Master's thesis

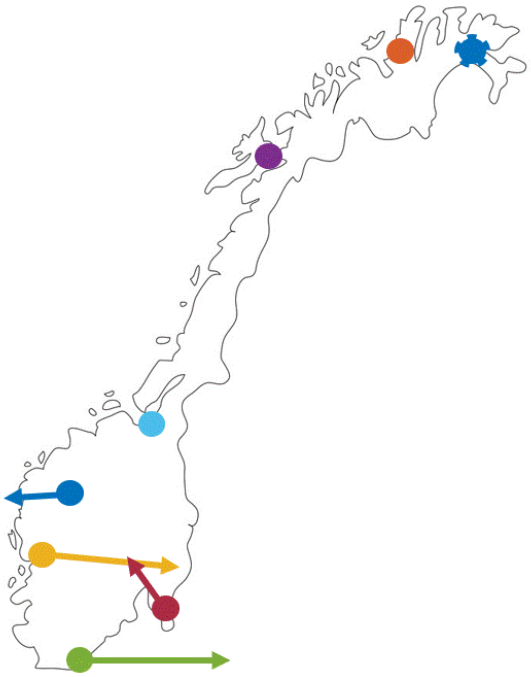
NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electric Power Engineering

Maiken Borud Omtveit

Visualizing Electromechanical Oscillations for Control Room Purposes

With Mode Shape Estimation Using N4SID+ DBSCAN

Master's thesis in Energy and Environmental Engineering
Supervisor: Kjetil Uhlen
June 2022



NTNU
Norwegian University of Science and Technology

Statnett
Fremtiden er elektrisk

Maiken Borud Omtveit

Visualizing Electromechanical Oscillations for Control Room Purposes

With Mode Shape Estimation Using N4SID+DBSCAN

Master's thesis in Energy and Environmental Engineering
Supervisor: Kjetil Uhlen
June 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electric Power Engineering

NTNU
Norwegian University of
Science and Technology

Abstract

In accordance with the Paris Agreement, renewable converter-based energy sources take a larger share of the power production as fossil-fuelled synchronous generators are phased out. In a power system with a decreasing share of synchronous generators, electromechanical oscillations are ever more prominent. Poorly damped modes may lead to multiple stability issues, and, in the worst case, islanding of parts of the grid. It is therefore important to control room operators to have a good overview of the modes that are present in the system. This is made possible with synchronized measurements from PMUs. This thesis seek to lay a foundation for further development of oscillation monitoring tools for the Nordic system operators.

The task of creating a good oscillation monitoring tool is divided in two. Firstly, it is important to make a visualization tool that fits the task. Electromechanical modes are a quite abstract phenomena, and it is important to monitor the essential parameters in a way that is easy for the operators to interpret in hasty situations in the control room. This will be addressed through a literature review of existing visualization tools and an interview with a Statnett control room operator. An overview of visualization methods is presented, and what makes a good visualization for monitoring electromechanical oscillations is discussed.

In a control room setting, there is a need for both system overview during ambient conditions, and during urgent situations to handle dangerous oscillations. Visualization methods that display information about the state of the modes in an intuitive way are favored for system overview, such as color coordinated text, compass plots, GIS and spectrograms for known modes. GIS with phasors, traditional and 3D spectrogram, compass plots and detailed text or graphs about the most critical modes in the system are recommended for urgent situations to help the operators to make good decisions.

Secondly, the algorithm that estimates the modal parameters must be accurate and fast. A literature study on this topic has also been conducted, and an algorithm based on system identification with N4SID and clustering of eigenvalues with DBSCAN is developed. N4SID traditionally needs the model order as input. This is avoided with use of clustering on an interval of model orders. This also greatly benefits the estimates, as true modes are clustered and noise is filtered out.

The algorithm is tested on simulated and real PMU data, and the results are promising. The algorithm quickly finds new modes in the system after a topology change with relatively good accuracy. It also finds modes in the real PMU data coherent with previous studies. Initial results favor rotor speed and angle as input signals to the algorithm. The major drawback is that the algorithm is rather slow, and that it struggles to separate modes that are close to each other in damping and frequency.

Sammendrag

Økt satsing på fornybare konverter-baserte energikilder og utfasing av eldre synkrongeneratorer basert på fossile energikilder fører til økt tilstedeværelse av elektromekaniske pendlinger i kraftsystemet. Elektromekaniske modier som er dårlig dempet medfører stabilitetsproblemer, og kan i verste fall føre til splittelse av nettet. Det er derfor svært viktig å overvåke elektromekaniske moder som befinner seg i kraftsystemet. Tidssynkroniserte målinger fra PMUer gjør dette mulig. Målet med denne masteroppgaven er å legge grunnlaget for videre utvikling av pendlingsovervåking i det Nordiske kraftnettet.

Å lage et godt pendlingsovervåking-verktøy er en todelt oppgave. Først må man finne frem til den beste måten å visualisere pendlingene på i kontrollrommet. Siden elektromekaniske pendlinger er et ganske abstrakt fenomen er det viktig at den avgjørende informasjonen om dem kommer frem på en konkret og forståelig måte. En litteraturstudie om eksisterende visualiseringsmetoder og intervjuer med en operatør fra Statnett er gjennomført for å kartlegge gode visualiseringsmetoder.

I kontrollrommet trengs det både pendlingsovervåking for bedre systemoversikt i normale operasjonstilstander, og i kritiske situasjoner for å håndtere alvorlige svingninger. I normale tilstander er verktøy som viser systemets tilstand på en intuitiv måte viktige. Det kan være ved hjelp av fargekoordinert tekst, viserdiagram, GIS og spektrogram for kjente modier. I kritiske situasjoner kan GIS med visere, spektrogram, viserdiagram og detaljert tekst og grafer om de mest alvorlige modene hjelpe operatører med å ta gode valg.

Den andre siden av oppgaven handler om å utvikle gode algoritmer som gir korrekte og raske estimater av modal-parameterne. En litteraturstudie er gjort på dette området også, og det ble utviklet en algoritme basert på systemidentifikasjon gjennom N4SID og klynging av egenverdier gjennom DBSCAN. N4SID trenger tradisjonelt at man velger modell-orden. Med DBSCAN slipper man dette, og man får i tillegg fordelen av å mer nøyaktig estimere de sanne modene i systemet og reduserer estimater av moder fra støy.

Algoritmen er testet på simulert og ekte PMU data, og resultatene er lovende. Algoritmen finner relativt nøyaktig estimater av modene, og oppdager raskt endring av moder som følge av systemendring. Den finner også moder i det ekte PMU-datasettet som samsvarer med tidligere studier. Den største ulempen ved algoritmen er at den bruker ganske lang tid, og den sliter med å separere moder som ligger nært hverandre i frekvens og demping.

Acknowledgement

Firstly, I want to thank Kjetil Uhlen for his valuable feedback, illuminating discussions about the workings of power systems, and for encouragements along the way of writing this master thesis. The weekly meetings of discussion have been essential to the work progress. Kjetil's kind humour and brilliant knowledge of power system dynamics and control makes it fun to work on this topic.

I would like to thank Hallvar Haugdal for allowing me to use the power system simulator he has developed in Python, and for answering questions about the algorithms. Being able to see everything that goes on behind the interface of a simulator makes the simulating process much more interesting and it has been a good learning experience both with regards to programming and to power system analysis.

The team at Statnett, notably Jørn Johnsen and Turid Siebenbrunner, must be greatly thanked for facilitating an interesting summer internship and this master thesis. Being a part of the NEWEPS project has been educative, and the team meetings in Oslo have inspired new ideas for this master thesis.

A special thanks goes to my fellow students that I have been lucky to share work space with for making the process of writing the thesis so much more enjoyable.

Lastly, I would like to thank my family for their constant support in my endeavors.

Maiken Borud Omtveit

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Today, our lives are largely dependent on electricity, and every year increasingly more so. In order to combat climate change, Norway has committed to decarbonization policies through the Paris Agreement, which additionally contributes to an increased stress on the grid. According to a study by Statnett, electricity demand will increase by 40% in the Nordic region by 2040, and the majority of the additional demand will be covered by wind and solar power [11]. A rising number of DC interconnections with neighboring countries help cover the additional demand, but can be a source of instability in the grid. All these factors make it important that the power grid operators have good tools in the control room to ensure a safe and stable electricity supply.

Oscillations are present in any power system, and are initiated by the small changes in load that happen all the time [1]. This is not inherently a bad thing. If the oscillations are properly damped, they will decay with time. However, a change as small as 10MW may determine whether the oscillations will decay or whether they will increase in magnitude and potentially lead to system collapse. The latter was the case in a tie-line in western USA/Canada in 1996 s displayed in Fig. 1.1. The decaying oscillations initiated at 400s and 550s turned into growing oscillations due to faults and protective relay actions. Similar events occur in power systems all over the world.

Previously, oscillatory behavior was estimated by linearizing a detailed parametric model of the system [12]. The event in western USA/Canada in 19996 sparked an interest around estimating modal properties through measured data as opposed to simulating the nonlinear system. After this event, more attention has been paid to finding data driven methods of oscillation monitoring. The Nordic TSOs have previously listed Oscillation detection assessment as the highest priority of applications that need to be developed [13].

In the Nordic system, the frequency quality has degraded over the past years. Min-

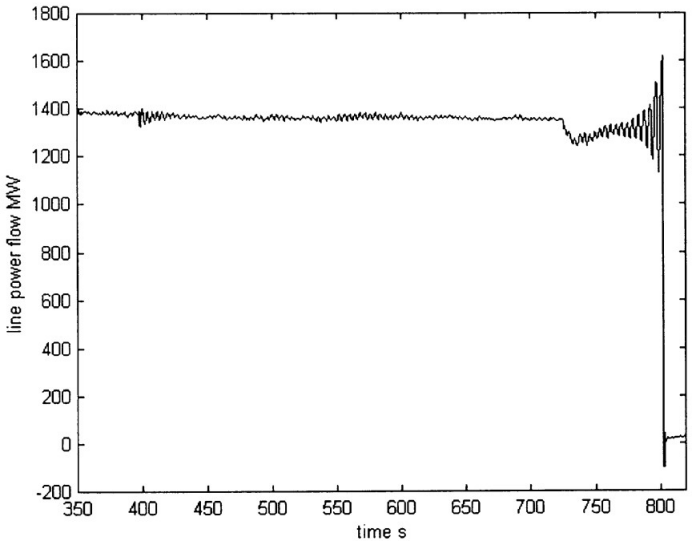


Figure 1.1: Collapse of a tie-line in western USA/Canada in 1996 due to oscillations [1].

utes outside the normal operating band are increasing, and there are more frequent oscillations [14]. The situation is worrisome for the TSO control room operators. To tackle the issue, Statnett and Svenska Kraftnät are cooperating with multiple research institutions and universities on a project called Nordic Early Warning Early Prevention System (NEWEPS). One of the goals of the NEWEPS project is to make a platform for visualizing oscillations in the system with data from phasor measurement units (PMUs). This specialization project aims to form the theoretical background to begin making such visualization tools, and develops one such tool.

The research topic lies in the intersection of data processing techniques, power system analysis, visualization methods and programming, so concepts from various disciplines need to be addressed. Chapter 2 presents relevant theory to the thesis, and begins with an introduction to PMU technology and how they are used in the control room today. Then, it focuses on relevant power system theory to describe the nature of oscillations with particular focus on the rotor angle stability as it is strongly connected to electromechanical oscillations. Linear analysis lays the foundation for most oscillation monitoring methods, and is therefore also explained.

The Theory chapter is followed by Chapter 3 that presents a literature review on visualization methods and algorithms for oscillation monitoring. The visualization methods and algorithms are categorized and discussed. The choice of the proposed algorithm is justified, and the N4SID algorithm is described in depth as it is the central part of the proposed algorithm. Chapter 4 presents the proposed algorithm

step by step. It consists of pre-processing and filtering, model estimation with N4SID, clustering and averaging of modes with DBSCAN, and then a calculation of the mode shapes which are plotted.

Following, in Chapter 5 the proposed algorithm is tested with simulated and real PMU data. The steps of the algorithm are thoroughly explained in the base case with ambient conditions. Then the algorithm performance is tested with respect to various input signals in Case 1. In Case 2, a significant change is made to the system topology which gives the system a different set of modes. The algorithms ability to detect new modes in the data set is evaluated and found to be satisfactory for control room purposes in this case. Case 3 gives the algorithm real data measured by PMUs in Norway during an event of sustained oscillations in the grid. The algorithm finds modes coherent with previous work on this data set, implying that the algorithm not only estimates correct modes and mode shapes on simulated data, but also on PMU data.

Chapter 6 discusses the results from the simulations and proposes areas for further work. Finally Chapter 7 gives concluding remarks about the findings of the thesis.

Chapter 2

Theoretical Background

In order to get a thorough understanding of the research questions of the thesis, it can be helpful to go through some central topics relevant to oscillation monitoring. The following sections will refresh or give an introduction to the topics of PMU technology, the root cause of electromechanical modes, and linear analysis. It is important to understand these concepts to see the usefulness and the legitimacy of the proposed algorithm in Chapter 4.

2.1 PMU Technology

PMU technology came around in the 1990s at Virginia Tech [15]. The PMU measures sinusoidal quantities in the power grid such as AC current. It measures the sinusoidal quantity

$$x(t) = X_m cos(\omega t + \theta) \tag{2.1}$$

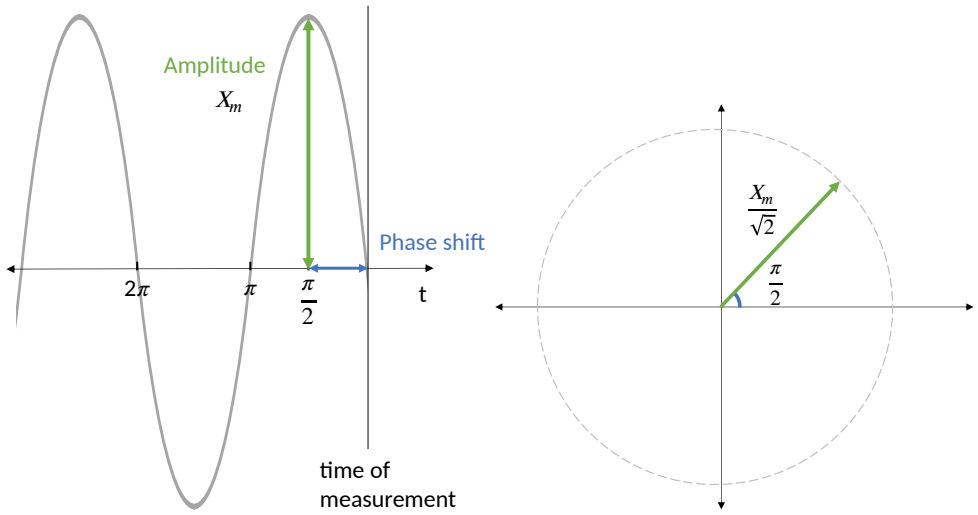
and estimates its phasor representation

$$X = \frac{X_m}{\sqrt{2}} e^{j\theta} = \frac{X_m}{\sqrt{2}} (cos\theta + jsin\theta) \tag{2.2}$$

where the magnitude is the root mean square (RMS) of the amplitude of the input signal, and the phase angle is the angle difference between the input signal at the measured instant and the peak of the input signal. The phasor representation is obtained through DFT of the measured signal

$$\bar{X} = X_r + jX_i = \frac{\sqrt{2}}{N} \sum_{n=1}^{N} \left(x_n cos\frac{2n\pi}{N} + jx_n sin\frac{2n\pi}{N} \right) \tag{2.3}$$

where X_r and X_i are the real and imaginary parts of the signal, x_n is the data sample, and N is the number of samples in one period.



(a) Sinusoidal signal of amplitude X_m and phase shift $\frac{\pi}{2}$

(b) Phasor representation of the signal. The magnitude is $\frac{X_m}{\sqrt{2}}$ and the phase shift is the same as the signal, $\frac{\pi}{2}$.

Figure 2.1

Each phasor gets a time stamp according to a common timing reference for the PMUs. The time stamp makes it possible to compare phasors from different locations and calculate the phase difference between them. This is very useful for oscillation monitoring, where the phase shift between the signals indicate which generators are oscillating against which.

PMU technology is increasingly used in the Nordic grid. There are PMUs installed in Norway, Finland, Sweden and Denmark and the information is shared between the countries. The PMUs measure frequency and voltage angles at sampling rates of 10 and 50 Hz. Statnett currently has a simple visualization tool that shows phasors in the control center. The PMUs have been used in a project to tune PSSs, for ramping down the power cable to Germany, and are occasionally used by the fault analysis team when the fault recordings are not enough [16], [17]. However, they are not in regularly for system overview. The NEWEPS project aims to integrate functionalities of oscillation and voltage stability monitoring based on PMUs in the control centers at Statnett and Svenska kraftnät.

2.2 Rotor Angle Stability

The stability of a power system is defined as its capability to regain equilibrium after a disturbance. Primarily, there are three quantities that are important for the

operation of a power system; the rotor angle, frequency and voltage. Stability issues can be classified to each of these categories as displayed in Fig. 2.9. Electromechanical oscillations are mainly related to rotor angle stability, but the effects manifest in the other quantities as well.

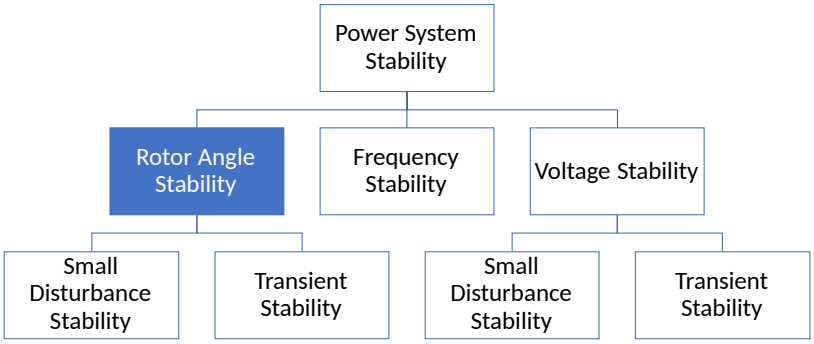


Figure 2.2: Classification of power system stability. Electromechanical oscillations are primarily related to rotor angle stability. Inspired by [2].

The rotor angle, also called power angle, is the angle of the rotor internal voltage with respect to the terminal generator voltage. They are both sinusoids, and the rotor angle δ is the phase shift between them. The internal voltage, also called electromotive force (emf), is expressed as a function of the terminal voltage and the internal reactance and generator current in the following way

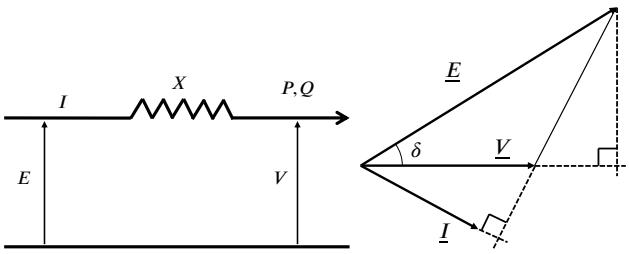
$$\underline{E} = \underline{V} + jX\underline{I} \tag{2.4}$$

Synchronous AC generators remain in synchronism after disturbances due to the self-regulating properties of their interconnection. When one machine deviates from nominal speed, the other generators compensate by transferring active power to the generator in order to reduce the difference in rotor speed. The inertia in the generators results in overshooting of the correction. This results in rotor swings which are the root cause of electromechanical oscillations [18], [1]. In order to understand the oscillatory generator response to a disturbance, it is beneficial to study the rotor dynamics.

The rotor dynamics are nonlinear and are described by the swing equation [2]

$$M\frac{d^2\delta}{dt^2} + D\frac{d\delta}{dt} = P_m - P_e \tag{2.5}$$

Where M is the inertia coefficient, D is the damping coefficient, P_m the mechanical power and P_e the electrical power, or the load. Equation 2.5 relates the rotor angle to the difference of mechanical and electrical power. By inspecting the equation it



(a) Equivalent circuit diagram of a simplified generator.

(b) Phasor diagram showing the rotor angle as the phase shift between the internal voltage E and terminal voltage V.

Figure 2.3: Simplified generator diagrams based on [2].

is easy to see that if there is more mechanical than electrical power, the rotor will accelerate, and vice versa.

Around the operating point, where there is an equilibrium of the powers, and assuming constant flux linkage, (2.5) can be linearized such that

$$M\frac{d^2\Delta\delta}{dt^2} + D\frac{d\Delta\delta}{dt} + K_{E'}\Delta\delta = 0 \tag{2.6}$$

where

$K_{E'} = \dfrac{\partial P_e}{\partial \delta}$ transient synchronizing power coefficient

$\Delta\delta = \delta'(t) - \hat{\delta}'_s$ change in transient rotor angle from the transient rotor angle defined in the operating point $\hat{\delta}'_s$

Equation 2.6 is a second order differential equation with the characteristic equation given by

$$\lambda^2 + \frac{D}{M}\lambda + \frac{K_{E'}}{M} = 0 \tag{2.7}$$

which has two roots,

$$\lambda_{1,2} = -\frac{D}{2M} \pm \sqrt{(\frac{D}{2M})^2 - \frac{K_{E'}}{M}} \tag{2.8}$$

A root can be either real or complex and inhibit the same qualities as an *eigenvalue*, which will be discussed in section **??**. Real roots signify an aperiodic response to a disturbance, whereas a complex root is oscillatory in nature.

For a complex root

$$\lambda = \alpha + j\omega \tag{2.9}$$

the damping factor is given by

$$\zeta = \frac{-\alpha}{\sqrt{\alpha^2 + \omega^2}} \tag{2.10}$$

ω is the frequency of the mode given in rad/s, so the frequency of mode i given in Hz as

$$f_i = \frac{\omega}{2\pi} \tag{2.11}$$

If the real part α of the root is positive, the damping is negative which means that the oscillations will grow if the mode is excited. Consequently, if the real part of the root is positive, the system is unstable.

The rotor angle stability can be described with the power-angle characteristic. The response to a disturbance may vary with the operating point as demonstrated in Fig. 2.4. It is desired that the rotor angle is less than the critical rotor angle in order to be stable. That means it has to be in the first quarter of the period. This allows the system to fall back into steady state. A generator operating in the second quarter would lose synchronism with the grid following a small disturbance.

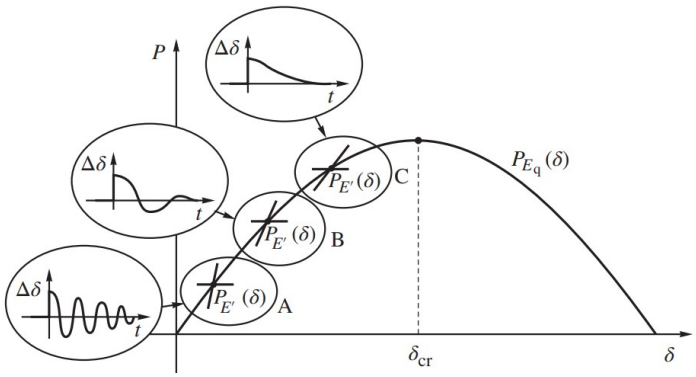


Figure 2.4: Rotor angle characteristic curve with power as a function of rotor angle. The figure shows that the damped responses depend on the operating point [2].

The relation between the phase angle and the frequency of the system becomes important in oscillation monitoring, as frequency might be an input to the monitoring algorithm. The relation is described by

$$\frac{d}{dt}\Delta\theta \propto \Delta F \tag{2.12}$$

In a phasor diagram, the frequency change would be seen as an acceleration of the phase angle. If the phasor is referenced to nominal frequency, 50 Hz in the Nordic synchronous area, the frequency change is seen as a change in phase angle with respect to the reference signal. This is visualized in Fig. 2.5.

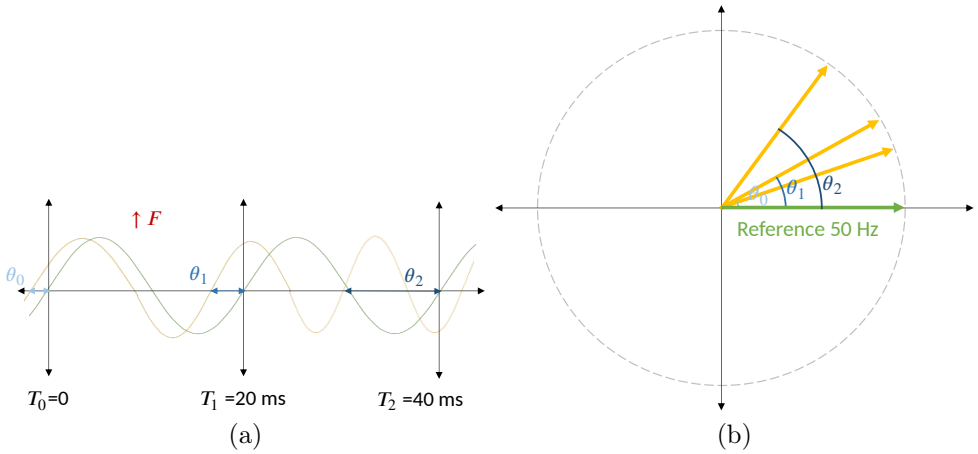


Figure 2.5: In T_0 the frequency of the signal is 50 Hz, and then the frequency increases such that the frequency is higher than 50 Hz in T_1 and even higher in T_2. This is seen in the phasor diagram by the change of phase angle.

