forward covariance matrix of the state vector:

$$K = (G - APC^T)(\Lambda_0 - CPC^T)^{-1}$$

(2.24)

$$P = E[x(t)x(t)^T] = APA^T - K(C^TPC + R)K^T + Q$$

(2.25)

The other models can also be used to represent a stochastic system and can be read from the references.

System and System Related Matrices:

Block Hankel matrices can be constructed from input/output data. In this case, there are no inputs to the stochastic system so there will only be output block Hankel matrices. The output data can be arranged into $i$ block rows and $j$ columns.

$$Y_p = \begin{bmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_j \\ \vdots & \vdots & \ddots & \vdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \\ y_i & y_{i+1} & \cdots & y_{i+j-1} \\ y_{i+1} & y_{i+2} & \cdots & y_{i+j} \\ \vdots & \vdots & \ddots & \vdots \\ y_{2i-1} & y_{2i} & \cdots & y_{2i+j-2} \end{bmatrix}$$

(2.26)

Although most data can be found in both matrices, the corresponding columns of $Y_p$ and $Y_f$ have no common data samples and are therefore called past and future respectively.

Matrices $Y_p^+$ and $Y_f^-$ are created from $Y_p$ and $Y_f$ by moving the first block row from $Y_f$ to
the end of $Y_p$.

$$\begin{pmatrix}
y_0 & y_1 & \cdots & y_{j-1} \\
y_1 & y_2 & \cdots & y_j \\
\vdots & \vdots & \ddots & \vdots \\
y_{i-1} & y_i & \cdots & y_{i+j-2} \\
y_i & y_{i+1} & \cdots & y_{i+j-1} \\
y_{i+1} & y_{i+2} & \cdots & y_{i+j} \\
\vdots & \vdots & \ddots & \vdots \\
y_{2i-1} & y_{2i} & \cdots & y_{2i+j-2}
\end{pmatrix}$$

(2.27)

The number of block row $i$ is selected such that $i \geq n$. The number of columns $j$ is approximately equal to given data samples. The dimensions of the block Hankel matrices are: $Y_p \in \mathbb{R}^{li \times j}$, $Y_f \in \mathbb{R}^{li \times j}$, $Y^+_p \in \mathbb{R}^{l(i+1) \times j}$ and $Y^-_f \in \mathbb{R}^{l(i-1) \times j}$.

System related matrices of the subspace identification algorithm are the observability and controllability matrices. These matrices are extensively used in this algorithm.

The extended observability matrix $\Gamma_i$ is defined as:

$$\Gamma_i = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{pmatrix} \in \mathbb{R}^{li \times n}$$

(2.28)

where $i$ is the number of block rows. The matrix is called extended because $i > n$, where $n$ is the system order. It is assumed that the matrix pair $(A, C)$ are observable, therefore the rank of $\Gamma_i$ is equal to $n$. The reversed extended controllability matrix $\Delta_i^s$ is defined as:

$$\Delta_i^s = \begin{pmatrix} A^{i-1}K & A^{i-2}K & \cdots & K \end{pmatrix} \in \mathbb{R}^{n \times li}$$

(2.29)

$(A, K)$ are assumed to be controllable.
Estimating Subspaces:

In the Stochastic identification problem, the row space of the state sequence $\hat{X}_i$ and the column space of the extended observability matrix $\Gamma_i$ are computed from output data. The system matrices are derived from $\hat{X}_i$ and $\Gamma_i$. Assumptions considered for stochastic identification:

1. The process noise $w(t)$ and $v(t)$ are not identically zero.

2. The number of measurements tends to infinity $j \to \infty$.

3. Two weighting matrices $W_1$ and $W_2$ defined by the user are chosen such that $W_1 \in \mathbb{R}^{li \times li}$ is full rank and $W_2 \in \mathbb{R}^{j \times j}$ obeys $\text{rank}(Y_p) = \text{rank}(Y_p W_2)$. In N4SID the weights $W_1$ and $W_2$ are $I_{ij}$ and $I_j$, respectively.

All the steps performed while estimating the subspaces of the stochastic system are listed as follows:

1. $O_i$ is defined as orthogonal projection of the future output space $Y_f$ into the past output space $Y_p$:

\[
O_i = \frac{Y_f}{Y_p}
\]  

(2.30)

2. The singular value decomposition (SVD) of $O_i$:

\[
IO_i J = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 S_1 V_1^T
\]  

(2.31)

3. The order of the system is equal to the number of non-zero singular values.

4. $O_i$ is equal to the product of the extended observability matrix $\Gamma_i$ and the forward Kalman filter state sequence $\hat{X}_i$ = $\begin{pmatrix} \hat{x}_i & \hat{x}_{i+1} & \ldots & \hat{x}_{i+j-1} \end{pmatrix}$:

\[
O_i = \Gamma_i \hat{X}_i
\]  

(2.32)
5. The extended observability matrix $\Gamma_i$ is equal to:

$$\Gamma_i = W_1^{-1}U_1S_1^2$$  \hspace{1cm} (2.33)

6. The extended controllability matrix $\Delta^c_i$ associated with the extended observability matrix $\Gamma_i$ is equal to:

$$\Delta^c_i = \Gamma_i^\dagger \Phi_{[Y_f,Y_p]}$$ \hspace{1cm} (2.34)

where $\Phi_{[Y_f,Y_p]}$ is the covariance between matrices $Y_f$ and $Y_p$.

Computing System Matrices:

From the previous section, the order $n$ of the system, the extended observability matrix $\Gamma_i$, and the state sequences $\hat{X}_i$ are determined.

1. The orthogonal projection $O_{i-1}$ calculated from the output data and is equal to:

$$O_{i-1} = \frac{Y_f - Y_p+}{Y_p} = \Gamma_{i-1} \cdot \hat{X}_{i+1}$$  \hspace{1cm} (2.35)

2. By removing the last $l$ rows of $\Gamma_i$, $\Gamma_{i-1}$ is defined.

$$\Gamma_{i-1} = \Gamma_i(1 : i - 1,:)$$  \hspace{1cm} (2.36)

3. $\hat{X}_{i+1}$ can be calculated as:

$$\hat{X}_{i+1} = \Gamma_{i-1}^\dagger O_{i-1}$$  \hspace{1cm} (2.37)

where $\dagger$ implies the Moore Penrose pseudo inverse.

4. A new set of equations is formed using the state sequences $\hat{X}_i$ and $\hat{X}_{i+1}$:

$$\begin{pmatrix} \hat{X}_{i+1} \\ Y_{ij_i} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} \hat{X}_i + \begin{pmatrix} \rho_w \\ \rho_v \end{pmatrix}$$ \hspace{1cm} (2.38)
where, $Y_{|i}$ is a block Hankel matrix with only one row of outputs and $\rho_w$ and $\rho_v$ are the Kalman filter residues. The system matrices $A$ and $C$ can be solved using least squares as $\rho_w$ and $\rho_v$ are uncorrelated to $\hat{X}_i$.

$$
\begin{pmatrix}
A \\
C
\end{pmatrix} = 
\begin{pmatrix}
\hat{X}_{i+1} \\
Y_{|i}
\end{pmatrix} \hat{X}_i^\dagger
$$

(2.39)

5. The output covariance matrix $G$ defined in (2.21) is determined as the last $l$ columns of $\Delta^c_i$.

Computing the Kalman Gain:

The covariance of the process and measurement noise can be recovered from the residuals $\rho_w$ and $\rho_v$ from (2.38) as:

$$
E[\begin{pmatrix}
\rho_w \\
\rho_v
\end{pmatrix} \cdot \begin{pmatrix}
\rho_w^T & \rho_v^T
\end{pmatrix}] = 
\begin{pmatrix}
Q_i & S_i \\
S_i^T & R_i
\end{pmatrix}
$$

(2.40)

It is assumed that estimated covariances are the non-steady covariances and are denoted using the subscript $i$. The non-steady state state Kalman filter equations are:

$$
P_{i+1} = AP_iA^T + Q_i
$$

$$
G = AP_iC^T + S_i
$$

(2.41)

$$
\Lambda_0 = CP_iC^T + R_i
$$

When $i \to \infty$, $Q_i \to Q$, $S_i \to S$ and $R_i \to R$. The matrices $Q$, $S$ and $R$ are found in this way. $A$ and $C$ were calculated from least squares previously. From these five matrices $G$ and $\Lambda_0$ are extracted by first solving the Lyapunov equation for $\Sigma^S$:

$$
\Sigma^S = A\Sigma^S C^T + Q
$$

(2.42)

$$
G = A\Sigma^S C^T + S
$$

(2.43)

$$
\Lambda_0 = C\Sigma^S C^T + R
$$

Finally, the model is converted to a Forward innovation form by solving the Riccati equation 2.25 and the Kalman gain can be easily computed using Equation 2.24.
2.4.3 Conversion of Discrete-time to Continuous-time Model

This section describes the relationship between continuous and discrete time state space models. State-Space models represent a system with a set of first order differential equations or linear difference using state variables [34]. Continuous State space model representation:

\[ \dot{x}(t) = Ax(t) + Bu(t) \]
\[ y(t) = Cx(t) + Du(t) \]  

(2.44)

Discrete State space model representation:

\[ x(kT + T) = A'x(kT) + B'u(kT) \]
\[ y(kT) = C'x(kT) + D'u(kT) \]

(2.45)

where, \( T \) is the sampling time, \( u(kT) \) and \( y(kT) \) are the input and output at time \( kT \) respectively. Conversion of the continuous time state space matrices to discrete time state space matrices are done in [33]. The relationship between the corresponding state matrices \( A, B, C \) and \( D \) with \( A', B', C' \) and \( D' \) are represented using equations (2.46) to (2.49). All these relationships assume the input is piecewise constant over a short interval.

\[ A' = e^{AT} \]  

(2.46)

\[ B' = \int_0^T e^{A\tau} B d\tau \]  

(2.47)

Matrices \( C \) and \( D \) are constant and does not depend on the sampling time \( T \), therefore:

\[ C' = C \]  

(2.48)

\[ D' = D \]  

(2.49)

State space representation of the transfer function \( G \), that takes the input \( u \) to \( y \), are expressed using the following equations:

\[ F(s) = C(sI - A)^{-1}B + D \]  

(2.50)
\[ F(z) = C'(zI - A')^{-1}B' + D' \]  

(2.51)

From both (2.50) and (2.51), it can be deduced that the poles of discrete and continuous time transfer functions are the eigenvalues of state matrix \((A \text{ or } A')\). In both cases, eigenanalysis of the state matrix that is \(\text{det}|\lambda I - A| = 0\) provides the poles. From equations (2.46) to (2.49) it is observed that discretization of continuous time poles can be done by \(z = e^{sT}\).

### 2.5 Mode Estimation

Electromechanical modes of oscillations occur due to natural oscillations of interconnected power systems and therefore cannot be eliminated. As power systems evolve, the frequency and damping of the existing modes change and new modes are created. The modal damping and frequency can be estimated and can be used to avoid oscillations from growing out of control. [20], [41].

Electric power systems have experienced problems with the following types of low frequency oscillations and are classified as follows:

1. Local mode oscillations: With local modes, one generator swings against the whole system. They have a frequency ranging from 1-2 Hz. This type of oscillation is usually located between the generators and the lines connecting them to the grid. The frequency and damping depend on the machine output and the impedance between the machine and the bus voltages. An example of a local mode is illustrated in Figure 2.5.

