

MULTI-AGENT ESTIMATION AND CONTROL OF
CYBER-PHYSICAL SYSTEMS

by

S M SHAFIUL ALAM

B.S., Bangladesh University of Engineering and Technology, 2008

M.S., Bangladesh University of Engineering and Technology, 2011

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Abstract

A cyber-physical system (CPS) typically consists of networked computational elements that control physical processes. As an integral part of CPS, the widespread deployment of communicable sensors makes the task of monitoring and control quite challenging especially from the viewpoint of scalability and complexity. This research investigates two unique aspects of overcoming such barriers, making a CPS more robust against data explosion and network vulnerabilities. First, the correlated characteristics of high-resolution sensor data are exploited to significantly reduce the fused data volume. Specifically, spatial, temporal and spatiotemporal compressed sensing approaches are applied to sample the measurements in compressed form. Such aggregation can directly be used in centralized static state estimation even for a nonlinear system. This approach results in a remarkable reduction in communication overhead as well as memory/storage requirement. Secondly, an agent based architecture is proposed, where the communicable sensors (identified as agents) also perform local information processing. Based on the local and underdetermined observation space, each agent can monitor only a specific subset of global CPS states, necessitating neighborhood information exchange. In this framework, we propose an agent based static state estimation encompassing local consensus and least square solution. Necessary bounds for the consensus weights are obtained through the maximum eigenvalue based convergence analysis and are verified for a radial power distribution network. The agent based formulation is also applied for a linear dynamical system and the consensus approach is found to exhibit better and more robust performance compared to a diffusion filter. The agent based Kalman consensus filter (AKCF) is further investigated, when the agents can choose between measurements and/or consensus, allowing the economic allocation of sensing and communication tasks as well as the temporary omission of faulty agents. The filter stability

is guaranteed by deriving necessary consensus bounds through Lyapunov stability analysis. The states dynamically estimated from AKCF can be used for state-feedback control in a model predictive fashion. The effect of lossy communication is investigated and critical bounds on the link failure rate and the degree of consensus that ensure stability of the agent based control are derived and verified via simulations.

MULTI-AGENT ESTIMATION AND CONTROL OF
CYBER-PHYSICAL SYSTEMS

by

S M SHAFIUL ALAM

B.S., Bangladesh University of Engineering and Technology, 2008

M.S., Bangladesh University of Engineering and Technology, 2011

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Dr. Balasubramaniam Natarajan

Copyright

S M Shafiqul Alam

2015

Abstract

A cyber-physical system (CPS) typically consists of networked computational elements that control physical processes. As an integral part of CPS, the widespread deployment of communicable sensors makes the task of monitoring and control quite challenging especially from the viewpoint of scalability and complexity. This research investigates two unique aspects of overcoming such barriers, making a CPS more robust against data explosion and network vulnerabilities. First, the correlated characteristics of high-resolution sensor data are exploited to significantly reduce the fused data volume. Specifically, spatial, temporal and spatiotemporal compressed sensing approaches are applied to sample the measurements in compressed form. Such aggregation can directly be used in centralized static state estimation even for a nonlinear system. This approach results in a remarkable reduction in communication overhead as well as memory/storage requirement. Secondly, an agent based architecture is proposed, where the communicable sensors (identified as agents) also perform local information processing. Based on the local and underdetermined observation space, each agent can monitor only a specific subset of global CPS states, necessitating neighborhood information exchange. In this framework, we propose an agent based static state estimation encompassing local consensus and least square solution. Necessary bounds for the consensus weights are obtained through the maximum eigenvalue based convergence analysis and are verified for a radial power distribution network. The agent based formulation is also applied for a linear dynamical system and the consensus approach is found to exhibit better and more robust performance compared to a diffusion filter. The agent based Kalman consensus filter (AKCF) is further investigated, when the agents can choose between measurements and/or consensus, allowing the economic allocation of sensing and communication tasks as well as the temporary omission of faulty agents. The filter stability

is guaranteed by deriving necessary consensus bounds through Lyapunov stability analysis. The states dynamically estimated from AKCF can be used for state-feedback control in a model predictive fashion. The effect of lossy communication is investigated and critical bounds on the link failure rate and the degree of consensus that ensure stability of the agent based control are derived and verified via simulations.

