AN ABSTRACT OF THE THESIS OF

Janhavi Kulkarni for the degree of Master of Science in Electrical and Computer Engineering presented on June 9, 2015.

Title: Rapid Grid State Estimation using Singular Value Decomposition.

Abstract approved:

Ted K. A. Brekken

Synchrophasor technology has gained great momentum with the widely adopted concept of smart grids in power systems at transmission and generation level. This has led to an improved state estimation of a power system with the advent of Phasor Measurement Units (PMU). PMUs contribute in providing real-time information about the grid state at a higher frequency than that has been historically available. This research aims to extend power system state estimation to a distribution level using the Oregon State University campus. Two state estimation techniques utilizing Singular Value Decomposition have been investigated in this research to estimate the state of the OSU power system with measurements from the sparsely deployed PMUs on the campus. The main objective is to build a state estimation technique for a system with incomplete observability based on static set of data compiled by valid power flow solutions and limited number of PMU measurements. Within this research both methods of estimating the state of the

grid are demonstrated on three different sized power grids comprising of 3 buses, 14 buses and 286 buses, respectively. The two methods - Similarity Matching and Filtering - have been discussed as potential methods to determine the state of a partially observable system at real-time or near real-time. Both the methods have been evaluated on the basis of computational speed and complexity as well as accuracy. The results obtained from both techniques of SVD are shown to have promising applications to rapid grid estimation at the distribution level of a system in which speed and sparseness are key.

Rapid Grid State Estimation using Singular Value Decomposition

by

Janhavi Kulkarni

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 9, 2015
Commencement June 2015

Master of Science thesis of Janhavi Kulkarni presented on June 9, 2015

APPROVED:

Major Professor, representing Electrical and Computer Engineering

Director of the School of Electrical Engineering and Computer Science

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Janhavi Kulkarni, Author

ACKNOWLEDGEMENTS

I would like to express my heartfelt and sincere gratitude to Dr. Ted Brekken, for his support and patience to make this research a success. His guidance as my academic advisor and confidence in my capabilities has always been my inspiration. I am extremely grateful to have had an opportunity to work with a professor with such great enthusiasm, knowledge and vision.

I would like to extend special thanks to all the professors who have taught me at Oregon State University. I also take this opportunity to thank Dr Eduardo Cotilla-Sanchez, Dr. Julia Zhang and Dr. Nathan Gibson for agreeing to be on my committee. Thank you to each and every fellow student of Energy Systems group both past and present for always being willing to provide suggestions and feedback.

Thank you Bonneville Power Administration (BPA) for making this research possible. I am extremely thankful to Anirban Roy for helping me through my difficulties. Special thanks to Vishvas Chalishazar for his constant motivation and moral support through the course of my studies and research. I also thank all my friends in Corvallis who have made the past two years and everything in it possible.

Most importantly I would like to thank my parents who have always supported, encouraged and believed in me to pursue my dreams. You deserve all the credit for who I am today. Special thanks to my younger brother for being my pillar of strength.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

<u>Figure</u> <u>Page</u>

LIST OF TABLES

1. Introduction

A power grid is mainly comprised of generating stations, transmission systems, distribution systems and substations. Each of these components of power system come into a picture at various levels of voltages. Transmission systems are responsible for transmitting the power generated at generation plants, at high voltages for multiple reasons including reducing the copper losses directly proportional to the square of current. The voltage is then stepped down using transformers at substations to various levels. The distribution system then helps in delivering power to larger loads like industrial sites and to feeders which provide power to remaining loads like residential and commercial loads. The power flow studies conducted to obtain grid measurements and the importance of state estimation have been introduced in following sections. The introduction of Phasor Measurement Units (PMUs) to further modify traditional state estimation is discussed and the goal of this research is introduced in the last section. Two algorithms using Singular Value Decomposition are developed through this research for state estimation with only a few PMUs deployed across the system.

1.1 Power Flow Studies

Power flow studies, also known as load flow studies are an integral part of any grid analysis to compute voltage magnitude and angle at each bus and in turn determine active and reactive power flow in all parts of a power system. Power flow studies are necessary to calculate transmission line losses and losses in equipments such as generators and transformers. Power flow studies also help in conducting system fault analysis and economic dispatch and load balancing for healthy grid. A one-line diagram of a power grid including information about buses, transmission lines, generators and loads is used as a power system model to perform load flow analysis. Figure1.1 is a simple one-line diagram of two buses with a load and a generator connected to the Bus 2. A single-line diagram is used to conduct power flow studies to determine the four important parameters of voltage magnitude - $|V|$, voltage angle - δ, active power - P and reactive power - Q.

.

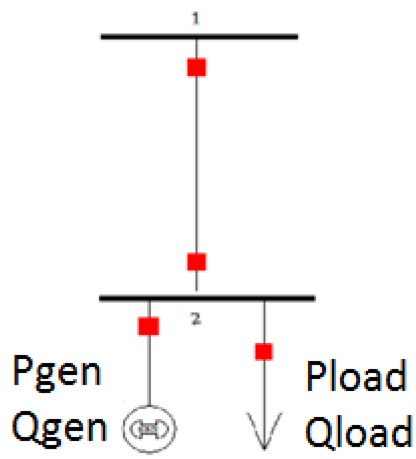

Figure 1.1 Single-line diagram of 2 bus system

Buses in a system are classified into three categories as follows

1. Slack Bus - A primary bus with voltage magnitude 1.0 per unit and 0 voltage angle used as a reference bus. P and Q are calculated for this bus using power flow.
2. PQ Bus - This is a load bus with active power P and reactive power Q known while power flow program calculates voltage magnitude $|V|$ and angle δ.
3. PV Bus- This is a voltage controlled bus with known inputs - active power P and voltage magnitude $|V|$ while the power flow program calculates voltage angle δ and reactive power Q [2].

To begin with power flow studies, power mismatches are calculated at every bus. For example, the power mismatches for both active and reactive power at Bus 2 shown in Figure 1.1 are calculated as

$$P_{Bus2} = P_{gen} - P_{load} \tag{1.1}$$

$$Q_{Bus2} = Q_{gen} - Q_{load} \tag{1.2}$$

Calculation of power flow requires series impedance Z and shunt admittance Y for every transmission line along with the winding series impedance and exciting branch shunt admittance for a transformer. Power flow studies are usually conducted on a balanced 3 phase steady state system which means all three phases of system have a phase shift of 120 degrees and have same amplitude and frequency [2]. Since it is a nonlinear problem, numerical methods are required for computation [1]. The two numerical iterative methods that are used to conduct power flow studies are Gauss Seidel and Newton Raphson. The Gauss Seidel method to obtain power flow solutions is described in next section.

1.1.1. Gauss Seidel

To perform Gauss Seidel power flow, power mismatches at all buses are calculated for both active and reactive power. The admittance matrix called the Y_{bus} for a given network is calculated based on the information from transmission line impedance and admittances. To begin with the first iteration, an arbitrary set of values are selected for unknown quantities at a bus which are updated with the result of every iteration of power flow equation (1.3) where *n* is total number of buses = 1,2,...*n* [3]. It is common to begin power flow iterations with the voltage set to 1 per unit and angle set to zero for a PQ bus. These iterations are repeated until the results converge to a preset value of error.

$$V_{i(j+1)} = \left(\frac{1}{Y_{ii}}\right)\left[\left(\frac{P_i - jQ_i}{V_{i(j)}}\right) - Y_{i1}V_1 - Y_{i2}V_2 - \cdots Y_{ii}V_{i(j)} - Y_{in}V_n\right] \tag{1.3}$$

Measurements for active and reactive power are given by equation,

$$P_i + jQ_i = V_i \left[\textstyle\sum_1^n Y_{in}V_n\right] \tag{1.4}$$

while

$$I_i = (P_i + jQ_i)/V_i \tag{1.5}$$

1.2. Power System State Estimation:

The increasing load demand and addition of various sources of renewable generation to the existing grid infrastructure is exerting stress on the power system. While the load and generation capacity has increased many fold over past years, capacity of transmission and distribution system has not followed. The interconnection of power system networks has become more intricate and complex which makes task of monitoring and operating the system challenging. To maintain continuity of services and meet the generation and load balance it is important to continuously monitor the system to prevent any outages leading to loss of services. Thus power system state estimation - an important aspect of the Energy Management System is critical in acquiring the current state of any power system network [4].

A system called the Energy Management System (EMS) at each power system control center, - is used for monitoring, controlling and optimizing the operation of the transmission system, generating stations and load. The main tool of EMS - Supervisory Control and Data Acquisition including a human computer interface, is responsible for recording and storing digital and analog measurements at various points in grid network used for analysis purpose by grid operators. Raw data in the form of measurements recorded at remote locations in field by Remote Terminal Units (RTU) is transmitted to control center to be fed to a state estimation tool to obtain an optimal state estimation. This is done to identify current state of a system. The measurements acquired include voltages at buses, power flows on the transmission lines, frequency, current along with circuit breaker positions and transformer tap status. A combination of SCADA and state estimation enhances the capabilities of EMS to provide a better platform for the monitoring and control of a system in real time [4].

Other functions of the state estimation based on the system network model and the measurements include

i. Identifying the gross errors in measurements caused due to the noise and fix them along with identification and removal of corrupted measurements,
ii. Create a one-line diagram of a system from the acquired data about circuits breaker and protective switches status,
iii. Perform an optimal state estimation to provide estimates of measurements at metered and unmetered locations of a system [5].

A power system at a given instance of time could exist either in Normal State, Emergency State or Restorative State depending upon the conditions of system equipments and health of the system along with the amount of deviation from optimal operating condition [5]. A system is in a Normal State when generation and load are balanced and there is very little or no deviation from the optimal operating condition. A system is said to be operating in Normal - Secure State when there are no operating constraints violated or when the system is maintained within its upper and lower limits of operation. Whereas, a system is said to be in Normal - Insecure state when one or more operating constraints might be violated but the continuity of service is maintained despite, a contingency occurring due to an outage on a line or failure of an equipment. Although the generation-load balance is not disturbed, system requires attention and some protective measures to prevent it from entering an Emergency State. The system experiences constant variations and when these changes are significant enough to cause contingencies leading to large power outages, the system enters an Emergency state. This calls for immediate attention for respective corrective measures to further avoid the cascading of outages leading to a partial or complete blackout by opening a faulty line or disconnecting a failed equipment. The state where these protective measures called restorative controls are taken to bring the system back to normal state is called Restorative State. Thus state estimation is necessary to aid the task of continuously monitoring and controlling a power system to maintain it in the normal state [5].

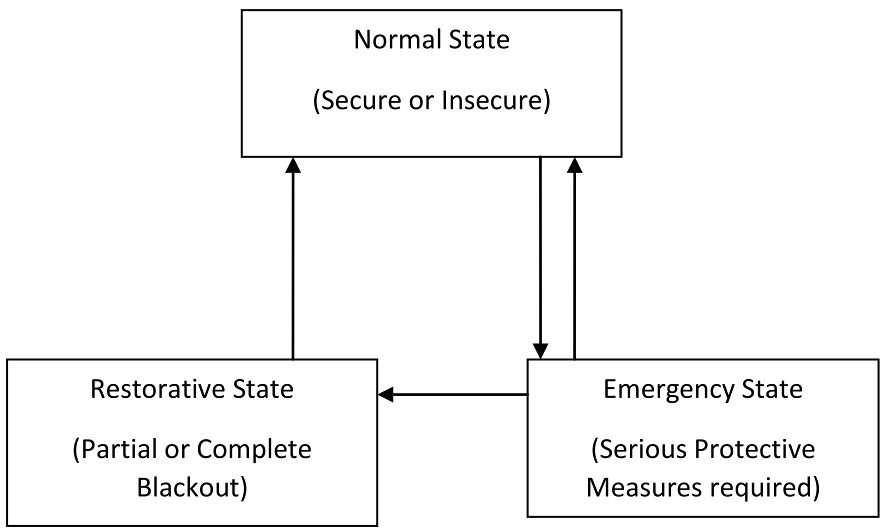

Figure1.2. States of a Power System [5].

1.3 Traditional Method of State Estimation - Weighted Least Squares (WLS)

The traditional method of state estimation involves measurements from SCADA system and is essential to compute and estimate unmeasured grid variables of active power, reactive power, current, voltage magnitude and angle, etc. while considering small discrepancies in the measurements due to noise and measurements that are inaccurate and missing. It is difficult to obtain measurements at various points of a grid at a same time due to the introduction of a certain time delay called time skew. Weighted Least Squares, Least Absolute Value and Weighted Least Absolute Value are a few methods that have been traditionally used for power system state estimation [6].

The error between measured value obtained from the measuring devices and true expected value of measurement can be expressed using equation (1.6) where e is the error, $Z_{measured}$ is the measurement from a device and Z is the expected value of the variable measurement [6].

$$e = Z_{measured} - Z \tag{1.6}$$

These errors are assumed to have a probability density function of Gaussian distribution, described in terms of its mean and standard deviation given by the equation (1.7) [5].

$$f(z) = \left(\frac{1}{\sqrt{2\pi\sigma}}\right) e^{-\left(\frac{1}{2}\right)\left(\frac{z-\mu}{\sigma}\right)^2} \tag{1.7}$$

Where

$f(z)$ - the probability density function for a variable z

σ - standard deviation of the variable z from the mean

μ - mean of z (expected value).

$Z_{measured}$ can then be described in terms of the number of standard deviations away from its mean which means that closer the $Z_{measured}$ is to its mean located at zero, smaller is the error between measured and expected value. This can be shown in following figure where for a higher value of standard deviation σ, higher is the error implying an inaccurate measurement. Similarly it can be seen that for a smaller value of σ, the measurement is more accurate as it is approaching towards the mean μ [6].

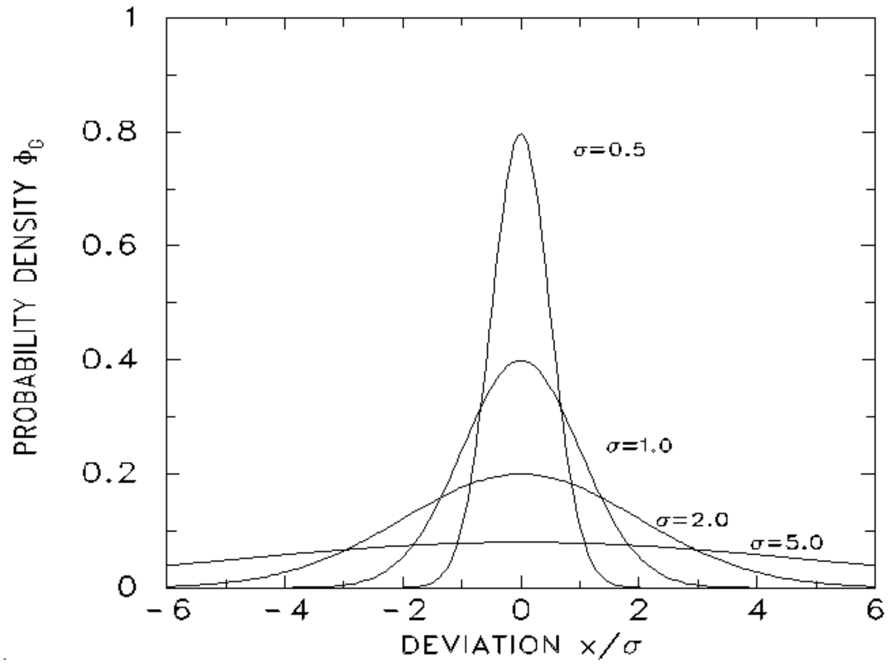

Figure 1.3. Gaussian distribution for different standard deviation [7].

Weighted Least Squares (WLS) method of state estimation uses voltages as static state variable to estimate current state of estimation which can be used to calculate any power flows or generator and load output variables. The static state variables are defined by a vector of voltage magnitudes and angle except for the voltage angle at swing or slack bus. The voltage angle at the swing bus is set to zero for reference. Similar to computation of power flows, state estimation is a nonlinear problem and follows a process of an iterative computation until the result converges to desired pre-set delta value between iterations [5]. Based on the expression for n number of measurements -

$$Z = h(x) + e \tag{1.8}$$

where

Z - a vector of the various measurements = $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_n \end{bmatrix}$

h(x) - the nonlinear function on the state variable

e - measurement error = $\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_n \end{bmatrix}$

x - state variable =$[x_1 \quad x_2 \quad x_3 \quad x_n]$.