2.3 Electromechanical Oscillations

There are various types of oscillations that can be observed in a power system. Normally, they are distinguished based on the nature of the input signal and the range of interaction with other generators in the system.

Oscillations that are observed may be ambient, transient or forced. Electromechanical modes are present in ambient and transient signals, also called free oscillations. Ambient conditions are produced by constant load changes in the system, whereas transient conditions are caused by large disturbances such as disconnecting a large load or generator. Forced oscillation refers to the systems response to a periodic external disturbance [19]. Precautions have to be taken to not confuse it with electromechanical modes. Forced oscillations may occur within the frequency range of electromechanical oscillations and may thus excite electromechanical modes [19]. System identification can be used to locate the source of forced oscillations, as presented in [20]. For an overview of methods of how to locate forced oscillations, see [21].

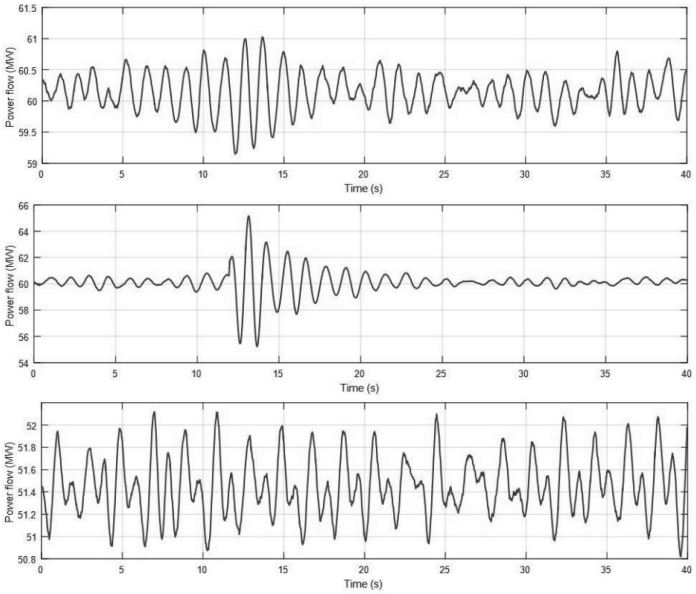


Figure 2.6: From top to bottom: ambient, transient and forced oscillations [1].

When classifying electromechanical oscillations based on location there are primarily three types of modes:

- Intraplant - generators on the same site oscillating against each other. The rest of the grid is unaffected.
- Local plant - single generation plant oscillating against the rest of the system. The oscillations are localized to the generator and the line connected to the grid. The rest of the grid is unaffected.
- Local area - generators within a relatively small area swinging against each other.
- Inter-area - two coherent groups of generators swinging against each other. The variation in tie-line power can be large.

The frequency of the different modes generally increase the closer the generators are to each other. The intraplant modes are at 2-3 Hz, and the inter-area modes are at 0.1-1 Hz [22]. The inter-area mode is the most troublesome, as it can lead to outages and islanding of parts of the grid [18]. The damping characteristic of an inter-area mode is determined by the tie-line strength, the nature of the loads, the power flow through the tie-lines, and the interaction between the loads and the dynamics of the generators and their controls [22]. Ususally, large interconnected

systems have two types of inter-area modes, a very low frequency mode in the order of 0.1-0.3 Hz that involves all the generators in the system split into two parts; and a low frequency mode in the order of 0.4-0.7 Hz that involves subgroups of generators swinging against each other [5].

2.3.1 Oscillation Parameters

Electromechanical oscillations can be described by the following parameters

- Frequency
- Damping
- Mode shape
- Excitation

The frequency of the oscillation depends on the system topology. The damping refers to how well damped the mode is. It is related to the system topology, but also control settings of components that provide damping in the system, such as capacitor banks and PSSs. The mode shape comprises information about the relative magnitude and the phase shift of the generators participating in the mode. The relative magnitude provides information about which generator is participating the most. The phase shift tells which generators are oscillating against which, or in other words the location of the oscillation. This tells us whether it's a single machine oscillating against the rest of the system, or a local area mode or an inter area mode. Mode shapes are usually displayed as phasors on a compass plot. The excitation refers to the absolute magnitude of the oscillations.

2.4 Linear Analysis

The power system is inherently nonlinear and stochastic by nature, making it challenging to analyze the output. However, the system behaves relatively linear around a specific operating point with small disturbances [23]. Consistent with power system theory, it can be linearized around an operating point and described by the state space form as

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} \tag{2.13}$$

$$\mathbf{y} = C\mathbf{x} + D\mathbf{u} \tag{2.14}$$

where $\mathbf{x}$ is the state variable vector, $\mathbf{u}$ is the input vector, $\mathbf{y}$ is the output vector, A is a real square matrix referred to as the state matrix, B is the input matrix, C

is the output matrix, and D is the feedforward matrix. The input $\mathbf{u}$ is normally conceptualized as noise from random load switching.,

The system can be discretized according to [24] as

$$\begin{aligned} \mathbf{x}(t+1) &= A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) &= C\mathbf{x}(t) + D\mathbf{u}(t) \end{aligned} \tag{2.15}$$

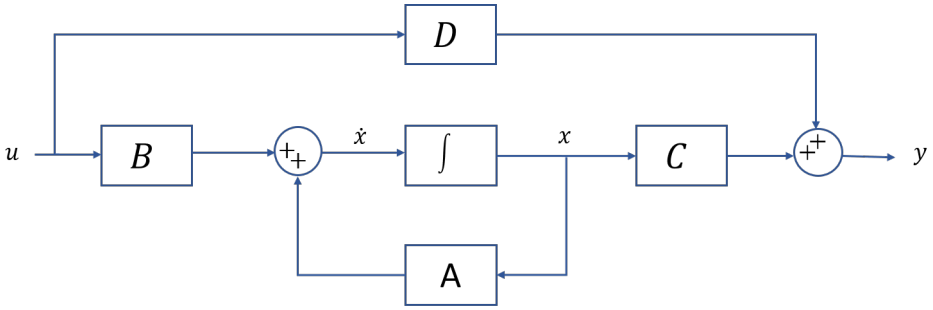


Figure 2.7: Block diagram of a general state space representation for a continuous time-invariant linear system.

The output in the time domain can then be expressed by the discretized system matrices as follows

$$\begin{aligned} y(t) &= Cx(t) + Du(t) \\ &= C(Ax(t-1) + Bu(t-1)) + Du(t) \\ &= CAx(t-1) + CBu(t-1) + Du(t) \\ &= CA^2x(t-2) + CABu(t-2) + CBu(t-1) + Du(t) \\ ... &= CA^tx(0) + Du(t) + \sum_{i=1}^{t} CA^{i-1}Bu(t-i) \end{aligned} \tag{2.16}$$

Where one sees that by expanding the recursive equations, the system response is calculated by use of the system matrices, the initial state vector, and the input vector at each time step.

For the algorithms implemented in this master thesis, it is useful to have an understanding of the relation between the transfer function and the state space model. A transfer function $G(s)$, relates the input to the output in the frequency domain such that

$$y(s) = G(s)u(s) \tag{2.17}$$

where $G(s)$ can be found by converting the system to the frequency domain through the Laplace transform. Applying the Laplace transform to equations **??** and 2.14 gives

$$sx(s) - x(0) = Ax(s) + Bu(s) \tag{2.18}$$
$$y(s) = Cx(s) + Du(s) \tag{2.19}$$

which implies

$$x(s) = (sI - A)^{-1}x(0) + (sI - A)^{-1}Bu(s) \tag{2.20}$$
$$y(s) = C(sI - A)^{-1}x(0) + C(sI - A)^{-1}Bu(s) + Du(s) \tag{2.21}$$

If the initial state vector $x(0)$ is zero, then 2.21 reduces to

$$y(s) = (C(sI - A)^{-1}B + D)u(s) \tag{2.22}$$

which comparing with 2.17 yields

$$G(s) = C(sI - A)^{-1} + D \tag{2.23}$$

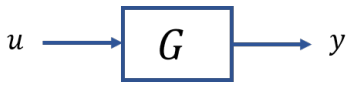


Figure 2.8: Block diagram of a general transfer function representation of a system.

2.4.1 Eigenvalues and Modes

The modes can be found by an eigenvalue analysis of the A matrix.

$$\det(\mathbf{A} - \lambda I) = 0 \tag{2.24}$$

The eigenvalues can then be found as

$$\lambda = \alpha + j\omega \tag{2.25}$$

Where ω is the frequency of the mode in radiance, and the frequency in Hz is given by

$$f_i = \frac{\omega}{2\pi} \tag{2.26}$$

The damping ratio is expressed as

$$\zeta = \frac{-\alpha}{\sqrt{\alpha^2 + \omega^2}} \tag{2.27}$$

Eigenvalues may be graphically represented in a plot such as in 2.10.

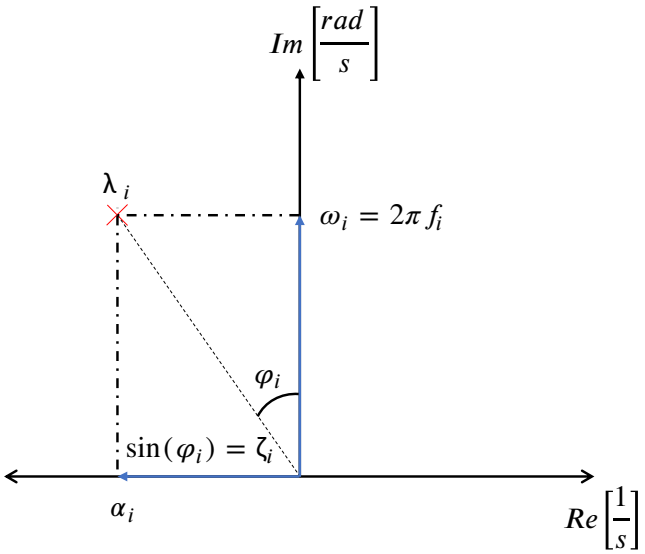


Figure 2.9: Eigenvalue $\lambda_i = \alpha_i + j\omega_i$ in an eigenvalue plot.

However, two modes of the same damping ratio may take different time to regain equilibrium. The modes of low frequency will sustain the oscillations for longer.

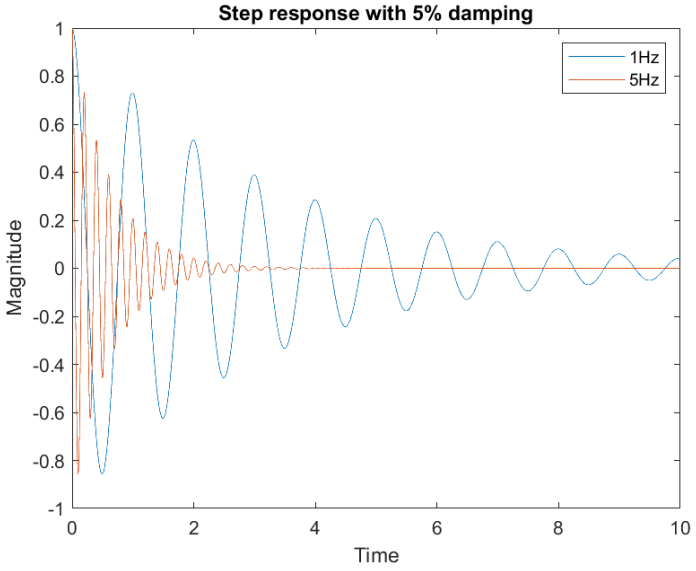


Figure 2.10: Observe that the lower frequency oscillation has longer settling time.

Linear analysis does not give information about whether the mode is excited or not, only which modes are latent in the system. Further, it is only possible to do the

eigendecomposition if you know all the parameters of the model, which one does not do in reality as there would have to be PMUs at every single node. Therefore, it can only be used for virtual models of the power system. However, the principle of the modes that are derived through linear analysis is the same when the modes are found through data analysis methods.

2.4.2 Eigenvectors and Mode Shapes

In power system analysis practice, a diagonalization of the state matrix in equation **??** is desired. If λ_i is an eigenvalue to $\mathbf{A}$, and $\mathbf{w}_i$ is a non-zero vector that fulfills the following relation

$$\mathbf{A}\mathbf{w}_i = \lambda_i \mathbf{w}_i \tag{2.28}$$

then $\mathbf{w}_i$ is the right eigenvector of $\mathbf{A}$ and is also known as the mode shape.

All the right eigenvectors and eigenvalues can be expressed on compact matrix form as

$$\mathbf{AW} = \mathbf{W}\boldsymbol{\Lambda} \tag{2.29}$$

where the right eigenvector matrix is given by

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 \dots & \mathbf{w}_n \end{bmatrix} \tag{2.30}$$

and the diagonal matrix of the eigenvalues is given by

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \tag{2.31}$$

Each column in the right eigenvector matrix corresponds to an eigenvalue λ_i, and the element $w_{j,i}$ is the mode shape of the corresponding state variable x_j. In general, a mode shape is a complex vector which determines the intensity level of participation in the oscillatory mode λ_i. The elements of a mode shape vector are relative to each other. This can be shown by multiplying a complex constant to both sides of (2.28).

$$\mathbf{A}(const * \mathbf{w}_i) = \lambda_i (const * \mathbf{w}_i) \tag{2.32}$$

Thus, normalization of a mode shape vector is convenient. A common way to normalize which is used in this thesis, is such that the largest mode shape equals 1 in magnitude and has 0 phase shift.

Pre-multiplying (2.29) with $\mathbf{W}^{-1}$ yields

$$\boldsymbol{\Lambda} = \mathbf{W}^{-1}\mathbf{AW} \tag{2.33}$$

Defining the left eigenvector matrix as

$$\mathbf{U} = \mathbf{W}^{-1} \tag{2.34}$$

The transformed state variables $\mathbf{x}$ can be transformed into modal variables $\mathbf{z}$

$$\mathbf{x} = \mathbf{W}\mathbf{z} \tag{2.35}$$

and the inverse transformation of (2.35) is given by

$$\mathbf{z} = \mathbf{U}\mathbf{x} \tag{2.36}$$

Substituting for $\mathbf{x}$ and in equation 2.13 with equation 2.35, and taking into consideration (2.31) gives

$$\dot{\mathbf{z}} = \mathbf{\Lambda}\mathbf{z} \tag{2.37}$$

which is called the modal form of the state space equation 2.13. The modal form is decoupled and thus considerably simpler than the original state space equation. Now each derivative of the state vector is only dependent on one modal variable, due to the diagonalized matrix where the off-diagonal elements are zero,

$$\dot{z}_i = \lambda_i z_i \tag{2.38}$$

which has the solution

$$z_i(t) = e^{\lambda_i t} z_{i0} \tag{2.39}$$

where z_{i0} is the initial value of modal variable i.

Through writing out the matrix multiplication in this equation, it is evident that each modal variable can be expressed as a linear combination of the state variables

$$z_i(t) = \sum_{i=1}^{n} u_{ij} x_j(t) \tag{2.40}$$

where u_{ij} is the (i, j) element of the left eigenvector matrix $\boldsymbol{U}$ made up by left eigenvectors. This shows that the left eigenvectors gives information about the controllability of individual modal variables by the state variables. A large coefficient u_{ij} means that a change in the state variable $x_j(t)$ has a big impact on the modal variable $z_j(t)$.

Likewise, each state variable can be expressed as a linear combination of the modal variables

$$x_k(t) = \sum_{i=1}^{n} w_{ki} z_i(t) \tag{2.41}$$

where w_{ki} is the (k, i) element of the right eigenvector matrix $\boldsymbol{W}$. This shows how well each modal variable is observed in each state variable. A large w_{ki} means that the modal variable $z_i(t)$ is observable in state variable x_k. In a graph one would then be able to see the oscillation of the mode of a significant magnitude if excited. The right eigenvector matrix thus describes the observability of the modes in each state variable.

If the eigenvectors are normalized, w_{kj} determines the magnitude and phase of the modal variables that participate in the activity of state variable $x_k(t)$. This is called the *mode shape*. Normally the eigenvectors are normalized with respect to the complex value with the largest magnitude in the right eigenvector.

Note that the mode shape describes the latent responses of the state variables in a linearized system, but does not describe whether the mode is excited or not. If a mode with a small coefficient is excited, it will not have a large impact on the state variable. The interesting modes are thus the ones with large controllability and observability coefficients.

Finally, remark that it may be easier to measure the output than the state variable in a power system. By substituting equation 2.14 with the expression for x in equation 2.35, one obtains

$$\dot{\mathbf{z}} = \mathbf{\Lambda z} + \mathbf{U u} \tag{2.42}$$

$$\mathbf{y} = \mathbf{CWz} + \mathbf{Du} \tag{2.43}$$

where the mode shapes can be found in the matrix

$$\mathbf{M} = \mathbf{CW} \tag{2.44}$$

This final result is particularly important in subspace methods.

Chapter 3

Literature Review on Oscillation Monitoring Visualization Methods and Algorithms

3.1 Visualization Methods and GUIs

In order to know what a good monitoring tool looks like, it is important to think about who will use it. In the future, there might be FACTS devices that ensure that all responses to changes in the system are well damped based on real-time oscillation monitoring. However, the first step is to monitor the oscillations such that the operators have a good system overview. Therefore, the operator has to be considered as the end user. The operators are constantly looking at information about the power system and only need information that they can act upon. An operator at Statnett reported that there are already too many alarms going off for the operators to be able to look at all of them [25]. Therefore the oscillation monitoring must be concise with only the most necessary details to not overload the operator with information. Clean design and appropriate color coordination are effective features of a good visualization tool.

A study of existing visualization tools was conducted in the specialization project leading to this master thesis. The various visualization methods can be grouped as displayed in Fig. 15. Larger images of the oscillation tools presented in the figure as well as sources to each image can be found in Appendix . Oscillation monitoring in control rooms often comprise multiple visualization methods, as it is hard to portray all of the modal parameters in one visualization tool. There is usually a compromise between giving a system overview and detailed information to base decisions on.

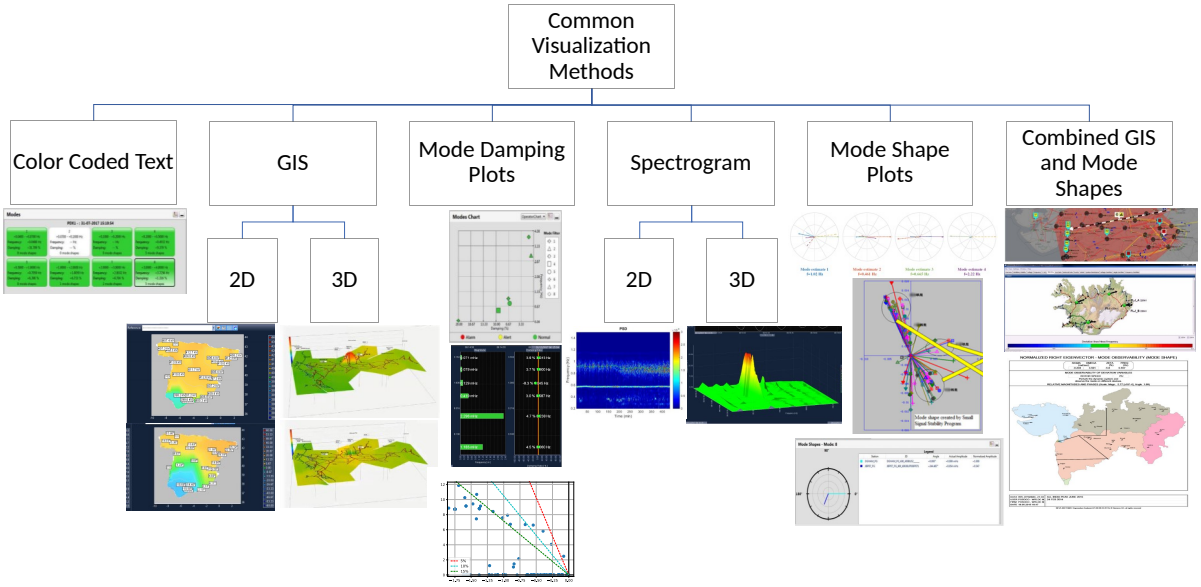


Figure 3.1: Categorization of visualization methods. Larger images of each method are found in Appendix A.

Color Coded text

Color coded text gives a quick system overview. It can be good as an initial overview, and when there are critical modes, more information could be displayed with other methods. The overview with color coded text used in Fig. 15 shows the eigenvalue of the mode, frequency, damping and number of participating mode shapes.

GIS

Geographic information system (GIS) is a very common tool used in control centres. On maps, colors are used to signify high and low values. In two dimensions, as used by the Spanish TSO [6], one would be able to see the colours pulsating when there is oscillation present in the system. In three dimensions, as developed by the Russian TSO [7], the spatial level of the map is also a function of the magnitude. A large magnitude in a point results in a peak in that point. This makes it easy to see how the oscillations propagate in the map, and can give an intuitive understanding of oscillations. GIS gives some information about the location and magnitude of the oscillations, but it does not do so very accurately. Therefore it is convenient to use for system overview, but not as decision basis for control actions. The operator at Statnett interviewed for this thesis, Birger Bjørland, suggested the 3D GIS visualization method for operator training [25]. Seeing oscillations visualized in this way can

give the operator more knowledge about the nature of oscillations, and thus make them more prepared for other visualization tools.

Mode Damping Plots

It is important to be aware of poorly damped modes. As shown in Fig. 15, there are multiple ways of visualizing the damping, for example by plotting relative damping against frequency of modes, as two connected bar charts for magnitude of excitation and damping, or in an eigenvalue plot. The first plot gives a good overview of which modes are poorly damped. The second plot gives information about the excitation of the mode in addition to the damping. In order to predict how the power system will develop it is useful to know the excitation and damping of the modes together. Therefore having a combined visualization of magnitude of excitation and damping of each mode, as in the second mode damping plot in the figure, can be beneficial. The eigenvalue plot displays the mode in the complex plane, and the frequency and damping is hence only displayed implicitly. The relative damping lines are usually displayed as a guideline to keep track of poorly damped modes. It is intuitive if the operators already know which modes to expect from the plot. However, as it only displays the frequency and damping implicitly, other mode damping plots are favoured.

Spectrogram

The spectrogram shows the magnitude of excitation of the modes in a data stream. It has time on one axis, and frequency on another. In a two-dimensional spectrogram, the color shows the magnitude of excitation. A warmer color means a higher excitation of the mode. Similarly, in three dimensions, the third dimension and the color shows the magnitude of excitation. The three-dimensional spectrogram might be a better real-time tool as the third dimension makes it easier to see the immediate change in magnitude. However, color is not a very accurate measure, and the spectrogram is only good to see the tendencies and to see whether a mode is excited or not. The height of the three-dimensional spectrogram might be difficult to read. It is therefore necessary with complementary tools to get an accurate picture of the excitation. However, the third dimension does make it a very intuitive visualization of the development of the magnitude. It may be a very useful tool for intuitively understanding the magnitude of excited modes. A spectrogram can only display data from one data stream at a time. Hence, if several modes involving different generators are excited, several spectrograms are needed to get the full picture of the oscillations present in the system.

Mode Shape Plots

A mode shape plot is a graphical representation of mode shapes where each arrow from a circle origin represents a PMU signal. It is an informative way of displaying the modal parameters frequency, damping, and location. One plot is made for each mode in the system, where the length of the arrow represents the participation of the component in the mode calculated from the right eigenvector. The angular separation between the arrows in a plot is the phase shift between the signals. If there are two clusters of arrows with a sufficiently large phase angle in between, it means that there are two groups of generators oscillating against each other with a phase shift. A compass plot is a particular mode shape plot where the mode shapes are normalized and displayed on a circle. Compass plots highlight the degree to which each generator participates in the mode. It is common to include the frequency, damping and excitation in the mode shape plot, such as in Fig. 3.2.

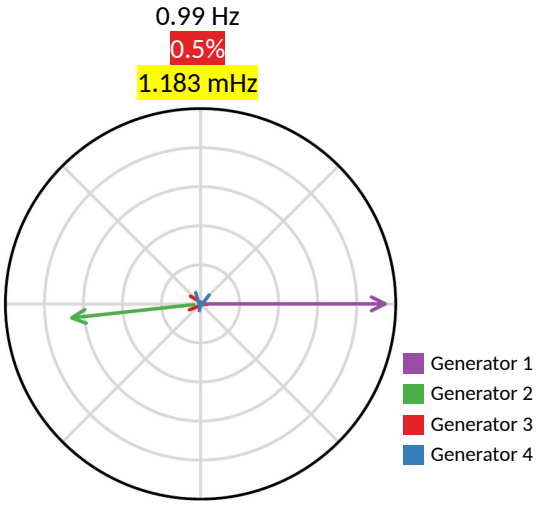


Figure 3.2: Arbitrary example of a compass plot of an inter-area mode with its frequency, damping and excitation displayed as text with colour coded highlights. Damping and amplitude of oscillation is highlighted ellow and red to display their criticality.

Combined GIS and mode shapes

Combining GIS and mode shapes is one of the most common ways to monitor oscillations in the control room of TSOs. One of the reasons is likely that it displays where the oscillations are present very intuitively. Remark that it can only show one mode at a time, so it does not give a complete overview of the oscillations present in the

grid. For that, multiple maps would be needed as it would probably be chaotic to have multiple phasors in the same GIS. It gives information about the most critical modes where measures should be taken. The Indian TSO POSOCO has a layer in the GIS that displays where the islanding is most likely to occur. This can help the operators decide which action fits the situation the best. The issue of islanding is outside the scope of this thesis, but still relevant to oscillation monitoring. There are many variants of combining GIS and mode shapes, but the most important takeaway from comparing this method is that it must be made intuitive and actionable for the operator. This can be done by making the phasors large and color coordinating the map to the frequency level, which will make the power oscillations observable.