Table of Contents

Table of Contents	viii
List of Figures	xiii
List of Tables	xvi
Acknowledgements	xvi
Dedication	xvii
1 Introduction	1
1.1 What is CPS?	1
1.2 Challenges in CPS	3
1.3 Research Questions	4
1.4 Related Work	5
1.4.1 Question 1: Correlated Information Aggregation	6
1.4.2 Question 2: State Estimation from Compressive Measurements	7
1.4.3 Question 3: Distributed Static Estimation	8
1.4.4 Questions 4 and 5: Distributed Dynamic Estimation	10
1.4.5 Question 6: Distributed and Decentralized Model Predictive Control	12
1.5 Contributions of This Dissertation	15
1.6 Organization of This Dissertation	21
2 Background	22

2.1	Compressed Sensing	22
2.2	State Estimation and Control	26
2.2.1	Static State Estimation	27
2.2.2	Dynamic State Estimation and Control	29
3	Impact of Correlated Distributed Generation on Information Aggregation in Smart Grid	35
3.1	Introduction	36
3.2	General System Model	37
3.3	Correlation Analysis in Distribution Grid	40
3.4	Information Aggregation in Distribution Grid	45
3.5	Results	48
3.6	Conclusions	53
4	Distribution Grid State Estimation from Compressed Measurements	55
4.1	Introduction	55
4.2	State Estimation from Compressed Measurements	59
4.2.1	Indirect Method	60
4.2.2	Direct Method	61
4.2.3	Computational Complexity	64
4.3	Results	66
4.4	Conclusions	70
5	Agent based State Estimation in Smart Distribution Grid	71
5.1	Introduction	72
5.1.1	Related Work	73
5.2	System Model	75

5.3	Proposed Method	78
5.3.1	Convergence Analysis	79
5.4	Simulation and Results	82
5.5	Conclusions	84
6	Distributed Agent based Dynamic State Estimation over a Lossy Network	87
6.1	Introduction	88
6.2	System Model	91
6.3	Proposed Method	94
6.4	Effect of Lossy Communication Network	96
6.5	Simulation and Results	98
6.5.1	Figure of Merit	98
6.5.2	Case Study: Perfect Communication	99
6.5.3	Case Study: Random Link Failure	102
6.6	Conclusions	104
7	Agent based Optimally Weighted Kalman Consensus Filter over a Lossy Network	105
7.1	Introduction	105
7.2	Shared and Unshared State Elements in Agent based model	106
7.3	AKCF with Adjustable Consensus Weights	109
7.4	Main Results	110
7.5	Stability under Lossy Communication	117
7.6	Case Study	119
7.7	Conclusions and Future Work	121
8	Agent based Kalman Filtering - Measurement vs Consensus?	122

8.1	Introduction	122
8.2	Incorporation of Flexible Policy in Basic AKCF	123
8.2.1	Random Measurement	124
8.2.2	Random Consensus	125
8.3	AKCF Stability in Case 2 and Case 3	127
8.4	Simulation and Results	130
8.5	Conclusions	132
9	Stability of Agent based Distributed Model Predictive Control over a Lossy Network	136
9.1	Introduction	137
9.2	System Model	139
9.3	Main Results	143
9.4	Effect of Lossy Network on Agent based control	146
9.5	Simulation and Results	148
9.6	Conclusions	151
10	Conclusion and Future Work	152
10.1	Summary	152
10.2	Future Work	154
	Bibliography	157
	Bibliography	157
A	SYS1 Parameters	173
B	SYS2 Parameters	175

C Positive Definiteness of $G_{t,k}$	178
D Least Square Estimate of AKCF Weights	180
E SYS3 Parameters	183
F Expected Value in Equation (8.10)	185
G Positive Definiteness of $B_{t,k}$	186
H Least Square Estimate of AMPC Weights	187