WLS state estimator tool minimizes the least square objective function given by (1.9) to get equation (1.10).

$$J(x) = \sum_{i=1}^{n} (z_i - h_i(x))^2 R \qquad (1.9)$$

where, R - the variance of measurement error equals $(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_n^2)$.

$$G(x^k)\Delta x^{k+1} = H^T(x^k)R^{-1}[z - h(x^k)] \qquad (1.10)$$

where

$\Delta x^{k+1} = x^{k+1} - x^k$

x^k - state variable solution at the end of *k* number of iterations.

H(x) - Jacobian matrix

$G(x^k)$ - Gain matrix = $H^T R^{-1} H$.

The gain matrix is a sparse positive definite matrix calculated from the Jacobian matrix *H* and the variance matrix *R* which is further decomposed into its lower triangular matrix and its transpose using Cholesky decomposition. The lower triangular matrix and its transpose are used to calculate Δx^{k+1} via the forward-backward substitution from which the grid variables, power flows on lines and the voltages at buses can be easily computed [5].

This method of system state estimation involves complex computation of the Jacobian matrix H and further decomposition of the gain matrix G into lower triangular matrix and its transpose based on Cholesky decomposition [5]. The traditional methods of system state estimation also have a disadvantage of being incapable of providing a real time estimate of the current state of a system, considering the challenges in measuring the real time data synced in time at higher frequency. These methods are difficult to apply for state estimation of a system with incomplete observability, due to increased computational complexity.

1.4 Synchrophasor Technology

Although traditional methods of state estimation can provide best possible guess for current state of the system, the dynamic state estimation of a system is challenging considering time skew - (i.e., the time delay in the measurements at various locations at the same time). With the advent of Synchrophasor technology, instantaneous phasor values of grid variables of voltage, frequency, current, active and reactive power, etc. can be recorded at transmission level of a power system. Grid monitoring devices called Phasor Measurement Units (PMUs) are an essential part of synchrophasor technology. These units record grid variables and provide dynamic data about state of a power system, at a higher frequency than that has been historically available. The sampling frequency or the frequency at which these units record measurements is over a 1000 times a second, while the time signals are synchronized by a common clock with the aid of the GPS systems. The timely aligned measurements provide a detailed image of the dynamic state of transmission systems to enhance maintenance of the system network [8].

Figure1.4 and Figure1.5 show a comparison between PMU measurements and SCADA measurements for system frequency and bus voltage. Thus it can be seen that PMU data is recorded at much higher resolution while compared to SCADA measurements.

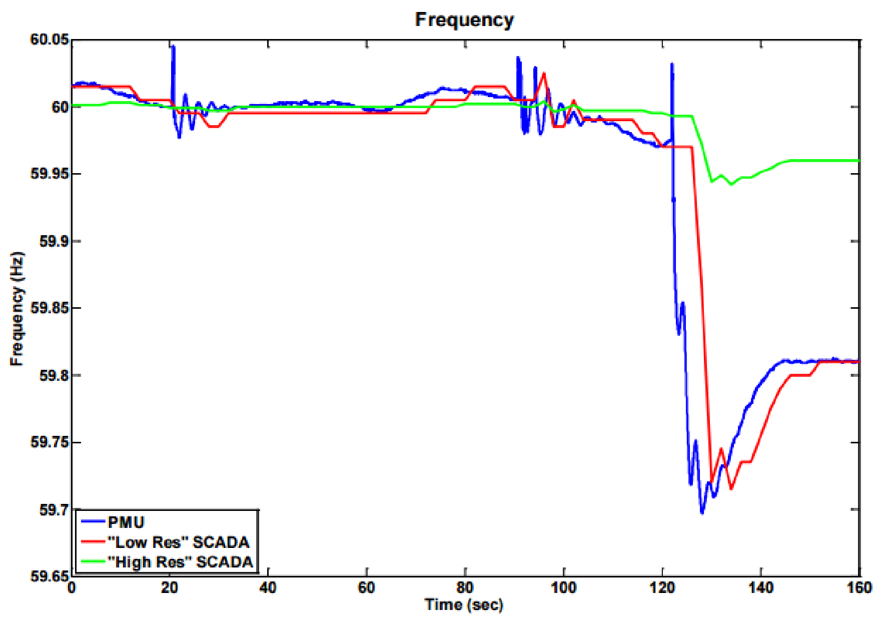

Figure 1.4. Comparison between PMU and SCADA measurements of system frequency [9].

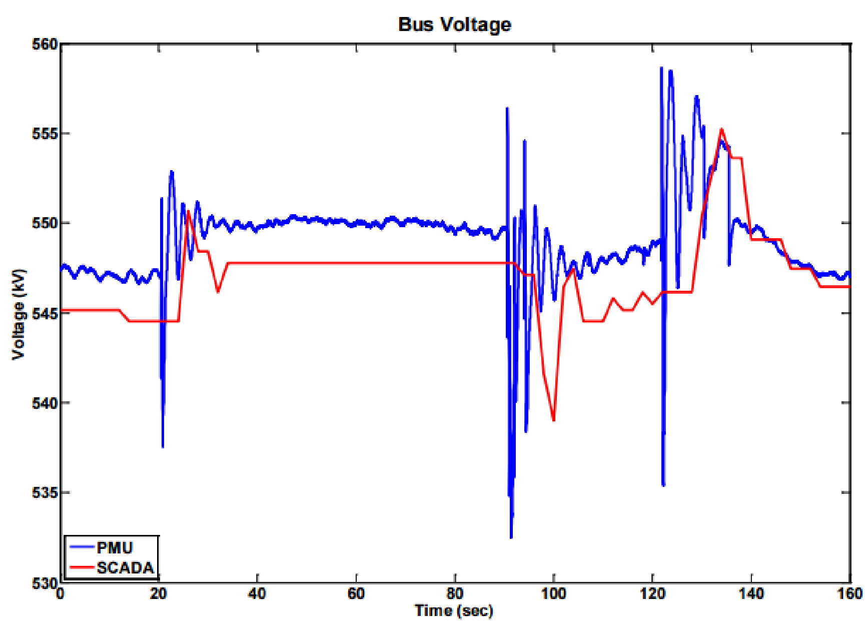

Figure 1.5. Comparison between PMU and SCADA measurements of bus voltage [9]

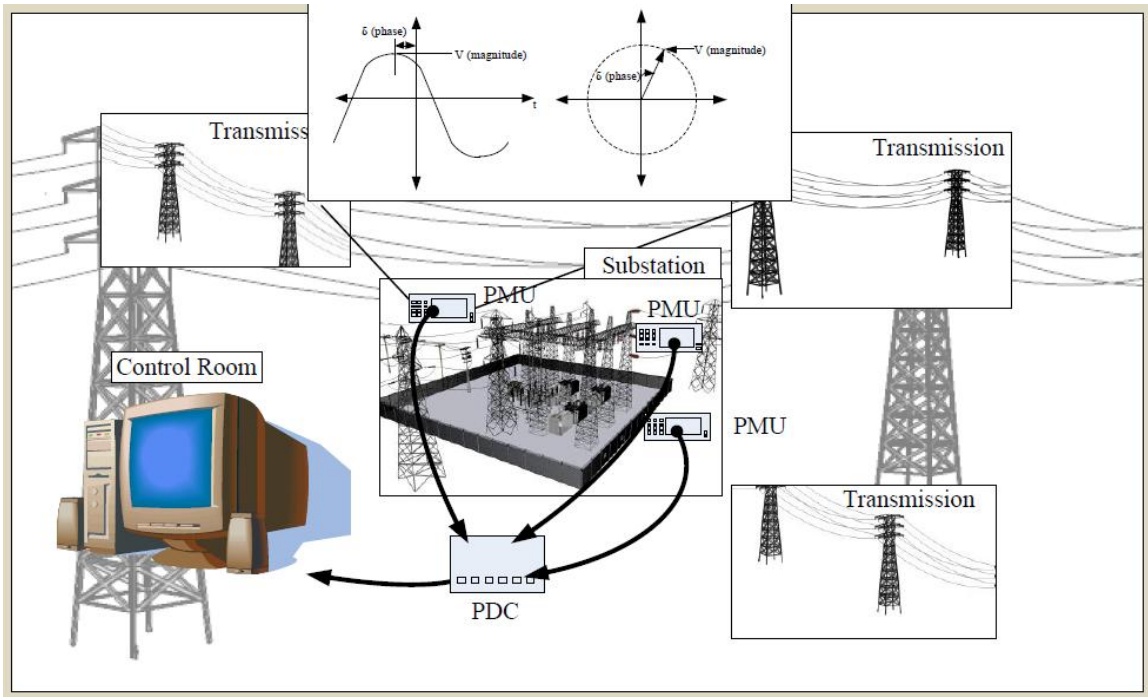

Figure 1.6. PMUs in a transmission level system [8].

Figure1.6 simplifies the demonstration of the use and placement of this technology in a power system at transmission level.

The major advantages and applications of Synchrophasor technologies can be broadly categorized in online applications or near real time applications and offline applications to analyze events causing partial or complete blackouts [8]. Online applications or near real-time applications are used to constantly monitor the system at steady and dynamic state to maintain and operate system within the operating limits and ensure a balance between the generation and load. With the help of high frequency data stream of PMU measurements, the oscillations occurring in the system frequency caused by bringing a generator online or failure of a generator can be detected. This is essential to determine the generation and load balance. Online applications could also include to check for spikes and dips in system voltages and to maintain the system voltage within the operating limits to prevent power system from collapsing. With constant monitoring of grid conditions at real-time or near real-time, faults in a system can be cleared by operating a relay and causing the circuit breaker to isolate the fault either by opening a line or disconnecting a faulty equipment. A great advantage of PMUs that can be explored for monitoring and sustaining the system, is to boost the sampling frequency on observing a variation in the grid behavior, which further assists the operators in taking protective measures to avoid a system collapse or a blackout. Offline applications include post fault analysis to study the causes of disturbance and restore the system back to normal state because PMUs provide historic data making the investigation of occurrence of events convenient which otherwise would require greater efforts and time.

As a part of Smart Grid Investment Grant (SGIG) and Smart Grid Demonstration Program, under the American Recovery and Reinvestment Act of 2009, twelve different organizations are major contributors towards the expansion of the synchrophasor technology network. This includes the deployment of various devices like PMUs, PDCs and communication systems to improve the operation and planning of the national grid [8]. The twelve grant recipients and the number of PMUs and PDCs installed by each are listed in Table 1.1.

SGIG and SGDP Synchrophasor Project	PMUs Installed*		PDCs Installed*	
	Recovery Act Project^	System Total	Recovery Act Project^	System Total
American Transmission Company	45	92	0	2
Center for Commercialization of Electric Technologies	15	18	4	4
Duke Energy Carolinas	98	98	2	2
Entergy Services Inc.	49	49	9	10
Florida Power & Light Company	45	45	13	13
Idaho Power Company	8	15	0	1
ISO-New England	77	77	8	8
Midwest Energy	7	7	1	1
Midwest Independent Transmission System Operator	148	148	21	21
New York Independent System Operator, Inc.	40	40	8	8
PJM Interconnection	56	56	15	15
Western Electricity Coordinating Council	336	481	49	69
TOTAL	924	1126	130	154

** As of 03/31/2013*

Table 1.1. Number of PMUs and PDCs installed [8].

With the current rate of growth in number of PMUs deployed and the number of regions adopting synchrophasor technology, the present coverage of synchrophasor network over transmission system is expected to increase 10 fold [8]. Figure. 1.7 shows PMUs installed and networked by March 2012.

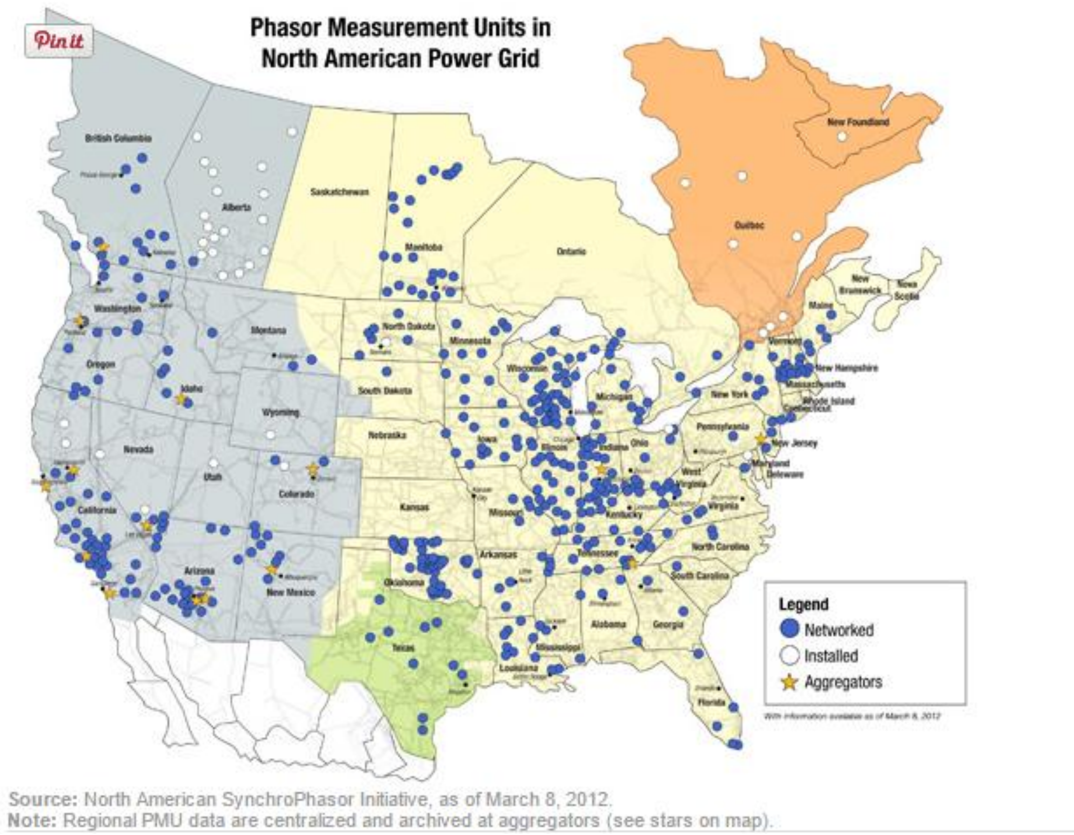

Figure 1.7. Location of PMUs in US [11]

In response to the August 1996 blackout, the first PMU was installed at Bonneville Power Administration (BPA). First PDC developed and installed at BPA was in May 1997 which makes BPA one of the pioneers in the field of synchrophasor technology [9]. BPA has built the largest and most sophisticated network of PMUs in the country which has improved grid operations at generation and transmission levels in the Northwest [10]. This network consisting of 126 PMUs at 50 locations, including substation and wind energy generation sites is the largest contributor to the Western Interconnection Synchrophasor Program (WISP) to have a network of more than 600 PMUs across the western grid [10].