3.1.1 Feedback from Operator

An interview with an operator at the control central was conducted in order to get feedback on the visualization tools found. The two visualization methods that use three dimensions, the GIS and the spectrogram, were highlighted as particularly interesting. As the experience with oscillations is rather low, the operator highlighted the need for visualization methods that are easy to interpret. That might be the reason for the operator's positive response to using 3D. No other screen in the control room uses 3D yet, and an important principle in control room ergonomics is to keep the visualization simple and informative. Having an extra dimension might confuse the operator in critical situations where they only need the most important information. However, it can be argued that visualization in 3D are most similar to the real world we live in, and therefore the brain understands it better. The operator suggests that it is a good idea to use 3D visualization displays for training the operators in the nature of oscillations, but for use in the control room more pragmatic approaches may be favorable.

As there is not much experience with oscillation monitoring in the Norwegian central control room today, the operator did not have specific preferences regarding the other visualization methods. It is worth noting though, that in the control center today there are only line diagrams and no maps. The advantage of this is that it gives a simpler overview of the system in terms of electrical parameters. This might suggest that using line diagram in stead of maps in the GIS visualization method could be easier to interpret for the operators.

When there are critical oscillations, the operator should over time learn how to respond to the oscillations. Normally the only advised thing to do is to lower the power transfer in the lines. FACTS devices can provide damping for oscillations, but that will happen automatically, and is outside the operators' control. When designing tools for the operators in the control room, it is important to have in mind how the tool is intended to be used. It is important that they know when there might be a collapse

due to large oscillations in order to for example reduce power transfer. Therefore, all the modal parameters should be displayed, such that the operator knows where the oscillations are occurring, at which frequency and with what damping.

Visualization Method	Mode Frequency	Excitation Magnitude	Damping	Mode Shape
GIS	x	x		
Mode shape plot	x		x	x
Spectrogram	x	x		
Damping plot	x		x	
Combined GIS and phasors	x	x		x
Color coded text	x	x	x	
Eigenvalue plot	x		x	

Table 3.1: Modal parameters displayed by the identified visualization method categories.

3.1.2 Choice of Visualization Method

Based on the operator's positive response to 3D, the 3D spectrogram was implemented in the specialization project. It is easy to adjust the tool to be in two dimensions, as the same spectral density algorithm can be used for both 2D and 3D, and the only difference is the plotting. No visualization method displays all the parameters, as shown in Table 3.1. In order to make a visualization tool that displays all the oscillation parameters it is necessary to combine multiple methods together. A method that complements the spectrogram made in the specialization project well is the mode shape plot. It gives information about the mode shapes, frequency and damping of the modes. As the spectrogram gives the excitation of the modes, they complement each other well and together display all the modal parameters. Therefore it was decided to make the mode shape plot in this thesis. Remark that as the algorithms behind the two methods together find all modal parameters of the system. Hence, with these two complementary algorithms, any visualization method can be implemented by changing the way it is plotted.

3.2 Oscillation Monitoring Algorithms

It is not enough to have good visualization tools if the information displayed is not accurate. In order for the tools to be good, good algorithms that process the PMU data and obtain the right mode parameters are needed. Algorithms that detect oscillations in signals is interesting in many applications outside power system analysis. Seismology, construction, communication and neuroscience are some examples the author has come across while learning about oscillation detection. This has influenced the development of the field of oscillation monitoring algorithms. There are a myriad of algorithms that can be used for electromechanical oscillation monitoring, and they can be separated on many bases. One proposal for how to categorize the algorithms is displayed in Fig 3.3.

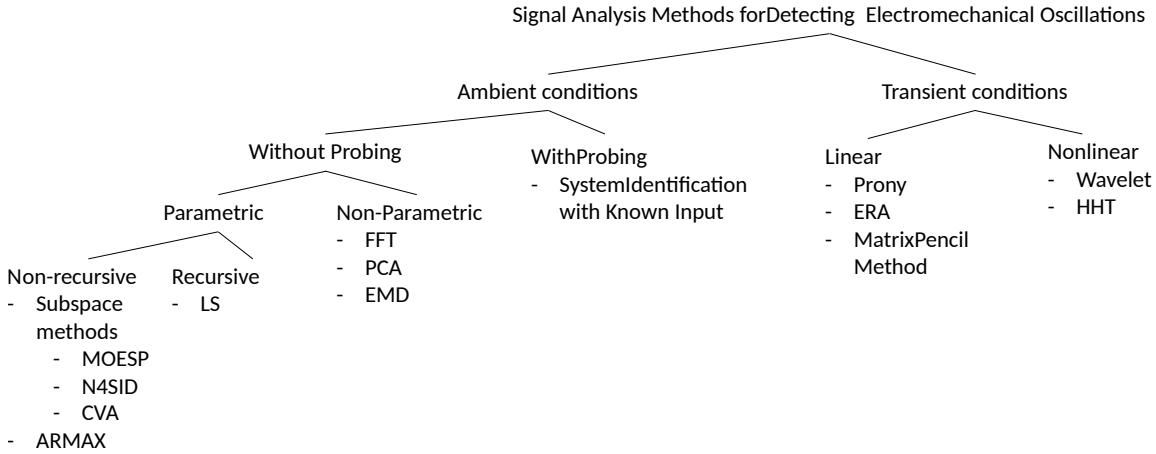


Figure 3.3: Categorization of data based oscillation monitoring algorithms based on [3].

Firstly, they are separated based on whether they work on ambient or transient conditions. Transient operation, as previously discussed in section 2.3 refers to the power system response to faults or large disturbances. It is assumed that the transient of the disturbance represents the true impulse response of the system. Transient analysis thus measures the stability of the system by determining the frequency and damping of the transient. Conversely, signals during ambient conditions are stochastic and dominated by broadband noise due to the load demand. The goal of ambient analysis is to track the damping and excitation of the modes present in the system.

The ambient algorithms can be separated on whether they use a known input, also called probing. Normally, this is not the case in the power system, as the loads are stochastic and unmeasured. Probing can be done by injecting an external disturbance through a large load or interconnection, and measure the output. Such methods are

for example discussed in [26].

The algorithms that do not use probing are divided into parametric and non-parametric methods. Parametric methods assume a model behind the oscillating behaviour, and seek to parameterize the model. These methods usually assume a functional form for the probability distribution functions of the signal. The input is usually assumed to be a Gaussian random load process [3].

Parametric methods can be distinguished as recursive or non-recursive methods. Recursive methods are methods with memory of past inputs, which with time converge towards a parameterized solution. A forgetting factor is deployed to discount previous data. Non-recursive methods, however, calculate the parameters from scratch for every new set of data. Another way to classify ambient condition algorithms is to distinguish between output-to-output and subspace methods, as done in [27]. For an overview of the algorithms in Fig. 3.3, and references to the original publications, see [28], [3].

3.2.1 Choice of Algorithm

As power systems are increasingly pushed to their stability limits, modes tend to be lower damped and can become unstable [3]. It is thus important to monitor the oscillations in real-time ambient conditions as well as during transient conditions. Monitoring the mode shape, in addition to frequency, damping and excitation, provides information about the source of the oscillations, enabling the operators to make better control decisions. A good way to ensure oscillation monitoring whatever the circumstances is to combine an ambient and a transient algorithm. It was decided to choose an algorithm that works on ambient conditions for the purpose of steady state monitoring in this thesis.

As probing conditions are rare, it is desired that the method is able to run without probing. Recursive algorithms require several minutes to converge towards a steady-state solution [28]. In a control room setting where a short time delay is crucial, it is important that the information comes in near real-time. Therefore, other algorithms are preferred for real-time monitoring.

In [28], N4SID was compared with FDD, Channel Matching Method, multchannel spectral method and an ARX transfer function method with respect to the accuracy of their mode shapes. N4SID proved to have a smaller bias when estimating mode shapes of high- and low-damped modes during ambient conditions compared to the other algorithms. It also proved to have less bias when there was a ringdown in the middle of the ambient data. It has the ability to take input data, in cases of probing, which gives a larger signal to noise ratio, and hence decrease standard deviation.

[29] compared the damping estimation of the wavelet method, N4SID and a method based on spectral independent component analysis and random decrement

called RD-ICA. RD-ICA and N4SID has a lower variance when the excitation of the mode is low, compared to the wavelet method. As the wavelet method is nonlinear it is more sensitive to averaging the estimate over a time window to obtain a single damping estimate, which is why it is categorized as a transient conditions method in Fig. 15. N4SID is expected to perform better than the other two methods in systems of non-Gaussian noise, as it does not rely on a Gaussian assumption, which is an advantage when using PMU data from the system.

The drawback of N4SID is that it has longer computational time than for example transfer function based methods, and there is a trade-off between bias and variance. Furthermore, N4SID has quite high standard deviations compared to other algorithms, though it should be noted that [28] suggests that the bias is a much larger contributor to error overall than standard deviation.

N4SID is deemed an interesting algorithm for the purposes of this thesis, and it is chosen as algorithm for the proposed method. The following sections will describe the algorithm in detail.

3.3 Subspace Method for System Identification

The subspace method for system identification was first presented by B. Gopinath at Bell Laboratories in 1969 [30]. The subspace identification problem is the following. Given a set of input-output measurements $u = [u_1, ..., u_m]$ and $y = [y_1, ..., y_p]$, determine the order of the unknown system and the system matrices (A, B, C, D).

The most common way to formulate the general subspace methods is to begin with considering a multivariate LTI system. For $t = 0, 1, ..., N-1$, the system is described by

$$x(t+1) = Ax(t) + Bu(t) + \eta(t) \tag{3.1}$$

$$y(t) = Cx(t) + Du(t) + \nu(t) \tag{3.2}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $y \in \mathbb{R}^p$. The residuals are $\eta \in \mathbb{R}^n$ and $\nu \in \mathbb{R}^m$ and describe the inaccuracy and noise in the model[1].

The state space model on matrix form is then

[1]Recall that for a system $Ax = b$, the residual, r, is [31]

$$r = Ax - b \tag{3.3}$$

Evidently, $r = 0$ only if x is a solution. For approximated solutions $r \neq 0$. Hence, η and ν represent the deviation of the model from the true system. This deviation is caused by both inaccurate modelling and noise in signal.

$$\begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} + \begin{bmatrix} \eta(t) \\ \nu(t) \end{bmatrix} \tag{3.4}$$

and the approximated model is given by

$$\underbrace{\begin{bmatrix} \bar{x}(t+1) \\ y(t) \end{bmatrix}}_{Y(t)} = \underbrace{\begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix}}_{\Theta} \underbrace{\begin{bmatrix} \bar{x}(t) \\ u(t) \end{bmatrix}}_{\Phi(t)} \tag{3.5}$$

which can be viewed as a linear regressor where Θ is treated as the parameter matrix, and Φ as the regressor by t. The equation can be solved as a standard LSE problem

$$\Theta = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} \tag{3.6}$$

$$= \left(\sum_{t=0}^{N-1} \begin{bmatrix} \bar{x}(t+1) \\ y(t) \end{bmatrix} \begin{bmatrix} \bar{x}^T(t) & u^T(t) \end{bmatrix} \right) \left(\sum_{t=0}^{N-1} \begin{bmatrix} \bar{x}(t) \\ y(t) \end{bmatrix} \begin{bmatrix} \bar{x}^T(t) & u^T(t) \end{bmatrix} \right)^{-1} \tag{3.7}$$

which exists on the condition that

$$rank \begin{bmatrix} \bar{x}(0) & \bar{x}(1) & \dots & \bar{x}(N-1) \\ u(0) & u(1) & \dots & u(N-1) \end{bmatrix} = n + m$$

is satisfied. In the equation for the system parameters, the only unknown is the state vector.

Moreover, the residuals η and ν are zero mean, and their covariance can be obtained by

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} = \frac{1}{N} \sum_{t=0}^{N-1} \begin{bmatrix} \eta(t) \\ \nu(t) \end{bmatrix} \begin{bmatrix} \eta^T(t) & \eta^T(t) \end{bmatrix} \tag{3.8}$$

Now, the key question is how to obtain the state variables, $x(t)$. This problem is called the System Realization Problem.

3.3.1 Realization of LTI Systems

Obtaining a state space representation from a given transfer function matrix is called the system realization problem. There are multiple sets of state and measurement equations that produce the same transfer function and thus the same output. What is the most interesting is the state space representation that uses the lowest dimension of state variables. This is called *minimal* realization. A summary of background knowledge on linear system theory which is useful for studying system realization is

found in Appendix C. The simplest form of system realization called the Ho-Kalman method that uses the impulse response as data, is described in Appendix G.

The input and output can be written in the following Hankel matrices

$$U_{0|k-1} = \begin{bmatrix} u(0) & u(1) & \dots & u(N-1) \\ u(1) & u(2) & & u(N) \\ \vdots & & \ddots & \vdots \\ u(k-1) & u(k) & \dots & u(k+N-2) \end{bmatrix} \tag{3.9}$$

$$Y_{0|k-1} = \begin{bmatrix} y(0) & y(1) & \dots & y(N-1) \\ y(1) & y(2) & & y(N) \\ \vdots & & \ddots & \vdots \\ y(k-1) & y(k) & \dots & y(k+N-2) \end{bmatrix} \tag{3.10}$$

Let $W_{0|k-1}$ be the Hankel data matrix that concatenates the input and output matrices such that

$$W_{0|k-1} = \begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} \tag{3.11}$$

From the state equations, the expression for the output after an arbitrary point in time t can be found by writing out the expression for each time step

$$\begin{aligned} y(t) &= Cx(t) + Du(t) \\ y(t+1) &= CAx(t) + CBu(t) + Du(t+1) \\ y(t+2) &= CA^2x(t) + CABu(t) + CBu(t+1) + Du(t+2) \\ &\vdots \\ y(t+k-1) &= CA^{k-1}x(t) + \sum_{i=t+1}^{t+k-1} CA^{i-1}Bu(t+k-1-i) + Du(t+k-1) \end{aligned} \tag{3.12}$$

This expression can be rewritten on matrix form as

$$\underbrace{\begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+k-1) \end{bmatrix}}_{y_k(t)\in\mathbb{R}^{pk\times 1}} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}}_{\mathcal{O}_k\in\mathbb{R}^{pk\times n}} x(t) + \underbrace{\begin{bmatrix} D & 0 & \dots & 0 \\ CB & D & & \\ \vdots & & \ddots & 0 \\ CA^{k-2}B & \dots & CB & D \end{bmatrix}}_{\Psi_k\in\mathbb{R}^{pk\times mk}} \underbrace{\begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+k-1) \end{bmatrix}}_{u_k(t)\in\mathbb{R}^{mk\times 1}} \tag{3.13}$$

where $\mathcal{O}_k$ is the observability matrix, and Ψ is is a block Toeplitz matrix derived from equation 3.12. This input-output relationship can be expanded to matrix form using the data matrices as follows

$$Y_{0|k-1} = \mathcal{O}_k X_0 + \Psi_k U_{0|k-1} \tag{3.14}$$

where the rank$X_0 = n$, rank$U_{0|k-1} = mk$ and rank$\begin{bmatrix} U_{0|k-1} \\ X_0 \end{bmatrix} = mk + n$. Rearranging this expression one can obtain the expression of the Hankel data matrix as follows

$$\begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} = \begin{bmatrix} I_{km} & 0_{km\times n} \\ \Psi_k & \mathcal{O}_k \end{bmatrix} \begin{bmatrix} U_{0|k-1} \\ X_0 \end{bmatrix} \tag{3.15}$$

In the simple Ho-Kalman method one relied on the impulse response of a system to find the system matrices through decomposition. A zero-input response can also be created from an arbitrary input-output data set through manipulation of the Hankel data matrix. Any input-output pair of length k that satisfies $y_k(0) = \mathcal{O}_k x(0) + \Psi_k u_k(0)$ can be expressed as a linear combination of the column vectors in the data matrix. The coefficients in the linear combination is given by a vector $\zeta \in \mathbb{R}^{N\times 1}$ such that

$$\begin{bmatrix} u_k(0) \\ y_k(0) \end{bmatrix} = \begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} \zeta \tag{3.16}$$

A zero-input response (zip) can be created by a linear combination of the column vectors of the data matrix. $\exists \zeta \in \mathbb{R}^{N\times 1}$ such that

$$\underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \tilde{y}(0) \\ \vdots \\ \tilde{y}(k-1) \end{bmatrix}}_{zip} = \begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} \zeta \tag{3.17}$$

By repeating this process, one can form a Hankel matrix by concatenating a sufficient amount of zero-input response vectors. The transformation of the data matrix $W_{0|k-1}$ to the zero-input response can be achieved through LQ-decomposition,

$$\begin{bmatrix} U_{0|k-1} \\ Y_{0|k-1} \end{bmatrix} \underbrace{\begin{bmatrix} \zeta_1 & zeta_2 & \dots & \zeta_{(m+p)k} \end{bmatrix}}_{Q} = \underbrace{\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}}_{L} \tag{3.18}$$

where the L matrix is a lower triangular matrix, which is the form of the zero input response. From this decomposition, the system matrices can be found with various methods. The two most well-known algorithms for this is MOESP and N4SID. The N4SID method will be explained below as that is the one used in this thesis.

N4SID

The N4SID algorithm was developed by Van Overschee and De Moor in 1992 [32]. The way N4SID differs from MOESP is primarily the form of the data matrix, and it applies oblique projection as opposed to MOESP's orthogonal projection. For N4SID, a particular Hankel matrix is formed of the output data, divided into past and future sub matrices, see [33]. The input and output is written in the following Hankel matrices where the subscripts denotes the past and future matrices,

$$U_{0|2i-1} = \begin{pmatrix} U_p \\ \hline U_f \end{pmatrix} = \left[\begin{array}{cccc} u(0) & u(1) & \dots & u(N-1) \\ u(1) & u(2) & & u(N) \\ \vdots & & \ddots & \vdots \\ u(k-1) & u(k) & \dots & u(k+N-2) \\ \hline u(k) & u(k+1) & \dots & u(k+N-1) \\ u(k+1) & u(k+2) & & u(k+N) \\ \vdots & & \ddots & \vdots \\ u(2k-1) & u(2k) & \dots & u(k+N-1) \end{array}\right] \tag{3.19}$$

$$Y_{0|2i-1} = \begin{pmatrix} Y_p \\ \hline Y_f \end{pmatrix} = \left[\begin{array}{cccc} y(0) & y(1) & \dots & y(N-1) \\ y(1) & y(2) & & y(N) \\ \vdots & & \ddots & \vdots \\ y(k-1) & y(k) & \dots & y(k+N-2) \\ \hline y(k) & y(k+1) & \dots & y(k+N-1) \\ y(k+1) & y(k+2) & & y(k+N) \\ \vdots & & \ddots & \vdots \\ y(2k-1) & y(2k) & \dots & y(k+N-1) \end{array}\right] \tag{3.20}$$

Writing the past data matrix as

$$W_p = \begin{bmatrix} U_p \\ Y_p \end{bmatrix} \tag{3.21}$$

the QR decomposition can be applied such that

$$\begin{bmatrix} U_f \\ W_p \\ Y_f \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & 0 \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix} \tag{3.22}$$

where the bottom right element is zero because past and future inputs are zero, and past outputs are zero, see [34]. From this, three equations can be derived

$$Y_f = R_{31}Q_1^T + R_{32}Q_2^T \tag{3.23}$$
$$Q_1^T = R_{11}^{-1}U_f \tag{3.24}$$
$$Q_2^T = R_{22}^{\#}(W_p - R_{21}Q_1^T \tag{3.25}$$

where $R_{22}^{\#}$ is the pseudo inverse of R_{22}, which is necessary because R_{22} is rank deficit. Substituting equation 3.24 and 3.25 into 3.23 gives

$$Y_f = (R_{31} - R_{32}R_{22}^{\#}R_{21})R_{11}^{-1}U_f + R_{32}R_{22}^{\#}W_p \tag{3.26}$$

Comparing equation 3.26 with the following equation

$$Y_f = \Psi_k U_f + \mathcal{O}_k X_f \tag{3.27}$$

and examining the subspace properties one can first observe that the first term in 3.26 and 3.27 relates to the same subspace. Secondly, as both equation 3.27 and 3.26 are direct sums, meaning that there is no overlap in bases, the second terms in both equations must be the same;

$$\mathcal{O}_k X_f = R_{32}R_{22}^{\#}W_p \tag{3.28}$$

Taking the singular value decomposition of the right hand side yields

$$R_{32}R_{22}^{\#}W_p = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1\Sigma_1 V_1^T \tag{3.29}$$

which can be split into

$$X_f = T^{-1}\Sigma_1^{1/2}V_1^T \tag{3.30}$$
$$\mathcal{O}_k = U_1\Sigma_1^{1/2}T \tag{3.31}$$

where the system parameters (A, B, C, D) can be found either through decomposition of X_f or $\mathcal{O}_k$. Here, the former is chosen as example. Begin by observing that X_f contains a series of states,

$$X_f = \begin{bmatrix} x(k) \\ x(k+1) \\ \vdots \\ x(k+N-1) \end{bmatrix} \tag{3.32}$$

which, together with equations 3.19 and 3.20 enables the forming of the following four matrices,

$$\overline{X}_k = \begin{bmatrix} x(k) & \dots & x(k+N-2) \end{bmatrix} \in \mathbb{R}^{n\times(N-1)} \tag{3.33}$$

$$\overline{X}_{k+1} = \begin{bmatrix} x(k+1) & \dots & x(k+N-1) \end{bmatrix} \in \mathbb{R}^{n\times(N-1)} \tag{3.34}$$

$$\overline{U}_{k|k} = \begin{bmatrix} u(k) & \dots & u(k+N-2) \end{bmatrix} \in \mathbb{R}^{m\times(N-1)} \tag{3.35}$$

$$\overline{Y}_{k|k} = \begin{bmatrix} y(k) & \dots & y(k+N-2) \end{bmatrix} \in \mathbb{R}^{p\times(N-1)} \tag{3.36}$$

These matrices are related in the state equations such that

$$\begin{bmatrix} \overline{X}_{k+1} \\ \overline{Y}_{k|k} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix} \tag{3.37}$$

which has the unique solution

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \overline{X}_{k+1} \\ \overline{Y}_{k|k} \end{bmatrix} \begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix}^2 \tag{3.38}$$

where $(.)^2$ denotes the Moore-Penrose pseudo inverse given by

$$\begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix}^2 = \begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix}^T \left(\begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix} \begin{bmatrix} \overline{X}_k \\ \overline{U}_{k|k} \end{bmatrix}^T \right)^{-1} \tag{3.39}$$

and thus all system matrices are found. Of course, the ones that are relevant for this master thesis are the A and C matrices. The A matrix is needed to find the eigenvalues and eigenvectors, and the C matrix is needed for the mode shapes.

The algorithm can be summarized as in Fig. 3.4. The first step is to get the input-output data on Hankel form. Then, through oblique projection, QR decomposition and SVD, the state sequences are obtained. Finally, least squares can be used to obtain the state space matrices.

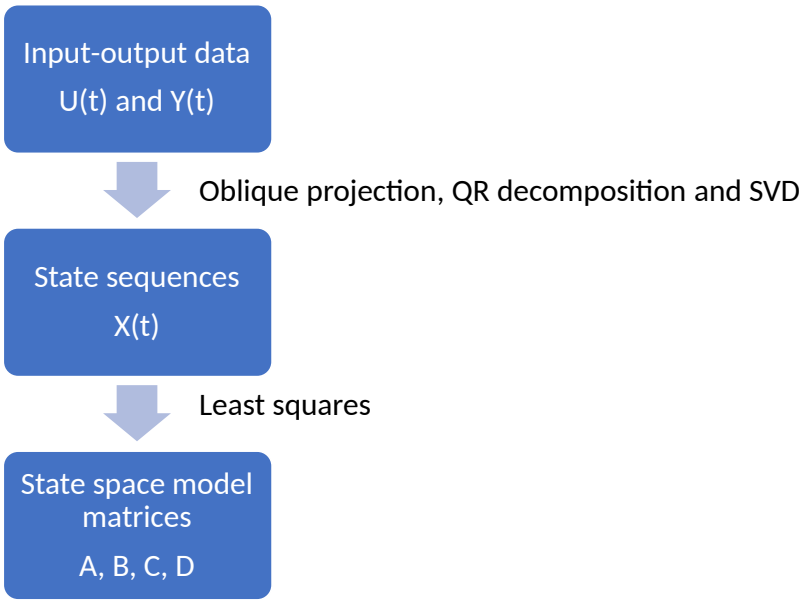


Figure 3.4: N4SID algorithm, where the boxes represent the quantities obtained in each step. The text explains how to get from step to step. Inspired by [4]

A more thorough step-by-step dividing Fig. 3.4 into smaller steps are found in Appendix H.

Chapter 4

Proposed Algorithm: N4SID+DBSCAN

4.1 Proposed Algorithm

The proposed algorithm is intended for ambient conditions, and is based on the subspace identification method N4SID to find the system parameters. The N4SID algorithm requires the user to decide model order. This is normally done by looking at the singular Hankel values, and deciding a fitting order. However, the output of the algorithm may give both true modes and modes deduced from noise or a too high model order. The algorithm has no way to distinguish between true system modes and other modes. Therefore, the combination N4SID+DBSCAN is proposed. DBSCAN is a clustering method that clusters nearby points. The assumption is that if running N4SID on an interval of model orders, the true modes will be repeated irrespective of the model order. By clustering all the modes found on the interval of model orders, estimates of the true modes are found. The clustering technique has been used with other oscillation monitoring methods before, for example with CPCA in [35] and with Prony's method in [36].