List of Figures

1.1	(a)A depiction of Cyber-Physical System. (b) Major CPS Challenges [1, 2, 3, 4].	2
1.2	(a) Holonic Multi-Agent System (HMAS), (b) Depiction of Power Distribution Network as an HMAS	14
2.1	Compressed Sensing	23
2.2	An Example of 1D Compressed Sensing	25
2.3	An Example of 2D Compressed Sensing	27
3.1	Single Phase Distribution Network with Renewable Energy Source.	38
3.2	IEEE 34-Node Test Feeder (indices in parenthesis are used in this research).	42
3.3	Spatial Correlation Matrix of Distributed Generation based on (3.5)	43
3.4	Correlation of Voltage Phasors for Spatially Correlated Generation and Random Load	44
3.5	Illustration of Spatial and Temporal Compressed Sensing	47
3.6	Variation of $INAE$ for Spatial, Temporal and Spatiotemporal Compressed Sensing	50
3.7	Reconstruction from Spatially Compressed Real Power Measurements	51
3.8	Reconstruction from Temporally Compressed Real Power Measurements	52
3.9	Voltage Phasors Etimated from Direct and Compressed Power Measurements	53
4.1	Centralized State Estimation: (a) Conventional, (b) Indirect Method, (c) Direct Method	58
4.2	Off-diagonal Entries of Normalized $\mathbf{J}^{(0)T} \Phi^T \Phi \mathbf{J}^{(0)}$	63

4.3	Spatial Correlation of $\mathbf{b}^{(0)}$	65
4.4	Voltage Phasors Estimated from <i>Direct</i> and <i>Indirect</i> Method	67
4.5	Variation of INAE and MIAE for <i>Indirect</i> and <i>Direct</i> Methods	69
5.1	Venn Diagram of State Elements for a Radial Power Distribution Network.	78
5.2	Effect of a : (a) Mean Convergence of Error in Voltage Magnitudes. (b) Mean Convergence of Error in Voltage Angles.	85
5.3	Effect of b : (a) Mean Convergence of Error in Voltage Magnitudes. (b) Mean Convergence of Error in Voltage Angles.	86
6.1	Inter-agent Information Exchange	96
6.2	Venn diagrams for SYS1 and SYS2	98
6.3	AKCF versus ADKF for SYS1	100
6.4	AKCF versus ADKF for SYS2	101
6.5	Effect of communication for SYS1	102
6.6	Effect of communication for SYS2	103
7.1	Extracting and Reordering of Shared and Unshared State Elements.	108
7.2	Venn Diagram of an Example 3-Agent System.	108
7.3	Relative Positioning of the Range of Quadratic Values.	111
7.4	AKCF Stability for SYS3 in Perfect and Lossy Network.	120
8.1	ϵ Upper bounds (ϵ^*) for 10 agent system. (a) $m = 10 - r$ agents are expected to use local measurements only and don't exchange information with neighbors. (b) $m = 5$ agents are expected to use local measurements while $1 \leq r \leq 10$ agents are expected to exchange information with neighbors.	131
8.2	AKCF stability performance of 10 agent system under case 2 (mutual events). The upper bound ϵ^* is mentioned in respective plots.	133

8.3	AKCF stability performance of 10 agent system under case 3 (independent events). The upper bound ϵ^* is mentioned in respective plots.	134
9.1	Agent based MPC Performance. (a) Perfect Network, (b) Lossy Network with Failure Probability = 0.4, (c) Lossy Network with Failure Probability = 0.8	150

List of Tables

4.1	CPU time for the proposed estimators	68
7.1	Steady State Bounds for SYS3	120
9.1	Steady State Bounds	149

Acknowledgments

First and foremost, I would like to thank my supervisor **Dr. Bala Natarajan** for his continuous guidance and support in my pursuit of the Doctorate of Philosophy. The four years of graduate research under his supervision was a memorable experience for me. I consider him as a great example of effective time management, thorough insight and untiring perseverance. Besides, his consistent and incentive feedback in every aspect of my professional development helped me to make significant improvement in communication and writing skills.