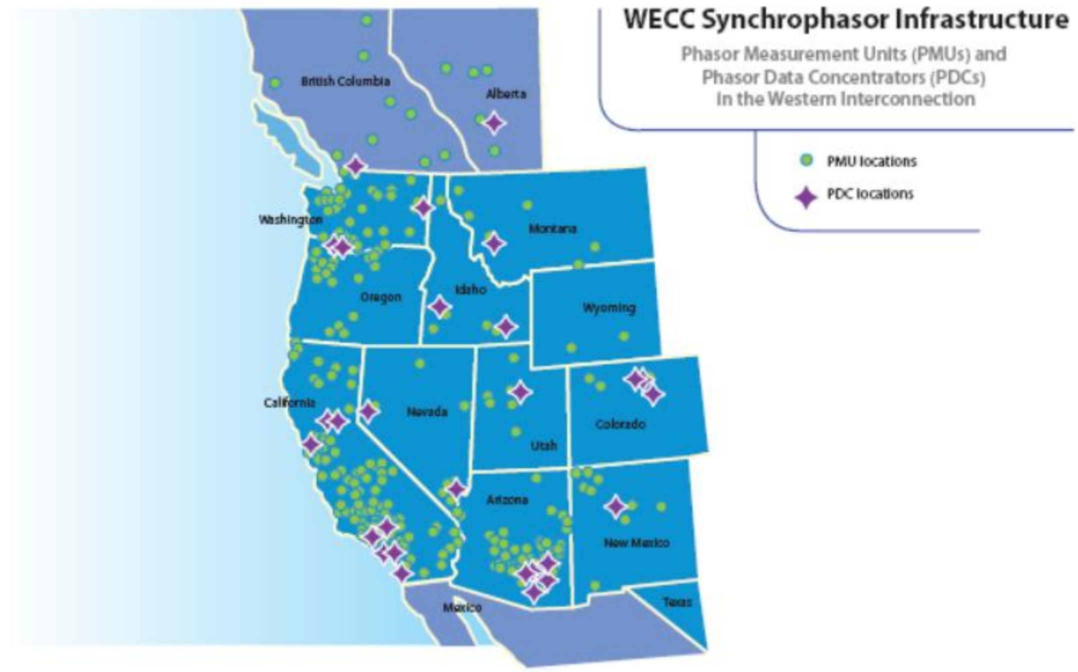

Figure 1.8. WECC Synchrophasor Infrastructure [9].

1.5 Traditional State Estimation with PMUs

As mentioned in Section 1.3, traditional method of WLS state estimation of a system is based on the nonlinear state equation given by (1.6), where x is the static state variable vector consisting of voltage magnitudes and angles with *(2N-1)* dimension, N being the total number of buses in system. Vector Z consisting of active power, reactive power, voltage magnitude and angle is further extended in order to accommodate PMU measurements [12]. To conduct state estimation using measurements from SCADA and PMUs, the voltage magnitudes from PMUs can conveniently replace SCADA measurements in WLS state estimator. But replacing SCADA measured voltage angles with PMU measured voltage angles to estimate the state of a system is a complicated problem [12].

One solution to this problem of using voltage angle measured by PMU, is to install a PMU at reference bus of the system's WLS estimator algorithm. This is helpful because voltage angles are estimated with respect to this reference [12]. The reference voltage angle is subtracted from voltage angles (obtained from SCADA). A large difference between the angles could lead to inaccurate estimates. But a change in system reference to minimize this difference could solve this accuracy problem [12]. This requires a PMU installed at the new reference which is challenging in terms of high installation cost. Another solution to replace the conventional measurements by PMU measured angles is to calculate the difference between the voltage angles at the beginning and end of a transmission line [12]. This could imply that the system would require more number of PMUs installed at the beginning and end of a line which involves the constraint of high cost. This method of using the WLS state estimator along with PMUs is a static state estimation technique and has no promising results for an application in real-time or near real-time.

1.6 Research Goal

There has been significant amount of research for state estimation of a grid using PMUs at transmission level. Planning and operation of systems at distribution level is equally critical to maintain the continuity of services and prevent local outages. This research aims at utilizing synchrophasor technology to estimate the state of a system at distribution level. Various universities are adopting smart grid technologies to be self-sustained for an added capability of operating in an islanded mode or as a micro grid. There has been significant progress towards moving the Oregon State University (OSU) system to a smart grid and making the campus self-sustained. This research funded by Bonneville Power Administration (BPA) includes installing several PMUs at significant locations on OSU campus to improve system monitoring and operations. This also includes determination of ideal locations for maximizing the benefits of installing PMUs. The ultimate goal is to develop a state estimation technique to estimate the state of campus with limited observability due to sparse deployment of PMUs across the campus. The anticipated application of this is to have an accurate load model of WECC [13]. The load flow model of OSU with a total load of 25 MW is further represented in terms of residential, commercial and industrial percentage at every bus.

2. System State Estimation of Oregon State University

2.1 Phasor Measurement Units on the Oregon State University Campus

Various locations for installing PMUs on OSU campus are predetermined based on number of loads served by a bus and the type of load because the loads are represented in terms of percentage of residential, commercial and industrial. The OSU campus consists of 286 buses in total and there are various buses which supply to either one or more kinds of loads in terms of residential, commercial and industrial. Residential loads on campus are buildings with lodging and student dormitories. Commercial loads are buildings with offices, classrooms, etc. and industrial loads are laboratories consisting of several 3 phase high power machines. The location of PMU placement is also determined based on sensitivity analysis by observing the electrical distance between various buses [14]. The sensitivity analysis in simulation is performed by increasing the active power at a load bus and observing the variation in voltage magnitude and angle at other buses. If the observed variation in voltage magnitude and angle at other buses exceeds a certain threshold value, the initial load bus responsible for this voltage variation is considered to be an ideal PMU location [14]. The following Figure 2.1 of OSU campus model, built for earlier research determines locations of installed PMUs and potential locations for PMUs.

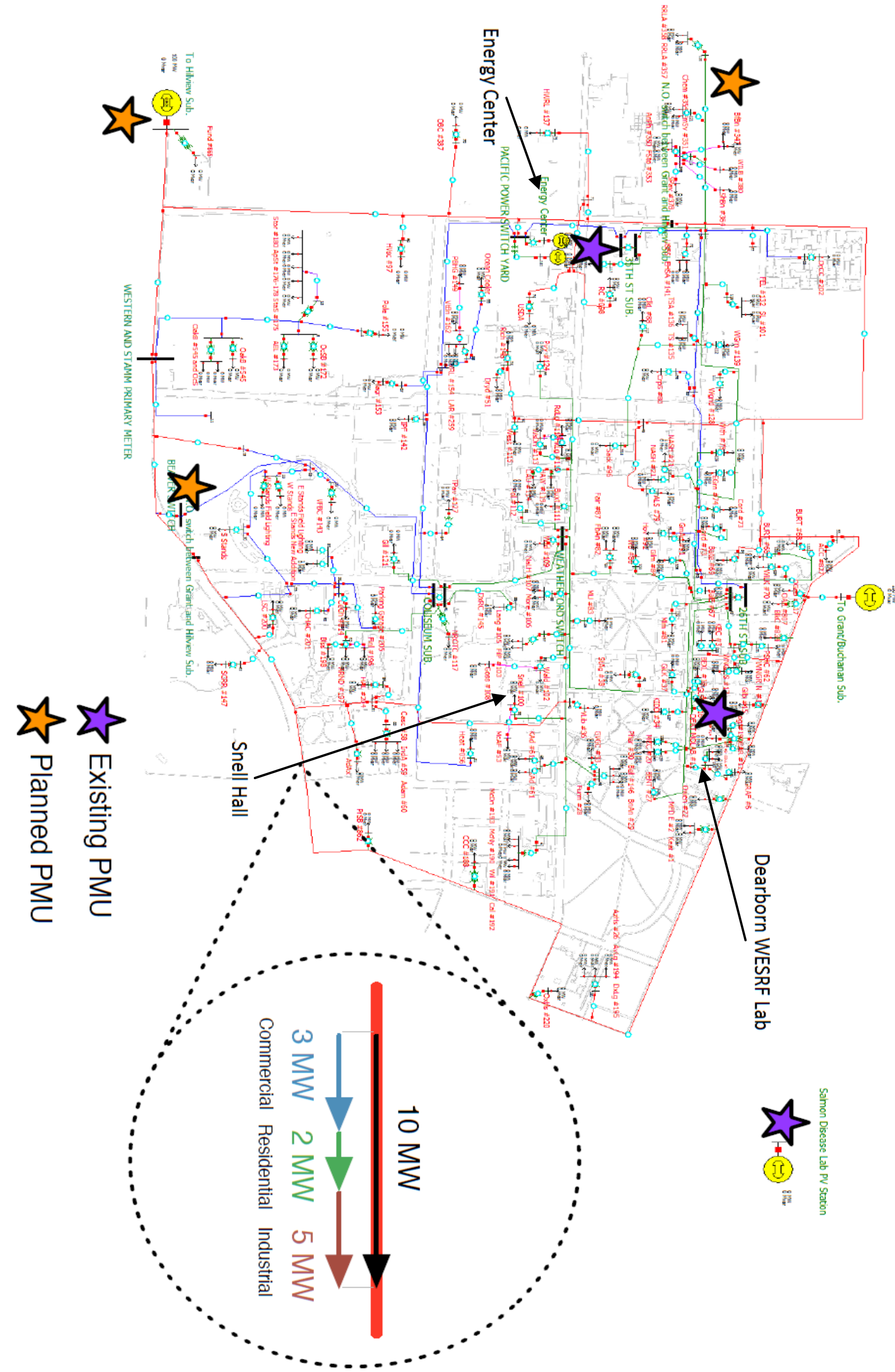

Figure 2.1. Existing and Planned PMUs on OSU campus

The initial locations selected for installing a PMU on the OSU campus include buildings - Snell Hall, Energy Center, Salmon Disease Lab and a metal fabrication plant in Albany, apart from WESRF lab at bus 244, which already has a PMU installed. The Snell Hall building supplied by bus 209, consists of various single phase loads like air conditioning systems which makes it an interesting location to study the behavior and impact on the remaining system. The Energy Center connected to bus 140, has the campus co-generation plant which is another interesting location for a PMU as it will allow the continuous monitoring of generation. Another important location selected for PMU installation is the Salmon Disease lab located at bus 6. The Salmon Disease lab has several kinds of blowers and compressors representing commercial load as well as pumps and variable frequency drives representing industrial loads. External to the campus, a metal fabrication plant serving as an industrial load is also chosen to have a PMU in association with Consumer Power in Albany which will provide a means for performance assessment of the communications. The measurements recorded at these various locations at high sampling frequency are transmitted to PDC to accumulate all the time stamped measurements in a synchronized order for analysis purpose.

2.2 Incomplete Observability Problem

Considering the high cost of installing PMUs it is practically impossible to install these units at every bus of the system. The cost constraint of PMUs limits total number of units that can be deployed across the system. Thus selecting right locations to install PMUs is extremely critical to provide necessary grid information. Although a few buses with PMUs are made observable, remaining of the system with no PMUs remains unobserved. This leads to the problem of incomplete observability.

Complete observability is essential to obtain the state of system in contrast to the incomplete observability provided by sparsely distributed PMUs. The traditional method of WLS to estimate state of system with incomplete observability involves complicated computations. These methods are also difficult to estimate the state of a system in real-time or near real-time. Therefore a faster method to estimate the state of a system at near real-time, with a few observable locations is studied through this research.

2.3 Dynamic State Estimation

Any electric grid is constantly evolving with variations in electric grid parameters which need to be monitored. Maintaining a balance between generation and load, preventing outages and taking right protective measures in case of a contingency are made significantly easier by a real time state estimation method. This provides system operators a comprehensive picture of the static and dynamic operations of a grid at real time. This research explores Singular Value Decomposition as a method which can allow for real-time and near real-time state estimation of OSU campus with the data obtained from PMU measurements at several locations.

Prior to having the planned PMUs installed and a real-time feed of the measurements, two algorithms to estimate the state of the system with incomplete observability are developed. A set of data using the OSU campus power system model is created offline which is used to test the algorithm. To ensure the algorithm can be used to estimate the state of a system with 286 buses, it is initially developed and tested on a small system of 3 buses and an IEEE 14 bus system which is further developed to apply to OSU campus power system model. Using SVD, the measurements for buses with no PMUs installed can be estimated based on a set of static data and time stamped measurements obtained from the PMUs installed.

2.4 Data for State Estimation

Power flows explained in Section 1.1 are conducted on above systems to create a set of offline data prior to having a real time stream of measurements from PMU. PowerWorld tool is used to run these power flows to build an offline set of data which essentially forms a library. The procedure of creating this offline library follows the procedure of running power flows in succession on a system by varying the load and generation at every bus. The generation and load power at a bus of a system is varied and the power flows solutions for voltage magnitude and angle, active and reactive power at all remaining buses are recorded. Thus the library consists of several records for the four measurements i.e.,$|V|$, δ, P and Q at all the buses of the system. This library forms the heart of state estimation using the concept of SVD.

As shown in Table 2.1, power flow library consists of$|V|$, δ, P and Q measurement channels for every bus along the columns and various records along the rows. The records of the library can be considered analogous to bus measurements at different instances of time. This format of library is in accordance to the format of data recorded by the PMU and transmitted to PDC. This library essentially is a lookup table used as reference to perform SVD for system state estimation. Apart from this power flow library, power flows are run on each system in similar fashion, to create six additional records to form sample test records which will be used for testing and analysis purpose. To perform state estimation using SVD, measurements at buses with PMUs are extracted for respective buses from a sample test record. This forms a new partial record. Measurements at remaining unobserved buses are filled in for this new record with placeholder values based on an approximation discussed in further sections.

	Bus1				**Bus2**				**...**	**Bus'n'**			
Record1	V_1^1	δ_1^1	P_1^1	Q_1^1	V_2^1	δ_2^1	P_2^1	Q_2^1	...	V_n^1	δ_n^1	P_n^1	Q_n^1
Record2	V_1^2	δ_1^2	P_1^2	Q_1^2	V_2^2	δ_2^2	P_2^2	Q_2^2	...	V_n^2	δ_n^2	P_n^2	Q_n^2
Record3	V_1^3	δ_1^3	P_1^3	Q_1^3	V_2^3	δ_2^3	P_2^3	Q_2^3	...	V_n^3	δ_n^3	P_n^3	Q_n^3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮	⋮
Record'm'	V_1^m	δ_1^m	P_1^m	Q_1^m	V_2^m	δ_2^m	P_2^m	Q_2^m	...	V_n^m	δ_n^m	P_n^m	Q_n^m

Table 2.1. Power flow library.

From above table, we see each bus has measurement channels for $|V|$, δ, P and Q for various records. Number of records and measurement channels for the 3 bus system library, IEEE 14 bus and OSU power system model will be elaborated in the following sections.

2.4.1 Simple 3 Bus System

Figure 2.2 shows a simple model created in PowerWorld for a 3 bus system with Bus 1 as reference bus of the system consisting of 2 generating units at Bus 2 and 3 and a load at bus 3. This 3 bus system is used to demonstrate the state estimation using SVD. Power flow library 'L' for this system is created by obtaining valid power flow solutions which span over a wide range to encapsulate different operating conditions of the system. This is done by running power flows in succession while varying the load and generation to obtain $|V|$, δ, P and Q at every bus. The L obtained for this system consists of 100 records and 12 channels with $|V|$, δ, P and Q measurements for each bus. To demonstrate state estimation for an incomplete observable system, Bus 3 is assumed to have a PMU. This makes Bus1 and Bus 2 unobservable and measurements of $|V|$, δ, P and Q are to be estimated for these buses.