2. Inter-area mode oscillations: It is very important to understand the nature and characteristics of inter-area oscillations. Inter-area oscillations refer to the swinging between a group of machines in one area against a group of machines in a different area. Usually the areas are connected via transmission
lines. These oscillations may result from small disturbances such as load variations or may occur as an aftermath of large disturbances such as faults or line tripping. These oscillations take place at very low frequencies from below 0.1 Hz to 1 Hz. The effects of these low frequency inter-area oscillations are not noticeable instantaneously but over a period of time they can grow in amplitude and cause system isolations.

Mode properties of these inter-area modes depends on the power system configuration, different types of generator excitation systems and load characteristics. That is they involve many parts of the system with highly nonlinear behavior. The natural frequency and damping of the inter-area modes depend on the strength of the tie-lines, the nature of the loads, power flow through the interconnections and the interaction of loads with the dynamics of generators. Characteristics of inter-area oscillations are analyzed and discussed in [62] and [63]. Figure 2.6 shows a typical example of an inter-area mode. The figure shows the power variations of a tie line and the oscillation frequency is approximately 0.3 Hz.

Figure 2.5: An example of a local mode
2.5.1 Power System Model

Power systems have many modes of oscillation. Most of these are due to generators. Power systems have multiple machines interconnected and therefore exhibit many modes of oscillation. The system dynamics are described by a set of n first order differential-algebraic equations [20]:

\[ \dot{x} = f(x_d, x_a, u) \] \hfill (2.52)

\[ 0 = g(x_d, x_a, u) \] \hfill (2.53)

\[ y = h(x_d, x_a, u) \] \hfill (2.54)

where, \( u \), \( y \) and \( x \) are the input, output and state vectors respectively. The state vector \( x^T = [x_d^T \ x_a^T] \), where \( x_d \) and \( x_a \) are the dynamic and algebraic state variables. Dynamic state variables provide a description of the system’s behavior. These variables are associated with the differential equations describing the system. Algebraic variables are physical quantities such as speed, angle, etc.

Equilibrium or singular points are defined as those points where all the derivatives of the
states \( \dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n \) are simultaneously zero. That is the system is at rest since all the state variables are constant and do not vary with time. In order to investigate small signal performance, linearization of Equation 2.52 is done. There is a stationary equilibrium point \( y_0 = h(x_{d0}, x_{a0}, u_0) \). The system is perturbed: If there are small deviations or perturbations from the stationary model can be approximated as:

\[
E' \frac{d}{dt} \begin{bmatrix} \Delta x_d \\ \Delta x_a \end{bmatrix} = A' \begin{bmatrix} \Delta x_d \\ \Delta x_a \end{bmatrix} + B' \Delta u \tag{2.55}
\]

\[
\Delta y = C' \begin{bmatrix} \Delta x_d \\ \Delta x_a \end{bmatrix} + D' \Delta u \tag{2.56}
\]

where, \( A', B', C' \) and \( D' \) are partial differentiation matrices.

\[
A' = \begin{bmatrix} \frac{\delta f}{\delta x_d} & \frac{\delta f}{\delta x_u} \\ \frac{\delta g}{\delta x_d} & \frac{\delta f}{\delta x_d} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{2.57}
\]

\[
B' = \begin{bmatrix} \frac{\delta f}{\delta u} \\ \frac{\delta g}{\delta u} \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \tag{2.58}
\]

\[
C' = \begin{bmatrix} \frac{\delta h}{\delta x_d} & \frac{\delta h}{\delta x_a} \\ \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{2.59}
\]

\[
D' = \begin{bmatrix} \frac{\delta h}{\delta u} \\ \end{bmatrix} = \begin{bmatrix} D_1 \end{bmatrix} \tag{2.60}
\]

\[
E' = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \tag{2.61}
\]

By eliminating the algebraic variables \( x_a \):

\[
A = A_{11} - A_{12}A_{22}^{-1}A_{21} 
\]

\[
B = B_1 - A_{12}A_{22}^{-1}B_2 
\]

\[
C = C_1 - C_2A_{22}^{-1}A_{21} 
\]

\[
D = D_1 - C_2A_{22}^{-1}B_2 
\]
This yields the following linear model.

\[
\dot{x}(t) = Ax(t) + Bu(t) \\
y(t) = Cx(t) + Du(t)
\]  

(2.63)

where, \(x(t) \in \mathbb{R}^n\) is a state vector; \(y(t) \in \mathbb{R}^l\) is the output vector; \(u(t) \in \mathbb{R}^m\) is the input vector; \(A \in \mathbb{R}^{n \times n}\) is the system state matrix; \(B \in \mathbb{R}^{n \times m}\) is the control input matrix; \(C \in \mathbb{R}^{l \times n}\) is the output matrix; \(D \in \mathbb{R}^{l \times m}\) is the control output matrix. Electromechanical mode estimation methods consider the fact that the nonlinearities of a power system is linearized about an operating point. The linearized state space model can be described as state space model equations as shown in Equation 2.63

2.5.2 Small Signal Analysis

The ability of a power system to maintain synchronism after being subjected to small disturbances is known as small signal analysis. An example of a small disturbance is random load variations in a power system. If power system oscillations caused by small disturbances can be suppressed, such that the deviations of system state variables remain small for a long time, the power system is stable. Small signal analysis provides an important insight into the electromechanical dynamics of the network. Eigenanalysis of the aforementioned linearized model of power system that is (2.63) is used to determine the small-signal dynamic behaviour of the system. The eigenvalues, frequency of oscillation, damping ratios and eigenvectors (mode shapes) can be determined by the eigenanalysis of the state matrix \(A\).

Eigenvalues:

The eigenvalues of a matrix are provided by the values of the scalar parameter \(\lambda\)

\[
Av = \lambda v
\]

(2.64)

where, \(A\) is a real asymmetric \(n \times n\) system state matrix and \(v\) is an \(n \times 1\) vector. Equation 2.64 has a nonsingular solution (i.e., \(v \neq 0\)), \(\lambda\) is the eigenvalue of the matrix \(A\). To calculate
the eigenvalues, Equation 2.64 can be written as:

\[(A - \lambda I)v = 0\]  \hspace{1cm} (2.65)

A necessary condition for existence of a nonsingular solution is:

\[det(A - \lambda I) = 0\]  \hspace{1cm} (2.66)

Expansion of the determinant in the above equation 2.66 gives the following polynomial equation:

\[\alpha_0 + \alpha_1 \lambda + ... + \alpha_{n-1}\lambda^{n-1} + (-1)^n\lambda^n = 0\]  \hspace{1cm} (2.67)

This is the characteristic equation, and the polynomial on the left side of the equation is the characteristic polynomial. The coefficient of \(\lambda^n\) is nonzero, therefore there are \(n\) number of roots. The \(n\) roots of \(\lambda = \lambda_1, \lambda_2, ..., \lambda_n\) are eigenvalues of \(A\).

The eigenvalues can be real or complex. Complex eigenvalues always appear in conjugate pairs. Moreover, similar matrices have the same eigenvalues and transposition of a matrix does not change its eigenvalues.

Eigenvectors:

For any eigenvalue \(\lambda_i\), the \(n \times 1\) vector \(v_i\) satisfying the Equation 2.64 is called the right eigenvector of \(A\) corresponding to eigenvalue \(\lambda_i\). The eigenvector:

\[v_i = \begin{bmatrix} \v_i_1 \\ \v_i_2 \\ \vdots \\ \v_i_n \end{bmatrix}\]  \hspace{1cm} (2.68)

Since Equation 2.64 is a homogenous equation, \(cv_i\) (\(c\) is any scalar) is also the solution to the equation. This eigenvector defines a one-dimensional subspace which is invariant under the operation of left multiplication of \(A\).

Similarly an \(n \times 1\) vector \(u_i\) satisfying the equation below:

\[u_iA = \lambda_i u_i\]  \hspace{1cm} (2.69)
is known as the left eigenvector of $A$ corresponding to eigenvalue $\lambda_i$. The left and right eigenvectors corresponding to different eigenvalues are orthogonal. $u_jv_i = 0$ where $\lambda_i \neq \lambda_j$.

In case of eigenvectors corresponding to the same eigenvalue: $v_iu_i = C_i$ where $C_i$ is any constant.

Modal Matrices:

The eigenproperties of $A$ can be expressed as the following matrices:

$$V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \quad (2.70)$$

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^T \quad (2.71)$$

$$\Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \quad (2.72)$$

$V$ is a matrix containing all the right eigenvectors arranged in columns, $U$ is a matrix containing all the left eigenvectors in rows and $\Lambda$ is a diagonal matrix consisting of all the eigenvalues of matrix $A$. These matrices are known as the modal matrices.

It was shown above that eigenvectors corresponding to the same eigenvalue result in a constant. These vectors can be normalized so that $u_iv_i = 1$. Using Equation 2.64 and this notation:

$$AV = VA \quad (2.73)$$

Since $UV = I$ it can be said that $U = V^{-1}$, therefore yielding:

$$V^{-1}AV = \Lambda \quad (2.74)$$

From the state equation, it can be seen that the rate of change of a state variable is a linear combination of all the state variables. Due to this coupling among the state variables it is difficult to see the system movement. To cancel the effect of coupling among the state
variables a new state variable, $z$ is introduced. The relationship between the original state
variable vector $\Delta x$ is:

$$\Delta x = Vz$$ \hfill (2.75)$$

The new state equation can be written as:

$$\frac{dx}{dt} = \Lambda z$$ \hfill (2.76)$$

In the above equation $\Lambda$ is a diagonal matrix and the equation can be simplified as:

$$\frac{dz_i}{dt} = \lambda_i z_i.$$ \hfill (2.77)$$

where $i = 1, 2, ..., n$. Equation 2.77 can be expressed in time domain as:

$$z_i(t) = z_i(0)e^{\lambda_i t}$$ \hfill (2.78)$$

where the initial values of $z_i$, $z_i(0)$ can be expressed from Equation 2.75 as:

$$z_i(0) = v_i^T \Delta x(0)$$ \hfill (2.79)$$

The original state vector in the time domain can be expressed by using Equations 2.78 and
2.79 in the following equation:

$$\Delta x = \sum_{i=1}^{n} v_i z_i(0) e^{\lambda_i t}$$ \hfill (2.80)$$

In other words the $i$th state variable in the time domain is:

$$\Delta x_i(t) = v_{i1} z_1(0)e^{\lambda_1 t} + v_{i2} z_2(0)e^{\lambda_2 t} + ... + v_{in} z_n(0)e^{\lambda_n t}$$ \hfill (2.81)$$

Equation 2.81 gives the times response of the system in terms of the eigenvalues and the left
and right eigenvectors. Therefore the free response is given by a linear combination of the $n$
dynamic modes of the system. The modes correspond to the eigenvalues of the system state
matrix, so the system stability can be determined by the eigenvalues:
1. A real eigenvalue corresponds to a non-oscillatory mode. A negative real eigenvalue is a decaying mode, where its magnitude defines how fast it is decaying, that is: the larger the magnitude, faster the decay. A positive real value states that the mode is in aperiodic instability.