My sincere gratitude to the Doctoral committee members **Dr. Anil Pahwa, Dr. Nathan Albin, Dr. Shelli Starrett** and **Dr. Andrew Ivanov** for their constructive feedback and advice in making this dissertation to the final shape.

I am also grateful to former WiCom lab members **Dr. Siddharth Deshmukh, Dr. Nicholas Roseveare** and **Dr. Mohammed Taj-Eldin** for their mentoring and guidance in conducting research and addressing new challenges in everyday life of a graduate student. In addition to that, I am thankful to current lab members **Kan Chen, Kumar Jhala, Wenji Zhang** and **Liu Chang** for their support and feedback about my research.

There are no words, which can express my feelings of gratitude and honor to my parents (**A H M Shah Alam & Shahida Alam**). I should not forget my parents' countless sacrifices to raise me up to what I am today. And, I am lucky to have **Syeda Rubaiyat Aziz** as my life partner, whose endless support and encouragement remind me of my capability and potential. Finally, I am grateful to all of my well-wishers, whose blessings and advices are the integral part of my success.

Dedication

This dissertation is dedicated to the three forerunners of science, engineering and technology in my family, who sacrificed their lives during the **Liberation War of Bangladesh** in **1971**.

- My paternal grandfather **Md. Khurshid Alam**, who was an Assistant Engineer at the East Pakistan Telegraph & Telephone (Currently known as **Bangladesh Telecommunications Company Limited**).
- My paternal uncle **A K M Mahbubul Alam (Bhashani)**, who was a *2nd* year MBBS student at **Mymensingh Medical College**.
- My maternal uncle **Md. Nurul Alam (Salim)**, who was a student at **Barisal Polytechnic Institute**.

Chapter 1

Introduction

1.1 What is CPS?

Cyber-Physical System (CPS) is a term used to describe “a ubiquitous and smart integration of sensing devices, computing processors and communication networks that reliably interacts with the physical world (with the probable involvement of humans in the loop) in real time”(Fig.1.1(a))[5]. Although this term was first introduced by Helen Gill in 2006 [6], the main concept dates back to post World War II era. In 1948, the famous mathematician Norbert Wiener coined the word “cybernetics” from the Greek “κυβερνητης” (meaning helmsman, governor, pilot, or rudder) to describe the integration of closed loop feedback control and communication [7]. CPS can be perceived as a set of embedded systems that closely interact with each other and beside information processing and control, are able to perform computational tasks with the knowledge of underlying physical dynamics. A successful implementation of CPS depends upon the synchrony at which various sequential tasks (e.g., data collection, processing, prediction and control) can be performed throughout a reliable network of diverse physical entities [8, 9]. CPS spans multiple applications such as (1) smart energy management with active end user participation in energy trade [10], (2) real-time traffic-aware autonomous transportation enabled with vehicular networks [11],