The proposed algorithm involves the following steps, which will be further explained in the following sections

1. Pre-processing and filtering
2. Model estimation
3. Clustering of eigenvalues
4. Mode Shape
5. Excitation analysis

4.1.1 Pre-Processing and Filtering

Firstly, it is important to make sure that the sampling frequency fulfills the Nyquist-Shannon criteria for the frequency range which is desired to monitor. The criteria states that the sampling frequency must be double the frequency which is to be observed in the signal [37]. As electromechanical modes are within 0.1-3 Hz, the lowest sampling frequency should therefore be 6 Hz, in order for them to be observable. As the sampling frequency of PMUs are usually 10 Hz or higher, this condition is normally satisfied. There are, however, advantages of downsampling the signal e.g. with respect to the speed of the algorithm. Then, this becomes an important consideration.

Potentially bad and missing data should be accounted for. In this algorithm, this is done by simply removing the inputs that are not numbers. There are, however, more complex data-driven ways to recover missing data, e.g. as done in [38] and [39].

Frequencies that are higher than the sampling frequency may be interpreted as contributions from the lower frequency spectrum. This phenomenon is called aliasing [40]. It can be beneficial for the speed of the algorithm to down-sample the signal to a lower frequency. In this case, it is important to filter the high frequencies outside the range of frequencies of interest. This is accommodated by a low-pass filter in this algorithm. The low-pass filter also serves to eliminate high frequency noise.

In order to remove the low frequency components outside the range of interest and to eliminate any DC-component, a high-pass filter is implemented. Both the low-pass and high-pass filters are first order, and together they form a band-pass filter between 0.1 and 3 Hz in order to accommodate the search for electromechanical modes. The bode diagram of the filter is displayed in Fig. 4.1.

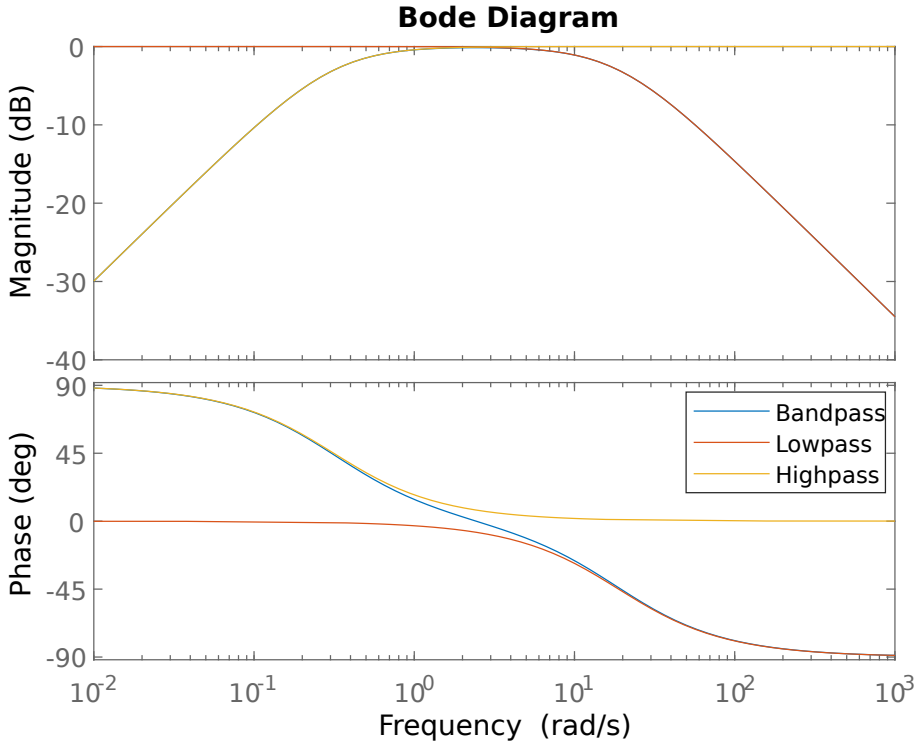


Figure 4.1: Bode diagram of combined lowpass and highpass filter (fiks x-akse).

4.1.2 Model Estimation with N4SID

The model estimation algorithm used was N4SID implemented in Matlab. The input to the algorithm is output data streams from the system and the order of the system. The order of the system is normally determined based on the Hankel singular values, where it is desirable to choose an order such that the value of the higher order are comparatively small. In this algorithm, however, the problem of choosing a model order is turned into an advantage by clustering. The model estimation is run over an interval of orders, in this case between 10 and 20. The model order is automatically increased to deal with negative poles. For each iteration on the model order, the eigenvalues of the **A** matrix are stored for later use in the clustering part [1].

[1]It can be noted that in some cases, using prediction error minimization (PEM) improves the model estimation. PEM updates the initial parameters of the model to fit the data. It can be used with N4SID to refine the system model. This was tried out, but it did not improve the model in this case, as an example in Fig.4.2 demonstrates. It was therefore not included in the algorithm.

4.1.3 Clustering and Averaging Eigenvalues Using DBSCAN

For each new order of model estimation, new eigenvalues are calculated. Some of the eigenvalues calculated are true eigenvalues of the system, however many of them are due to noise in the signal or due to the order of the algorithm being higher than the number of observable modes. It is assumed that if there exists an observable mode in the signal and the model estimation is reasonably correct, its eigenvalue will stay relatively constant through the changes of model order. Therefore, by observing which modes are repeated in multiple model orders, one can distinguish the true modes of the system. In order to separate the groups of true modes from the noise, a clustering algorithm is used.

The clustering algorithm used is Density-based spatial clustering of applications with noise (DBSCAN) implemented in Matlab. The principle is to index each data point as either a part of a cluster, or as noise. A cluster is defined by the maximum distance to a neighbour, ϵ, and the minimum number of points of the cluster. The advantage of DBSCAN compared to other clustering methods is that the shape of the cluster is arbitrary. Firstly, it assumes that none of the points belong to a cluster. It picks one point as starting point, and considers whether there are neighbours within ϵ distance. If there are no neighbours, it classifies the point as noise and proceeds to the next point. If there are neighbours, however, it repeats the process with the neighbours until there are no neighbours within ϵ distance. If the number of points in the cluster is above the minimum cluster size, it labels all the points in the cluster with the cluster value. The first cluster found will be indexed as one, the second as two and so on. Points that are not assigned to a cluster get the index -1. The clustering is completed when all points are assigned to a value. For a thorough description of how the clustering is carried out, the interested reader is referred to [41].

Through trial and error, ϵ was chosen to be 0.5 and minimum cluster size depended on the type of data used in Chapter 5, and was 8 in the case of simulated data and 4 with the real PMU data. After clustering, the averaged modes are found from the centroids of the clusters. The averaged modes are, finally, the estimated modes of the algorithm. From these eigenvalues, the estimated frequency and relative damping of

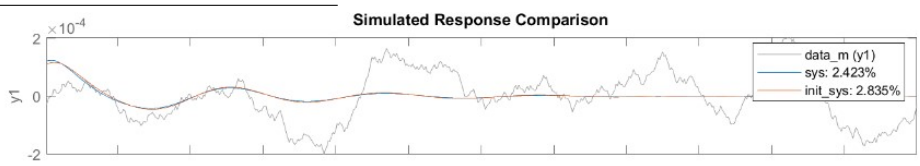


Figure 4.2: Estimated signal with N4SID compared to N4SID+PEM. init_sys refers to the model estimated with N4SID, and sys refers to the model estimated with PEM in addition to N4SID.

the modes are calculated with equations 2.26 and 2.27.

4.1.4 Mode Shapes and Plotting

As demonstrated in 2.4.2, the mode shape is dependent on the system matrices. Therefore, the mode shapes must be obtained from the state matrices of one or multiple orders. In this algorithm the largest order in the interval is used to obtain the mode shapes. As the true modes have been found through clustering, only the mode shapes of these modes are interesting to look at. These are plotted as compass plots.

Chapter 5

Case Studies with Simulated and Real PMU Data

In order to investigate the different attributes of the algorithm, some case studies were conducted. Firstly, a base case with simulated ambient data was used to verify that the algorithm detects modes with some accuracy. The true modes in the simulated data were obtained through linearization of the model as explained in 2.4. Knowing the true modes made it possible to evaluate the accuracy of the estimates of the proposed algorithm. The base case section also serves the purpose of leading the reader carefully through the algorithm.

Case 1 investigates whether the algorithm works on various typical outputs from the power system, namely electric and mechanical power, rotor speed and angle, voltage magnitude and angle, and generator current. It is important to know which outputs the proposed method performs better on in order to achieve as high accuracy as possible in real-time monitoring. The performance of each output is measured based on the accuracy of the estimation of the modes and of the mode shapes.

In Case 2, a change is done to the topology such that the modes are displaced. This is achieved through disconnecting an interconnecting line between the areas. The goal of the test is to see how quickly the algorithm finds the new mode after the change.

Real PMU data is introduced to the algorithm in Case 3. The PMU data is obtained from 8 PMUs located across Norway during an event in 2018 where standing oscillations were observed for 2 min. The data set has previously been used in [35] with an algorithm based on CPCA. The results from Case 3 are compared with the findings of [35].

It must be noted that all the cases are done on only one set of data, which means that they only give demonstrations of the workings of the algorithm. In order to get conclusions of statistical significance, a much higher number of Monte Carlo tests

would have to be run. Thus no numerical conclusions about the algorithm can be drawn from these test cases. The tests can, however, give an indication as to what are the important aspects of this algorithm.

5.1 Simulation Script: DynPSSimPy

An open source code called DynPSSimPy is used for modelling the power system [42]. The model was implemented in Python by Hallvar Haugdal as a contribution to his PhD work at NTNU. The script called DynPSSimPy is developed from scratch and involves both linear and nonlinear solvers, time domain simulation and modal analysis.

In order to simulate a PMU used in the Nordic grid, the sampling frequency is set to 50 Hz. In order to achieve a constant sampling frequency, it is important that the time step of the solver is constant. The solver used is Runge Kutta, and the time step is kept constant by setting the maximum step size to 0.02 s.

5.2 Test Model: Kundur's Two Area Model

In order to test the algorithm it is convenient to produce a set of data where the modes are already known. With small disturbances around the linearized operating point, the modes from modal analysis of a system should be observable in the state variables. The model used for testing is the Kundur two area model, as presented in [5] and displayed in Fig. 5.1. The system consists of four generators, of which G1 and G2 are hydroelectric generators, and G3 and G4 are steam turbine generators. Governors, excitation systems and power system stabilizers are connected to all generators. This is a good system to observe modes in, as it has three electromechanical modes, of which one is inter-area and two are local area. The governor values were slightly modified in order to reduce the damping of the modes.

The eigenvalues of the system are plotted in Fig. 5.2 (a). A closeup of the eigenvalues where the local and inter-area modes are highlighted is displayed in Fig. 5.2 (b). A general rule in power systems is that the damping should be more than 5%, which is satisfied in this situation.

The mode shapes of the system are displayed in compass plots in Fig. 5.3. The plots show that there are two local area modes at around 1 Hz, one in area 1 between G1 and G2, and one in area 2 between G3 and G4. G1 is oscillating almost exactly in antiphase with G2, and the same goes for the second local mode where G3 is in antiphase with G4. The mode shapes show that the mode at 0.56 Hz is an inter-area mode between area 1 and area 2. The generators in area 2 is participating the most

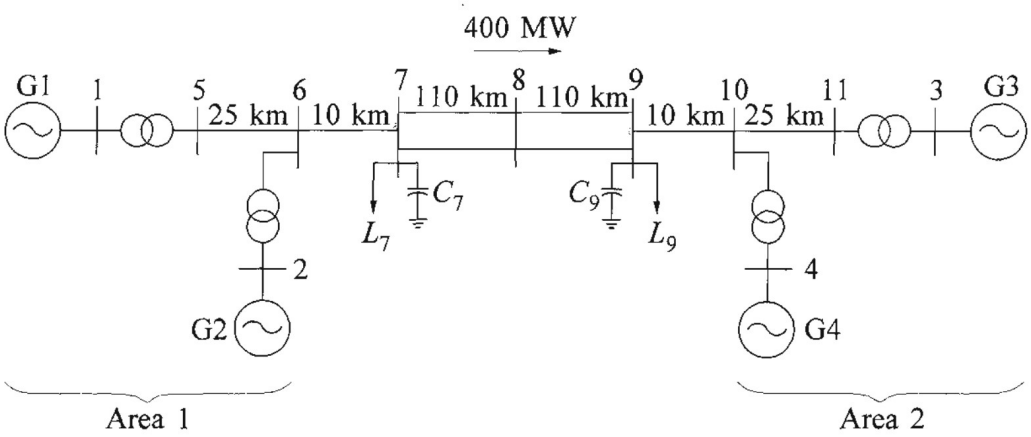
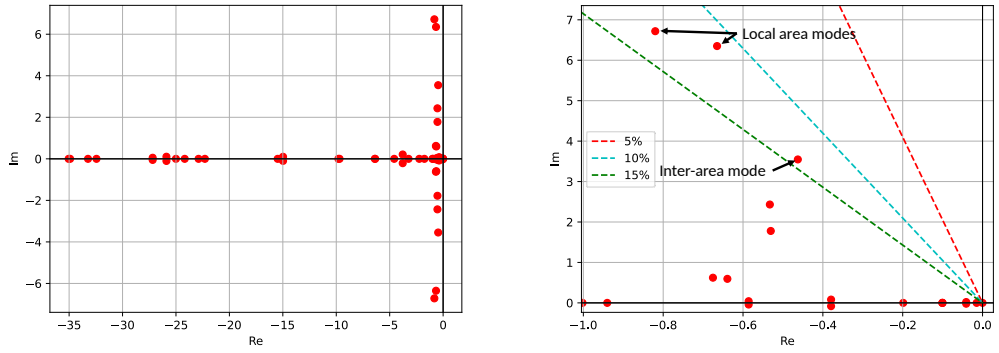


Figure 5.1: Kundur two-area system [5]



(a) Eigenvalues of the Kundur two area system.

(b) Electromechanical modes in the system.

Figure 5.2

in this mode. The two control modes at 0.28 Hz and 0.39 Hz involve the generators in area 1 oscillating against the generators in area 2, with a moderate phase shift.

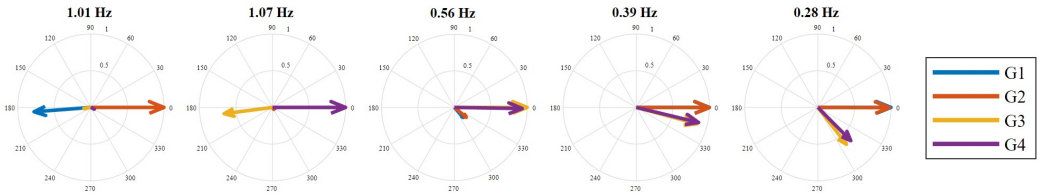


Figure 5.3: Mode shapes of the Kundur two-area system.

The dominant modes of the modified model are summarized in Tab. 5.1. It should be noted that the damping of the electromechanical modes are the lowest of the modes in the system. They might therefore be easier to observe, although as displayed in Fig. 2.10 low frequencies are sustained for longer, and hence might be easier to detect.

Frequency [Hz]	Damping ratio [%]	Generator	Vs.	Generator
1.07	12.11	G3		G4
1.01	10.41	G1		G2
0.56	12.93	G1, G2		G3, G4
0.39	21.37	G1, G2		G3, G4
0.28	28.61	G1, G2		G3, G4

Table 5.1: Modes of the Kundur two area system.

5.2.1 Modelling Load Variation in Ambient Conditions

Load variation happens constantly in the power grid due to the nature of consumption of electricity. Even small load variations during ambient conditions may excite electromechanical modes. It is impossible to measure all the loads in the system, so only outputs can give information about excited modes. The frequency is one such output. How much the frequency varies as a response to load variation depends on where in the grid it is measured and at what time during the year, week and day. An example of the frequency during a 2 minutes interval during ambient conditions is given in Fig. 5.4, as measured by a Statnett PMU at Hasle. It is seen that the frequency varies about an interval of 0.1 Hz. In the simulation script, the variation is modelled as normally distributed load changes between 0 and 0.1 pu. It is implemented by modifying the reduced Y-bus diagonal element after Kron reduction. The random load change happens at every time step, which is set to be 0.02 seconds. The

frequency can be seen to vary about 0.0005 Hz in this interval. Large fluctuations at low frequencies that would be present in a real system due to larger load changes are not present in the signal. As seen by comparing Fig. 5.4 and Fig. 5.5, the variations are smaller by a factor of 200. There are many ways to model load variations, but it is a generally accepted simplification to assume load variation during ambient conditions to be uncoloured noise on short time intervals. This means that if one was to filter the real ambient signal through an appropriate high-pass filter, one would obtain something like the simulated signal.

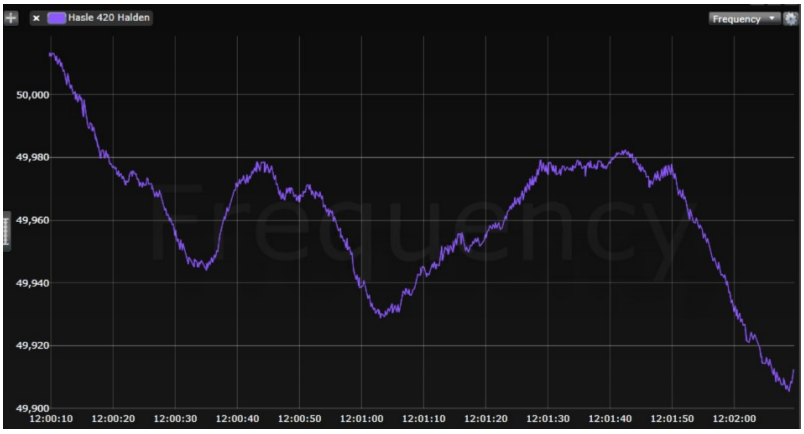


Figure 5.4: Frequency variations during ambient conditions at the interconnection Hasle-Halden on a 2 minute interval measured by PMU.

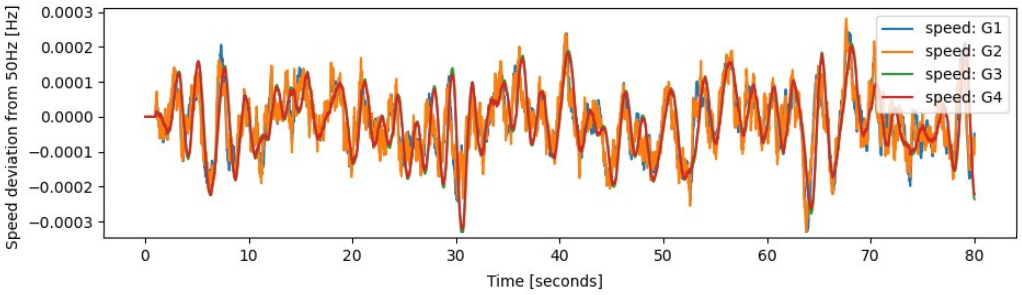
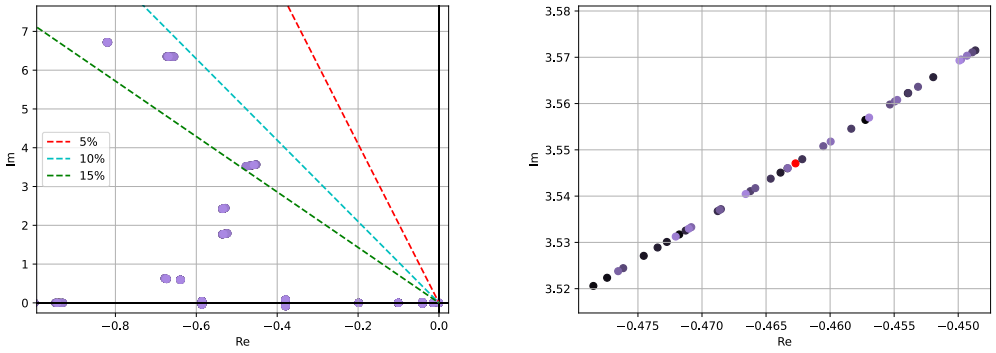


Figure 5.5: Frequency variations and variation of power production and load during ambient conditions on a 2 minute interval from simulations.

5.3 Base Case: Ambient Conditions

A base case consisting of only ambient data was made in order to test that the algorithm works. The simulation lasts 80 seconds, and the whole simulation period is used as input to the algorithm. The output data used is the rotor speed of the four generators in the system. The load variation will have a slight impact on the eigenvalues throughout the simulation. This is shown in Fig. 5.6, where the eigenvalues are plotted every other second of the simulation. As the closeup of the 0.5 Hz inter-area mode shows, it varies between 0.56 Hz and 13.47% damping to 0.57 Hz and 12.46% damping. The first local area mode remains on 1.01 Hz, and varies between 10.58% and 10.25% damping throughout the simulation period. The second local area mode remains on 1.07 Hz and varies between 12.14% and 12.09%.



(a) Eigenvalues plotted every other second during the 80 s simulation. The original eigenvalue of the system is plotted in red, then the eigenvalues are plotted in gradual colors.

(b) Closeup of the inter-area mode variations. The mode varies between 0.56 to 0.57 Hz, and the damping varies between 13.47% and 12.46%.

Figure 5.6

5.3.1 Filtering

The effect of the band-pass filter is displayed in Fig. 5.7. The filtered data has less high frequency noise due to the low pass filter, and the trend is removed due to the high pass filter. It is difficult to observe the settling time of the filter on the rotor speed in ambient conditions, as it is already centered around 0. This will be more prominent in Case 1.

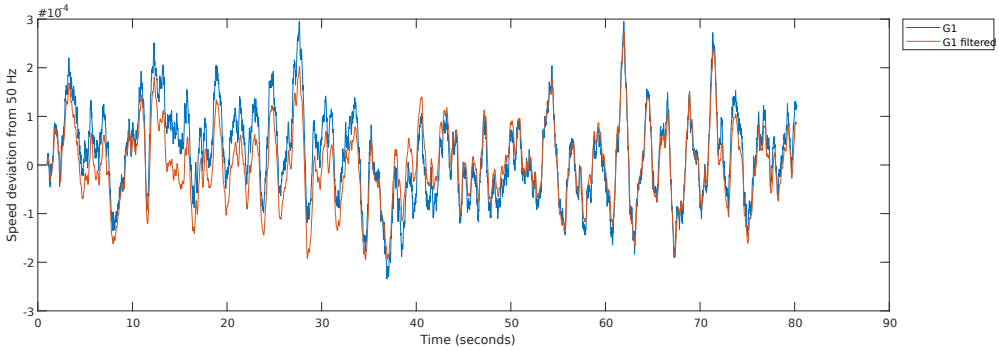
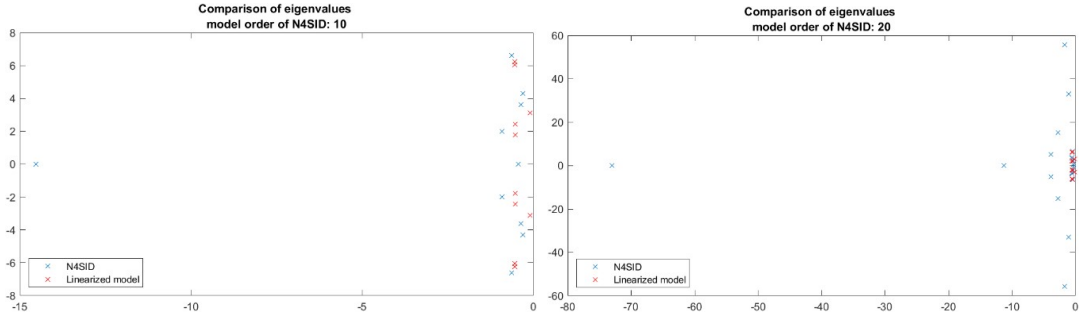


Figure 5.7: Data of the rotor speed at G1 before and after filtering in blue and orange respectively.

5.3.2 Model Estimation

For each model order, a new estimate of eigenvalues are made. Of course, a higher model order will result in more modes, and there seems to be a tendency that higher model orders give additional eigenvalues of higher frequencies, as Fig. 5.8 demonstrates. The additional modes found with higher orders are either of such high frequency that they are uninteresting, or they are purely real eigenvalues, which means they do not hold oscillatory properties and are therefore uninteresting.



(a) Eigenvalues of N4SID model order 10 in blue, and modes from the linearized nonlinear model in red.

(b) Eigenvalues of N4SID model order 20.

Figure 5.8: The figures show how higher model orders find more eigenvalues in general, and specifically find more modes of higher frequencies and lower damping.

5.3.3 Clusters and Mode Shapes

The clustering algorithm found 4 eigenvalues, which are displayed in the cluster plot in Fig. 5.10. The estimated modes are listed in Table 5.2. All the modes that are found are in the vicinity of the known true eigenvalues of the system. In addition to the oscillatory modes, the algorithm also finds a purely real mode at -0.4. The algorithm does not manage to distinguish between the two local area eigenvalues that exist at about 1 Hz, nor does it distinguish between the two control modes at 0.28 and 0.39 Hz. The frequency of the estimated modes are fairly accurate, but the damping seems to be consistently underestimated, even when accounting for the change of damping due to the load variations.

