UNIVERSITY OF CALIFORNIA SAN DIEGO

The Power of Synchrophasors: Electric Grid Monitoring, Control and Optimization

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Applied Mechanics)

by

Sai Akhil Reddy Konakalla

Committee in charge:

Professor Raymond A. de Callafon, Chair
Professor Robert R. Bitmead
Professor Jan Kleissl
Professor Bhaskar D. Rao
Professor George R. Tynan

2019

The dissertation of Sai Akhil Reddy Konakalla is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2019

DEDICATION

To my parents.

EPIGRAPH

That which seems like poison at first, but tastes like nectar in the end; this is the joy of sattva, born of a mind at peace with itself

—Bhagavad Gita, verse 18.37

O my Lord! Increase me in my knowledge

—Quran 20:114

For wisdom will come into your heart, and knowledge will be pleasant to your soul; discretion will watch over you, understanding will guard you

—Proverbs 2:10-11

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Raymond A. de Callafon for his support as the chair of my committee. Through multiple drafts and many long nights, his guidance has proved to be invaluable.

I would also like to acknowledge my lab mates Amir Valibeygi and Yangsheng Hu without whom my research would have taken much longer time. It is their support that helped me in an immeasurable way.

Chapter 2, in full, is a reprint of the material as it appears in ”Sai Akhil R. Konakalla, and Raymond A. de Callafon. Feature Based Grid Event Classification from Synchrophasor Data, Procedia Computer Science 108 (2017): 1582-1591” and ”Sai Akhil R. Konakalla, and Raymond de Callafon. Optimal filtering for grid event detection from real-time synchrophasor data,” Procedia Computer Science 80 (2016): 931-940”. The dissertation author is the primary investigator and author of these articles.

Chapter 3, in full, is a reprint of the material as it appears in ”Sai Akhil R. Konakalla, Raymond A. de Callafon, Synchrophasor Based Recursive Power Prediction with Real-Time Demand Optimization, to appear in IEEE Transactions on Smart Grids” and ”Sai Akhil R. Konakalla, Raymond A. de Callafon, Recursive power demand prediction based on multi-level clustering of power demand data, Smart Grid Communications (SmartGridComm), 2017 IEEE International Conference on. IEEE, 2017”. The dissertation author is the primary investigator and author of these articles.

Chapter 4, in full, is a reprint of the material as it appears in ”Sai Akhil R. Konakalla, Raymond A. de Callafon, Microgrid Dynamic Modeling and Islanding Control with Synchrophasor Data, to appear in IEEE Transactions on Smart Grids”, ”Yangsheng Hu, Sai Akhil R. Konakalla, Raymond A. de Callafon, Covariance Based Estimation for Reduced Order Models of Microgrid Power Flow Dynamics, IFAC-

PapersOnLine, 51.15 (2018): 903-908 and ”Sai Akhil R. Konakalla, Raymond A. de Callafon, Optimal hybrid power dispatch for distributed energy resources with dynamic constraints, Power and Energy Systems (ICPES), 2017 IEEE 7th International Conference on, 2017”. The dissertation author is the primary investigator and author of the first and third article and is the secondary author of the second article.

Chapter 5, in full, is a reprint of the material as it appears in ”Sai Akhil R. Konakalla, Raymond A. de Callafon, Spectrum Based Optimal Filtering for Short-term Phasor Data Prediction, submitted to IEEE Transactions on on Industry Applications” and ”Sai Akhil R. Konakalla, and Raymond de Callafon, Robustness to missing synchrophasor data for power and frequency event detection in electric grids, Renewable Energy Research and Applications (ICRERA), 2017 IEEE 6th International Conference on. IEEE, 2017”. The dissertation author is the primary investigator and author of these articles.

VITA

2015	B.S. in Electrical Engineering, Indian Institute of Technology, Patna
2016	M.S. in Engineering Sciences (Applied Mechanics), University of California San Diego
2019	Ph.D. in Engineering Sciences (Applied Mechanics), University of California San Diego

PUBLICATIONS

Amir Valibeygi, **Sai Akhil R. Konakalla**, Prabhakar Kumaraguru, Raymond A. de Callafon, “Reduced-order Modeling of Dynamic Coupling and Uncertainties in Microgrid Power Flow,” in progress.

Amir Valibeygi, **Sai Akhil R. Konakalla**, Prabhakar Kumaraguru, Raymond A. de Callafon, “Hardware-In-the-Loop Power Control for Microgrids with Hybrid Distributed Energy Resources,” in progress.

Sai Akhil R. Konakalla, Raymond A. de Callafon, “Spectrum Based Optimal Filtering for Short-term Phasor Data Prediction,” submitted to IEEE Transactions on Industry Applications.

Sai Akhil R. Konakalla, Raymond A. de Callafon, “Synchrophasor Based Recursive Power Prediction with Real-Time Demand Optimization,” to appear in IEEE Transactions on Smart Grids.

Sai Akhil R. Konakalla, Amir Valibeygi, Raymond A. de Callafon, “Microgrid Dynamic Modeling and Islanding Control with Synchrophasor Data,” to appear in IEEE Transactions on Smart Grids.

Hu, Yangsheng, **Sai Akhil R. Konakalla**, and Raymond A. de Callafon. ”Covariance Based Estimation for Reduced Order Models of Microgrid Power Flow Dynamics,” IFAC-PapersOnLine 51.15 (2018): 903-908.

Sai Akhil R. Konakalla, and Raymond A. de Callafon. ”Optimal hybrid power dispatch for distributed energy resources with dynamic constraints,” Power and Energy Systems (ICPES), 2017 IEEE 7th International Conference on. IEEE, 2017.

Sai Akhil R. Konakalla, and Raymond de Callafon. ”Recursive power demand prediction based on multi-level clustering of power demand data,” Smart Grid Communications (SmartGridComm), 2017 IEEE International Conference on. IEEE, 2017.

Sai Akhil R. Konakalla, and Raymond de Callafon. ”Robustness to missing synchrophasor data for power and frequency event detection in electric grids,” Renewable Energy Research and Applications (ICRERA), 2017 IEEE 6th International Conference on. IEEE, 2017.

Sai Akhil R. Konakalla, and Raymond A. de Callafon. ”Feature Based Grid Event Classification from Synchrophasor Data,” Procedia Computer Science 108 (2017): 1582-1591.

Sai Akhil R. Konakalla, and Raymond de Callafon. ”Optimal filtering for grid event detection from real-time synchrophasor data,” Procedia Computer Science 80 (2016): 931-940.

Sai Akhil R. Konakalla, and Ahmad Ali. ”Tuning rules for fractional order controllers for integrating and unstable process models,” IET Digital Library (2015): 3-5.

ABSTRACT OF THE DISSERTATION

The Power of Synchrophasors: Electric Grid Monitoring, Control and Optimization

by

Sai Akhil Reddy Konakalla

Doctor of Philosophy in Engineering Sciences (Applied Mechanics)

University of California San Diego, 2019

Professor Raymond A. de Callafon, Chair

Recent growth in deployment of distributed energy resources (DERs), energy storage systems, and advanced grid control schemes have increased the levels of variability in generation and load conditions over the electric transmission and distribution system. Microgrids are distribution level power systems that operate as a single controllable system in a grid connected or islanded (disconnected) mode. The balance between supply and demand of power is one of the most important requirements of microgrid management. The ability to control such a system hence becomes a very important task in order to stably and smoothly operate the microgrid. One such elegant method to overcome these challenges is the use of the phasor measurement units (PMUs) for wide-area distributed networks capable of measuring time-

synchronized phasor (synchrophasor) data at high rates in the multiples of the main AC frequency (up to 240 Hz for a 60 Hz signal). The importance to use synchrophasors for this application is two-fold: Firstly, the ability for high frequency and synchronized sampling significantly enhances (transient) grid monitoring and also improves performance of the state estimation or (recursive) prediction that can be used for real-time control. Secondly, synchrophasor measurements can be used to coordinate the power exchange between cooperative multiple loads (microgrids) using the time-synchronized measurements.

However, due to the high sampling rates of synchrophasor data, it becomes an unwieldy task to process such huge data from multiple locations. Local (signal) processing of PMU data to detect, store and classify events in the electric grid at the point of measurement is one of the main contributions of the dissertation. To address the demand and supply problem in a microgrid, (recursive) demand prediction and (optimal) generation dispatch techniques using synchrophasors are discussed in the dissertation. In order to design high speed control algorithms for such microgrid systems, derivation of a system model via system identification using data driven techniques would be more convenient than derivation via swing equations. Hence, the dissertation contributes in the estimation of a low order linear time invariant (LTI) model for active and reactive power flow (derived from the PMUs) through the microgrid. Using the derived models, it is shown how real-time PMU measurements can further be used in the islanding (power flow and angle) control of microgrids. Finally, robustness of missing PMU data for control applications is addressed in the dissertation.

Chapter 1

Introduction

1.1 Motivation for Synchrophasor Data

Recent growth in deployment of distributed energy resources (DERs), energy storage systems, and advanced grid control schemes have increased the levels of variability in generation and load conditions over the transmission and distribution system. Large scale decentralization of electricity production and rise in the adoption of independently operable (micro-) grids has made it even more difficult to monitor load or generation perturbations possibly caused by DER power production variability, load switching, (unintended) islanding, faults or line tripping. To overcome these challenges, many distribution systems operators (DSOs) have been focusing on developing methodologies that rely on robust information layers for the grid [1–3]. Complementary to the traditional Supervisory Control And Data Acquisition (SCADA) systems, synchrophasor vector processing systems implemented in (protection) relays, digital fault recorders and specialized Phasor Measurement Units (PMU) can produce time synchronized measurements of 3 phase AC amplitude and phase angle (phasor) of voltage and currents at rates up to 240 Hz compared to SCADA with samples rate of 4 seconds [4]. With increasing speeds of operability of new devices such as smart inverters, knowledge of SCADA sub-interval data becomes very important to measure and react to fast transients (within 4 s). Distribution networks have also increased a need for distributed control methods that require time synchronization between several DERs and loads in the network. The ability of these PMUs to time-synchronized high sampling three phase AC voltage and current phasor measurements from various points in a wide geographical distribution in a power grid with a precise global positioning system (GPS) clock has improved the feasibility of analyzing the operating status of a wide area power grid and have made these inevitable and ubiquitous. Widespread installation of PMUs has hence inspired the development of (real-time) synchrophasor applications [5,6] for system monitoring, event (disturbance) detection and automated grid control.

1.2 Synchrophasor Data

1.2.1 Synchrophasor Data Representation

To explain the basic terminology and the use of phasor data, consider a measurement of a AC signal $x(t)$, which is either a voltage or current signal. In an ideal steady-state operating mode, the signal $x(t)$ is a pure sinusoidal signal $x(t) = A\cos(2\pi ft + \delta)$ with an amplitude A, frequency f, and phase offset δ. Using the phasor representation, the signal $x(t)$ is represented by the real part of the complex number $x(t) = \text{Re}\{Ae^{2\pi ft}e^{j\phi}\}$ where $X = Ae^{j\phi}$ denotes the phasor of $x(t)$.

Although this is an accurate representation of a steady-state sinusoidal signal, any changes in the signal $X(t)$ with respect to the pure sinusoidal representation would mean that the amplitude A and frequency f would change with respect to time. To characterize these time variations we may write $x(t) = A(t)\cos(2\pi f_0 t + \phi(t))$

where $\phi(t) = f_e(t)t + \delta$ denotes a time varying phase shift formulated as a time varying frequency error $f_e(t)$ with respect to the nominal frequency f_0 of the AC signal $x(t)$. This defines the (time varying) frequency of the AC signal as

$$f(t) = f_0 + \frac{d}{dt}\phi(t) = f_0 + f_e(t) + t\frac{d}{dt}f_e(t) \tag{1.1}$$

and makes the phasor a time-varying signal

$$X(t) = A(t)e^{j\phi(t)} \tag{1.2}$$

with respect to the nominal AC frequency f_0.

1.2.2 Power Calculations using Synchrophasor Measurements

For notational brevity and to conform to the development of the modeling and control approaches that are presented in this thesis, the power flow equations given here use a balanced single-phase-equivalent network representation. Phasor (voltage, current) measurements at time intervals $t_k = k*T_s$, where T_s is the sampling time and k is the positive integer used for the time index and recursion are used to calculate real (reactive) power at the point of common coupling at each interval t_k. Typically, T_s is a multiple of $1/60$ seconds for synchrophasor sampling. The (total) apparent power is given by

$$\mathbf{S} = \mathbf{V}\mathbf{I}^* = V_m I_m e^{j(\theta_v - \theta_i)} = P + j\ Q \tag{1.3}$$

The real power at a given instant t_k can be calculated as

$$P(t_k) = V_m(t_k)I_m(t_k)cos(\theta_v(t_k) - \theta_i(t_k)) \tag{1.4}$$

Similarly, the reactive power at a given instant t_k can be calculated as

$$Q(t_k) = V_m(t_k)I_m(t_k)sin(\theta_v(t_k) - \theta_i(t_k)) \tag{1.5}$$

where $\mathbf{V} = V_m e^{j\theta_v}$ and $\mathbf{I} = I_m e^{j\theta_i}$ are the voltage and current phasors respectively and $\mathbf{V}, \mathbf{I} \in \mathbb{C}$ where $|\mathbf{V}| > 0, |\mathbf{I}| > 0 \in \mathbb{R}^n$ and $\theta_v, \theta_i \in (-\pi, \pi]$.

1.3 Problem Formulation

Although there are plethora of synchrophasor applications, there are also a few issues associated with using PMU data. These issues formulate some of the problems that are addressed in this thesis.

- First, handling or storing the enormous amount of data from multiple PMUs at high data rates becomes non-trivial for grid monitoring. Manual observations of PMU data time sequences to observe trends or possibly detect anomalies in the data quickly becomes a unwieldy task. Hence automated or semiautomated data analysis techniques to detect and identify faults become essential. Most of the power blackouts are typically caused by a single event that leads to cascading outages. Detecting and classifying these faults (events) can be not only used for operator knowledge but also for control or decision making to prevent potential cascading grid failures.

- Load prediction is a critical component of load and generation balance. Due to the dynamic behavior of power consumption and volatile load changes, a static approach does not always provide accurate power demand prediction results. This can be due to errors in (input) forecasts such as weather and unforeseen events that can drastically affect the power demand. For improvement of dynamic power demand prediction performance, the static predictions of power data have to be (online) recursively updated using fast measurements such as power derived from PMUs.

- Power flow in microgrids is typically controlled using SCADA measurements. However, they are limited by update rates and high latencies which degrade the performance of the control system. Time synchronization poses a big challenge especially in a centralized control architecture. Hence PMUs can be used for high updates rate and also obtain time synchronized GPS measurements to not only control the power flow but also voltage angles to enable fast and smooth intentional switching between islanding and grid connected modes with small voltage and frequency transient fluctuations.

- PMUs measurements may become unreliable due to the presence of different reasons such as electrical sensor failures, protocol failures, interface errors or network communication failures between PMUs and data servers. Such data losses can degrade the performance of a real-time controller and/or lead to system instabilities. Hence methods to automatically predict and replace missing or invalid data become important task in order to reliably use PMU data.

1.4 Contributions

1.4.1 Event Detection

One of the contributions of the thesis is a local synchrophasor based real-time data processing algorithm to automatically detect dynamic transient effects in AC voltage and current signals analyzed by a synchrophasor vector processing systems. The basic output of the algorithm is the time stamp when a dynamic event was detected. The approach is based on real-time discrete-time filtering of phasor data (amplitude and phase angle) to create optimal Filtered Rate of Change (FRoC) signals for each phasor and postulate an event detection algorithm based on the dynamic response of the FRoC signals. The idea of using phasors directly for event detection has been addressed in earlier work [6–8]. However, optimality of the FRoC signals is addressed in this thesis by (recursive) least squares estimation of the parameters of a linear discrete-time filter that minimizes the variance of the FRoC signals for both angle and amplitude data of the phasor for event detection.

1.4.2 Event Classification

The next contribution combines estimation of dynamic event parameters combined with standard clustering methods to build classifiers to quantify power grid oscillation events on synchrophasor data. The event parameters included frequencies, damping and participation factors of each oscillation mode in an event and can be estimated by realization algorithm. Application of the feature based grid event classification to actual PMU synchrophasor data obtained from a microPMU shows realistic clustering based classification of various detected disturbance events. Our future work will implement the event classification algorithm in real-time on a PMU connected client computer to detect and automatically classify grid events.

1.4.3 Demand Prediction and Cost Optimization

Due to the inherent dynamic behavior of power consumption, a static offline approach might not always result in an accurate power prediction. This is more likely seen in commercial buildings where unanticipated demand is often encountered due to a sudden increase or decrease of occupants, unexpected changes in weather or other miscellaneous activities [9]. For improvement of prediction performance, a method for (online) recursive estimation of power is developed in this thesis. Power demand profiles are updated in real-time using a Kalman filter, at every instant any new power or demand measurements are available. Such a recursive estimation of power demand would be efficient because the prediction algorithm would adapt to sudden changes in load to minimize the error between the offline predictions and measured values. One of the important applications of power prediction is for planning and optimal utilization of distributed energy resources (DERs). A part of a contribution of the thesis is a modification of the existing dispatch strategies to run in an operational setting using synchrophasor measurements with possibly 60 times per second resolution.

1.4.4 Dynamic Modeling

In order to design high speed control algorithms for such microgrid systems, derivation of a system model via system identification using input/output system data would be more convenient than derivation via swing equations [10]. Hence, the estimation of a low order linear time invariant (LTI) model for active and reactive power flow (derived from the PMUs) through the microgrid becomes an important control pre-requisite in both grid connected and islanded modes. Given that there are several power generating and consuming elements in the microgrid, each with their own dynamics, the ability to observe the power flow at the main point of common coupling in the microgrid using a PMU forms the basis of modeling in the microgrid from a control stand-point. One of the main contributions of the thesis is to apply system identification techniques from input-output data obtained from the PMUs for two cases: when a step excitation is possible and not when a step excitation is not possible. Step realization method and covariance based realization (CoBRA) methods are presented.

1.4.5 Coordinated Dispatch Optimization

Optimal Hybrid Power Dispatch (OHPD) algorithm that uses a convex multi-objective optimization to optimally dispatch power among multiple hybrid distributed energy resources (DERs) that are subjected to DER limitations in the form of power amplitude, power rate and energy storage constraints is presented in this thesiss. The solution is computed by minimizing an objective function motivated by economical incentives and subjected to constraints to satisfy a given power demand while taking into account the DER limitations. As a result, the OHPD algorithm uses a net power demand signal that obeys limitations on power ramp rate and power amplitude to formulate either a power following dispatch strategy or a cycle charging dispatch strategy, where excess power from DERs is used to charge batteries in a BESS.

1.4.6 Microgrid Power Flow and Angle Control

In this thesis, a detailed switching and control algorithm is proposed for a microgrid consisting of multiple distributed energy resources including both synchronous generators and inverter-based DERs. The controller enables fast and smooth intentional switching between islanding and grid connected modes with small voltage and frequency transient fluctuations. Modeling of the plant for control purposes is performed by model estimation techniques using DER inputs, PMU measurements, and realization algorithms. The controller is designed based on the estimated models. The data-based controller design methodology can be easily extended to accommodate other combinations of DERs in the microgrid. The developed control algorithm is tested on a HIL setup that models the microgrid and proves fast and smooth transition.

1.4.7 Short-term Prediction of PMU Data

Most of the applications of synchronized phasor measurements have been mainly concerning power system model validation, event analysis (detection, classification and localization), real-time display, etc. However, synchrophasors have a much greater potential than just monitoring and visualization such as reliable and economical operation of power systems in real-time control and protection schemes. However, the use of syncrophasors in control is prone to missing measurements. Hence, a method to predict the phasor data in real-time for short periods based on the historical measurements becomes important. The final part of the thesis describes a spectral based phasor data prediction and recovery algorithm that replaces invalid synchrophasor data and transitions smoothly to valid data using a Kalman filter based approach.

1.5 Outline of the Thesis

The thesis is organized as follows: Chapter 2 discusses the contributions of the thesis for grid monitoring which comprises of grid event detection and classification. Chapter 3 explores the contributions of the thesis for recursive power demand prediction and use for planning and dispatch purposes. Chapter 4 entails the need for synchrophasors for power flow dynamics estimation of a microgrid along with dispatch method and power flow and angle control. Chapter 5 discusses the contributions of the thesis for robustness to missing synchrophasor data for use in control purposes by short-term prediction of PMU data. The final chapter 6 concludes the thesis by addressing the formulated problem, gathering conclusions from all the sub-problems and set forth suggestions for further research and contributions.

Chapter 2

Grid Event Detection and Classification

2.1 Need for Grid Monitoring

The enormous volumes of synchronized time stamped data produced at 60Hz sampling by PMUs provides a clear challenge for data management and provides new opportunities for power systems control and protection [11–13]. Manual observations of PMU data time sequences to observe trends or possibly detect anomalies in the data quickly becomes a unwieldy task. It has been recognized that automated or semi-automated data analysis techniques to identify faults [5], out-of-step conditions [14], power generation anomalies [15], detect PMU data events on multiple PMUs [16], state estimation [17] and possibly extract (dynamic) knowledge from such events [18] are highly desirable. Such applications greatly automate the data management task associated to analyzing PMU data and improve automated grid monitoring capabilities [19].

Algorithms for calculation of phasors, local system frequency and rate of change of frequency (RoCoF) that follow the guidelines of the (recent) IEEE Standard C37.118 [20, 21] are abundant. Especially the accuracy of phasor measurement under dynamic conditions, that include transient effects due to load switching, has been improved by dynamic phasor estimates that use (weighted) least squares, discrete-time (moving average) filtering and/or advanced algorithms based on discrete Fourier transforms [4]. Local signal processing (edge processing) of PMU data that exploits the same computational and processing capabilities of PMUs is important to reduce the need to transmit high frequency PMU data to a central repository for analysis and event detection.

This chapter describes local synchrophasor based real-time data processing algorithm to automatically detect dynamic transient effects in AC voltage and current signals analyzed by a synchrophasor vector processing systems. The basic output of the algorithm is the time stamp when a dynamic event was detected. The approach is based on real-time discrete-time filtering of phasor data (amplitude and phase angle) to create optimal Filtered Rate of Change (FRoC) signals for each phasor and postulate an event detection algorithm based on the dynamic response of the FRoC signals. The idea of using phasors directly for event detection has been addressed in earlier work [6–8]. However, optimality of the FRoC signals is addressed in this chapter by (recursive) least squares estimation of the parameters of a linear discrete-time filter that minimizes the variance of the FRoC signals for both angle and amplitude data of the phasor for event detection.

It is shown that such optimal filtering of phasor data will lead to FRoC signals that have much better variance properties than the RoCoF signals based on the rate of change of the bus frequency [22] produced by the PMU. The smaller variance properties are achieved by the optimal filter that estimates the dynamics of the noise on the phasor data due to sensor and grid dynamics and can be used to detect the start time of dynamic transient effects in phasors more accurately. For the practical illustration of the algorithm, real-time PMU data acquisition is implemented on a Raspberry PI computer running Python packages under Linux, receiving C37.118 format data from the microPMU system [23], developed by Power Standards Lab as an extension to the well-established low voltage PQube instrumentation by Power Sensors Ltd. It is shown how real-time processing of the phasor data received by C37.118 can be used for local event detection on the basis of data obtained during several local events measured by the microPMU.

To be able to fully leverage event detection, automated or semi-automated analysis techniques [24–26] to automatically classify grid events into relevant load or generation perturbations such as faults, out-of-step conditions, power generation anomalies are highly desirable. Such techniques greatly automate grid monitoring by extracting relevant information from PMU data and classify grid disturbance events.

Most of the power blackouts are typically caused by a single event that leads to cascading outages [27]. Clearly, it is critical to develop a PMU synchrophasor data analysis scheme that not only detects but also identifies the cause of these events to prevent potential cascading grid failures. This chapter builds on a systematic data analysis procedure that uses the following 3 subsequent steps: event detection [2,28,29], parameter estimation [18,30] and event classification [25,26]. More specifically, the event detection involves detecting unusual occurrences of critical events in a power grid such as transient switching, outages, faults, line trips and intentional system test events like the Chief Joseph brake test. The parameter estimation involves estimating critical parameters of the disturbance event and extracting the features that relate these parameters to a specific type of event. Event classification involves grouping events according to its type, cause and extent of effect on the grid and attributing such events to specific components in the grid.

Data based event classification techniques such as neural networks, multiple linear regressions, generalized linear modeling, regression trees have been extensively studied [31]. Also, dimensionality reduction techniques such as principal component analysis (PCA) [32], [15] have become a great interest due to their faster computation features. However, the separate steps of event detection, parameter estimation and event classification have to be carried accurately and in a timely manner in order to prevent propagation of such outages or blackouts. Additional event localization [2] can be used to identifying the approximate location of the origin of the event.

For the illustration and application of this methodology, real-time PMU data acquisition has been implemented for IEEE C37.118 formatted data from a microPMU system developed by Power Standards Lab (PSL). Firstly, event detection is performed in a distributed sense on a Raspberry PI processing the C37.118 data stream of a single PMU. Once the detected events are available, it is shown how real-time processing of the phasor data containing detected events is used to extract features which then can be used as a training set to classify recorded events. The training and testing data sets consist of event instances created from measurements of the single phase phasors (frequency) from the wall outlet in the Synchrophasor Grid Monitoring and Automation (SyGMA) lab at UCSD over several months.

2.2 Signal Processing for Event Detection

2.2.1 Popular Detection Method: Rate of Change

The phasor data $(A(t), \phi(t))$ are estimated and exported by a PMU measuring device at regular time interval $t_k = k\Delta_T$, where $f_s = 1/\Delta_T$ denotes the sampling frequency. The typical sampling frequency is given by $f_s = 60$Hz for updates on the phasor at each cycle of an AC sigal $x(t)$ in an electricity grid with a nominal AC frequency of $f_0 = 60$Hz. The phasor data consisting of amplitude $A(t_k)$ and phase angle $\phi(t_k)$ are communicated, along with an accurate GPS measurement time t_k for data processing.

A popular signal for event detection, also often exported by the PMU, is the rate of change of frequency (RoCoF) signal

$$RoCoF(t_k) = 60 \cdot (f(t_k) - f(t_{k-1})) \tag{2.1}$$

that can be represented as the discrete-time filtered signal $RoCoF(t_k) = F(q)f(t_k)$ where $F(q)$ is a discrete-time derivative filter represented by the "transfer" function

$$F(q) = \frac{1 - q^{-1}}{60} = \frac{q - 1}{60q} \tag{2.2}$$

where q is use to denote the discrete-time shift operator $qx(t_k) = x(t_{k+1})$ and $q^{-1}x(t_k) = x(t_{k-1})$. Although the filter $F(q)$ and the resulting RoCoF signal is indeed a viable signal for detecting changes in the frequency $f(t_k)$, it is highly susceptible to noise on the actual phase angle estimate $\phi(t_k)$.

It is clear from (1.1) that differentiation or a discrete-time derivative of the phase angle $\phi(t_k)$ is required to obtain an estimate of frequency $f(t_k)$, whereas an additional discrete-time derivative is needed to obtain $RoCoF(t_k)$. A more careful design of the filter $F(q)$ is needed to obtain a signal suitable for event detection that is robust to noise on the phase angle estimate $\phi(t_k)$. Furthermore, time varying changes in the phase angle estimate $\phi(t_k)$ is only one part of the phasor. Time varying changes in the amplitude $A(t_k)$ of the phasor $X(t_k)$ in (5.3) must also be taken into account to properly detect events.

2.2.2 Optimal Filtering for Detection Signal

To set up the framework for the construction of optimal filters to process the discrete-time sampled phasor data $X(t_k)$ in (5.3), we consider an unobservable discrete-time event signal $d(t_k)$. The event signal $d(t_k)$ can only be observed via noisy observations of the phasor data $(A(t_k), \phi(t_k))$ characterized by

$$\begin{aligned} A(t_k) &= G_A(q)d(t_k) + n_A(t_k) \\ \phi(t_k) &= G_\phi(q)d(t_k) + n_\phi(t_k) \end{aligned} \tag{2.3}$$

where $G_A(q)$ and $G_\phi(q)$ denote unknown dynamic systems that filter the effect of the common event signal $d(t_k)$ on both the amplitude $A(t_k)$ and phase $\phi(t_k)$ of the phasor $X(t_k)$. The noise $n_A(t_k)$, $n_\phi(t_k)$ on the signals in (2.3) is characterized by filtered white noise signals

$$\begin{aligned} n_A(t_k) &= H_A(q)e_A(t_k) \\ n_\phi(t_k) &= H_\phi(q)e_\phi(t_k) \end{aligned} \tag{2.4}$$

where $H_A(q)$ and $H_\phi(q)$ denote unknown dynamic systems that model filter the white noises $e_A(t_k)$, $e_\phi(t_k)$ and model the spectral content of the output noise on the data in (2.3). As such, we may assume that $H_A(q)$ and $H_\phi(q)$ are spectral factorizations of the noise spectrum and are stable and stably invertible filters [33]. We assume that the white noises $e_A(t_k)$, $e_\phi(t_k)$ are uncorrelated with an unknown variance without loss of generality to formulate the estimation problem of the optimal filters to process the phasor data $(A(t_k), \phi(t_k))$.

To illustrate the idea of optimal filtering, we consider the amplitude signal $A(t_k)$ only, as the approach to filter $\phi(t_k)$ will be similar. Formulating a filtering $F_A(q)$ of the amplitude data $A(t_k)$ leads to the estimate $\hat{d}(k) = F(q)A(t_k) =$

$F(q)G_A(q)d(t_k) + F(q)H_A(q)e_A(t_k)$ where it can be observed that choosing $F_A(q) = G_A(q)^{-1}$ would lead to a perfectly constructed event signal $d(t_k)$, but susceptible to a filtered noise signal $F(q)H_A(q)e_A(t_k)$ that may be arbitrary bad. Furthermore, such a choice is only possible if the dynamics $G_A(q)$ is known and invertible. An example of this approach is the choice of the filter $F(q)$ for the RoCoF signal in (2.2), where an attempt is made to approximate the inverse of integration (differentiation) to process a step-wise change in the bus frequency due, resulting in high frequency noise amplification.

Instead, choosing $F_A(q) = H_A(q)^{-1}$ would lead to

$$\hat{d}_A(t_k) = F(q)G_A(q)d(t_k) + e_A(t_k) \tag{2.5}$$

and constitutes a filtered version of the event signal $d(t_k)$ perturbed by only a white noise signal. The properties of a white noise signal can now be used to formulate an event detection algorithm that exploits the correlation between subsequent measurements of $\hat{d}(t_k)$ over time. More details on the actual event detection will be given in the next section, first we focus on the construction of the filter $F_A(q) = H_A(q)^{-1}$.

The possibility to choose a filter $F_A(q) = H_A(q)^{-1}$ is motivated by the fact that $H_A(q)$ is a spectral realization of the spectrum of the noise $n_A(t_k)$ in (2.3). It is well known that such a filter $H_A(q)$ can always be realized by stable and stably invertible filters $F_A(q) = H_A(q)^{-1}$ [33], guaranteeing the existence of the stable filter $F_A(q)$. As $H_A(q)$ is unknown, we proposed two crucial steps to compute a filter $F_A(q)$ that is able to approximate the inverse of the noise dynamics $H_A(q)$:

1. Select N data points of phasor data of $A(t_k)$ where *no* event was present, e.g. $d(t_k) = 0$.

2. Select an (optional) filter $L_A(q)$ to filter the phasor amplitude data. The filter

$L_A(q)$ is used to emphasize certain frequency ranges where event detection is important. For example, a high pass filter $L_A(q)$ will avoid detection of offsets on the amplitude $A(t_k)$.

The selection of N data points where no event is present in step 1 can be by manual inspection of the data or based on the event detection algorithm summarized later. Clearly, this first step is required for the initialization and calibration of the event detection algorithm. The ability to include a user-specified filter $L_A(q)$ in step 2 above provides an extra design step in the event detection. In addition, a carefully designed filter $L_A(q)$ can make the approximation of the inverse of the noise dynamics more easier to achieve. With the help of the two steps above, (2.5) reduces to

$$\hat{d}_A(t_k) = F_A(q)L_A(q)A(t_k) = F_A(q)L_A(q)H_A(q)e_A(t_k) \tag{2.6}$$

where is can be observed that an approximation of the inverse of the (filtered) noise dynamics $L_A(q)H_A(q)$ by $F_A(q)$ would now lead to a white noise $\hat{d}_A(t_k)$. Parametrizing the filter $F_A(q)$ in an Moving Average format

$$F_A(q) = 1 - f_1q^{-1} - f_2q^{-1} - \cdots - f_nq^{-1} \tag{2.7}$$

allows $\hat{d}_A(t_k)$ in (2.6) to be written in a linear regression form

$$\begin{aligned}\hat{d}_A(t_k,\theta) &= A(t_k) - f_1A(t_{k-1}) - f_2A(t_{k-2}) - \cdots - f_nA(t_{k-n}) \\ &= A(t_k) - \theta\varphi_A(t_k)\end{aligned}$$

where $\theta = \begin{bmatrix} f_1 & f_2 & \cdots & f_n \end{bmatrix}$ and $\varphi_A(t_k) = \begin{bmatrix} A(t_{k-1}) & A(t_{k-2}) & \cdots & A(t_{k-n}) \end{bmatrix}^T$.

A maximum likelihood estimator $\hat{\theta}_A^N$ that minimized the variance

$$\hat{\theta}_A^N = \min_{\theta} \frac{1}{N} \sum_{k=1}^{N} \hat{d}_A(t_k, \theta)^2$$

over N time samples is now simply given by the Least Squares estimate

$$\hat{\theta}_A^N = \left[\frac{1}{N} \sum_{k=1}^{N} A(t_k) \varphi_A(t_k)^T \right] \left[\frac{1}{N} \sum_{k=1}^{N} \varphi_A(t_k) \varphi_A(t_k)^T \right]^{-1} \tag{2.8}$$

and reduces the optimally filtered detection signal $\hat{d}_A(t_k, \hat{\theta}_{LS}^N)$ to a white noise signal if indeed $F_A(q, \hat{\theta}_{LS}^N) \approx (L_A(q) H_A(q))^{-1}$. Increasing the order n of the filter $F_A(q, \theta)$ in (2.7) increases the design freedom of achieving this approximation, while the user chosen filter $L_A(q)$ can simplify the objective to achieve the approximation for a given value of the order n [30].

2.2.3 Proposed Detection Method: Filtered Rate of Change

The properties of a white noise signal can now be used to formulate an event detection algorithm that exploits the correlation between subsequent measurements of $\hat{d}(t_k)$ over time. Using $E\{\cdot\}$ to denote the expectation operator, the white noise signal $e_A(t_k)$ in (2.4) satisfies

$$E\{e_A(t_k) e_A(t_k - \tau)\} = \begin{cases} \lambda & \tau = 0 \\ 0 & \tau \neq 0 \end{cases}$$

indicating that subsequent values of $e_A(t_k)$ are uncorrelated.

The optimal filter $F_A(q, \hat{\theta}_A^N)$ found by minimizing the variance of the filtered event signal $\hat{d}_A(t_k, F_A(q, \hat{\theta}_A^N))$ in the case of no event ($d(t_k) = 0$) and an order n

large enough to satisfy $F_A(q, \hat{\theta}_A^N) \approx (L_A(q)H_A(q))^{-1}$ ensures that

$$\begin{aligned} \hat{d}_A(t_k, \hat{\theta}_A^N) &= F_A(q, \hat{\theta}_A^N)L_A(q)A(t_k) \\ &\approx e_A(t_k) \end{aligned}$$

is also a white noise signal in the case of no event. This allows us to formulate an event detection algorithm on the premise of assuming normal distributions for the $\hat{d}_A(t_k, F_A(q, \hat{\theta}_A^N))$ signal and exploiting the following information.

Consider N data points of phasor data of $A(t_k)$ where *no* event was present, e.g. $d(t_k) = 0$. These N data points are typically *not* the same as the N data points without event on which the parameter $\hat{\theta}_A^N$ was estimated to allow cross validation [30]. With the computed optimal filter $F_A(q, \hat{\theta}_A^N)$ we can now compute a variance estimate

$$\hat{\lambda}_A = \min_{\theta} \frac{1}{N} \sum_{k=1}^{N} \hat{d}_A(t_k, \hat{\theta}_A^N)^2, \quad \hat{d}_A(t_k, \hat{\theta}_A^N) = F_A(q, \hat{\theta}_A^N)A(t_k) \tag{2.9}$$

Assuming that the numerical values of $\hat{d}_A(t_k, \hat{\theta}_A^N)$ are generated by a normal distribution, the probability that

$$|\hat{d}_A(t_k, \hat{\theta}_A^N)| > 3\sqrt{\hat{\lambda}_A} \tag{2.10}$$

is less then 0.3% for a particular value of t_k. Although checking if (2.10) is satisfied for event detection, there is still a small probability of false event detection at each time stamp t_k that may lead to may false event alarms over a large number of data points that is generated by a PMU. Assuming, in addition, that $\hat{d}_A(t_k, \hat{\theta}_A^N)$ is a white noise signal with uncorrelated samples, the probability that

$$|\hat{d}_A(t_l, \hat{\theta}_A^N)| > 3\sqrt{\hat{\lambda}_A}, \quad l = k, k+1, \ldots, k+m-1 \tag{2.11}$$

for m *consecutive* time stamps $t_l,\ l = k, k+1, \ldots, k+m-1$ will be even smaller, typically $0.3^m\%$. The event criterion in (2.11) clearly will lead to much less false event alarms at the price of a small delay of m consecutive samples. The delay is often negligible, as choosing $m = 6$ would only lead to 0.1 sec delay at $f_s = 60$ Hz sampling, while reducing false alarm probability significantly.

If an event does occur on the amplitude measurement $A(t_k)$ of the phasor $X(t_k)$, the filtering leads to a signal

$$\begin{aligned} \hat{d}_A(t_k) &= F_A(q, \hat{\theta}_A^N)L_A(q)A(t_k) \\ &\approx F_A(q, \hat{\theta}_A^N)L_A(q)G_A(q)d(t_k) + e_A(t_k) \end{aligned}$$

with an optimal filter $F_A(q, \hat{\theta}_A^N)$ accurate enough to satisfy $F_A(q, \hat{\theta}_A^N) \approx (L_A(q)H_A(q))^{-1}$. Since $\hat{d}_A(t_k)$ is now the sum of a filtered event signal $F_A(q, \hat{\theta}_A^N)L_A(q)G_A(q)d(t_k)$ and a white noise $e_A(t_k)$, we may expect that not only (2.10), but also (2.11) will be satisfied. Depending on the dynamics of $F_A(q, \hat{\theta}_A^N)L_A(q)G_A(q)$ and the duration of the event signal $d(t_k)$, the absolute value of the signal $|\hat{d}_A(t_k)|$ may stay out of the bound $3\sqrt{\hat{\lambda}}$ for a larger number of consecutive samples. This allows one to increase the value of m in (2.11), while reducing the probability of false event detection.

It should be pointed out that the LS estimate in (2.8) can be updated (recursively) each time a set of N data points is available where no event was detected. For now, the LS estimate serves as a calibration of the optimal filter $F_A(q, \hat{\theta}_A^N)$ to reduce the amplitude $A(t_k)$ measurements of the phasor $X(t_k)$ to white noise $\hat{d}_A(t_k, \hat{\theta}_A^N)$ for event detection. The exact same procedure for choice of a data filter $L_\phi(q)$ and computation of optimal filter $F_\phi(q, \hat{\theta}_\phi^N)$ with event detection can be applied to the angle $\phi(t_k)$ measurements of the phasor $X(t_k)$ in parallel.

2.3 Power Oscillations in Grid Dynamics

Modern power grids are mostly interconnected and operate close to the stability limits in transient and steady state modes [27]. Natural response of one group of closely coupled machines oscillate against one another in large interconnected grids. Higher frequency modes are localized with small groups that oscillate against each other. A small disturbance in any part of the system can initiate an inter-area oscillation, in a heavily interconnected grid. These oscillations, especially if lightly damped, can lead to a line trips due to protective relay trips and inducing a grid system collapse [34].