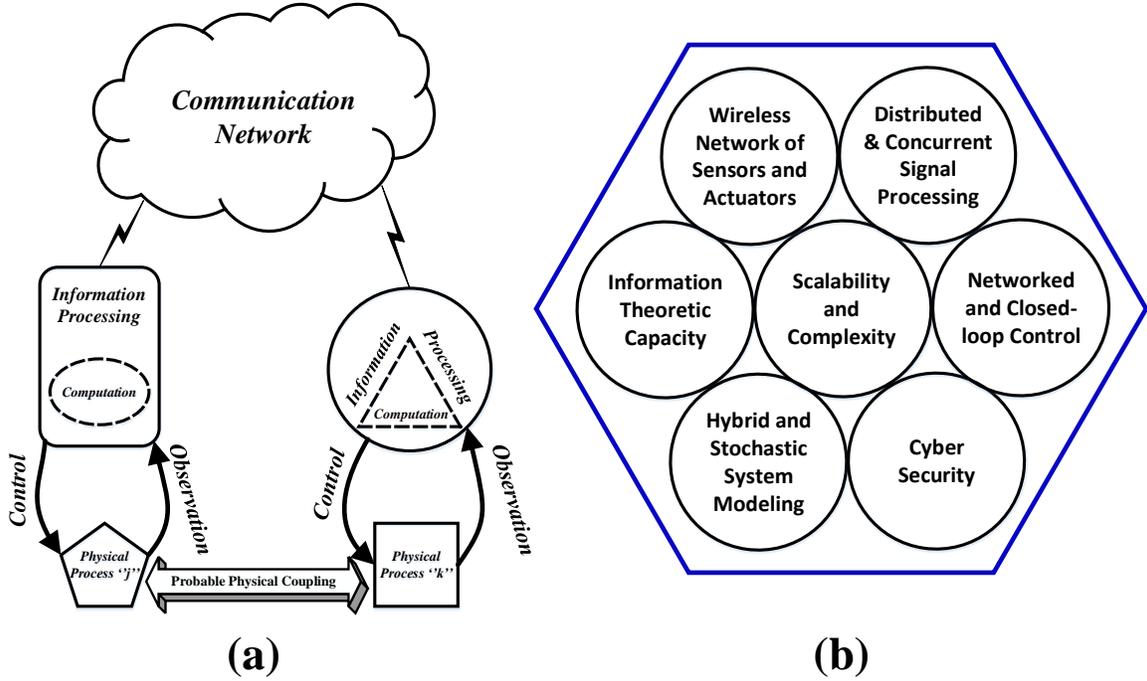


Figure 1.1: (a) A depiction of Cyber-Physical System. (b) Major CPS Challenges [1, 2, 3, 4].

(3) ubiquitous healthcare enhancing societal well being [12], (4) exploration and disaster management in hazardous or inaccessible environments, and (5) climate and pest-smart agriculture allowing efficient cultivation in available arable lands [13, 14]. A broader and more structured view of CPS, its impacts and applications is presented in [4]. The inherent heterogeneous characteristics and complexity of CPS demands an adaptive, reconfigurable and near realistic modeling. Unlike an embedded system, the model should incorporate physical dynamics along with communication reliability and information processing so that it can imitate the temporal evolution of CPS states and predict responses to any unusual circumstances [15]. In the next section, we shed light on some of the major CPS challenges which are especially important from sensing, computation and control perspectives.

1.2 Challenges in CPS

A successful, reliable and *green* implementation of a cyber-physical system presents at least seven major challenges as depicted in Fig.1.1(b). Specifically, the issue of “Scalability and Complexity” intersects with every other major challenge from the perspective of (a) sensing, (b) communication and (c) computation. We briefly discuss the mentioned challenges from these perspectives.

- *Sensing Big Data*: Technological advances are making wireless communicable sensors cheaper and more powerful. As a result, a large-scale physical system can now be observed in real time through massive deployment of such sensors [16]. These sensors open the door to the “cyber” aspect of an intelligent self-healing dynamical system. However, such universal usage of high resolution sensors can generate a significant volume of raw data. Consequently, with increasing numbers of communicable sensors, the effective throughput of underlying network can be drastically impacted if necessary remedies are not taken [17]. To give an idea about this “big data” problem, we highlight an ongoing smart grid project in Austin, Texas [18]. Here, with the deployment of communicable smart energy meters with 15 second resolution, about 2.7 billion units of raw data were collected in the year of 2011 [19]. Such an exponential growth in raw measurements demands an efficient and robust information aggregation scheme along with economic storage for successful monitoring and optimization.
- *Computation Time and Concurrency*: In many CPS, spatially distributed sensors collect raw data in a periodic manner and convey them to a fusion center mostly through a wireless communication network [20]. Based upon a mathematical model of the system, the fusion center uses these measurements to obtain estimates of underlying system states and derive necessary control decisions. The control signals are then fed back to the respective actuators that ensure expected behavior of the CPS. Besides algorithmic complexity, speed and architecture of the estimation and control

unit (ECU), the overall computation and processing time is also dependent upon the volume of raw data involved. This is critical since the physical process may evolve at a rate much faster than the rate at which ECU can make control decisions. Thus, the concurrency in cyber-space information processing and real-world physical dynamics is a vital issue for desired behavior of the integrated cyber-physical system [6].