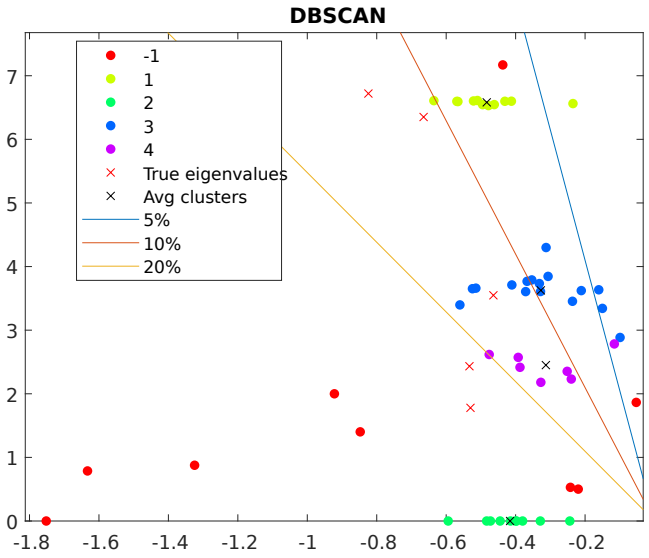


Figure 5.9: Result of clustering the eigenvalues with DBSCAN. The red crosses are the true eigenvalues obtained through linearization of the nonlinear model. The black crosses are the average of each cluster from the DBSCAN algorithm.

Frequency [Hz]	Damping ratio [%]	Generator	Vs.	Generator
1.05	7.32	G1		G2
0.58	8.99	G1, G2		G3, G4
0.39	2.67	G1, G2		G3, G4

Table 5.2: Estimated modes in base case.

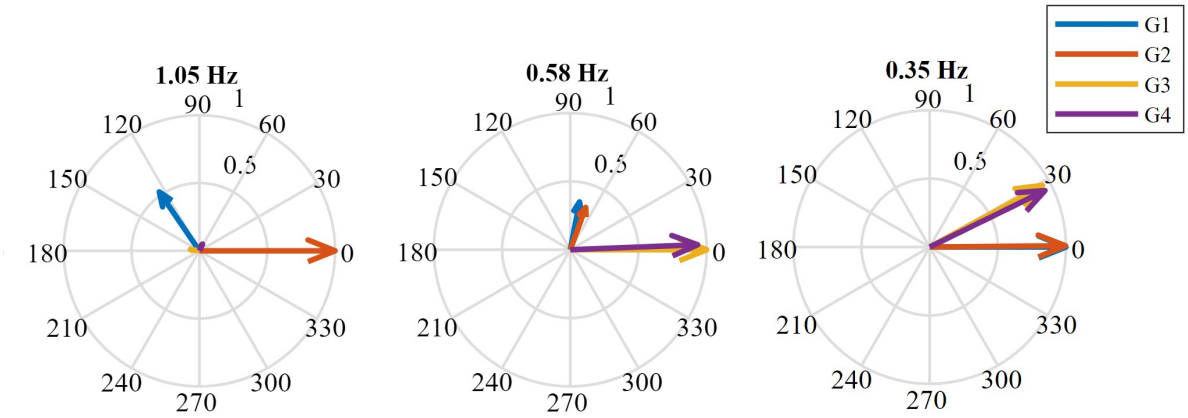


Figure 5.10: Phasors obtained for the estimated modes.

5.4 Case 1: Comparing Performance of the Proposed Algorithm with Respect to Various Output Signals

It is useful to know which data output gives the best estimation of the electromechanical modes in the system. In order to see the effect of output signal on the estimated modes, the algorithm was run on the same ambient scenario as in the base case, but with various output signals. The signals are obtained from simulation of ambient condition as described in 5.3. The output signals used as input to the algorithm are electric and mechanical power, rotor speed and angle, voltage magnitude and angle and generator current.

The ambient time domain data of all the output signals from the four generators in the modelled power system are displayed in Fig. 5.11. The output signals vary significantly with respect to amplitude and noise. The mechanical power at G3, for example, is subject to a significant amount of high frequency noise. It can be noted that the voltage angle is similar in trend to the rotor angle, though it has more high frequency components.

During the simulation of this case, it was observed that the filter has quite a long settling time. It takes nearly 20 s for the filter to settle on the output signal of the electrical power, as displayed in Fig. 5.12. This is not a major issue, but something one must be aware of when using the filter. When using the band-pass filter, one should take care to not use the first 20 s of the data for the oscillation monitoring algorithm as it may distort the results. The simple band-pass filter is one of the simplest forms of filters, and there are other filters with shorter settling time that

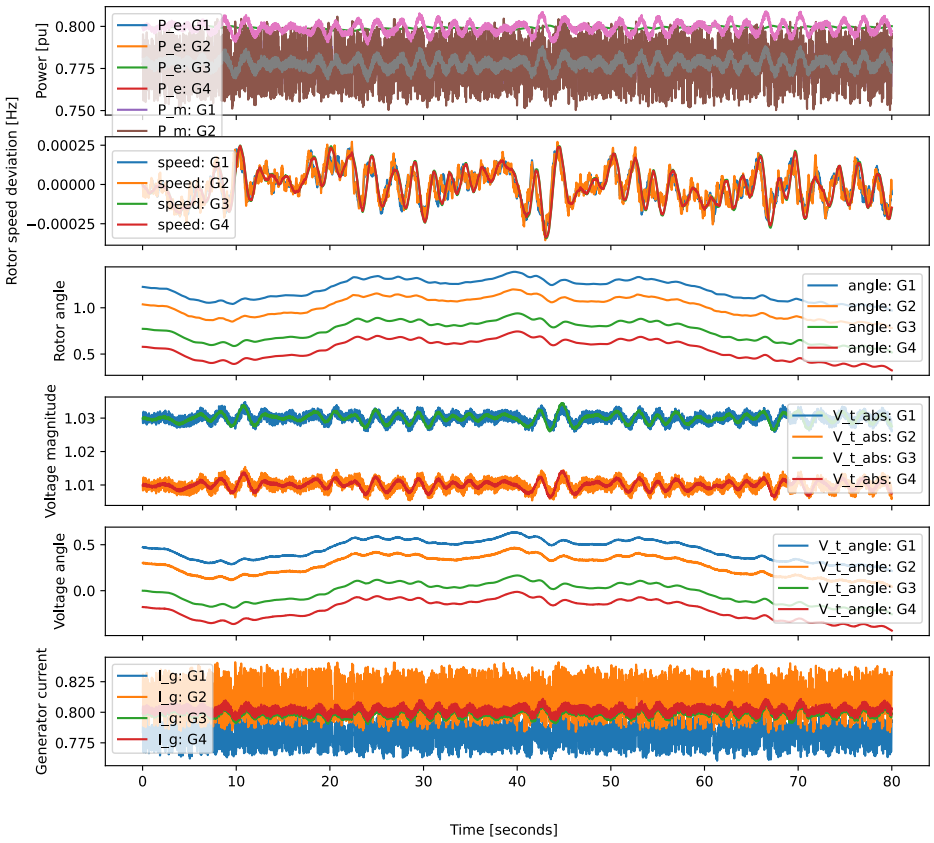


Figure 5.11: The time domain signal of all the outputs that are fed to the algorithm are displayed in this figure.

would solve this problem, for example EMD as used in [43]. There is, however, often a forced compromise between short settling time and steady gain in the frequency band that needs to be addressed and suited to each situation. In this situation, it might be more important with a steady gain, as an oscillating gain could potentially interfere with the oscillation monitoring algorithm.

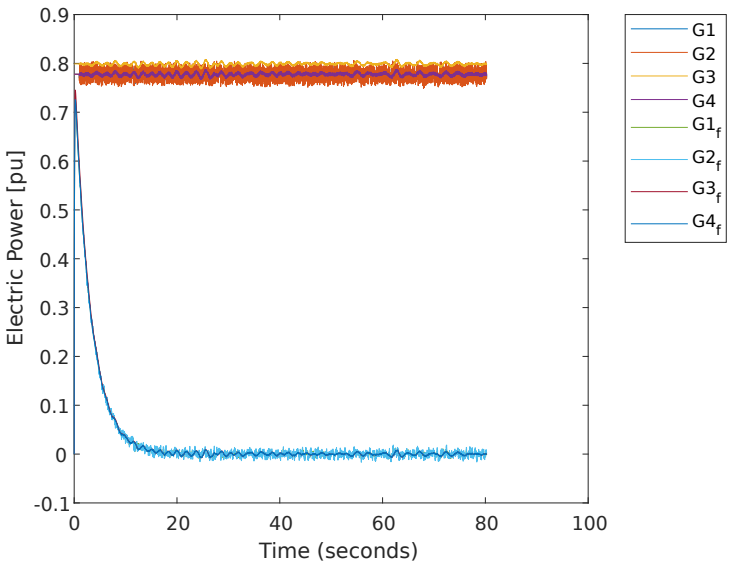


Figure 5.12: Electrical power before and after filtering. Observe that it takes nearly 20 s for the filter to settle on the signal.

The averaged estimated modes are displayed in Fig. 5.13. As the figure shows, most outputs produce an estimate of three separate modes in vicinity to the clusters of true modes. In this single test case, the rotor speed and angle output proved most successful for estimating the modes. They both find quite accurate estimates of the three modes. Interestingly, the current produces an estimate of much better damping of the inter-area mode at 0.55 Hz, and the voltage angle estimates the damping of the control mode at 0.30-0.40 Hz to be much less. A potential explanation of this is that the parameters of the clustering algorithm might need finer tuning.

The mode shapes found with each output is displayed in Fig. 5.14. The first row shows the true mode shapes found by linearization of the Kundur system. As previously discussed, the Kundur system has two local area modes at 1 Hz. As Fig. 5.14 shows, the algorithm only captures one of them, namely the local mode in area 1. G1 participates the most in this mode, and this is captured by the algorithm regardless of the system output. The phase shift between G1 and G2 is most accurately measured with the rotor angle. Many of the estimates find a shift of about 120 degrees. In some

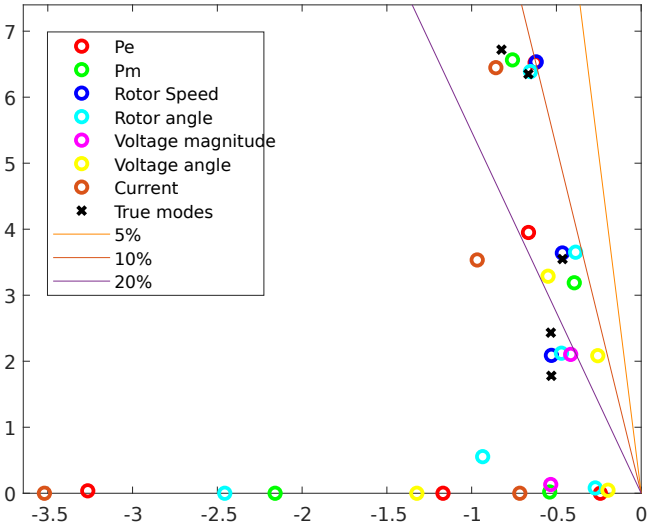


Figure 5.13: Estimated modes with varying outputs.

compass plots this is represented as a negative angle, or a phase shift of 240 degrees. This is due to the fact that modes are always found in pairs of complex conjugates.

All the mode shapes found of Mode 2 at 0.56 Hz correctly estimate the two major components, G3 and G4. However there is some variation as to the participation and the phase shift of G1 and G2. The two control modes are similar in mode shape, and the algorithm detects similar mode shapes. Remark that using electrical power and current as output did not yield a mode for the control mode.

This is just a single case, and the simulation should be repeated for a high number of random Monte Carlo simulations in order to obtain information of statistical significance. However, the electromechanical modes are found with most outputs, except for the voltage magnitude, as expected. This indicates that the algorithm works with multiple outputs from the power system.

5.5 Case 2: Change in System Topology

A large change in the system topology was made by disconnecting one of the lines between bus 7 and bus 8, as showed in Fig. 5.15. This is an interconnecting line between the two areas, and the disconnection leads to lower transfer capacity between the areas, and a higher loading on the remaining weak line between bus 7 and 8. The topology change produces new eigenvalues of the linearized model, and the goal of

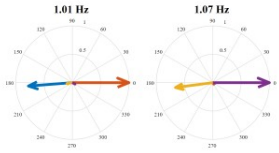
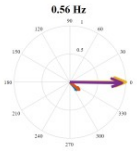
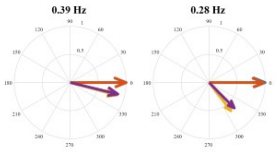
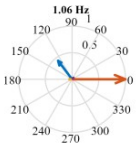
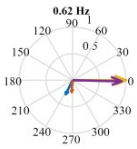
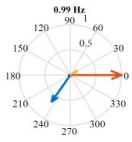
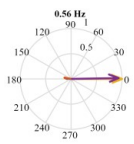
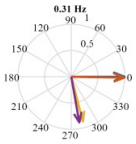
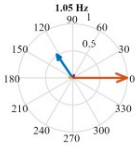
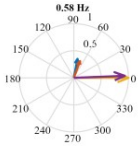
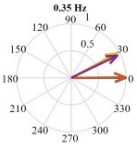
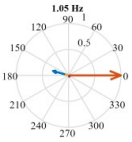
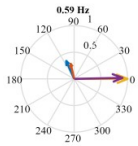
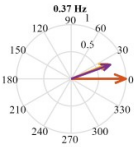
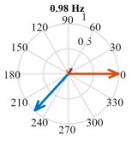
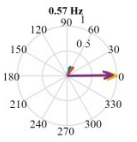
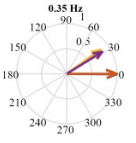
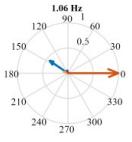
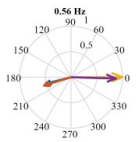
G1 G2 G3 G4	**Mode 1** **~1 Hz**	**Mode 2** **~0.55 Hz**	**Mode 3** **~0.30-0.40**
Kundur System			
Electric Power			
Mechanical Power			
Rotor Speed			
Rotor Angle			
Voltage Magnitude			
Voltage Angle			
Current			

Figure 5.14: Mode shapes found with varying outputs.

Case 2 is to investigate how long it takes for the algorithm to detect the change. This is achieved by implementing the algorithm as a sliding window method, and observing when the new modes are detected.

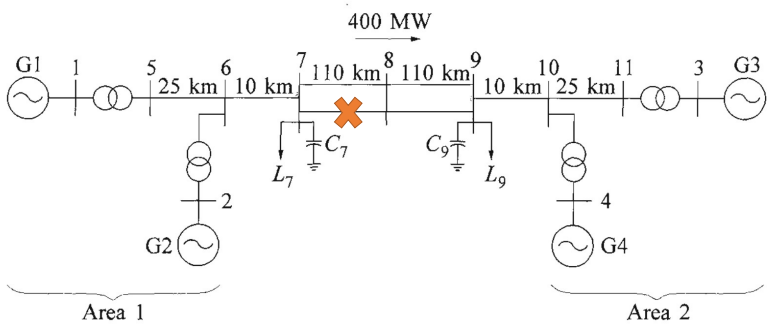


Figure 5.15: This image shows where the line tripping occurs.

The new system topology has a new set of eigenvalues as displayed in Fig. **??** and in Table 5.3. A higher line loading leads to a higher impedance in the line, which reduces the damping of the system. It is particularly the lower control control mode that is poorly damped after the topology change, dropping from 28.61% to 17.65% relative damping. This mode is expected to be detected by the algorithm. Fig. 5.16 shows the variations of the mode during the simulation due to load variation.

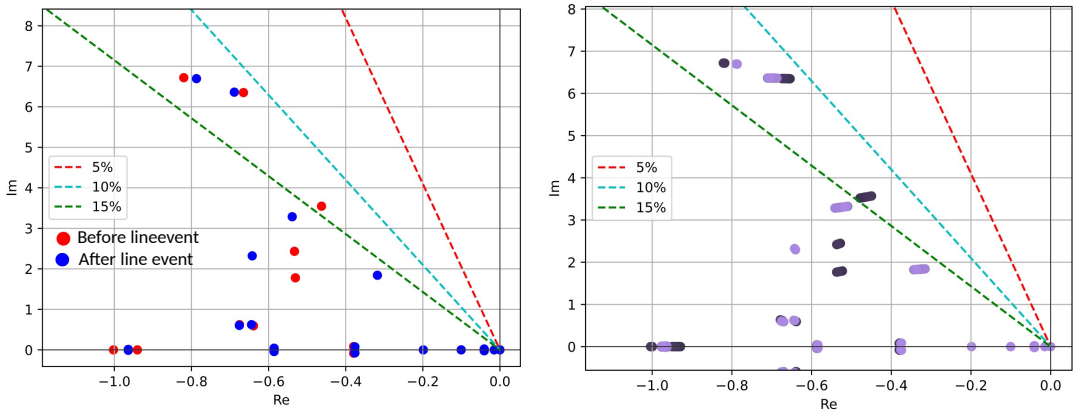


Figure 5.16

The simulation data set is 4 min long, and the line event happens halfway through at 2 min. The input to the algorithm is the rotor speed at all four generators in the

Before Line Event		After Line Event	
Frequency [Hz]	Damping ratio [%]	Frequency [Hz]	Damping ratio [%]
1.07	12.11	1.07	11.69
1.01	10.41	1.01	10.90
0.56	12.93	0.53	15.73
0.39	21.37	0.37	26.75
0.28	28.61	0.29	17.65

Table 5.3: True modes before and after the line event obtained from system linearization.

system, which is plotted in Fig. 5.17. The rotor speed of all generators spike at the time of disconnection of the line, and then stabilize on a higher level due to the increased impedance on the interconnection.

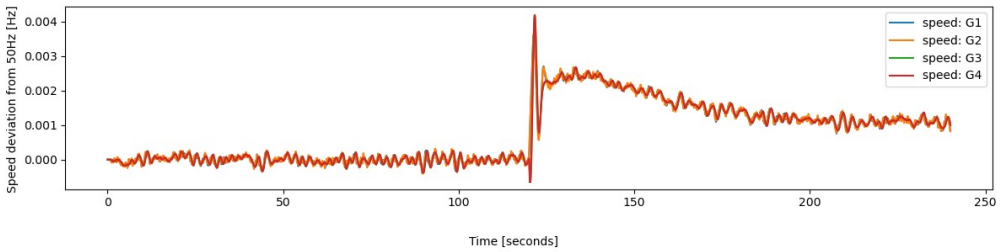


Figure 5.17: Rotor speed and powers during simulation of line tripping at t = 120 s.

In the case of a large disturbance in the time series, the effect of the filter is clearly observable. Low frequencies are removed reducing large fluctuations in the signal, and high frequencies are removed causing a smoother signal with less noise.

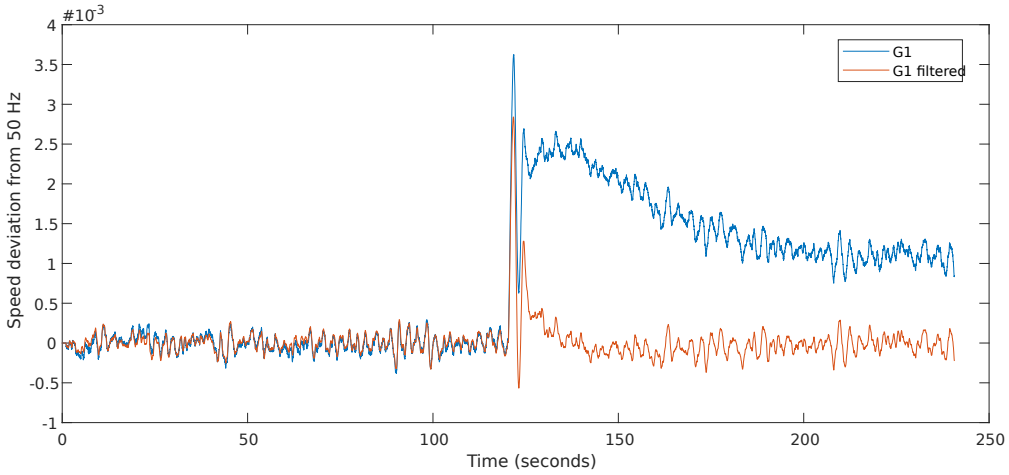
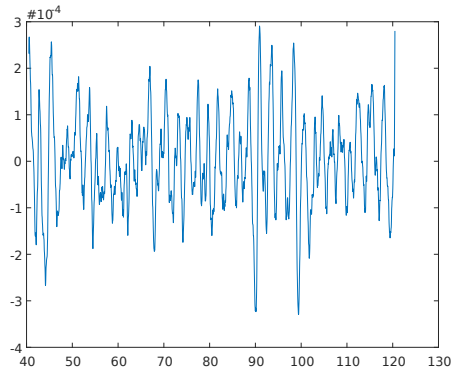


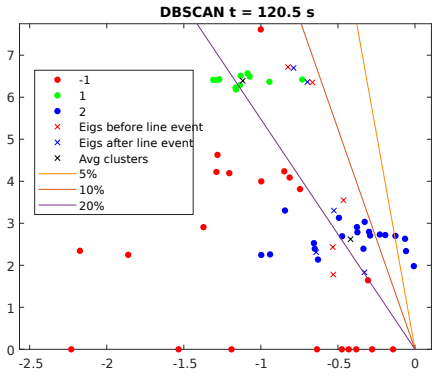
Figure 5.18: Filtered and unfiltered signal of the rotor speed at G1.

For each sliding time window, the clustering algorithm finds a new set of modes. The time domain signal of four windows during and right after the transient are displayed in Fig. 5.19 with the corresponding clusters of the windows. At t = 120.5 s, the algorithm only finds the 1 Hz mode, and seems to struggle to find the lower frequency modes. In the second window at t = 122 s, the algorithm identifies the 1 Hz mode, the 0.5 Hz mode, and a low frequency mode some distance from the true control mode. At t = 125 s, the algorithm correctly identifies the three low-damped modes in the system, and continues to do so in the time span after, as the example of t = 130 shows in Fig. **??**.

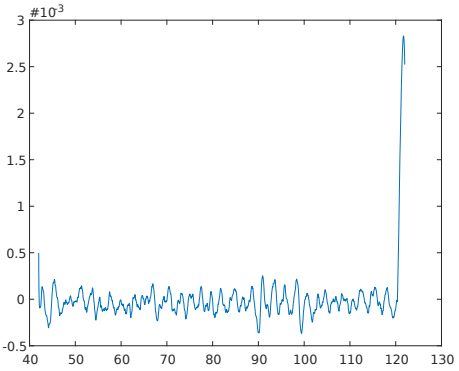
The algorithm continues to not distinguish between the two modes around 1 Hz and the two modes around 0.3-0.4 Hz. This could be due to choice of parameters in the clustering algorithm. The algorithm does, however, consistently find the three clusters of modes to a relatively good accuracy already 5 seconds after a change of system topology. The results indicate that the algorithm may detect changes in topology rapidly. This is particularly important in control room settings, where it may be needed to detect modes quickly.



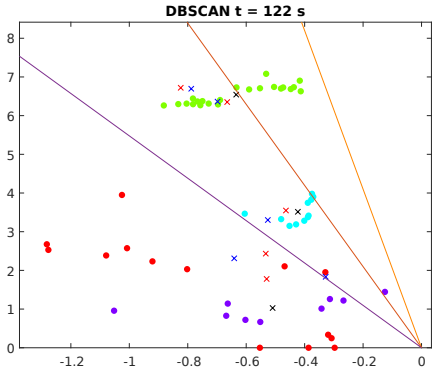
(a) t = 120.5 s.



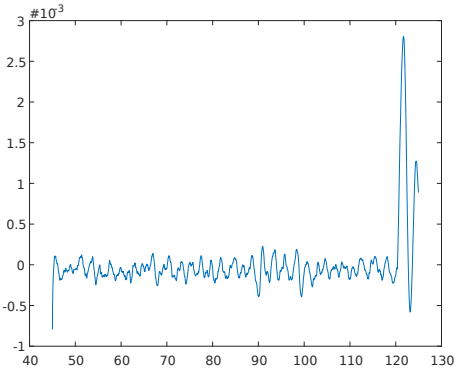
(b)



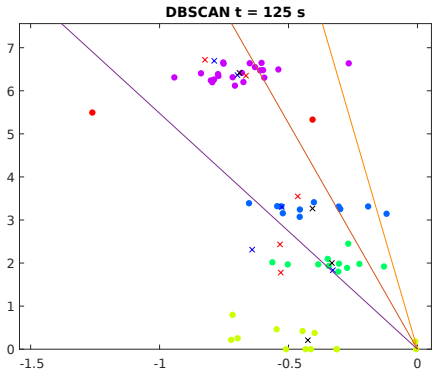
(c) t = 122 s.



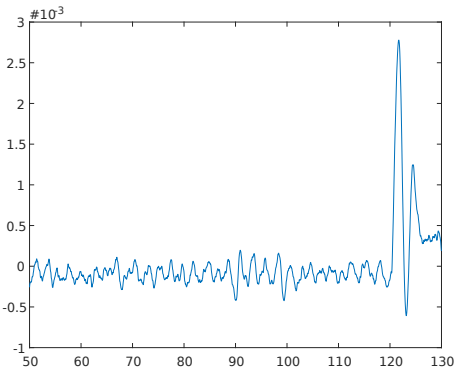
(d)



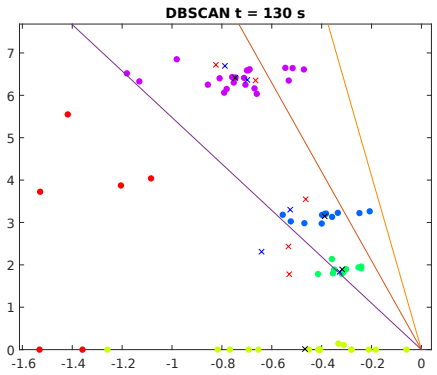
(e) t = 125 s.