In a power system with multiple machines, measurements at nearby buses show coherent oscillation characteristics. As a result, changes in real power can be seen on frequency measurements while changes in reactive power can be seen in voltage measurements [27]. To illustrate the transient phenomenon in a largely interconnected power grid, consider a simple two-area power system model with two machines M1, M2 (each modeling generation and load) connected via a single bus.

Assuming that the machines M1 and M2 operating at voltages V_1 and V_2, start at initial angles δ_1 and δ_2 respectively with respect to reference nominal frequency and that M1 accelerates while M2 de-accelerates from their nominal rotation frequency. Excluding random variation of load caused by demands, let us consider only the frequency dependent load variation linearized with a coefficient β, allowing us to write

$$\frac{2H_1}{\omega_s}\ddot{\delta}_1 = P_{m1} - P_{e1} + \beta_1\dot{\delta}_1 \quad \text{and} \quad \frac{2H_2}{\omega_s}\ddot{\delta}_2 = P_{m2} - P_{e2} + \beta_2\dot{\delta}_2 \tag{2.12}$$

where P_{mi} indicates the mechanical power output in area i, P_{ei} indicates real power flow in/out of area i, H_i indicates the area inertia constant, ω_s indicates the synchronous frequency and subscripts $i = 1, 2$ refer to areas of M1 and M2 respectively. Assuming that the transmission line is lossless, with line impedance X, then power flows from M1 to M2 can be represented by

$$P_{e1} = -P_{e2} = \frac{V_1 V_2}{X} \sin(\delta_{12}) \tag{2.13}$$

where $\delta_{12} = \delta_1 - \delta_2$. In steady state operation for the lossless transmission, the generated power at M1 is absorbed through M2 and therefore $P_{m1} = -P_{m2}$.

Combining the equations (2.12) and (5.3), one obtains

$$\begin{aligned} \tfrac{2H_1}{\omega_s}\ddot{\delta_1} - \tfrac{2H_2}{\omega_s}\ddot{\delta_2} &= 2P_{m1} + P_{e1} - P_{e2} + \beta_1\dot{\delta_1} - \beta_2\dot{\delta_2} \\ \tfrac{2H_1}{\omega_s}\ddot{\delta_1} - \tfrac{2H_2}{\omega_s}\ddot{\delta_2} &= 2P_{m1} - \tfrac{V_1V_2}{X}\sin(\delta_{12}) + \beta_1\dot{\delta_1} - \beta_2\dot{\delta_2} \end{aligned} \tag{2.14}$$

from which it can be seen that a two-area system with $H_1 = H_2 = H$, $\beta_1 = \beta_2 = \beta$, the normalized voltage units $V_1 = V_2 = 1$ and $P_{m1} = P_{m2} = 0$ follows

$$\ddot{\delta_{12}}(t) = -\omega^2\delta_{12}(t) - \beta_k\dot{\delta_{12}}(t) \tag{2.15}$$

under the assumption $\sin(\delta_{12}) \approx \delta_{12}$, valid for small values of the time t dependent $\delta_{12}(t)$. From (2.15) one recognizes a standard 2nd order differential equation that will have a sinusoidal solution with an (undamped) oscillation frequency $\omega = \sqrt{\frac{\omega_s}{HX}}$ rad/s and damping ratio $\beta_k = \frac{\omega_s\beta}{2H}$.

The purposes of the above-mentioned analysis is to show that the power flow (5.3) of a two-area power system behaves much like a damped mass-spring system with natural (undamped) oscillation frequency $\omega > 0$ and damping ratio $0 \leq \beta \leq 1$. Since an electromechanical oscillation is an exchange of kinetic energy through electrical power between various machines, the oscillation takes a specific path through the grid [34]. In a large grid, e.g. the Western Electric Coordinating Council (WECC), there may be several such areas and the transmission system may have several resonance modes with coupling and possibly non-linear effects. The unavoidable phenomena of power oscillations due to rotating machine inertia will be exploited in the feature extraction whenever power disturbances occur on the grid.

2.4 Parameter Estimation for Feature Extraction

2.4.1 Free Response Dynamic Model

In a large interconnected grid (e.g. WECC) that can consist of several load-generation areas, it becomes very difficult to derive the swing equations due to highly coupled non-linear dynamic behavior between areas. As an alternative, we deploy system identification techniques to extract information from the grid each time an event occurs [30]. The realization algorithm [35] is used here to identify modal parameters of the power oscillation dynamics, assuming the free response of a linear dynamic behavior excited by a (im)pulse signal.

For the realization algorithm, consider a linear dynamic model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{2.16}$$

is used to capture the free response of the power oscillation or grid frequency oscillations $y(k)$ at discrete times, $t = k\Delta T, k = 0, 1, \ldots$ with a constant sampling time ΔT. In this chapter, we choose to use $y(k)$ to be the supply frequency $f(k)$ of the AC voltage signal as provided by a PMU. The motivation to use the measured AC supply frequency $f(k)$ is due to the strong correlation between $f(k)$ and real power flow $P(k)$ in a grid with rotating machinery as discussed in the the previous section. In addition, it is much easier to measure fluctuation in the AC supply frequency $f(k)$ via a PMU, without having to measure both AC voltage and AC current for power flow calculation. That said, feature extraction can also be applied to power measurements without loss of generality.

In (2.16), the vector $x(k) \in \mathbb{R}^{n\times 1}$ denotes a n-dimensional state vector, $A \in \mathbb{R}^{n\times n}$ is the state matrix, $B \in \mathbb{R}^{n\times 1}$ is the input matrix, $C \in \mathbb{R}^{1\times n}$ is the output matrix and $D \in \mathbb{R}^{1\times 1}$ is the feedthrough scalar/matrix with C and D having (only) a single row, the analysis presented here assumes a single (output) measurement. However, the approach can be easily extended to multiply synchronized PMU measurements where C and D have may multiple rows. With the model (2.16), the (impulse) input $u(k)$ and power or frequency measurement $y(k)$ are related via the Markov parameters $g(k) = CA^{k-1}B$ for $k > 0$ and allow the input/output dynamic relation to be written as a discrete-time convolution

$$y(k) = Du(k) + \sum_{j=1}^{\infty} g(j)u(k-j)$$

where $y(k) = g(k)$ due to the (im)pulse input $u(k)$. Our goal is to estimate the order n of the model and then compute (realize) the matrices A, B and C of the model based on discrete-time power or frequency swing observations $y(k)$. These matrices are then used to extract the relevant parameters to quantify the event.

2.4.2 Realization Algorithm for Model Estimation

For the realization algorithm, one uses $2N$ discrete-time measurements $y(k),\ k = 1, 2, \ldots, 2N$ to construct a Hankel matrix H and the corresponding shifted Hankel matrix $\bar{H}$ of the same size given by

$$H = \begin{bmatrix} y(1) & y(2) & \cdots & y(N) \\ y(2) & y(3) & \cdots & y(N+1) \\ \vdots & \vdots & \ddots & \vdots \\ y(N) & y(N+1) & \cdots & y(2N-1) \end{bmatrix}$$

$$\bar{H} = \begin{bmatrix} y(2) & y(3) & \cdots & y(N+1) \\ y(3) & y(4) & \cdots & y(N+2) \\ \vdots & \vdots & \ddots & \vdots \\ y(N+1) & y(N+2) & \cdots & y(2N) \end{bmatrix}$$

where it can be shown that $\bar{H} = H_1 A H_2$ where $H_1 \in \mathbb{R}^{N \times n}$, $H_2 \in \mathbb{R}^{n \times N}$ are a rank n decomposition $H = H_1 H_2$ of the Hankel matrix H. Such a rank n decomposition can be computed via a Single Value Decomposition (SVD) $H = U \Sigma V^T$ where $U, V \in \mathbb{R}^{N \times N}$ are orthonormal matrices and $\Sigma \in \mathbb{R}^{N \times N}$ is a diagonal matrix with N singular values ordered in a non-decreasing magnitude on the main diagonal.

Choosing a fixed model order n, the SVD allows a rank n approximation $H \approx U_n \Sigma_n V_n^T$ of the full rank N matrix H, where $U_n \in \mathbb{R}^{N \times n}$ indicates the first n columns in U, $\Sigma_n \in \mathbb{R}^{n \times n}$ the first n rows and n columns of Σ and $V_n \in \mathbb{R}^{N \times n}$ indicates the first n columns in V. The rank n approximation rewrites $H \approx H_1 H_2$ where $H_1 = U_n \Sigma_n^{1/2}, H_2 = \Sigma_n^{1/2} V_n^T$ and the state space matrices A, B, C can be computed using $A = \Sigma_n^{-1/2} U_n^T \bar{H} V_n \Sigma_n^{-1/2}$, $B = H_2(:,1)$ and $C = H_1(1,:)$, where $H_2(:,1) \in \mathbb{R}^{n \times 1}$ denotes the first column in H_1 and $H_1(1,:) \in \mathbb{R}^{1 \times n}$ denotes the first row of H_1.

2.4.3 Model Based Feature Extraction

For feature extraction we consider transient dynamic behavior stability of the model, determined by the eigenvalues $\lambda_i,\ i = 1, 2, \ldots, n$ of the estimated state matrix A. In addition to the eigenvalues of A, also the nullspace solutions to $[A - \lambda_i I]\phi_i = 0$ and $\psi_i[A - \lambda_i I] = 0$ are computed to find the left ψ_i and right ϕ_i eigenvectors. The information on the (complex) eigenvalues λ_i, left ψ_i and right ϕ_i eigenvectors are now used to compute the main dynamic features of a power of frequency oscillation event $y(k),\ k = 1, 2, \ldots$: the oscillation frequencies f_i, damping ratios ζ_i and relative participation factors P_i of the event, respectively given by

$$f_i = \frac{|s_i|}{2\pi}, \zeta_i = \frac{-a_i}{2\pi f_i}, P_i = |\phi_i \psi_i| \tag{2.17}$$

where $s_i = 60 \ln(\lambda_i) = a_i \pm j b_i$ for 60Hz sampling of the data $y(k)$.

The frequency f_i, damping ζ_i and participation factor P_i, collectively called the (dynamic event) parameters in (2.17) are used to characterize the dynamic aspects of an observed event and motivated as follows. Firstly, oscillating areas may be characterized by different modes or oscillation frequencies f_i. Secondly, the damping factor ζ_i plays an important role in determining the nature of the oscillation mode and will help in deciding whether the disturbance event occurred in the main or local grid area. For example, it was found that undamped oscillations at about 0.33 Hz was the major restraint on a larger transfer [34] in the WECC system caused by generator high gain automatic voltage regulators (AVR). Next, the participation factors P_i can be used to characterize the relative contribution of each oscillation mode and decide the source of the oscillation event. Finally, next to the listed (dynamic event) parameters in (2.17) that capture the dynamic behavior of a power oscillation event, also the static behavior is used in event characterization. The static parameter is characterized by the post-event deviation of the measured frequency amplitude ΔF before and after the event and computed via

$$\Delta F = \frac{1}{N_1} \sum_{k=N_1-k_0}^{k_0-1} y(k) - \frac{1}{N_2} \sum_{k=k_0+N}^{k_0+N+N_2-1} y(k)$$

where N_1 is the number of data points before the event at time index $k = k_0$ and N_2 is the number of data points after the event has settled at time index $k = k_0 + N$.

2.5 Feature Based Clustering for Grid Event Classification

Computational techniques such as distribution based cluster analysis, a type of unsupervised machine learning can be used to identify different patterns. In this thesis, the k-nearest neighbor algorithm is used to find patterns and classify the event. The k-means clustering method is a method used in a p-dimensional space, with an L_1-norm distance measure used for minimization of the distance from the point to centroid. The L_1 is the sum of absolute differences, also known as the 'city block distance'.

The goal of the algorithm is to partition the given data $\{x_1, x_2, ..., x_n\} \in \mathbb{R}^d$ into k disjoint clusters $\mathcal{C} = \{C_1, ..., C_k\}$ such that the sum of the cluster elements from the cluster centroid is minimized. Each centroid is the component-wise median of the points in that particular cluster. Formally, the problem can be formulated. Given a set of k cluster centers $C = \{c_1, ..., c_k\}$, where $x, c \in \mathbb{R}^d$, the k-means objective function for L_1 distance minimization is given by

$$J(\mathcal{C}, C) = \sum_{j=1}^{k} \sum_{x_i \in C_j} |x_i - c_j| \tag{2.18}$$

and the cost function $J(\mathcal{C}, C)$ (2.18) is non-convex. As a result, the problem of finding the global minimum of the k-means objective needs to be solved via a non-linear minimization, initialized at an initial centroid value c_j.

Once clustering has been achieved, it is worthwhile to give the computed clusters C a meaningful interpretation. The interpretation can be based on the fact that power system oscillations are classified by the system components that they effect. Traditionally, oscillation modes are classified based on oscillating frequencies alone [36], although damping is included in most calculation. The classical frequency distinction results in so-called interarea mode oscillations ($f <\leq 1$ Hz and typically around 0.3 Hz), local plant mode oscillations ($1.0 < f \leq 2.0$ Hz), intraplant mode oscillations ($2.0 < f <\leq 3.0$ Hz), torsional mode oscillations ($10 < f < 16$ Hz) and control mode oscillations. But the participation P_i of each such oscillating frequencies is often overlooked. In some cases, there might be an event caused by more than one of such modes of oscillation. Hence consideration of participation factor of each of such participating mode would provide better insights for classification.

2.6 Event Detection and Classification on Phasor Data from a μPMU

2.6.1 Data collection from μPMU system

The Power Standards Lab (PSL) μPMU (μPMU) includes a PQube instrument that contains measurement, recording, and communication functionalities along with a remotely-mounted micro GPS receiver, and a power supply. These devices can be connected to single- or three phase secondary distribution circuits up to 690V (line-to-line) or 400V (line-to-neutral), either into standard outlets or through potential transformers (PTs). The devices continuously sample AC voltage and current waveforms at 256 or 512 samples per cycle and can produce 3 phase phasor data $X(t_k)$ at sampling rates of 60 or 120 Hz.

The phasor data data is streamed in real-time to a client computer using IEEE C37.118 standard over the Ethernet using TCP port 4713 and the client computer used for data acquisition here is a Raspberry PI model B+ that acts as an interface between the data source (μPMU) and the data archive server (OSIsoft server). An overview of the hardware setup used for real-time C37.118 data acquisition from the μPMU into the Raspberry PI is shown in Fig. 2.1.

Due to the flexible computing environment on the Raspberry PI, real-time event detection can directly be implemented via Python. Socket programming is used to read data from the TCP port 4713 in a predefined frame-size. The μPMU-Raspberry PI act as a server-client application that communicate C37.118 data using sockets.

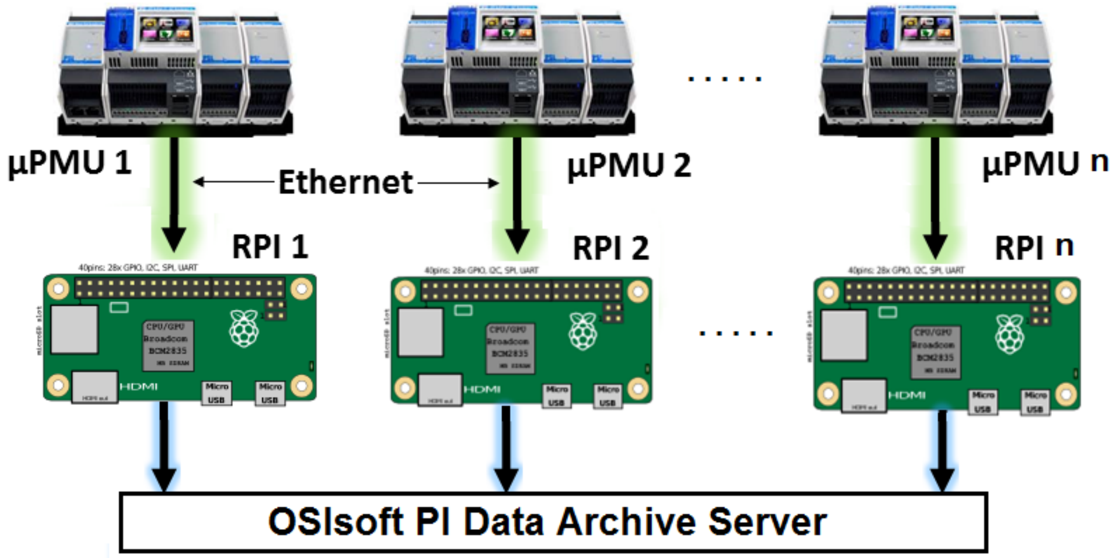

Figure 2.1: Hardware setup for data processing and archival.

2.6.2 Decoding the C37.118 Data in Real-time

A popular way to process C37.118 data is the use of OpenPDC[1] by Grid Protection Alliance (GPA), but here we decode data directly in the Python application used to read data from the TCP port 4713. The Synchrophasor measurements are tagged with the UTC time corresponding to the time of measurement usually consisting of three numbers: a second-of-century (SOC) count, a fraction-of-second (FRACSEC) count, and a message time quality flag as in [20]. The synchrophasor consists of four message types: data, configuration, header, and command. The first three message types are transmitted from the μPMU that serves as the data source, and the last (command) is received by the μPMU. All message frames start with a 2-byte SYNC word then followed by FRAMESIZE word (2-byte), IDCODE (2-byte), a time stamp consisting of a second-of-century (SOC, 4-byte)[2] and FRACSEC (4-byte), which includes a FRACSEC integer (24-bit) and a Time Quality flag (8-bit).

[1] http://openpdc.codeplex.com/

[2] The SOC count is a four (4) byte binary count of seconds from UTC midnight (00:00:00) of January 1, 1970, to the current second representing a 32-bit unsigned integer.

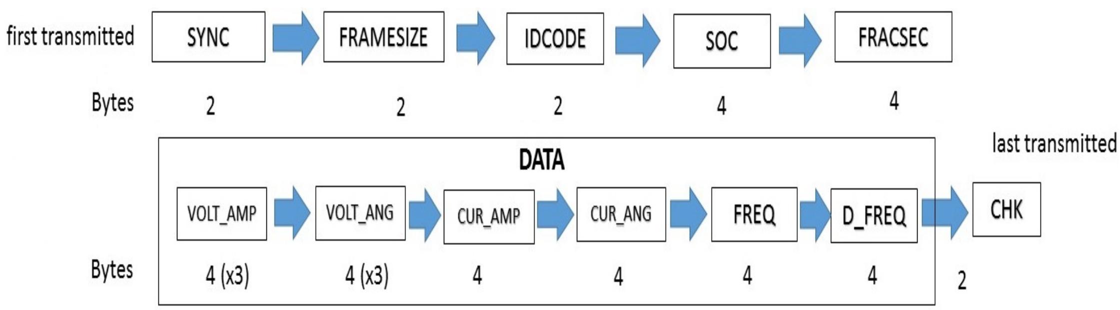

Figure 2.2: Data frame byte transmission of C37.118 data

The SYNC word provides synchronization and frame identification. The IDCODE positively identifies the source of a data, header, or configuration message, or the destination of a command message. All data frames terminate in check word (CHK) which is a CRC-CCITT. This CRC-CCITT uses the generating polynomial $X^{16} + X^{12} + X^5 + 1$ with an initial value of 1 (hex FFFF) and no final mask. All frames are transmitted exactly as described with no delimiters and an illustration of an example frame transmission order is shown in Fig.2.2.

2.6.3 Application of Event Detection

For demonstration purposes, single phasor voltage amplitude $A(t_k)$ and angle $\phi(t_k)$ data are collected from the Engineering Building at UCSD via a μPMU. The data used for demonstration of event detection was collected over the course of 48 hours at 60Hz sampling (approx. 10^6 data points) while a tornado warning was issued in San Diego in the afternoon of January 6, 2016.

For the optimal filter estimation, the data filter $L_A(q)$ on the voltage amplitude data $A(t_k)$ was chosen as a first order high pass Butterworth filter with a cut-off frequency of 0.1Hz to avoid detection of low frequency (off-set) disturbance events. The voltage phase data $\phi(t_k)$ limited between $[-180, 180]$ deg was first properly un-

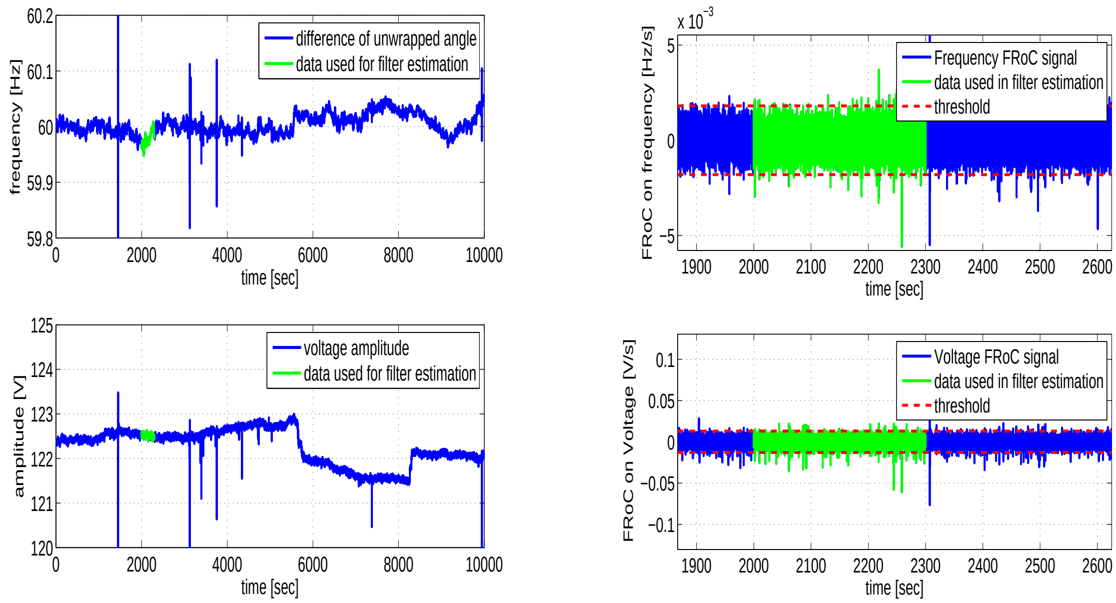

Figure 2.3: Left: part of the full phasor data (frequency and amplitude) with indication of data use for filter estimation. Right: zoomed-in version of the FRoC signal for both frequency and amplitude data in the neighborhood of the data used for filter estimation.

wrapped to $\phi_u(t_k)$ and then reduced to

$$f_e(t_k) = 2\frac{f_s}{160}(\phi_u(t_k) - \phi_u(t_{k-1}) + 60$$

to obtain a frequency estimate on which the same data filter $L_\phi(q) = L_a(q)$ was applied for optimal filter estimation. Only a small part of the data of just 5 minute length ($N = 18000$) is used for the estimation of the optimal filter paremeters $\hat{\theta}_A^N$ and $\hat{\theta}_\phi^N$ of an $n = 20$th order MA filter, while a neighboring set of points of also 5 minute length is used to estimate the variance estimates in (2.9) that will serve as event detection bounds. An zoomed-in version of the available data set along with the estimate of the event detection bounds is given in Fig. 2.3.

Based on the estimated optimal filters that generated the FRoC signal and the variance bounds, event detection is initiated if $m = 20$ consecutive samples of

$\hat{d}_A(q, \hat{\theta}_A^N)$ or $\hat{d}_\phi(q, \hat{\theta}_\phi^N)$ are outside the bounds $3\sqrt{\hat{\lambda}_A}$, $3\sqrt{\hat{\lambda}_\phi}$. Based on the requirement of $m = 20$ consecutive samples, 3 significant events are detected in the phasor data and an overview of the events are depicted in Fig. 2.4.

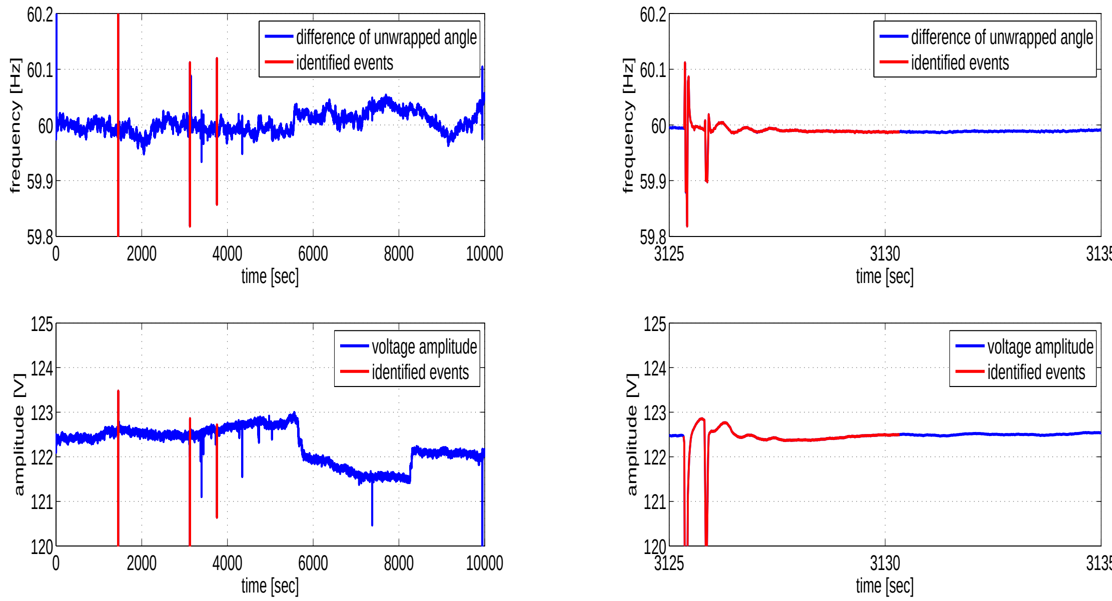

Figure 2.4: Left: part of the full phasor data (frequency and amplitude) with indication of events detected on the data (colored red). Right zoomed-in version of the 2nd event detected in the data showing oscillations on the frequency and several voltage dips.

Although there are other "spikes" seen in the data, these do not signify the detect events as they do not satisfy the event detection criteria. These events signify the disturbances on the grid that can explicitly be seen on the phasors of the power signal.

2.6.4 Frequency and Participation Factor Clustering

To illustrate the clustering of oscillation modes based on frequency domain data, PMU data acquisition has been implemented for IEEE C37.118 data from the microPMU system developed by Power Standards Lab (PSL). Event detection is performed as described in [29] and detect events based on the single phase AC supply frequency measurements from the wall outlet in the lab over several months. Using the set of measured events, first the k-means clustering algorithm is applied to a 2-dimensional space with natural frequency f_i of oscillation and the participation factor P_i of that oscillation frequency f_i. The resulting number of clusters are chosen to be 4 and the results of the feature (frequency f_i and partition factor P_i) extraction is depicted in Fig. 2.5. From the automated clustering results depicted in in Fig. 2.5, the following interesting conclusions may be drawn. Cluster 1 may be defined as the

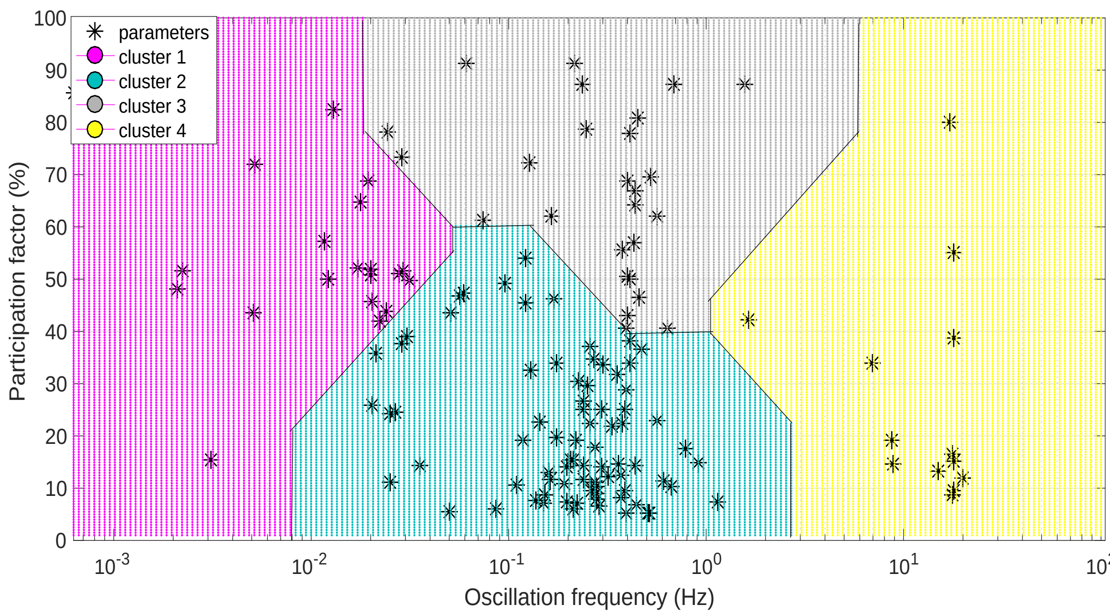

Figure 2.5: Demonstration of two-dimensional k-means clustering of oscillation modes based on oscillation frequency f_i and participation factor P_i.

cluster containing the very-low frequency interarea (VLFI) oscillations of the grid. These oscillations occur in almost every detected event usually at a higher participation factors $P_i > 40\%$. These oscillations are a result of large coherent generation areas in the WECC system swinging against each other and hence can be seen for almost every excitation in the WECC system. These oscillations do not provide any useful information in event classification as they inadvertently occur in almost every event. Cluster 2 may be defined as the cluster containing the low participation frequency interarea (LPFI) oscillations of the grid. These consist of the typical interarea oscillations in the WECC system that exhibit frequency oscillations of $f_i \sim 0.3$ and have a lower participation P_i in the events. Moreover, this is usually associated with a frequency component from Cluster 1. Cluster 3 may be defined as the cluster containing the high participation frequency interarea (HPFI) oscillations of the grid. These consist of the typical interarea oscillations in the WECC system that exhibit frequency oscillations of $f_i \sim 0.3$ Hz and have a significant participation P_i in any grid event. Finally, Cluster 4 may be defined as the cluster containing the frequency torsional mode (FT) oscillations of the grid. Usually these modes are excited when a multi-stage turbine generator is connected to the grid system through a series compensated line.

2.6.5 Frequency, Participation Factor and Damping Clustering

Although the above description may provide an intuitive clustering of events, the clustering approach in this chapter can be extended to allow further classification based on damping ratio. This method of classification helps differentiate between oscillations initiated in the local or the main grid in the WECC system due to variation of damping levels in each of these cases. Hence, three-dimensional k-means clustering involving natural frequency of oscillation f_i, participation factor P_i and damping factor ζ_i is performed. We keep the same number of 4 clusters as in Fig 2.5, but include damping ratio as the third dimension for classification. The clustering results are shown in Fig. 2.6. The 3-dimensional clustering provides an intuition about the events that occur locally (close to the measurement point) or non-locally (far-off from the measurement point). Cluster 1 contains very low frequency components with very high damping which does not contain any signature of an event. Hence compon-

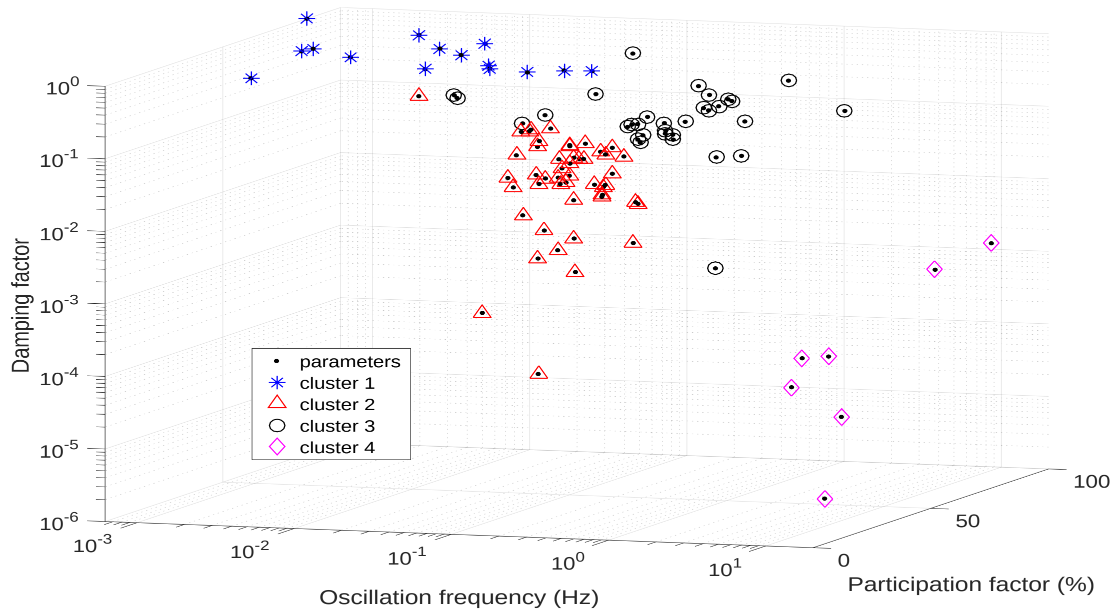

Figure 2.6: Demonstration of three-dimensional k-means clustering of oscillation modes based on oscillation frequency f_i, participation factor P_i and damping factor ζ_i for classification of oscillation modes at different damping levels.

ents falling in cluster 1 can be ignored for classification of events into local/non-local events. Cluster 2 may be defined as the cluster containing frequency components around 0.3 Hz WECC frequency with low participation, low damping factor signifying that these may be the components present in the local grid events. Cluster 3 may be defined as the cluster containing frequency components around 0.3 Hz WECC frequency with high participation, high damping factor signifying that these may be the components present in the non-local events such as the Chief Joseph break test. Cluster 4 depicts high frequency torsional oscillations that occur at very low levels of damping signifying local events.

2.6.6 Steady State Frequency Deviation Clustering

Finally, single dimensional classification of grid events resulting in steady state frequency deviations can also be performed. Application of the k-means clustering on the difference of the frequency before and after the event ΔF in (2.4.3) will provide additional information on a particular grid event in terms of steady state power loss or power surplus. These events may be caused due to unexpected generation trip or sudden load loss in an area. The resulting clusters based on ΔF are shown in Fig. 2.7. The events that fall in Cluster 1 do not show any significant

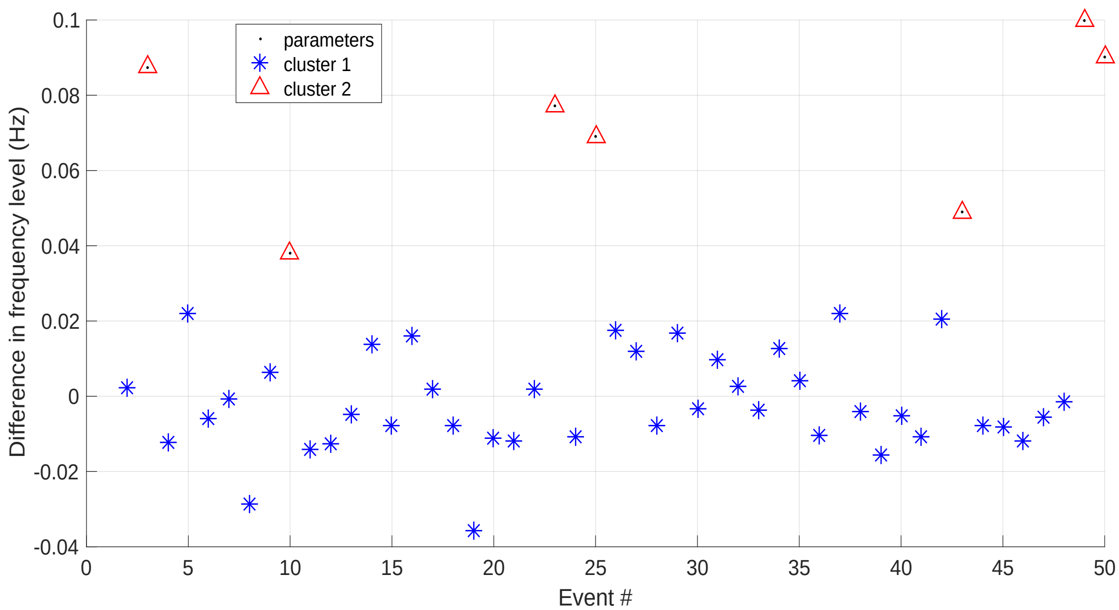

Figure 2.7: Demonstration of single dimensional k-means clustering of events based on difference of frequency amplitude ΔF before and after the event as in (2.4.3) for classification of events.

steady state frequency level ΔF deviation and hence do not fall under generation loss events. Whereas the events in Cluster 2 show a steady state frequency deviation $\Delta F > 0.04$ Hz and can be interpreted as generation surplus (or load loss) during the event.

2.6.7 Validation on PMU Data Events

The deduced rules for event classification of grid event can be applied on a set of events not used for (the training of) the clustering algorithm. This validation will demonstrate the robustness of the classification algorithm. For this purpose frequency PMU data from three measured known events are used: (1) the Chief Joseph Brake test on April 14^{th} 2015 (WECC event), (2) the disturbance caused during a major Tornado warning issued in San Diego on January 6^{th} 2016 (local event) and (3) the loss of major resources in Montana (generation loss) on January 21^{st} 2016 measured from Coronado substation. Table 1 shows the classification of these three events based on the clustering computed on the training set of events. The cluster components of the first event fall in cluster 1 (VLFI) and cluster 3 (HPFI) of the first classifier, cluster 3 (non-local) of the second classifier and cluster 1 of the third classifier. Hence, this event can be classified as a HPFI, non-local event without any power loss. Similarly, the cluster components of the second event fall in cluster 1 (VLFI) and cluster 2 (LPFI) of the first classifier, cluster 2 (local) of the second classifier and cluster 1 of the third classifier. Hence, this event can be classified as a LPFI, local event without any power loss. Finally, the cluster components of the third event fall in cluster 1 (VLFI) and cluster 3 (HPFI) of the first classifier, cluster 3 (non-local) of the second classifier and cluster 2 of the third classifier. This event can be classified as a HPFI, non-local event with generation loss, affirming the credibility of the algorithm.

Table 2.1: Assessment of classification algorithm based on test events

Test Events (Cluster)	**Classifier 1**				**Classifier 2**				**Classifier 3**	
	$\mathcal{C}$**1**	$\mathcal{C}$**2**	$\mathcal{C}$**3**	$\mathcal{C}$**4**	$\mathcal{C}$**1**	$\mathcal{C}$**2**	$\mathcal{C}$**3**	$\mathcal{C}$**4**	$\mathcal{C}$**1**	$\mathcal{C}$**2**
Event 1 (April 14^{th}, 2015)	X	-	X	-	X	-	X	-	X	-
Event 2 (January 6^{th}, 2016)	X	X	-	-	X	X	-	-	X	-
Event 3 (January 21^{st}, 2016)	X	-	X	-	X	-	X	-	-	X

2.7 Summary of Contribution

This chapter shows that optimal filters which can be used to formulate filtered rate of change phasor signals can be estimated on the basis of phasor data where no disturbances are present via a straight-forward least squares optimization. The resulting variance bounds give rise to event detection algorithms that check for the number of consecutive samples for which the filtered phasor signals are outside the variance bounds.The resulting procedure is applied to actual phasor measurement data obtained from a μPMU system developed by Power Standards Lab and shows realistic event detection for various frequency (angle) and voltage disturbance events.

This chapter also shows that combining estimation of dynamic event parameters combined with standard clustering methods can be used to build classifiers to quantify power grid oscillation events on synchrophasor data. The event parameters included frequencies, damping and participation factors of each oscillation mode in an event and can be estimated by realization algorithm. Application of the feature based grid event classification to actual PMU synchrophasor data obtained from a μPMU shows realistic clustering based classification of various detected disturbance events.

The text and data in Chapter 2, in full, is a reprint of the material as it appears in "Optimal filtering for grid event detection from real-time synchrophasor data", Sai Akhil R. Konakalla and Raymond A. de Callafon, *Procedia Computer Science*, 80 (2016): 931-940 and "Feature Based Grid Event Classification from Synchrophasor Data", Sai Akhil R. Konakalla and Raymond A. de Callafon, *Procedia Computer Science*, 108 (2017): 1582-1591. The dissertation author is the primary investigator and author of these articles.