2. Complex eigenvalues always occur in conjugate pairs, where each pair represents an oscillatory mode.

\[ \lambda_i = \sigma_i + j\omega_i \quad (2.82) \]

Eigenvectors and \( z(0) \) are complex valued for complex eigenvalues. Hence:

\[ (a + jb)e^{(\sigma - j\omega)t} + (a - jb)e^{(\sigma + j\omega)t} = e^{\sigma t}(2a \cos \omega t + 2b \sin \omega t) \quad (2.83) \]

which exhibits \( e^{\sigma t}\sin(\omega t + \theta) \). The real component gives the damping of the oscillation. A negative real part represents a damped oscillation whereas a positive real part represents an oscillation with increasing amplitude. The damping ratio \( \zeta_i \) for mode \( i \) is defined as:

\[ \zeta_i = -\frac{\sigma_i}{\sqrt{\sigma_i^2 + \omega_i^2}} \times 100 \quad (2.84) \]

The imaginary component provides the frequency of oscillation. Oscillation frequency \( f_i \) (in Hz) is:

\[ f_i = \frac{\omega_i}{2\pi} \quad (2.85) \]

Mode shape:

From the above discussion, it can be deduced that the relationship between the system time response and vectors \( \Delta x \) and \( z \) are:

\[ \Delta x = Vz(t) = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} z(t) \quad (2.86) \]

\[ z(t) = U\Delta x(t) = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^T \Delta x(t) \quad (2.87) \]
The variables $\Delta x_1, \Delta x_2, ..., \Delta x_n$ are the original state variables and variables $z_1, z_2, ..., z_n$ are the transformed state variables which represent the modes the system and each variable corresponds to one mode. That is the variables $z$ are directly related to modes. From Equation 2.86, the right eigenvectors determine the form of exhibition of each mode. Thus is when a specific mode is excited, the relative activity of each state variable is described by the right eigenvector which is known as the mode shape. For example: the degree of activity of the state variable $x_k$ in the $i$th mode is given by the element $v_{ki}$ of the right eigenvector $v_i$. The magnitudes of the element $v_i$ provides the level of activity of each of the $n$ state variables resulting from the $i$th mode and the angles provide the phase shift of the state variables with regard to the mode.

From Equation 2.87, the left eigenvector $u_i^T$ represents the way the original state variables combine to effect the $i$th mode. Therefore the $k$th element in the right eigenvector $v_i$ measures the level of activity of the state variable $x_k$ in the $i$th mode; while the $k$th element of the left eigenvector $u_i^T$ is the contribution of the activity to the $i$th mode.

It is discussed above that the eigenvalues and eigenvectors of a linearized state space model are the modes and mode shapes, respectively. Damping provides direct evaluation of the mode stability. The mode shape (amplitude and phase) helps in identifying the most powerful/energetic oscillations.

Electromechanical modes cannot be measured directly due to random load variation of the power system so they are estimated from eigenanalysis of the (estimated) state matrix $A$. Estimation of electromechanical modes means measurement of the frequency, damping ratio, amplitude and phasing of natural oscillations. The mode is usually referred as a complex value, whose real and imaginary parts contribute for the frequency and damping of the mode. Eigenanalysis of the state matrix $A$ provides the eigenvalues $\lambda_i$.

The modal frequency $f_i$ and damping ratio $\zeta_i$ of the particular mode $\lambda_i$ can be determined by the real and imaginary parts of the eigenvalue.
The mode shape determines the amplitude and phasing of the mode. The mode shape $v_i$ is calculated from the right eigenvector.

The mode shape of measured output $y_m$ relative to $y_n$ is given by the ratio of $m$th and $n$th elements of $v_i^{(y)}$.

$$\text{Modeshape} = \frac{v_{m,i}}{v_{n,i}}$$  \hspace{1cm} (2.88)

In the paper [7], the authors provide a comparison of five existing algorithms for estimating the electromechanical modes of power systems. The five different methods used are: Transfer function method, Spectral Method, Frequency Domain Decomposition (FDD) Method, Channel Matching Method and subspace identification Method using Numerical algorithm for Subspace State Space System IDentification (N4SID). The algorithms are compared using simulated data from a 17 machine model of Western Electricity Coordinating Council (WECC) under ambient conditions for both high and low damping, as well as cases of ambient conditions where there are ringdown disruptions. The performance of these methods are compared using Monte Carlo simulations with WECC measured data results. A discussion about the practical issues faced during implementation is also provided. It also concludes with recommendations of using appropriate methods during different conditions are provided.

A brief explanation of the methods used was given in Chapter 1. A detailed explanation of subspace identification using N4SID algorithm was described in section 2.4.

In order to compare the five different methods a 17-machine simulation model was used. The mode shape with a high and low damping were estimated. Voltage angles from the buses of the 17-machine model were simulated during ambient conditions (for 10 minutes) and reported at 30 samples per second to resemble the output of a typical PMU.

For the statistical analysis and comparison, Monte Carlo simulations were used. Bias
and standard deviation were estimated. Bias is equal to the difference of the true value of
the mode shape and sample mean of the Monte Carlo trial results and standard deviation is
estimated by the sample standard deviation of the Monte Carlo trial results.

Simulation studies conducted for estimating low damping mode shapes had very clear
and good results. All the algorithms had low bias and standard deviation in the Monte Carlo
simulations, the results were thus easy to deduce. All of the five methods were applicable
for poor damping. It was acknowledged that none of the estimators could perform properly
for estimating mode shapes for high damping. The conclusions were drawn based on biases.
For example N4SID had the lowest bias but the highest standard deviation. The authors
suggested that bias has a larger contribution towards overall error based on their previous
results.

A ring-down type data analysis was done. The results seemed very appropriate as Spectral
Method , FDD and Channel matching method all rely on the fact that the $ith$ eigenvalue $\lambda_i$
has a negligible real part $\sigma_i$. These methods fail to estimate ring-down data modes. whereas
N4SID algorithm gives a proper result for both high and low damping. Transfer Function
method provides good results for low damping but does not fail to estimate mode shapes.

Different advantages of N4SID were illustrated. The N4SID algorithm estimates a state-
space model that gives both modes and mode shapes, so it requires no separate algorithm
unlike the TF methods. For the simulation studies in this paper, since the true modal
values are known, they were used where necessary with the TF Methods. There are a lot
of assumptions concerned with the transfer function based methods used for power system
modal estimation. N4SID can directly incorporate measured system inputs so it could easily
estimate mode shapes for probed data. The results in the tables support that N4SID is
clearly the best to estimate modes for both ambient and ring-down data compared to the
TF methods.
2.6 Summary

In this chapter, a brief summary of necessary background of system identification were discussed. Linear system identification method, subspace identification was introduced stating that stochastic subspace identification method was used as a blind MIMO system identification method. Different algorithms can be used for subspace identification and emphasis on N4SID was given.

Power systems can be linearized into a set of state space equations. A brief summary of this linearization was provided. The power system modal properties were obtained by eigenanalysis of the the state matrix of these set of equations. The power system model can be estimated by subspace identification.
Chapter 3
NONLINEAR IDENTIFICATION

3.1 Introduction

Identifying the electromechanical modes of a power system is very important as explained earlier. Power systems always experience small disturbances due to increased interconnectivity, increased demands and random load changes. If these modes persist unidentified within the system, they can grow out of control in the power system, resulting in stability and reliability issues and in severe cases, blackouts and system separation. Since the electromechanical modes cannot be measured directly, they must be identified or estimated.

Some nonlinearities are always present in power system data as power systems are highly nonlinear. They consist of an interconnection of nonlinear elements such as generators, motors, loads etc. With any nonlinearity present in the data, any attempt to fit a linear model will be unsuccessful. In this thesis, ambient data was used and shown to have nonlinearities present in it. Ambient data measured from power systems are known to have noisy non-stationary random fluctuations resulting from small magnitude excitations. Usually, modes are identified using linear models from ambient data but the accuracy of the results maybe compromised if significant nonlinearities are present in the measured data. Using nonlinear models to identify modes from nonlinear data is more convenient.

An example of a nonlinear model is the Wiener model. This chapter provides an introduction to Wiener models. To deal with nonlinearities within the power system data and identifying it’s electromechanical modes, a Wiener model can be used. A Wiener model with a multiple input nonlinearity is an universal model for nonlinear dynamic systems [3]. The nonlinearities of a power system can be anywhere, but by using a Wiener model it can be represented by a linear dynamic system with all the nonlinearities at the output.
Usually system identification methods are applied to systems with both known inputs and outputs. However, in many cases, the input is not known or cannot be measured. In these cases only the output is known. These kinds of identification are known as blind identification. The main objective of this chapter is to detect the nonlinearities in power system data and perform blind identification of electromechanical modes of a simulated power system using a Wiener Model.

This chapter is outlined as follows: First of all different kinds of nonlinearities present in power systems are discussed. The effect of these nonlinearities on power system mode estimation is recognized. This leads to the conclusion of using nonlinear analysis of nonlinear power system data. In order to find out the presence of nonlinearities, linearity tests must be carried out. Description of the linearity tests used in this thesis is provided. At the end, nonlinear systems and introduction to Wiener systems is given. The usage of Wiener systems to provide nonlinear estimates is described.

3.2 Power System Nonlinearity

Power systems are huge and complex dynamic systems consisting of different nonlinear elements including different types of generator sets, electric loads, substations and different transmission lines. Details of the nonlinear element models can be found in [44] and [20]. While estimating modes from a power system, a linear model is derived to represent the dynamics of a power system. Chapter 2 uses eigenanalysis of a power system model that had been linearized about a steady operating point to estimate the power system modes. Electromechanical oscillations can easily and accurately be characterized using this linear model. But the system is actually nonlinear, so a fundamental understanding of how the nonlinearities can alter the actual system oscillations is required. In certain rare cases nonlinearities can cause undamped oscillations [41]. In this section the different nonlinearities present in a power system is stated and how they can influence the oscillations are pointed
Generator models contribute several nonlinearities to a power system. Three of the most prevalent nonlinearities are represented by the swing equation, excitation limiting and magnetic saturation.

Nonlinearities in the swing equation are introduced by the electrical torque calculation. Under normal operating conditions, the relative position of the rotor axis and the resultant magnetic field axis is fixed. During any disturbance, the rotor decelerates or accelerates with respect to the synchronously rotating air gap. This results in a relative motion. Thus the equation describing this relative motion is known as the swing equation. That is, these rotational inertia equations describe the balance between the electromagnetic torque and the mechanical torque of a machine. These equations are not linear function for all synchronous machine models (in the classical model of a synchronous machine it is represented by trigonometric equations [20]). In [55], the effects of swing equations on power system oscillations were simulated. This was carried out to determine the effect of generator nonlinearities on oscillations. The results showed that the linear eigenanalysis estimations vary very little with swing equation nonlinearities.

Excitation limiting occurs in the first swings of severe power system disturbances. In different excitation system models used to model power systems, the only nonlinearity introduced is the exciter ceiling of the exciter output voltage, that is the maximum and minimum limits. For modal analysis, these limits are not considered as the linearized model operates within an operating point such that the exciter limits are not exceeded. The exciter limiting can cause changes in the modal estimation results when considering a linearized model instead of nonlinear analysis [55]. The major difference is noticed in the damping. For generators with strong excitation limiting results in less damping in their modes. Usage of Power System Stabilizers (PSS) significantly reduces these exciter performance problems [20].

In a synchronous generator, if the field current increases, the flux linkage also increases but
with a decreasing slope. This is due to magnetic saturation. The method of reprising this effect is by using mutual inductances [20]. These modelling equation of the magnetic saturation are nonlinear. The magnetic saturation modelling adds damping to nonlinear analysis when compared to linear analysis [55].

Power system loads also contribute to huge amount of nonlinearities to the system. In order to show the effects of nonlinearity in load models. Load models can be classified into different kinds: constant impedance, constant current, constant power and loads depending on frequencies [55]. Introducing nonlinear load models slightly alters the modal estimations. These are the few nonlinearities in power system that can change the estimation results if not taken into account. Linear models become less reliable for system analysis which tend to have nonlinearities. Any attempt to fit a linear model to nonlinear data will yield unreliable and inaccurate results. Thus the need of nonlinear analysis to estimate power system modes form nonlinear power system data is necessary.