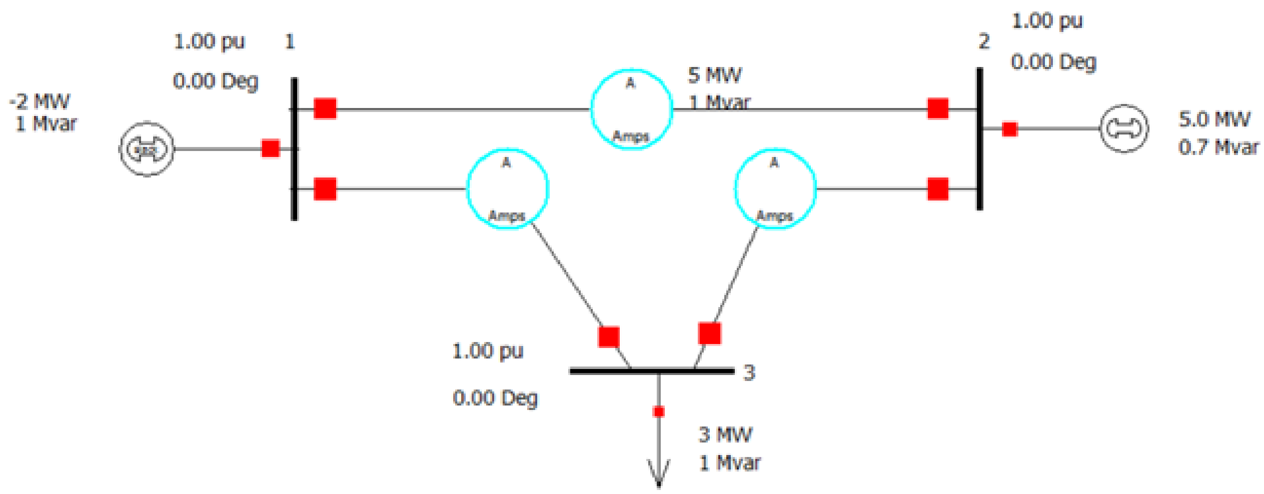

Figure 2.2. Simple 3 Bus system used to demonstrate grid estimation using SVD

2.4.2 IEEE 14 Bus System

SVD method is next applied to an IEEE 14 bus system to test its functionality and accuracy to further use it to estimate grid state of the OSU campus system consisting of 286 buses. Figure 2.3 shows IEEE 14 bus system adopted to create library L for 14 bus system. The 14 bus system comprises of 10 load buses with a total load of 260 MW and four generators with a total capacity of 262 MW [18]. Similar to 3 bus system the power flow library L for this system is built by solving for power flow solutions. It is desired that the library provides a broad coverage of valid solutions as it effectively embodies the underlying power flow equations. The library is created by varying the load and generation at buses in succession until a large number of records are created. A total of 111 records are obtained from the power flow solutions which form the rows of power flow library L. Each of the 14 buses of the system have channels for voltage magnitude and angle, active and reactive power. Since there are a few buses which have both load and generator connected, the total number of channels in library is 60. Thus the power flow library for 14 bus system has a dimension of 111 by 60. Similarly additional 6 records with 60 channels are created apart from the library 111 records and maintained separately to test the SVD algorithms. This system is assumed to have PMU s installed at three locations - buses 12, 13 and 14. Thus the measurements for these buses are extracted from a single test record out of the six test records to form a partial new record. But the data for the remaining buses of this new partial record is to be approximated which is explained in the next chapter. The state of grid is to be estimated by estimating measurements at all the remaining buses from 1 to 11 which are unobserved.

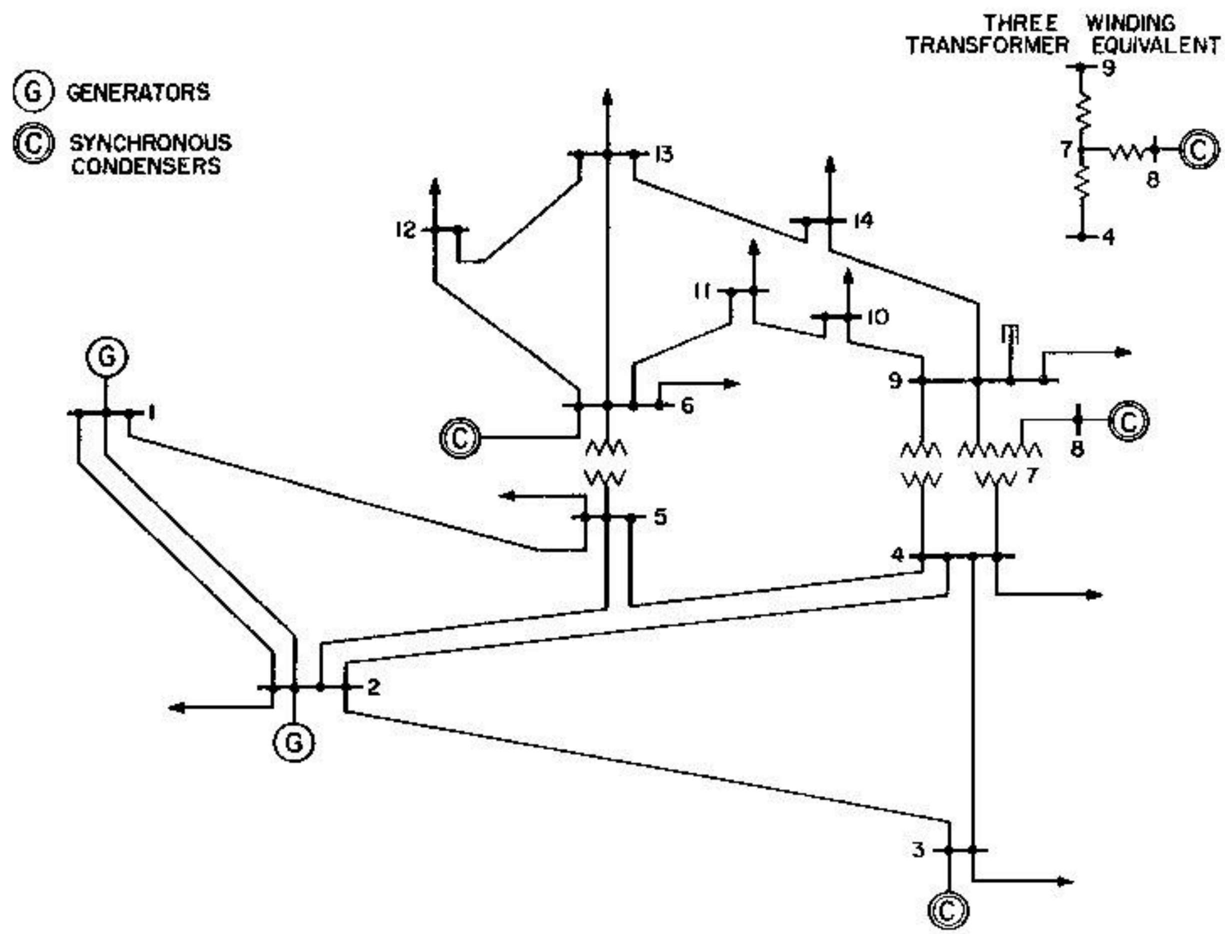

Figure 2.3. IEEE 14 Bus System [15].

2.4.3 OSU Power System (286 Bus System)

The ultimate goal of this research is to estimate the state of OSU power system with limited observability due to sparse deployment of PMUs across the campus. Next figure shows the power flow model of OSU system created for earlier research with a total of 286 buses [14]. Two generators shown on the north and west of campus are used to represent the equivalence of two main feeders/substations while the red/blue and green colored lines represent the 20.8 kV and 4.16 kV lines owned by different entities supplying to 185 load buses of total capacity of 25 MW [14].

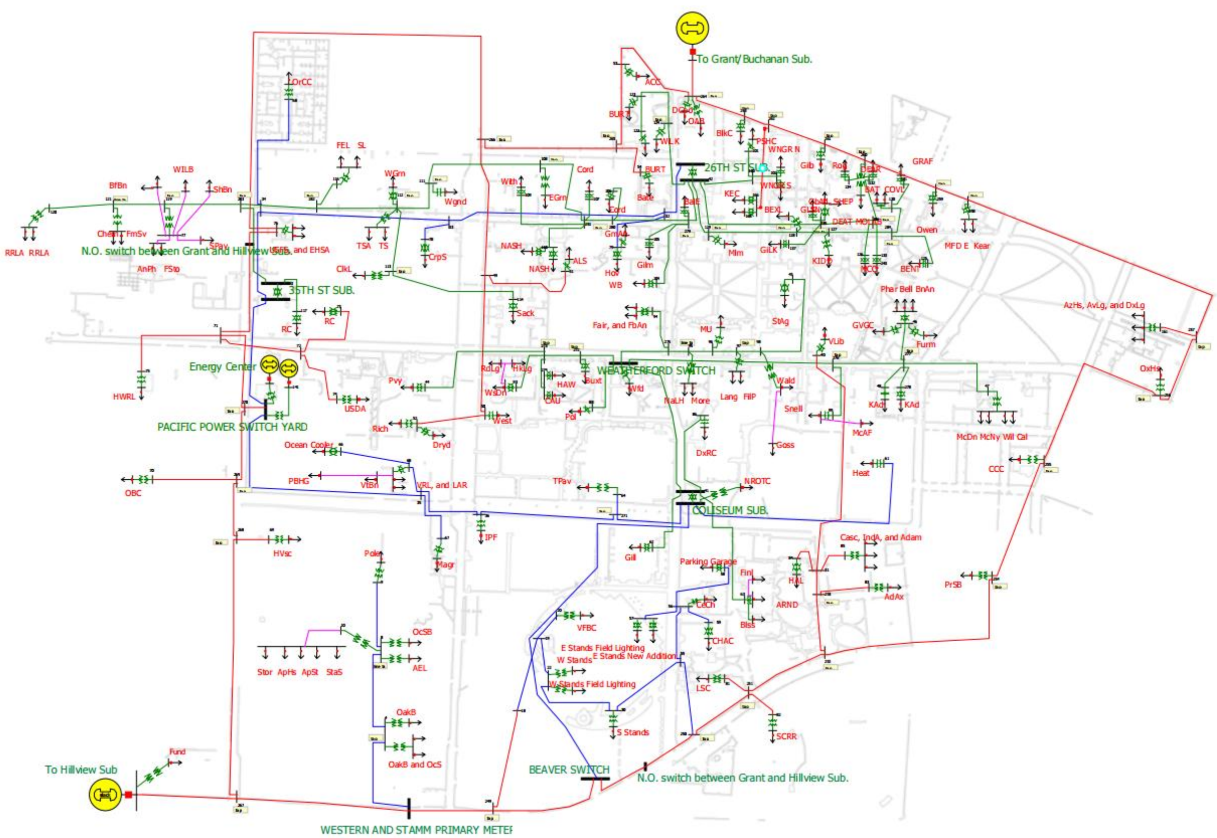

Figure 2.4. OSU campus power flow model.

As discussed in previous sections, the library comprising of valid power flows for 286 buses is built to provide a broad coverage of power flow solutions. 125 records are obtained in the library, as a result of successive power flows conducted. In addition to these 125 library records, 6 test records are obtained in the same manner. Thus the 286 buses power flow library L, consists of 125 rows as records and 844 columns as channels. The number of channels is 844 due to several buses in system with no load or generator connected which means each of these buses would have only two channels $|V|$ and δ. The test library comprising of the 6 test records which is not a part of the library L has 844 channels for every record. The power flow library L which is essentially used for training (i.e. SVD) is kept separate from test library. The present PMUs installed on the campus of OSU at buses 140, 209 and 244 will be used for testing the SVD algorithm. Thus for testing purpose, measurements corresponding to these buses are extracted from a single test record to form a new partial record. The remaining elements of this partial new record are essentially filled with placeholder values as explained in the next chapter.

3 Singular Value Decomposition (SVD)

3.1 Theory

Singular Value Decomposition is a matrix analysis technique most commonly used for data reduction and similarity matching applications. SVD is defined by the equation-

$$A = U * S * V' \tag{3.1}$$

where

> *A* is a matrix of dimension $m \times n$, decomposed into an orthogonal space that helps in clustering the data to develop a relation between the rows and columns of *A*.
>
> *U* is a unitary matrix of dimension $m \times m$ which forms an orthonormal basis for columns of matrix *A*.
>
> *V* is an $n \times n$ dimensional unitary matrix which forms orthonormal basis for the representation of rows of matrix *A*.

S is an $m \times n$ dimensional diagonal matrix with singular values describing the strength of concepts connecting the rows and columns of *A*. The matrix *S* consists of singular values along its diagonal in a descending order where $S_1 > S_2 > S_3 > S_4$ and so on as shown in the matrix below-

$$\begin{bmatrix} S_1 & 0 & 0 & 0 & \cdots \\ 0 & S_2 & 0 & 0 & \ldots \\ 0 & 0 & S_3 & 0 & \ldots \\ 0 & 0 & 0 & S_4 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

SVD helps in clustering data with an underlying similarity to form 'concepts.' SVD requires that we keep the singular values or the concepts which are stronger to further reduce the *U*, *S* and *V* matrices. Data reduction for eliminating concepts which do not contribute in providing useful information about the data, can be obtained by retaining larger singular values and setting the smaller singular values to zero thus eliminating the corresponding columns and rows of U and V matrices [16]. Eliminating concepts with lower values can be considered to be analogous to filtering of data to minimize the noise.

One of the most common applications of SVD is its utilization for recommender systems [17]. In recommender systems the basic idea is to make a huge set of data, with millions of rows and columns informative of the relation between inputs and outputs. One example that can be used to explain this application of SVD is a large table consisting of *users* and *shows* as in Netflix. The table with *users* along the rows and TV *shows* along the columns comprises of ratings provided for all the *shows* by every *user*. The high possibility of a single *user* favoring a particular genre of *shows* implies that multiple *shows* belonging to a single genre like comedy or action are rated similarly by the *user*. This can be utilized to cluster data belonging to similar group or the *shows* rated in a similar fashion. Performing SVD on this table allows clustering of such data, which identifies a fundamental pattern of the data and decomposes it to reveal similarity between *users* and *shows*. Concepts are essentially the similarly rated groups or genres of *shows* here. The matrix U obtained on performing SVD links the *users* belonging to original table to concepts. Similarly matrix V links the shows to concepts and the singular values of S describe the strength of concepts connecting *users* and *shows*. These matrices are further reduced to retain high energy singular values. The reduced matrix U obtained after setting weaker singular values to zero consists of *users* along rows and the concepts along columns. Similarly reduced matrix V consists of TV *shows* along rows and the concept scores along columns. The importance of SVD application is realized after performing matrix operations on these matrices to provide a rating for a new *user*. These matrix operations on U, S and V will be discussed in the following sections as applied to estimate the state of a power system [19].

3.2 Methodology

Above technique of SVD is now applied to estimate state of a system which has incomplete observability with limited number of locations of a grid made observable with PMU measurements. Prior to estimating state of OSU campus power system, SVD is applied to smaller systems consisting of three and 14 buses. The first step towards estimating state of system is to build a library by running power flow equations several times to form a library and 6 test records as explained in the previous chapter. The two approaches using SVD namely - Similarity Matching and Filtering Method, - will be discussed for these three systems in the following sections.

3.2.1 Data Scaling

Scaling the power flow library before applying SVD for state estimation is a critical step to have all the measurements with different units on a single uniform scale. It is important to normalize the data to flatten the scale of$|V|$, δ, P and Q measurements to maintain a single scale of reference common to different units of these measurements. This makes comparison of various units on a common scale highly convenient. For example, in case of power flow library, analysis of measurement data belonging to voltage per unit , angle degrees, MW and MVAR units of $|V|$, δ, P and Q becomes possible with the normalized data.

There are various methods of normalizing data including Z-scores Transformation, Feature Scaling and Quantile Normalization. For this research, the power flow library is normalized using the technique of Quantile Normalization. In this method, the original data library is transformed to the source cumulative distribution function data - using the Cumulative Distribution Function (CDF) with minimum and maximum values set for every channel. Source CDF is then transformed using inverse of cumulative distribution function to obtain the target CDF with minimum and maximum values set at 0 and 100. This is a great normalization technique to normalize all the channels of library to a scale of 0 to 100. The following figures show a simple demonstration of normalizing a single measurement belonging to a channel of power (of 2 MW), to the scale of 0 to 100.