Chapter 3

Phasor Based Recursive Demand Prediction for Economic Planning

3.1 Introduction

Steady growth in the global population and improvement of building services have resulted in a steady increase in electricity consumption. In the United States more than 44% of residential and commercial energy consumption corresponds to heating, ventilating and air conditioning (HVAC) systems in buildings [37], [38]. Also, recent rise in deployment of environmental friendly systems such as electric vehicles (EVs) has added to the increased energy consumption. Hence, planning and operational management of energy consumption and supply is a growing topic of interest and increasingly gaining higher traction due to fast, precise power measurement devices such as phasor measurement units (PMUs) [39] and also higher penetration of energy storage systems (ESS) such as batteries.

Short-term prediction of energy consumption [40], [41] is essential for demand side management, energy resource allocation, renewables energy integration and achieving financial benefits not only from a utility but also a micro-grid point of view. Due to the dynamic behavior of power consumption and volatile load changes, a static approach based on clustering does not always provide accurate power demand prediction results [42], [43], [44]. Dynamic power demand changes are more prevalent in commercial buildings where unanticipated demand is often encountered due to a sudden increase or decrease of occupants, unexpected changes in weather or other miscellaneous activities [45]. For improvement of dynamic power demand prediction performance, the static clustering of power data is combined with a Kalman filter [46] for (online) recursive updates of power demand predictions. This combination provides more accurate power demand prediction updates, as the state estimation of the Kalman filter can provide updates in case of sudden changes in load demand to minimize the error between the offline predictions and measured values [47]. In the later part of this chapter, the effectiveness of the algorithm is tested by predicting the load profile of two separate situations for a commercial load for both low and high power demand profiles.

One of the important applications of power prediction is for planning and optimal utilization of distributed energy resources (DERs) that are generally a combination of renewable and non-renewable resources along with storage such as battery energy storage system (BESS) [48]. The DERs may be used for achieving economical and non-economical objectives such as stability and reliability in a power grid. It is therefore essential to design dispatch optimization schemes [49], [50] to satisfy microgrid stability and reliability requirements, achieve economic benefits amidst non-linear and dynamics DER characteristics and limitations. A dynamic dispatch optimization that computes an optimal solution for the power dispatch of DERs at a pre-defined dispatch interval to achieve one such application i.e. minimizing electricity costs of a system load, is presented as part of the main contributions of this chapter.

Time-of-use (TOU) energy cost management and demand charge management (DCM) are the most general parts of the total electricity costs [51]. Utilities price electricity and charge end consumers based on total energy consumption ($/kWh) and demand consumption ($/kW), in which consumers pay for their peak demand on the grid. Utilities use a TOU pricing structure, wherein higher electricity prices coincide with periods of higher demand. Using optimal strategies, TOU management can reduce energy charges via energy time-shifting and price arbitrage, while DCM can reduce demand charges via peak load shifting. In TOU management, BESS is charged with electricity purchased from the grid when prices are low and discharged to offset energy use when prices are high to result in reduction of energy charges. Similarly, in DCM applications, BESS is charged when demand is low - ideally when energy prices are also low - and discharged to mitigate the peak load when demand is high. reduction in demand charges. Because price arbitrage via TOU management does not generally yield a return-on-investment for an energy storage system, DCM is often sought as an additional value stream (sometimes of much higher value) to justify investment in energy storage [52].

The proposed approach in this chapter is a modification of the existing dispatch strategies to run in an operational setting using synchrophasor measurements with possibly 60 times per second resolution. The approach incorporates a dynamic demand target threshold to which the load can optimally be peak shaved (reduced) with dynamically changing operational constraints and recursively updated load forecasts. The approach also focuses on optimal battery charge/discharge behavior to reduce non-coincident peak demand with TOU arbitrage. To maximize its peak reduction capability while preventing battery degradation or failure, specific modifications (i.e., factors of safety) are added to the dispatch scheme based on the observed net load and frequent changes in dispatch. Modifications for separate applications could be accomplished by including extra objectives or constraints in the algorithm with iteration over measured data during the previous days or weeks to determine optimal factors of safety.

3.2 Existing Techniques for Static Day Ahead Prediction

Aside from white-box prediction approaches, which solely rely on detailed information models using dedicated simulation engines, there have been a number of grey-box and black-box prediction schemes that use data mining and data analytics to predict the energy consumption of a particular load [53]. Several demand prediction approaches include regression methods such as simple regression model (SRM), multiple linear regression (MLR), artificial neural network (ANN) methods, decision trees (DT), support vector machine (SVM) and other time series approaches and statistical methods that be used for both regression and classification.

Further modifications of such standard machine learning techniques such as least squares SVM (LS-SVM) and parallel SVM [54] have also been implemented in the field of demand prediction. Neural networks perform well for benchmarking purposes with high prediction capability after selecting important model input parameters. ANN and SVM require many parameters and might become computationally expensive, but their prediction accuracy is usually better than decision trees and statistical algorithms. Decision trees and other statistical algorithms, however, are generally easy to use and computationally inexpensive [55].

There have also been numerous statistical based methods to predict demand previously proposed in the literature. Some of these methods include conditional demand analysis (CDA), auto regressive moving average with exogenous inputs (ARMAX) and Gaussian mixture models (GMM). These prediction methods rely on historical load data and exogenous variables such as weather forecasting (temperature, humidity, cloud cover etc.) [56]. Since it is difficult to model the (non-linear) weather dependencies on load accurately, concepts such as heating degree-days (HDD) and cooling degree-days (CDD) and cloud cover Oktas for cloud cover have been previously proposed. Clustering algorithms have been widely used in the literature to perform a first step in finding representative buildings (centroids) and develop archetypes in the white-box and grey-box prediction methods. Moreover, forecasting models based on similarity to group centroids have been proposed to forecast energy consumption of a certain year on the basis of measured values and influence factors.

3.3 Demand Prediction based on Historical Load Observations

3.3.1 Principal Component Analysis

Principal component analysis (PCA) is a technique of identifying prevalent patterns in data by decomposing the data into features that highlight the similarities and differences. Data with a high dimensionality for which visualization is cumbersome, PCA is a powerful tool for analyzing the data into a finite number of relationship between data dimensions with meaningful (physical) interpretation. PCA is generally performed via an eigenvalue decomposition of the covariance matrix [57]. In addition, a singular value decomposition (SVD) of a data matrix also provides the principal components. The singular value decomposition (SVD) is a matrix factorization method in linear algebra used to decompose a $m \times n$ matrix A such that

$$A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T \tag{3.1}$$

where U is an $m \times n$ orthogonal matrix, $U = (u_1, u_2, \ldots, u_n)$, whose columns are the eigenvectors of AA^T, V is an $n \times n$ orthogonal matrix, $V = (v_1, v_2, \ldots, v_n)$ whose columns are the eigenvectors of A^TA; Σ is an $n \times n$ diagonal matrix, $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$ with non-negative real values k, where $k = 1, 2, \ldots, n$ called singular values, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$.

An SVD has the merit of providing the principal (main) components in a data matrix. This is realized by considering only a limited number of singular values (called the order) in the diagonal matrix Σ in (3.1) and corresponding to the principal components of the data matrix A. The other non-principal components are attributed to noise effect on the measured data. The idea of principal components is applied to an $n \times p$ power (demand) data matrix P, where n is the number of power samples over a day, and p is the number of days over which power was measured. Performing an SVD yields a decomposition

$$P = U\Sigma V^T$$

where, singular values $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ are the square roots of the eigenvalues of the covariance matrix P^TP, the right singular vectors V^T are principal directions and the principal components are given by $U\Sigma$ under the assumption that P is centered i.e. the column mean is equal to zero.

3.3.2 Component based Clustering for Load Classification

Clustering is one of the most widely used unsupervised learning classification methods to organizing objects into groups whose members are classified to be similar.

Cluster analysis groups data objects based only on information found in the data to distinguish the objects and their properties.

The method of k-means clustering partitions n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Consider m points $P = \{p_1, p_2, \ldots, p_m\} \subseteq R^n$ and an integer k denoting the number of clusters. The objective of k-means is to find a k-partition of P such that points that are close neighbors to the same cluster and points that are far from each other belong to different clusters based on an L_2- or L_1-norm. A k-partition of P is a collection $S = S_1, S_2, \ldots, S_k$ of k non-empty pairwise disjoint sets which covers P. Let $s_j = |S_j|$ be the size of S_j where $j = 1, 2, \ldots, k$. For each set S_j , let $\mu_j \in R^n$ be its centroid such that

$$\mu_j = \frac{1}{s_j} \sum_{p_i \in S_j} p_i$$

The k-means objective function is

$$F(P,S) = \sum_{i=1}^{m} |p_i - \mu(p_i)|_r$$

where $\mu(p_i) \in R^n$ is the centroid of the cluster to which p_i belongs $r = 1$ or 2 based on L_1 or L_2 minimization. The objective of k-means clustering is to compute the optimal k-partition of the points in P,

$$S_{opt} = \text{arg}\min_{S} F(P,S)$$

3.3.3 Demand Prediction Based on Basis Functions for Independent Clusters

The combination of PCA of the power demand data followed by clustering leads to the cluster groups denoted by C, containing particular power demand sets of data. Performing SVD on data in each such cluster yields basis functions for each individual cluster. Formally, for a $n_c \times p$ (demand) data matrix P_c, where n_c is the number of samples and p is the number of variables in cluster $c \in C$, performing SVD yields a decomposition $P_c = U_c \Sigma_c V_c^T$ where singular values $(\sigma_{c1}, \sigma_{c2}, \ldots, \sigma_{cn})$ are the square roots of the eigenvalues of the covariance matrix $P_c^T P_c$, the right singular vectors V_c^T are principal directions and the principal components are given by $U_c \Sigma_c$ under the assumption that P_c is centered i.e. the column mean is equal to zero.

For each $P_c^{n_c \times p}$ belongs to c, the predicted power, P_c can be formulated according to

$$P_c = \theta_c V_c^T + E_c \tag{3.2}$$

where $V_c^{n_c \times p}$ is the computed basis or the principal direction based on training (data) set, $\theta_c = U_c \Sigma_c$ is the principal component of power data , P_c belonging in cluster $c \in C$ and E_c is the error in prediction resulted in lower order approximation of the training data. The initial value of θ_c for prediction can be equal to the mean value of all the previously recorded θ_c in that cluster, and then can be recursively estimated as discussed in the following sections.

3.4 Refinement of Power Demand Prediction based on Weather Variability

Variations in weather conditions are important factors to be considered in modeling electricity demand [58]. Parametric and non-parametric models of temperature, humidity and wind speed have been accounted for and well documented in the literature [41], [45], [59]. However, since it is difficult to model the (nonlinear) weather dependencies on load accurately, concepts such as heating degree-days (HDD) and cooling degree-days (CDD) have been previously proposed. In this chapter, similar concepts are used to formulate a binary distinction of demand data depending on the degree of weather dependency (hot or cold) on the load in a particular data set. In addition, other variables such as calendar effects (including extracurricular events with anticipated load demands) that tend to happen during unusual time of the day are also considered for more accurate load forecasting.

To include the weather and other variables in the prediction algorithm, subspace clustering method are incorporated to distinguish load data according to exogenous variables. Subspace clustering is an extension of traditional clustering that separates clusters in different subspaces within a dataset. For subspace clustering, let $w_{ij} = 1$ if point j belongs to subspace i and $w_{ij} = 0$ otherwise. Let us assume that the number of subspaces equals to n and the subspace dimensions d_i are known. Our goal is to find the subspace bases $\{V_i \in R^{D\times d_i}\}_{i=1}^n$, the low-dimensional representations $\{\mu_i \in R^D\}_{i=1}^n$ and the segmentation of the data w_{ij} for $j = 1, \ldots, N$, $i = 1, \ldots, n$. This can be done by minimizing the sum of the city-block distances (L_1) or squared distances (L_2) from each data point to its own subspace

$$\arg\min_{V_i,\mu_i,w_{ij}} \sum_{j=1}^{N}\sum_{i=1}^{n} w_{ij}|x_j - \mu_i V_i^T|_{L_1,L_2} \tag{3.3}$$

subject to $w_{ij} \in \{0, 1\}$. Given μ_i, V_i, the optimal value for w_{ij} is

$$w_{ij} = \begin{cases} 1, & \text{if } i = \arg\min_{k=1,\dots,n} |x_j - \mu_k V_j^T|_{L_1,L_2} \\ 0, & \text{otherwise} \end{cases}$$

The predicted power for subspace i, P_i can be formulated according to (3.2) with the subspace bases as computed in (3.3).

3.5 Real-time Prediction of Demand Based on Synchrophasor Measurements

To meet this computation in real-time, the data up to time k is condensed into a memory vector $S(k)$ so that the recursive estimate $\hat{\theta}^{k+1}$ of the parameter $\hat{\theta}^k$ can be recursively updated according to

$$\hat{\theta}^{k+1} = T(\hat{\theta}_k, S(k), P_c(k)) \tag{3.4}$$

where T is a Kalman recursive filter, satisfying the discrete state equations:

$$\theta(k+1) = \theta(k)F(k) + v(k) \tag{3.5}$$

$$P_c(k+1) = \theta(k+1)V^T(k) + E(k) \tag{3.6}$$

where $\theta(k)$, $F(k)$, $P_c(k)$, $V(k)$, $v(k)$ and $E(k)$ are vectors of $n \times 1$ system states, the $n \times n$ dimensions of the state transition matrix, $m \times 1$ measurement vectors, the $m \times n$ output matrix, $n \times 1$ system error and $m \times 1$ measurement error, respectively. The noise vectors, $v(k)$ and $E(k)$, are drawn from white Gaussian noise that has a mean of zero and no time correlation and Q_1 and Q_2 are the covariance matrices of

$v(k)$ and $E(k)$ respectively. The state matrix F can be determined by solving least squares minimization as described in [57].

Phasor (voltage, current) measurements at time intervals $t_k = k * T_s$, where T_s is the sampling time and k is the positive integer used for the time index and recursion are used to calculate real (reactive) power at the point of interconnect at each interval t_k. Typically, T_s is a multiple of $1/60$ seconds for synchrophasor sampling.

$$P(t_k) = V_m(t_k) I_m(t_k) cos(\theta_v(t_k) - \theta_i(t_k)) \tag{3.7}$$

where $\mathbf{V} = V_m \angle \theta_v$ and $\mathbf{I} = I_m \angle \theta_i$ are the voltage and current phasors respectively.

Given an *a priori* estimate of the state vector $\hat{\theta}(0) = \hat{\theta}_0$ and the corresponding error covariance matrix $P_c(0) = P_{c_0}$, the index k is set to $k = 0$ and then the Kalman estimator is used to estimate the next state by recursively computing the following equations for measurement and time updates:

$$\begin{gathered}
\hat{\theta}(k+1) = \hat{\theta}(k) + K(k)[P_c(k) - V^T(k)\hat{\theta}(k)] \\
K(k) = [P_c(k)V(k)][H(k)P_c(k)V(k) + Q_1]^{-1} \\
\hat{P}_c(k+1) = [I - K(k)V^T(k)]P_c(k)[I - K(k)V^T(k)]^T + \\
K(k)Q_2K^T(k)
\end{gathered}$$

where $K(k)$ is the Kalman gain. For the Kalman filter, the choice of an *apriori* estimate of the state $\hat{\theta}_0$ and its covariance error P_{c_0} is made from the off-line prediction values. It should be noted that the parameter estimate $\hat{\theta}$ is being recursively updated using measurements P_c but not V^T because the degree of freedom for the estimates of V^T is higher than $\hat{\theta}$ when with small number of principal base load components.

3.6 Dynamic Dispatch Optimization using Predicted Data

In this section, a specific example of an objective function is chosen to minimize the total electricity cost of a microgrid system. A microgrid may be defined as distribution level power systems that consist of a group of power consuming loads interconnected with a set of power producing DERs that operate as a single controllable system in the main grid connected mode or islanded (grid disconnected) mode. In such a microgrid, the objective is to minimize the total electricity cost by controlling the power flow at the point of interconnect of the system. For the sake of the dispatch, the combination of DERs such as batteries, diesel generators, solar PVs, etc. are assumed to act as a single lumped DER called lumped distributed energy resource (LDER).

3.6.1 Energy and Cost Calculation from Real-Time Synchrophasor Power Measurements

Assuming $P(t_k) < 0$ when real power is produced from the microgrid to the point of interconnection at time t_k, energy usage at time t_k is denoted by $E_{kWh}(t_k)$ and with sampling time T_s, energy usage in the units of kilowatt-hour [kWh] can be computed by

$$E_{kWh}(t_k) = \frac{T_s}{3600 \cdot 1000} \cdot P(t_k) + E(t_{k-1}) \tag{3.8}$$

However, for the computation of energy cost, down-sampled energy usage $E_{kWh}(t_m)$ is instead used where

$$E_{kWh}(t_m) = E_{kWh}(t_k) \qquad k = \frac{T_m \cdot 60}{T_s} \cdot m \tag{3.9}$$

and $m = 1, 2, \ldots$ is now the positive index for the time index t_m. With $t_m = m \cdot T_m$, it is clear that the down-sampled energy usage $E_{kWh}(t_m)$ has a sampling time of T_m.

The energy cost takes into account a time dependent rate $R_u(t_m)$ for electric energy usage (negative for generation) at a seasonally dependent time-of-day $t_m = m \cdot T_m$. Hence the time-of-use energy cost $C_{tou}(t_m)$ is defined as

$$C_{tou}(t_m) = R_u(t_m) \cdot E_{kWh}(t_m) \tag{3.10}$$

The total energy cost C_{tou}^N over a billing period can now be computed as

$$C_{tou}^N = \sum_{m=0}^{n} C_{tou}(t_m) + \sum_{m=n}^{N-1} C_{tou}(t_m) \tag{3.11}$$

where n is the number of completed intervals in the billing cycle at any given time t_n and N is the total number of intervals in the billing cycle.

Similarly, down-sampled real power, denoted by $P(t_m)$ at the time intervals $t_m = m \cdot T_m$ is used for demand cost calculation and is computed from down-sampled energy usage $E_{kWh}(t_m)$ via

$$P_{kW}(t_m) = 3600 \cdot \frac{E(t_m) - E(t_{m-1})}{T_m} \tag{3.12}$$

It is important to recognize that the down-sampled real power $P(t_m)$ is not the down-sampled real power $P(t_k)$, but computed from down-sampling the integral of the real power $P(t_k)$.

The demand cost takes into account a time dependent price $D_u(t_m)$, that is used to penalize the maximum electric power usage

$$P_{max}(t_p) = \max_{m=0,1,2,\ldots,n,\ldots,N-1} P_{kW}(t_m) \tag{3.13}$$

(also known as demand) over a billing period $t_p = t_m$ for $m = 0, 1, 2, \ldots, N-1$ at a seasonally dependent time-of-day $t_m = m \cdot T_m$.

Hence the demand cost at the end of the billing period C_{dem}^N is the cost of the maximum electric power usage $P_{max}(t_p)$, weighted by a season dependent demand pricing $D_u(t_p)$ given by

$$C_{dem}^N = P_{max}(t_p) \cdot D_u(t_p) \tag{3.14}$$

3.6.2 Objective Formulation for Real-Time Scheduling Updates

The finite horizon objective function J_n at any given time t_n for the cost minimization is the sum of two costs and is defined as

$$J_n := \text{arg} \min_{\bar{p}_{kW}(t_n)} \{C_{tou}^{N-n-1} + C_{dem}^N\} \tag{3.15}$$

subject to the dynamically changing constraint set $\mathbf{C_n}$ that includes resource and operational constraints where $\bar{p}_{kW}(t_n)$ is the power generated/consumed from the LDER to reduce the total cost J_n. Typically $\bar{p}_{kW}(t_n) < 0$ for power consumed from the point of interconnect.

The part of J_n is a minimax estimation of $\bar{p}_{kW}(t_m)$ for C_{dem}^N. However, assuming that $D_u(t_m) \gg R_u(t_m)$, the C_{dem}^N term has more weight to the optimization than the C_{tou}^N term. Hence, the following transformation can be used to explicitly find the global minimum.

Minimax Optimization

The minimax optimization problem can be stated as:

$$\min_x M_f(x)$$

where

$$M_f(x) = \max_{1 \leqq i \leqq m} f_i(x)$$

and $x = [x_1, x_2, \cdots, x_n]^T$.

Remark 1. *This minimax problem can be transformed into an equivalent one via a epigraph transformation to obtain an explicit solution and for algorithmic purposes. The epigraph transformation preserves convexity of the optimization problem [60].*

In the epigraph form, the $C_{dem}^{\prime N}$ can be transformed into an inequality constraint where the upper bound of $P_{kW}(t_m)$, $P_{kW}^{opt}(t_m)$ can be iteratively computed until the optimization finds a local minimum. However, due to existence of non-unique minima, the freedom on $P_{kW}(t_m)$ for computing minima results in ill-conditioning of the problem. Hence an extra penalty term on the rate of change on $\bar{p}_{kW}(t_m)$ needs to be added to make the problem well-conditioned and to prevent losses due to round-trip efficiencies to/from distributed energy resources. Hence, the resulting optimization becomes:

Hence, the objective function J_n for the cost minimization at any given time t_n is now defined as:

$$\min_{\bar{p}_{kW}(t_n)} \quad \sum_{m=n}^{N-n-1} R_u(t_m) \cdot E_{kWh}(t_m) + \alpha(P_{kW}(t_n) - P_{kW}(t_{n-1}))^2 \tag{3.16a}$$

$$\text{s.t.} \quad P_{kW}(t_n) = \bar{p}_{kW}(t_n) + P_c(t_n) \tag{3.16b}$$

$$P_{kW}(t_n) \leq P_{kW}^{opt}(t_n) \tag{3.16c}$$

$$\eta \cdot p_{kW}^{min} \leq \bar{p}_{kW}(t_n) \leq 1/\eta \cdot p_{kW}^{max} \tag{3.16d}$$

$$C_{kWh}^{min} \leq C_{kWh}(t_n) \leq C_{kWh}^{max} \tag{3.16e}$$

$$C_{kWh}(t_{n_i}) = C_{kWh}(t_{n_f}) \tag{3.16f}$$

where $P_c(t_n)$ is the predicted load at time t_n, η is the effective round-trip efficiency of the LDER, p_{kW}^{min} and p_{kW}^{max} are the charging and discharging (power) limits respectively of the LDER, C_{kWh}^{min} and C_{kWh}^{max} are the minimum and maximum energy stored respectively in the LDER, $C_{kWh}(t_{n_i})$ and $C_{kWh}(t_{n_f})$ are the initial and final energy levels at the beginning and the end of the dispatch period of the LDER respectively and α is the weighting factor associated with penalizing the subsequent power flow changes to/from the LDER. The energy stored in the LDER at a given time t_n is given by

$$C_{kWh}(t_n) = -T_d \cdot \bar{p}_{kW}(t_n) + C_{kWh}(t_{n-1}) \tag{3.17}$$

where T_d is the dispatch interval and $\bar{p}_{kW}(t_n) > 0$ means the power is generated from the LDER.

3.6.3 Computation of $P_{kW}^{opt}(t_m)$

Since $P_{kW}(t_m)$ is now the constraint of the minimization, the upper bound of $P_{kW}(t_m)$, $\bar{P}_{kW}^{opt}(t_m)$ can be computed iteratively using the bisection algorithm. Initially, the bounds of $P_{kW}(t_m)$ are set for the bisection algorithm [61]. These values typically are $[0, P_{kW}^{max}]$ where P_{kW}^{max} is the value of $P_{kW}(t_m)$ when $\bar{p}_{kW}(t_m) = 0$. The optimization is first run with the constraint $P_{kW}(t_m) \leq \bar{P}_{kW}^{opt}(t_m) = 0$. If the optimization does not converge or find a solution, the new value of $\bar{P}_{kW}^{opt}(t_m)$ is set to the mean of 0 and $P_{kW}^{max}(t_m)$. The iteration is repeated until the optimization converges to find a solution for J_n. However, because C_{dem}^{N} is only dependent on the maximum value of $P_{kW}(t_m)$ over the period $m = 1, 2, \ldots, N$, the previous information of the maximum should be carried over at each interval when the optimization is run. Hence, the actual value of $P_{kW}^{opt}(t_m)$ would be the supremum of the computed value from the bisection algorithm $\bar{P}_{kW}^{opt}(t_m)$ and the maximum of the demand $P_{kW}(t_n)$ for

$n = 0, 1, \ldots, m-1$ intervals of the billing period,

$$P_{kW}^{opt}(t_m) = sup\{\bar{P}_{kW}^{opt}(t_m), \max_{n=0,1,\ldots,m-1} P_{kW}(t_n)\} \tag{3.18}$$

3.7 Application to Demand Data from a Building

3.7.1 Metered Data

To illustrate the clustering based prediction algorithm on power demand data, 15-minute interval load (kW) data for one calendar year (2015-2016) has been collected from a revenue meter of a building. Pre-processing of the demand data is performed in order to clean the power data by eliminating data with dropouts ('0' elements) and anomalies (low power elements) in the real measurements.

Weather data for the same time frame was obtained from Weather Underground[1]. The hourly historical measurements of temperature (^{0}C) and relative humidity ($\%$) were downloaded from this weather station to allow further correlation analysis with the power demand data.

3.7.2 Principal Component Analysis For Clustering Based On Load Data

Pre-processing the data yielded 363 days of 15-minute interval demand (load) data for the commercial load. The training set consist of 360 days of power data and 3 days as testing/validation data. Hence the size of the resulting data matrix (P) is 96×360. To analyze this power data matrix, the principal components in the data are computed via the singular value decomposition (SVD) and the results are depicted in

[1]Weather Underground is a commercial weather service providing real-time weather information via the Internet, additional information at www.wunderground.com

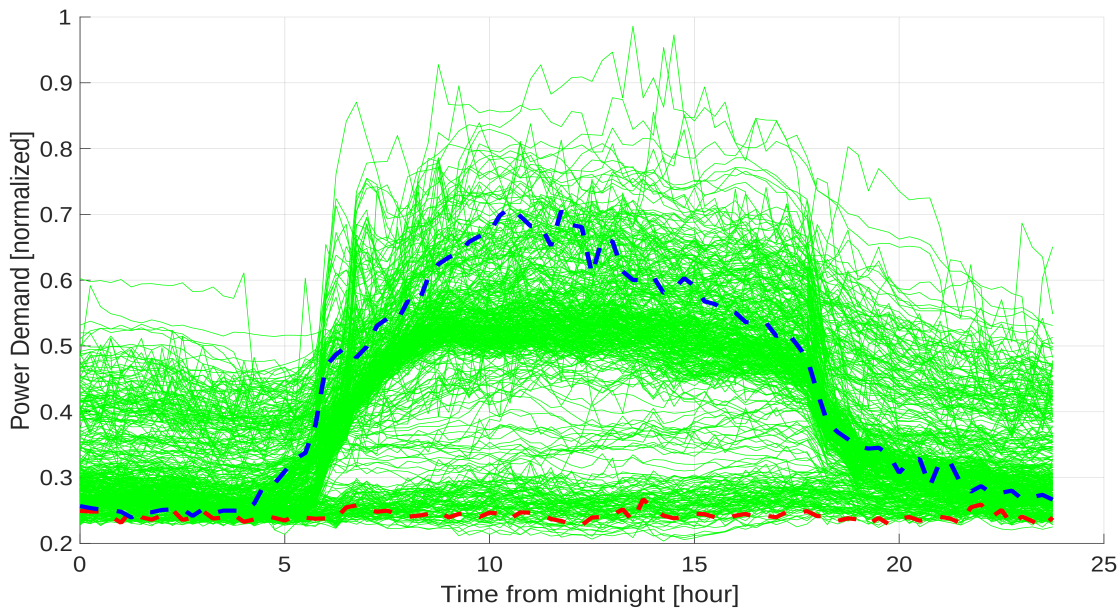

Figure 3.1: Historical normalized daily power demand data (green lines) measured at 15 minute intervals for the entire year and used to create data matrix P, where dashed blue and dashed red lines indicate a typical "high demand" and "low demand" days respectively.

Fig. 3.2. From Fig. 3.2(a), it can be seen that the singular values are insignificant after order 2, indicating that the shape of the historical power data can be reduced to two main principal components. The shape of the two principal components in depicted in Fig. 3.2(b) and identified as the two bases V_1 and V_2 for the power data P. Subsequently, each of the power data sequences $P_c,\ c = 1, 2, \ldots, p = 360$ (columns in the power data matrix P) can be written as a linear combination

$$P_j = \theta_j(1)V_1 + \theta_j(2)V_2 + E_j$$

where the values of the parameters $\theta_j(1)$ and $\theta_j(2)$ for $j = 1, 2, \ldots, p$ are found by minimizing the Least Squares error on E_j. The (spread of the) numerical values for $\theta_j(1), \theta_j(2)$ for $j = 1, 2, \ldots, p$ is depicted in Fig. 3.2(c) and the resulting fit $\theta_j(1)V_1 + \theta_j(2)V_2$ on the actual power data P_j for $j = 1, 2, \ldots, p$ is depicted in

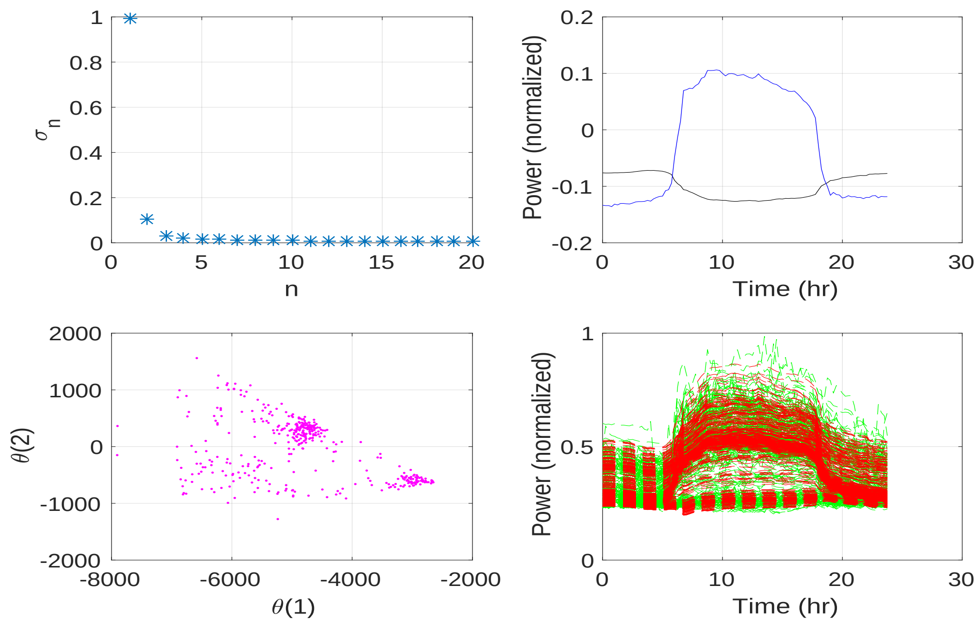

Figure 3.2: (a) Top left figure shows singular values of the data matrix P, consisting of the normalized daily power data displayed in Fig. 3.1 showing that it has two main directions, (b) Top right figure shows the two bases for P, (c) Bottom left shows principal components of the two dimensional vector θ, for a two dimensional PCA of the data matrix P, and (d) Bottom right figure shows actual measured (dashed green) and reduced dimensional fitted (dashed red) power in kW.

Fig. 3.2(d). It can be observed that the reduced two-dimensional combinations of the bases V_1 and V_2 depicted in Fig. 3.2(b) gives an excellent coverage of the measured power data.

Based on the parameter values $\theta_j(1), \theta_j(2)$ for $j = 1, 2, \ldots, p = 360$, the k-means clustering algorithm can be performed to separate the power data into two distinct clusters to be able to classify the data and use those classifiers for prediction. Fig. 3.6 shows the results of k-means L_1-norm "city block" clustering of the principal components into two clusters. The resulting clusters can now be interpreted as 'high-demand days or working days' and 'low-demand days or non-working days'.

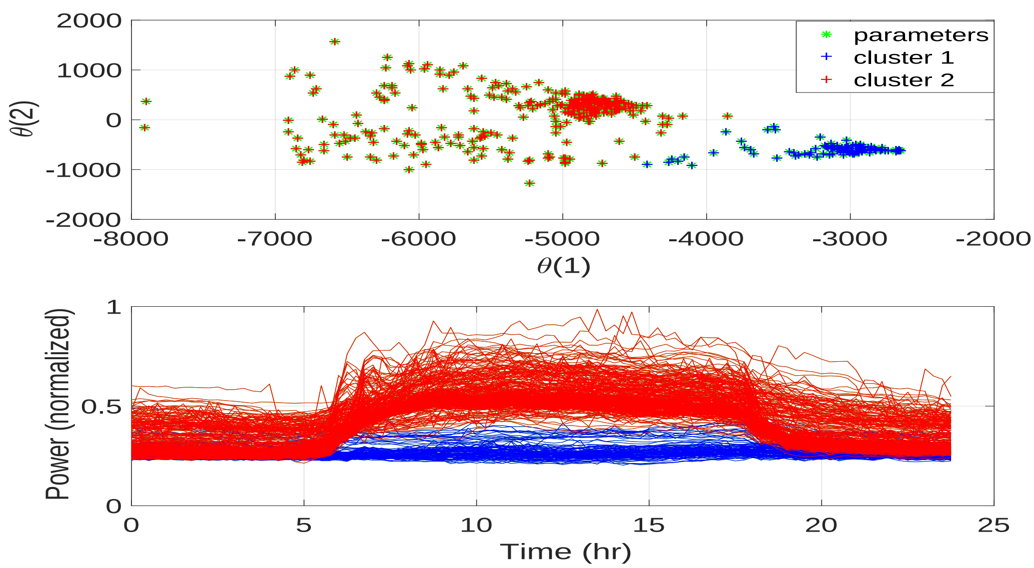

Figure 3.3: (a) Figure on the top shows the results of k-means clustering resulting in two clusters, b) Bottom figure shows the actual power demand data of each cluster indicating the relevance of cluster analysis and automatic classification of data in "high demand" (working) and "low demand" (non-working) days.

The characterization of the two dimensional basis V_1, V_2 and the k-means clustering can now be used for a static day-ahead prediction of the power demand. For the prediction one only needs to have prior information about which cluster the day would fall into. Once this information is known, the power can be predicted by multiplying the basis V_1, V_2 by a proper choice of initial estimate of a parameter θ_1, θ_2.

3.7.3 Sub-Space Clustering Based on Load, Weather and Miscellaneous Data

Fig. 3.6 shows the non-conformity of data around the cluster centroid. The non-conformity is especially clear for working days, indicating that there is an effect of other variables (weather, other activities) on the power demand of the commercial load during the working days. Thus, classification solely based on two clusters resembling working and non-working days for the commercial load does not yield the best prediction results. A multi-dimensional subspace clustering technique can be used to account for other important features such as temperature, humidity and extra-curricular events for refined classification and prediction.

The weather dependency can be seen only on working days due to (higher) human activity related to the the commercial load and therefore sub-clustering may be performed only on the working days. Furthermore, relative humidity plays a vital role in the electricity consumption of a commercial load due to increased load on HVAC, especially during the summer. The correlation between temperature and load was found to be insignificant as compared to humidity H and hence only humidity H was used for sub-space clustering of the load data on working days. Furthermore, it was found that a logarithmic humidity $log_{10}(H)$ has a stronger correlation than the absolute value H of humidity, hence $log_{10}(H)$ was used for analysis. Also, since the HVAC system was switched on every Monday morning in the commercial load, the load on Mondays was higher than the other working days and therefore Mondays are classified as other miscellaneous days affecting the load of the commercial load by assigning a higher weighting factor (λ) for Mondays (0.6) as compared to the other working days (0.4).

Now the sub-space clustering algorithm is implemented using 3 sub-spaces: principal components of load (weighted by the parameter θ), logarithmic humidity ($log_{10}(H)$) and a weighting on miscellaneous days (λ). The subspace clustering results of working days are depicted in Fig. 3.7.

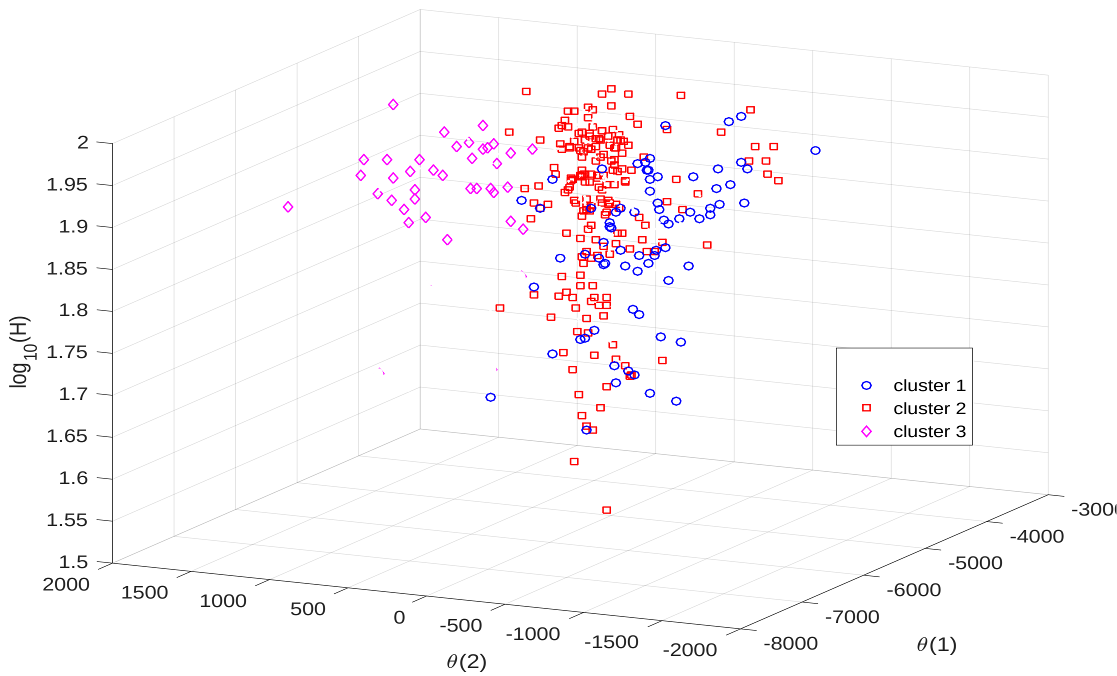

Figure 3.4: Sub-space clustering of working days in three different clusters. This figure shows the subspace clustering projected onto the 3-D space of θ and $log_{10}(H)$, showing clear distinction between clusters.

The number of clusters were chosen to be 3. Cluster 3 includes load days with higher relative humidity ($> 85\%$) and hence the load on these days is observed to be higher than the rest of the working days. Cluster 2 includes all other working days that are not affected by weather or other miscellaneous activities in the commercial load. Cluster 1 includes miscellaneous days such as Mondays where the load is higher than the other normal working days. Once these clusters have been identified, an

SVD can be performed to obtain basis (V_c) and components (θ) to use for prediction. For a day ahead prediction, one needs to have prior information about which cluster the day would fall into, usually determined by weather forecasts and miscellaneous activities. Once this information is known, the power can be predicted by multiplying the calculated sub-basis by a proper choice of initial estimate of θ (such as the mean value of theta of that particular cluster).

3.7.4 Prediction of Demand Based on Clustering and Recursive Estimation

To illustrate the effectiveness of the clustering based prediction algorithm, the prediction of the power demand of the commercial load is demonstrated for two days with different profiles using the basis obtained from the training set above. For an offline day-ahead prediction using single criterion (1-level) clustering i.e. working or non-working days, we consider the basis and multiply it with θ_{mean} of that particular cluster. The prediction results for 1-level (offline) clustering for a working day and a non-working day are shown in Fig.3.8(a) and Fig.3.8(b) respectively. It can be seen that 1-level clustering results in good prediction for a non-working day because of the non-correlation of load with exogenous variables such as weather.

However, for a working day (in this case, a high temperature and high humidity Monday) the load heavily depends on these exogenous variables and hence baseline 1-level prediction would not yield good results. It is clear that sub-space clustering for offline prediction is necessary for prediction of load on such a working day.