(f)



(g) t = 130 s.



(h)

Figure 5.19: Time domain and clusters of sliding windows during a line event.

Before Line Event		After Line Event	
Frequency [Hz]	Damping ratio [%]	Frequency [Hz]	Damping ratio [%]
1.01	9.93	0.98	9.48
0.59	26.30	0.51	19.31
0.45	16.59	0.33	10.88
0.20	62.70		

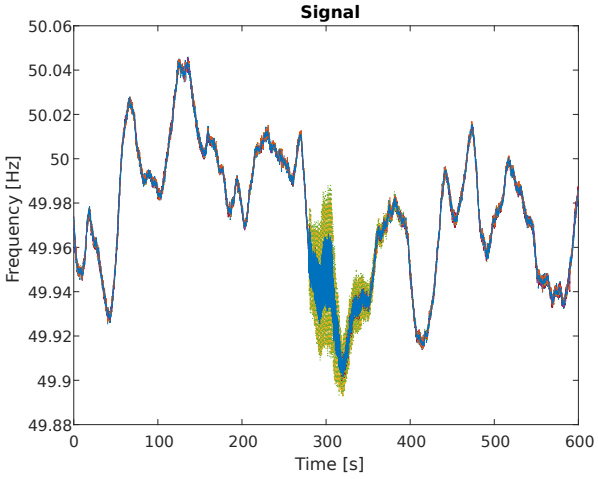
Table 5.4: Modes before and after the line event obtained from the N4SID-DBSCAN algorithm.

The algorithm successfully finds the new low-damped mode after the disturbance. The estimated modes before and after the event are listed in Table 5.4. It can be observed that the algorithm generally is better at finding the low-damped modes. The lower damping might make them sustained for a longer period of time in the signal, which make them easier to detect. This is not necessarily a negative feature of the algorithm, as the low-damped modes are the most critical to system stability anyways. It would be interesting to track the estimated modes continuously after a system topology change to get a more accurate answer to exactly when the algorithm finds the new modes. However, the initial results presented here suggest that the algorithm is relatively quick at detecting new modes.

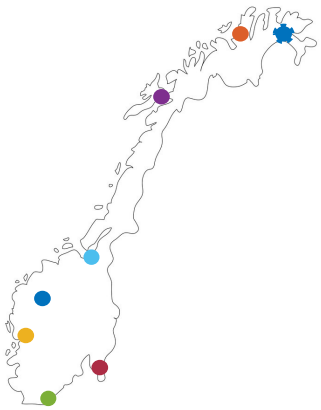
5.6 Case 3: PMU data

On September 24th 2018, there were standing oscillations present in the Norwegian grid. The oscillations lasted for nearly 2 minutes, and involved both inter-area and local area modes. In [35], a study of the event was done using a novel oscillation monitoring method based on CPCA. Case 3 will use the same PMU data as in [35] and investigate the modes using the N4SID+DBSCAN algorithm. The presentation of the case is inspired by the way done in [35], using the same color and linetype scheme to make it easier to compare the result from the two methods.

The signals obtained with 8 PMUs in Norway are displayed in Fig. 5.20 (a). The PMUs are located approximately as displayed in Fig. 5.20 (b).



(a) Frequency signal during oscillations in 2018 measured by PMUs. The oscillations start about 280 s in.



(b) Map of the approximate location of the 8 PMUs. Inspire by [35].

Figure 5.20

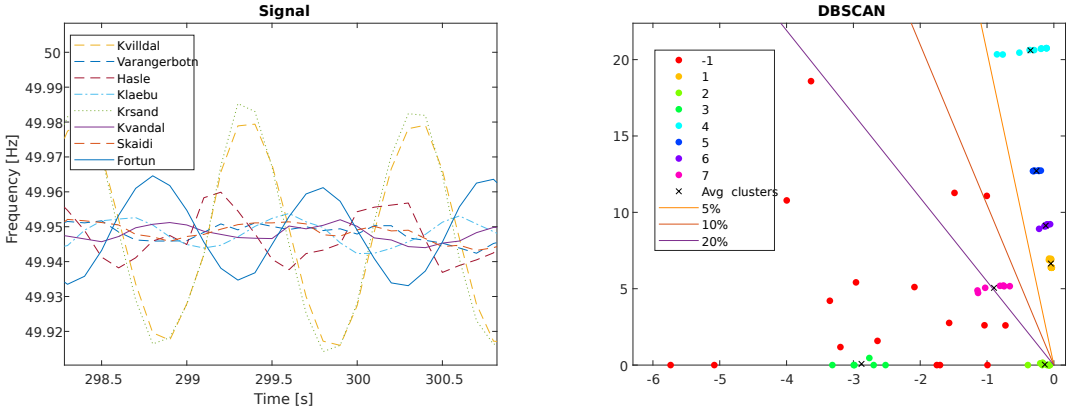
A detail extracted from the signal shown in Fig. 5.21 (a) shows that there is a mode visible to the naked eye at about 1 Hz. From the figure, one can see that Kristiansand and Kvilldal have the highest excitation and are in phase. Fortun is participating in the mode, almost 90 degrees phase shifted, and Hasle is also phase shifted with respect to Kristiansand and Kvilldal. The other signals might participate in the mode, but it is not as obvious by visual inspection only. N4SID+DBSCAN produced the clusters displayed in Fig. 5.21 (b) when run on the oscillation period of 2 min.

The frequency and damping of the estimated modes are presented in Table 5.5. The mode at 1.06 Hz is extremely low-damped with 0.68% relative damping. This is likely to be the mode that is visible throughout the sample.

Frequency [Hz]	Damping ratio [%]
1.06	0.68
1.45	1.33
3.28	1.7
2.02	2.05
0.80	17.44

Table 5.5: Modes found with N4SID+DBSCAN of 2018 oscillation event.

The estimated mode shapes with N4SID+DBSCAN are compared with the estimated mode shapes from [35] in Fig. 5.22. They are consistent with each other. All



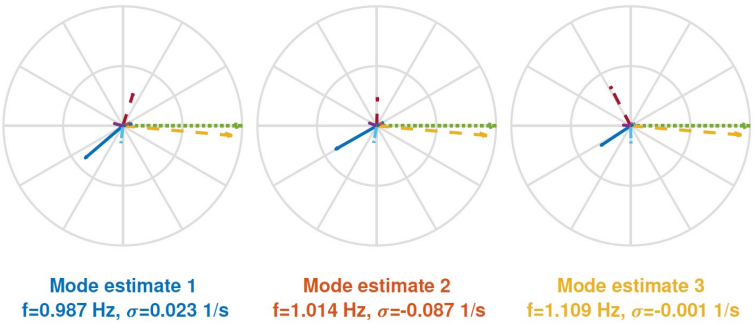
(a) (b)

Figure 5.21: Detail extraction of signal and clusters produced by the N4SID+DBSCAN algorithm.

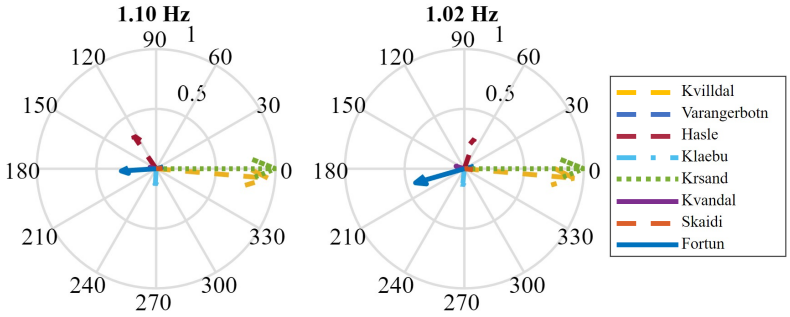
mode shapes present the expected mode shapes by visual inspection described above.

The other mode shapes that were found are presented in 5.23. Based on the estimated mode shapes, Hasle seems to be oscillating against Klæbu at 3.3 Hz. As the frequency is quite high, however, this is more likely to be the case of two local area, or even local plant, modes of the same frequency. There seems to be a local area mode between Hasle and Kristiansand at about 2 Hz. The mode at 1.47 Hz seems to be a local area mode where Kristiansand swings against the rest of the system. The mode at 0.81 Hz seems to be an inter-area mode that involves many of the generators in the system. These modes were not commented on in [35], and they can therefore not be compared, but it was deemed as interesting results to include in this thesis.

The algorithm developed in this thesis successfully detects the same inter-area mode as detected with CPCA in [35]. This may signify that the algorithm works not only on simulated data, but also on real system data from PMUs. This is are good initial results that motivate further work on the algorithm, and testing on real-time PMU data could be a next step.



(a) 1 Hz mode shapes estimated in [35].



(b) 1 Hz mode shapes estimated with N4SID+DBSCAN.

Figure 5.22

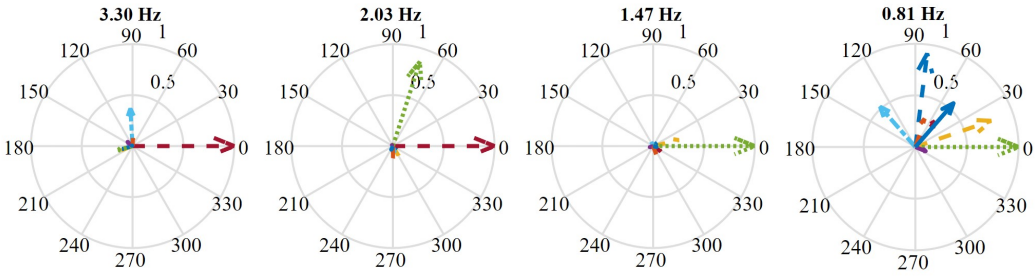


Figure 5.23: Estimated mode shapes from PMU data.

Chapter 6

Discussion of Simulation Results and Further Work

The results of the simulations are promising for the method. More testing of the method is needed in order to ensure that the trends found in these simulations are statistically representative of the algorithm. The algorithm seems to work on many different output signals, with rotor angle and speed having the best performance. The results might be improved if multiple types of output signals are fed to the algorithm simultaneously.

The length of the window has not been thoroughly tested out, but from experience during the testing, it seems that there exists a sweet spot for optimal window size. Too long window size may lead to other trends than electromechanical oscillations distorting the output. Too short window might not provide enough data for the algorithm to detect any modes with accuracy. Anecdotally, 10 min of ambient data from the Kundur system as described in 5 was used as input to the algorithm, and the algorithm did not find any of the true modes in the system.

An improvement to the algorithm that would perhaps give better estimates of the mode shapes is to instead of using the mode shapes obtained from the highest model order on the order interval, use an average of the mode shapes from multiple model orders. This could give the operators better information about which generators are participating in the modes.

The run time of the base case, where the input was four data streams of 4000 data points, was calculated with Matlabs stopwatch timer to be 46 seconds. The stopwatch was started before importing the data from the comma separated file, and ended at the end of the algorithm after calculating the frequencies and damping of the averaged clusters. The test was conducted without plotting. This reveals that although the algorithm may detect new modes within 5 seconds after an event, the delay of the algorithm itself is the major obstacle to real-time detection of oscillations. It might be

interesting to compare the accuracy and latency across a specter of various oscillation monitoring algorithms in a methodical manner.

Per now, the algorithm is too slow for automatic remedial actions. However, it is worth investigating remedial actions that could be supervised by oscillation monitoring. As discussed in [44], a remedial action that gives better damping of oscillations is lowering the power transfer of weak tie lines. This might be a costly measure, and oscillation monitoring plays a key part in the governing of the stability limits of such tie lines. Previously, PMU data has been used to tune PSSs in real time, e.g. in [45]. Oscillation monitoring algorithms such as N4SID+DBASCAN might aid the performance of real time tuning of PSSs.

As the speed of N4SID+DBSCAN is quite slow, it might be a good oscillation method to use during ambient situations. Algorithms developed to obtain oscillation parameters during transient conditions, such as Prony, are much faster. A combination of the two could give benefits to the system operator by allowing a quick notice of new modes during transients, when time is critical, and otherwise give a system overview during ambient conditions. One way to implement Prony with N4SID+DBSCAN could be to have a ROCOF limit that decides which algorithm to use. With large ROCOF, there is likely to be a transient situation and Prony should be run. Transient algorithms might also be used to cross-check the ambient mode estimation when there are transients.

Another combination of algorithms that could provide better system awareness is N4SID+DBSCAN and a FFT based algorithm, as the one developed in the author's specialization project. The FFT based 3D Welch's mode meter gives intuitive information about the excitation of the modes. It could be implemented as an extension of the N4SID+DBSCAN algorithm such that for each estimated mode, the Welch's method is run on the largest element of the mode shapes belonging to the mode. The mode shape with the largest magnitude is the one most observable in the signal, and running Welch's method on it will give the excitation of the mode in that signal. The largest mode shape will have the highest excitation, and the estimate therefore gives the maximum excitation of that mode.

The Welch's mode meter implemented in the specialization project takes only 0.06 s to compute and plot the spectrum, as displayed in Fig.6.1. In the figure, the colors represent the magnitude. It could be investigated whether having the colors correspond to the damping of each mode would be a good visualization method for the control room. This way, the operators would get information about frequency, damping and excitation at once.

When developing oscillation monitoring tools, it is important to keep in mind the goal of the tool - to improve control room operators' system awareness. In the control room the information presented needs to be simple and concretely related to what the operators can do. For example are well-damped modes not as interesting

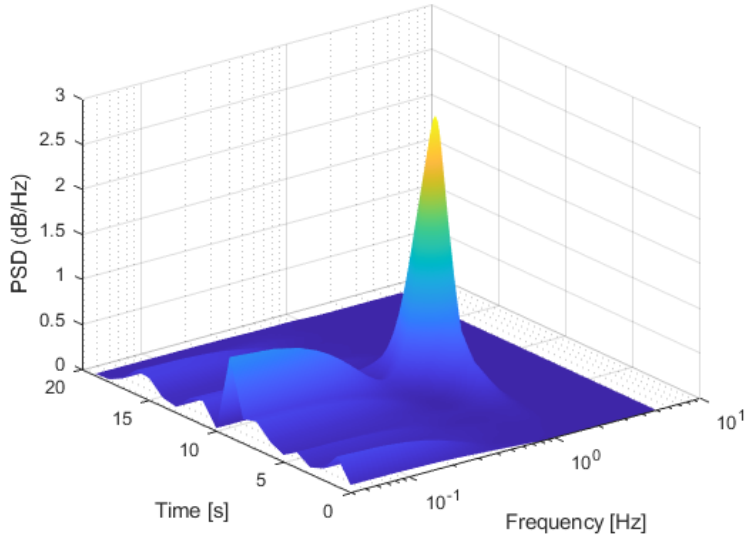


Figure 6.1: Welch's mode meter on the dataset described in 5.6.

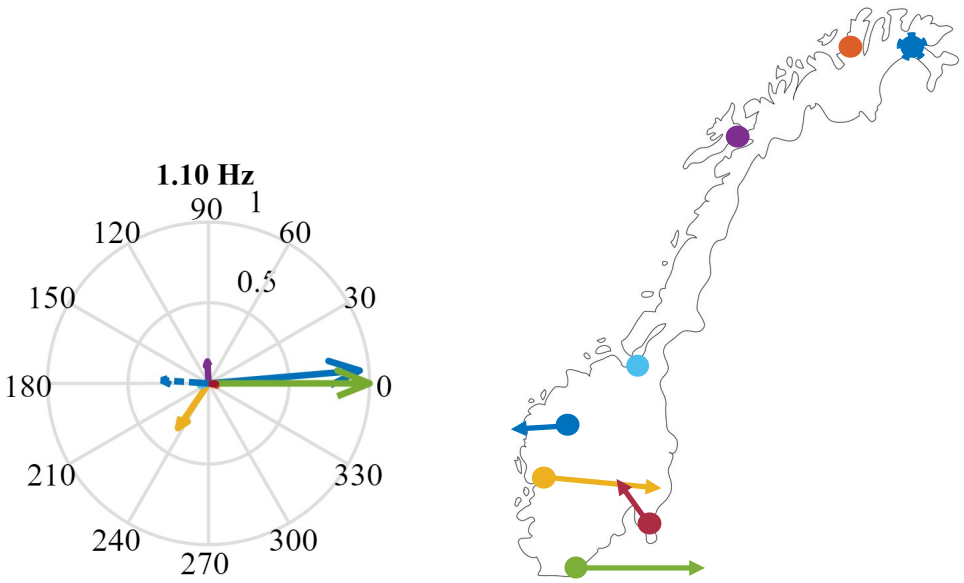
as low-damped ones. It is only the critical information that needs to be on display.

In the proposed oscillation monitoring method in this thesis, the visualization method used has been compass plots and text to display the frequency, damping and mode shapes. A next step of the algorithm can be to implement a GIS interface with the critical modes represented on a map. In this case, the 1 Hz mode from Case 3 would look something like in Fig. 6.2 (b). Displaying the mode shapes with use of GIS may, as previously argued, make mode shapes more intuitive to the operators.

Developing appropriate oscillation monitoring and control tools will likely be an iterative process where control room operators along with scientists and developers test out new functions to see how functional they are in the control room. One cannot do something about a problem one cannot see, therefore the initial step is to make operators aware of when, where and how oscillations are present in the power system.

This can be done through workshops on oscillations where the developers learn about the important factors of visualization in control room applications, and operators improve the level of knowledge of oscillations. In the beginning it might be a good introduction to the topic to use intuitive visualization tools without too much focus on the numerical aspects such as 2D and 3D GIS to demonstrate how the oscillations behave in the grid. When operators are more comfortable with oscillation frequency, damping, excitation and mode shapes, it is a good idea to use visualization methods that focus on visualizing the numerical aspects of the modal parameters.

A final remark is that there seems to be potential for adopting algorithms devel-



(a) 1.10 Hz mode shapes estimated with N4SID+DBSCAN.

(b) Illustration of GIS and phasor visualization of the estimated inter-area mode at 1.10 Hz.

Figure 6.2

oped in other fields than power system engineering for oscillation monitoring. This has already been done with for example FDD which originated from civil engineering to monitor vibrations in constructions, and there might be potential for more interdisciplinary learning on the topic [3]. Oscillation monitoring becomes more and more important as synchronous machines take a smaller share of the generation, and contributions from other areas are very welcomed.

Chapter 7

Conclusion

This thesis has explored oscillation monitoring with regards to a system operator's perspective. Creating good oscillation monitoring tools is a two-sided task. Firstly, it is important to create appropriate visualization tools for the user and the area of application. Secondly, creating reliable algorithms that provide the right information to the visualization tool is essential to the success of the monitoring tool. When designing visualization tools, it is important to keep in mind what they will be used for. Intuitive visualization methods without numerical information, such as 2D and 3D GIS, might be used for educational attainment in the initial stages. Afterward, a numerical representation of the modal properties should be in focus. GIS and phasors may give an intuition of where oscillations are present in the system. A 3D spectral analysis as displayed in Fig.6.1 may prove a good visualization tool for excitation of critical modes, with the potential to include damping as color coordinates. GIS with phasor representation might give the operators a good overview of where the oscillations are located in the system. Accurate and fast algorithms must be the basis of these visualization methods.

In this thesis, an algorithm called N4SID+DBSCAN has been presented. The applied algorithm based on N4SID has been implemented, and its weakness in choosing model order is turned into its strength through clustering with DBSCAN. The algorithm finds the frequency, damping and mode shape of modes in a signal, and the output of the algorithm is visualized as compass plots and text. The algorithm has been tested on simulated and real PMU data, and proved to be a promising tool for oscillation monitoring. The algorithm works on many types of input from the simulated power system with relatively good accuracy. It detected new modes after a change in topology within only 5 seconds after the event. When testing the algorithm on real PMU data, modes and mode shapes found were coherent with earlier work on the data set, signifying that the algorithm can be suitable for system monitoring.

The biggest drawback of the algorithm is the long computation time, which was

about 46 seconds when using 4 data streams and 80 seconds of data. There is a compromise between the amount of input data and the time the algorithm takes to estimate modes and mode shapes. In a control room setting, it is important that the data is up to date. Therefore, further work should be focused on optimizing the speed of the algorithm. Further, It struggles to separate modes that are close in frequency and damping. This might lead to the algorithm not detecting separate low-damped modes when they are close to each other.

Bibliography

[1] G. Rogers, *Power Systems Oscillations*, 1st ed. New York, New York: Springer Science + Business Media, 1999.

[2] J. Machowski, Z. Lubosny, J. W. Bialek, and J. R. Bumby, *Power Systems Dynamics Stability and Control*, 3rd ed. Hoboken, New Jersey: John Wiley & Sons, 2020.

[3] J. Thambirajah, E. Barocio, and N. Thornhill, "Comparative review of methods for stability monitoring in electrical power systems and vibrating structures," *IET Generation, Transmission and Distribution*, vol. 4, no. 10, pp. 1086–1103, 2010.

[4] M. Bakke, "Subspace identification using closed-loop data," 2009.

[5] P. Kundur, *Power System Stability and Control*. Cambridge MA: McGraw-Hill Inc., 1994.

[6] CIGRE, "Wide area monitoring systems - support for control room applications," Technical brochure, 2018.

[7] IEEE, "Virtual event - international conference on smart grid synchronized measurements and analytics (sgsma)," https://www.sgsma2021.org/, May 2021.

[8] J. Seppänen, "Methods fr monitoring electromechanical oscillations in power systems," Aalto University Doctoral Dissertation, 2017.

[9] H. Haugdal and K. Uhlen, "Power oscillation monitoring using statistical learning methods," in *Proc. of the IEEE Milan PowerTech Conference*. Milan, Italy: IEEE, 2019.

[10] POSOCO, "Report on low frequency oscillation in indian power system," https://posoco.in/download/report-on-low-frequency-oscillation-in-indian-power-system-march-2016/?wpdmdl=523, 2016.

[11] "Believes in european power system without emissions | statnett," https://www.statnett.no/en/about-statnett/news-and-press-releases/news-archive-2020/believes-in-european-power-system-without-emissions/, (Accessed on 10/04/2021).

[12] J. Hauer, D. Trudnowski, G. Rogers, B. Mittelstadt, W. Litzenberger, and J. Johnson, "Keeping an eye on power system dynamics," *IEEE Computer Applications in Power*, vol. 10, no. 4, pp. 50–54, 1997.

[13] M. Chenine, K. Zhu, and L. Nordström, "Survey on priorities and communication requirements for pmu-based applications in the nordic region," in *Proc. of the IEEE Bucharest PowerTech Conference*. Bucharest, Romania: IEEE, 2009, pp. 1–8.

[14] H. Kuisti, M. Lahtinen, M. Nilsson, K. Eketorp, E. Ørum, D. Whitley, A. Slotsvik, and A. Jansson, "Nag - frequency quality report," 4 2015.

[15] A. G. Phadke and T. Bi, "Phasor measurement units, wams, and their applcations in protection and control of power systems," *Journal of Modern Power Systems and Clean Energy*, vol. 6, pp. 619–629, 2018.

[16] A. B. Leirbukt, J. O. Gjerde, P. Korba, K. Uhlen, L. K. Vormedal, and L. Warland, "Wide area monitoring experiences in norway," in *Proc. of the IEEE International Symposium on Power Electronics, Electrical Drives, Automation and Motion*. Amalfi, Italy: IEEE, 2018.

[17] D. Örs and M. Omtveit, "Interview with olve mogstad, senior engineer in statnett fault analysis team," July 2021.

[18] J. Paserba, "Analysis and control of power system oscillations," *IFAC Proceeding Volumes*, vol. 30, no. 17, pp. 75–83, 1997.

[19] M. Ghorbaniparvar, "Survey on forced oscillations in power system," *Journal of Modern Power Systems and Clean Energy*, vol. 5, pp. 671–682, 2017.

[20] T. D. Duong and S. D'Arco, "Locating generators causing forced oscillations based on system identification," in *Proc. of the IEEE PES Innovative Smart Grid Technologies Europe*. The Hague, Netherlands: IEEE, 2020.

[21] B. Wang and K. Sung, "Location methods of oscillation sources in power systems: A survey," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 2, pp. 151–159, 2017.

[22] B. Pal and B. Chaudhuri, *Robust Control in Power Systems*, 1st ed. New York, New York: Springer Science + Business Media, 2005.

[23] D. J. Trudnowski and J. E. Dagle, "Effects of generator and static-load nonlinearities on electromechanical oscillations," *IEEE Transactions on Power Systems*, vol. 12, no. 3, pp. 1283–1289, 1997.

[24]

[25] M. Omtveit, "Interview with birger bjørland, system operator at statnett national control centre," 8 2021.

[26] L. Ljung, *System Identification Toolbox: User's Guide*. Natick, Massachusetts: MathWorks, 1995.

[27] N. Zhou, Z-Huang, L. Dosiek, D. Trudnowski, and J. W. Pierre, "Electromechanical mode shape estimation based on transfer function identification using pmu measurements," in *IEEE Power & Energy Society General Meeting*. Calgary, Canada: IEEE, 2009.

[28] L. Dosiek, N. Zhou, J. W. Pierre, Z. Huang, and D. J. Trudnowski, "Mode shape estimation algorithms under ambient conditions: A comparative review," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 779–787, 2013.