### 3.3 Nonlinearity Detection

It is important to detect the nonlinearities in the power system data in order to find out which kind of model should be used. In order to detect possible nonlinearities in the data, Gaussianity and Linearity tests are carried out. In this section, Gaussianity and Linearity tests from [12], [40] are derived. The simulation results from [2] show that these tests are powerful and successful in detecting nonlinearity in different kinds of time-series. Hinich’s tests [12] are based on bispectra (third order spectra). The power system data obtained is in time series. The assumptions made in time series analysis are

- the process is stationary
- the process is described as a linear model
A linear process $z(t)$ can be expressed as:

$$z(t) = \sum_{s=0}^{\infty} h(s)\epsilon(t - s) \quad (3.1)$$

where, $h(s)$ is an impulse response and $\epsilon(t)$ is considered to be independent identically distributed (i.i.d) with:

$$E(\epsilon(t)) = 0$$

$$E(\epsilon^2(t)) = \sigma^2_e$$

$$E(\epsilon^3(t)) = \mu_3$$

The third order cummulants $c_{zzz}(a, b) = E[z(t)z(t+t_1)z(t+t_2)] \neq 0$ for most values of $t_1$ and $t_2$, if $\epsilon(t)$ is not normal and $E[\epsilon^3(t)] \neq 0$. In this method, an estimator of the bispectrum, the Fourier transform of $[c_{zzz}(t_1, t_2)]$ is utilized. The bispectrum $S(\omega_1, \omega_2)$ is used to construct a statistic to test whether the bispectrum of $z(t)$ is non-zero. The null hypothesis being: if the bispectrum is zero, then $z(t)$ is Gaussian. If the relationship between $z(t)$ and $\epsilon(t)$ is nonlinear then $z(t)$ will be non Gaussian even if $\epsilon(t)$ is normal. In order to test whether $z(t)$ is linear another test statistic considering constant bicoherence is determined. If the bicoherence is constant for a certain range, then $z(t)$ will be linear. An overview of these tests is as follows:

The auto covariance function $R(v)$ of $z(t)$ is:

$$R(v) = E[z(t)z(t+v)]$$

$$= E[(\sum_{s=0}^{\infty} h(s)\epsilon(t-s))(\sum_{s'=0}^{\infty} h(s')\epsilon(t+v-s')))$$

$$= \sum_{s} \sum_{s'} h(s)h(s')E[\epsilon(t-s)\epsilon(t+v-s')] \quad (3.2)$$

But, $\epsilon(t)$ is independent so:

$$R(v) = \sigma^2_e \left[ \sum_{s=-\infty}^{\infty} h(s)h(s+v) \right] \quad (3.3)$$
Third order central moment of \( z(t) \) is:

\[
C(t_1, t_2) = E[z(t)z(t + t_1)z(t + t_2)]
\]

\[
= \sum_{s_1} \sum_{s_2} \sum_{s_3} h(s_1)h(s_2)h(s_3)E[\epsilon(t - s_1)\epsilon(t + t_1 - s_2)\epsilon(t + t_2 - s_3)]
\]

\[
= \mu_3 \left[ \sum_s h(s)h(s + t_1)h(s + t_2) \right]
\]

(3.4)

Let \( H(\omega) = \sum_{t=0}^{\infty} h(t) \exp(-j\omega t) \) be the filter transfer function, then the bispectral density function of \( z(t) \) is given by:

\[
S(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{t_1 = -\infty}^{\infty} \sum_{t_2 = -\infty}^{\infty} C(t_1, t_2)e^{-j\omega_1 t_1 - j\omega_2 t_2}
\]

\[
= \frac{\mu_3}{(2\pi)^2} H(\omega_1)H(\omega_2)H^*(\omega_1 + \omega_2)
\]

(3.5)

\( H^*(\omega) \) is the complex conjugate of \( H(\omega) \).

The equation 3.5 shows that if \( \mu_3 = 0 \), then the bispectral density function \( S(\omega_1, \omega_2) = 0 \) for all values of \( \omega_1 \) and \( \omega_2 \). Now if \( \epsilon(t) \) is Gaussian, then \( \mu_3 = 0 \) and \( S(\omega_1, \omega_2) = 0 \). Under suitable conditions the normality of \( \epsilon(t) \) implies the normality of \( z(t) \). That is: if \( z(t) \) is Gaussian then the bispectrum \( S(\omega_1, \omega_2) = 0 \) for all values of \( \omega_1 \) and \( \omega_2 \). An estimator of the bispectrum is constructed using the Fast Fourier transform (FFT) of \( z(t) \)

\[
F(j, k) = \frac{Z(\frac{j}{N})Z(\frac{k}{N})Z^*(\frac{j+k}{N})}{N}
\]

(3.6)

where \( Z(\frac{j}{N}) = \sum_{t=0}^{N-1} z(t) \exp\left(\frac{-j\omega t}{N}\right) \) is the fast Fourier transform of \( z(t) \) and \( N \) is the number of samples. An estimator of the bispectrum is obtained by averaging \( F(j, k) \) over \( M^2 \) adjacent frequency pairs as

\[
\hat{S}(\omega_1, \omega_2) = \frac{\omega_1 M - 1}{j = (\omega_1 - 1)M} \frac{\omega_2 M - 1}{k = (\omega_2 - 1)M} \frac{F(j, k)}{M^2}
\]

(3.7)

where \( M \) is an integer \( M = N^c \) for \( \frac{1}{2} \leq c \leq 1 \). The average skewness \( \bar{\gamma} \) is

\[
\bar{\gamma} = \int_{\Omega} \int_{\Omega} [S_z(\omega_1)S_z(\omega_2)S_z(\omega_1 + \omega_2)]^{-1} |\hat{S}_z(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2
\]

(3.8)
\[ S_z = \sigma^2 |H(\omega)|^2 \] is the spectrum of \( z(t) \). The probability \( P \) of false alarm is calculated as follows:

\[ P = 1 - \Phi(-\bar{\gamma}N/2^{1/2}) \quad (3.9) \]

\( \Phi \) is cumulative distribution function of a normal variate. Using this probability \( P \), it can be found out whether the data is Gaussian. If the probability is small then the assumption of zero bispectrum can be rejected and the data is non-Gaussian.

The squared bicoherence is defined as

\[ \Psi^2(\omega_1, \omega_2) = \frac{|S_z(\omega_1, \omega_2)|}{S_z(\omega_1)S_z(\omega_2)S_z(\omega_1 + \omega_2)} = \frac{\mu_3^2}{\sigma_6^6} \quad (3.10) \]

The squared bicoherence is equal to the square of the skewness of \( z(t) \). An estimator of the squared bicoherence is:

\[ \hat{\Psi}^2_{a,b} = \frac{|\hat{S}_Z(a,b)|^2}{(\frac{N}{M^2})\hat{S}_Z(\frac{(2a-1)M}{2N})\hat{S}_Z(\frac{(2b-1)M}{2N})\hat{S}_Z(\frac{(2(a+b)-1)M}{2N})} \quad (3.11) \]

For the data to be linear, the squared bicoherence is constant for all \( \omega_1 \) and \( \omega_2 \). In practice, it is not usually flat: a constant value is estimated by computing the mean \( \beta \) of the bicoherence over a non redundant region. If \( z(t) \) is Gaussian then the spectrum, \( S_z \) is chi-squared distributed with two degrees of freedom, therefore the squared bicoherence is chi-squared distributed with two degrees of freedom and centrality parameter of \( \beta \). The sample interquartile range \( R \) can be estimated. It is compared with the theoretical interquartile range of the chi-squared distribution with two degrees of freedom and centrality parameter \( \beta \). If the estimated interquartile range \( R \) is close to the value of the theoretical value, the linearity hypothesis of constant bicoherence is accepted and the data is linear.
3.4 Nonlinear Systems

Using a nonlinear model to describe the relationship between inputs and outputs can have possibilities of describing a system more accurately. Constructing nonlinear models from measured input-output data is an important system identification problem. There are several different kinds of model structures for nonlinear systems. For example: Volterra series, Wiener models, nonlinear state space model shown in Equation 2.2. Other approaches include more general block structured models, where the system is represented by combination of alternating linear filters and memoryless nonlinearities.

A brief description of a simple block structured model, the Wiener model is discussed. The discussion will require a basic introduction to Volterra theory. Volterra series forms the base for Wiener theory. A description of how the Wiener model can be used to estimate system nonlinearity is also provided.

3.4.1 Volterra and Wiener Series

The Volterra series represents the output of a system as a summation of terms of different nonlinear orders. It is a generalization of the convolution form as shown in [65] and [3]:

\[ y(t) = \sum_{n=1}^{\infty} \int_0^\infty \cdots \int_0^\infty h^{(n)}(\tau_1, \ldots, \tau_n)x(t - \tau_1)\ldots x(t - \tau_n)d\tau_1\ldots d\tau_n \]  

(3.12)

where \( n \) is the order of the nonlinear term, and \( t \) is the continuous time variable. For \( n = 0 \) that is zero order term a constant output is created: \( y^{(0)}(t) = h^{(0)} \). For the first-order, a linear term is generated: \( y^{(1)} = \int_0^T h^{(1)}(\tau)x(t - \tau)d\tau \). The second-order generates a bilinear term described by nonlinear interactions between two different inputs: \( y^{(2)} = \int_0^\infty \int_0^\infty h^{(2)}(\tau_1, \tau_2)x(t - \tau_1)x(t - \tau_2)d\tau_1d\tau_2 \). The response of a system contains contributions from all the given kernels in a system. Similarly, the \( n \)th order Volterra series kernel is described by nonlinear interactions between the input at \( n \) different times.

A more efficient way to represent the system in discrete time and replace the integrals with
A finite, discrete Volterra model is given as:

\[ y(t) = \sum_{n=0}^{Q} \sum_{\tau_1=0}^{T-1} \ldots \sum_{\tau_n=0}^{T-1} h^{(n)}(\tau_1, \ldots, \tau_n)x(t - \tau_1) \ldots x(t - \tau_n) \]  

(3.13)

where, \( Q \) is the maximum kernel order and \( T \) is the kernel memory length, and are both finite. According to [3], Volterra series can be used as an universal approximator for systems. Let \( K \) be any uniformly bounded equicontinuous set in \( C(\mathbb{R}) \). \( N \) is any time invariant fading memory operator defined on \( K \). Then, for any \( \epsilon > 0 \) and for all \( x \in K \), there is a Volterra operator \( \hat{N} \) such that:

\[ |N(x(t)) - \hat{N}(x(t))| \leq \epsilon \]  

(3.14)

In [3], the extended Equation 3.14 shows that finite dimensional Volterra series can be used to approximate the output of any fading memory system. The problem with Volterra series is that, with increasing order the kernel terms result in high order convolutions which is difficult to compute. Again the outputs of the Volterra terms are not orthogonal and that results in difficulty during identification of systems.

Volterra series were to be orthogonalized, its coefficients could then be estimated using projections [26]. The Wiener series is an orthogonal expansion of the Volterra series whose terms are orthogonal only when the input is a white Gaussian noise signal [65]. Complete derivation of the Wiener series can be found in [26] and [45]. The Wiener series model is:

\[ y(t) = \sum_{n=0}^{\infty} G_n[k^{(n)}(\tau_1, \ldots, \tau_n); x(t'), t' \leq t] \]  

(3.15)

This equation applies to a set of kernels, the Wiener kernels \( k^{(n)}, n = 0, \ldots, \infty \) to the input history \( x(t'), t' \leq t \), to generate the output, \( y(t) \).