Fig. 3.8(c) depicts the offline prediction results using sub-space clustering with variables such as humidity, load on Mondays etc. and then performing SVD. It can be seen that the prediction yields much better results as other factors such as humidity and day of the week have been accounted into the basis and principal components. However, use of recursive Kalman filter for prediction update would result in even better prediction results as depicted in Fig. 3.8(d). The Kalman filter uses power measurement updates every 15 min to recursively correct the prediction values to be closer to the actual profile of the day. Due to the extra information of actual power measurements throughout the day, the combination of the static clustering of power data and the dynamic Kalman filter for on-line recursive updates of power demand predictions results in the most accurate prediction of power demands in the commercial load. The predictions on a given day do require prior information of the other variables such as weather forecasts and day of week, also observed by [62], [57].

3.7.5 Cost Optimization using Predicted Data

In this section, the dispatch optimization is demonstrated at a given (15-min) interval for a daily (24-hr) dispatch to minimize the total electricity cost of the microgrid over a (monthly) billing period. Two different scenarios are shown where: (1) The dispatch at the (first) interval of the 24 hour period where the actual POI measurement power has higher disturbances and variations than that to the static day-ahead prediction and (2) The dispatch at the (mid-day) interval at $n = 48$ where the actual POI measurement power has much higher offset than that to the static day-ahead prediction for a certain period. Once the load prediction is available from Sec. 3.5, the dispatch optimization problem is solved as in Sec. 3.6.2 to minimize demand charges and energy time of usage. In this example, to demonstrate the optimization results, the LDER consists of a single 620 kWh/250 kW Li-ion battery. The parameters assumed for the analysis are: $C_{kWh}^{full} = 620$ kWh, $\eta = 87\%$, $p_{kW}^{min} = -250$ kW (charging), $p_{kW}^{max} = 250$ kW (discharging), $C_{kWh}^{min} = 62$ kWh, $C_{kWh}^{max} = 558$ kWh

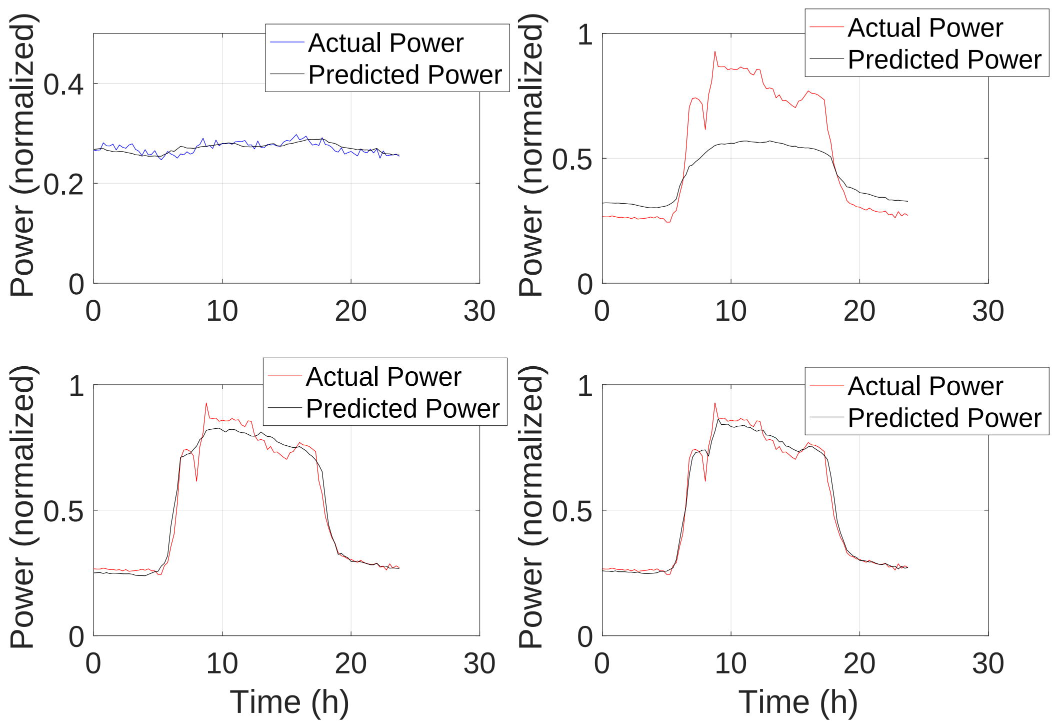

Figure 3.5: Power prediction based on clustering: (a) Top left figure shows the (offline) power prediction on a non-working day (1-level clustering), (b) Top right shows power (offline) prediction on a hot Monday with 1-level clustering, (c) Bottom left figure shows (offline) prediction on a hot Monday with 2-level subspace clustering and (d) Bottom right figure shows (online) prediction with 2-level clustering and recursive estimation with measurements updates using a Kalman filter.

($10\% - 90\%$ of full charge), $C_{kWh}(t_{n_i}) = 310$ kWh and $C_{kWh}(t_{n_f}) = 310$ kWh (50%) and $T_d = 15$ min. The weighting factor α is tuned to minimize power chattering to/from the LDER and is set at 0.0007. The maximum dispatch period has 96 intervals i.e. $m = 0, 1, ...95$. For the billing period of one-month, the demand pricing $D_u(t_p)$ is assumed to be constant at 29$/kW. The time-of-use energy price per kWh for a working day (weekdays) is given by

$$R_u(t_m) = \begin{cases} 6 \text{ ¢}, \; for \; p \in [0, 23] \cup [72, 87] \\ 10 \text{ ¢}, \; for \; p \in [24, 39] \cup [88, 95] \\ 12 \text{ ¢}, \; for \; p \in [40, 71] \end{cases} \tag{3.19}$$

where $m = 96 \cdot (d-1) + p$; d is the working day of the billing month and $R_u(t_m) = 6$ ¢ for a non-working day (weekends and holidays).

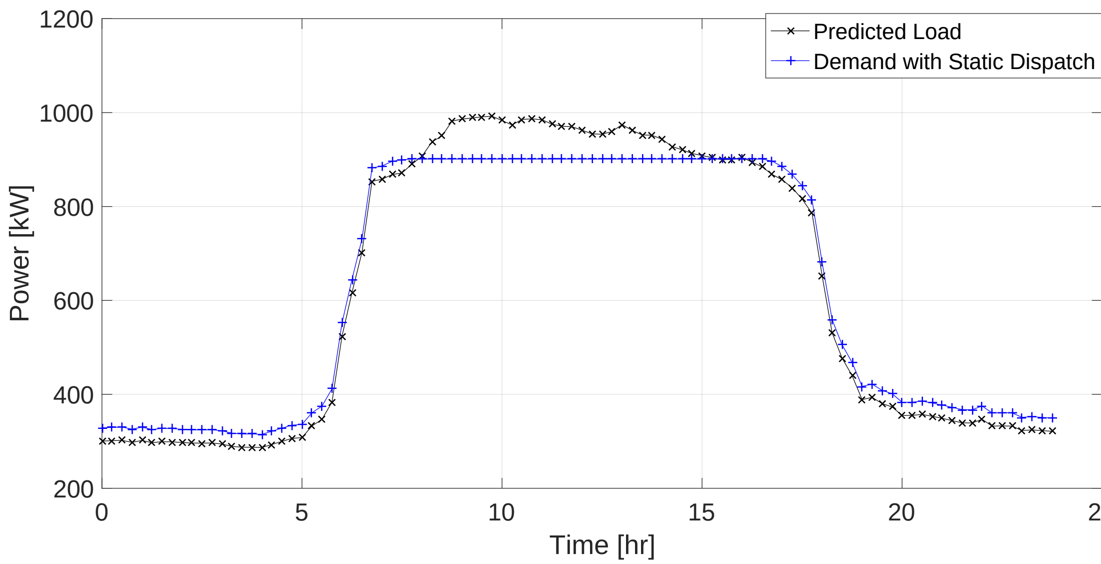

Figure 3.6: Dispatch optimization results for a static day-ahead prediction profile. The cross (black) curve shows the demand at the POI based on offline prediction and dashed (blue) curve shows the estimated demand at POI with the generated dispatch strategy.The optimization is executed at midnight for the next 24 hour period taking into consideration the electricity costs and constraints.

In the first test case, the dispatch is calculated at the first interval, i.e. $n = 1$ for the billing period consisting of 30 days at 15-min intervals and hence 96×30 intervals in the billing period. Using the load prediction, $P_c(t_m)$, of the microgrid from Sec. 3.5 for a day ahead case i.e. with 96 intervals, the dispatch is calculated for $m = 1, 2, \ldots, 96$ by minimizing Eq. (3.16) with constraints in Eq. (3.16b)-(3.16f). The optimization results in this case are shown in Fig. 3.6 and 3.7. Fig. 3.6 shows the estimated optimized power and measurement at the POI for the 24 hour period as evaluated in Sec. 3.5 with the dispatch schedule executed at midnight for the 24 hour period taking into consideration the electricity costs and constraints. It can be seen that the optimization clearly results in peak shaving of power at the POI to primarily minimize demand charges while shifting the energy for time-of-usage pricing. The transformation used in Sec. 3.6 results in the maximum savings in the total electricity cost. Fig. 3.7 shows the state of charge (SoC) of the battery in $\%$ at time t_n which is defined

$$SoC(t_n) = C_{kWh}(t_n)/C_{kWh}^{full} \times 100 \tag{3.20}$$

In this case, since the peak demand period is coincident with the highest time-of-use rates, as in Fig. 3.7, the battery charges during the low time-of-usage rate in the earlier part of the day to be able to produce power to minimize demand and also reduce the consumption during high rates. Fig. 3.7 also shows the battery (output) power at each interval of the dispatch period to meet the objectives and constraints. It could be seen that the battery discharges at the maximum rate around the peak to shave the peak demand and result in maximum achievable projected[2] savings. Fig. 3.8 shows the comparison of the estimated demand at the POI using a planned (static) dispatch schedule and (real-time) dynamic dispatch updated every 15-min. It can be

[2]The projections are based on the assumption that the predicted load is similar for all the working days and a similar dispatch schedule is executed for every working day of the billing period. For the non-working days, the battery is assumed to be offline i.e. no power is generated or absorbed.

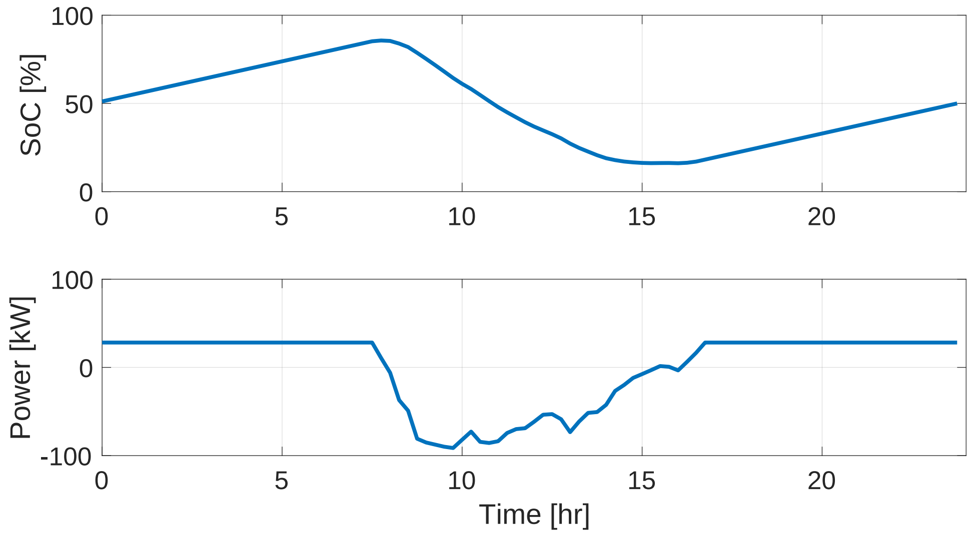

Figure 3.7: Static dispatch optimization results for a static day-ahead prediction. The figure on the top shows the planned battery state of charge (SoC) for the 24 hour dispatch period and the figure on the bottom shows the planned power schedule of the battery for the dispatch period to obtain optimized demand as shown in Fig.3.6.

seen that the dispatch schedule generated by the static optimization as in Fig. 3.7 does not result in expected savings due to not peak shaving the unanticipated peaks as in Fig. 3.8 but however the dynamic optimization is able to peak shave at a level to yield better savings.

In the second test case, the dispatch is calculated at the middle of the dispatch period at $n = 48$ for the same billing period as in case (1). However, the actual load is much higher than predicted for a period $n = 49$ to $n = 63$. Using the recursively updated load prediction, $P_c(t_m)$, of the microgrid from Sec. 3.5 at 49^{th} interval, the dispatch is calculated for $m = 49, 50, \ldots, 96$ by minimizing Eq. (3.16) with constraints in Eq. (3.16b)-(3.16f). The optimization results in this case are shown in Fig. 3.8 and Fig. 3.9. It shows the estimated optimized power and measurement at the POI for the 24 hour period as evaluated in Sec. 3.5 with the static dispatch schedule executed at the midnight of the 24 hour period taking into consideration the electricity costs and constraints. The figure also shows the dynamic optimization

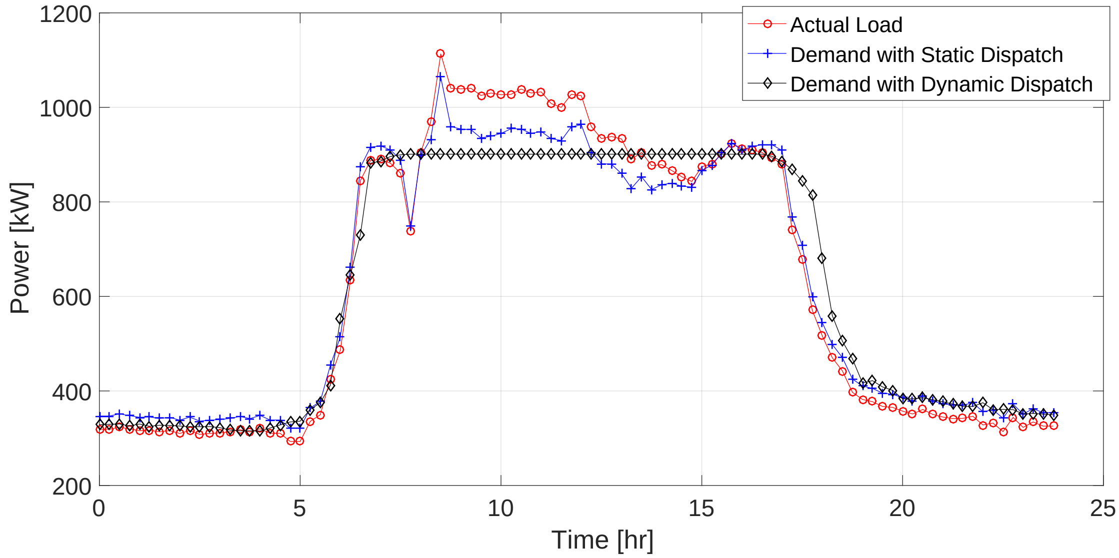

Figure 3.8: Real-time dispatch optimization results comparing the static and dynamic prediction and dispatch for case 1. The red curve shows the actual measured demand at the POI and blue curve shows the estimated demand at POI with the generated dispatch strategy with static predictions. However, the black curve shows the actual demand of POI with real-time predictions and optimization. The optimization is executed every 15-mins for the 24 hour period taking into consideration the electricity costs and constraints.

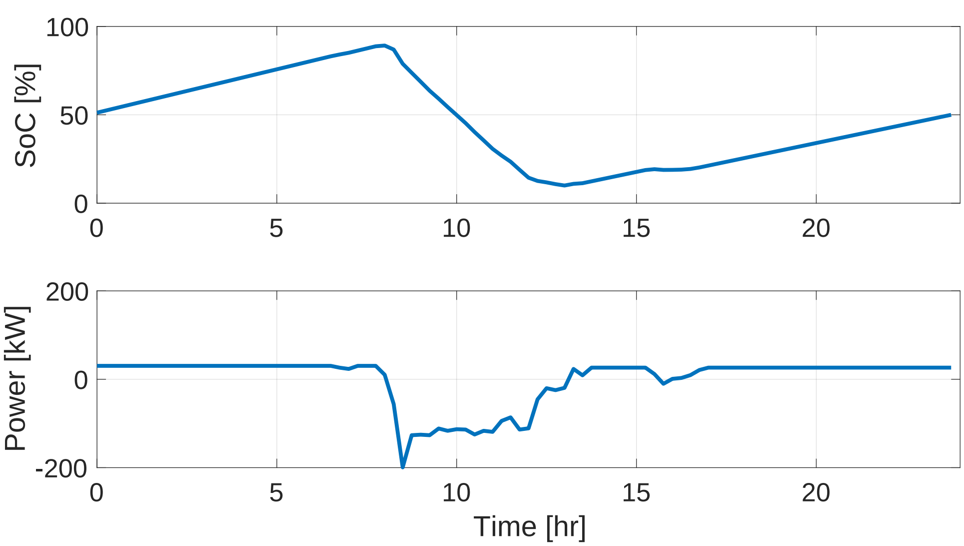

Figure 3.9: Real-time dispatch optimization results for a dynamic dispatch based on a real-time predicted profile for case 1. The figure on the top shows the real-time dynamic battery state of charge (SoC) for the 24-hour period and the figure on the bottom shows the real-time dynamic power schedule of the battery for the 24-hour period to obtain optimized demand as shown in Fig. 3.7.

result executed at $n = 48$ for the remaining period considering electricity costs and constraints. Fig. 3.9 show the state of charge (SoC) and the battery output power in $\%$ at time t_n for the dynamic dispatch updated at $n = 49$.

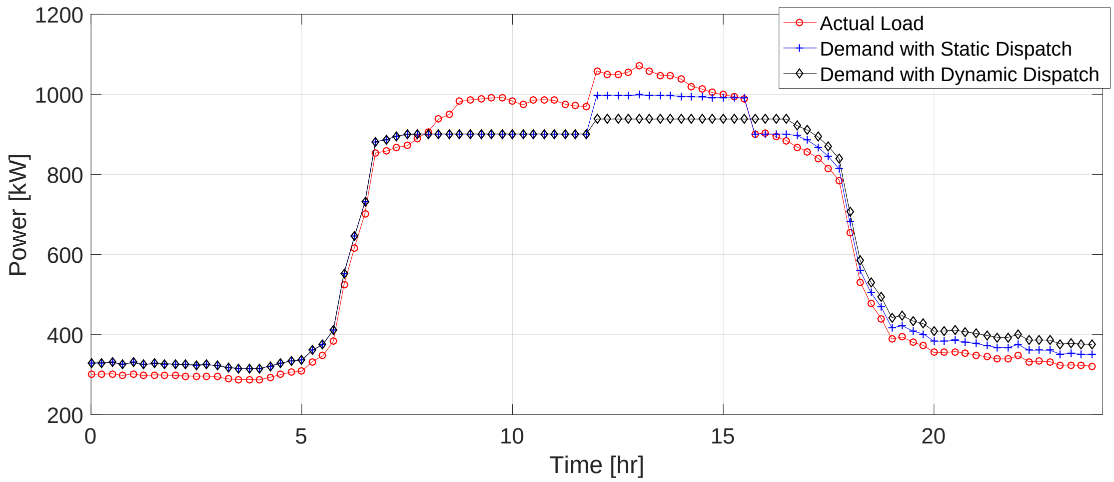

Figure 3.10: Real-time dispatch optimization results comparing the static and dynamic prediction and dispatch for case 2. The red curve shows the actual measured demand at the POI and blue curve shows the estimated demand at POI with the generated dispatch strategy with static predictions. However, the black curve shows the actual demand of POI with real-time predictions and optimization. The optimization is executed every 15-mins for the 24 hour period taking into consideration the electricity costs and constraints.

In this case, it can be seen that the updated prediction is higher from the day ahead prediction from $n = 49$ to $n = 63$ meaning that the demand of the microgrid is higher during this period than usually anticipated due to some unforeseen event or condition that was not predicted without the synchrophasor update. The optimized power was peak shaved at the same level as in Fig. 3.6 until $n = 48$ but since the demand is higher now for the next few intervals, the demand level cannot be maintained due to limited energy stored in the battery. The optimization hence results in an updated dispatch power of the battery to be able to still minimize the electricity costs with the new prediction data and constraints with a higher level of optimized demand. In lieu of the updated optimization, it could be clearly seen that the demand

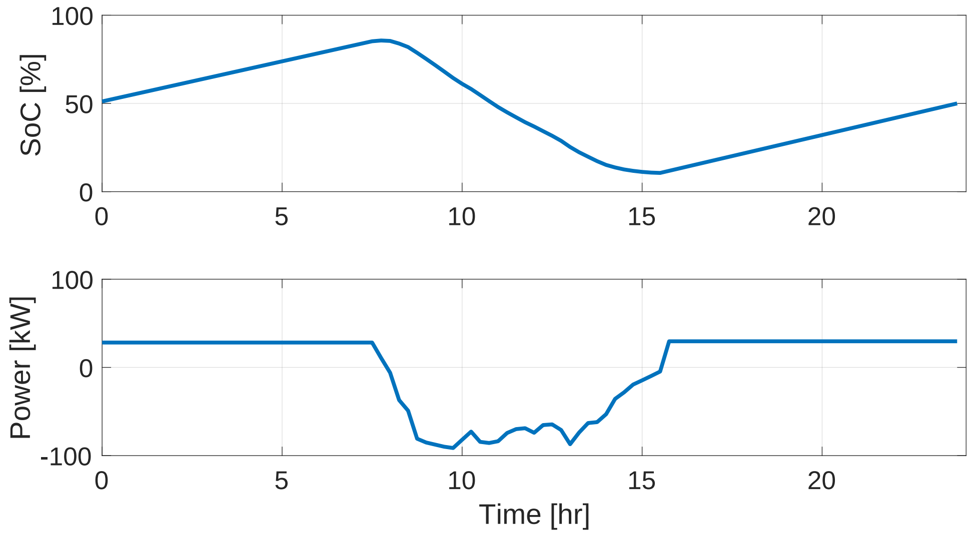

Figure 3.11: Real-time dispatch optimization results for a dynamic dispatch based on a real-time predicted profile for case 2. The figure on the top shows the real-time dynamic battery state of charge (SoC) for the 24-hour period and the figure on the bottom shows the real-time dynamic power schedule of the battery for the 24-hour period to obtain optimized demand as shown in Fig.3.10.

shaving at a much higher level and hence resulting in undesired levels of demand charges for the intervals of high load in this period. Fig. 3.11 shows the state of charge and battery (output) power at each interval of the dispatch period to meet the objectives and constraints with prediction update and dynamic dispatch.

3.8 Summary of Contribution

In this chapter, it is shown how real-time synchrophasor data can be used for recursive prediction of dynamic power demand of a system and how these predictions are used for real-time demand optimization for economic benefits. A static day-ahead baseline load profile is first created using a principal component analysis and clustering algorithm for power load consumption. A dynamic prediction method is then formulated using a recursive Kalman filter that uses phasor data to provide real-time updated load forecasts. The power dispatch algorithm aims to achieve maximal economic benefits mainly by demand reduction and energy arbitrage subject to the resource constraints and operational limits of a load. The final part of the chapter affirms the credibility of the combination of real-time prediction and optimization algorithms on real load data collected from a site for two different test cases. The first case shows the comparison of static vs dynamic dispatch when the predicted load has disturbances and variations from the actual load throughout the day and the optimization is being executed at the beginning of the day. The second case shows the comparison of static vs dynamic dispatch when the predicted load has an offset from the actual load for a given period in the day and the optimization is being executed at the when the demand offset begins.

The text and data in Chapter 3, in full, is a reprint of the material as it appears in "Synchrophasor Based Recursive Power Prediction with Real-Time Demand Optimization", Sai Akhil R. Konakalla, Raymond A. de Callafon, *to appear in IEEE Transactions on Smart Grids* and "Recursive power demand prediction based on multi-level clustering of power demand data", Sai Akhil R. Konakalla, Raymond A. de Callafon, *Smart Grid Communications (SmartGridComm)*, 2017 IEEE International Conference on. IEEE, 2017. The dissertation author is the primary investigator and author of these articles.

Chapter 4

Synchrophasor based Microgrid Dynamic Modeling and Control

4.1 Introduction

Proliferation of Phasor Measurement Units (PMUs) [63] in both transmission and distribution grids has enabled high speed critical grid monitoring and control. This is important for the microgrid control for two strong reasons: availability of higher sampling rate data (60 Hz or higher) that can be used for high-speed power disturbance rejection and time synchronized nature of the phasors (voltage and current) to control multiple microgrids [64]. In order to design high speed control algorithms for such microgrid systems, derivation of a system model via system identification using input/output system data would be more convenient than derivation via swing equations [10]. Hence, the estimation of a low order linear time invariant (LTI) model for active and reactive power flow (derived from the PMUs) through the microgrid becomes an important control pre-requisite in both grid connected and islanded modes. Given that there are several power generating and consuming elements in the microgrid, each with their own dynamics, the ability to observe the power flow at the main point of common coupling in the microgrid using a PMU forms the basis of modeling in the microgrid from a control stand-point. However, not always, step-excitation on the system is possible. Hence a multivariate identification method called covariance based realization algorithm (CoBRA) for a microgrid is used to estimate a low order LTI model for active and reactive power flow (derived from the PMUs) through the microgrid during operating conditions in the presence of load fluctuations, generation variations and power disturbances in both grid connected and islanded mode. Given that there are several power generating and consuming elements in the microgrid, each with their own dynamics, the ability to observe the power flow at the main point of interconnection in the microgrid using a PMU forms the basis of modeling the microgrid from a control stand-point. The ability of CoBRA to perform system identification in presence of multiple (correlated) power system inputs, disturbances and outputs in operating conditions given that the inputs are persistently excited makes it viable for identifying the dynamics of the microgrid.

This chapter also presents a dynamic dispatch multi-objective optimization that computes an optimal solution for the power dispatch of a hybrid collection of DERs. The solution is computed by minimizing an economically formulated objective function, subjected to constraints to satisfy a given power demand and taking into account DER limitations [65] such as (power) ramp rate limits, amplitude saturation, energy levels, battery efficiency and number of battery cycles [66], [67]. As such, the load following dispatch scheme [68], i.e. a dispatch strategy which allows available DERs to produce power only to meet the net demand, is addressed in this chapter. In addition, the optimization algorithm presented here may also be extended to cycle charging dispatch strategy [68], where excess power from DERs is used to charge the batteries. The algorithm for Optimal Hybrid Power Dispatch (OHPD) in this chapter uses a net power demand signal that obeys limitations on power ramp rate and power amplitude and then dispatches the power signals to each individual DER based on dynamic and auxiliary variables such as battery SoC and fossil fuel cost. The advantages of assuming a rate limited power demand are two-fold: a) It always facilitates in computation of an optimal control signal as per dynamic (ramp and amplitude) rates thus maintaining grid stability [69], [70], b) It guarantees that the net demand can always be met with the available DER power. The algorithm also outputs the dynamic bounds of the demand signal for guaranteed dispatch to be fed back into a possible real-time microgrid controller.

An additional refinement of the OHPD algorithm is included in this chapter that allows the use of a BESS with multiple batteries and inverters (denoted as sub-units) connected in parallel for simultaneous operation [71], [72]. Equal power sharing between the sub-units with possibly different power ramp rate and amplitude constraints might not be optimal, as it may result in battery draining or overcharging in case of mismatched SoC levels between the batteries. It is important to balance the SoC levels for full-load operation at all times [72], [73] and included as a constrained convex optimization problem in the OHPD algorithm. Linear quadratic regulation (LQR) [74] concepts with a penalty on the sub-unit dispatch signals are used to establish a trade off between the objective function that includes SoC balancing and the need for (real) power dispatch tracking.

In the last part of the chapter, a detailed switching and control algorithm is proposed for a microgrid consisting of multiple distributed energy resources including both synchronous generators and inverter-based DERs. The controller enables fast and smooth intentional switching between islanding and grid connected modes with small voltage and frequency transient fluctuations. The controller is designed based on the estimated models. The data-based controller design methodology can be easily extended to accommodate other combinations of DERs in the microgrid. The developed control algorithm is tested on a HIL setup that models the microgrid and proves fast and smooth transition.

4.2 Physical Grid Component Modeling

In this chapter, a microgrid that can be either connected or disconnected (islanded) from the main grid and consisting of a hybrid collection of DERs interconnected with loads as illustrated in Fig. 4.9 is considered. The DERs include: (a) a source switchable battery/inverter system which can operate either in grid following (current source) or grid forming (voltage source) mode, (b) Emergency synchronous diesel generator to meet the loads when disconnected from the main grid, (c) Photovoltaic (PV) or solar inverter that outputs AC power from DC solar power. The loads include day-to-day residential or industrial appliances that are modeled as resistive and inductive (R-L) elements . The simulation is performed for the microgrid in the grid connected and islanded cases separately to estimate the dynamics in each case. For the purpose of demonstration of system identification, first in the grid connected mode where the microgrid is connected to the main utility grid via the point of interconnection (POI), the real and reactive power desired from the (battery) inverter, the solar PV inverter and the real and reactive power consumed by the loads are considered as inputs to the microgrid. The real and reactive power flow at the POI is considered as an output of the microgrid. It is important to note that in grid connected case, except for the desired power set-points to the battery/inverter, all the other input/output power flow data used for identification are obtained from the PMU measurements at 60 samples per second. Hence the grid dynamics obtained are from the input of the battery/inverter, measured load power and measured solar power to the

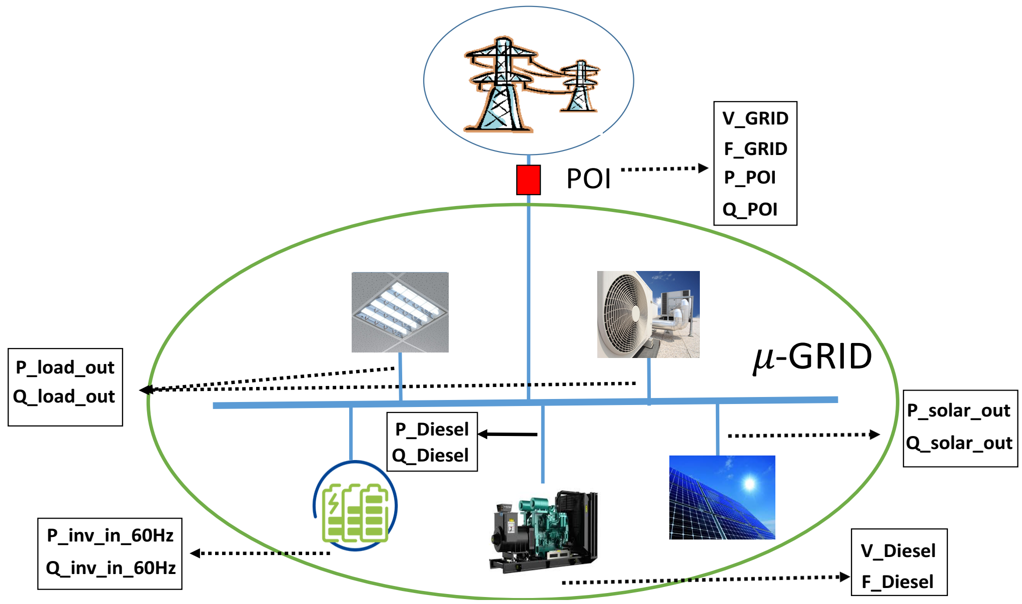

Figure 4.1: Overview of the microgrid system for simulation.

measured power at the POI.

Similarly, in the islanded mode where the microgrid is disconnected from the main utility grid via the POI, the real and reactive power desired from the (battery) inverter, the solar PV inverter and the real and reactive power consumed by the loads are considered as inputs to the microgrid. Additionally, the bus voltage and frequency desired to maintain by the diesel generator is also considered as an input to the microgrid. The real and reactive power flow to/from the diesel generator and the voltage (V) and frequency (F) of the microgrid are considered as outputs of the microgrid. In islanded case, except for the desired power set-points to the battery/inverter and V-F set-points to the diesel generator, all the other input/output power flow data used for identification are obtained from the PMU measurements at 60 samples per second. Hence the grid dynamics obtained are from the input of the battery/inverter, input of the diesel generator, measured load power and measured solar power to the measured power flow at the diesel generator and the measured V-F of the microgrid.

A collection of hybrid distributed energy resources are modeled using two different approaches. In the first approach, Simscape Power Systems™, a toolbox to design control systems for electrical power system in MATLAB Simulink is used as a modeling platform. Simscape simulator allows for detailed modeling of different system components present in the microgrid; allowing the user to run identification or control experiments with the simulated model in order to validate the operation of microgrid under realistic conditions. DER models are built on top of the available libraries and local primary control loops are designed and tuned for different DERs.

As seen in Fig. 4.9, the system to be identified consists of a battery inverter, a solar inverter, a synchronous diesel generator and loads operating inside their operational limits. These have highly non-linear behavior in their response for a given set-point input. In grid connected mode, the model dynamics are estimated between solar (P_solar_out, Q_solar_out) and load (P_load_out, Q_load_out) power flow along with the controllable inverter ($P\_inv\_in\_60Hz$, $Q\_inv\_in\_60Hz$) set-point as system inputs and the (real and reactive) power flow at POI (P_POI, Q_POI) as the system outputs. Similarly, in islanded mode, the model dynamics are estimated between solar (P_solar_out, Q_solar_out) and load (P_load_out, Q_load_out) power flow along with the controllable diesel generator (V_Diesel, F_Diesel) set-point as system inputs and the power flow at diesel generator (P_Diesel, Q_Diesel), voltage and frequency (V_GRID, F_GRID) of the microgrid as the system outputs as shown in Fig. 4.9. These set-point values and power flow measurements, which form the input/output data for system identification, are nonlinearly related due to either the non-linearity in the physical system such as batteries or non-linearities due to local control loops in these DERs such as voltage, frequency droop controllers in diesel generators or maximum power point tracking of solar PV. Since we consider the case where the microgrid system is operating around a nominal point, it can be approximated as a piece-wise linear model with a high order sufficient to approximate such non-linearities in order to use the model for designing the control algorithm.

Although the first approach is helpful for modeling and control design with a given simulation time, it does not have the capability to run in real-time. The hardware-in-the-loop (HIL) simulation is a technique used to develop and test complex, real-time embedded systems and includes the electrical emulation of sensors and actuators. In the second physical modeling approach, a collection of hybrid distributed energy resources are modeled using the HIL environment of Real Time Digital Simulator (RTDS) [75]. The RTDS® simulator allows for detailed modeling of the power and control system components present in microgrids. The RTDS simulator operates in real time, allowing the user to interface physical equipment, in our case microgrid controller, with the simulated model in order to test and validate the operation of microgrid protection and control devices under realistic conditions. RTDS hardware is digital signal processor based and utilizes advanced parallel processing techniques. RTDS software includes accurate power system component models required to represent many of the complex elements which make up physical power and control systems. DER models are built on the top of the available RSCAD [75] libraries and the individual local control loops are also designed in the RTDS.

In this chapter, a microgrid that can be either connected or disconnected (islanded) from the main grid and consisting of a hybrid collection of DERs interconnected with loads as illustrated in Fig. 4.9 is considered. The DERs include: (a) a source switchable battery/inverter system which can operate either in grid following (current source) or grid forming (voltage source) mode [76], (b) Emergency synchronous diesel generator to meet the loads when disconnected from the main grid [77], (c) Photovoltaic (PV) or solar inverter that outputs AC power from DC solar power [78]. The loads include day-to-day residential or industrial appliances that are modeled as

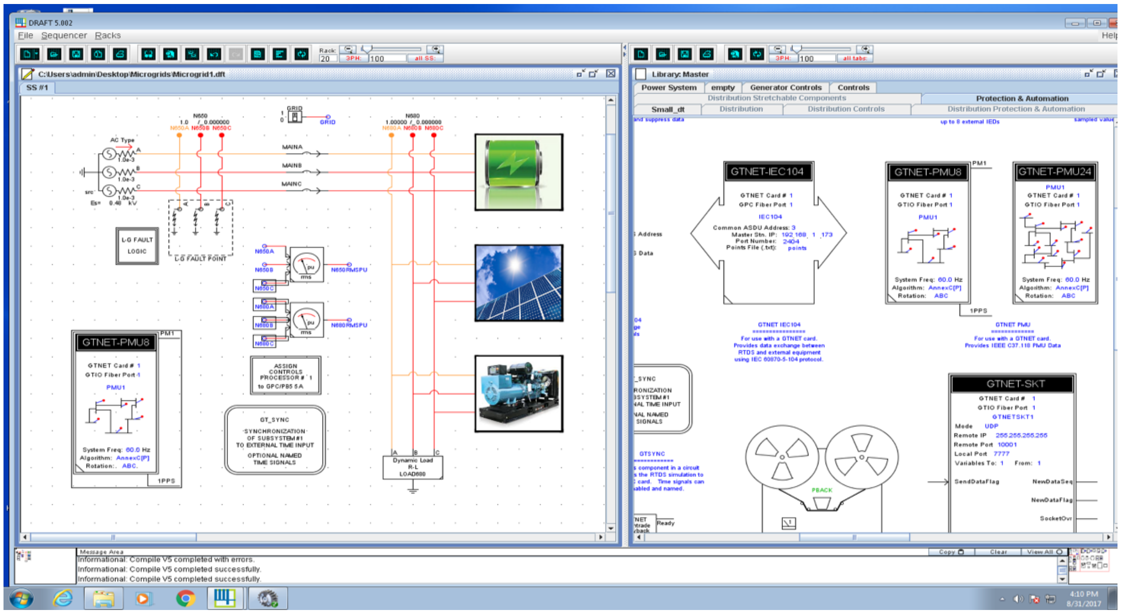

Figure 4.2: Dynamic physical modeling of the microgrid in RTDS using RSCAD libraries.

resistive and inductive (R-L) elements. It is important to understand the model of the microgrid inverter to be able to control the power flow, voltage and frequency in different modes of operation. The RTDS model of the microgrid is illustrated in Fig. 4.2.

Various control strategies for inverter-based DERs have been widely studied in the literature. [79–81] provide a review of the most common methodologies. We consider two kinds of control strategies used to operate an inverter with power-electronic interfaces [82]. The inverter model is derived for these two following control strategies. First, PQ inverter control or current source inverter (CSI) where the inverter is used to supply a given amount of current from a given real and reactive power set point at a given (measured) voltage, and second, voltage source inverter (VSI) control where the inverter is controlled to feed the load with predefined values for voltage and frequency. For the problem under consideration, in the grid connected mode, the inverters are set to operate in the current source mode. In the islanded operation, one of the inverters will be operating as a voltage source that regulates voltage magnitude and frequency of the microgrid and other DERs will follow the voltage and frequency set by that inverter.

The VSI control in the inverter closely mimics the governor and automatic voltage regulator (AVR) of a synchronous generator by measuring active and reactive power to droop voltage and frequency. In a conventional power system, synchronous generators will share any increase in the load by decreasing the frequency (rotor speed) according to their governor droop characteristic [83]. This principle is implemented in inverters by decreasing the reference frequency when there is an increase in the load. Similarly, reactive power is shared by introducing a droop characteristic in the voltage magnitude. Instantaneous real and reactive power components $\tilde{p}, \tilde{q}$ and are calculated from the measured output voltage and output current as in

$$\begin{aligned} \tilde{p} &= v_{od} i_{od} + v_{oq} i_{oq} \\ \tilde{q} &= v_{od} i_{od} - v_{oq} i_{oq} \end{aligned} \tag{4.1}$$

where v_{od}, i_{od} and v_{oq}, i_{oq} are the direct and quadrature three-phase voltage and current vectors in the d-q frame respectively. The instantaneous power components are passed through low-pass filters, shown in Eq. (4.2), to obtain the real and reactive powers P and Q and corresponding to the fundamental component. ω_c represents the cut-off frequency of low-pass filters.

$$P = \frac{\omega_c}{s + \omega_c} \tilde{p},\ Q = \frac{\omega_c}{s + \omega_c} \tilde{q}, \tag{4.2}$$

The frequency f_{inv} is set according to the droop gain (k_{dp}) and phase is set by integrating the frequency. This mimics the governor and inertia characteristics of conventional (diesel) generators and provides a degree of negative feedback. For instance, if the power drawn from a generator increases then the rotation of its voltage angle slows and f_{inv} lowers. In the following equations f_0 will represent the nominal frequency set-point whereas δ is the angle of the inverter frame seen from a reference frame. From Eq.(4.3) it can be seen that the angle of the inverter voltage, δ, changes in response to the real power flow in the required negative sense and with a gain set by the droop

$$\begin{aligned} f_{inv} &= f_0 - k_{dp}P \\ \dot{\delta} &= f_{inv},\ \ \delta = f_0 t - \int k_{dp} P dt \end{aligned} \tag{4.3}$$

For the reactive power a droop is introduced in the voltage magnitude as given in Eq. (4.4). Here, V_0 stands for the nominal set-point of d-axis output voltage. With k_{dq} as the reactive droop gain, the control strategy is chosen such that the output voltage magnitude reference V_{inv} is aligned to the d-axis of the inverter reference frame, and the q-axis reference is set to zero

$$V_{inv} = v_{od}^* = V_0 - k_{dq}Q,\ \ v_{oq}^* = 0 \tag{4.4}$$

Hence voltage angle of the inverter δ can be altered by controlling f_0 and the voltage magnitude V_{inv} by controlling V_0 for VSI.