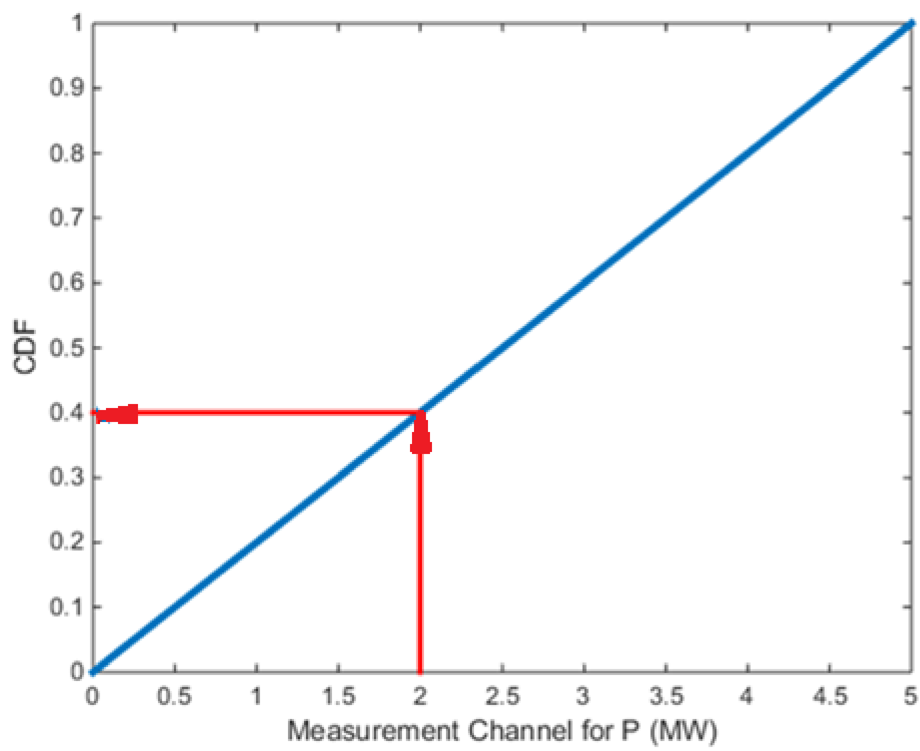

Figure 3.1. CDF calculated for single measurement channel of active power P = 2 MW. Minimum and maximum set for this CDF transformation are the channel minimum and maximum, i.e., 0 MW and 5 MW respectively.

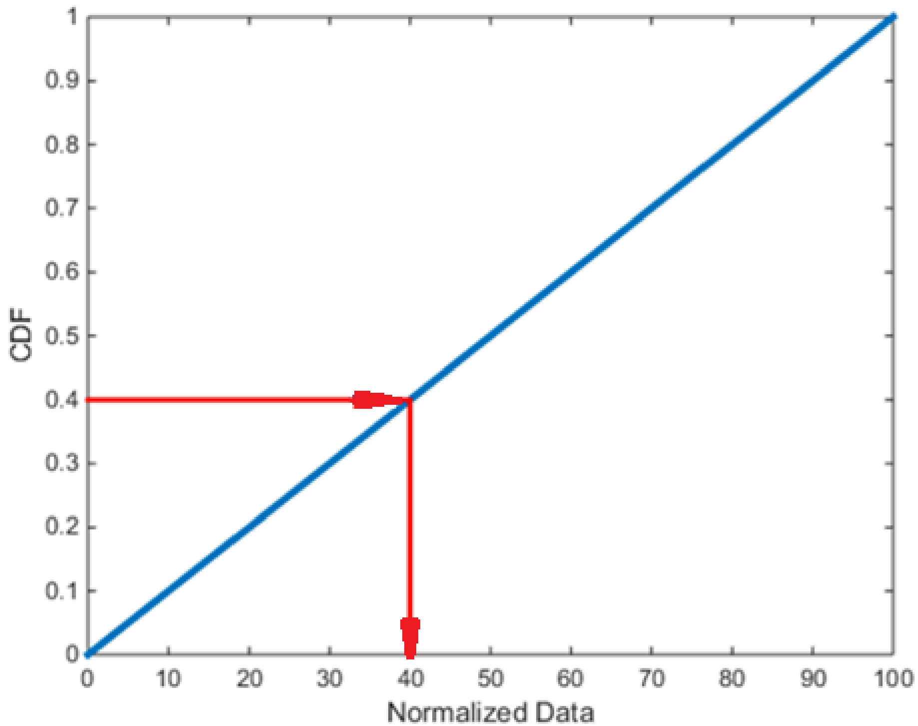

Figure 3.2. Inverse CDF calculated to normalize source CDF data. Minimum and maximum set to 0 and 100.

This normalization is achieved by using the 'cdf' and 'icdf' functions in Matlab given by the equation,

$$L_{normalized} = F_{target}^{-1}(F_{source}(L)) \tag{3.2}$$

where F_{source} is the CDF of original measurements of library L and F_{target} is the desired CDF of the processed data. The CDF for target data is set with minimum and maximum value of 0 and 100. An example of 2 records for 3 bus system library and the respective normalized values, are shown in following tables.

Table 3.1 consists of excerpts from the 3 bus system library, for two records of raw measurements for all 12 channels.

	V_1	δ_1	P_1	Q_1	V_2	δ_2	P_2	Q_2	V_3	δ_3	P_3	Q_3
Record1	1	0	-0.78	-2.2	1	-0.02	-2.25	-0.22	1	0.01	-3.03	-2.42
Record2	1	0	-1.25	-2.2	1	-0.02	-2.25	-0.22	1	0.01	-3.5	-2.42

Table 3.1. Excerpts from 3 bus system original data measurements for all channels of 2 records.

Table 3.2 shows the normalized values for respective channel measurements for the above two records.

	V_1	δ_1	P_1	Q_1	V_2	δ_2	P_2	Q_2	V_3	δ_3	P_3	Q_3
Record1	50	50	41.9	0	0	16.66	0	40.9	99.99	74.99	16.80	19.73
Record2	50	50	38.12	0	0	16.66	0	40.88	99.99	74.99	11.42	19.73

Table 3.2. Normalized values for the above channel measurements

This library with data normalized to scale of 0 to 100 is then used for SVD. Another important criteria of choosing this method of normalization is to have a uniform scale for error calculation for all four units. Limiting the data to range between 0 to 100, represents the errors in terms of percentage which is simple to comprehend. This also limits the results of SVD calculations to a single scale, by ensuring all the measurements are maintained within specified bounds, especially larger units of active and reactive power in MW and MVAR respectively.

3.2.2 SVD Method

SVD is performed on power flow library using the Matlab function 'svd' to decompose library matrix and obtain U, S and V' with matrix dimensions as given in the following equation.

$$A_{n_r X n_m} = U_{n_r X n_r} * \; S_{n_r X n_m} * \; V'_{n_m X n_m} \qquad (3.3)$$

where

n_r is the number of records in library.

n_m is the number of bus measurements in library - active power, reactive power, voltage magnitude and angle for all buses.

These matrices are further reduced by retaining larger concepts while the smaller insignificant concepts are set to zero. A common rule used to determine the total number of concepts that must be retained is - to keep singular values that account for 90% of the energy in *S*. This implies that the sum of squares of retained singular values must be 90% of the sum of squares of all the singular values [16]. The following equation shows the matrices after dimension is recued to n_o concepts -

$$\bar{A}_{n_m X n_r} = \bar{U}_{n_r X n_o} * \bar{S}_{n_o X n_o} * \bar{V}'_{n_0 X n_m} \tag{3.4}$$

where

$\bar{U}_{n_r X n_o}$ is the reduced matrix with records in rows and concept score in columns.

$\bar{V}'_{n_0 X n_m}$ is the reduced matrix with measurement channels in columns and concept score in rows, as $\bar{V}_{n_m X n_o}$ would have concepts along columns and channels along rows.

$\bar{S}_{n_o X n_o}$ is a matrix with the concepts and their strengths.

A single row of matrix $\bar{U}_{n_r X n_o}$ is the strength of concepts for the corresponding record of the library. Every element of the first row in $\bar{U}_{n_r X n_o}$ is essentially the strength of every element of first record in *A* for all the concepts retained. Similarly for $\bar{V}_{n_m X n_o}$, every row is a measurement channel which is connected to the concepts. Thus elements in the first row of this matrix describe the strength of first measurement channel for every concept.

3.2.3 Similarity Matching

Above form of the reduced library has its measurements and records projected in a concept space. This reduced library is used for similarity matching to estimate measurements for the unobserved buses of system. When PMU measurements at a single or multiple observed buses are recorded, the measurements for all other unobserved buses, are approximated. The measurement channels for unobserved buses are approximated with normalized placeholder data. The normalized placeholder data is an average of normalized measurements for the corresponding buses in library.

Thus a new record with PMU measurements from an observed bus and placeholder values for unobserved buses is formed. The new record is then projected to concept space using the equation

$$C_{1Xn_0} = r_{1Xn_m} * \overline{V}_{n_mXn_o} * \bar{S}^{-1}_{n_0Xn_0} \tag{3.5}$$

where C_{1Xn_0} is the concept score of the new record r_{1Xn_m}.

The closest match or the most similar record obtained on comparing this new record with records of library L in concept space, is then used to approximate unobserved measurements of the new record. Similarity matching is done by finding the smallest 2-norm value between new record concept score C_{1Xn_0} and the concept scores of library records. The 2-norm is calculated between $C_{1Xn_0}$and every row in$\overline{U}_{n_rXn_o}$. The closest library record is used to estimate the state of system at unobserved buses. Mean Absolute Error between estimated new record and original test record (out of the six test records created for reference purpose) is calculated to demonstrate the accuracy of similarity matching technique.

3.2.4 Similarity Matching Algorithm

This section attempts to simplify the description of estimating state of a power grid using SVD technique - similarity matching. This algorithm, used for estimating the state of all three systems including 3 buses, IEEE system consisting14 buses and OSU campus model consisting of 286 buses is explained as following -

i. Perform power flow equations on power system for various values of load and generation to create a library L of multiple records and 6 additional records for voltage magnitude and angle, active and reactive power measurements at all buses.

ii. Normalize the power flow library using Quantile Normalization as explained in section 3.2 -Data Scaling.

iii. Run SVD on normalized power flow library using equation 3.1 to obtain U, S and V matrices.

iv. Top two concepts are chosen to retain high energy concepts as explained in Section 3.1 to obtain reduced matrices $\bar{U}$, $\bar{V}$ and $\bar{S}$.

v. A single record R_{test}, is selected from the set of six test records and normalized as explained in Section 3.2.

vi. New PMU measurements from a single bus or multiple buses for $|V|$, δ, P and Q are recorded. These are measurements for the respective observed buses extracted from R_{test} to form a vector '*m*'.

vii. The remaining unobserved channels for $|V|$, δ, P and Q measurements are approximated by placeholder values determined by calculating mean of the normalized measurements for every channel in L to form vector '*p*'.

viii. A new record '*r*' is then obtained from m and p by $r = [p\ m]$.

ix. Concept scores C for this new record 'r' are computed using Equation 3.5 which translates 'r' to concept space.

x. C is compared with the concept scores of library records by computing 2-norm between C and every row of $\overline{U}$.

xi. Record with smallest 2-norm value is the most similar library record to 'r'.

xii. This most similar record from L is used to estimate measurements to form record- R_{est}.

xiii. Steps v. to xi. are repeated for all the six test records.

To determine the accuracy of SVD Similarity Matching algorithm the Mean Absolute Error is calculated by determining error between estimated new record R_{est} and original test record R_{test} given by -

$$MAE = mean(|R_{test} - R_{est}|) \tag{3.6}$$

It is important to note that the estimated new record R_{est} is a result of similarity matching algorithm. Whereas, r is the approximated new record which forms an input to the similarity matching algorithm.

3.2.5 Filtering Method

Another approach to estimate the state of a power system - called filtering method using SVD is explored as a part of this research to compare its accuracy with similarity matching method. As explained in previous section, concept scores are calculated for new record r. This new record r is approximated using PMU measurements at observed locations and the normalized placeholder values for unobserved bus measurements. The placeholder values are calculated by computing average of every channel in the normalized library L for unobserved buses, as explained earlier.

In this method, concept scores are run backwards through SVD operations. This is given by equation

$$R_{est\ 1Xn_m} = C_{1Xn_0} * \bar{S}_{n_oXn_o} * \bar{V}'_{n_mXn_o} \qquad (3.7)$$

where the subscripts denote the dimension of each matrix and $R_{est\ 1Xn_m}$ is the estimated new record. By substituting the expression for concept score C_{1Xn_0} given by (3.5), the above equation is further simplified into -

$$R_{est\ 1Xn_m} = r_{1Xn_m} * \overline{V}_{n_mXn_o} * \bar{S}^{-1}_{n_0Xn_0} * \bar{S}_{n_oXn_o} * \bar{V}'_{n_mXn_o} \qquad (3.8)$$

which is reduced to derive the following equation used for filtering method.

$$R_{est\ 1Xn_m} = r_{1Xn_m} * \overline{V}_{n_mXn_o} * \bar{V}'_{n_mXn_o} \qquad (3.9)$$

The Mean Absolute Error between estimated new record R_{est} and original test record R_{test} is calculated using equation (3.6). MAEs for Filtering Method and Similarity Matching are further compared and discussed in the next chapter.

3.2.6 Filtering Method Algorithm

Algorithm of the filtering method using SVD for grid estimation is explained in this section.

i. Perform power flow on power system for various values of load and generation to create a library *L* of multiple records and 6 additional records for voltage magnitude and angle, active and reactive power, measurements at all buses.
ii. Normalize power flow library *L* as explained in section 3.2 -Data Scaling.
iii. Run SVD on normalized power flow library using equation (3.1) to obtain *U*, *S* and *V* matrices.

iv. Top two concepts are chosen to retain high energy concepts as explained in Section 3.1 to obtain reduced matrices $\overline{U}$, $\overline{V}$ and $\bar{S}$.

v. A single record R_{test}, is selected from the set of six test records and normalized as explained in Section 3.2.

vi. New PMU measurements from a single bus or multiple buses for $|V|$, δ, P and Q are recorded. These are measurements for the respective observed buses extracted from R_{test} to form a vector 'm'.

vii. Remaining unobserved channels for $|V|$, δ, P and Q measurements are approximated by placeholder values determined by calculating mean of the normalized measurements for every channel in L to form a vector 'p'.

viii. A new record 'r' is then obtained from m and p by $r = [p\ m]$.

ix. New record r is run backwards using equation (3.9) to obtain R_{est}.

x. Steps v. to ix. are repeated for all six test records.

MAE is calculated by determining error between R_{est} and R_{test} to check the accuracy of system state estimation with limited number of PMU measurements. MAE for filtering method are then compared with Similarity Matching technique in the next chapter.

4 Results

4.1 Similarity Matching Results for 3 Bus System

The 3 bus system power flow library consists of 100 records along rows and 12 channels for the measurements of $|V|$, δ, P and Q at every bus. With a single PMU located at bus 3, state of the system at remaining buses of 1 and 2, is to be estimated. Measurements at buses 1 and 2 are approximated by using the placeholder data. A new record consisting of these placeholder measurements and PMU measurements is obtained. Similarity matching algorithm finds a record most similar to this new record from library L in concept space. Records of the library translated to the concept space are shown in Figure 4.1. The concept scores of library records are plotted for the top two concepts retained as discussed previously.

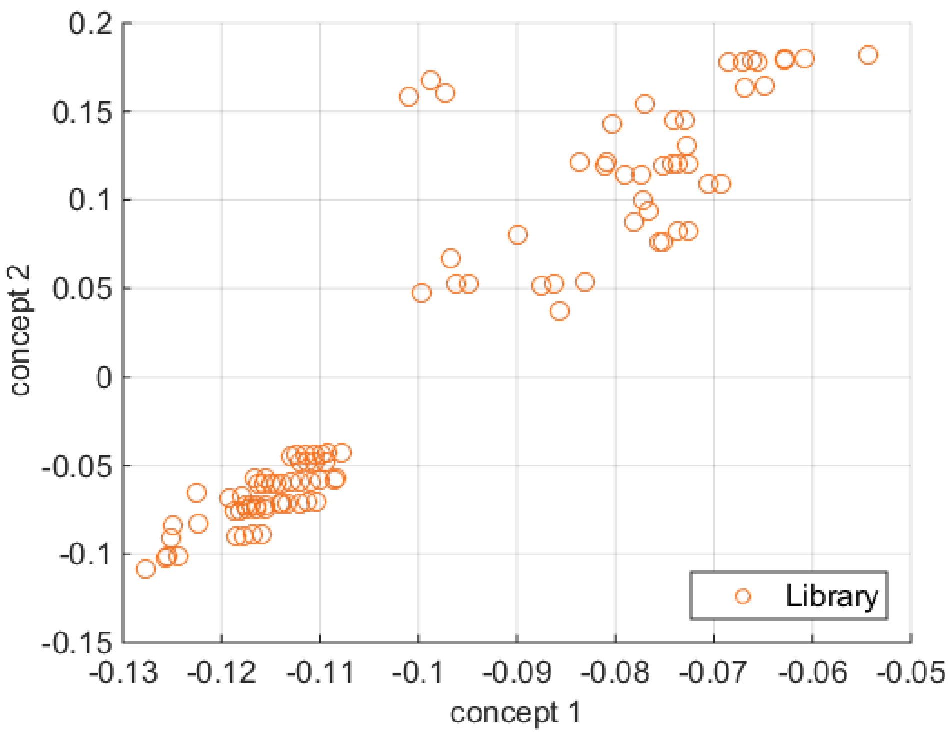

Figure 4.1. Concept scores for 100 records - 3 bus system.