The output of the zero-order operator is:

\[ G_0[k^{(0)}; x(t)] = k^{(0)} \]  

(3.16)
This output is a constant and is independent of the input.

The output of the first-order operator is given by:

\[ G_1[k^{(1)}(\tau); x(t)] = \sum_{\tau=0}^{T-1} k^{(1)}(\tau)x(t - \tau) \] (3.17)

Equation 3.17 is the same as first-order Volterra kernel. Therefore if the input is zero mean Gaussian, then the first-order output will also be zero-mean Gaussian. It is orthogonal to any constant that is the output of the zero-order operator. The output of the second order Wiener operator is given by:

\[ G_2[k^{(2)}(\tau_1, \tau_2); x(t)] = \sum_{\tau_1=0}^{T-1} \sum_{\tau_2=0}^{T-1} k^{(2)}(\tau_1, \tau_2)x(t - \tau_1)x(t - \tau_2) - \sigma^2 \sum_{\tau=0}^{T-1} k^{(2)}(\tau, \tau) \] (3.18)

The expected value of the output of \( G_2 \) is 0 since \( E[x(t - \tau_1)x(t - \tau_2)] = \sigma^2 \delta_{\tau_1,\tau_2} \). Thus, \( G_2 \) is orthogonal to \( G_0 \) (constant). All the terms in \( G_2G_1 \) have expected values of zero as they each involve an odd number of Gaussian random variables. Consequently \( G_2 \) and \( G_1 \) are also orthogonal. For further information [65] can be read.

### 3.4.2 Wiener Model

The Wiener model consists of linear dynamic element followed by static nonlinearity. The output model is represented by:

\[ y(t) = f(g(t), \theta_n) \] (3.19)

where \( \theta_n \) is the parameter vector and \( g(t) \) is the output of the linear block.

\[ g(t) = G(q)u(t) \] (3.20)

where \( G(q) = \frac{B(q)}{F(q)} \) is a linear digital filter, \( u(t) \) is the input to the system and \( q \) is the forward shift operator.

The nonlinear function \( f() \) is often represented by a polynomial:

\[ y(t) = f[x(t)] = \sum_{l=0}^{M} c^l u^l(t) \] (3.21)
in which case the output of the Wiener model will become:

\[ y(t) = \sum_{n=0}^{M} c^{(l)} \left( \sum_{\tau_1=0}^{T-1} \cdots \sum_{\tau_n=0}^{T-1} h(\tau_1), \ldots, h(\tau_n) u(t - \tau_1) \cdots u(t - \tau_n) \right) \]  

(3.22)

Any nonlinear system with a fading memory can be represented by a Wiener model. The block diagram in Figure 3.1 represents the Wiener model with a single input multiple output (SIMO) LTI system followed by memoryless multiple input single output (MISO) nonlinearity.

![Block diagram of Wiener series approximator](image)

Figure 3.1: Block diagram of Wiener series approximator

### 3.4.3 Blind Identification of Wiener Model

A blind identification of a single input single output (SISO) Wiener model is carried out in [61]. The Wiener model had SISO linear and nonlinear parts. A cost function for PEM identification is constructed which generated the initial estimates of the linear and nonlinear parts of the Wiener model. This iterative optimization was carried out using the maximum likelihood principle. This was done using a few assumptions:

1. The nonlinear function \( y = f_0(i) \) is a monotonic bijective function. As shown in Figure 3.2, the static nonlinearity of the Wiener systems can be described as a mapping function from \( i \) to \( y \).

2. \( G_0 \) is a causal, stable and inversely stable monic transfer function. The linear time invariant (LTI) system is characterized by a transfer function \( G_0 \).
3. The nonlinear parameter vector $\theta_{NL}$ gives an unique representation of the nonlinearity. Since the identification uses a parametric approach, parameter vectors are divided into linear and nonlinear part, $\theta_L$ and $\theta_{NL}$ respectively. The LTI transfer function $G(z, \theta_L)$ can be parameterized as $G(z, \theta_L) = \frac{C(z, \theta_L)}{D(z, \theta_L)} = \frac{1+\sum_{r=1}^{n_c} c_r z^{-r}}{1+\sum_{r=1}^{n_d} d_r z^{-r}}$. Therefore, $\theta_L^T = [c_1, c_2, ..., c_{n_c}, d_1, d_2, ..., d_{n_d}]$. For the nonlinear parameterization an inverse function approximate $h$ is used where $i = h(y, \theta_{NL})$. Since $\theta_{NL}$ is unique, $h(y, \theta_{NL}) = h(y, \theta_{NL}^*)$, $\forall y$.

4. The inverse function $g$ is twice continuously differentiable.

5. The unknown input $e(t)$ is zero mean, white Gaussian noise.

6. The output $y(t)$ is known exactly that is, it is observed without any errors.

The system parameters are identified using the maximum-likelihood principle. First of all, the likelihood function is set up using the above assumptions. The maximum-likelihood estimates are then defined as the maximizing argument of the likelihood function. Under the assumptions 1 – 6 stated above, the Gaussian negative log-likelihood (NLL) function of the observations $y^T = [y(0), ..., y(N - 1)]$ is given by:

$$L(y|\theta, \lambda) = \frac{N}{2} \log(2\pi) + \frac{N-1}{2} \log(\frac{1}{N} \sum_{t=0}^{N-1} (G^{-1}(q, \theta_L)h(y(t), \theta_{NL}))^2) - \frac{N-1}{2} \sum_{t=0}^{N-1} \log|h'(y(t), \theta_{NL})|$$

(3.23)

given the model parameters $\theta$, and the input variance $\lambda$. $q$ is the forward shift operator and $(.)'$ is the derivative w.r.t $y$.

$\lambda$ can expressed by minimizing the NLL that is, $\delta L(y|\theta, \lambda)/\delta \lambda = 0$:

$$\lambda = \sum_{t=0}^{N-1} (G^{-1}(q, \theta_L)h(y(t), \theta_{NL}))^2$$

(3.24)

After eliminating $\lambda$ the NLL becomes:

$$L(y|\theta) = \frac{N}{2} \log 2\pi + \frac{N}{2} \log(\frac{1}{N} \sum_{t=0}^{N-1} (G^{-1}(q, \theta_L)h(y(t), \theta_{NL}))^2) + \frac{N}{2} - \sum_{t=0}^{N-1} \log|h'(y(t), \theta_{NL})|$$

(3.25)
A cost function can be easily obtained:
\[
V(y, \theta) = h^2(y, \theta_{NL}) \sum_{t=0}^{N-1} (G^{-1}(q, \theta_L)h(y(t), \theta_{NL}))^2
\]  
(3.26)

where, 
\[
g(y, \theta_{NL}) = e^{-(1/N) \sum_{t=0}^{N-1} \log|h'(y(t), \theta_{NL})|}
\]
is correction factor to the cost function. The cost function is written in sum of squares form.

\[
\hat{\theta} = \arg \min_{\theta} V(y, \theta)
\]  
(3.27)

The minimizer in Equation 3.27 can be calculated using the classical Gauss-Newton based iterative methods. As an iterative method can be used to find the Maximum Likelihood Estimator (MLE), good initial estimates are required to start the iterative procedure. This is done in two steps: First the monotonically increasing static nonlinearity is estimated and then the linear part is estimated. After the initial estimates of the model are identified the cost function is minimized.

Nonlinearity Estimation:
As shown in Figure. 3.2, the intermediate outputs \(i(t)\) are estimated using a nonlinear function \(f_0\) using the system outputs \(y(t)\). In this case, there are no inputs or the inputs are unknown. So it is necessary to determine \(i(t)\), which is then used to determine the linear model \(g_0\). This is done in two step: in the first step, an estimate of the nonlinearity function \(f_0\) is determined; in the second step, the intermediate signal \(i(t)\) is estimated. As the nonlinearity is bijective, it is assumed to be monotonically increasing. Hence for every \((i_0, y_0)\) such that \(i_0 = h(y_0)\):

\[
P(I \leq i_0) = P(Y \leq y_0)
\]  
(3.28)
where, $I$ and $Y$ are static stochastic processes related to the signals $i(t)$ and $y(t)$ respectively. $P(*)$ is the probability. The probability of finding a point within this region results in cumulative distribution of $i(t)$ and $y(t)$ being equal at $(i_0, y_0)$

$$F_I(i_0) = F_Y(y_0)$$  (3.29)

According to assumption 2, the input is Gaussian so the intermediate signal $i(t)$ is also Gaussian. Hence,

$$F_I(i_0) = \Phi(i_0)$$  (3.30)

The Gaussianity of the input implies that the intermediate signal $i(t)$ is Normally distributed. $\Phi(*)$ is the Normal standard distribution. The cumulative distribution function at $y_0$ is calculated by sorting out the measured $y$ values, then taking the ratio of the observations smaller than or equal to $y_0$ to the total number of observations, $N$.

$$F_Y(y_0) = \frac{\text{number of } y(t) \leq y_0(t)}{N} - \frac{1}{2N}$$  (3.31)

$\frac{1}{2N}$ is a finite sample correction, which is equal to half the number of the smallest discontinuity introduced by $F_Y(y_0(t))$. The range of $F_Y(y_0)$ is from $\frac{1}{N}$ to 1 and that of the distribution function of $y(t)$ is from 0 to 1.

Thus, the inverse nonlinearity function $h(*)$ is estimated, the next step would be to estimate the intermediate signal $i(t)$ The intermediate signal $i(t)$ is obtained by:

$$i(t) \cong \hat{i}_0(t) = \Phi^{-1}(F_Y(y(t)))$$  (3.32)

From (3.32) it can concluded that the nonlinearity can be removed from output data $y(t)$ and the intermediate signals $i(t)$ is estimated from the inverse nonlinearity function $h(*)$.

Linear Part Estimation and Minimization of Cost Function:

After this, the second part is to estimate the LTI part from the intermediate signal. This can be done by obtaining the estimate of the intermediate signal $i(t)$ and performing a blind
linear MLE of $\theta_L$ in Equation 3.26. Let the estimate be $\hat{\theta}_L^{(0)} y(t)$ of the LTI is inverse filtered using $G^{-1}(q, \hat{\theta}_L^{(0)})$, an estimate of the LTI input is obtained. The order of $G$ is selected as to whiten the power spectrum of the input signal. After the initial parameter estimates are obtained, the maximum likelihood sum of squares cost function is minimized using Gauss-Newton based iterative algorithm. This algorithm requires the use of the Jacobian matrix, defined as partial derivatives of the residuals vector (with entries defined as the terms added up quadratically in the cost function) w.r.t. the parameters.

In the general case the computation of Jacobian requires a calculation of mixed second order derivatives which are very difficult to compute. The process of obtaining the estimate of the linear part is as follows:

A linear-in-the-parameters model is used to describe the nonlinearity. This model basis function chosen was Hermite polynomials which are orthogonal w.r.t. the Gaussian probability density function. Explicit expressions for the Jacobian matrix is derived for this parameterization.

The expressions derived from the data and the parameterization helps in computation of the Jacobian. This Jacobian is used to minimize the cost function $V(y, \theta)$. The minimization is carried out using Levenberg-Marquart algorithm. This method helps yield the most likely parameters.