In CSI control mode, the AC current output from the inverter is calculated based on the given P, Q set-points and the feedback measurement of the (bus) voltage. Hence $\tilde{p}$, $\tilde{q}$, v_{od}, v_{oq} are known and i_{od}, i_{oq} are calculated from Eq. (4.1).

4.3 Estimation of Low Order Grid Dynamics with Step Excitation Inputs

With several power generating and consuming assets in the microgrid, the ability to regulate and stabilize power flow at the PCC of the microgrid requires understanding of the dynamics of power flow of each asset. The approach to modeling asset dynamics is driven by observing and modeling power flow with synchrophasor data instead of a model derivation via power system swing equations. After all, it is the same synchrophasor data that will be used for real-time feedback to automate the regulation and stabilization of power flow at the PCC of the microgrid.

4.3.1 Realization Algorithm for Model Estimation

For modeling and control design, it is assumed that dynamics of assets in the microgrid under both operating conditions can be modeled using a minimal LTI discrete-time system $\mathbf{y}(k) = C\mathbf{x}(k) + \mathbf{v}(k)$ with $\mathbf{x}(k+1) = A\mathbf{x}(k) + B\mathbf{u}(k)$, where the input $\mathbf{u}$ is a vector denoting the real and reactive power dispatch commands to the p different controllable assets in the microgrid. The output $\mathbf{y}$ is the observed real/reactive power flow at the PCC, possibly perturbed by load, generation and measurement perturbations $\mathbf{v}$ at the time index k. When the microgrid is operating around an equilibrium, the system can be approximated by a linear time invariant (LTI) model, in which the matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times 2p}, C \in \mathbb{R}^{2\times n}, D \in \mathbb{R}^{2\times 2p}$ and the order n needs to be determined from the synchronized input/output data $\mathbf{u}$ and $\mathbf{y}$.

Leveraging the ability to provide step-wise excitation signals to the p assets in the microgrid, the step-response realization algorithm of [84] is used to determine an estimate of the order n using a Singular Value Decomposition (SVD) and estimates of matrices $\{A, B, C, D\}$ using an orthogonal projection. More details on the estimation procedure can be found in [84], but it suffices to mention here that each asset is dispatched with a (unit) step-wise change in real and reactive power stored in the input vector $\mathbf{u}$, whereas the resulting change in the relevant synchrophasor data (real/reactive power or frequency and voltage) measured by a PMU at the PCC is stored in the output vector $\mathbf{y}$. The subsequent estimation of the model order n and the state matrices A, B, C and D will not only reveal the linear dynamics of power flow from the asset to the PCC, but also the dynamic and static coupling between real/reactive power flow of the asset to the real/reactive power flow at the PCC.

For the estimation procedure, the relevant synchrophasor data $\mathbf{y}$ produced by the PMU at the PCC over the time index $k = 1, \ldots, 2N-1$ is stored in a (weighted) Hankel matrix

$$
\begin{aligned}
R(1, 2N-1) = &\begin{bmatrix} \mathbf{y}(1) & \mathbf{y}(2) & \cdots & \mathbf{y}(N) \\ \mathbf{y}(2) & \mathbf{y}(3) & \cdots & \mathbf{y}(N+1) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}(N) & \mathbf{y}(N+1) & \cdots & \mathbf{y}(2N-1) \end{bmatrix} \\
&- \begin{bmatrix} \mathbf{y}(0) & \mathbf{y}(0) & \cdots & \mathbf{y}(0) \\ \mathbf{y}(1) & \mathbf{y}(1) & \cdots & \mathbf{y}(1) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}(N-1) & \mathbf{y}(N-1) & \cdots & \mathbf{y}(N-1) \end{bmatrix}
\end{aligned}
\tag{4.5}
$$

where it can be shown that

$$
R(1, 2N-1) = H(1, 2N-1) \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{4.6}
$$

where $H(1, 2N-1)$ is a Hankel matrix

$$H(1, 2N-1) = \begin{bmatrix} \mathbf{m}(1) & \mathbf{m}(2) & \cdots & \mathbf{m}(N) \\ \mathbf{m}(2) & \mathbf{m}(3) & \cdots & \mathbf{m}(N+1) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{m}(N) & \mathbf{m}(N+1) & \cdots & \mathbf{m}(2N-1) \end{bmatrix}$$

containing the Markov parameters or impulse response coefficients $\mathbf{m}(k) = CA^{k-1}B$ for which it is known that rank$\{H(1, 2N-1)\} = n$, provided $2N - 1 \geq n$ [84]. Due to the rank preservation of the (weighted) Hankel matrix $R(1, 2N-1)$ in (4.6), we see that rank$\{R(1, 2N-1)\} = n$ and a standard SVD of $R(1, 2N-1)$ given by

$$R(1, 2N-1) = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U_1 \Sigma_1 V_1^T$$

with $\Sigma_1 \in \mathbb{R}^{n \times n} > 0$ can be used to find a rank n decomposition of $R(1, 2N-1) = R_1 R_2$ with $R_1 = U_1 \Sigma_1^{1/2}$ and $R_2 = \Sigma_1^{1/2} V_1^T$.

Similar to the definition of $R(1, 2N-1)$ in (4.5), the matrix $R(2, 2N)$ can be created and it is easy to show that $R(2, 2N) = R_1 A R_2$. This allows the state matrices to be extracted and estimated from the SVD via the projection $\hat{A} = \Sigma_1^{-1/2} U_1^T R(2, 2N) V_1 \Sigma_1^{-1/2}$, $\hat{B} = R_2(:, 1 : 2)$ and $\hat{C} = R_1(1 : 2, :)$, where $R_2(:, 1 : 2) \in \mathbb{R}^{n \times 2}$ denotes the first 2 columns in R_2 and $R_1(1 : 2, :) \in \mathbb{R}^{2 \times n}$ denotes the first 2 rows of R_1.

The end result of the step-response realization algorithm is a dynamic model of the form

$$G : \begin{cases} \hat{\mathbf{x}}(k+1) = \hat{A}\hat{\mathbf{x}}(k) + \hat{B}\mathbf{u}(k) \\ \hat{\mathbf{y}}(k) = \hat{C}\hat{\mathbf{x}}(k) \end{cases} \tag{4.7}$$

for which both a simulation of the relevant synchrophasor data $\hat{\mathbf{y}}(k)$ produced by the PMU at the PCC and a frequency response $G(e^{j\omega}) = \hat{C}(e^{j\omega}I - \hat{A})^{-1}\hat{B}$ can be created. Any static or DC coupling between real and reactive power is characterized by $\hat{C}(I - \hat{A})^{-1}\hat{B}$ at $\omega = 0$. We will illustrate the estimation of the dynamic model and the possible dynamic and static coupling between real and reactive power for different assets in a microgrid in the next section.

4.3.2 Synchrophasor based Dynamic System Modeling

In grid connected mode, an input (real/reactive) power step dispatch is applied independently to the CSI and the diesel generator in PQ mode available in the microgrid. The dynamics from the synchronized dispatch to the power flow measured by the synchrophasor data at the PCC are estimated for both the controllable CSI and the diesel generator in PQ mode. In case of islanding mode, power flow dynamics to the PCC is not applicable. Instead, only the dynamics between the frequency set point $f(k)$ of the controllable VSI (or diesel generator in voltage-frequency droop mode) and the voltage angle $\theta_{v_A}(k)$ at the PCC is relevant and simply given by the discrete-time integrator model

$$\theta_{v_A}(k) = 2 \cdot 180 \cdot T_s(f(k) - 60) + \theta_{v_A}(k-1)$$

for a 60Hz AC grid, where $T_s = 1/60$ is the sampling time of the PMU data. For estimation of synchronized power flow dynamics, PMU measurements are also sampled at 60 samples per second. For the step-wise excitation of the assets, ± 40kW step excitation are used for real power and ± 10kVAR for reactive power for each of the assets to limit voltage perturbations on the main bus of the microgrid.

A comparison of the excellent agreement between the synchronized measurements of real/reactive power flow $\mathbf{y}$ and the modeled or simulated real/reactive power flow $\hat{\mathbf{y}}$ at the PCC due to the LTI model in (4.7) is given in Figure 4.3. The modeling results in Figure 4.3 have been captured by a $n = 7$ order multivariable model and illustrate the fast dynamics and short settling time of the CSI, while the dynamics of the diesel generator in PQ mode is dominated by an additional time delay of 15 samples (0.25 sec) and a much longer settling time.

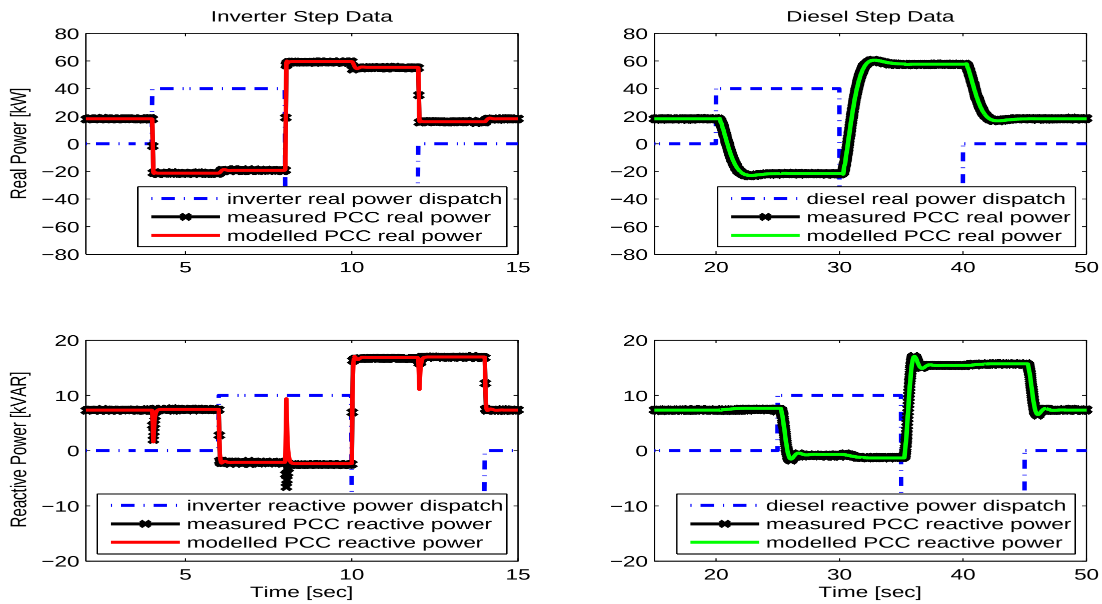

Figure 4.3: Time-domain data of the step-wise real/reactive power dispatch $\mathbf{u}(k)$ for the Current Source Inverter (CSI) and the diesel generator. Measurements of PMU-based power flow $\mathbf{y}(k)$ at the PCC due to step-wise power dispatch compared with simulated power flow $\hat{\mathbf{y}}(k)$ at the PCC using a $n = 7$th order multivariable LTI model G in (4.7) with additional output time delays.

In the time-domain plots of Fig. 4.3 also the dynamic and static coupling between real and reactive power flow at the PCC from the CSI can be observed. Such static and dynamic coupling is more clearly visible from the amplitude Bode response of the multivariable LTI model depicted in Figure 4.4. It can be observed that the diagonal elements provide a DC-gain of approximately 1, but DC coupling of approximately -7dB exists between real and reactive power at the PCC for the CSI.

Additional dynamic coupling between real/reactive power flow can be observed in the off-diagonal elements.

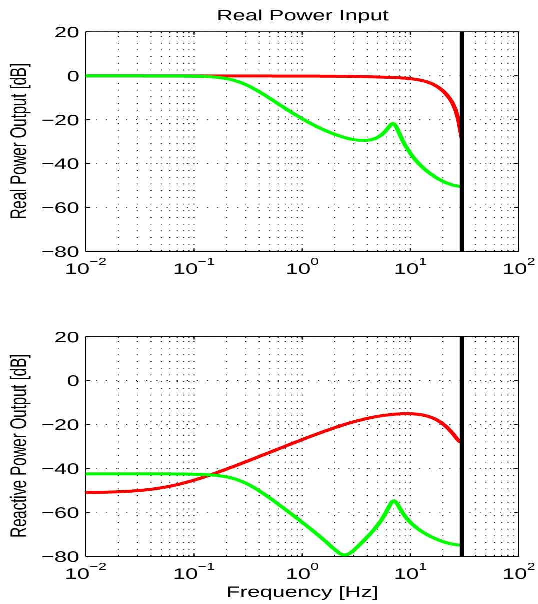

Figure 4.4: Amplitude Bode response of the $n = 7$th order multivariable LTI model $G(e^{j\omega})$ in (4.7) of PCC power flow for the Current Source Inverter (CSI) and the diesel generator.

4.4 Low Order Modeling without Step Excitation Inputs

Applying step-excitation inputs on a power system during normal operation is a non-trivial task as it effects the system operation. Hence, system identification using conventional identification techniques that model dynamics using step or impulse excitation on inputs becomes limited. Hence, the main motivation in using such a multivariate CoBRA for a microgrid is to estimate a low order LTI model for active and reactive power flow (derived from the PMUs) through the microgrid during operating conditions in the presence of load fluctuations, generation variations and power disturbances in both grid connected and islanded mode. Given that there are several power generating and consuming elements in the microgrid, each with their own dynamics, the ability to observe the power flow at the main point of interconnection in the microgrid using a PMU forms the basis of modeling the microgrid from a control stand-point. The ability of CoBRA to perform system identification in presence of multiple (correlated) power system inputs, disturbances and outputs in operating conditions given that the inputs are persistently excited makes it viable for identifying the dynamics of the microgrid.

4.4.1 Problem Formulation

Discrete-time signals are considered in this chapter and the mean of all signals can be removed beforehand if we only focus on the dynamics. In order to simultaneously handle both deterministic and stochastic signals, we denote the (cross) covariance function of two quasi-stationary vector signals $\mathbf{s}(t)$ and $\mathbf{w}(t)$ as follows:

$$\begin{aligned} R_{sw}(\tau) =& \bar{E}\left\{\mathbf{s}(t+\tau)\,\mathbf{w}(t)^{\mathrm{T}}\right\} \\ =& \lim_{N\to\infty}\frac{1}{N}\sum_{t=0}^{N-1} E\left\{\mathbf{s}(t+\tau)\,\mathbf{w}(t)^{\mathrm{T}}\right\}, \end{aligned} \tag{4.8}$$

whose sample counterpart can be written as below:

$$\hat{R}_{sw}(\tau) = \frac{1}{N}\sum_{t=0}^{N-1}\mathbf{s}(t+\tau)\,\mathbf{w}(t)^{\mathrm{T}}, \tag{4.9}$$

where N is the length of available data and the value of the signal at the extended time instant (beyond $[0, N-1]$) is set to be zero. Typically, a finite record of the values of sample covariance function $\hat{R}_{sw}(\tau)$ is sufficient for parameter estimation. Thus, we may produce the data just for $|\tau| \leq \bar{\tau} \ll N$ with $\bar{\tau}$ being some finite constant.

In practice, when the magnitudes of the input and output signals are not relatively large and the system is operating around an equilibrium, the system can be well described by an LTI model. In this chapter, we consider a minimal LTI discrete-time system with unknown order n, which can be described as follows:

$$\begin{cases} \mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t) + \mathbf{v}(t), \end{cases} \tag{4.10}$$

where $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times p}, C \in \mathbb{R}^{m\times n}, D \in \mathbb{R}^{m\times p}$ are the unknown matrices and $\mathbf{v}(t)$ is the quasi-stationary (probably colored) Gaussian noise with zero mean and unknown spectrum. Although the model is simple, it's typically enough to provide acceptable fitting in modelling an actual system such as electromechanical system, etc. For simplicity, no input noise is considered, which is often the case for the open-loop identification.

The objective is to determine the model order n and estimate matrices $\{A, B, C, D\}$ in (4.10) up to a similarity transformation from the input and output signals $\{u(t), y(t)\}$.

4.4.2 Covariance Based Realization Algorithm

It can be observed from (4.10) that the output $y(t)$ is contaminated by the noise $\mathbf{v}(t)$. In many cases, we are more interested in the deterministic dynamics of the system than the stochastic part due to the noise. Thus, an instrumental variable $\xi(t)$ satisfying the following conditions can be introduced to eliminate the noise effect of $\mathbf{v}(t)$:

- $\xi(t) \in \mathbb{R}^{p\times 1}$ is correlated with $\mathbf{u}(t)$.

- $\xi(t) \in \mathbb{R}^{p\times 1}$ is uncorrelated with $\mathbf{v}(t)$.

A simple choice is letting $\xi(t) = \mathbf{u}(t)$. Note that the dimension of $\xi(t)$ can also be larger than $\mathbf{u}(t)$. In practice, only finite data points are available and thus we can come up with the following equations using sample expectation defined in (4.9):

$$\begin{cases} \hat{R}_{x\xi}(\tau+1) = A\hat{R}_{x\xi}(\tau) + B\hat{R}_{u\xi}(\tau), \\ \hat{R}_{y\xi}(\tau) = C\hat{R}_{x\xi}(\tau) + D\hat{R}_{u\xi}(\tau) + \hat{R}_{v\xi}(\tau), \end{cases} \tag{4.11}$$

where the initial condition becomes $\hat{R}_{x\xi}(-\bar{\tau}) = \hat{R}_{\bar{\tau}}$ for $|\tau| \leq \bar{\tau} \ll N$. In ideal case, $R_{v\xi}(\tau) = 0$ since $\mathbf{v}(t)$ is with zero mean and uncorrelated with $\xi(t)$. As the data length N is sufficiently large, the components of its sample counterpart $\hat{R}_{v\xi}(\tau)$ should also be around zero but with unknown variance. In practice, the noise part $\hat{R}_{v\xi}(\tau)$ is small enough to be ignored.

Then, the objective in Sec. 4.4 is transformed into the following identification problem: given the values of sample covariance functions $\{\hat{R}_{u\xi}(\tau), \hat{R}_{y\xi}(\tau)\}$ for $|\tau| \leq \bar{\tau} \ll N$, determine the model order n and estimate matrices $\{A, B, C, D\}$ in (4.11) up to a similarity transformation. This problem can be similarly solved using the deterministic (MOESP-like) subspace methods by constructing Hankel matrices, and this class of subspace methods using covariance values is referred to as CoBRA. Different frameworks of CoBRA can be specified for different purposes such as pole location constraints and optimal estimation.

4.4.3 Excitation and Power Flow Data

See table 4.1 for the input and output signals used for identification under the grid connected mode.

See table 4.2 for the input and output signals used for identification under the

Table 4.1: Notations under the grid connected mode.

Input	Description	Unit
P_inv_in_60Hz	Real Power Set-point (inverter)	kW
Q_inv_in_60Hz	Reactive Power Set-point (inverter)	kVAR
P_load_out	Real Power Measurement (load)	kW
Q_load_out	Reactive Power Measurement (load)	kVAR
P_solar_out	Real Power Measurement (solar)	kW
Q_solar_out	Reactive Power Measurement (solar)	kVAR
Output	Description	Unit
P_POI	Real Power Measurement (POI)	kW
Q_POI	Reactive Power Measurement (POI)	kVAR

islanded mode.

Table 4.2: Notations under the islanded mode.

Input	Description	Unit
V_Diesel	Voltage Set-point (diesel)	V
F_Diesel	Frequency Set-point (diesel)	Hz
P_inv_in_60Hz	Real Power Set-point (inverter)	kW
Q_inv_in_60Hz	Reactive Power Set-point (inverter)	kVAR
P_load_out	Real Power Measurement (load)	kW
Q_load_out	Reactive Power Measurement (load)	kVAR
P_solar_out	Real Power Measurement (solar)	kW
Q_solar_out	Reactive Power Measurement (solar)	kVAR
Output	Description	Unit
P_Diesel	Real Power Measurement (diesel)	kW
Q_Diesel	Reactive Power Measurement (diesel)	kVAR
V_GRID	Voltage Measurement (microgrid bus)	V
F_GRID	Frequency Measurement (microgrid bus)	Hz

4.4.4 Application of CoBRA

With the aid of CoBRA, we can identify low order models for both grid connected and islanded mode using given input and output data. The data length is $N = 6000$ but only the segment after $t = 600$ is used. The instrumental variable $\xi(t)$ is simply chosen as the input. The Hankel matrix used for CoBRA is constructed to make sure $R_{u\xi}(0)$ exists in each block row. Also, the values of sample covariance functions $\{\hat{R}_{u\xi}(\tau), \hat{R}_{y\xi}(\tau)\}$ for $\tau < 0$ are also used for identification since they are also containing system information. For simplicity, the identification from arbitrary data segments is not considered here. Finally, a 2-nd order model for grid connected mode and 6-th order model for islanded mode can be obtained to give good approximation. The excitation signals are shown in Fig. 4.5 and Fig. 4.7. The comparison between the real (solid lines) and estimated (dashed lines) outputs is shown in Fig. 4.6 and Fig. 4.8. It can observed that the estimated low order models can capture all the microgrid power flow dynamics.

4.5 Microgrid Islanding Control

Seamless islanding and seamless grid reconnection are major requirements for microgrids that are designed to operate in both modes. Additionally, in certain islanding or reconnection cases, fast switching between islanding and grid connection is a strict requirement for the microgrid to survive under grid or microgrid faults. For instance, in the event of a grid fault and in order for the microgrid to continue to supply power locally, it should be able to quickly transition to islanded mode of operation. On the other hand, for an islanded microgrid where one DER trips off due to a fault, fast grid reconnection should take place to prevent power outage in the microgrid. In many microgrid islanding applications, the two requirements of smooth transition and fast transition are conflicting. While transitioning from islanding to grid connection has been widely studied in the literature [85] the inverse transition, i.e. from grid connected to islanded has been often overlooked. In this work, we propose controllers that ensure *fast* and *smooth* transitions between islanding and grid connected modes of microgrid operation upon detection of faults in either or due to other requirements. The controllers will be designed based on models obtained using phasor data in the previous section. Also, the control operation will be performed using high update rate feedback of phasor data from PMU measurements.

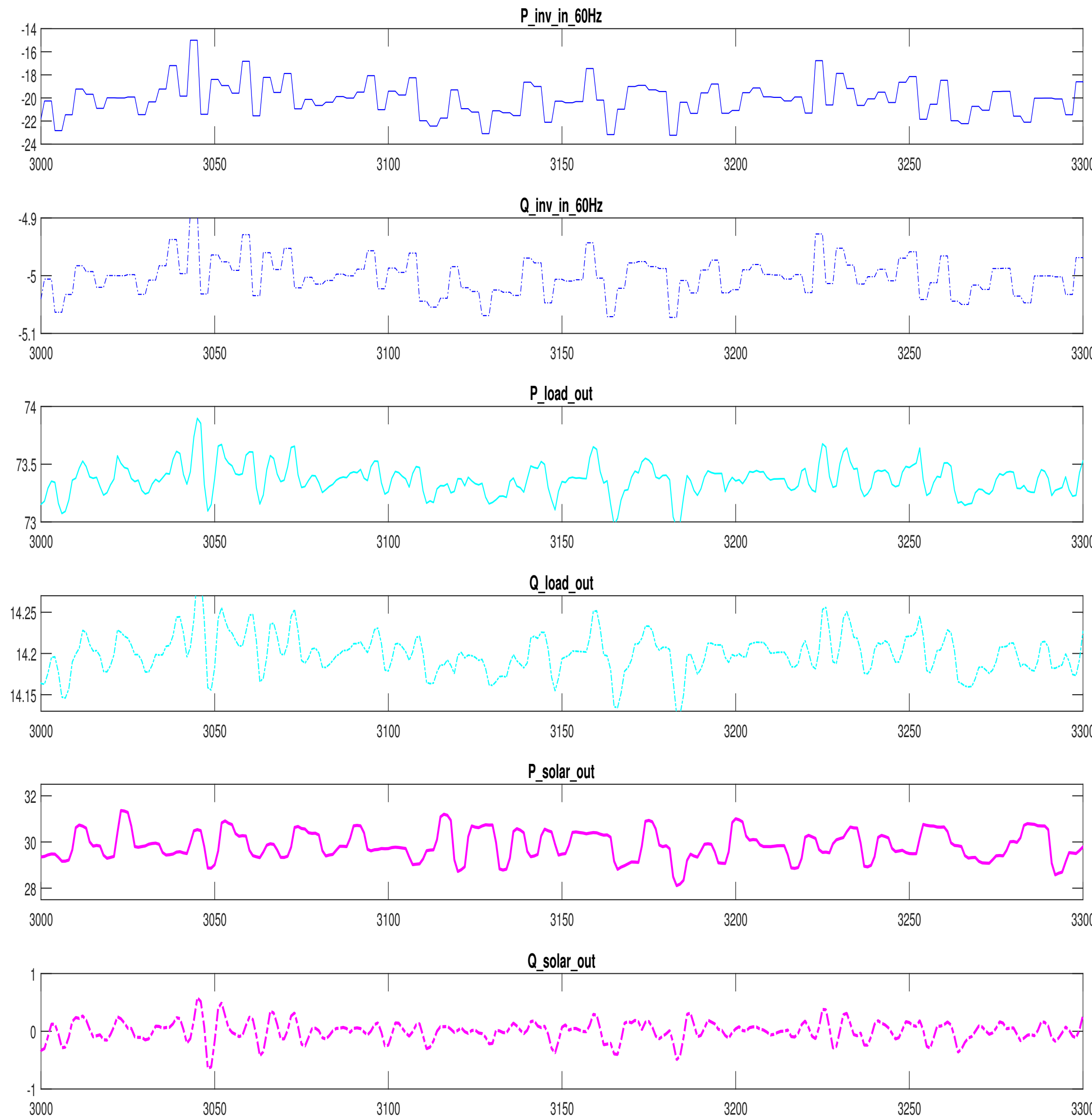

Figure 4.5: Input signals under the grid connected mode.

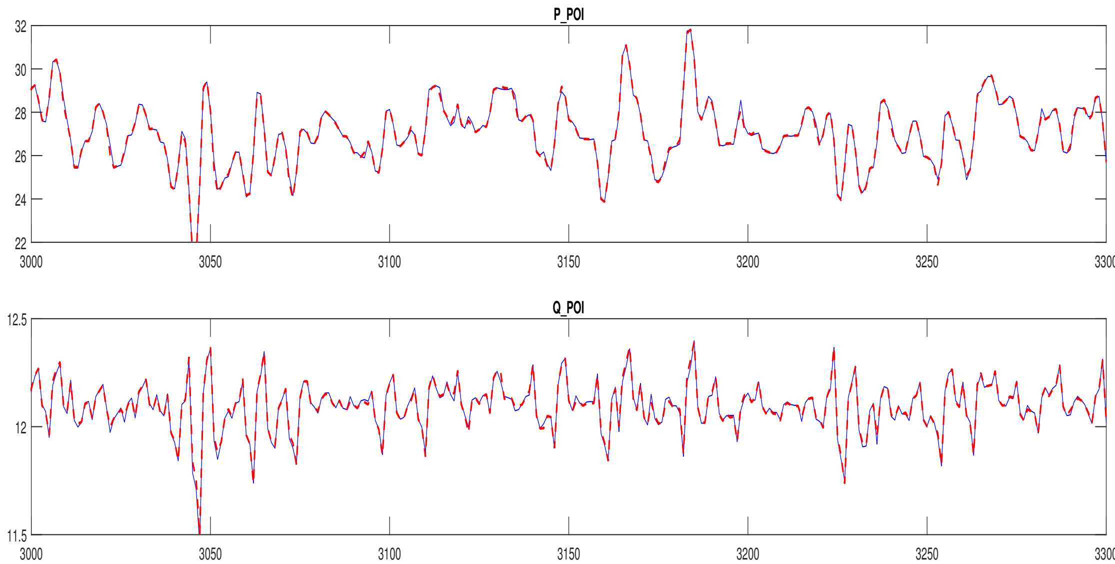

Figure 4.6: Output signals (solid lines) under the grid connected mode. The dashed lines represent the outputs of a 2-nd order model estimated by CoBRA.

4.5.1 Microgrid Reconnection using Synchrophasors

Fig. 4.9 shows a microgrid with multiple distributed energy resources and loads. The microgrid is connected to a main grid feeder via a switch or a circuit breaker (CB). The status of the switch or CB determines the operating state of the microgrid i.e. when open the microgrid is isolated from the utility grid or islanded, and when closed, the microgrid is grid connected. In islanded mode, due to the difference between supply and demand, the voltage and frequency fluctuate and are not synchronous with that of the main grid. At any given time t_k, the voltage across the interconnection switch (dropping time indexing) is given by

$$|V_{AB}| = \sqrt{V_A^2 + V_B^2 - 2\ V_A \cdot V_B\ cos(\theta_{v_A} - \theta_{v_B})} \tag{4.12}$$

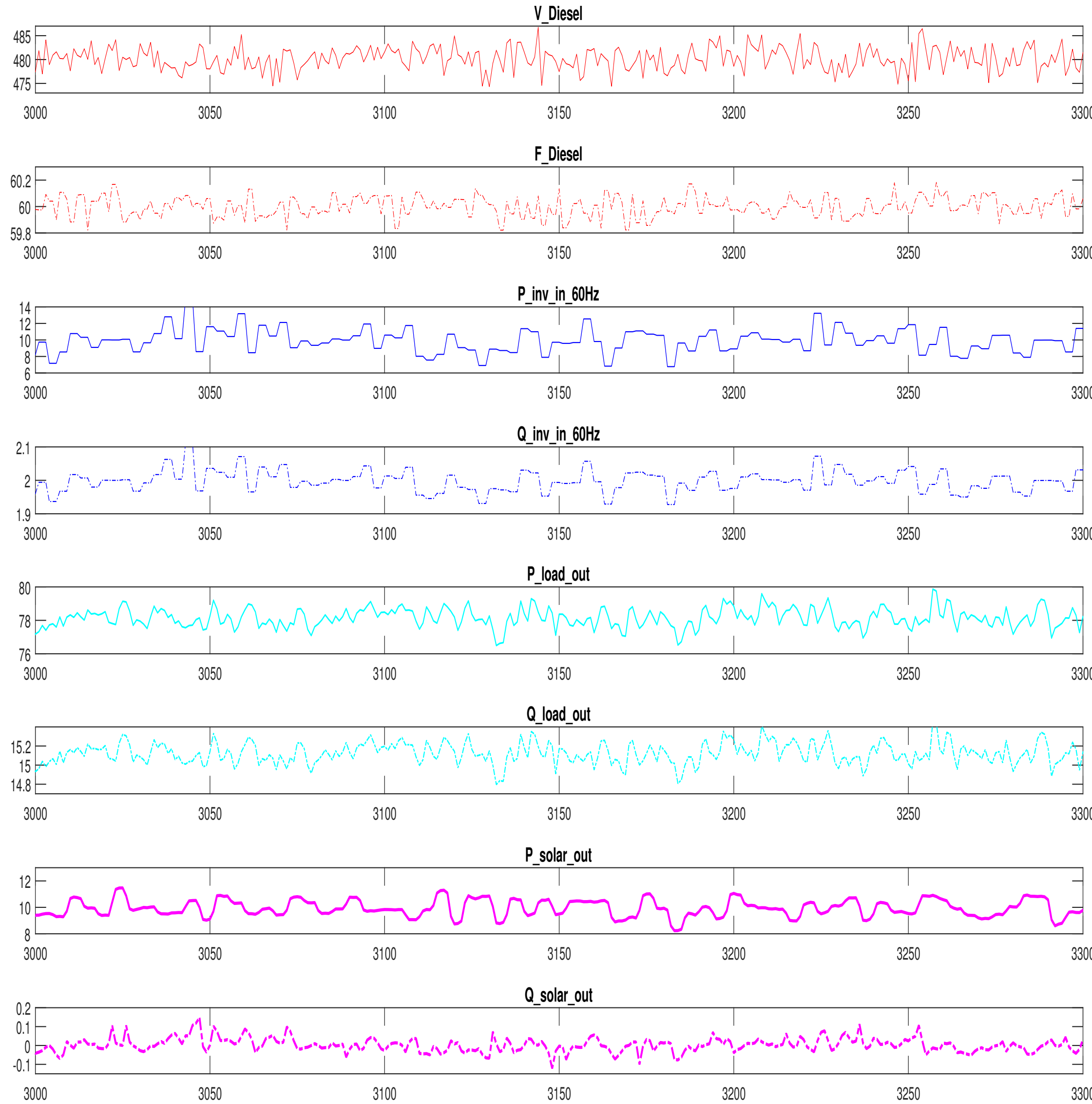

Figure 4.7: Input signals under the islanded mode.

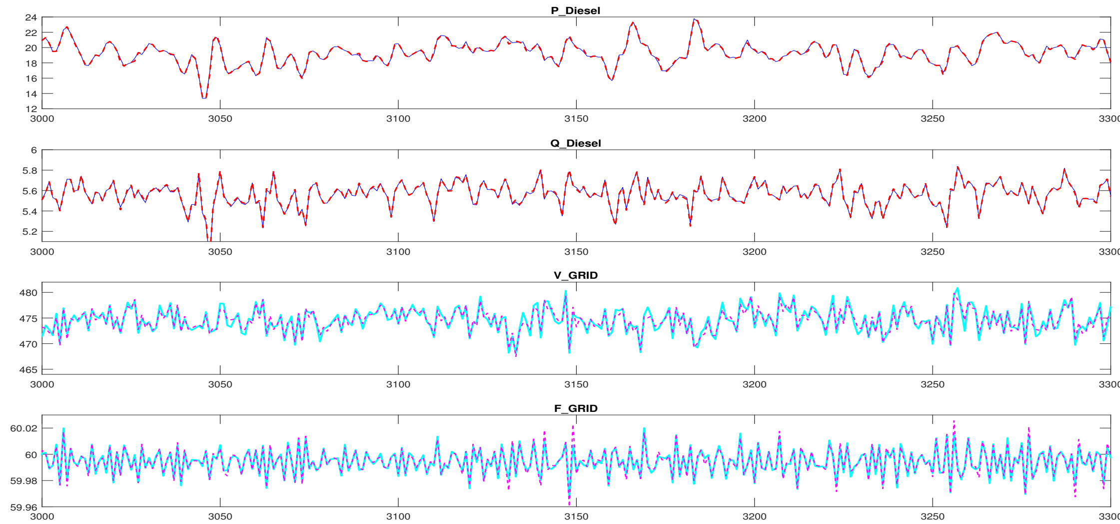

Figure 4.8: Output signals (solid lines) under the islanded mode. The dashed lines represent the outputs of a 6-th order model estimated by CoBRA.

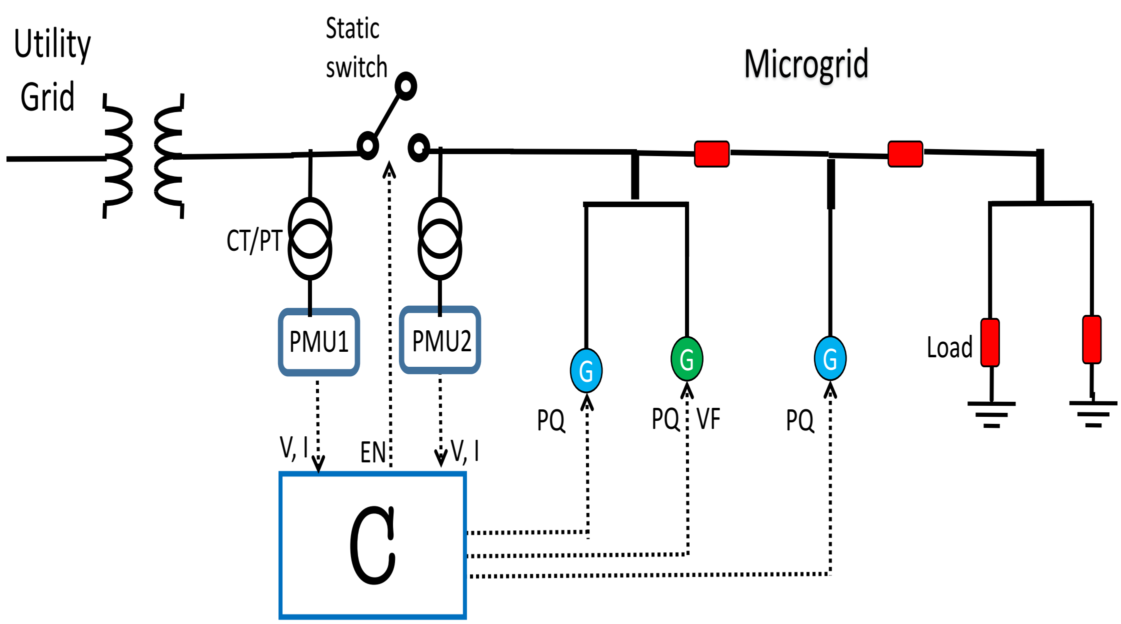

Figure 4.9: The figure shows the microgrid that can disconnect from the main utility grid with a controllable static switch. The controller issues P, Q or V, F commands to the generators depending on the state of the switch and feedback from PMUs 1 and 2. Inside the microgrid, the (green) generator operates as a voltage source in island mode which defines the microgrid voltage angle and the other (blue) generators follow the defined microgrid voltage angle.

where V_A, V_B and θ_{v_A}, θ_{v_B} are voltage magnitudes and phase angles of the phasors of main grid and microgrid feeders respectively. Hence, for a smooth transition at the time of reconnection, the magnitude of V_{AB} should be close to ~ 0 meaning voltages on both sides of the switch should have the same magnitude, phase angle, and frequency to avoid huge inrush current as calculated in [86]. Typically, voltage magnitudes at bus/node A and B are regulated at close to 1 p.u. Therefore, the voltage difference is mainly determined by angle difference $\theta_{v_{AB}}$ at the time of reconnection. Hence, the objective at the time of reconnection is to minimize the difference between the angles of nodes A-B while assuming the voltage magnitudes are close to each other (at 1 p.u.). The acceptable angle difference for reconnection is specified in IEEE 1547-Standard for Interconnecting Distributed Resources with Electric Power System [87].

4.5.2 Grid Connected to Islanding Transition

Intentional transition to islanding is a strategy to isolate a microgrid from possible disturbances and faults in the main grid. Such transition would require the operation of the microgrid DERs to switch from grid following to grid forming mode of operation. DERs' frequency and voltage controllers in the islanded mode will then activate and control the operation of the microgrid. While such frequency and voltage control is necessary for a microgrid to successfully operate in islanded mode, smooth switching from grid connected to islanded may require more than activating frequency and voltage control. If power flow across the PCC is significant, opening the breaker and switching to islanded mode could result in unacceptable frequency deviation or possibly DERs tripping. For DERs that possess inertia like diesel and gas turbine, a sudden increase in power demand could result in violation of their operational limits and may cause unexpected shut downs. This issue is often neglected in the literature and transition to islanding by just opening the breaker is assumed trivial. We propose one possible approach to smooth transition to islanding by performing power control at the PCC. The main objective would then be to quickly control the real and reactive power at the PCC and bring them to zero. Such control will be performed by utilizing the fast response of inverter-based DERs and possibly other DERs and by observing the operational limits of different DERs. This will ensure that the microgrid will not experience out of limit frequency deviations upon islanding.