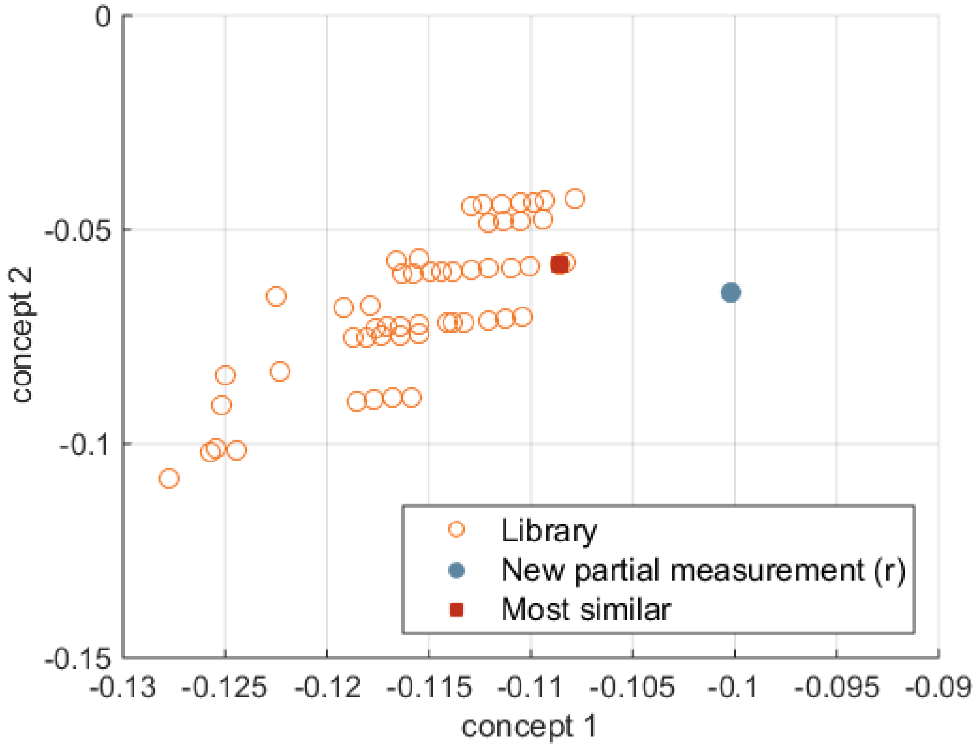

Figure 4.2. Enlarged image to show the new record '*r*' and its most similar library record.

Figure 4.2 shows how the new record is matched to its closest record of library in concept space by determining 2-norm. This figure is an enlarged version of Figure 4.1 to show library record - 42, which is the closest match to test record 5.

Table 4.1 shows results tabulated for all the six test records. It shows most similar record obtained from library L for each test record along with the corresponding 2-norm distance between them. This table lists the MAE values calculated by finding error between every test record and its estimated record. As we know that test records are normalized to a scale with minimum and maximum bound set to 0 and 100 respectively, the MAE values calculated are indicative of percentage error. While we can see that the first test record is estimated with an error percentage of 0.35, the second test record is estimated with an error percentage of 25.67.

Thus for three bus system power flow library consisting of 100 records, the percent error for all six test records lies between a range of 0 to 26.

Test Record #	Most Similar Record	2-Norm Distance	MAE in %
1	47	0.0203	0.35
2	98	0.0071	25.67
3	69	0.0223	11.69
4	46	0.0077	8.96
5	42	0.0107	13.65
6	22	0.0087	13.66

Table 4.1. Results for 3 bus system with 1 observable bus.

4.2 Case 1 (3 Observable Buses)

4.2.1 Similarity Matching Results for IEEE 14 Bus System

Similarity matching algorithm is next demonstrated for the IEEE 14 bus system. For the first case of 3 observable buses, PMUs are assumed to be located at buses 12, 13 and 14. The state of remaining unobserved buses - from 1 to 11 is to be estimated. The 111 records of power flow library are translated to a concept space as shown in Figure 4.3. This figure shows concept scores of all library records for the top two concepts.

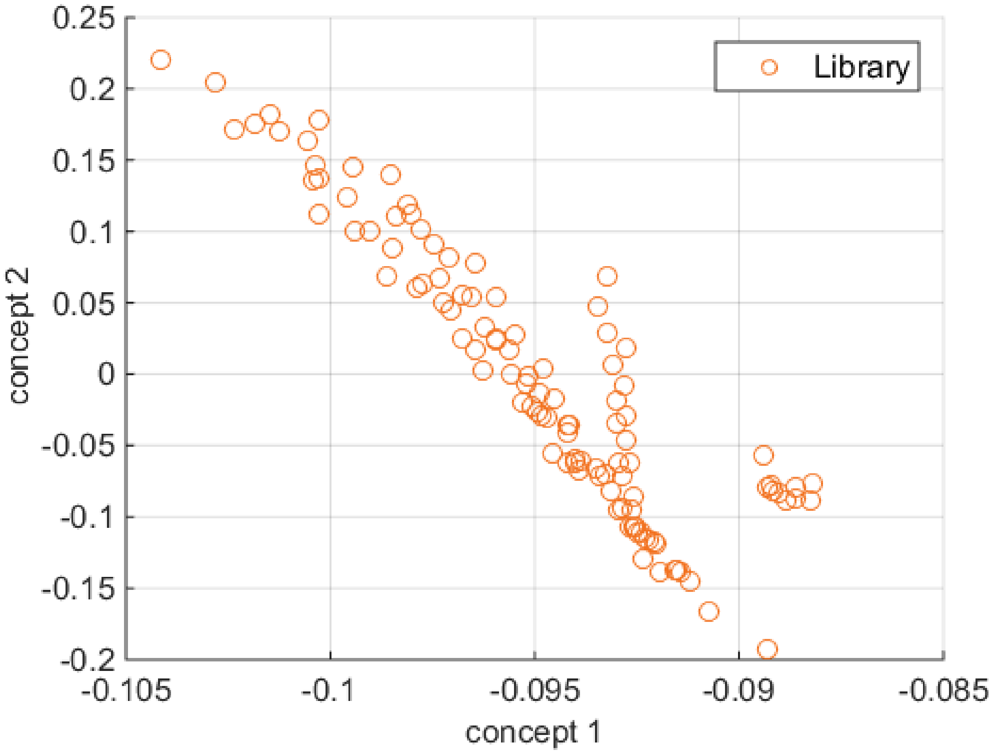

Figure 4.3. Concept scores for 111 records - 14 bus system.

Figure 4.4 shows the closest library record to fifth test record obtained by determining the 2-norm distance. It is a zoomed version of Figure 4.3 with the approximated new record and its most similar record from library in concept space. Approximated new record is then replaced by this similar record to estimate the state of system for buses from 1 through 11.

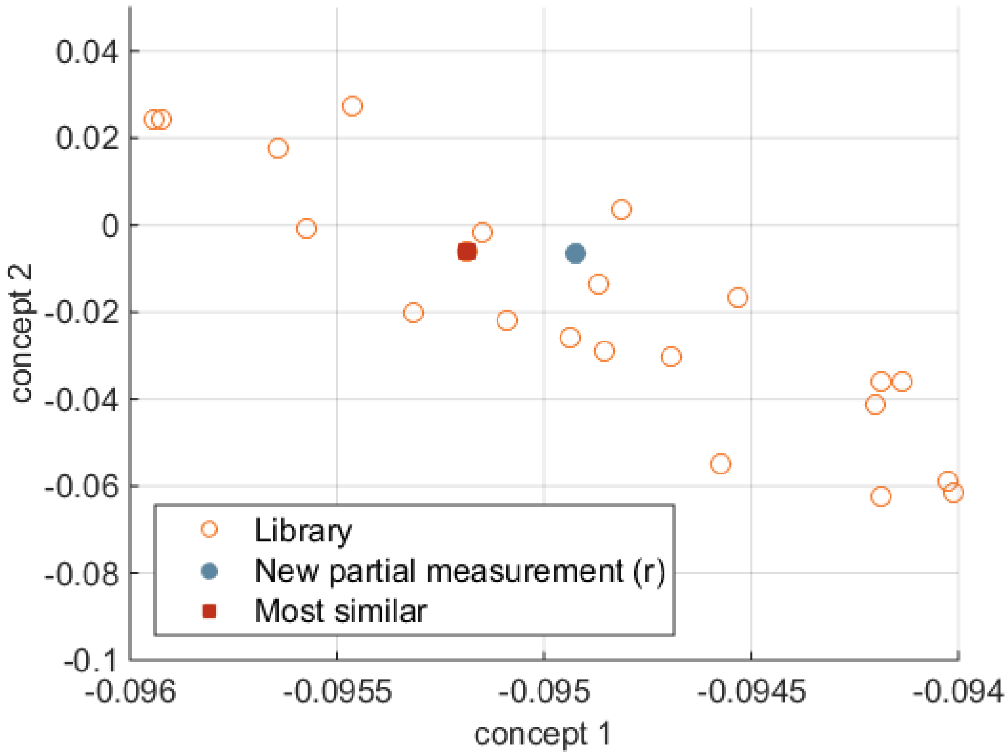

Figure 4.4. Enlarged image to show the new record '*r*' and the most similar library record: Case 1 - 14 bus system.

The least 2-Norm distance between new record and library record are tabulated in Table 4.2 for all six test records. The table also shows MAE calculated for every test record. It can be seen from this table that the errors for all six test records lie in a range of 4 to 13 percent. This means that the accuracy of estimating the state of system for unobserved buses of 1 through 11 is in range of 87 to 96 percent.

Test Record #	Most Similar Record	2-Norm Distance	MAE in %
1	65	0.002	4.62
2	35	0.0007	4.97
3	7	0.0028	5.96
4	34	0.0007	8.93
5	18	0.0006	4.52
6	49	0.0015	13.43

Table 4.2 Results for 14 bus system: Case 1.

4.2.2 Similarity Matching Results for OSU Power System

The OSU system power flow library consists of 125 records with a total number of 844 channels. Similar to case 1 of 14 bus system, 3 PMUs are considered located at buses 140, 209 and 244 on OSU campus. The state of system at remaining buses is to be estimated. Figure 4.5 shows the 125 records of power flow library translated to concept space for top two concepts. The most similar record obtained on running similarity matching algorithm for the fifth test record is shown in Figure 4.6. The plot is an enlarged version of Figure 4.5 to show approximated new record and its most similar record from library *L* in concept space. The approximated new record is then replaced by the most similar record, to obtain estimates for all 283 unobserved buses of the system.

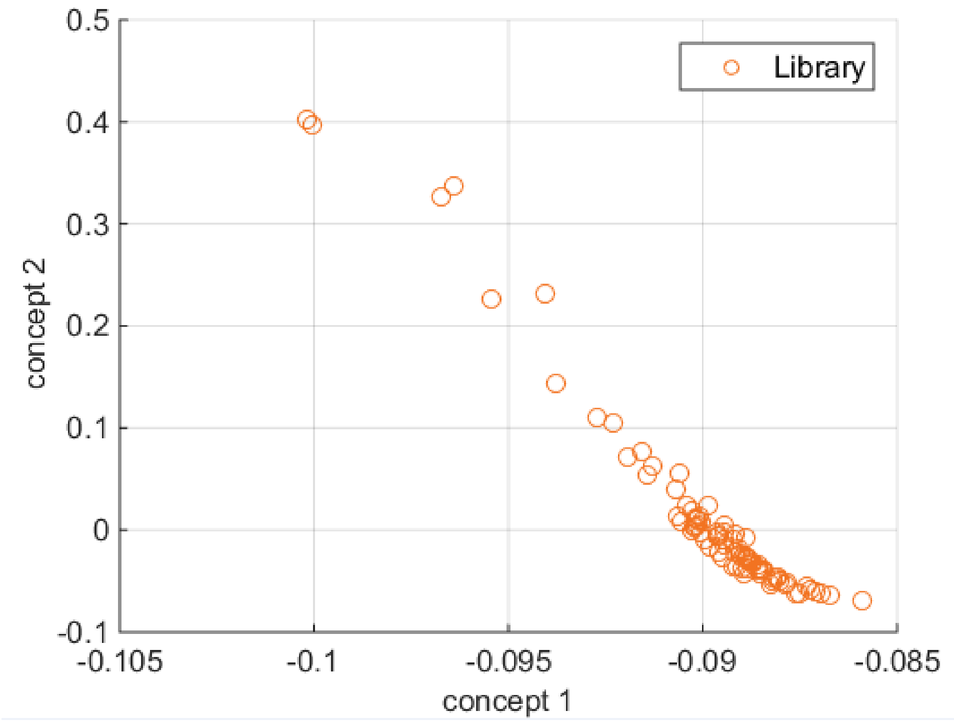

Figure 4.5. Concept scores for 125 records - 286 bus system.

The results obtained for all six test records for case 1 are tabulated in Table 4.3. MAE calculated for all six test records lie between 2 and 20 percent. This implies that accuracy of estimating the state of system is between 80 and 98 percent for six test records.

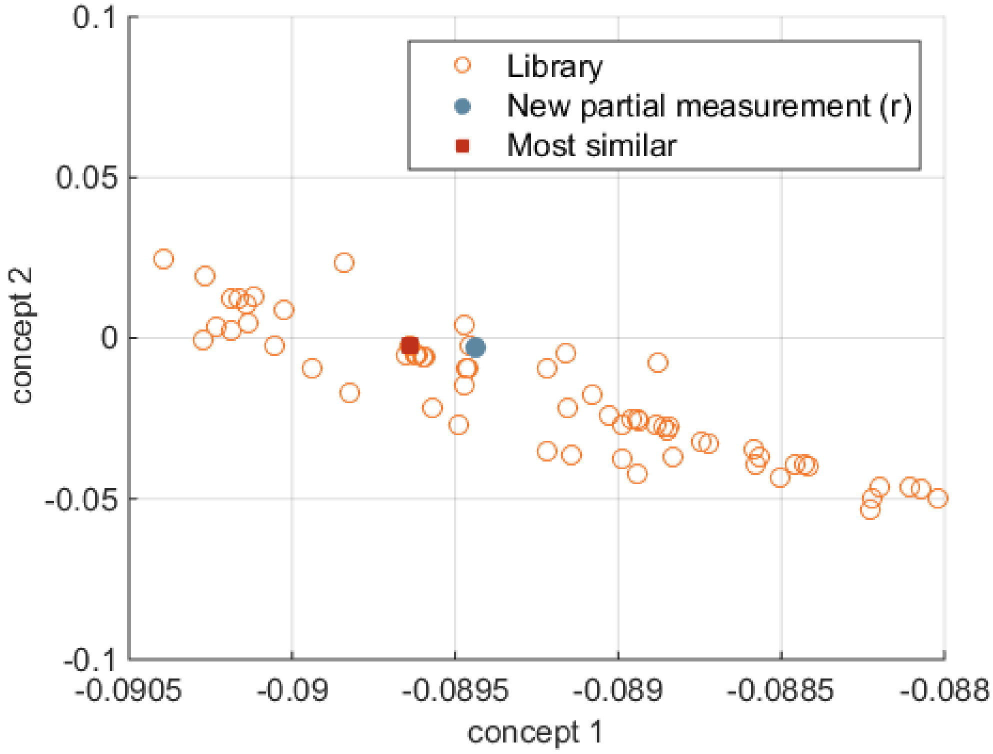

Figure 4.6. Enlarged image to show the new record '*r*' and the most similar library record: Case 1 - 286 bus system.

Test Record #	Most Similar Record	2-Norm Distance	MAE in %
1	31	0.0007	2.74
2	77	0.0009	11.05
3	9	0.0015	19.54
4	31	0.0007	2.68
5	97	0.0004	5.53
6	30	0.0006	4.43

Table 4.3. Results for 286 bus OSU System: Case 1.