In this paper [61], a maximum likelihood estimator is used to estimate the SISO nonlinear and linear parts of a Wiener model. The MLE cost function is derived in the sum of squares form (3.27) and iterative Gauss-Newton method is used to calculate the estimates. This method is only applied to SISO blind identification and would be difficult to use in MIMO case. The ability to estimate the high quality initial estimates of the nonlinear function is utilized in this thesis.
3.5 Summary

Power system consists of many nonlinearities. This chapter identifies some of these nonlinearities and also summarizes how these nonlinearities can alter the mode estimates of a power system. Power system data has a possibility of containing nonlinearities. The detection of nonlinearities are using linear and Gaussian tests are given.

An introduction to nonlinear systems leading to Wiener systems is given. Wiener models are universal approximators for nonlinear systems. Blind identification of Wiener models to estimate the nonlinear function of a given nonlinear system is discussed.
Chapter 4

BLIND IDENTIFICATION OF
ELECTROMECHANICAL MODES USING A
WIENER MODEL

4.1 Introduction

In this chapter, the proposed algorithm: blind identification of electromechanical modes of a power system using a Wiener model is detailed and tested with simulated data. This identification is performed for MIMO power system with only the outputs being available. Simulated data obtained from test power system network are used to determine the numerical results. The data are tested for presence of nonlinearity. If there are any nonlinearities present in the data it is fitted into a nonlinear model, Wiener model, to estimate the power system modes. For comparison and validation, the Wiener model estimates of the power system modes are compared with a linear estimator and the real values of the simulated system.

4.2 Proposed Algorithm

In this section, the proposed algorithm is detailed. Power systems have a lot of nonlinearity so even ambient data are prone to consist certain amount of nonlinearity. As discussed in Chapter 3, it is accurate to use a nonlinear model to estimate power system modes from nonlinear data. When the measured data is available, the next step is to test the data for any nonlinearity. The steps can be summarized as follows:

1. Determination of nonlinearity in power system: Prior to identifying the power
system modes, it is necessary to determine whether the power system data have any nonlinearity present in it. This is done using linearity tests.

2. Estimation of the power system nonlinearity: After carrying out the linearity tests on the power system data, if the data is nonlinear, the nonlinearity is estimated using Wiener model. Blind identification of Wiener model provides an estimate of the nonlinearity function of the nonlinear data. Using this nonlinearity function, a set of linear intermediate data can be estimated.

3. Subspace identification: The intermediate data derived from the Wiener model is used for linear analysis to estimate the power system model. This is carried out using subspace identification. The algorithm N4SID is used.

4. Conversion of discrete to continuous time state space models: A discrete time power system model is estimated by N4SID, thus it is necessary to convert it to continuous time model.

The following sections introduces the method used to estimate power system modes from nonlinear data using Wiener models:

4.2.1 Linearity Test

In order to detect possible nonlinearities in power system Gaussianity and Linearity Tests are carried out. The tests were described in details in Chapter 3. The data is collected in the form of time series and assumed to be a linear process:

\[ y(t) = \sum_{i=0}^{\infty} h(s) \epsilon(t - s) \]  

(4.1)

In this case, \( \epsilon(t) \) is considered to be independent identically distributed (i.i.d) with zero mean. Therefore, third order cummulants \( c_{(a,b)} \) = \( E[y(t)y(t + a)y(t + b)] = 0 \) for most values of a and b, if \( \epsilon(t) \) is normal and \( E[\epsilon^3(t)] = 0 \). In this method, an estimator of the
bispectrum, the Fourier transform of \[ c_{(a,b)} \] is determined. The bispectrum \( S_y \) is used to construct a statistic to test whether the bispectrum of \( y(t) \) is non-zero. The null hypothesis being: if the bispectrum is zero, then \( y(t) \) is Gaussian. If the relationship between \( y(t) \) and \( \epsilon(t) \) is nonlinear then \( y(t) \) will be non-Gaussian even if \( \epsilon(t) \) is normal.

In order to test whether \( y(t) \) is linear another test statistic considering constant bicoherence is determined. If the bicoherence is constant for a certain range, then \( y(t) \) will be linear.

In the Gaussianity tests, the null hypothesis is that the data must have zero bispectrum. The test statistic is considered approximately chi-squared. The probability of false alarm (Pfa) is probability that the value of the chi-squared random variable will exceed the computed test statistic. The Pfa value indicates the possibility of being wrong in assuming that the data has a non-zero bispectrum. If this probability is small, say 0.95, the assumption of zero bispectrum is accepted.

In the Linearity test, the range \( R_{est} \) of values of the estimated bicoherence is computed and compared with the theoretical range \( R_{theory} \) of a chi-squared random variable with two degrees of freedom and non-centrality parameter \( \lambda \). \( \lambda \) is directly proportional to the mean value of the bicoherence value. If the estimated and theoretical ranges are very different from each other, the data is not linear. The tests are carried out using the High-Order Spectral Analysis Toolbox (HOSA) in MATLAB. Usage of the toolbox is given in [47].

Using this tests if the power system data consist of nonlinearity, then it is necessary to estimate this nonlinearity. This nonlinearity can be determined using blind identification of Wiener models. The next section describes how it is done.

### 4.2.2 Estimating Nonlinearity

The nonlinearity can be detected using linearity tests. After the data is proven to have nonlinearities present in it, the next step would be to estimate the equation representing this nonlinearity. Blind identification of Wiener models was described in Chapter 3. During blind identification of Wiener models an initial estimate of nonlinearity is constructed. Using this
nonlinearity an estimate of intermediate signals was received. The method is briefly outlined here:

First of all, it is assumed that the nonlinearity function $g()$ is considered to be bijective and monotonically increasing/decreasing and the unknown intermediate input $i(t)$ is considered to be Gaussian. The estimation is carried with multiple outputs vectors $[y_0, y_1, ..., y_M]$ where $y_k \in \mathbb{R}^{t \times 1}$ for $k = 0, 1, ..., M$. but with one output at a time. For the first output vector $y_0$, the nonlinearity function is estimated in the following lines. According to the assumptions, for any point $(i_{01}, y_{01})$ such that $i_{01} = g(y_{01})$ The probability of finding a point within this region results in cumulative distribution of $i_0(t)$ and $y_0(t)$ being equal at $(i_{01}, y_{01})$

$$F_{i_0}(i_{01}) = F_{y_0}(y_{01}) = \Phi(i_{01})$$

(4.2)

where $\Phi(*)$ is the Normal standard distribution.

The cumulative distribution function at $y_{01}$ is calculated by sorting out the measured $y$ values, then taking the ratio of the observations smaller than or equal to $y_{01}$ to the total number of observations, $N$.

$$F_{y_0}(y_{01}) = \frac{\text{number of } y(t) \leq y_{01}(t)}{N} - \frac{1}{2N}$$

(4.3)

Thus, the nonlinearity function $g(*)$ is estimated the next step would be to estimate the intermediate signal $i_0(t)$ The intermediate signal $i_0(t)$ is obtained by:

$$i(t) \cong \hat{i}_{01}(t) = \Phi^{-1}(F_{Y}(y(t)))$$

(4.4)

Similarly for every output $y(t)$ the intermediate signals are estimated using Equation 4.4. These signals are tested with linearity tests and should be linear. These linear data are used to carry out subspace identification of the power system model.

4.2.3 Subspace Identification

Stochastic subspace identification was explained in details in Chapter 2. Numerical Algorithm for Subspace State Space IDentification (N4SID) implementation of subspace identi-
Subspace methods can be used to identify blind MIMO power systems. For blind identification using subspace method, the power system model is expressed as:

\[
\dot{x}(t) = \hat{A}\hat{x}(t) + Ke(t) \tag{4.5}
\]
\[
y(t) = \hat{C}\hat{x}(t) + e(t)
\]

where \( \hat{A} \) is the estimated state matrix and \( \hat{C} \) is the estimated output matrix and \( K \) is the Noise gain matrix which is chosen by the algorithm automatically. The input control signals \( u \in \mathbb{R}^m \) are unknown and its effect can be considered merged into the errors. The estimation errors can be denoted by \( e(t) \).

Eigenanalysis of the estimated state matrix \( \hat{A} \) is used to determine the power system modes. To estimate the appropriate mode shape the estimated output matrix \( \hat{C} \) is multiplied by the \( \phi_i \), the \( i \)th right eigenvector of \( \hat{A} \), as shown in Equation 4.6.

\[
\phi_i = \hat{C}\phi_i \tag{4.6}
\]

The resultant mode shape of measured output \( y_m \) relative to \( y_n \) is the ratio of:

\[
\text{Mode Shape} = \frac{\phi_{m,i}}{\phi_{n,i}} \tag{4.7}
\]

Implementation of N4SID algorithm based subspace identification is done using the function \textit{n4sid}() function in System Identification Toolbox of MATLAB. The system carries out subspace identification using discrete time state space models. In order to estimate modes the discrete time state space models should be converted to continuous time. The next section explains how this is done.

4.2.4 Conversion of Discrete to Continuous Time State Space Models

The relationship between continuous and discrete time state space models was given in Chapter 2. It was deduced that the poles of discrete and continuous time are the eigenvalues of their respective state matrices. In both cases, eigenanalysis of the state matrix that is
\[ \text{det} | \lambda I - A | = 0 \] provides the poles. It was observed that discretization of continuous time poles can be done by:

\[ z = e^{sT} \]  

(4.8)

In this thesis, discrete time state space identification is performed. Therefore, the eigenvalues \( \lambda_{\text{dis}} \) calculated from the discrete state matrix must be converted to continuous time. This is done by:

\[ \lambda_{\text{con}} = \frac{\ln(\lambda_{\text{dis}})}{T} \]  

(4.9)

These are the power system modes.

In Chapter 2, section 2.5 details of the electromechanical modal properties of a power system and how are they determined. The modes are estimated by eigenanalysis of linearized power system state space equations. The modal properties of power system modes from measured data requires the estimates of the state matrix \( A \) and the output matrix \( C \) of the linearized power system. This is achieved by estimating \( A \) and \( C \) and then performing eigendecomposition.

4.3 Test Data

The Kundur two area network shown in Figure 4.1 was used to generate the simulation data. The system contains 11 buses and two different identical areas connected by weak tie-line. There are two load buses (buses 7 and 9). Each area has two identical generators. Each generator is represented using sub-transient model with exciters, turbine governors and power system stabilizers. These devices add a lot of dynamics and are highly nonlinear to the power system. For modal analysis, the Power System Toolbox (PST) was used. PST is a collection of MATLAB files that performs different power system analysis such as power-flow, transient stability analysis and small signal stability analysis. All these analysis methods are coded in MATLAB functions [4]. PST can be downloaded from: http://www.eps.ee.kth.se/personal/vanfretti/pst/Power_System_Toolbox_Webpage/Downloads.html
For estimating modes, the system is perturbed with small disturbances such as random load changes. Small signal stability analysis using PST is done. A detailed explanation of small signal stability analysis in PST can be studied from [32]. In PST, the generators are modelled as 6th order models consisting of 6 states for each machine. A classical generator model with two states, models the generator as a constant voltage behind a transient reactance. In case of a 6th order model the machine is modelled with voltage behind subtransient reactance. This model includes the effects of the field windings and damper windings. The effect of damper windings are divided with one effecting the direct-axis and two on the quadrature-axis. The different generator states are defined as shown in Table 4.1.

PST uses the standard exciter models outlined by the Institute of Electrical and Electronic Engineers (IEEE). There are several exciter models explained in details in [4]. In this thesis, a static exciter is used. A static exciter receives its excitation power from the generator terminal bus. There are up-to four exciter states in the power system.