Next, we explain the design procedure for a high bandwidth controller that brings the power flow at PCC to zero utilizing fast response of inverter-based DERs. PID control action is designed for this purpose using the identified DER models in the previous section with the consideration of available DER limits. Denoting by $P_{err} = P_{PCC}$ the power error signal and by P the controller output, the control command will be computed as

$$P(k) = k_p P_{err}(k) + k_i x_i(k) - k_d x_d(k) \tag{4.13}$$

where x_i and x_d are integrator and filtered derivative states of the controller updated according to

$$x_d(k) = \frac{P_{err}(k-1) - P_{err}(k)}{T_s} - a_1 x_d(k-1) \tag{4.14a}$$

$$x_i(k) = T_s.P_{err}(k) + x_i(k-1) \tag{4.14b}$$

Tuning of the controller is performed by using models obtained in the previous section between inverter power setpoint and power measurement at the PCC. While the controller is designed to achieve fast power regulation as required by the microgrid, it should also observe the operational limits of the inverter and other DERs in charge of implementing the computed power. Such limitations include maximum and minimum ramp rate and amplitude limits of the DERs. Denoting the DERs ramp rate and amplitude constraints by $r_{min,max}$ and $A_{min,max}$ respectively, the controller output at each step is limited to respect the following constraints:

$$A_{min} \leq P(k) \leq A_{max} \tag{4.15a}$$

$$r_{min} \leq \frac{P(k) - P(k-1)}{T_s} \leq r_{max} \tag{4.15b}$$

If the controller output computed in Eq. (4.13) violates any of the above constraints, it will be updated with its marginal value given above to respect the limitations of the DERs and avoid windup in the controller. Also, in such case the integrator's state in the controller should be updated to prevent windup according to

$$x_i(k) = [P(k) - k_p P_{err}(k) + k_d x_d(k)]/k_i \tag{4.16}$$

This will ensure that inverter amplitude or ramp rate saturation will not result in integral windup in the controller.

While the controller is tuned for the dynamic response of the inverter as obtained from models in the previous section, the power limitation of the inverter might be insufficient for bringing PCC power flow to zero. In such case, other DERs will also assist in achieving the setpoint determined by the controller and the amplitude and ramp rate limits of all the involving DERs should be considered while computing the upper and lower limits in Eqs. (4.15a) and (4.15b) and updating the limits online. Fast power regulation is initiated by fast inverter response and gradually accompanied by slower diesel response. To implement such logic, the control output is low-pass filtered and the filtered value is passed through rate and amplitude limit gates that apply the limitations of diesel generator before commanding to the diesel. The difference between the total controller output and the amount allocated to the diesel is commanded to the inverter. The inverter signal therefore constitutes the entire output of the controller upon inception of the control while diesel gradually ramps up and takes over low frequency trends seen within the control signal. A similar controller is designed for reactive power control and tuned using the model obtained in the previous section.

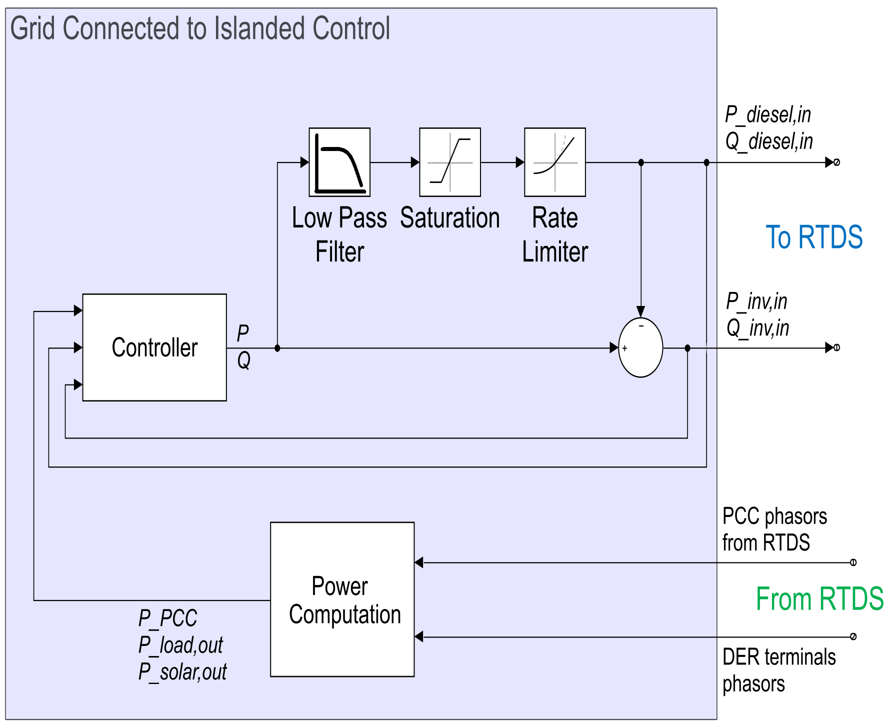

Figure 4.10: Overview of power control loop shown within the HIL setup. The power controller involves two major steps of power setpoint computation and power allocation between DERs. The first step is performed to minimize response time using PMU data and while utilizing knowledge of the outputs of the second step and operational limits of the DERs.

4.5.3 Islanded to Grid Connected Transition

The main goal of microgrid operation in the islanded mode is to supply power to local loads while controlling voltage and frequency within the acceptable limits at the microgrid. In certain industrial and critical applications, the requirement for fast grid re-connection is essential to reliability and resiliency of the microgrid. For example in microgrids where multiple DERs are supporting the load, in the event of one of the DERs' failure, fast grid re-connection should take place to prevent power outage in the microgrid. The following equations indicate the induced power surges at the PCC due to phase angle or voltage difference between the two sides of the breaker at the time of re-connection.

$$P_c = \frac{V_1.V_2.sin(\delta)}{X}, \quad Q_c = \frac{V_1.(V_1 - V_2).cos(\delta)}{X} \tag{4.17}$$

where V_1 and V_2 are are voltage magnitudes at the two sides before reconnection, X is the connection impedance and δ is the phase angle difference between voltages at the two sides. As seen in these equations, phase angle difference between the two sides of the breaker can lead to large real power surges while voltage magnitude difference can lead to reactive power surges. Since the voltages magnitudes are normally kept within acceptable limits by the main grid or by the DERs primary control, voltage differences will have negligible effects on reactive power flow and therefore voltage tracking is not necessary for smooth transition. This will be later verified in the results section. We now aim to design a controller to prevent power surges due to phase angle difference at the time of reconnection. While the microgrid relies on the DERs' primary control to adjust microgrid voltage and frequency, this additional control layer is designed to enable synchronization of the microgrid with the main grid and enable fast grid reconnection when required. Phase angle control is realized by manipulating the input frequency to the inverter. Various controllers have been proposed in the literature for phase angle control [88]. These controllers use either angle error or a combination of angle and frequency error as the input to the controller. In this work, we only use angle error and will show that a high bandwidth angle controller that employs PMU measurements can achieve phase angle synchronization in a few cycles. Since the model of the system between frequency input and phase angle output includes integral action and some delay, it suffices to provide only a proportional feedback to the phase controller. We further design this controller based on the estimated model of the system. The identified model relates the frequency input to the inverter with the phase angle of the microgrid under nominal loads in the islanded mode.

4.5.4 Mode Transitioning States

We now discuss different microgrid operation states in relation to the mode transition between grid connected and islanded modes. Intentional transition from grid connected to islanded will be initiated once potential request for islanding is identified. This may happen for example when the grid frequency has unusual fluctuations but is still within limits. At this time, "`init_island`" signal will be enabled by the supervisory controller and sent to the power controller. Such signal indicates the potential request for islanding and activates the power controller to quickly bring PCC power to zero. This could be accomplished in a time interval spanning only a few cycles depending on the inverter capacity and the performance of the designed controller. Once PCC power is brought within a predefined threshold around zero, islanding will be allowed by the supervisory controller. Islanding can then be performed upon the receipt of "`go_island`" signal by opening the breaker at PCC and switching DER controls mode from grid following to grid forming.

A similar procedure can be executed for transition to grid connected mode.

The process will be initiated once potential request for re-connecting is identified. At this time, "`init_gc`" signal will be enabled by the supervisory controller and sent to the angle controller. This signal indicates potential request for grid connection and activates the angle controller to quickly reduce the angle difference between the microgrid and main grid to zero. Once the angle difference is brought within a predefined threshold around zero, re-connection can be executed upon the receipt of "`go_gc`" by closing the breaker and switching DERs control from grid forming to grid following. This state transition algorithm is also presented in Fig. 4.11. A similar diagram could be shown for grid connected to islanded transition.

4.6 Power Tracking in Grid Connected Mode with Multiple DERs

4.6.1 Hybrid Power Dispatch Allocation

To formalize the concept of the Optimal Hybrid Power Dispatch (OHPD) algorithm, we consider a set of n DERs, each with their own set of parameters that characterize the dynamics of power generation/absorption and limitations on energy storage and summarized in Table 4.3. In the set of n DERs, we will also make a distinction between fossil energy generation (FEG), renewable energy generation (REG) and the storage of energy in a battery energy storage system (BESS).

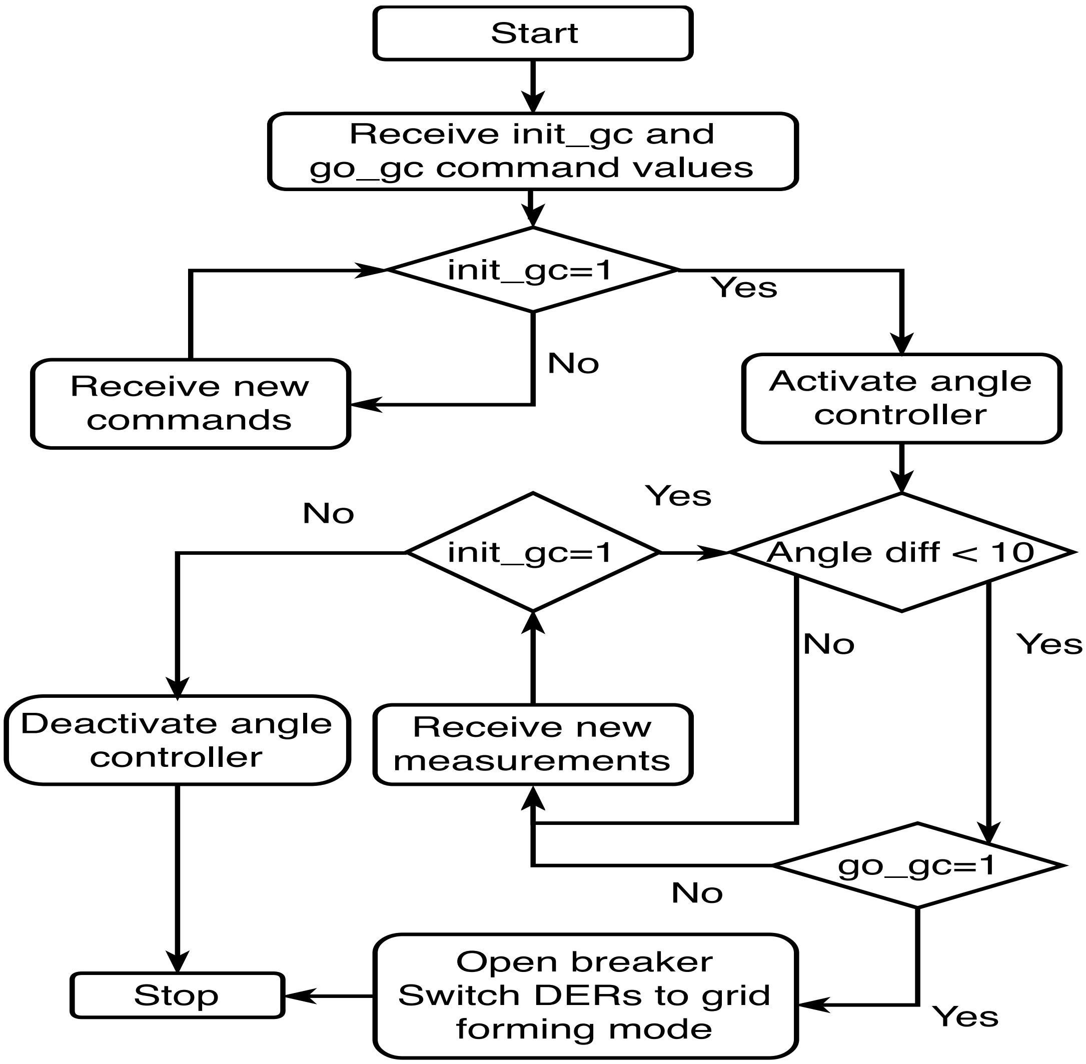

Figure 4.11: The sequential steps performed before transition to grid connection takes place

Table 4.3: System Constraints and Parameters

System Constraints	
PSL_f^{max}	Maximum power saturation limit of the FEG (kW)
PSL_f^{min}	Minimum power saturation limit of the FEG (kW)
SRL_f^{max}	Maximum power slew rate limit of the FEG (kW/s)
SRL_f^{min}	Minimum power slew rate limit of the FEG (kW/s)
t_d	Time delay between FEG demand signal and FEG power response (s)
$BPSL_i^{max}$	Maximum power saturation limit of i^{th} battery sub-unit (kW)
$BPSL_i^{min}$	Minimum power saturation limit of i^{th} battery sub-unit (kW)
$BSRL_i^{max}$	Maximum power slew rate limit of i^{th} battery sub-unit (kW/s)
$BSRL_i^{min}$	Minimum power slew rate limit of i^{th} battery sub-unit (kW/s)
SoC_i^{lb}	Minimum storage battery state of charge bound of i^{th} battery sub-unit (%)
SoC_i^{ub}	Maximum storage battery state of charge bound of i^{th} battery sub-unit (%)
BC_i	Storage battery capacity of i^{th} battery sub-unit and $BC = \sum_i^n BC_i$ for total battery capacity of the BESS (kWh)
Parameters	
SoC_{TAR}	Target Storage Battery State of Charge for the storage battery system (%)
R	Rate of fossil burnt ($/kWh)
$BSFC$	Brake Specific Fuel Consumption of the FEG
T_s	Sampling time used for the SoC model (s)
SoC_{lsb}	Lower safety bound for the battery system for safe BESS operation (%)
SoC_{usb}	Upper safety bound for the battery system for safe BESS operation (%)
SoC_{mean}	Mean value of SoC of all the individual battery sub-units (%)
η	Battery roundtrip efficiency for charging and discharging cycles
λ	Weighting factor used for battery or fossil optimization
β	Penalty factor used to account for fossil delay time
γ	Penalty factor used to account for power chattering and roundtrip efficiency

4.6.2 Multi Objective Problem Formulation

The discrete-time (real) power signals $p_i(t_k)$ for DERs $i = 1, 2, ..., n$, $k = 0, 1, ...$ sum up to a total (real) power demand

$$p(t_k) = \sum_{i=0}^{n} p_i(t_k)$$

for n available DERs. To formulate the OHPD algorithm, we first define the power tracking error, i.e. the error between the anticipated power demand (generation) $P(t_k)$ and the DER total achievable power demand (generation) $p(t_k)$,

$$e(t_k) = P(t_k) - p(t_k)$$

The time-dependent signal $P(t_k)$ may be a control signal from a controller, scheduling signal obtained from Economic Load Dispatch (ELD) [65], [89], or a result of a prediction algorithm such as solar or wind prediction. For formulation of the problem, $P(t_k)$ is assumed to be a discrete variable $\forall k \in I$. In this chapter, we focus on a guaranteed power allocation algorithm where the tracking error

$$e(t_k) = 0 \qquad \forall\ t_k,\ k = 0, 1, ...$$

assuming that the signal $P(t_k)$ satisfies the PSL, SRL constraints of all such available DERs as mentioned in Table 4.3.

Hence for a system consisting of multiple DERs, our goal would be to determine an optimal dispatch of (real) power over such multiple energy systems with interdisciplinary objectives and constraints [90]. To demonstrate such a multi-DER dispatch problem, we consider BESS and FEG systems each with a different set of objectives. The objective of the BESS would be to minimize the SoC error,

$$e_{SoC}(t_k) = SoC(t_k) - SoC_{TAR}$$

i.e to keep the total SoC of system at a target level for optimal usage and conditional flexibility to both charge and discharge the batteries as per demand. Hence the battery objective function can be formulated as a least squares error minimization problem formally stated as

$$J_b = \arg\min_{P_b} \{e_{SoC}(t_k)\}^2$$

where $P_b := P_b(t_k)$ is the instantaneous BESS power at time t_k.

However, the objective of a FEG system would be to minimize the total amount of fuel burnt for both economic and environmental reasons. The fuel burnt for the fossil generator depends on various factors such as power drawn, operation time, type of fossil, rated power of the generator and the load factor or brake specific fuel consumption ($BSFC$) of the generator [91], [67]. Hence the total economic fossil fuel cost at any time t_k would be

$$C_f = f(P_f, t_u, R, BSFC)$$

where P_f is the instantaneous fossil generation power, t_u is the generator usage time, R is the $\$$/kWh rate of the fossil burnt and $BSFC$ is the operating load factor. However, due to the delay time for fossil generators as mentioned in Table 4.3, the cost function would not only be a function of power until time t_k but also a function of cost paid for the fossil delay period t_d as a penalty for using fossil [92]. And hence the objective function for FEG system can be formulated as a quadratic norm cost minimization with an added penalty on future fossil cost due to present use as

$$J_f = \arg\min_{P_f}\{C_f^2 + \beta \cdot C_f\}$$

where β is a penalty factor for penalizing use of fossil delay period t_d. The second cost term is linear so that the emphasis remains on minimizing the fossil at present time t_k.

The total objective function of such multi-DER system would minimize both the objective functions

$$J_T(x) = \arg\min_{x}[J_b(x), J_f(x)]$$

where $x = [P_b \ P_f]^T$ Since both J_b and J_f are convex functions, the resulting non-negative weighted-sum would also be convex and would minimize the weighted scalar function of the two objective functions. Hence, the cost function for the dual peaks objective problem would be

$$J_t(x, \lambda) = \arg\min_{x}\{\lambda \cdot \bar{J}_b + (1 - \lambda) \cdot \bar{J}_f\}$$

where $\lambda = [0, 1]$ where $J_T(\lambda)$ is the Pareto objective function, $\bar{J}_b$, $\bar{J}_f$ are the scaled objective functions. The minimizer of $J_T(\lambda)$ is an efficient solution for the original multi-objective problem i.e. its image belongs to the Pareto curve [90]. The weighting factor λ emphasizes if the optimization favors $\bar{J}_b$ or $\bar{J}_f$. Hence the resulting optimization can be stated as

$$\min_{x} \lambda \cdot \{e_{SoC}(t_k)\}^2 + (1 - \lambda) \cdot \{C_f^2 + \beta \cdot C_f\}$$

s.t.

$$P_b(t_k) + P_f(t_k) = P(t_k) \quad \forall t_k$$

$$0 \leq P_b(t_k) \leq PSL_b^{max} \; if \; P(t_k) \geq 0, \; else \; PSL_b^{min} \leq P_b(t_k) \leq 0$$

$$SoC_{lb} \leq SoC(t_k) \leq SoC_{ub}$$

$$SRL_b^{min} \cdot T_s \leq P_b(t_k) - P_b(t_{k-1}) \leq SRL_b^{max} \cdot T_s$$

$$0 \leq P_f(t_k) \leq PSL_f^{max}$$

$$SRL_f^{min} \cdot T_s \leq P_f(t_k) - P_f(t_{k-1}) \leq SRL_f^{max} \cdot T_s$$

where $x = [P_b \; P_f]^T$ and T_s is the sampling time.

The SoC model is realized by the method of coulomb counting [72] as follows:

$$SoC(t_k) = -T_s \cdot (\sum_{k=1}^{m} P_b(t_k)/BC) \cdot 100 + SoC(t_{k-1})$$

4.6.3 State of Charge Balancing for Multi-Battery Systems

The guaranteed OHPD algorithm in Section 4.6.1 focuses on optimal power dispatch based on economic, safety and environmental factors among multiple non-similar DERs such as BESS, FEG and other REGs/non-REGs. However, once the power is split among such DERs, there might be another step of decision making process on how to sub-split the power signals among similar subsystems such as multiple batteries in a BESS, different types of fuel subsystems in the FEG, etc. In this chapter, we focus on slew-rate constrained energy level based optimal splitting of (real) power in the BESS containing multiple battery sub-units.

In this problem, the objective of the optimization is to balance the SoC levels of all the battery units i.e. to bring the SoCs of all the battery units to close to each other. Mathematically, the convex optimization problem can be stated $\forall \; k = 0, 1, ...$

as

$$J : \min_{P_i} \sum_{i=1}^{n} \{SoC_i(t_k) - SoC_{mean}(t_k)\}^2 + \gamma \cdot \{P_i(t_k) - P_i(t_{k-1})\}^2$$

s.t.

$$\sum_{i=1}^{n} P_i(t_k) = P_{dem}(t_k) \quad \forall t_k$$

$$BPSL_i^{min} \leq P_i(t_k) \leq BPSL_i^{max}$$

$$T_s \cdot BSRL_i^{min} \leq P_i(t_k) - P_i(t_{k-1}) \leq T_s \cdot BSRL_i^{max}$$

$$SoC_i^{min} \leq SoC_i(t_k) \leq SoC_i^{max}$$

$$P_i(t_k) \geq 0 \ if \ P_{dem}(t_k) > 0;$$

$$P_i(t_k) <= 0 \ if \ P_{dem}(t_k) < 0, \ P_i(t_k) = 0 \ if \ P_{dem} = 0$$

where,

$$SoC_i(t_k) = -T_s \cdot (\sum_{k=1}^{m} P_i(t_k)/BC_i) \cdot 100 + SoC_i(t_{k-1})$$

$$SoC_{mean}(t_k) = \frac{1}{n} \cdot \sum_{i=1}^{n} (SoC_i(t_k))$$

where P_{dem} is the total power demand for the BESS.

The first term in the optimization is a quadratic norm minimization of the SoC mean error for each such battery unit and the second term represents the additional penalty imposed on the control signal itself i.e. in the (real) power signals to avoid power chattering effects and also to compensate for efficiency losses in the battery units, with a penalty factor γ similar to weights in LQR control [74]. Typically, $\gamma < 1$ as the optimization is focused on SoC balancing.

4.7 Simulation Test Case Scenarios

4.7.1 Illustration of Islanding Control

The operation of the microgrid controller is demonstrated in different operation modes starting from grid connected and then switching to islanded mode due to a grid fault event and to provide uninterrupted power to the microgrid. Finally, fast grid reconnection capability is indicated. The considered microgrid comprises both diesel generator and inverter-based DERs. Due to the narrow control bandwidth of the diesel generation, control objectives such as fast power control and grid angle tracking would not be possible.

The control operation is performed using feedback of high-resolution synchrophasor data from PMUs and inverters with high update rate. Fig.4.12 shows power flow results realized by the power controller and dispatch of diesel and energy storage inverter. As shown in this figure, power flow at PCC (purple) is able to follow the set-point despite sudden load switching events of different sizes and other variable load disturbances. For negative power commands, the diesel produces zero output and the inverter will charge the battery while reacting to disturbances and smoothing PCC power flow. For positive power command signals, the inverter picks up high frequency components by quickly reacting to the power set-point. However, the inverter dispatch signal is gradually transferred to the diesel generator set. This transition leaves enough reaction margin for the inverter to quickly react to future disturbances.

Next, we present phase angle control results and discuss how angle control can provide fast reconnection. As shown in Fig.4.13, following an islanding event at time $t = 1s$, the phase difference between the two sides of the breaker at the PCC of the microgrid deviates from zero. Such switching to the islanded mode can be either intentional or unintentional. In the case of intentional switching, the microgrid su-

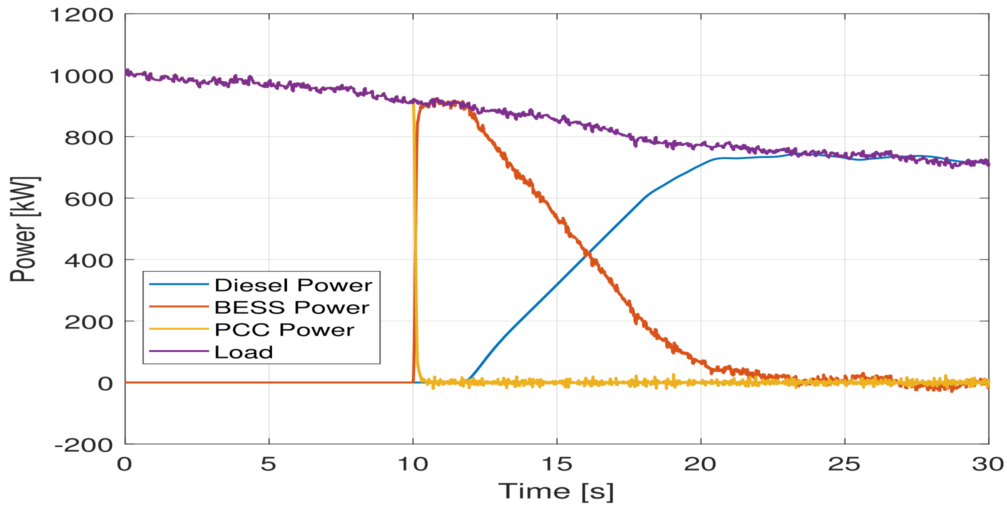

Figure 4.12: Microgrid power control at PCC in grid-connected mode. The microgrid controller attempts to follow the power set-point by allocating proper power set-points to the BESS inverter and diesel generator. The power set-point is reached by a combination of power output from the inverter and the diesel where high frequency fluctuations are allocated to BESS inverter and low frequency trends are picked up by the diesel.

pervisory controller will command lower level control to switch from grid connected mode to islanded mode at the time of switching. In the case of unintended switching, an intermediate islanding detection algorithm will command the microgrid controller to switch from grid connected mode to islanded mode. Once control is switched to islanded control and by keeping the setpoint of the inverter-based DER at $60Hz$, the microgrid can continue operation and provide power to its loads. However, motivated by the requirement for fast grid reconnection, the angle tracking controller should be activated to allow tracking the phase angle of the main grid and therefore enable fast grid reconnection upon request. Angle tracking controller is activated at $t = 3s$ where it quickly reduces phase angle difference and brings it to $< \pm 10^o$.

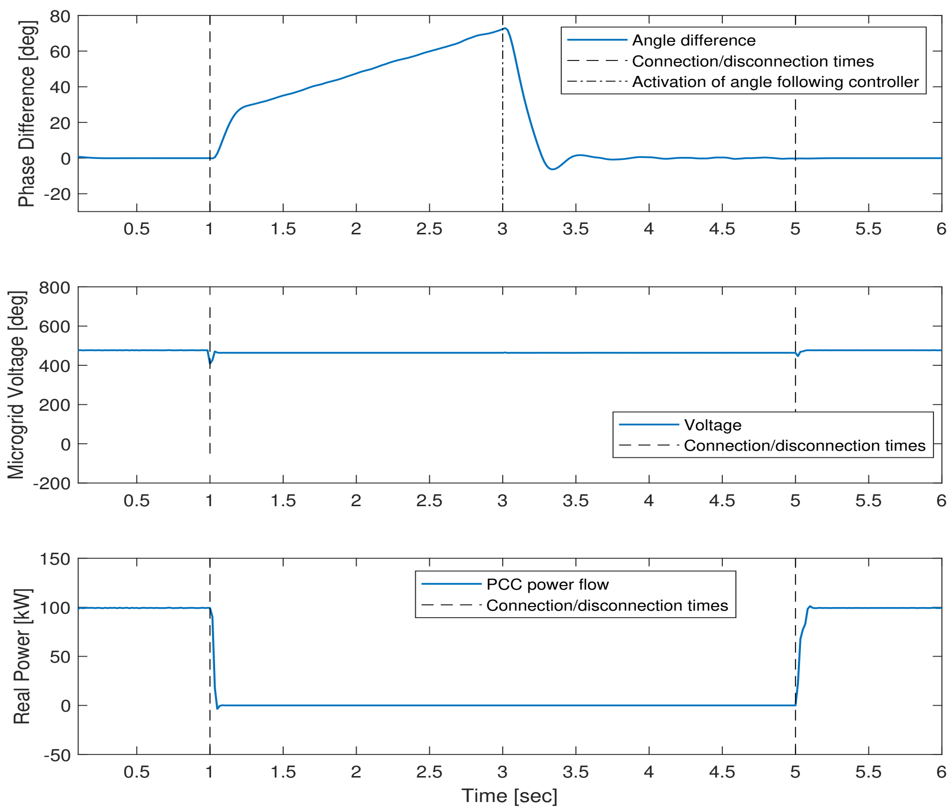

Figure 4.13: Microgrid voltage angle tracking control in the islanded mode. The microgrid controller attempts to follow the voltage angle set-point by manipulating the frequency setpoint to the inverter in voltage source mode. Islanding occurs at $t = 1s$ but angle controller does not activate until $t = 3s$. Grid reconnection can happen any time after the angle difference is below 10 degrees.

4.7.2 Illustration of Real Power Dispatch

To illustrate the results of the OHPD algorithm for different DERs with distinct energy levels, dynamics, limitations and economic factors, a hybrid micro-grid system with BESS and FEG systems is considered. The constraints and parameters are chosen in this simulation study based on optimal operational planning [91] for BESS and FEG systems. The constraints used in the simulation are as follows: PSL_b^{max}=20 kW, PSL_b^{min}=-20 kW, PSL_f^{max}=30 kW, PSL_f^{min}=0 kW (discharge only for FEG), SRL_b^{max}=1.5 kW/s, SRL_b^{min}=-1.5 kW/s, SRL_f^{max}=0.1 kW/s, SRL_f^{min}=-0.1 kW/s, SoC_{lb}=10%, SoC_{ub}=90%, BC=240 kWh, $t_d = 30s$. The parameters used in this study are: $SoC_{TAR} = 50\%$, $R = 5\$/kWh$, $T_s = 0.1666s$, $BSFC$ is considered 90% at all loads, SoC_{lsb}=10%, SoC_{usb}=20% and η=87%.

For simulation purposes in this chapter, a power demand curve to be tracked is assumed to be well within the power and energy limits of the available DERs. Hence the power demand curve considered is an amplitude, rate limited finite time signal P(t_k) that has a ramp rate limitation at $\pm$1.5 kW/s and is amplitude bounded by -20/+50 kW. The simulation runs for a 20 min period consisting of zero, positive and negative portions of the power demand curve.

Initially, the OHPD algorithm is initialized for a battery with an initial SoC of 70% and the value of lambda is chosen to be 0.5, i.e. equally favoring BESS and FEG objectives and leads to the results as depicted in Fig.5.8. It can be seen that the dispatch tracks the power demand curve (blue curve) with zero error (hence guaranteed power tracking) by equally accomplishing two objectives i.e. bringing the battery SoC to the target SoC and minimizing the total fuel cost of fossil while respecting the DER constraints. It can be observed that the power during negative power demand period i.e. power generation, is fully absorbed by the storage battery system because of energy irreversibility FEG system. The simulations illustrated in

Fig.5.8 show that the BESS would be at 45% SoC and the cost of fossil burnt would be $\$20$ by the end of 20 mins of operation. The value of β was chosen to be $0.1 \cdot t_d$ to penalize fossil delay time. However, for the value of $\lambda = 0$, the emphasis of

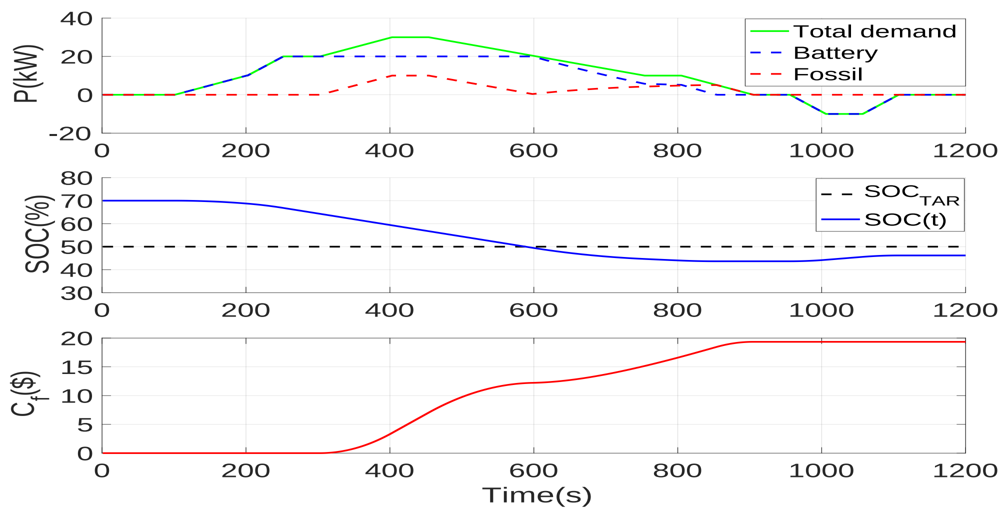

Figure 4.14: Results of dual optimization when the value of $\lambda = 0.5$. It can be seen that the optimization focuses both on minimizing e_{SoC} and C_f hence resulting in moderate performance.

the optimization would be completely on minimizing fuel costs and hence the energy level of the storage battery system degrades faster compared to the case of $\lambda = 0.5$. The illustration in Fig.5.4 shows the resulting power dispatch, battery SoC and fuel costs over time for an initial SoC of 70%, $\lambda = 0$ for an identical demand curve as in the previous case. It can be seen that all the power comes from BESS except when the demand goes over the maximum PSL of the battery system at $t_k = 300s$. The FEG runs until the demand goes back below the maximum PSL at $t_k = 600s$. Hence the battery SoC degrades to 40% until it charges back to 42% in the generation period before $t_k = 1000s$, as seen in Fig.5.4. However, the fuel cost is minimized and is at 12$ by the end of simulation period. Similarly, for the value of $\lambda = 1$, the emphasis

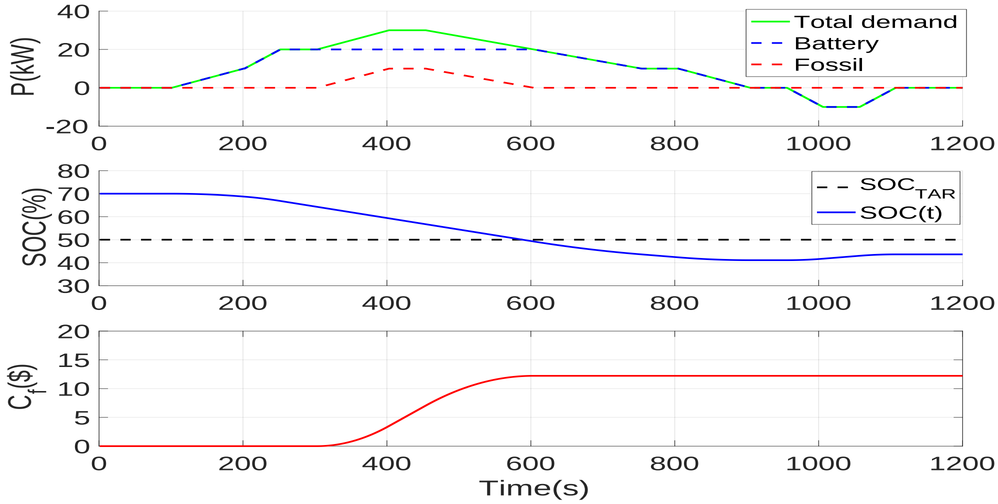

Figure 4.15: Results of dual optimization when the value of $\lambda = 0$. It can be seen that the optimization focuses entirely on minimizing C_f hence resulting in worse performance of SoC error, e_{SoC}.

of the optimization would be completely on minimizing SoC error, e_{SoC} and hence the fuel consumption goes up. The illustration in Fig.5.6 shows the resulting power dispatch, battery SoC and fuel costs over time for an initial SoC of 70%, $\lambda = 1$ for an identical demand curve as in the previous cases. It can be seen that all the power comes from BESS until it reaches SoC_{TAR} and then the demand is fulfilled by the FEG, except for the period when the demand goes above PSL_b^{max} of BESS. Hence the battery SoC settles close to $SoC_{TAR} = 50\%$, however performing worse on the fossil costs over 30$ as seen in Fig.5.6. Fig.5.7 illustrates the Pareto optimality map for different power demand levels. The colormap in Fig.5.7 depicts the total objective function value for different values of SoC error e_{SoC} and fossil costs C_f. Imaginary $-45°$ slope intercept lines running across the map represent different levels of power demand. In Fig.5.7, the box marked with 1 shows a function optimum for a demand signal of 1 kW and the box marked with 2 shows a function optimum for 0 kW

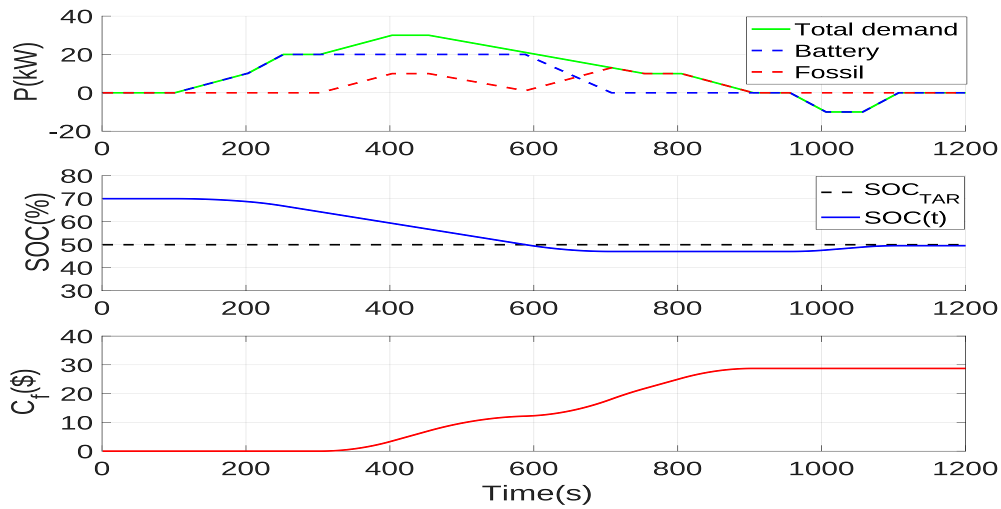

Figure 4.16: Results of dual optimization when the value of $\lambda = 1$. It can be seen that the optimization focuses entirely on minimizing e_{SoC} hence resulting in worse performance of fossil costs, C_f.

demand signal, indicating that the optimal function value increases with increase in demand.

However, if the value of lambda is considered to be a function of SoC as in (2.12) with $SoC_{lsb} = 10\%$ and $SoC_{usb} = 20\%$, the results are illustrated in Fig.4.18. An initial SoC of 30% is considered to demonstrate the effect of the SoC dependency of λ, for an identical demand curve as in the previous cases. It can be seen that all the power comes from BESS until it reaches SoC_{usb} and then the demand is shared proportional to $SoC(t_k) - SoC_{lsb}$ between BESS and FEG, until it reaches SoC_{lsb} below which all the power comes from the FEG. Hence the BESS SoC settles inside the SoC_{lsb} and SoC_{usb} bounds by the end of simulation period as shown in Fig.4.18.