[29] J. Turunen, J. Thambirajah, M. Larsson, B. C. Pal, N. F. Thornhill, L. C. Haarla, W. W. Hung, A. M. Carter, and T. Rauhala, "Comparison of three electromechanical oscillation damping estimation methods," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2398–2407, 2011.

[30] B. Gopinath, "On the identification of linear time-invariant systems from input-output data," *The Bell System Technical Journal*, vol. 48, no. 5, pp. 1101–1113, 1969.

[31] E. Kreyszig, *Advanced Engineering Mathematics*, 10th ed. Hoboken, New Jersey: John Wiley & Sons, 2011.

[32] P. V. Overschee and B. D. Moor, "N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica, Special Issue on Statistical Processing and Control*, vol. 30, no. 1, pp. 75–93, 1994.

[33] K. de Cock and B. de Moor, "Analysis and damping of interarea oscillations in the ucte/centrel power system," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.

[34] H. H. Asada, "Lecture notes in identification, estimation and learning," 2020.

[35] H. Haugdal and K. Uhlen, "Mode shape estimation using complex principal component analysis and k-means clustering," in *Proc. of the International Conference on Smart Grid Synchronized Measurements and Analytics*. College Station, USA: IEEE, 2019.

[36] S. Føyen and M.-E. Kvammen, "A signal analysis toolbox for power system identification in smart grids," 2018.

[37] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[38] Z. Yang, H. Liu, T. Bi, Z. Li, and Q. Yang, "An adaptive pmu missing data recovery method," *International Journal of Electrical Power Energy Systems*, vol. 116, 2020.

[39] J. H. Chow, H. Lev-Ari, M. Weng, S. Konstantinopoulos, and G. M. D. Mijolla, "Synchrophasor missing data recovery via data-driven filtering," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4321–4330, 2020.

[40] L. Ljung, *System Identification Toolbox: Theory for the User*. Upper Saddle River, New Jersey: Prentice Hall PTR, 1999.

[41] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A density based algorithm for discovering clusters in large spatial databases with noise," in *Proc. of the 2nd International Conference on Discovery and Data Mining (KDD'96)*. Portland, Oregon: AAAI, 1996, pp. 226–231.

[42] H. Haugdal, "Dynpssimpy," https://github.com/hallvar-h/DynPSSimPy, 2021.

[43] J. Khodaparast, O. Fosso, and M. Molinas, "Emd-prony for phasor estimation in harmonic and noisy condition," in *Proc. of the IEEE International Symposium of Power Electronics, Electrical Drives, Automation and Motion*. Amalfi, Italy: IEEE, 2018.

[44] M. Klein, G. J. Rogers, and P. Kundur, "A fundamental study of inter-area oscillations in power systems," *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 914–921, 1991.

[45] K. Uhlen, L. Vanfretti, M. M. de Oliveira, A. B. Leirbukt, V. H. Aarstrand, and J. O. Gjerde, "Wide-area power oscillation damper implementation and testing in the norwegian transmission network," in *Proc. of the IEEE Power and Energy Society General Meeting*. San Diego, California: IEEE, 2012.

Appendices

A Literature Review: Oscillation Monitoring Tools

Color Coded Text

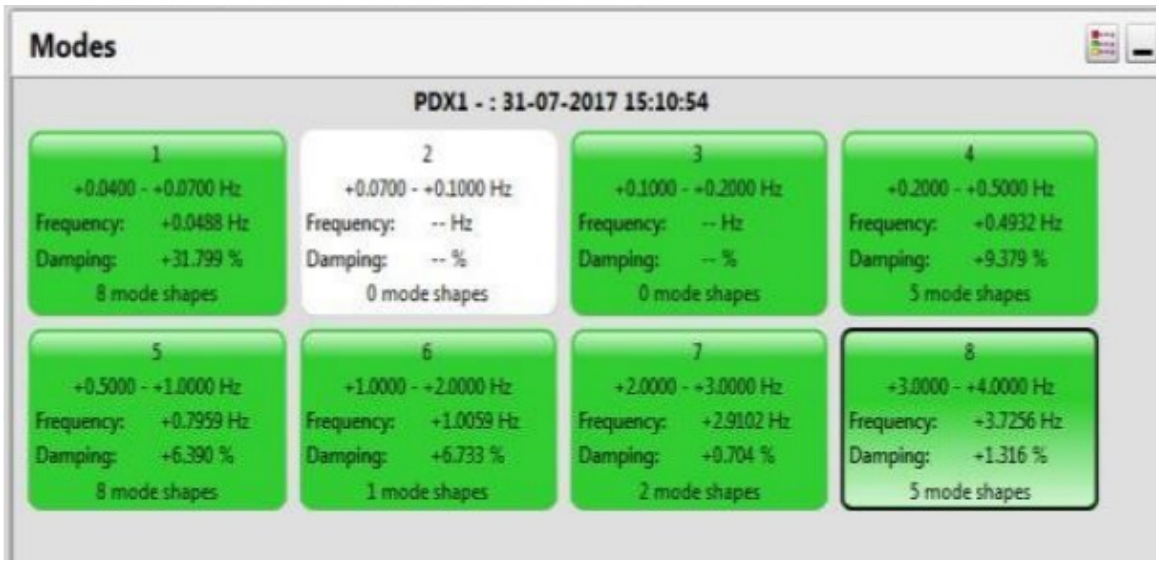


Figure 1: Color coded boxes that represents modes [6]. The information displayed is the frequency and damping of the mode as well as the number of participating mode shapes.

GIS

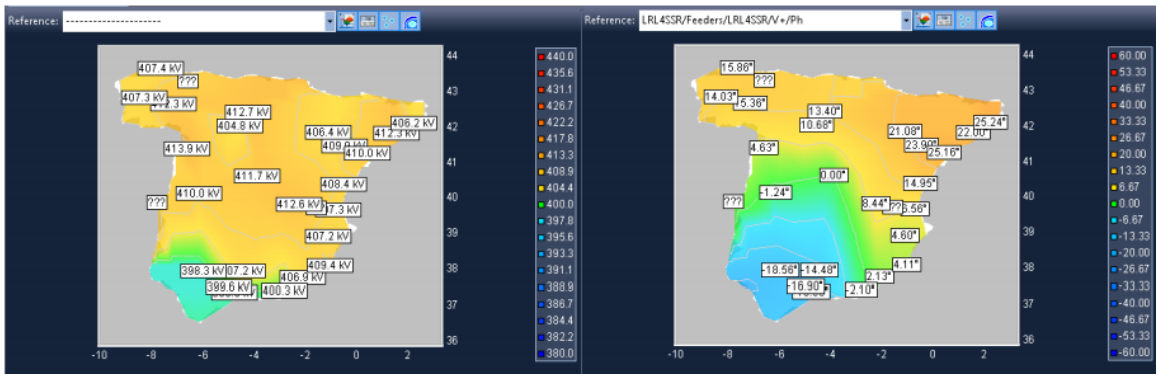
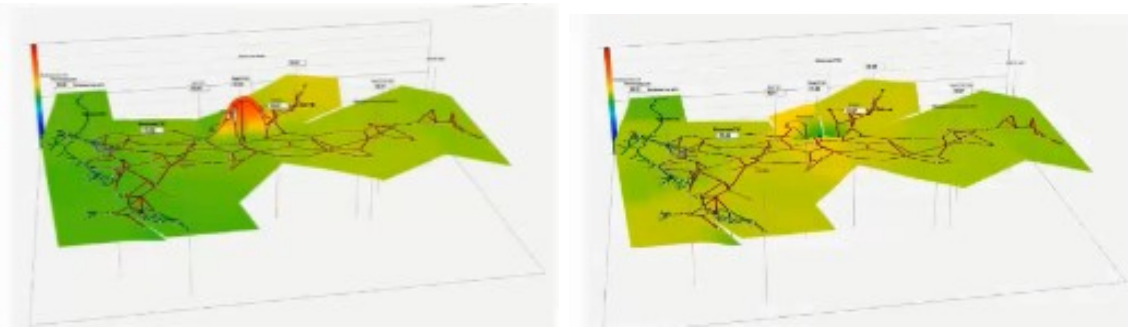


Figure 2: Two-dimensional GIS tool by the Spanish TSO [6]. Oscillations will be observed as pulsating colors.



(a) (b)

Figure 3: Three-dimensional GIS tool demonstrated at the IEEE SGSM conference in 2021 [7]. Observe the oscillatory mode in the middle of the map.

Mode Damping Plots

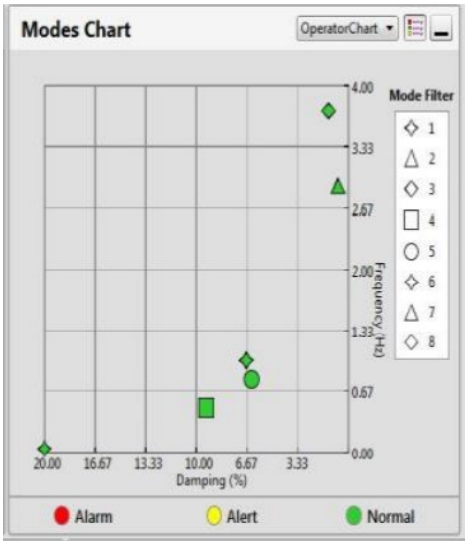


Figure 4: Modes chart that shows the modes on a frequency-damping plot [6]. The shapes refer to the specific mode number, which is used consequently in other visualization methods in the same tool.

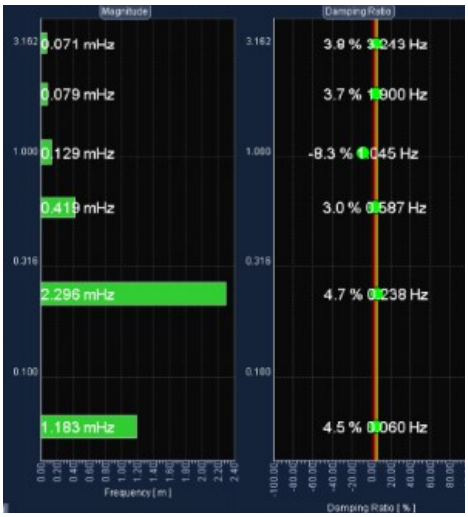


Figure 5: A close-up of the damping and excitation monitoring of Fig 8 [6].

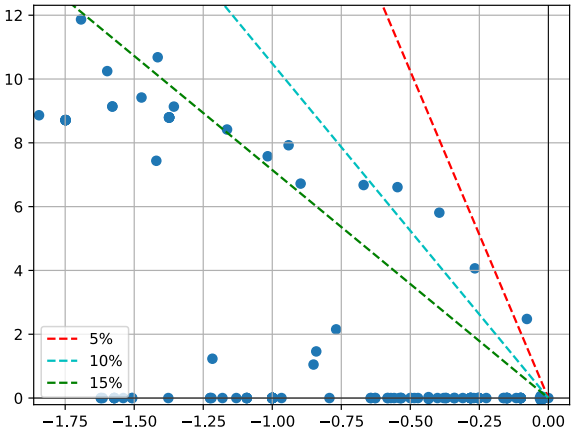


Figure 6: Eigenvalue plot created in DynPSSimPy using the n44 model with damping lines displayed.

Spectrogram

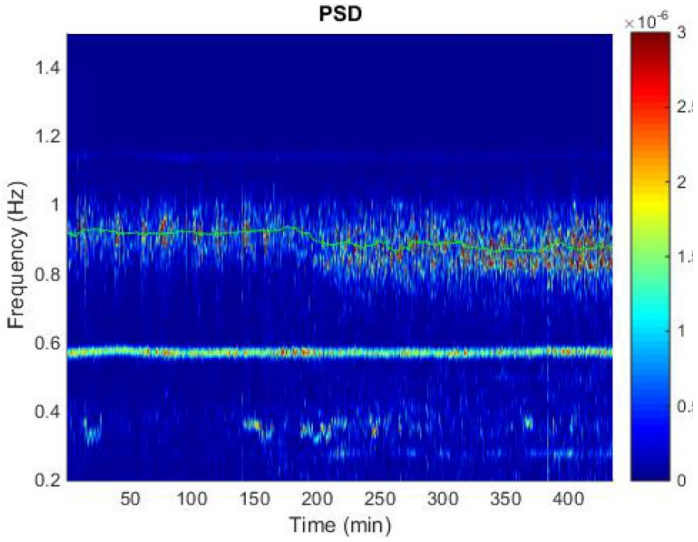


Figure 7: The spectrogram shows the magnitude of excitation in colors. In this spectrogram, there is electromechanical oscillation at 0.8-0.9 Hz and a forced oscillation at 0.6 Hz [8].

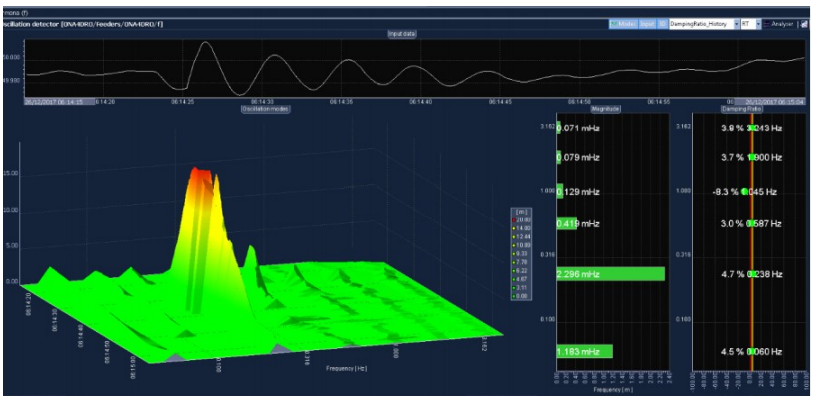


Figure 8: The three-dimensional spectrogram shows the same information as the two-dimensional one, but with a third dimension to display the magnitude [6]. In this application, the spectrogram is complemented with a graph on damping and a graph on exact magnitude of the mode.

Mode Shape Plots

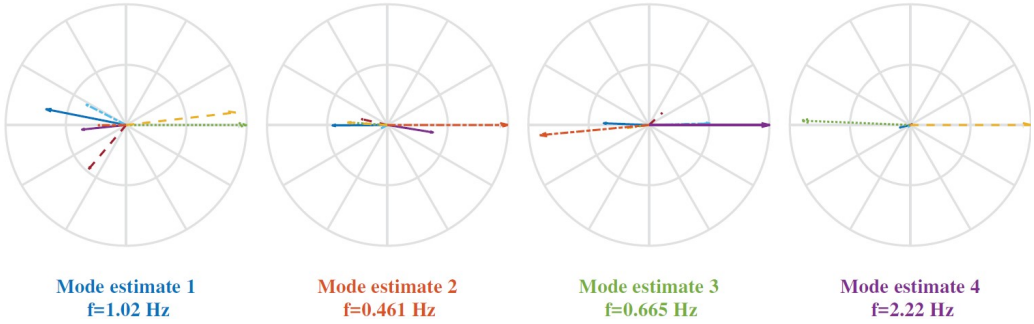


Figure 9: Compass plots as presented in a paper about using statistical learning methods for mode identification [9].

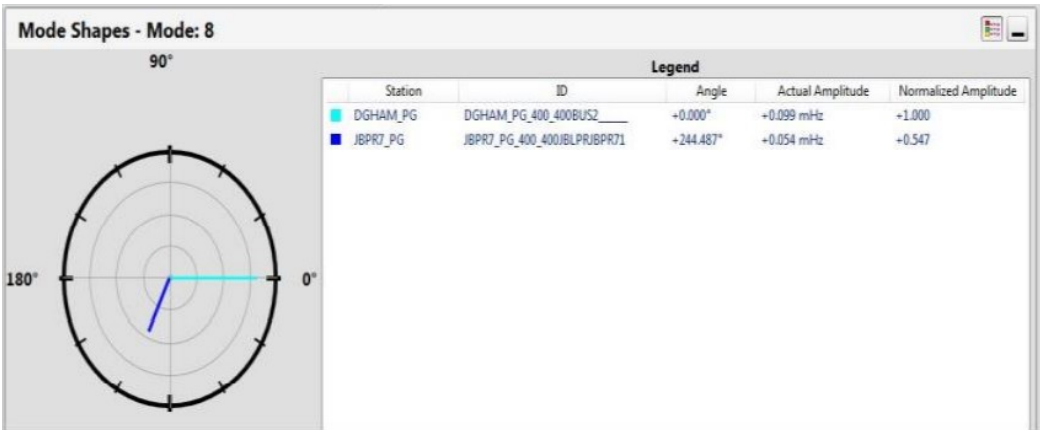


Figure 10: Compass plot with information presented as text about the phase shift, magnitude and normalized magnitude of oscillation [6].

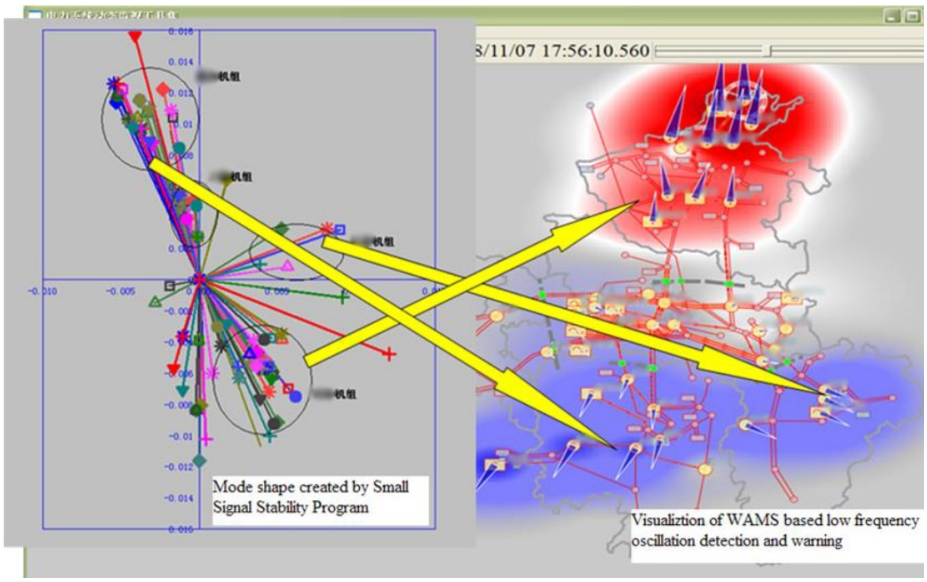


Figure 11: Mode shapes presented as phasors which is connected with the combined GIS and mode shape visualization [6]. Each cluster in the phasor plot is connected to an area participating in the inter-area mode.

Combined GIS and Mode Shapes

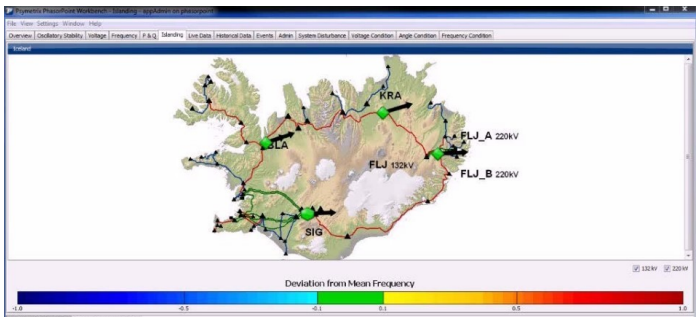


Figure 12: Phasors on GIS of Iceland with color displaying deviation from mean frequency in this case [6]. In an oscillation monitoring case, the phasors would be the mode shapes, and the colour could be connected to damping.

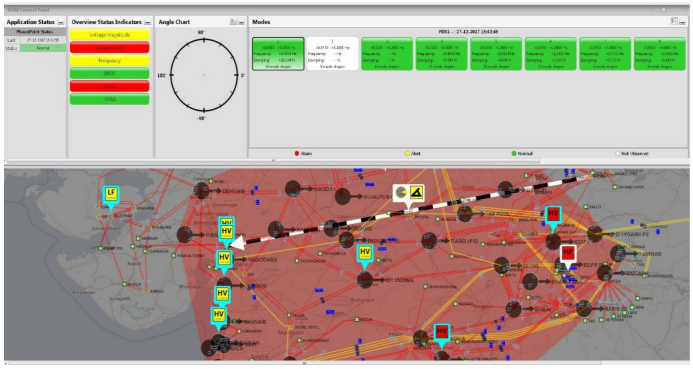


Figure 13: Indian monitoring tool displaying phasors as well as other stability indicators on the GIS [6].

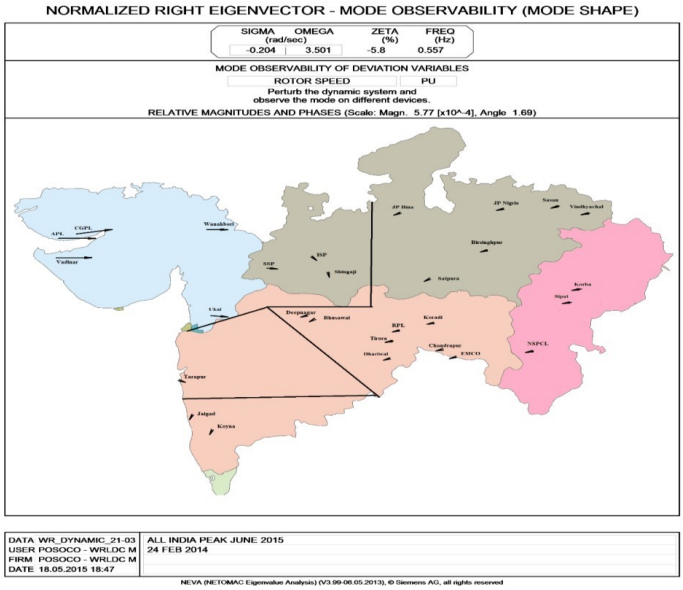


Figure 14: Specific visualization of the mode shape of a poorly damped mode in the western region of the Indian power grid [10]. The arrowsdisplay the relative magnitude and phase of the mode shape, and the lines display where the grid is most likely to separate into islands.

B Visualization Methods Categorized

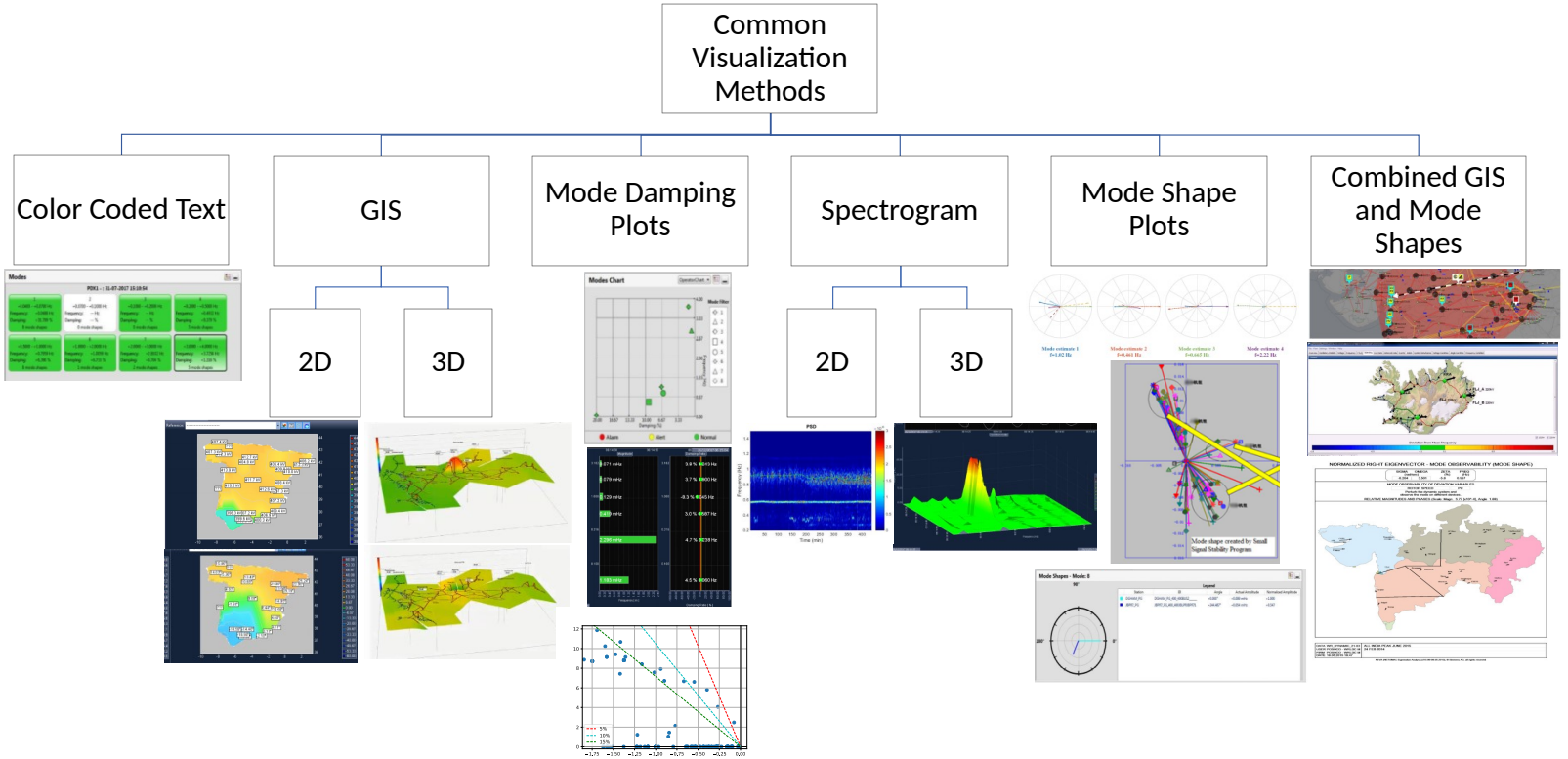


Figure 15: Categorization of visualization methods.