Power system stabilizers (PSS) are connected to the exciter models. The input to the PSS is the normalized bus frequency or generator power. The output of the PSS is fed to the reference summing junction of the exciter.

Turbine governors are needed to drive the synchronous generators. Turbines have a speed governing control. A governor controls the speed of the isolated generator but once it is connected to the system the speed is synchronized. The action of a governor is then used to control the power output of the generator. It determines the amount of load change by the generator. Two types of turbines are used in PST: hydraulic and thermal. In this simulation

<table>
<thead>
<tr>
<th>State Variable</th>
<th>State description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>rotor angle</td>
</tr>
<tr>
<td>$\omega$</td>
<td>rotor speed</td>
</tr>
<tr>
<td>$\psi_f$</td>
<td>field winding flux linkage</td>
</tr>
<tr>
<td>$\psi_{kd}$</td>
<td>d-axis and damper winding flux linkage</td>
</tr>
<tr>
<td>$\psi_{kq1}$</td>
<td>first q-axis damper winding flux linkage</td>
</tr>
<tr>
<td>$\psi_{kq2}$</td>
<td>second q-axis damper winding flux linkage</td>
</tr>
</tbody>
</table>

Table 4.1: Dynamic states of a sub transient generator
thermal turbines were used. The model of the thermal turbine is described in details in [4].

Loads can be modelled in PST as constant power (active or reactive), constant impedance, constant current or a combination of all three. The modelling of loads are of significant importance as they contribute to accurate estimation of the system modes.

PST can be used to carry out small signal stability analysis as stated in Chapter 2. A single driver, 'svm_mgen', for small signal stability is provided. Details of carrying out small signal analysis is provided in the PST manual available in the download site.

Small perturbations can be applied to the system model. In this simulation, random load variations are added. To imitate random load fluctuations 1 per cent of Gaussian white noise was introduced into the loads. The loads were modelled as constant current load with both active and reactive loads. Both the active and reactive loads were perturbed with white noise of zero mean and 0.01 variance. The power system was simulated for 10 minutes with a time interval of 0.01 seconds. The voltage angles from all the 11 buses are collected. The voltage angle data collected from a single bus (Bus 3) is shown in Figure 4.2.
4.3.1 Preprocessing

The raw data collected from the Kundur two-area network power system is preprocessed using the methods proposed in [7], except that in that study, the initial data were taken from Phasor Measurement Units sampled at 30Hz, whereas in this study, the data were generated by a numerical simulation with a sample time of 0.01 seconds. The preprocessing steps along with their respective figures are summarized below:

1. The data obtained is in terms of phase angles so it necessary to obtain the bus frequency data. This is done by passing the data through a first order derivative filter.

2. The derivative filter introduces high frequency amplifications. These should be removed by passing the data through a low pass filter. This also provides
3. Phasor measurements units use data at 5 samples per second. This 5 Hz frequency is appropriate for the Kundur two area network because there are no low-frequency inter-area modes around this frequency.

4. The data is then passed through a high pass filter. This filter removes low frequency trends introduced by power system elements such as governor dynamics.

5. After the data is passed through different filters high amplitude transients are
introduced at both ends of the data. These remaining transients are chopped off/removed using impulse response filtering. This filter minimizes the initial start-up and ending transients. It also preserves the features in the filtered time waveform exactly where the features in the unfiltered waveform.

After the data is preprocessed it is used for identifying modes.

4.3.2 Monte Carlo Simulations

The performance of the method introduced is evaluated in a statistical method using Monte Carlo simulations due to the stochastic and randomness of the ambient excitations. Monte
Carlo trials helps estimating the statistical performance of any system by providing several identically distributed independent random samples to use as its inputs. In this thesis, the mean and standard deviation of the estimated mode frequencies and damping ratios are calculated. For comparing the mode shapes, the bias is calculated. The Monte Carlo simulation procedures are as follows:

- All the load buses are set as random by introducing additive white Gaussian noise. Dynamic simulation of the system is carried out for 10 minutes. The procedure is repeated for 100 times, each time perturbing the loads with a different sequence of white Gaussian noise.

- The generated response data is collected. In this case, the time-series of the voltage phase angles from each individual bus are gathered. This results in 100 sets of data.

- Each trial is then processed as described in Section 4.3.1. First, the signals are preprocessed, as described in Section 4.3.1, and then a MIMO Wiener system is fitted to the preprocessed data using the proposed algorithm. The statistics are calculated based on 100 times of the experiment.

- The statistics, of the identified modal parameters, are then estimated by com-
puting averages over the ensemble of Monte Carlo trials. Thus, the mean is estimated by calculating the sample mean. The standard deviation is estimated by the sample standard deviation of the Monte Carlo trials and the bias is estimated by subtracting the actual value from the sample mean of the Monte Carlo trial results.

4.4 Numerical Validation of Proposed Algorithm

Simulated data from the Kundur two area network is used to test the algorithm described earlier. The data is initially tested for nonlinearity. After the presence of nonlinearity is detected, the data is used to determine the nonlinearity function using blind identification of Wiener model. Using these nonlinearity function a set of intermediate data are derived. These data are tested again for nonlinearity. If these signals are linear they are used to estimate mode frequency, damping ratio and mode shape using eigenanalysis of power system determined by N4SID algorithm. These modal properties are compared with linear analysis method. This comparison shows how the results vary if nonlinear data are fitted to linear models. To further determine the use of nonlinear model, Wiener model for mode estimation the prediction errors are determined. The prediction errors are tested for Gaussianity as Gaussian inputs were provided to the simulated systems. The models were tested to see which of them provided a better model fit for the power system data. At the end, the number of inputs provided to the system are determined. The system is perturbed with white Gaussian noise at two of the load buses.

4.4.1 Nonlinearity Detection

The data collected from the simulated power system is checked for nonlinearity using linearity tests. The probability of false alarm (Pfa) that is, the probability of being wrong in assuming a nonzero bispectrum and the difference between \( R_{\text{est}} \) and \( R_{\text{theory}} \), is calculated for all of the
time series. The Pfa calculated for the data are very small for each bus, about 0.1. The Pfa results are displayed in Figure 4.8. The null hypothesis for Gaussianity is rejected and the bus data are not Gaussian. The deviations in the estimated and theoretical ranges are unreliable for data with non-zero bispectrum (Gaussian). Therefore none of the power system data were Gaussian nor linear.

For each time-series of voltage angle, the inverse nonlinearity is estimated using Wiener model’s inverse nonlinear function. After this, the intermediate signal $i(t)$ is estimated for every individual voltage angle time-series. These set of data are again tested for Gaussianity and linearity.

The Gaussianity test results displayed in Figure 4.9 shows that almost all of the probabilities computed were larger than 0.95. In this case, all the signals are Gaussian. The difference between the estimated and theoretical ranges, $R_{est} - R_{theory}$ are computed and plotted in Figure 4.10 for each of the signals. The estimated and theoretical ranges are close to one another; that is the data are linear.

Figure 4.11 shows the nonlinear function $i = g_0(y)$ constructed by the Wiener model. In order to determine the nonlinearity function using different other methods, third order
Figure 4.9: Gaussianity test results of intermediate signal data from the Wiener model

Figure 4.10: Linearity test results of intermediate signal data from the Wiener model
polynomial and Hermite polynomial are used in Figure 4.12. Details of these methods are given in [65]. Both the methods provide good estimates of the nonlinear function. They look exactly the same as the Wiener model derived nonlinearity function with small amount of fitting error. These linear intermediate data are used to estimate modes using subspace identification. The real modes are compared with the estimated modes of Wiener and Linear model.

4.4.2 Mode Estimation

After the data are tested for nonlinearity and passed through the Wiener model’s inverse nonlinearity function to derive linear intermediate data, it is used to estimate power system modes. The power system modes estimated by the Wiener and Linear model are compared with the real values of the simulated power system modes. First the estimated mode frequency and damping ratio are compared, then the mode shape. In both the case statistical comparison using Monte Carlo trials are used.
Figure 4.12: Estimating the inverse nonlinear function using different methods and their corresponding fit errors.

Modal Frequency and Damping Ratio

A comparison of the estimated results, using Wiener and linear model with the real modes of the test bus system is shown in Table 4.2. The real modes are calculated by small signal analysis in PST (Power System Toolbox) in MATLAB. The toolbox considered 60 states. The system had two modes. One inter-area mode at 0.61 Hz and the another local mode at 1.4 Hz.

Table 4.2: Comparison of estimated results with the real modes of Kundur two area network

<table>
<thead>
<tr>
<th>Method</th>
<th>Mode</th>
<th>Frequency (Hz)</th>
<th>Damping ratio(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>Real values</td>
<td>1</td>
<td>0.6125</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.4882</td>
<td>-</td>
</tr>
<tr>
<td>Linear Model</td>
<td>1</td>
<td>0.6027</td>
<td>0.0069</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.5000</td>
<td>0.0017</td>
</tr>
<tr>
<td>Wiener Model</td>
<td>1</td>
<td>0.6120</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.5000</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

The results show that both the methods provide nearly identical estimates of the modal
frequencies, but the Wiener model provided a much better estimate of the damping ratios. In case of the inter-area mode (mode 1), the Wiener model estimates had an error (mean) 20 times smaller than that from the linear model, and the standard deviation was 3 times smaller. For the local mode (mode 2), error was 13 times smaller and the standard deviation was almost 500 times smaller using the Wiener model.

While estimating the damping ratio, both the methods could provide estimates close to the damping of the real inter-area mode. On the other hand, the local mode damping estimates produced by the two models were very different. The Wiener model provided a much better estimate when compared with the linear model. The Wiener model had accurate results for both high and low damping and both inter-area and local modes. The linear model estimates were unreliable for local mode with a low damping ratio.

Therefore overall the Wiener model estimates were more reliable and accurate with less standard deviation and closer mean when compared to the real mode values. Linear models could detect inter-area modes with high damping but provided poor results for low damped local mode estimates.

Mode Shape

A comparison of the mode shape 1 estimates with the real mode shapes are depicted in Figure 4.13. The real mode shape 1 shows that buses 1 and 2 in area 1 are swinging against buses 3 and 4 in area 2. Thus this mode is an inter-area mode. Mode shape 2, the mode shape compared with the estimated mode shapes are depicted in Figure 4.14. In case of mode shape 2, the generation bus 3 is swinging against bus 4 at 1.4Hz. The area-1 buses are swinging in the same direction. Bus 1 is chosen as the reference bus and the mode shapes are normalized by dividing each vector with bus 1 mode shape.

The numerical results for mode shape 1 and mode shape 2 estimates are tabulated in Table 4.3 and Table 4.4 respectively. Both methods provided good estimates for mode shape 1, but in case of the Wiener model, the error was reduced 5 times for the magnitude
estimates and 2 times for the angle estimates. The biases calculated for the Wiener model were much reduced when compared to that of the linear model.

In case of mode shape 2, the Wiener model provided excellent estimates almost identical to the real mode shape. The errors and standard deviation for both angles and magnitudes were significantly reduced. The standard deviation of the magnitudes were reduced to at least 50 times and the angles about 10 times.