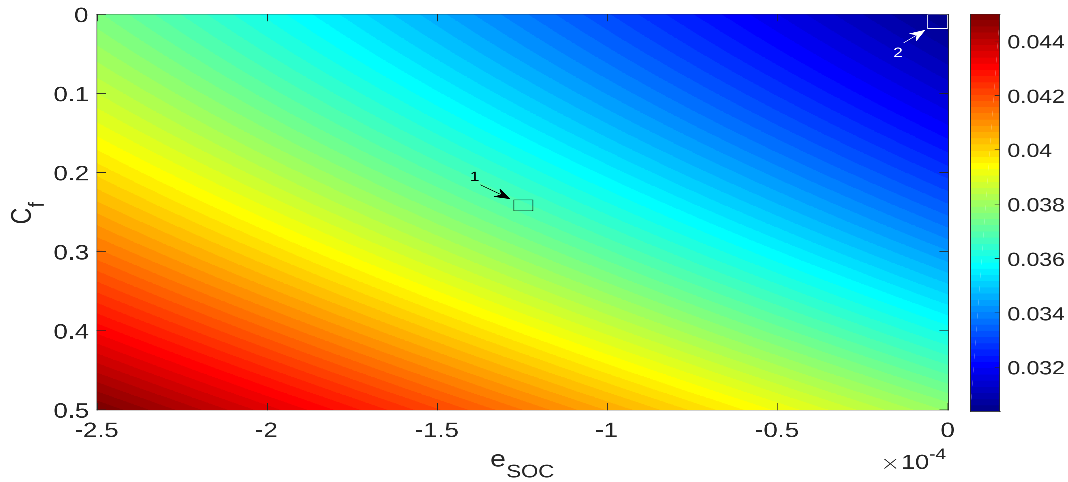

Figure 4.17: Colormap representing Pareto optimal criterion the color of the map represents the objective function value $J = \lambda \cdot J_b + (1 - \lambda) \cdot J_f$. Box marked 1 represents the objective function value when the demand $P(t_k) = 0$ and box marked 2 represents the objective function value when $P(t_k) = 1$ kW. $-45°$ imaginary lines represent objective function values at different levels of demand, $P(t_k)$.

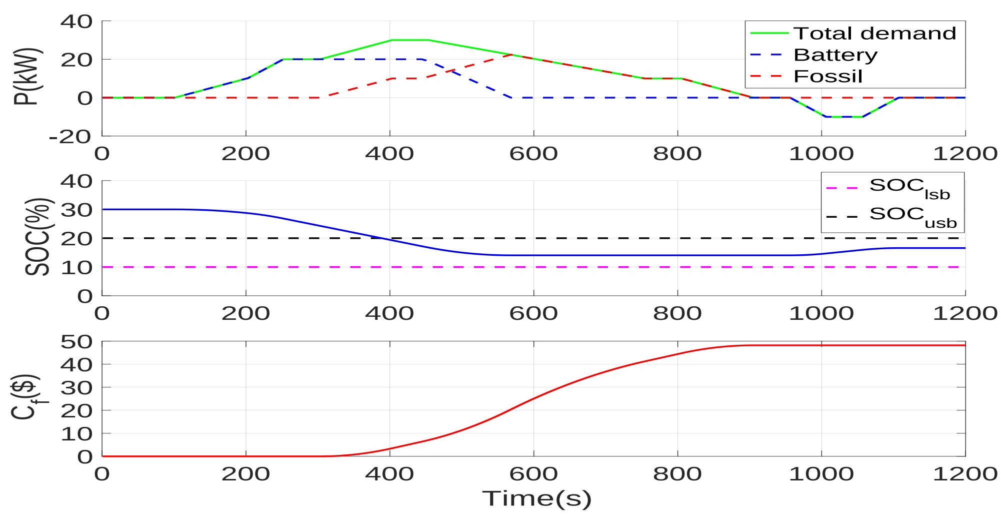

Figure 4.18: Results of dual optimization when the value of $\lambda = f(SoC)$. It can be seen that the optimization focuses entirely on minimizing C_f until $SoC(t_k) < SoC_{usb}$ but then changes to minimizing e_{SoC}.

4.7.3 SoC Balancing Power Splitter

To illustrate the SoC balancing algorithm, three storage battery units with equal storage capacity BC_i and power limits $BPSL_i$ but distinct $BSRL_i$ values are considered. For simulation purposes, the battery capacity of each unit $BC_i = 80$ kWh, $BPSL_i = 6.667$ kW and $BSRL_1 = 1$ kw/s, $BSRL_2 = 0.667$ kW/s and $BSRL_3 = 0.333$ kW/s respectively. A step signal with a ramp rate of 1.5 kW/s and step size of $\pm$ 9 kW is considered as a total power demand signal for simulation just over 12 min period. To demonstrate the functionality of the BESS SoC balancing algorithm, initially the SoCs of the units are considered to be at 55%, 60% and 80% respectively. The SoC bounds of all the units are considered to be $SoC_i^{lb} = 10\%$ and $SoC_i^{ub} = 90\%$. The simulation in Fig.4.19 shows the balancing of battery SoCs over

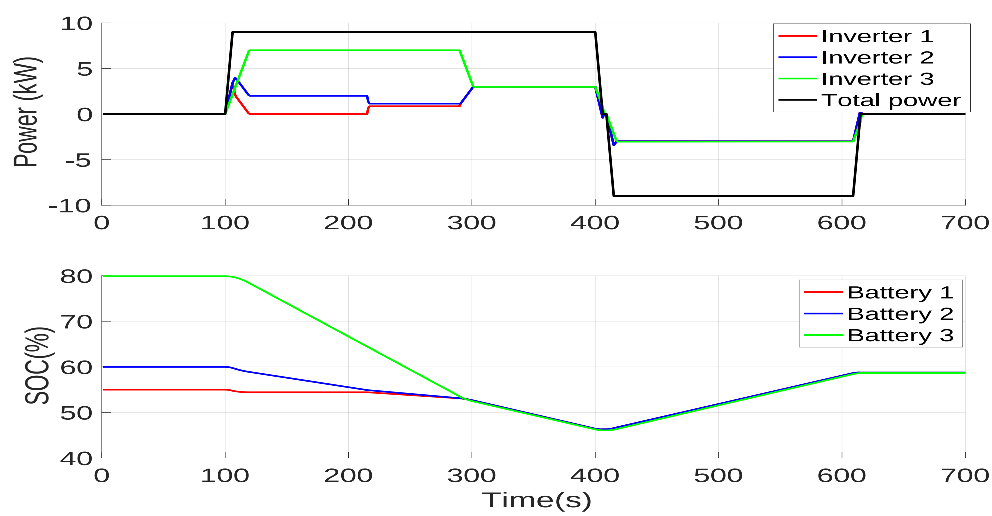

Figure 4.19: Illustration of the SoC balancing algorithm for three battery sub-units. It can be seen that SoCs of all three sub-units converge to the mean SoC value at $t = 300s$ and dispatch power equally satisfying the real power demand and the DER constraints after $t = 300s$.

time to their mean value. Initially, the power is drawn from Battery 1 and Battery 2 to track the ramp rated demand signal until it stays constant at 9 kW. Then, the maximum power is drawn from Battery 3 to bring the SoC to the mean level, and the remaining power comes from Battery 2 and Battery 1. However, when all three SoCs converge to the mean, it results in equal power sharing when the demand signal is constant. It can be seen at $t = 300s$ that Battery 1 and Battery 2 contribute partially higher power than Battery 3 to satisfy the ramp-rated demand signal (1.5 kW/s) but then all three split power signals become equal when the demand remains constant at -9 kW. It can be seen after $t = 300s$ that all three SoC become close to each other equal to SoC_{mean} hence validating the SoC balancing algorithm with a given battery and inverter constraints. The value of γ was chosen to be equal to 0.33 for optimal results.

4.8 Summary of Contribution

A detailed switching and control algorithm is proposed for a microgrid consisting of multiple distributed energy resources including both synchronous generators and inverter-based DERs. The controller enables fast and smooth intentional switching between islanding and grid connected modes with small voltage and frequency transient fluctuations. Modeling of the plant for control purposes is performed by model estimation techniques using DER inputs, PMU measurements, and realization algorithms. The controller is designed based on the estimated models. The data-based controller design methodology can be easily extended to accommodate other combinations of DERs in the microgrid. The developed control algorithm is tested on a HIL setup that models the microgrid and proves fast and smooth transition.

The text and data in Chapter 4, in full, is a reprint of the material as it appears in "Microgrid Dynamic Modeling and Islanding Control with Synchrophasor Data", Sai Akhil R. Konakalla, Raymond A. de Callafon, *to appear in IEEE Transactions*

on Smart Grids, “Covariance Based Estimation for Reduced Order Models of Microgrid Power Flow Dynamics”, Yangsheng Hu, Sai Akhil R. Konakalla, Raymond A. de Callafon, *IFAC-PapersOnLine*, 51.15 (2018): 903-908 and ”Optimal hybrid power dispatch for distributed energy resources with dynamic constraints”, Sai Akhil R. Konakalla, Raymond A. de Callafon *Power and Energy Systems (ICPES), 2017 IEEE 7th International Conference on*, 2017. The dissertation author is the primary investigator and author of the first and third article and is the secondary author of the second article.

Chapter 5

Short-term Prediction of Missing Synchrophasor Data for Control

5.1 Introduction

Many applications such as real-time control involve computations based on instantaneously obtained PMU measurements [4] [93]. The performance of any such state estimator or controller would largely depend on the availability and quality of PMU measurements. For PMU measurements, IEEE standards require the total vector error (TVE) between a measured phasor and its reference value to be less than 1% under steady state operation to be able to mark the data as good quality data [21]. However, in practice, this criterion may not be met due to the presence of different measurement uncertainties in PMU data, missing PMU data samples due to errors in measurement data such as framing errors, sync word errors, CRC errors, packet size errors or data loss that may occur due to problems in the communication between the PMU and the data server [21], [4]. Such data losses often make the states unobservable and can degrade the performance of a real-time controller and/or lead to system instabilities. Manual observations of PMU data time sequences to observe data errors or possibly detect missing samples in the data quickly becomes an unwieldy task.

Clearly, automated or semi-automated data analysis techniques to identify bad quality or missing PMU data are highly desirable [21]. Furthermore, it is essential to develop methods to allow data extrapolation at the sampling rate of the PMU in case of temporarily invalid or missing PMU measurements. The short-term prediction can also be useful to interpolate between slow SCADA measurements when the fast PMU data might be missing. One classic method to address data missing problem is to hold the data constant equal to the last valid sample (sample and hold) until the measurements become valid again [21]. This method is not optimal since it only uses the knowledge of the last valid sample, not the signal spectral contents to predict the next 'P' missing samples of data. However, there are some existing studies on how the signal spectral contents and previous observations can be better used for prediction [94], [95], [96], [97]. In one of the previous works, the authors have proposed one such prediction methods that use grid dynamics and model to estimate prediction filters [97]. However, this method uses the entire dataset of length 'L' to estimate only one prediction filter to predict next 'P' samples using P-sample ahead prediction. The filter works well to predict a few samples ahead but it could be seen that predictions asymptotically converge to the unconditional mean. Fig. 5.1 illustrates this behavior of convergence of the forecasts to a mean value as the predictions become longer. This limits the prediction only to a few samples by capturing fast (high frequency) trends and the prediction is not capable of capturing slow (low frequency) trends in the data.

One method to capture both higher and lower frequency trends is by decimation of the dataset (signal) and subsequent estimation of prediction filters as discussed in [97] by least squares minimization (fitting) at different intervals of the spectral con-

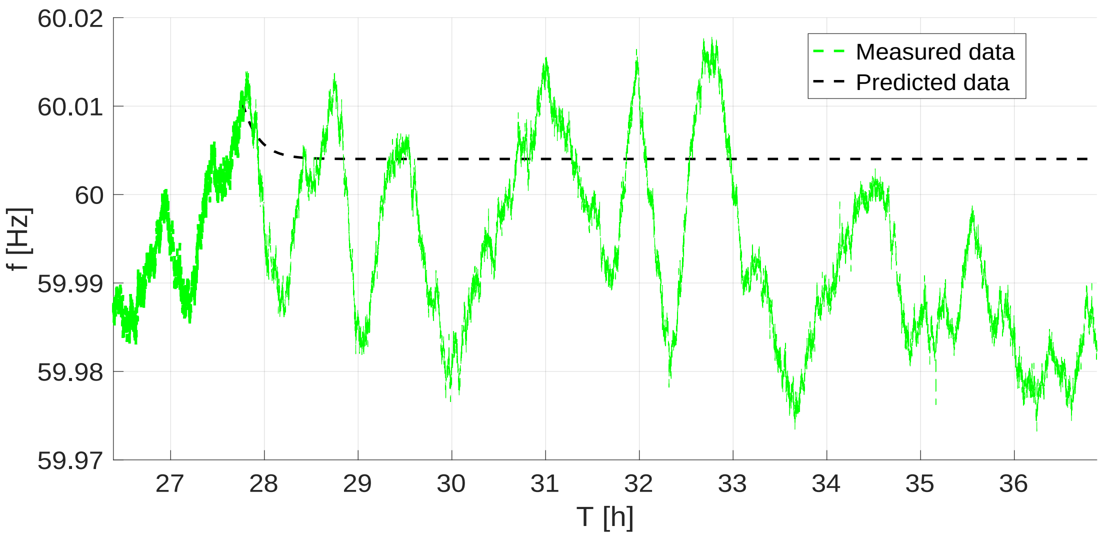

Figure 5.1: Prediction results using 'L' step ahead predictors from a single filter.

tent of the dataset [30]. This will result in predictors that will be able to capture the full spectral content of the signal. In this chapter, non-uniform downsampling and reconstruction is used for prediction and interpolation of the predicted samples, resulting in unevenly spaced predictors [98], [99]. In addition, the chapter presents interpolation or reconstruction methods to be able to compute the extrapolated data at the same sample rate of the PMU i.e 60 or 120 Hz. Two such interpolation methods have been discussed in this chapter [100], [101], [102] and result in uniformly sampled prediction data that can be used as prediction to extrapolate temporarily missing or invalid PMU measurements until they become valid again.

For applications such as real-time phasor (or power flow) control, it is very important and be able to smoothly transfer to the actual PMU measurements in the case when PMU data becomes valid again to avoid any discontinuities (jumps) in the signals. In this chapter, a Kalman filter based smoothing approach is used to fuse the prediction data and newly valid observations with weights so as to gradually transition from extrapolated data to actual PMU measurements [103], [104]. For a practical illustration of the prediction algorithm, real-time PMU data acquisition is implemented on a Raspberry PI device running python packages like NumPy and SciPy under Linux environment, receiving C37.118 format data [20] from the microPMU system [23], developed by Power Standards Lab (PSL) as an extension to the well-established low voltage PQube instrumentation by Power Sensors Ltd. It is shown how real-time processing of the phasor data received by C37.118 can be used for data prediction on the basis of data obtained before and after several local missing samples of data as measured by the μPMU.

5.2 Motivation for Non-uniform Downsampling

In many signal processing problems, it is necessary to compute the Fourier transform (FT) of a function or measured response that rises quickly then decays slowly with increasing abscissa value. In such cases, it is desirable to derive the function, or sample the response, with a small abscissa interval during the rise, but then increase the interval with increasing abscissa value [105]. Various approximate techniques have previously been presented to accomplish effectively the required FT of a function or response that is sampled in such a manner. These include fitting the response with a set of functions for which the analytical FTs are known, discrete Fourier transform (DFT) of a suitably interpolated response, decomposition of a polynomial approximation of the response into partial fractions, and FTs of segments of the response. Applying this idea for prediction, it translates to estimating predictors that are highly concentrated (small abscissa interval) in the beginning and far placed with increasing abscissa. For a given discrete signal $f(k)$, $k = \{1, 2, ..., L\}$ in order to be able to predict P samples ahead at a given sampling rate n, designing a single prediction filter would not be optimal because of the inability to capture the complete frequency spectrum of $f(k)$. As a result, the prediction error can only be

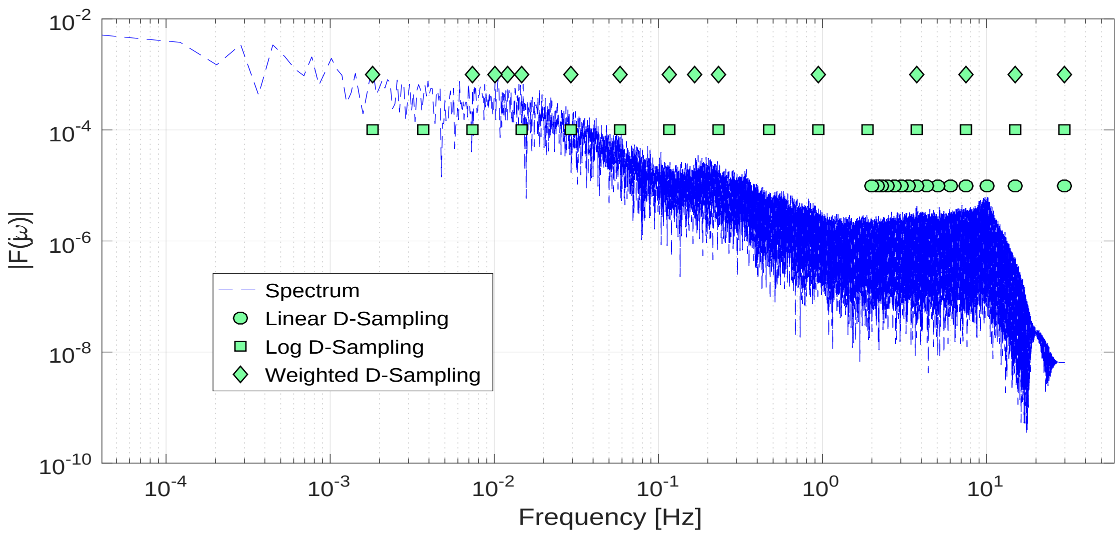

Figure 5.2: Figure illustrating the frequency spectrum $|F(j\omega)|$ of the signal $f(k)$ with comparison of three different choices of downsampling factors: (a) Linear downsampling, (b) Logarithmic downsampling and (c) Spectrally weighted downsampling.

small either at $p << P$ or $p \approx P$ but not both because such unbiased filter can estimated only by fitting the spectra (by frequency domain) in a limited range. Hence multiple prediction filters may be required to be able to capture all the frequency contents of the signal defined by the signal spectrum. Fig. 5.8 shows the frequency spectrum of one such signal $f(k)$ along with the intervals of down-sampling to be used to estimate prediction filters given a fixed number of filters. It can be seen that uniformly spaced down-sampling i.e. taking a signal and downsampling it by a factor of $n = 1, 2, 3, ..., N$ does not capture the entire spectra mainly due to the logarithmic nature of the spectrum. However, choosing a downsampling factor of $n = \{1, 2, 4, 8, ..., N\}$ i.e. $n = 2^b$ where $b = \{0, 1, 2, ..., \log_2 N\}$ (logarithmic downsampling) would capture the complete spectrum of $f(k)$ with the same number of filters. An alternate choice of selecting a downsampling factor n weighted according a known spectrum of the signal $f(k)$ would result in optimal prediction filters capturing the most significant parts of the spectrum of $f(k)$.

For the downsampling of data $f(k)$, two well known solutions exist. First, if the downsampling factor is an integer n, then the decimation method by interleaving n samples to form down-sampled data vector $f_n(m)$ where $m = \{1, 2, ..., L/n\}$. Second, if the downsampling factor is an a rational number $r = p/q$ then the frequency domain method that includes zero padding to the required length, computing the discrete Fourier transform (DFT), truncating the DFT a factor q, and finally computing the inverse DFT (IDFT) is used. However, in this chapter, only an integer downsampling factor is considered.

5.3 Spectrum Based Optimal Filter Estimation for Phasor Data Prediction

5.3.1 General Filter Estimation for Prediction Error Minimization

To set up the framework for the construction of optimal filters to process the discrete-time sampled phasor data $x(k)$ we consider the noisy observations of the phasors and frequency for an unobserved event signal $d(k)$,

$$
\begin{aligned}
A(k) &= G_A(q)d_A(k) + H_A(q)e_A(k) \\
\phi(k) &= G_\phi(q)d_\phi(k) + H_\phi(q)e_\phi(k) \\
f(k) &= G_f(q)d_f(k) + H_f(q)e_f(k)
\end{aligned}
$$

But since the prediction algorithm would be meaningful only during the ambient period of the grid i.e. when there are no significant events [29], only consider the noisy observations of the phasor data $A(k)$, $\phi(k)$ and frequency data $f(k)$ without any main grid event signal $d(k)$, the observations can be then characterized by the filtered white noise signals

$$\begin{aligned} A(k) &= H_A(q)e_A(k) \\ \phi(k) &= H_\phi(q)e_\phi(k) \\ f(k) &= H_f(q)e_f(k) \end{aligned} \tag{5.1}$$

where q is the shift operator, $G_A(q)$, $G_\phi(q)$ and $G_f(q)$ denote unknown grid dynamics that filter the effect of main event signal $d(t)$. $H_A(q)$, $H_\phi(q)$ and $H_f(q)$ denote unknown dynamic systems that model filter the white noises $e_A(k)$, $e_\phi(k)$ and $e_f(k)$ and model the spectral content of the output noise on the data. As such, we may assume that $H_A(q)$, $H_\phi(q)$ and $H_f(q)$ are spectral factorizations of the noise spectrum and are stable and stably invertible filters. We assume that the white noises $e_A(k)$, $e_\phi(k)$ are uncorrelated with an unknown variance without loss of generality to formulate the estimation problem of the optimal filters to process the phasor data $(A(k), \phi(k))$. To illustrate the idea of optimal filtering, we only consider the derivative of phase angle $\phi(k)$: $f(k)$ as in [22] as it is synonymous to prediction of $\phi(k)$ (differentiation) and the approach to predict $A(t)$ would need a similar exercise.

Our goal of estimating $H_f(q)$ reduces to a system identification problem and special case of Box-Jenkins model structure [30], without an input i.e. without a stimulus signal $d(k)$. It involves time-series modeling with model structures such as the AutoRegressive Integrated Moving-Average (ARIMA) model [96]. To illustrate the parameterization of any such time series regression model, consider the structure where the parameterization in terms of θ is chosen so that the family of models $g(k, \theta, \phi(k)), \theta \in \mathbf{D_m}$ covers an exclusive set of models [30]. From Eq. (5.1) we have

$$f(k) = H_f(q, \theta)e_f(k)$$

The prediction error for generalized p step ahead predictor of the system is

$$\epsilon_p(k,\theta) \;\; = \;\; f(k+p) - \hat{f}(k+p|\theta)$$

and thus the p step ahead prediction of the system is

$$\hat{f}(k+p|\theta) \;\; = \;\; (1 - W_p(q,\theta))f(k) \tag{5.2}$$

where $W_p = \bar{H}_p(q)H_f^{-1}(q)$, $\bar{H}_p(q) = \sum\limits_{l=0}^{p-1} h(l)q^{-l}$, $h(l), [l = 0, 1, 2, ...]$ is the impulse response of the disturbance model $H_f(q)$.

Without loss of generality, by dropping θ $H_f(q,\theta)$ can be written as

$$H_f(q) \;\; = \;\; L(q)\tfrac{C(q)}{A(q)} \tag{5.3}$$

where polynomials $A(q)$, $C(q)$ are defined as

$$\begin{aligned} A(q) &= 1 + a_1q^{-1} + a_2q^{-2} + ... + a_{na}q^{-na} \\ C(q) &= 1 + c_1q^{-1} + c_2q^{-2} + ... + c_{nc}q^{-nc} \end{aligned} \tag{5.4}$$

and $L(q)$ is an optional transformation operator that makes the time series stationary [96].

Using the quadratic norm minimization of the prediction error [30], the parameter estimate $\hat{\theta}_N$ can be formulated as

$$\hat{\theta}_N \;\; = \;\; \arg\min_\theta \tfrac{1}{N} \sum_{t=t0}^{N} \tfrac{1}{2}\epsilon_p^2(t,\theta)$$

This minimization can be recursively computed using recursive search algorithms such as

$$\hat{\theta}_N^{m+1} = \hat{\theta}_N^m - J_N(\frac{1}{N}\sum_{t=t0}^{N}\frac{1}{2}\epsilon_p^2(t,\theta_N^m))$$

where, $J_N(.)$ is a directional search function to update the parameters. However, it has been shown in [106] that a normal one step ahead prediction error criterion is optimal for predicting each step in the p step ahead prediction if the model error is dominated by the variance part.

The model with the parameters obtained from such search algorithms can be validated by investigating the residual whiteness such as through sample covariance [30].

$$R_\epsilon^N(\tau) = \frac{1}{N}\sum_{k=1}^{N-\tau}\epsilon(k,\hat{\theta}_N)\epsilon(k+\tau,\hat{\theta}_N)$$

for different time lags τ. The model parameters thus constitute the coefficients of the numerator and denominator of the filter $H_f(q)$.

5.3.2 Computation of Filters from Downsampling of Data

The filter $H_f(q)$ derived from the complete measurements of data phasor data $f(k)$ captures the complete spectral content of $f(k)$, $|F(j\omega)|$ and hence can only either capture the response at either lower or higher frequencies depending on the spectral fit of the filter. Hence the idea here is to create N independent down-sampled filters $H_f^n(q)$, $n = \{1,2,...,N\}$ from the down-sampled versions of the data $f(k)$, each designed to capture the spectral content of the n down-sampled data. In other words, the spectral content is sampled in a way that it can be perfectly reconstructed back from the samples. The choice of the spacing of n has been discussed in Sec. 5.1. In this case, non-uniform downsampling at sampling times m and signal sample values $f_n(m)$ for $m = 1,...,L/n$. The aim is to compute the filtered version of $f(k)$ with the anti-aliasing filter $AF(q)$ for downsampling and ensure that the original sampling is faster than at least twice the signal frequency of the down-sampled signal. The aim of the filtering is to remove frequencies above $1/(2T)$, which becomes the effective Nyquist frequency after downsampling. Using the method in Section. 5.2, filters $H_f^n(q)$ are computed using filtered down-sampled data $f_n(m)$.

5.3.3 Filter Based 1-Sample Ahead Predictors

The estimated optimal filter $H_f(q)$ can now be used to forecast the future samples of the phasor and frequency data. According to the p step ahead forecasting equation of any such assumed model assuming $L(q)$ is stable and invertible and substituting Eq. (5.4) in Eq. (5.3), we have

$$\begin{aligned} L^{-1}(q)[f_n(m) + a_1 f_n(m-1) + ... + a_{na} f_n(m-na)] \\ = e(m) + c_1 e(m-1) + ... + c_{nc} e(m-nc) \end{aligned}$$

With the conditional expectation, the forecasted value based on such model would be

$$\begin{aligned} \hat{f}_n(m+p) = E[-(\sum_{i=1}^{na}(a_i f(m+p-i)) \\ +L(q)(\sum_{j=1}^{nc}(c_j e(m+p-j)))] \end{aligned}$$

One step ahead prediction of $f_n(m)$ is obtained by substituting $p=1$ in the above equation,

$$\begin{aligned} \hat{f}_n(m+1) = E[-(\sum_{i=1}^{na}(a_i f(m+1-i)) \\ +L(q)(\sum_{j=1}^{nc}(c_j e(m+1-j)))] \end{aligned}$$

5.4 Data Interpolation from Irregular Sample Predictors

In conventional sampling theory, the periodic signal is sampled at a uniform rate at a minimum of twice the signal bandwidth. However, the problem of reconstructing a band-limited signal from its values taken at a sequence of irregularly spaced sample points has also attracted much attention in signal processing literature. A number of computational methods have been proposed for different aspects of the subject. A key result in this field is the theorem of Duffin and Schaeffer [99] which says that if the density of the sequence of sample points is strictly larger than the Nyquist density, then stable reconstruction is possible. Assuming that the signal to be reconstructed is periodic, and that the signal samples correspond to one or more periods of the periodic signal, it can be consistently assumed that the signal is band-limited. This enables the (periodic, band-limited) signal to be exactly reconstructed from the signal according to Shannon-Wittaker-Kotelnikov sampling theorem [100]. Finite-length signal reconstruction with irregular discrete samples may be possible by imposing band-limited filters before sampling. A band-limited discrete signal can be exactly reconstructed from a finite number of arbitrarily spaced samples which can vary between samples. Here, the idea is to use the knowledge of the frequency content of the predicted signal to be able to use the reconstruction theorem for interpolation of data from unevenly spaced predictors.

Any periodic, M-band-limited discrete signal $\hat{f}(k)$ with period S, can be reconstructed exactly from $R \geq 2M+1$ samples $\hat{f}(k)$ where $\hat{f}(k) = \hat{f}(dT)$ with d is the integer sample spacing and kT mod S distinct. Ideally, $R = S/d$. The $R = 2M + 1$

irregular samples $\hat{f}(k)$ can be represented in matrix form as

$$\hat{f}(k) = \mathbf{D}_\delta \mathbf{a}$$

where $\mathbf{D}_\delta$ is the $R(2M+1)$ matrix constructed from the R rows of $\mathbf{D}$ corresponding to the distinct sample points $n_l \in \{0, ..., S-1\}$, $l = \{1, 2, ..., R\}$ with $\hat{f}(k)$ having the R elements $(\hat{f}(k))_l = f(n_l)$. The elements of the matrix $\mathbf{D}_\delta$ can be written explicitly as

$$(\mathbf{D}_\delta)_{l,m} = \mathbf{D}_{M,N}(n_l - md)$$

In order to compute the $2M+1$ values of **a** necessary to fully reconstruct the signal, D_δ^{-1} must exist. While it is possible to write D_δ^{-1} in closed form, the closed-form inverse is impractical for numerical computation. Instead, well-known conventional numerical inverse methods can be used to compute D_δ^{-1} or solve the corresponding linear system. Also, D_δ and D_δ^{-1} depend only on the sample locations and not the sample values. Thus, if the sample locations remain the same, multiple reconstructions can be accomplished with only one matrix inversion.

Since D_δ^{-1} exists, it is possible (at least theoretically) to compute

$$\mathbf{a} = \mathbf{D}_\delta^{-1} \hat{f}(k) \tag{5.5}$$

no matter what the precise values of the n_l are, so long as they are distinct. Once **a** is computed, it can then be used to reconstruct the signal using the discrete sampling equation. Reconstruction from irregular samples can thus be viewed as a two-step process: first compute the frequency coefficients of the signal using Eq. (5.5), then reconstruct the full signal. Once **a** is computed from the locations of the irregular samples, the reconstruction of $\hat{f}(k)$ is independent of the original sample locations.

Further, when there is no sampling noise, the **a** computed from either irregular or regular sampling is identical.

In the oversampled case where $R > 2M+1$, $\mathbf{D}_\delta$ is rectangular but remains full column rank. It is overdetermined and therefore has a unique pseudo-inverse. While the extra samples could be discarded to make $\mathbf{D}_\delta$ square, retaining all of the samples improves the performance in the presence of sampling noise [100]. When $R > 2M + 1$, the pseudo-inverse is used in Eq. (5.5) to compute **a**. Iterative approaches to matrix inversion or linear system solution can be very effective for reconstruction, and iterative methods are commonly used to refine matrix inverses computed using QR factorization or other methods. We note that iterative methods can also be used to compute approximate reconstructions. These are particularly useful for very large problems with thousands or millions of samples.

Another method of interpolation is the nearest neighborhood interpolation called the Voronoi-Allebach-algorithm [102]. In this method, given the sampling values at known positions a step function is created which is constant from midpoint to midpoint of the given sampling sequence. In the case where the number of missing points between subsequent sampling positions is odd, the midpoint itself is assigned the arithmetic mean of the two adjacent sampling points. Because of the assumed smoothness of the signal, this step function will not be too far away from the original signal, in the mean-squared sense. Having used the sampling values, the information concerning the band-width M can be used next. The step function is filtered, i.e. only those Fourier coefficients which belong to M are preserved, the rest is filtered. This method can be iteratively repeated to minimize the error between the spectral content of the actual signal and the reconstructed signal.

5.5 Recovery to Valid Data using Kalman Smoothing Filter

For the smooth transfer from predicted data to actual valid measurements, consider two sensors with different covariances. The system dynamics and the measurements can modeled by the following discrete-time state-space model:

$$\begin{aligned} x(k) &= A(k)x(k-1) + v(k) \\ f^j(k) &= C^j(k)x(k) + w^j(k) \end{aligned} \tag{5.6}$$

where k represents the discrete-time index, $j = \{1, 2\}$, $x(k)$ is the state-vector, $\{f^1(k), f^2(k)\}$ are measurement and prediction vectors respectively, $v(k)$ and $\{w_1(k), w_2(k)\}$ zero-mean white Gaussian noise with covariance matrices Q and $\{R_1, R_2\}$, respectively. By prediction and measurement fusion, the two models can be integrated into the following single model:

$$f(k) = C(k)x(k) + w(k) \tag{5.7}$$

where $w(k) \sim N(0, R(k))$. In this section, the two measurement fusion methods are formulated and their performance in state estimation are compared.

The measurement fusion method obtains the fused measurement and prediction information by weights as

$$\begin{aligned} f(k) &= [\sum_{j=1}^{N} W_j]^{-1}[\sum_{j=1}^{N} W_j f^j(k)] \\ C(k) &= [\sum_{j=1}^{N} W_j]^{-1}[\sum_{j=1}^{N} W_j C^j(k)] \\ R(k) &= [\sum_{j=1}^{N} R^{j^{-1}}]^{-1} \end{aligned} \tag{5.8}$$

where $\{W_1, W_2\}$ are the weights for measurement and prediction data respectively for $N = 2$.

Given the state-space model described by Eq. (5.5) and (5.5), the Kalman filter provides an unbiased and optimal estimate of the state-vector in the sense of minimum estimate covariance, which can be described by the following equations.

Prediction:

$$\begin{aligned} \hat{x}(k|k-1) &= A\hat{x}(k-1) \\ P(k|k-1) &= AP(k-1)A^T + Q \end{aligned} \tag{5.9}$$

Estimate (Correction):

$$\begin{aligned} \hat{x}(k) &= \hat{x}(k|k-1) + K(k)[f(k) - C(k)\hat{x}(k|k-1)] \\ K(k) &= P(k|k-1)C(k)^T[C(k)P(k|k-1)C^T(k) + R(k)]^{-1} \\ P(k) &= [I - K(k)C(k)]P(k|k-1) \end{aligned} \tag{5.10}$$

where $\hat{x}(k)$ represents the estimate of the state-vector $x(k)$, $P(k)$ is the state estimate covariance matrix, and $K(k)$ is the Kalman gain matrix.

5.6 Application of Filter Estimation and Prediction on Actual Measurements

For demonstration purposes, a single phasor amplitude $A(k)$, phasor $\phi(k)$ and frequency $f(k)$ data are collected from the San Diego Super Computer center of UCSD from a micro-PMU (μPMU). The Power Standards Lab (PSL) μPMU [23] includes a PQube instrument that contains measurement, recording, and communication functionalities along with a remotely mounted micro GPS receiver, and a power supply. These devices can be connected to single- or three phase secondary distribution circuits up to 690V (line-to-line) or 400V (line-to-neutral), either into standard outlets or through potential transformers (PTs). The devices continuously sample AC voltage and current waveforms at 256 or 512 samples per cycle and can produce 3 phase phasor data $X(k)$ at sampling rates of 60 or 120 Hz. In this chapter, for the purpose of illustration of the prediction and interpolation methods, only frequency data $f(k)$ measured at 60 $samples/sec$ are considered. The data was collected during the ambient period of the day i.e. when no events were detected on the event detection algorithm in [29]. The obtained ambient data thus containing no events ($d_f(k) = 0$) over the course of 48 hours at 60Hz sampling ($L \sim 10^6$ data points).

First, the complete set of frequency data $f(k)$ is used to compute the prediction filter for the 1-step ahead predictor, $H_f^1(q)$. Using 'L' samples of data, the filter $H_f^1(q)$ is estimated to minimize the least squares of the fit using the maximum likelihood estimation method. In this case, an $ARMA(p,q)$ method was used for 1-step ahead prediction error minimization using least squares. For $H_f^1(q)$, the values of (p,q) were (8,4). Any non-stationarity in the data can be removed using a differencing filter or by simply extending the estimation method to $ARIMA(p,d,q)$. Augmented Dickey Fuller test (ADF) was used to reject the null hypothesis of a unit root against the autoregressive alternative and auto correlation function (ACF) test. Fig. 5.3 shows the residuals, auto-correlation function (ACF) and partial auto-correlation function (PACF) by fitting 'L' samples with a $ARMA(6,2)$ filter. Now the goal is to find next $N-1$ filters to be able to obtain 1-step ahead predictors $\hat{f}_n(m+1)$ from each of such filter $H_f^n(q)$, $n = \{2,3,...,N\}$. As discussed in Sec. 5.1, there are several methods of choosing the downsampling factor. Two methods: (a) logarithmic downsampling factor and (b) spectrum based downsampling factor are illustrated in this chapter. First for logarithmic sampling, the down-sample factors are chosen as

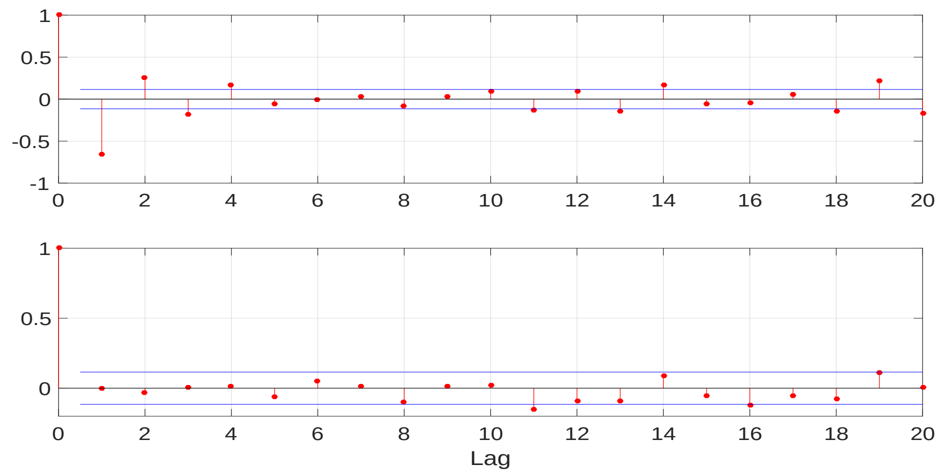

Figure 5.3: Illustration of autocorrelation test (with 95% bounds) to validate the estimated $ARMA$ model: (a) $ARMA$ (5,2) (b) $ARMA$ (6,3). It can be clearly seen that $ARMA$ (6,3) passes the whiteness test.

$n = 2^b$ i.e. $n = \{2, 4, 8, 16, ..., 2^{15}\}$. In this case, the data $f(k)$, $k = \{1, 2, 3, ..., L\}$ is first filtered through anti-aliasing filter $AF_2(q)$ before downsampling by the factor of 2 resulting in $f_2(m)$, $m = \{1, 2, 4, ...\}$. Now, $f_2(m)$ is used to compute the prediction filter for the 1-step ahead predictor, $H_{1f}^{2}(q)$. Using '$L/2$' samples of data, the filter $H_{1f}^{2}(q)$ is estimated to minimize the least squares of the fit using the maximum likelihood estimation method. In this case, an $ARMA$ (6,3) method was used for 1-step ahead prediction error minimization using least squares. This process (filtering, downsampling, least square minimization) was repeated until all the 16 filters are computed. For case (b), the down-sample factors are chosen with weighting to capture the spectral power of $f(k)$ as seen in Sec. 5.1. In this case, the downsampling factor was $n = 2^b$ i.e. $p = \{2, 3, 4, 2^5, 2^6, 2^7, 2^8, 2^9, 2^{10}, 2^{11}, 3000, 8192, 2^{14}, 25000, 2^{15}\}$. Repeating the process as in (a), the end result would be 16 estimated filters $\{H_{2f}^{1}(q), H_{2f}^{2}(q), H_{2f}^{3}(q), ...H_{2f}^{16}(q)\}$. Once the filters are determined, 1-step ahead forecasts as discussed in Sec. 5.3.2 are computed resulting in the prediction vector $\hat{f}_1 = [\hat{f}_1(m+1), \hat{f}_2(m+1), \hat{f}_4(m+1), ..., \hat{f}_{15}(m+1)]$ in case (a) and $\hat{f}_2 =$

$[\hat{f}_1(m+1), \hat{f}_2(m+1), \hat{f}_3(m+1), ..., \hat{f}_{15}(m+1)]$ in case (b). The resulting predictors $\hat{f}_1$ and $\hat{f}_2$ are illustrated in Fig. 5.4.