C Quick review of Linear System Theory

Impact of Impulse Response on State Equation

Given the system state space model

$$x(t+1) = Ax(t) + Bu(t) \tag{1}$$
$$y(t) = Cx(t) + Du(t) \tag{2}$$

we have that expanding the terms of $y(t)$ gives an expression solely dependent on the initial state variables $x(0)$ and the past control input u.

$$\begin{aligned} y(t) &= Cx(t) + Du(t) \\ &= C(Ax(t-1) + Bu(t-1)) + Du(t) \\ &= C(A^2x(t-2) + ABu(t-2)^+Bu(t-1)) + Du(t) \\ &\vdots \\ &= CA^tx(0) + \sum_{i=1}^{i=t} A^{i-1}Bu(t-i) + Du(t) \end{aligned}$$

The impulse response is defined as

$$x(0) = 0 \tag{3}$$
$$u(t) = \begin{cases} k, & \text{for } t = 0 \\ 0, & \text{for } t > 0 \end{cases} \tag{4}$$

As the only input happens in $t = 0$, the rest of the time series will depend on the initial values of the input $u(0)$. The output response to the impulse input is

$$y(t) = \begin{cases} Du(0), & \text{for } t = 0 \\ CA^{t-1}Bu(0), & \text{for } t > 0 \end{cases} \tag{5}$$

Figure 16: Realization.

Let G(0), G(1), G(2), ... be an infinitely large impulse response coefficient matrices of the system that goes back to the beginning of the system

$$y(t) = \sum_{k=0}^{\inf} G(k)u(t-k) \tag{6}$$

such that

$$y(t) = G(t)u(t) \tag{7}$$

By comparison of Equation 5 and 7, the expression for $G(t)$ is found to be

$$G(t) = \begin{cases} D, & \text{for } t = 0 \\ CA^{t-1}B, & \text{for } t > 0 \end{cases} \tag{8}$$

where the difference from $y(t)$ is that the input term is eliminated, giving $G(t)$ the role as the transfer function of the system.

Observability

The observability of the system describes how much each state variable participates, or is observable, in the output signal. The observability matrix describes the linear transform from the state matrix to the output matrix. The inverse observability matrix can therefore be used to find the state variables when the output is given. Assuming no input signal for all t, the output sequence from $t = 0$ to $t = k - 1$ is given by

$$\begin{aligned} y(0) &= Cx(0) \\ y(1) &= Cx(1) = CAx(0) \\ y(2) &= CA^2x(0) \\ &\vdots \\ y(k-1) &= CA^{k-1}x(0) \end{aligned}$$

which written on matrix form is

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(k-1) \end{bmatrix} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}}_{\mathcal{O}_k} x(0) = \mathcal{O}_k x(0) \tag{9}$$

where $\mathcal{O}_k$ is called the extended observability matrix, and $\mathcal{O}_k \in \mathbb{R}^{kp \times n}$. The time interval stops at $k-1$ such that the dimension of the extended observability matrix is dependent on k and not $k+1$. Recall that the dimension of the output vector i p and the dimension of the state vector, also known as the order of the system, is n.

Note that the initial state $x(0)$ is determined only if $\mathcal{O}_k$ is of full rank. Recall that a matrix of full rank is a matrix where the rows are linearly independent, i.e. it is not possible to write any row as a linear combination of the other rows.

In the context of system identification, the system order n is often unknown. Therefore, k is set to be strictly larger than n.

If $p = 1$, $k = n$ and $\mathcal{O}_k$ is non singular (the inverse matrix exists), then the state vector is uniquely determined by

$$x(0) = \mathcal{O}_n^{-1} \begin{bmatrix} y(0) \\ \vdots \\ y(n-1) \end{bmatrix} \tag{10}$$

Reachability and Controllability

The reachability of a system determines the way in which the initial states are still a part of the output expression for any time t. Consider a discrete-time LTI system of order n. If the initial state $x(0) = 0$ can be transferred to any state at time n by means of a sequence of input, $u(0), u(1), ..., u(n-1)$, such that

$$x(n) = A^n x(0) + A^{n-1}Bu(0) + A^{n-2}Bu(1) + ... + Bu(n-1) \tag{11}$$

$$= A^n x(0) + \underbrace{\begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix}}_{\mathscr{C}} \begin{bmatrix} u(n-1) \\ u(n-2) \\ \vdots \\ u(0) \end{bmatrix} \tag{12}$$

then the system is reachable, and the reachability matrix $\mathscr{C} \in \mathbb{R}^{n \times nm}$. If the reachability matrix is of full column rank, then a sequence of inputs exists that brings the initial state to an arbitrary state in n steps.

If A is non-singular (invertible), then the system is not only reachable, but also controllable. A system being controllable means that it is possible to go from an arbitrary state back to the origin.

Figure 17: Reachability refers to the ability to go from initial state to an arbitrary state by means of a sequence of control inputs and the controllability matrix. Controllability refers to the ability to go from an arbitrary state to the initial state, and is only possible if A is invertible.

When the order of the system is unknown, the extended reachability matrix $\mathscr{C}_k$ is used, and $\mathscr{C}_k \in \mathbb{R}^{k \times km}$ where $k > n$.

D Geometric Tools

Tools from the mathematical branch of geometry is useful in the subspace identification framework. Projections is the basis of least squares linear regression in higher dimensions, which is used in N4SID. The reason linear regression is used is that often the matrix equations, e.g. $Ax = b$ do not have a solution because there are more equations than there are state variables, and consequently A is not square. In order to reduce the matrix A to a square matrix which is solvable, one can project the excess rows in the matrix into the first square part of the matrix. Firstly, projections will be introduced in one and two dimensions, and then the general equations for orthogonal and oblique projection will follow.

Projection onto a Subspace

The orthogonal projection of a vector $v \in \mathbb{R}^n$ is its transformation to a linear combination of the n basis vectors spanning the n-dimensional space. Perhaps the simplest form of projection is to project a two-dimensional vector p onto a line a, as in Fig 18.

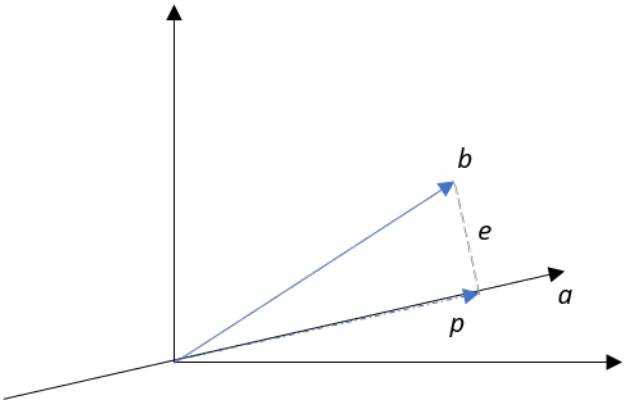


Figure 18: The two-dimensional vector b projected onto the one-dimensional space as p.

Let the line going perpendicularly from b onto a be defined as

$$e = b - p \tag{13}$$

and let the projection be a constant x times the direction of the line a it is projected onto such that

$$p = xa \tag{14}$$

The key in order to have the line e go from b perpendicularly onto line a is to make sure the dot product between a and the e is zero

$$a^T e = a^T(b - p) = a^T(b - xa) = 0 \tag{15}$$

Rearranging this equation gives

$$x = \frac{a^T b}{a^T a} \tag{16}$$

which by substituting equation 14 renders

$$p = a\frac{a^T b}{a^T a} \tag{17}$$

where the matrix P that projects b onto a is given by

$$P = \frac{aa^T}{a^T a} \tag{18}$$

As a is not a square matrix, $aa^T \neq a^T a$, and therefore P is not the identity matrix, even though it might seem like it at first sight.

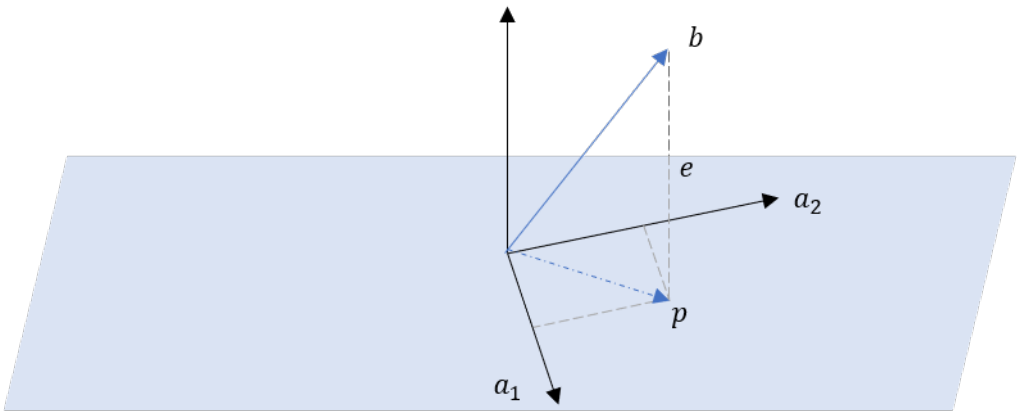


Figure 19: The 3-dimensional vector b projected onto the 2-dimensional space as p.

Now let's look at the projection of a matrix onto a subspace, which is what is used in subspace methods described in this master thesis. Let a three-dimensional vector b be projected onto a two-dimensional plane consisting of a_1 and a_2. Let the plane of a_1 and a_2 be the column space of A, such that

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \tag{19}$$

The projection of b on the plane is a linear combination of the directions a_1 and a_2 with corresponding coefficients such that

$$p = \hat{x_1}a_1 + \hat{x_2}a_2 = A\hat{x} \tag{20}$$

In order to project b onto A, let the line e be the line to the plane from b to p

$$e = b - p = b - A\hat{x} \tag{21}$$

In order for it to be perpendicular to the line, the dot product condition must be met, which can be described by two equations,

$$a_1^T(b - A\hat{x}) = 0 \tag{22}$$
$$a_2^T(b - A\hat{x}) = 0 \tag{23}$$

which on matrix form is expressed equivalently as

$$A^T(b - \hat{x}) = 0 \tag{24}$$

Rearranging the equation gives

$$\hat{x} = (A^TA)^{-1}A^Tb \tag{25}$$

which gives as an expression for the projection

$$p = A\hat{x} = A(A^TA)^{-1}A^Tb \tag{26}$$

where the projection transform matrix P is given by

$$P = A(A^TA)^{-1}A^T \tag{27}$$

Orthogonal Projection

In the following sections it is assumed that three dummy matrices, A, B and C are given with the following dimensions

$$A \in \mathbb{R}^{p\times j}, B \in \mathbb{R}^{q\times j}, C \in \mathbb{R}^{r\times j}$$

Mark that all matrices have j columns. The elements of a row can thus be considered as a vector on the j-dimensional ambient space. It is assumed that $j \geq max(p, q, r)$, which is always the case in identification algorithms.

For any dimension, the general orthogonal projection of the row space of A onto the row space of B is denoted A/B and is on matrix form described by

$$A/B = AB^T(BB^T)^{\dagger}B = A\Pi_B \tag{28}$$

where $(.)^\dagger$ denotes the Moore-Penrose pseudo-inverse of the matrix, which is the generalized inverse of a matrix that can be computed using the singular value decomposition. It is possible to project A onto the orthogonal complement of the row spaces in B, called $B^\perp$

$$A/B^\perp = A - A/B = A(I_j - B(BB^T)^\dagger B) = A\Pi_{B^\perp} \tag{29}$$

The two projections Π_B and $\Pi_{B^\perp}$ decompose A into two matrices with orthogonal row spaces

$$A = A\Pi_B + A\Pi_{B^\perp} \tag{30}$$

and can be computed via the LQ decomposition of $\begin{bmatrix} B \\ A \end{bmatrix}$, as proposed by [33], and briefly described in section F. Finally, the matrix representations of the orthogonal projections are given as

$$A/B = L_{21}Q_1^T \tag{31}$$

$$A/B^\perp = L_{22}Q_2^T \tag{32}$$

Oblique Projection

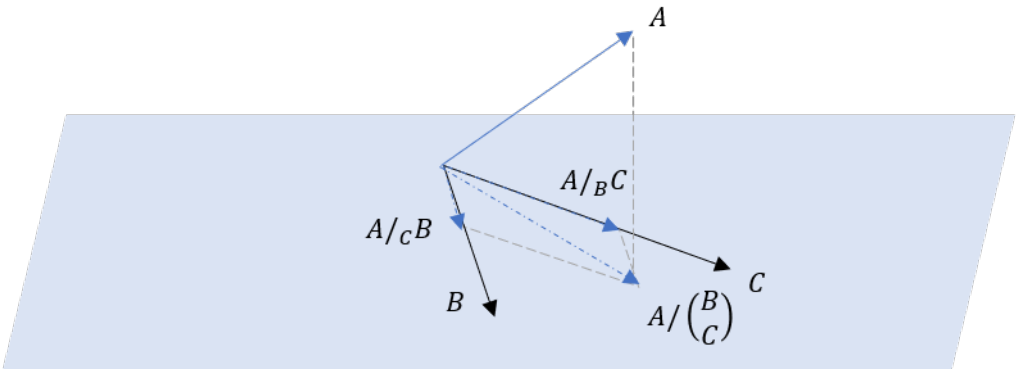


Figure 20: Oblique projection of A onto the subspace of B and C where $j = 3$, and $p = q = r = 1$.

In a system of linear equations, the parameter matrix may not happen to be orthogonal. In these cases one can use oblique projection. The rows of A can be decomposed as a linear combination of the rows of two non-orthogonal matrices B and C, and of their orthogonal complement, denoted by

$$A = L_B B + L_C C + L_{B^\perp, C^\perp} \begin{bmatrix} B \\ C \end{bmatrix}^\perp \tag{33}$$

where the oblique projection of the row space of A along the row space of B into the row space of C is denoted by

$$A/_B C = L_C C \tag{34}$$

Given the LQ decomposition of the matrices to be

$$\begin{bmatrix} B \\ C \\ A \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix} \tag{35}$$

the orthogonal projection of A onto the joint row space of B and C is denoted by

$$A/\begin{bmatrix} B \\ C \end{bmatrix} = \begin{bmatrix} L_{31} & L_{32} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \tag{36}$$

which can also be expressed as a linear combination of the rows of B and C

$$A/\begin{bmatrix} B \\ C \end{bmatrix} = L_B B + L_C C = \begin{bmatrix} L_B & L_C \end{bmatrix} \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \tag{37}$$

Comparing equation 36 and 37 gives

$$\begin{bmatrix} L_{31} & L_{32} \end{bmatrix} = \begin{bmatrix} L_B & L_C \end{bmatrix} \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \tag{38}$$

and finally, the oblique projection of the row space of A along the row space of B onto the row space of C is given by

$$A/_B C = L_C C = L_{32} L_{22}^{-1} \begin{bmatrix} L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \tag{39}$$

E Matrices

In subspace methods, the properties of certain matrices are used for computation. The following sections present central matrix types and their properties.

Symmetric Matrices

The Laplace matrix is a symmetric matrix. A symmetric matrix is a matrix in which the elements are mirrored across the diagonal. These matrices have particular properties when it comes to eigenvectors and eigenvalues. An n $\times$ n symmetrical matrix, A will have the following properties:

- A has exactly n (not necessarily distinct) eigenvalues
- There exists a set of n eigenvectors, one for each eigenvalue, that are mutually orthogonal

Toeplitz Matrix

A Toeplitz matrix is a rectangular matrix where each descending diagonal from left to right is constant. For example

$$\begin{bmatrix} a & b & c & d \\ e & a & b & c \\ f & e & a & b \\ g & f & e & a \\ h & g & f & e \end{bmatrix}$$

By using a Toeplitz matrix, two discrete signals can be convoluted by converting one signal to a Toeplitx matrix and taking the matrix multiplication

$$y = f * g = \begin{bmatrix} f_1 & 0 & & 0 & 0 \\ f_2 & f_1 & \ddots & 0 & 0 \\ f_3 & f_2 & \ddots & 0 & 0 \\ \vdots & f_3 & \ddots & f_1 & 0 \\ f_{m-1} & \vdots & \ddots & f_2 & f_1 \\ f_m & f_{m-1} & \ddots & f_3 & f_2 \\ 0 & f_m & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & f_{m-1} & f_{m-2} \\ \vdots & \vdots & \ddots & f_m & f_{m-1} \\ 0 & 0 & \ddots & 0 & f_m \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \\ \vdots \\ g_n \end{bmatrix} \tag{40}$$

Hankel Matrix

The Hankel matrix is a square matrix in which each descending diagonal from right to left is constant. For example

$$\begin{bmatrix} a & b & c & d \\ b & c & d & e \\ c & d & e & f \\ d & e & f & g \end{bmatrix}$$

If the entries in the Hankel matrix are matrices themselves, it is called a *block Hankel* matrix.

Let J_n be the $n \times n$ exchange matrix. if H is a $m \times n$ Hankel matrix, then $H = TJ_n$ where T is a $m \times n$ Toeplitz matrix. If T is real symmetric, then $H = TJ_n$ will have the same eigenvalues as T.

For subspace methods, a particular Hankel matrix is formed of the output data, divided into past and future sub matrices, see [33]. The input and output is written in the following Hankel matrices where the subscripts denotes the past and future matrices.

$$U_{0|2i-1} = \begin{pmatrix} U_p \\ \hline U_f \end{pmatrix} = \left[\begin{array}{cccc} u(0) & u(1) & \dots & u(N-1) \\ u(1) & u(2) & & u(N) \\ \vdots & & \ddots & \vdots \\ u(k-1) & u(k) & \dots & u(k+N-2) \\ \hline u(k) & u(k+1) & \dots & u(k+N-1) \\ u(k+1) & u(k+2) & & u(k+N) \\ \vdots & & \ddots & \vdots \\ u(2k-1) & u(2k) & \dots & u(k+N-1) \end{array}\right] \tag{41}$$

$$Y_{0|2i-1} = \begin{pmatrix} Y_p \\ \hline Y_f \end{pmatrix} = \left[\begin{array}{cccc} y(0) & y(1) & \dots & y(N-1) \\ y(1) & y(2) & & y(N) \\ \vdots & & \ddots & \vdots \\ y(k-1) & y(k) & \dots & y(k+N-2) \\ \hline y(k) & y(k+1) & \dots & y(k+N-1) \\ y(k+1) & y(k+2) & & y(k+N) \\ \vdots & & \ddots & \vdots \\ y(2k-1) & y(2k) & \dots & y(k+N-1) \end{array}\right] \tag{42}$$

Hankel Singular Values

Hankel singular values provide a measure of the energy of each state in a system.

Hilbert Matrix

The Hilbert matrix is a square Hankel matrix with the entries being the unit fractions. The 4×4 Hilbert matrix is for example

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{bmatrix}$$

F Decomposition of Matrices

LQ and QR Decomposition

The LQ decomposition of full rank matrices A and B is denoted by

$$\begin{bmatrix} A \\ B \end{bmatrix} = LQ^T = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} \tag{43}$$

where Q is a rectangular matrix where the columns are orthogonal, calculated for example through Gram-Schmidt's orthogonalization procedure or Householder Reflection [34]. L is lower triangular. Note that the orthogonal property of Q means that $QQ^T = Q^TQ = I$.

The LQ decomposition is the transpose of the RQ decomposition, where R is an upper triangular matrix.

Singular Value Decomposition

The singular value decomposition of a matrix, M is denoted as

$$M = U\Sigma V^T \tag{44}$$

where σ contains the vectors that to the largest degree describes the initial matrix, M.

The SVD of a Hankel matrix is a special case due to the symmetrical properties

$$H = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1\Sigma_1 V_1^T \tag{45}$$

where

$$\Sigma_1 = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

G Ho-Kalman's Method for System Realization

The basis of subspace methods is the Ho-Kalman method, where the state space system parameters (A, B, C and D) are obtained from impulse response coefficient matrices. The objective of the master thesis is to obtain these matrices from input-output data. However, it can be useful to look at the simpler case first.

Consider a LTI system that is observable and reachable (see C for definitions of observable and reachable), meaning that the matrices A, B and C are minimal.

Recall that the expression for the transfer function G for an impulse response is given as

$$G(t) = \begin{cases} D, & \text{for } t = 0 \\ CA^{t-1}B, & \text{for } t > 0 \end{cases} \tag{46}$$

which makes D fairly simple to obtain, as it is just the transfer function between the input and output in $t = 0$. Constructing a block Hankel matrix from the impulse response coefficient matrices from $t = 1$ to $t = 2k - 1$ yields

$$H = \begin{bmatrix} G_1 & G_2 & \dots & G_k \\ G_2 & G_3 & & G_{k+1} \\ \vdots & & \ddots & \vdots \\ G_k & G_{k+1} & \dots & G_{2k-1} \end{bmatrix} = \begin{bmatrix} CB & CAB & \dots & CA^{k-1}B \\ CAB & CA^2B & & CA^kB \\ \vdots & & \ddots & \vdots \\ G_k & G_{k+1} & \dots & G_{2k-1} \end{bmatrix} \tag{47}$$

A property of H is that it can be decomposed to the extended observability matrix and the extended reachability matrix [34].

$$H = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}}_{O_k} \underbrace{\begin{bmatrix} B & AB & \dots & A^{k-1}B \end{bmatrix}}_{C_k} \tag{48}$$

However, as the parameter matrices are unknown, O_k and C_k have to be found through the singular value decomposition.

The singular value decomposition of H yields

$$H = U\Sigma V^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_1 V_1^T \tag{49}$$

where

$$\Sigma_1 = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

and the eigenvalues are in decreasing order, such that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

The observability and reachability matrices can be constructed as

$$O_k = U_1 \Sigma_1^{1/2} \tag{50}$$

$$C_k = \Sigma_1^{1/2} V_1^T \tag{51}$$

where

$$\Sigma_1 = \Sigma_1^{1/2} \Sigma_1^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} \tag{52}$$

The minimal system matrices A, B and C can be determined from the observability and reachability matrices above. By inspection of equation 48, $C \in \mathbb{R}^{p \times n}$ can be found as the first p rows of O_k, and $B \in \mathbb{R}^{n \times m}$ as the first m columns of C_k. The $A \in \mathbb{R}^{n \times n}$ matrix is found by post-multiplying O_k with A, and pre-multiplying with O_k^T such that

$$A = (O_k^T O_k)^{-1} O_k^T O'_{k+1} \tag{53}$$

where

$$O'_{k+1} = O_{k+1}(p+1 : p(k+1), :) \in \mathbb{R}^{pk \times n}$$

Thus, all the minimal system matrices A, B, C and D are found.

H N4SID Algorithm: Step-by-Step

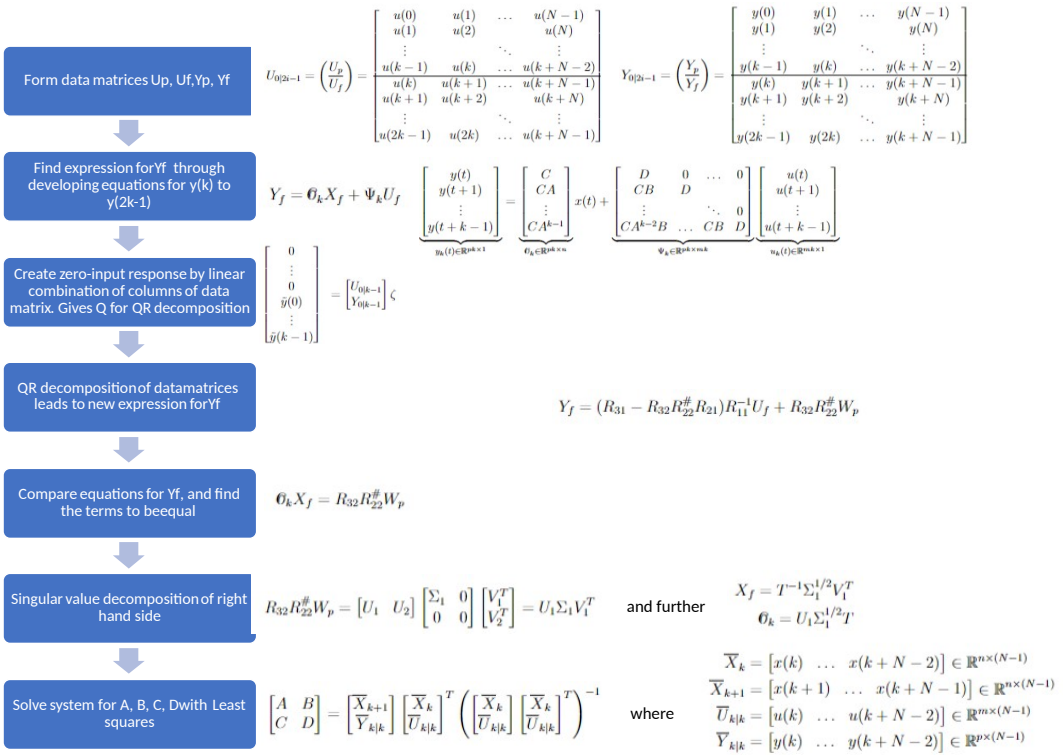


Figure 21: A stepwise instruction to the N4SID method.

Malken Borud Omtveit

Visualizing Electromechanical Oscillations for Control Room Purposes

NTNU
Norwegian University of
Science and Technology

Statnett
Fremtiden er elektrisk