The results suggest that the Wiener model is a better estimator for both the mode shapes. It had lower deviation and biases. It had more similar estimates compared to the real mode shapes. In general the Wiener model estimates were good for both the local and inter-area mode shape. In case of the linear model, the mode shape estimates were good for the inter-area mode but failed to produce as good estimates as the Wiener model.
Table 4.3: Estimated values of Mode shape 1 normalized w.r.t Bus 1

<table>
<thead>
<tr>
<th>Bus</th>
<th>Linear Model</th>
<th>Wiener Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>1</td>
<td>Mag.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Mag</td>
<td>0.742</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>3.71</td>
</tr>
<tr>
<td>3</td>
<td>Mag</td>
<td>0.926</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>154.6</td>
</tr>
<tr>
<td>4</td>
<td>Mag</td>
<td>0.749</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>153.4</td>
</tr>
</tbody>
</table>

Table 4.4: Estimated values of Mode shape 2 normalized w.r.t Bus 1

<table>
<thead>
<tr>
<th>Bus</th>
<th>Linear Model</th>
<th>Wiener Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std.</td>
</tr>
<tr>
<td>1</td>
<td>Mag.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Mag</td>
<td>2.020</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>14.94</td>
</tr>
<tr>
<td>3</td>
<td>Mag</td>
<td>0.3934</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>39.05</td>
</tr>
<tr>
<td>4</td>
<td>Mag</td>
<td>0.7684</td>
</tr>
<tr>
<td></td>
<td>Angle (°)</td>
<td>178.78</td>
</tr>
</tbody>
</table>

4.4.3 Comparison of Wiener and Linear Model

In this section, the Wiener model is compared with linear model. The modal estimation of both the models were good but as a model for power system data, which model had better results will be shown here. First of all, which model provides a better data fit is determined. Then the prediction errors of each of the models are found out and compared.
Table 4.5: Comparing model fits with linear and Wiener Model

<table>
<thead>
<tr>
<th>Outputs</th>
<th>Goodness Fit(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear Model</td>
</tr>
<tr>
<td>$y_1$</td>
<td>73.00</td>
</tr>
<tr>
<td>$y_2$</td>
<td>68.47</td>
</tr>
<tr>
<td>$y_3$</td>
<td>66.19</td>
</tr>
<tr>
<td>$y_4$</td>
<td>66.03</td>
</tr>
<tr>
<td>$y_5$</td>
<td>70.58</td>
</tr>
<tr>
<td>$y_6$</td>
<td>72.78</td>
</tr>
<tr>
<td>$y_7$</td>
<td>68.07</td>
</tr>
<tr>
<td>$y_8$</td>
<td>64.64</td>
</tr>
<tr>
<td>$y_9$</td>
<td>64.33</td>
</tr>
<tr>
<td>$y_{10}$</td>
<td>67.31</td>
</tr>
<tr>
<td>$y_{11}$</td>
<td>67.45</td>
</tr>
</tbody>
</table>

Model Fits

A comparison of the two different models were done by comparing the model output and the measured output. This is done using MATLAB function `compare()`. This function provides the normalized root mean square (NRMSE) measure of the goodness of the fit. NRMSE fitness value is measured using (4.10)

$$fit = 100(1 - \frac{||y - \hat{y}||}{||y - \text{mean}(y)||}) \quad (4.10)$$

where $y$ is the power system data and $\hat{y}$ is the estimated output of the model. The goodness fit is calculated for each of the outputs for the data obtained from the Wiener model and linear model. Table 4.5 provides the comparison results. The fit using the Wiener model had a higher percentage for each and every output. The NRMSE fitness value for the Wiener model outputs were almost 100 percent and on the other hand the linear model had a value around 70 percent.
Prediction Errors:

Using the identified model the one-step ahead prediction of the outputs $y_{p1}$ are calculated for both the linear and Wiener model. The estimation errors are calculated by deducting the outputs $y$ from $y_{p1}$. The probability density functions (PDFs) of the prediction errors are shown in Fig.4.15. The estimation errors are normally distributed. Therefore the errors are Gaussian. The distribution obtained from Wiener model has a more smother distribution when compared with the distribution obtained from the linear model. Therefore, the results suggest that Wiener model prediction errors are more Gaussian. The errors obtained from the Wiener model are much smaller. The prediction errors obtained from the two different algorithms were compared. When using the Wiener model, the prediction errors were much smaller almost one-third of that of the linear model errors. The distribution graphs of the Wiener model errors were perfectly normally distributed. The distribution graph of the errors obtained from the linear model were less Gaussian when compared to that of the Wiener model.

Therefore it can be stated that the Wiener model has less prediction errors and the errors
are more Gaussian.

Autocorrelation of the Prediction Errors:
Tests are carried out to determine the residuals $e(t)$ of the identified model while using both linear and Wiener model. The autocorrelation function of the prediction errors $e(t)$ are computed and displayed in Figure 4.16. The 99 per cent confidence intervals for these values are computed and represented by the pink region. Neither sets of the residuals show that they are white noise as there are significant correlations at non-zero lags that is 1, 5 and 6. There is not much significant difference between the two plots. This reason could be the simulation is under-ordered for a power system. There could be some unmodelled dynamics present in the system. Therefore summing up the comparison results, it can be stated that using a Wiener model the prediction errors are minimized and more Gaussian.

![Correlation function of residuals of Output $y_5$ using Wiener Model](image1)

![Correlation function of residuals of Output $y_5$ using Linear Model](image2)

Figure 4.16: Correlation plots of the residuals from Output $y_5$ using a Wiener and Linear Model Data
Estimating the number of Inputs:
The simulated power system had two load buses. The load buses were perturbed with Gaussian white noise. That is two different inputs were provided to the simulated system.

In order to determine the number of inputs, a matrix of the prediction errors were constructed. The matrix consisted of 11 columns, each containing the prediction errors from their respective buses. The prediction errors were deduced in the previous subsections. Singular Value Decomposition (SVD) of these prediction errors is determined. The number of singular values determines the number of independent components in the data, and hence the number of inputs.

In Figure 4.17 the number of singular values obtained by performing SVD of the prediction errors is shown. In this case, two large singular values are obtained which states that there were two inputs to the system.

4.5 Discussions

In this chapter, estimates of the modal frequency, damping ratio and mode shapes are determined using both a Wiener model and a linear Model. Both the model estimates are compared with the real modes of the simulated power system. The Wiener model can be used to identify power system mode properties more accurately than is possible using a
linear model. The estimation and prediction errors were relatively less when compared with
the linear model. Wiener models had a very good overall performance of providing mode
estimates for both inter-area and local modes. They also had minimum prediction errors
which were normally distributed. The errors were more Gaussian when using a Wiener
model, which is expected since the noise introduced to the power system were Gaussian.
Lastly, the power system data was a better fit with the nonlinear Wiener model.
Furthermore, given that although both the Wiener and linear models yield the same modal
frequencies and damping ratio, a more accurate estimation of the mode shapes are obtained
using a Wiener model.
Chapter 5

CONCLUSION

5.1 Overview

This thesis presented a method to identify electromechanical modes of a power system using a Wiener model. The method estimates the modes by fitting nonlinear data to a nonlinear model. The model initially estimates the nonlinearity present in the power system, then determines estimates of the mode frequency, damping ratio and mode shape and at last, prediction errors of the estimation are deduced.

Chapter 2 gave an overview of the system identification theory and mode estimation. Mathematical formulation of the system identification theory was provided. This introduced the problem to be solved in this identification of power system modes as blind identification of multiple input multiple output (MIMO) system. The identification tools required in this thesis were also introduced. In case of blind MIMO identification, subspace identification was utilized. The algorithm used in subspace identification was N4SID.

Finally modal estimation using eigenanalysis is discussed. This concluded that power system modes are the eigenvalues of the state matrix of a linearized power system state equations.

Chapter 3 provided information about the nonlinearities present in a power system which can affect the modal estimation results. Wiener theory background was introduced. The usage of the Wiener model to provide initial estimates of the nonlinearity as used in [61] was discussed. This method was used in this thesis to provide an initial estimate of the nonlinearity in the power system data.

Chapter 4 dictated the original work of the thesis. It described the whole methodology used in this thesis and a simulation example was provided. Initially, the simulated power system data was tested with linearity and Gaussianity tests to determine whether there are
nonlinearities present in the data. The data, that is the output data, was used in blind identification of power system modes using a Wiener model. Using these output data the nonlinear function was derived and a set of linear data were achieved. These data were used to identify the power system model using subspace identification with N4SID algorithm. Using this model, the modes of the power system were estimated by eigenanalysis. Simulation results obtained were displayed and validation of the Wiener model was done by comparing its estimation results, prediction and estimation errors with the linear model.

5.2 Contributions

The main contributions of the thesis can be summarized as follows:

1. Linearity and Gaussianity tests were applied to the simulated power system data and showed the presence of measurable nonlinearities. These power system nonlinearity was estimated using blind identification of a Wiener model. An estimate for the nonlinearity function of each of the power system data sets were achieved. These nonlinearity functions were used to derive a new set of linear intermediate data. These data were passed through a linear dynamic model in order to identify the power system modes. The nonlinearity function derived from the nonlinearities present in the simulated power system data were shown in Chapter 4.

Power systems are highly random and nonlinear. Nonlinearity exists since power system networks are interconnection of several nonlinear elements such as generators, motors loads etc. Measured power system data is likely to exhibit nonlinearity due to random load changes, nonlinear trends from slow control actions in operating conditions, faults or major cascading events. Nonlinearities can affect the estimation results of power system oscillation. Therefore it is necessary to determine whether the data consist of any nonlinearity.
2. In the simulation results the power system data were shown to have nonlinearity present in them. These data were used to estimate the nonlinearity using a Wiener model and then used to estimate power system modes using subspace identification. The results were exactly similar to the real values; estimation results were accurate even with nonlinearities present in the simulation data. Using a Wiener model accurate estimates of the modes were achieved for both high and low damping local and inter area modes. The electromechanical modes of a power system were estimated accurately from nonlinear power system data. Power system data are prone to have nonlinearity present in them. Subspace identification alone cannot provide good estimation results when there are nonlinearities. Wiener models with multiple input nonlinearity is a universal model for nonlinear dynamic systems. The nonlinearities in the power system can be anywhere, but using Wiener systems it can be represented as a linear dynamic system with nonlinearities placed at the output.

3. Using a Wiener model was shown to dramatically reduce the prediction errors in the blind identification of power systems. The numerical results verify that while using a Wiener model the prediction errors were reduced. The simulation data were a better fit to the Wiener model than the linear model. The simulated system was perturbed with Gaussian white noise. Wiener model errors were more Gaussian compared to the linear model stating that the nature of the errors were more similar to the original input. System identification means building mathematical models of dynamic systems using measured data. There will always be a certain amount of prediction error in the model used. The model is a better fit when the errors are minimum.
5.3 General Conclusions

Wiener models had a very good overall performance of providing mode estimates for both inter-area and local modes. The mode shapes estimations of the Wiener model were exactly the same as the real mode shapes. They had minimum prediction errors which were normally distributed.

Nonlinearity present in power system data should be detected otherwise linear models are unable to provide reliable and accurate estimates for the electromechanical modes. On the other hand, if nonlinear models are used these power system nonlinearities are taken in account and estimation results are improved.

5.4 Future Work

In order to gain a greater understanding of the performance of the Wiener models as power system mode estimators, much work must be completed. Wiener models showed a promising result when compared to formerly used linear models.

The method should be applied to real power system data, that is data from PMUs.

The algorithm was only applied to ambient power system data. The performance of the method can be tested with different types of power system data such as ring-down and probing data. The accuracy and reliability of the results obtained should be determined.

Finally, the inputs of the blind identification should be determined. It should be recalled from chapter 4, that there were two inputs in this simulated system. The next step would be to determine these inputs.
Bibliography