4.3 Case 2 (6 Observable Buses)

4.3.1 Similarity Matching Results for IEEE 14 Bus System

It is interesting to further see an impact of a higher number of observed buses for a substantial power network on the similarity matching algorithm. To test similarity matching algorithm with more number of observed buses, a second case of six observed buses is considered. For this case, buses 9 through 14 are considered to have PMUs installed. The state of system for buses 1 through 8 is to be estimated. Similar to Case 1, the approximated new record is plotted with its most similar library record obtained from similarity matching. Figure 4.7 shows concept scores of approximated new record and the most similar library record plotted for fifth test record. We can see from Table 4.4, record 34 is closest library record for the new approximated record.

Similarly, results of most similar record obtained for all six test records are tabulated in Table 4.4 for the 14 bus system with 6 observed buses. The corresponding 2-Norm distance and MAE are shown in this table. It is clearly observed by comparing Table 4.2 and Table 4.4 that the percent MAE reduces significantly with more number of observed buses in a system. This means the percent accuracy of estimating measurements at unobserved buses increases. Practically, this is an ideal situation in real time where more number of PMUs makes the state estimation more accurate. Thus the algorithm developed here follows this ideal real time scenario.

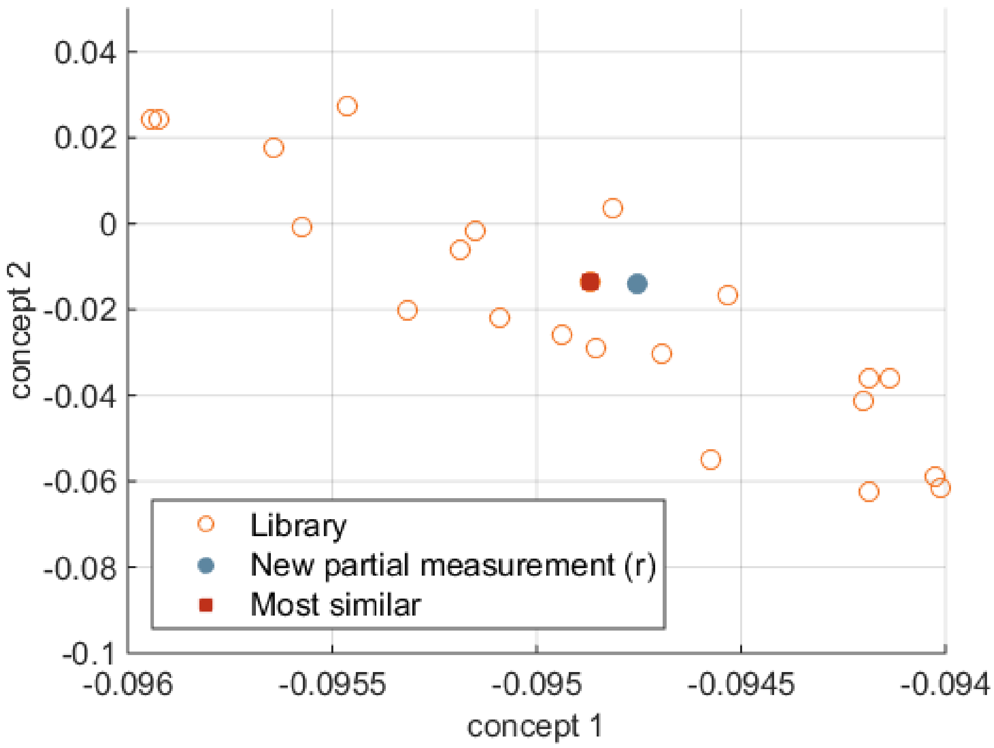

Figure 4.7. Enlarged image to show the new record 'r' and the most similar library record: Case 2 - 14 bus system.

Test Record #	Most Similar Record	2-Norm Distance	MAE in %
1	12	0.0030	6.12
2	58	0.0008	3.67
3	49	0.0012	1.90
4	64	0.0017	5.92
5	34	0.0004	1.75
6	5	0.0035	9.10

Table 4.4. Results for 14 bus system: Case 2.

4.3.2 Similarity Matching Results for OSU Power System

Similar to case 2 of 14 bus system, the impact of larger number of buses with PMUs is demonstrated on 286 buses system. Along with PMUs located at buses 140, 209 and 244, buses 245, 246 and 247 are assumed to be observable with PMUs. The optimal placement of PMUs for these additional 3 buses location has not been considered but is a potential topic of future research. With PMUs located at above 6 buses, the state of OSU system is to be estimated at remaining 280 buses. The closest library record to an approximated new record for test record 6 is shown in Figure 4.8. This closest record is then used to replace the approximated new record to estimate the state at 280 buses. Results for all 6 test records are tabulated in Table 4.5.

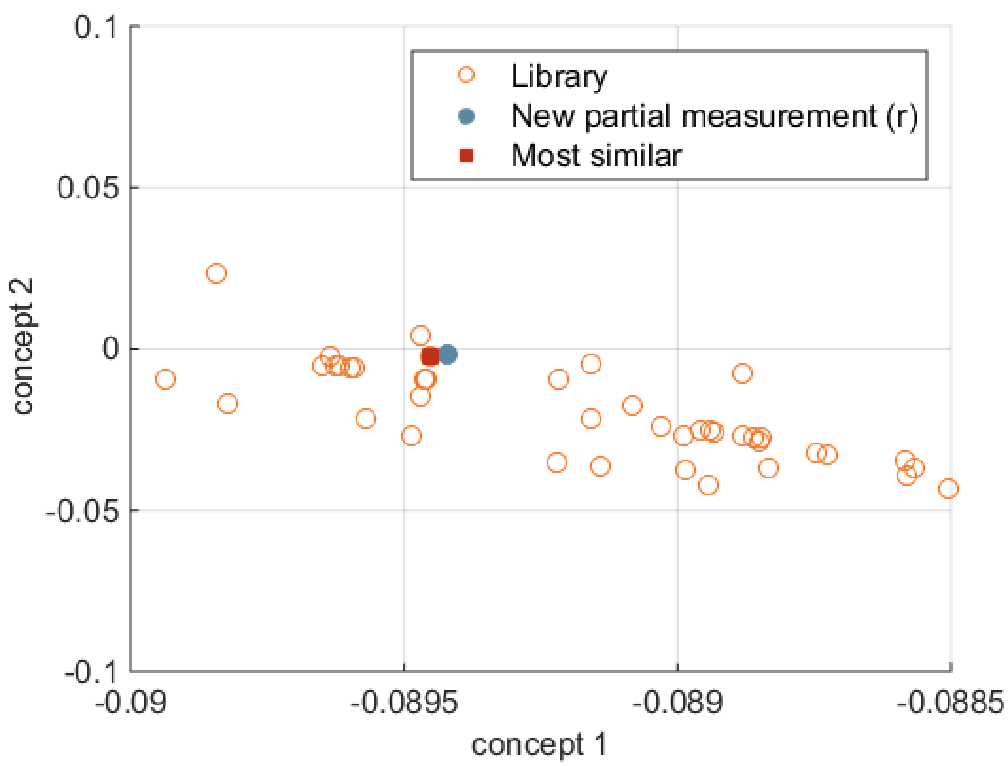

Figure 4.8. Enlarged image to show the new record '*r*' and the most similar library record: Case 2 - 286 bus system.

For 6 observable buses in a 286 bus system, the percent error between estimated record and test record is in the range of 2 to 17. This implies that the accuracy of estimating the state of system is in a range of 83 to 98 percent. On comparing results for 286 bus system with 3 and 6 observable buses we see an improved accuracy in estimating the state of system. It is interesting to note this behavior of the similarity matching algorithm developed, which follows an ideal scenario of better accuracy of state estimation with more number of PMUs in a system.

Test Record #	Most Similar Record	2-Norm Distance	MAE in %
1	31	0.0008	2.73
2	9	0.0017	10.42
3	49	0.0004	16.99
4	31	0.0008	2.67
5	97	0.0003	5.47
6	30	0.0004	4.43

Table 4.5 Results for 286 bus system: Case 2.

To observe an impact of increasing the number of PMUs on similarity matching state estimation algorithm, the average MAEs for all six test records are plotted for IEEE 14 bus system and OSU 286 buses system. This figure also shows the minimum and maximum MAE of the six test records. For the 14 bus system, average MAE for six test records is plotted for 7 observed buses. It is observed that with an increase in number of observed buses the average MAE decreases. This implies that with more number of observed buses, the accuracy of grid estimation increases. The values for minimum and maximum MAE for the six test records also reduced towards the right side of the plot. Similarly for 286 bus OSU system, average MAE for six test records is plotted against 140 observed buses. It is clearly observed that although there is a very slow decrease, the slope shows a reducing trend of MAE when a large number of PMUs are placed in system.

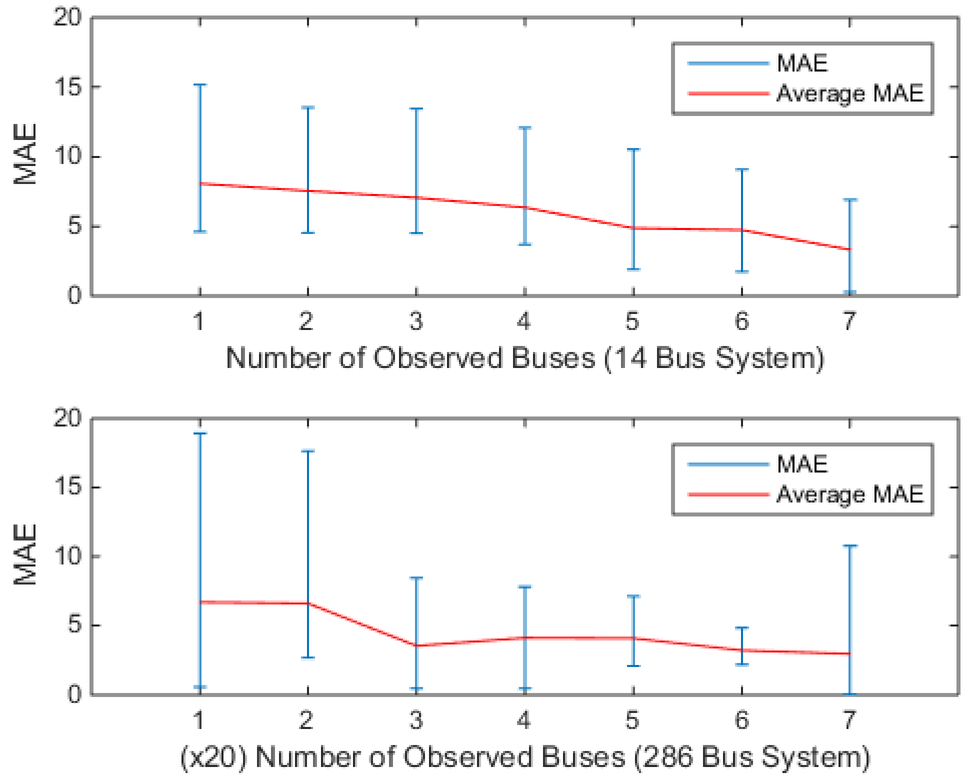

Figure 4.9. Average MAE for increasing number of observed buses.

4.4 Filtering Method Results

4.4.1 3 Bus System

The same 3 bus system with 12 measurement channels and 100 records is used to demonstrate the second algorithm - Filtering Method using SVD. In this method, simple matrix operation is performed on the approximated new record (with PMU measurements and placeholder values) to estimate the state of system at all unobserved buses as explained in section 3.2.6. The results obtained for the error between estimated record and test record are tabulated in Table 4.6. As we know, MAE can be interpreted in terms of percentage because the library data is Quantile Normalized with a minimum set to 0 and maximum set to 100. From Table 4.6 we can see that, the percent error lies in a range of 6 to 22 percent. With an accuracy of 78 to 94 percent, filtering method of SVD is proved to be an equally useful technique to estimate the state of an incomplete observable system.

Test Record #	MAE in %
1	6.07
2	16.05
3	21.85
4	15.65
5	14.46
6	14.47

Table 4.6. Filtering Method MAE for 3 bus system.

4.4.2 IEEE 14 Bus System

The IEEE 14 bus system used for Similarity Matching is now used to demonstrate the second algorithm - Filtering method based on SVD technique. The same power flow library consisting of 111 records and 60 channels along with 6 additional test records are used for filtering method. The state of system for all unobserved buses is estimated by performing simple matrix operations on a new partial record using the algorithm explained in Section 3.2.6. This algorithm is demonstrated for both the cases described for similarity matching method. Case 1 shows results obtained from state estimation for 3 observable buses and case 2 shows results obtained for 6 observable buses. It is interesting to note from Table 4.8, that MAE values for case 2 are smaller than MAE values for case 1 which proves an improved accuracy for 6 PMUs in a 14 bus system. In an ideal situation with more number of observed buses in a system, state estimation is more accurate. This comparative study for both cases is conducted to check the performance of algorithm and compare its behavior to the ideal real time scenario.

Test Record #	Case 1- MAE in % (3 Observable Buses)	Case 2- MAE in % (6 Observable Buses)
1	4.05	3.80
2	9.13	8.32
3	4.89	4.04
4	10.53	9.13
5	4.30	3.71
6	15.16	12.74

Table 4.7. Filtering Method MAE for 14 bus system.

4.4.3 OSU 286 Bus System

Results for MAE values obtained from filtering method of SVD for the OSU system are tabulated in Table 4.6. Results are obtained for case 1 and case 2 with 3 and 6 observable buses respectively. It is observed that the highest error for case 1 and case 2 is 17 percent. This shows filtering method is an equally good technique for estimating the state of system. It is comparatively a simpler technique as it involves simple matrix operations after decomposing the power flow library using SVD. Comparative study for case 1 and case 2 can be used to show the increase in state estimation accuracy with increase in the number of PMUs from 3 to 6. Case 1 for 286 bus OSU system with 3 observable buses shows that MAE lies in a range of 2 to 20 percent. Whereas in case 2, it can be seen that MAE for all six test records lies in a range of 2 to 17 percent. Although there is an extremely small decrease in the error percent with an increase in the number of PMUs from 3 to 6 for a large 286 bus system, the reduced MAE is evident for first four test records.

Test Record #	Case 1- MAE in % (3 Observable Buses)	Case 1- MAE in % (6 Observable Buses)
1	4.14	4.13
2	8.79	8.37
3	17.07	17.03
4	4.17	4.15
5	2.95	2.95
6	4.97	4.97

Table 4.8. Filtering method MAE for OSU power system.

To further check the impact of increasing number of PMUs in the system on the accuracy of state estimation using Filtering method, average MAEs for all six test records are plotted against the number of observed buses. Figure 4.9 shows the plot for average MAEs and number of buses with PMUs. This plot also shows the minimum and maximum MAE for the six test records. Optimal placement of PMUs based on the electrical distance is not considered for this plot and is a potential topic for future research. It can be seen in the first plot of Figure 4.10, that average MAE for six test records has reduced with an increase in number of buses with PMUs. This clearly confirms the ideal situation of an increased state estimation accuracy with a higher number of observed locations in a power grid. Also, the minimum and maximum MAE for one observed bus is seen to be close to 5 and 15 respectively. As the number of observed buses increases along the right side of plot, the minimum and maximum MAE is also seen to reduce. For the second plot in Figure 4.8, although the decrease is in extremely small steps, it can be seen as the number of observed buses increased to 140 for the larger system of 286 buses, the MAE clearly follows a decreasing trend. With these two plots it can be shown that filtering method algorithm confirms the idea of achieving a better accuracy of grid estimation with higher number of PMUs in a system.