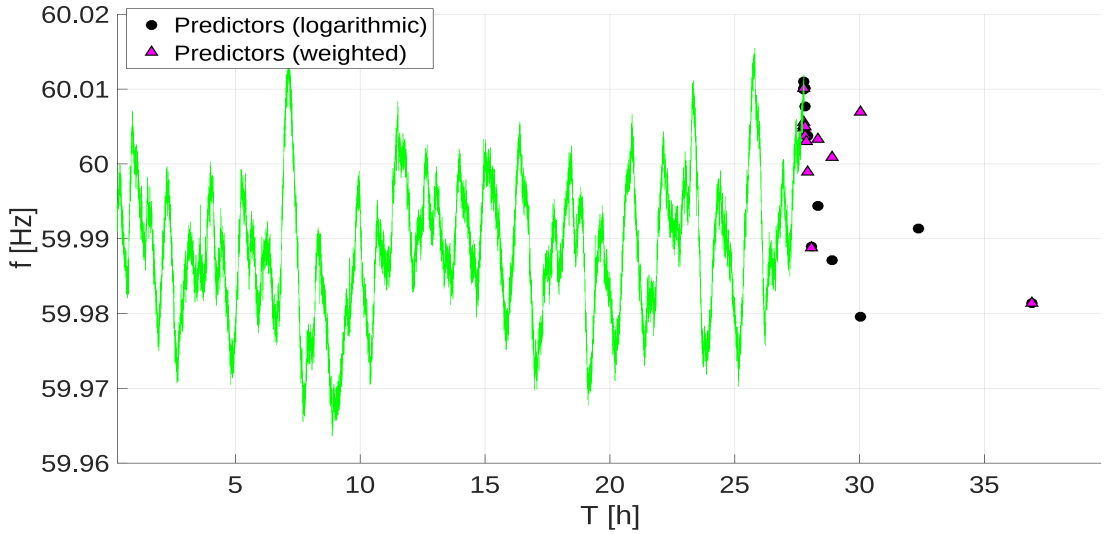

Figure 5.4: Resulting predictors from filters estimated using logarithmic and spectrally weighted downsampling factors.

Once the predictors are obtained, the next task is to interpolate between the non-uniformly spaced predictors to obtain uniform sampled (interpolated) predictors at 60 Hz sampling to be able to be used for applications such as missing data etc. In this chapter, two methods have been proposed for interpolation as discussed in Sec. 5.3.3. The first method is using the band-limited signal reconstruction from irregular samples using the Shannon-Wittaker-Kotelnikov sampling theorem discussed in Sec. 5.3.3. In this method, the frequency domain information of $f(k)$ is used to reconstruct the time domain data using the implementation of one-dimensional irregularly-spaced samples with using the approach and notation described in [100]. However, this is method is implemented in a piece-wise manner to ensure that the number of samples available for reconstruction $R \geq 2M + 1$ where $f(k)$ is piece-wise M-band-limited. This method is applied to the predicted data $\hat{f}_1$ and $\hat{f}_2$ obtained from the prediction filters $H_{1f}^{p}(q)$ and $H_{2f}^{p}(q)$. The results are illustrated in figures Fig. 5.6 and 5.7 in the dashed line (spectral).

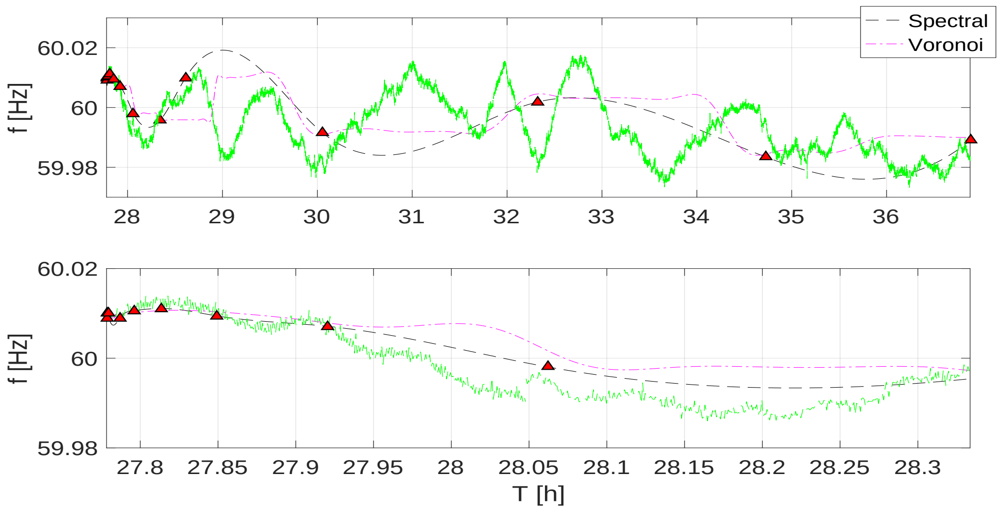

Figure 5.5: Interpolation using Spectral and Voronoi methods for predictors for non-uniform one step ahead predictors with sampling rates determined by spectrally weighted downsampling: (a) complete prediction data, (b) zoom-in version at the beginning of prediction.

The second method is using the available (assumed) knowledge of the signal bandwidth starting with nearest neighborhood interpolation. Given the sampling values at known positions a step function is formed which is constant from midpoint to midpoint of the given sampling sequence. In the case where the number of missing points between subsequent sampling positions is odd, the midpoint itself is assigned the arithmetic mean of the two adjacent sampling points. Because of the assumed smoothness of the signal, this step function will not be too far away from the original signal, in the mean-squared sense. Having used the sampling values, the information concerning the bandwidth can be used next. The step function (which apparently contains many high frequencies outside of M) is filtered, i.e. only those Fourier coefficients which belong to M are preserved, the rest is discarded. Fig. 5.6 and 5.7 show the filtered version of the step-function, the original sampling points plotted with markers. This filtering process can also be made iterative to minimize the er-

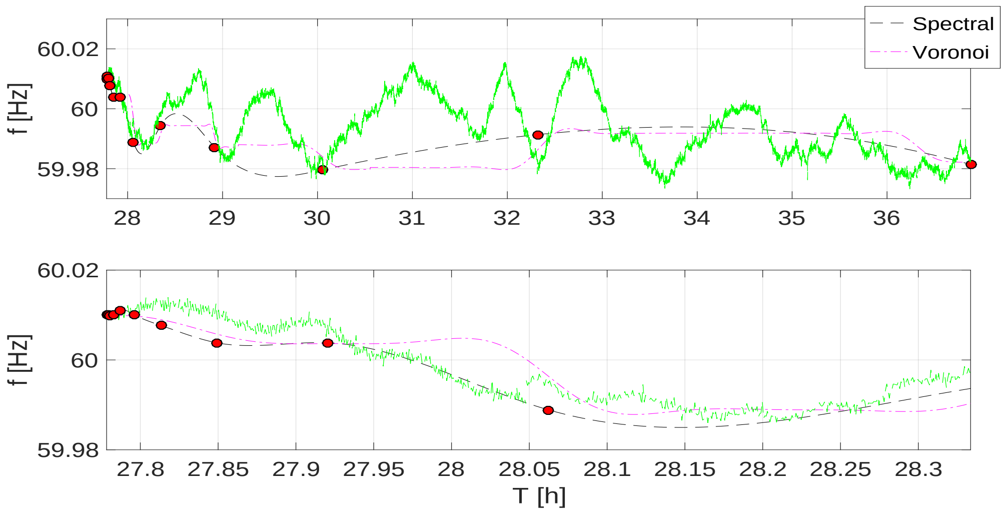

Figure 5.6: Interpolation using Spectral and Voronoi methods for non-uniform one step ahead predictors using a logarithmic distribution of downsampling rates: (a) complete prediction data, (b) zoom-in version at the beginning of prediction.

rors between the spectrum of the training and testing data. In this method too, the frequency domain information of $f(k)$ is used to reconstruct the time domain data using the implementation of one-dimensional irregularly-spaced samples. However, the filters are designed in a piece-wise manner with decreasing cut-off frequencies with increasing p to be able to predict at different ranges of the spectrum. This method is applied to the predicted data $\hat{f}_1$ and $\hat{f}_2$ obtained from the prediction filters ${H_1}_f^p(q)$ and ${H_2}_f^p(q)$. The results are illustrated in figures Fig. 5.6 and Fig. 5.7 in the dashed line (Voronoi).

The bottom plots in Fig. 5.6 and Fig. 5.7 illustrate the prediction for small p to show the goodness of prediction up to a certain prediction range. However, the variance of prediction increases as the prediction range increases however capturing the slower dynamic changes in the signal $f(k)$. It can also be observed that the spectral based interpolation gives a better prediction compared to Voronoi interpolation for both prediction methods. It is important to note that for both prediction methods, fixed filters ${H^1}_f(q)$ and ${H^2}_f(q)$ derived from a given set of training data are used to generate 1-step ahead predictors to reduce computational time. However, it is trivial to recursively estimate the $H^1{}_f(q)(k)$ and $H^2{}_f(q)(k)$ with new incoming data or after every 'u' samples. This might be computationally expensive as a trade-off for better predictions.

For this prediction of phasor data to be viable for real-time control, it is important to make a smooth transition from the predicted data to actual valid measurements. This is important to avoid any undesired jumps in the control signal due to a sudden change of data streams. The method involving fusion of multiple data streams using an output weighted-Kalman filter is discussed in Sec. 5.4. Two independent streams of samples: predicted and actual are considered as two sensor measurements for the Kalman filter design. It is clear that the covariance of the predicted data is always greater than the covariance of the measured data, the weights W_1, $W_2 = 1-W_1$ are considered on predicted and measured data respectively, to ensure a smooth transition. For the demonstration of the use of Kalman filter three different values of weight W_1 are considered. The choice of W_1 dictates the rate of transition from the predicted data to valid measurements. A smaller value of W_1 means that the transition or convergence of the filtered signal to actual measurements is faster. Fig. 5.7 shows the output of the Kalman filter to be used for real-time control applications with three different choices of weight W_1: $W_1 = 0.3$, $W_1 = 0.4$ and $W_1 = 0.5$ from the point where the measurement data is valid. It can be seen that the Kalman filter with weight $W_1 = 0.3$ converges faster than $W_1 = 0.4$ and $W_1 = 0.5$.

The overview and the logical steps of filtering, prediction and interpolation methods that can be implemented in real-time extrapolation of missing PMU data is shown in Fig. 5.8. First, data of length 'L' samples is stored as a vector. Next, this data is downsampled $p = \{1, 2, ...P\}$ times after filtering and used for estimation of $H^p{}_f(q)$ filters, $p = \{1, 2, ...P\}$. The estimated filters can now be used for 1-step

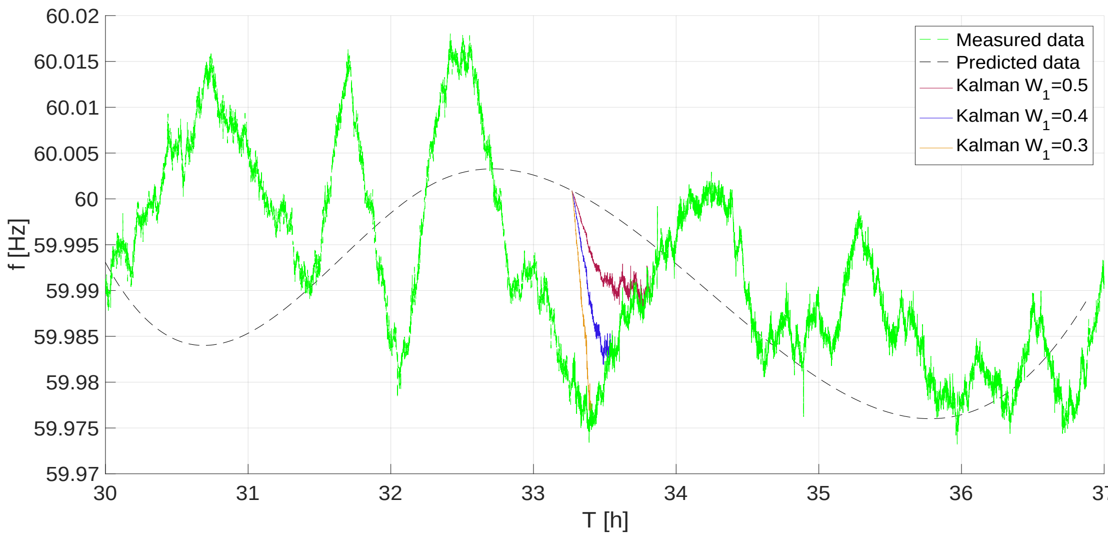

Figure 5.7: Comparison of Kalman based transitioning of data for real-time control applications for three different weights on prediction and measurement data.

ahead predictions with the data of length 'L'. However, the filters may be recomputed every $k > T_{wait}$ s to account for the change in the spectral content over time k. Once the predictors are obtained, one of the interpolation methods can be used to result in uniformly sampled prediction data (every 60 $samples/s$ in case of grid frequency data used in this chapter). The missing data can be replaced with the obtained predicted (and interpolated) data until the PMU measurements are valid. Once the PMU data is valid again, actual measurements can be instead used in place of predicted data. For use of real-time predicted data for control purposes, this data can now be filtered using Kalman filter for smooth transition from predicted to actual data. If the missing data is longer than a certain period $k > T_{miss}$ the variance on predicted data increases and is no longer a good source of information to be used for control. This however, can be complemented using SCADA measurements if available.

The prediction filters are system dependent that varies according to the dynamics of the grid, filtering by the PMU, events such as generation or load loss as seen in the grid and effect of external control systems such as droop control or auto-

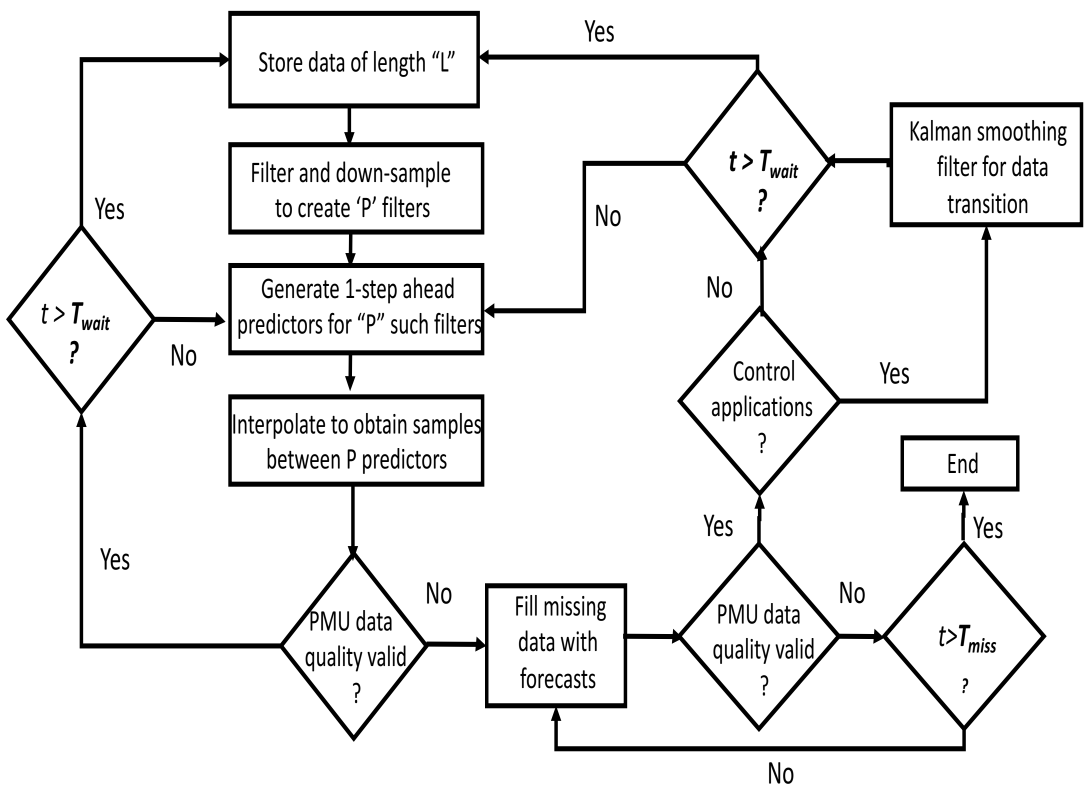

Figure 5.8: Flowchart illustrating the logic flow of steps to replace missing PMU data.

matic voltage regulation (AVR) systems. Hence, the prediction filters may need to change with time by recursively estimating the filters $H_f^p(q)$ with time. The order of this filter can be fixed for a particular point of measurement for a given PMU as the order is insensitive to the minor changes in the grid dynamics. Moreover, the filter coefficients may be sensitive to the grid dynamics hence require online estimation.

5.7 Summary of Contribution

Most of the applications of synchronized phasor measurements have been mainly concerning power system model validation, event analysis (detection, classification and localization), real-time display, etc. However, synchrophasors have a much greater potential than just monitoring and visualization such as reliable and economical operation of power systems in real-time control and protection schemes. However, the use of syncrophasors in control is prone to missing measurements. Hence, a method to predict the phasor data in real-time for short periods based on the historical measurements becomes important. This chapter describes a spectral based phasor data prediction and recovery algorithm that replaces invalid synchrophasor data and transitions smoothly to valid data using a Kalman filter based approach. The drawback of the method is that it becomes unreasonable after long periods of missing data and hence a data fall-back such as SCADA might be used for control purposes in such scenarios.

The text and data in Chapter 5, in full, is a reprint of the material as it appears in "Spectrum Based Optimal Filtering for Short-term Phasor Data Prediction", Sai Akhil R. Konakalla, Raymond A. de Callafon, *submitted to IEEE Transactions on on Industry Applications* and "Sai Akhil R. Konakalla, and Raymond de Callafon, Robustness to missing synchrophasor data for power and frequency event detection in electric grids, *Renewable Energy Research and Applications (ICRERA), 2017 IEEE 6th International Conference on. IEEE, 2017*". The dissertation author is the primary investigator and author of these articles.

Chapter 6

Conclusion

The thesis describes the use of synchrophasors in a wide range of applications involving electric grid monitoring, demand prediction, dispatch optimization and control. It also addresses main problems in using synchrophasor data for these variety of applications.

The second chapter shows how the optimal filters, which can be used to formulate filtered rate of change phasor signals, can be estimated on the basis of phasor data where no disturbances are present via a straight-forward least squares optimization. The resulting variance bounds give rise to event detection algorithms that check for the number of consecutive samples for which the filtered phasor signals are outside the variance bounds. It is also shown that combining estimation of dynamic event parameters combined with standard clustering methods can be used to build classifiers to quantify power grid oscillation events on synchrophasor data.

The third chapter shows how real-time synchrophasor data can be used for recursive prediction of dynamic power demand of a system and how these predictions are used for real-time demand optimization for economic benefits. A dynamic prediction method is formulated using a recursive Kalman filter that uses phasor data to provide real-time updated load forecasts. The power dispatch algorithm discussed aims to achieve maximal economic benefits mainly by demand reduction and energy arbitrage subject to the resource constraints and operational limits of a load, while using recursively predicted demand data.

The fourth chapter discusses detailed switching and control algorithm for a microgrid consisting of multiple distributed energy resources including both synchronous generators and inverter-based DERs. Modeling of the plant for control purposes is performed by model estimation techniques using DER inputs, PMU measurements, and realization algorithms. The developed control algorithm is tested on a HIL setup that models the microgrid and proves fast and smooth transition. Dispatch algorithms to control the microgrids with hybrid DERs are also discussed in the third chapter.

In the fifth chapter, a method to predict the phasor data in real-time for short periods based on the historical measurements is discussed. A spectral based phasor data prediction and recovery algorithm that replaces invalid synchrophasor data and transitions smoothly to valid data using a Kalman filter based approach is described.

There are a numerous applications or contributions that can be easily extended from the scope of this thesis. In the first chapter, event detection and classification are discussed. However, event localization is an other important step to be able to fully understand the event to be able to take a control action. Hence, event localization is a future work or contribution of this thesis. Similarly, optimal control algorithms can be designed to formulate both the dispatch and control sub-problems into one optimal control problem. Finally, a complete method to be able to detect and react upon grid events (example, island a microgrid) can be considered as one of the immediate future scope of this thesis.

Bibliography

[1] A. Allen, M. Singh, E. Muljadi, PMU data event detection: A user guide for power engineers, Tech. rep., NREL/TP-5D00-61664 (2014).

[2] O. Ardakanian, Y. Yuan, R. Dobbe, A. von Meier, S. Low, C. Tomlin, Event detection and localization in distribution grids with phasor measurement units, arXiv preprint arXiv:1611.04653.

[3] M. Biswal, S. Brahma, H. Cao, Supervisory protection and automated event diagnosis using pmu data.

[4] F. Aminifar, M. Fotuhi-Firuzabad, A. Safdarian, M. Shahidehpour, Synchrophasor measurement technology in power systems: Panorama and state-of-the-art, IEEE Access 2 (2015) 1607 – 1628.

[5] J. Roberts, D. Tziouvaras, Fault type selection system for identifying faults in an electric power system, U.S. Patent 6,525,543.

[6] J. Liu, G.and Quintero, V. Venkatasubramanian, Oscillation monitoring system based on wide area synchrophasors in power systems, in: Proc. iREP Symposium on Bulk Power System Dynamics and Control, 2007, pp. 1 – 13.

[7] A. Tiwari, V. Ajjarapu, Event identification and contingency assessment for voltage stability via PMU, in: Proc. 39th North American Power Symposium, 2007, pp. 413 – 420.

[8] J. Tate, Event detection and visualization based on phasor measurement units for improved situational awareness, Ph.D. thesis, University of Illinois at Urbana-Champaign (2008).

[9] P. Samadi, H. Mohsenian-Rad, V. W. Wong, R. Schober, Tackling the load uncertainty challenges for energy consumption scheduling in smart grid, IEEE Transactions on Smart Grid 4 (2) (2013) 1007–1016.

[10] S. A. R. Konakalla, R. A. de Callafon, Feature based grid event classification from synchrophasor data, Procedia Computer Science 108 (2017) 1582–1591.

[11] E. Schweitzer, A. Guzmán, H. Altuve, D. Tziouvaras, J. Needs, Real-time synchrophasor applications in power system control and protection, in: Proc. 10th IET International Conference on Developments in Power System Protection, 2010, pp. 1–5.

[12] E. Schweitzer, D. Whitehead, A. Guzmán, Y. Gong, M. Donolo, Advanced real-time synchrophasor applications, SEL Journal of Reliable Power 2.

[13] P. Castello, P. Ferrari, A. Flammini, C. Muscas, S. Rinaldi, A new IED with PMU functionalities for electrical substations, IEEE Trans. Instrum. Meas. 62 (2013) 32093217.

[14] E. Schweitzer, A. Guzmán, Synchrophasor processor detects out-of-step conditions, in: IEEE International Conference on Smart Grid Communications, 2011, pp. 576 – 581.

[15] F. Tobar, L. Yacher, R. Paredes, M. Orchard, Anomaly detection in power generation plants using similarity-based modeling and multivariate analysis, in: Proc. American Control Conference, San Francisco, CA, USA, 2011, pp. 1940–1945.

[16] L. Xie, Y. Chen, P. Kumar, Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis, IEEE Trans. on Power Systems 29 (2014) 2784–2794.

[17] L. Schenato, G. Barchi, D. Macii, R. Arghandeh, K. Poolla, A. von Meier, Bayesian linear state estimation using smart meters and PMUs measurements in distribution grids, in: Proc. IEEE International Conference on Smart Grid Communications, 2014, pp. 572–577.

[18] P. Tripath, S. Srivastava, S. Singh, An improved prony method for identifying low frequency oscillations using synchro-phasor measurements, in: Proc. International Conference on Power Systems, 2009, pp. 1–5.

[19] V. Salehi, A. Mazloomzadeh, O. Mohammed, Development and implementation of a phasor measurement unit for real-time monitoring, control and protection of power systems, in: PRoc. Power and Energy Society General Meeting, 2011, pp. 1–7.

[20] IEEE, C37.118.2-2011 – standard for synchrophasor data transfer for power systems, Tech. rep., IEEE Power and Energy Society (2011).

[21] K. Martin, An overview of the IEEE standard C37.118.2 – synchrophasor data transfer for power systems, IEEE Trans. on Smart Grid 5 (2014) 1980–1984.

[22] T. Lobos, J. Rezmer, Real-time determination of power system frequency, IEEE Trans. Instrum. Meas. 46 (1997) 877–881.

[23] A. von Meier, D. Culler, A. McEachern, R. Arghandeh, Micro-synchrophasors for distribution systems, in: Proc. IEEE Innovative Smart Grid Technologies Conference, 2014.

[24] R. Buyya, R. N. Calheiros, A. V. Dastjerdi, Big Data: Principles and Paradigms, Morgan Kaufmann, 2016.

[25] Y. Chen, L. Xie, P. R. Kumar, Power system event classification via dimensionality reduction of synchrophasor data, in: 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM), IEEE, 2014, pp. 57–60.

[26] D. Nguyen, R. Barella, S. A. Wallace, X. Zhao, X. Liang, Smart grid line event classification using supervised learning over pmu data streams, in: Green Computing Conference and Sustainable Computing Conference (IGSC), 2015 Sixth International, IEEE, 2015, pp. 1–8.

[27] H. Saadat, Power system analysis, WCB/McGraw-Hill, 1999.

[28] D.-I. Kim, T. Y. Chun, S.-H. Yoon, G. Lee, Y.-J. Shin, Wavelet-based event detection method using pmu data.

[29] S. A. R. Konakalla, R. de Callafon, Optimal filtering for grid event detection from real-time synchrophasor data, Procedia Computer Science 80 (2016) 931–940.

[30] L. Ljung, System Identification - theory for the user, Prentice Hall, Upper Saddle River, NJ, 1999.

[31] G. W. Stagg, A. H. El-Abiad, Computer methods in power system analysis, McGraw-Hill, 1968.

[32] O. Antoine, J.-C. Maun, Inter-area oscillations: Identifying causes of poor damping using phasor measurement units, in: 2012 IEEE Power and Energy Society General Meeting, IEEE, 2012, pp. 1–6.

[33] K. Åšträm, Introduction to Stochastic Control Theory, Academic Press, New York, 1970.

[34] F. Schleif, J. White, Damping for the northwest-southwest tieline oscillations-an analog study, IEEE Transactions on Power Apparatus and Systems (12) (1966) 1239–1247.

[35] R. A. De Callafon, B. Moaveni, J. P. Conte, X. He, E. Udd, General realization algorithm for modal identification of linear dynamic systems, Journal of engineering mechanics 134 (9) (2008) 712–722.

[36] Y. Yuan, Y. Sun, L. Cheng, G. Chen, P. Wang, Power system low frequency oscillation monitoring and analysis based on multi-signal online identification, Science China Technological Sciences 53 (2010) 2589–2596.

[37] G. Tardioli, R. Kerrigan, M. Oates, O. James, D. Finn, Data driven approaches for prediction of building energy consumption at urban level, Energy Procedia 78 (2015) 3378–3383.

[38] V. Cherkassky, S. R. Chowdhury, V. Landenberger, S. Tewari, P. Bursch, Prediction of electric power consumption for commercial buildings, in: Neural Networks (IJCNN), The 2011 International Joint Conference on, IEEE, 2011, pp. 666–672.

[39] A. Phadke, B. Pickett, M. Adamiak, M. Begovic, G. Benmouyal, R. Burnett, T. Cease, J. Goossens, D. Hansen, M. Kezunovic, et al., Synchronized sampling and phasor measurements for relaying and control, IEEE Transactions on Power Delivery 9 (1) (1994) 442–452.

[40] J. Hinman, E. Hickey, Modeling and forecasting short-term electricity load using regression analysis, Journal of IInstitute for Regulatory Policy Studies [ ].

[41] R. Ramanathan, R. Engle, C. W. Granger, F. Vahid-Araghi, C. Brace, Short-run forecasts of electricity loads and peaks, International journal of forecasting 13 (2) (1997) 161–174.

[42] R. Cottet, M. Smith, Bayesian modeling and forecasting of intraday electricity load, Journal of the American Statistical Association 98 (464) (2003) 839–849.

[43] J. Yoo, K. Hur, Load forecast model switching scheme for improved robustness to changes in building energy consumption patterns, Energies 6 (3) (2013) 1329–1343.

[44] K. Nur'aini, I. Najahaty, L. Hidayati, H. Murfi, S. Nurrohmah, Combination of singular value decomposition and k-means clustering methods for topic detection on twitter, in: Advanced Computer Science and Information Systems (ICACSIS), 2015 International Conference on, IEEE, 2015, pp. 123–128.

[45] C.-L. Hor, S. J. Watson, S. Majithia, Analyzing the impact of weather variables on monthly electricity demand, IEEE transactions on power systems 20 (4) (2005) 2078–2085.

[46] R. E. Kalman, et al., A new approach to linear filtering and prediction problems, Journal of basic Engineering 82 (1) (1960) 35–45.

[47] P. R. Kumar, P. Varaiya, Stochastic systems: Estimation, identification, and adaptive control, SIAM, 2015.

[48] A.-H. Mohsenian-Rad, V. W. Wong, J. Jatskevich, R. Schober, A. Leon-Garcia, Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid, IEEE transactions on Smart Grid 1 (3) (2010) 320–331.

[49] R. Hanna, J. Kleissl, A. Nottrott, M. Ferry, Energy dispatch schedule optimization for demand charge reduction using a photovoltaic-battery storage system with solar forecasting, Solar Energy 103 (2014) 269–287.

[50] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, J. Lavaei, A survey of distributed optimization and control algorithms for electric power systems, IEEE Transactions on Smart Grid 8 (6) (2017) 2941–2962.

[51] J. Eyer, G. Corey, Energy storage for the electricity grid: Benefits and market potential assessment guide, Sandia National Laboratories 20 (10) (2010) 5.

[52] P. Jacquot, O. Beaude, S. Gaubert, N. Oudjane, Analysis and implementation of an hourly billing mechanism for demand response management, IEEE Transactions on Smart Grid.

[53] A. Foucquier, S. Robert, F. Suard, L. Stéphan, A. Jay, State of the art in building modelling and energy performances prediction: A review, Renewable and Sustainable Energy Reviews 23 (2013) 272–288.

[54] A. Ahmad, M. Hassan, M. Abdullah, H. Rahman, F. Hussin, H. Abdullah, R. Saidur, A review on applications of ann and svm for building electrical energy consumption forecasting, Renewable and Sustainable Energy Reviews 33 (2014) 102–109.

[55] G. K. Tso, K. K. Yau, Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks, Energy 32 (9) (2007) 1761–1768.

[56] H. R. Khosravani, M. D. M. Castilla, M. Berenguel, A. E. Ruano, P. M. Ferreira, A comparison of energy consumption prediction models based on neural networks of a bioclimatic building, Energies 9 (1) (2016) 57.

[57] J. V. Tsyganova, M. V. Kulikova, Svd-based kalman filter derivative computation, IEEE Transactions on Automatic Control 62 (9) (2017) 4869–4875.

[58] E. A. Feinberg, D. Genethliou, Load forecasting, in: Applied mathematics for restructured electric power systems, Springer, 2005, pp. 269–285.

[59] G. A. Darbellay, M. Slama, Forecasting the short-term demand for electricity: Do neural networks stand a better chance?, International Journal of Forecasting 16 (1) (2000) 71–83.

[60] C. Charalambous, A. Conn, An efficient method to solve the minimax problem directly, SIAM Journal on Numerical Analysis 15 (1) (1978) 162–187.

[61] R. B. Kearfott, Interval computations: Introduction, uses, and resources, Euromath Bulletin 2 (1) (1996) 95–112.

[62] E. Kyriakides, M. Polycarpou, Short term electric load forecasting: A tutorial, in: Trends in Neural Computation, Springer, 2007, pp. 391–418.

[63] A. G. Phadke, J. S. Thorp, Synchronized phasor measurements and their applications, Vol. 1, Springer, 2017.

[64] H. Wang, J. Huang, Incentivizing energy trading for interconnected microgrids, IEEE Transactions on Smart Grid 9 (4) (2018) 2647–2657.

[65] N. Agrawal, K. Swarnkar, D. S. Wadhwani, D. A. Wadhwani, Economic load dispatch problem of thermal generators with ramp rate limit using biogeography based optimization, International Journal of Engineering and Innovative Technology (IJEIT) 1 (3) (2012) 98–133.

[66] Z. Li, W. Wu, B. Zhang, B. Wang, Decentralized multi-area dynamic economic dispatch using modified generalized benders decomposition, IEEE Transactions on Power Systems 31 (1) (2016) 526–538.

[67] N. J. Williams, P. Jaramillo, J. Taneja, Pv-array sizing in hybrid diesel/pv/battery microgrids under uncertainty, in: PowerAfrica, 2016 IEEE PES, IEEE, 2016, pp. 189–193.

[68] C. D. Barley, C. B. Winn, Optimal dispatch strategy in remote hybrid power systems, Solar Energy 58 (4-6) (1996) 165–179.

[69] Y. Lin, P. Barooah, J. L. Mathieu, Ancillary services to the grid from commercial buildings through demand scheduling and control, in: American Control Conference (ACC), 2015, IEEE, 2015, pp. 3007–3012.

[70] Y. Ma, A. Kelman, A. Daly, F. Borrelli, Predictive control for energy efficient buildings with thermal storage: Modeling, stimulation, and experiments, IEEE Control Systems 32 (1) (2012) 44–64.

[71] W. F. Pickard, D. Abbott, Addressing the intermittency challenge: Massive energy storage in a sustainable future, Proceedings of the IEEE 100 (2) (2012) 317.

[72] L. Maharjan, S. Inoue, H. Akagi, J. Asakura, State-of-charge (soc)-balancing control of a battery energy storage system based on a cascade pwm converter, IEEE Transactions on Power Electronics 24 (6) (2009) 1628–1636.

[73] J. Qin, R. Rajagopal, Dynamic programming solution to distributed storage operation and design, in: Power and Energy Society General Meeting (PES), 2013 IEEE, IEEE, 2013, pp. 1–5.

[74] E. Lavretsky, K. A. Wise, Optimal control and the linear quadratic regulator, in: Robust and Adaptive Control, Springer, 2013, pp. 27–50.

[75] R. Kuffel, J. Giesbrecht, T. Maguire, R. Wierckx, P. McLaren, Rtds-a fully digital power system simulator operating in real time, in: WESCANEX 95. Communications, Power, and Computing. Conference Proceedings., IEEE, Vol. 2, IEEE, 1995, pp. 300–305.

[76] S.-H. Ko, S. R. Lee, H. Dehbonei, C. V. Nayar, Application of voltage-and current-controlled voltage source inverters for distributed generation systems, IEEE Transactions on Energy Conversion 21 (3) (2006) 782–792.

[77] K. Yeager, J. Willis, Modeling of emergency diesel generators in an 800 megawatt nuclear power plant, IEEE Transactions on Energy Conversion 8 (3) (1993) 433–441.

[78] S. Jain, V. Agarwal, A single-stage grid connected inverter topology for solar pv systems with maximum power point tracking, IEEE transactions on power electronics 22 (5) (2007) 1928–1940.

[79] H. Han, X. Hou, J. Yang, J. Wu, M. Su, J. M. Guerrero, Review of power sharing control strategies for islanding operation of ac microgrids, IEEE Transactions on Smart Grid 7 (1) (2016) 200–215.

[80] T. Vandoorn, J. De Kooning, B. Meersman, L. Vandevelde, Review of primary control strategies for islanded microgrids with power-electronic interfaces, Renewable and Sustainable Energy Reviews 19 (2013) 613–628.

[81] A. Mohd, E. Ortjohann, D. Morton, O. Omari, Review of control techniques for inverters parallel operation, Electric Power Systems Research 80 (12) (2010) 1477–1487.

[82] J. P. Lopes, C. Moreira, A. Madureira, Defining control strategies for microgrids islanded operation, IEEE Transactions on power systems 21 (2) (2006) 916–924.

[83] N. Pogaku, M. Prodanovic, T. C. Green, Modeling, analysis and testing of autonomous operation of an inverter-based microgrid, IEEE Transactions on power electronics 22 (2) (2007) 613–625.

[84] D. Miller, R. de Callafon, Identification of linear time-invariant systems via constrained step-based realization, in: Prepr. 16th IFAC Symposium on System Identification, IFAC, 2012, pp. 1155–1160.

[85] R. Majumder, A. Ghosh, G. Ledwich, F. Zare, Control of parallel converters for load sharing with seamless transfer between grid connected and islanded modes, in: Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE, IEEE, 2008, pp. 1–7.

[86] S. Arefifar, C. Almeida, W. Xu, Reclosing impact on power distribution systems with distributed generators.

[87] I. S. C. Committee, et al., Ieee standard for interconnecting distributed resources with electric power systems, IEEE Std1547-2003.

[88] S. D'Arco, J. A. Suul, A synchronization controller for grid reconnection of islanded virtual synchronous machines, in: Power Electronics for Distributed Generation Systems (PEDG), 2015 IEEE 6th International Symposium on, IEEE, 2015, pp. 1–8.

[89] N. Li, L. Chen, S. H. Low, Optimal demand response based on utility maximization in power networks, in: Power and Energy Society General Meeting, 2011 IEEE, IEEE, 2011, pp. 1–8.

[90] M. Caramia, P. Dell'Olmo, Multi-objective management in freight logistics: Increasing capacity, service level and safety with optimization algorithms, Springer Science & Business Media, 2008.

[91] G. G. Moshi, C. Bovo, A. Berizzi, Optimal operational planning for pv-wind-diesel-battery microgrid, in: PowerTech, 2015 IEEE Eindhoven, IEEE, 2015, pp. 1–6.

[92] R. A. de Callafon, A. H. Habib, J. Kleissl, Scheduling of dynamic electric loads using energy storage and short term power forecasting, in: Control Applications (CCA), 2016 IEEE Conference on, IEEE, 2016, pp. 1531–1536.

[93] A. Valibeygi, R. de Callafon, M. Stanovich, M. Sloderbeck, K. Schoder, J. Langston, I. Leonard, S. Chatterjee, R. Meeker, Microgrid control using remote controller hardware-in-the-loop over the internet, arXiv preprint arXiv:1804.05998.

[94] M. He, V. Vittal, J. Zhang, Online dynamic security assessment with missing pmu measurements: A data mining approach, IEEE Transactions on Power Systems 28 (2) (2013) 1969–1977.

[95] M. Brown, M. Biswal, S. Brahma, S. J. Ranade, H. Cao, Characterizing and quantifying noise in pmu data, in: Power and Energy Society General Meeting (PESGM), 2016, IEEE, 2016, pp. 1–5.

[96] R. Kohn, C. F. Ansley, Estimation, prediction, and interpolation for arima models with missing data, Journal of the American Statistical Association 81 (395) (1986) 751–761.

[97] S. A. R. Konakalla, R. de Callafon, Robustness to missing synchrophasor data for power and frequency event detection in electric grids, in: Renewable Energy Research and Applications (ICRERA), 2017 IEEE 6th International Conference on, IEEE, 2017, pp. 808–812.

[98] F. Eng, F. Gustafsson, Algorithms for downsampling non-uniformly sampled data, in: Signal Processing Conference, 2007 15th European, IEEE, 2007, pp. 1965–1969.

[99] Z. Song, B. Liu, Y. Pang, C. Hou, X. Li, An improved nyquist–shannon irregular sampling theorem from local averages, IEEE Transactions on Information Theory 58 (9) (2012) 6093–6100.

[100] D. G. Long, R. O. Franz, Band-limited signal reconstruction from irregular samples with variable apertures, IEEE Transactions on Geoscience and Remote Sensing 54 (4) (2016) 2424–2436.

[101] K. M. Flornes, Y. Lyubarskii, K. Seip, A direct interpolation method for irregular sampling, Applied and Computational Harmonic Analysis 7 (3) (1999) 305–314.

[102] T. Werther, A first guided tour on the irregular sampling problem, Dept. of Mathematics, University of Vienna. http://www. unet. univie. ac. at/ a9301647/fsp/irsatour/irsampl. html.

[103] J. Zhang, G. Welch, G. Bishop, Z. Huang, A two-stage kalman filter approach for robust and real-time power system state estimation, IEEE Transactions on Sustainable Energy 5 (2) (2014) 629–636.

[104] Q. Gan, C. J. Harris, Comparison of two measurement fusion methods for kalman-filter-based multisensor data fusion, IEEE Transactions on Aerospace and Electronic systems 37 (1) (2001) 273–279.

[105] G. Haines, A. G. Jones, Logarithmic fourier transformation, Geophysical Journal 92 (1) (1988) 171–178.

[106] J. Zhao, Y. Zhu, R. Patwardhan, Identification of k-step-ahead prediction error model and mpc control, Journal of Process Control 24 (1) (2014) 48–56.