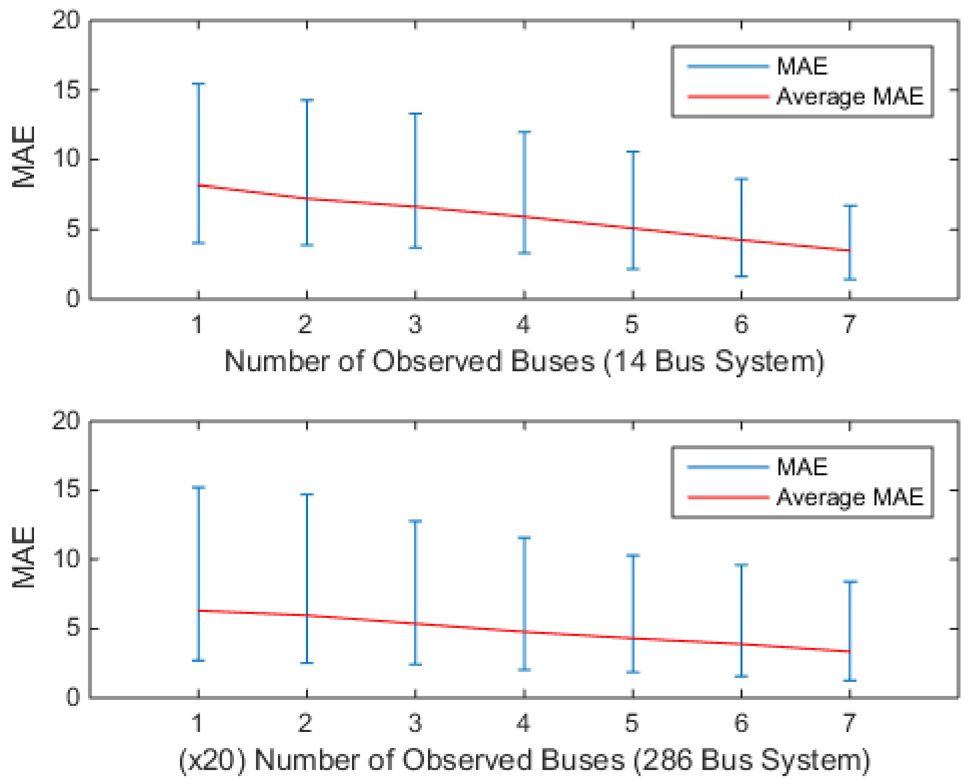

Figure 4.10. Average MAE for 7 and 140 observed buses for each system.

4.5 Summary

The results of similarity matching and filtering method for substantial sized grid consisting 14 bus as discussed above, are summarized in the plot shown in Figure 4.11. A comparison between both the methods – similarity matching and filtering, - applied to an IEEE 14 bus system can be obtained from this plot. It is clearly seen, that filtering method is equally accurate while compared to similarity matching.

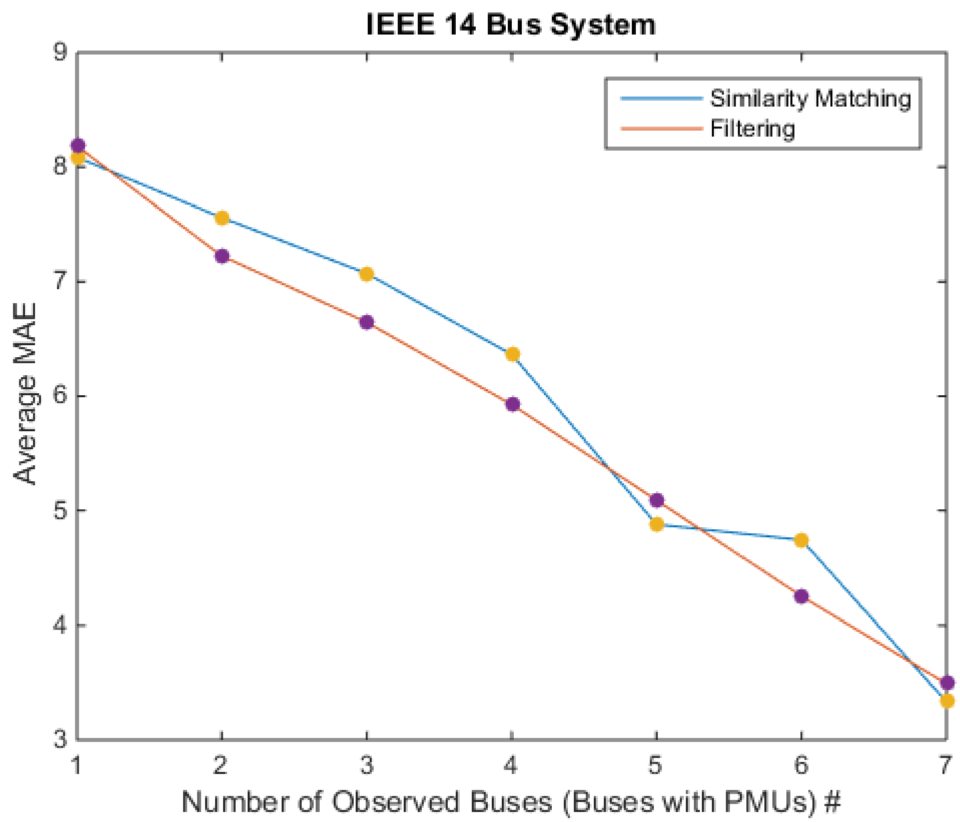

Figure 4.11. Average MAE for 7 observed buses - 14 bus system. This figure shows the plots for both similarity matching and filtering method.

Similarly the results of similarity matching and filtering methods for OSU system with 286 buses are shown in Figure 4.12. It shows the decreasing trend in the MAE for 140 observed buses considered. The dip seen in similarity matching curve could be due to the optimal placement of

PMU at bus 60. On comparing similarity matching and filtering for 286 bus OSU system, it is seen that although similarity matching has marginally better accuracy, filtering is an equally accurate technique suitable for state estimation.

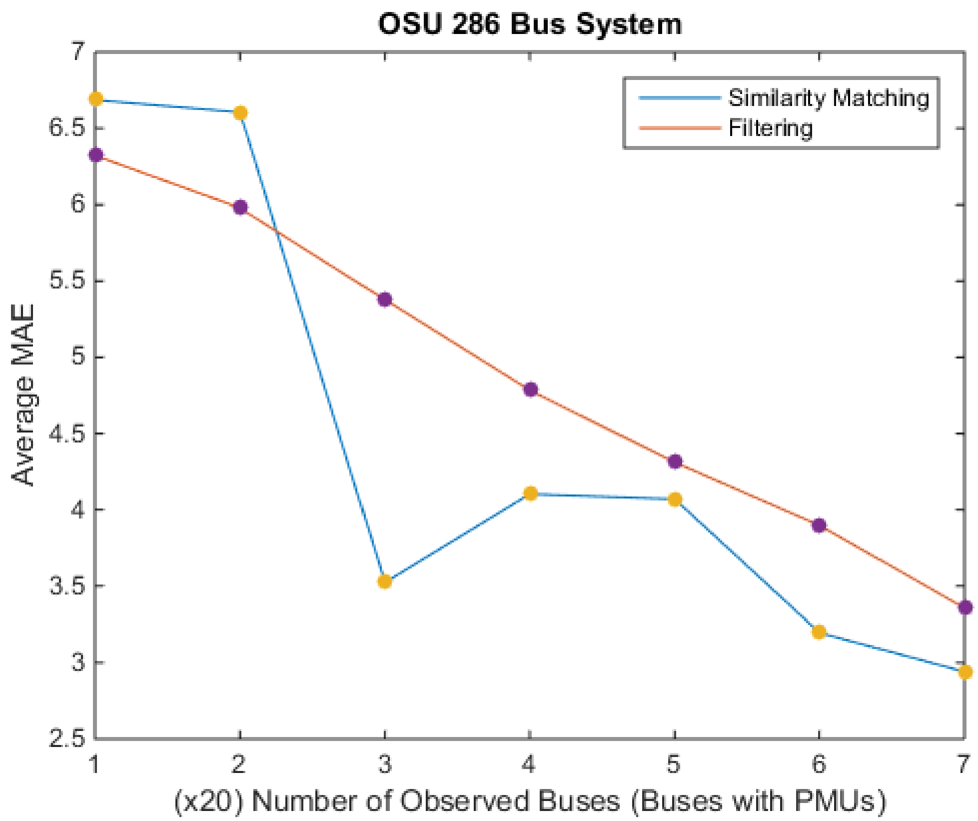

Figure 4.12. Average MAE for 140 observed buses - 286 bus system. This figure shows the plots for both similarity matching and filtering method.

5. Conclusion

Grid estimation for a system with incomplete observability at real time is not a simple task considering the limited number of PMU measurements available. This research explores SVD methods to estimate the state of a grid at real-time or near real-time with a few PMUs deployed across the system. Two approaches of state estimation of grid investigated in this research are Similarity Matching and Filtering employing the technique of SVD.

An important conclusion that can be drawn from this study is, that normalization of the power flow library prior to using it for state estimation is not a trivial concept. It is shown that normalizing the data of $|V|$, δ, P and Q belonging to different measurement units to a uniform and more comprehensible scale is critical. Scaling the data using Quantile Normalization with the minimum bound set to zero and the maximum set to 100 allows the interpretation of mean absolute errors calculated in terms of percentage.

The 3 bus system used to demonstrate the grid estimation method for a partially observable record shows that when a single bus of the system is made observable, the errors lie within a range of 0 to 26 percent for the six test records. The first test record had a mean absolute error of almost zero percentage which means the accuracy of estimating state of the system was almost 100%. At the same time, the MAE of the second record is 25.6 percent which means the accuracy with which the state of unobserved buses is estimated is 74.4 percent. For the filtering method, MAE lies in a range of 6 to 22 percent. This implies a 78 percent to 94 percent accuracy in estimating the state of a 3 bus system with a single observable bus. Thus both the methods - similarity matching and filtering method using SVD are equally accurate in estimating the state of system.

The IEEE 14 bus system is next used to demonstrate state estimation using SVD for a comparatively larger substantial grid before applying it to the OSU system. For the first case of 3 observable buses for 14 bus system, accuracy of estimating the state of system is in the range of 86 percent to 95 percent. Similarly for filtering method, the least MAE is 4.05 percent for the first test record and the highest MAE is 15.16 percent for the test record 6. With an error percent lying in a range of 4 to 16 percent for the 14 bus system with 3 buses observable, the best accuracy of estimating the state of 11 unobserved buses is 96 percent. Thus similarity matching method and filtering method can both be valuable methods to estimate the state of a system.

Further the algorithms are applied to estimate the state of OSU campus with sparse distribution of PMUs across the campus. For 286 bus system we can see that percentage of mean absolute error between estimated record R_{est} and test record R_{test} lies between 3 and 20 percent. This indicates 97 percent accuracy can be obtained using similarity matching method. While the results of filtering method for the same case with 3 observable buses show a maximum MAE of 17 percent. This means the accuracy of estimating state for unobserved 283 buses of system is at least 83 percent. Thus both the SVD methods prove to be a great tool to estimate the state of a grid at real time or near real-time with partial observability.

The next conclusion that can be drawn from this research is that with an increase in the number of observed buses with PMUs in a system, the MAE error accuracy reduces, thus the state estimation for $|V|$, δ, P and Q measurements is more accurate. This is an ideal real-time scenario which is confirmed by the algorithms developed through this research. For both systems of 14 buses and 286 buses, a second case with 6 observable buses each was studied. The results clearly show a reduced MAE for this case as compared to first case with 3 buses with PMUs. Figure 4.8 and 4.9 confirm that both the algorithms for SVD developed through this research successfully follow the concept of better state estimation of a grid with more number of observed measurements. Although a slow decrease in average MAE for the larger 286 bus system is observed, the clear decreasing trend with a larger number of PMU observed buses is evident. From Figure 4.11 and Figure 4.12, it can be concluded that similarity matching method although marginally, has better accuracy compared to filtering method.

Finally it can be concluded that both state estimation methods based on SVD technique are relatively faster compared to other machine learning techniques. The time taken to decompose power flow library by performing SVD is 0.485 seconds. Similarity matching method requires comparatively more number of computations after decomposing the power flow library to calculate concept scores and then find a closest matching library record. This algorithm used for estimating state of system requires a computation time of about 0.57 seconds. On the other hand filtering method involves simple matrix operations after the library has been decomposed which makes it comparatively faster. Thus the time taken to obtain an estimate using filtering method is estimated to be 0.52 seconds. The computation speed of both the methods can be compared based on time taken to estimate the state of system. It can be observed that the decomposing of large power flow data using SVD is computationally expensive and comparatively more time consuming. Once the library is decomposed, following matrix operations are comparatively simple and fast. Based on the time taken by each algorithm, it can be clearly concluded that filtering method is relatively faster compared to similarity matching method. Therefore it is demonstrated that the SVD techniques have a valuable application to power system estimation particularly for applications in which speed and sparseness are key.

References

[1] Wikipedia - Power Flow Studies.

[2] Glover, J. Duncan, Mulukutla Sarma, and Thomas Overbye. *Power System Analysis & Design, SI Version*. Cengage Learning, 2011.

[3] Electrical Engineering Power System Analysis [Online] Available - http://nptel.ac.in/courses/Webcourse-contents/IIT-KANPUR/power-system/ui/Course_home-4.htm.

[4] GE - Digital Energy "Energy Management Systems" [Online] Available https://www.gedigitalenergy.com/products/brochures/uos/XA21_EMS.pdf.

[5] Abur, Ali, and Antonio Gomez Exposito. *Power system state estimation: theory and implementation*. CRC Press, 2004.

[6] Ashwani Kumar Chandel Associate Professor, NIT - Hamirpur - "Power System State Estimation" [Online] Available. http://www.iitmandi.ac.in/ciare/files/9_Ashwani_Power_System_State_Estimation.pdf.

[7] Advanced Study Power - Gaussian Distribution [Online] Available - http://www.asp.ucar.edu/colloquium/1992/notes/part1/node15.html.

[8] "Synchrophasors Technology and their Deployment in the Recovery Act Smart Grids Program" [Online] Available - https://www.smartgrid.gov/sites/default/files/doc/files/Synchrophasor%20Report%2008%2009%202013%20DOE%20%282%29%20version_0.pdf.

[9] "Synchrophasor Projects and Data Challenges at BPA", Tony Faris, [Online] Available - http://cialab.ee.washington.edu/nwess/2014/speakers/6B_NWESS%202014_Synchrophasors%20BPA_Faris.pdf.

[10] Bonneville Power Administration - News Releases - Synchrophasor break throughsmark quantum leap in grid reliability, [Online] Available - http://www.bpa.gov/news/newsroom/Documents/20141229-NWPPA-EO-Column-Synchrophasor.pdf.

[11] "US Energy Information Administration", Energy - Phasor Measurement Units in North American Power Grid [Online] Available - http://www.eia.gov/todayinenergy/detail.cfm?id=5630

[12] Zivanovic, R., Cairns, C., " Implementation of PMU Technology in State Estimation: an Overview", IEEE Africon, 1996

[13]BPA, TIP 328: Real Time Load Composition Estimation- [Online] available- http://www.bpa.gov/Doing%20Business/TechnologyInnovation/TIPProjectBriefs/2015-TIP-328.pdf.

[14] Chen Huo, Jiajia Song, K Wagner, G Harrold, E Cotilla-Sanchez. "Integrating Synchrophasor Technology with the Oregon State University Campus Smart Grid Project", IEEE SusTech, 2014.

[15] F. G. Longatt [Online] Available- http://fglongatt.org/OLD/Test_Case_IEEE_14.html

[16] Leskovec, Jure, Anand Rajaraman, and Jeffrey David Ullman. *Mining of massive datasets*. Cambridge University Press, 2014.

[17] Ana Belen Barragans Martinez, Jose J. Pazos Arias, Ana Fernandez Vilas, Jorge Garcia Duque and Martin Lopez Nores, "What's on TV Tonight? An Efficient and Effective Personalized Recommender Systems of TV Programs", IEEE Consumers Electronics Society 2009.

[18] Elektrisidad pilipinas IEEE test systems in PowerWorld - Edwin Cano [online] available -http://elektrisidadpilipinas.blogspot.com/2011/03/ieee-test-systems-in-powerworld.html.

[19] C. Madtharada, S. Premrudeepreechacharnb, N. Watson," Power system state estimation using singular value decomposition", Electric Power Systems Research, Vol.67, No. 2, pp. 99-107, 2003.