- *Robustness to Communication Network Impairments:* The communication network is an integral part of a CPS. For sensing, estimation and control purposes, a large-scale CPS requires a robust network that can handle the large volume of information with least amount of delay as well as packet drops. However, faulty reception of packetized data specially from distant sensor(s) is quite possible resulting in incorrect estimation and control actions leading to unstable operation of the physical system [20]. Furthermore, centralized schemes for information processing require more power for wireless transmission resulting in rapid decay of sensor’s battery life. Thus, avoiding long distance communication will not only make the CPS more robust but will also conserve individual sensor’s power consumption. Hence, an efficient information aggregation scheme as well as distributed estimation and control strategies can pave the way to a greener and more robust cyber-physical system.

Based on these three perspectives, we seek to address a few fundamental research questions in this dissertation. These questions and the previous attempts to address them are discussed in the following sections.

1.3 Research Questions

Question 1. *For large scale cyber-physical systems with wide-spread deployment of communicable sensors, how can we minimally aggregate measurements at a fusion center while ensuring a low communication burden and preserving adequate system information?*

Question 2. *Can this minimal information be effectively utilized for centralized estimation of the physical system states?*

Question 3. *Is it possible for sensors (with communication and computation capabilities) to estimate only a partial set of system states instead of estimating the entire system at a fusion center? If such distributed estimation is feasible, how should the sensors interact with each other so that centralized processing can be avoided? What information should they share? What would be the performance of such estimation strategies?*

Question 4. *How does the degree of inter-sensor information exchange impact estimation stability? What is the impact of a lossy communication network on such distributed estimation methods?*

Question 5. *What if sensors are given the freedom to use measured data and/or information from other sensors to derive local system awareness? How does such autonomous sensor choices affect the state estimation stability?*

Question 6. *Can these distributed estimator design scenarios be applied to distributed control decision making?*

1.4 Related Work

This dissertation attempts to address the research questions listed above in the context of one of the most complex and challenging CPS - the electric power system [21]. However, the methods, algorithms and theoretical results/insights developed in this work can be applied to any CPS with minor application specific modifications.

The current electric grid will presumably evolve as a smart and intelligent network, a futuristic model of energy cloud and information web for interested clients. The conventional electric power system is established over a hierarchical structure with generation and distribution network at the top and bottom level, respectively. A smart implementation can

be achieved with advanced metering infrastructure (AMI), which includes communicable meters and actuators throughout the grid to collect data and act on it (e.g., use it for state estimation and / or control). It will also enable end-user participation in system monitoring, operation and control through two-way communication with electrically coupled neighbors and exchanging relevant encrypted information [22]. Additionally, the future grid will experience a significant penetration of cost effective and renewable energy sources especially at the end-user premises [23]. Also, with advances in technology the usage of electric vehicles is expected to increase in near future. Beside high volume data management, these components will make the electric power generation and usage more stochastic in nature that demands for more sophisticated and distributed implementation of power system estimation and control. From a smart grid perspective, the following subsections summarize the prior work related to the research questions of interest.

1.4.1 Question 1: Correlated Information Aggregation

As mentioned earlier, the high penetration of renewable energy sources and rapid growth in electric vehicle usage will make the power system sensing, estimation and control task more challenging. One key characteristic such distributed generation (DG) is the spatial correlation resulting from the geographical proximity of the renewable energy sources [24, 25]. Furthermore, a distribution network consists of loads characterized in several categories such as residential, commercial etc. Typically, the load demands of similar class of customers (e.g., residential, commercial) is expected to be highly correlated. It should be noted that the inherent correlations of a signal corresponds to low variability or some degree of smoothness across it, resulting in compressible coefficients especially in the wavelet domain [26, 27]. Hence, it is expected that, with high level of correlation, the signal will be more compressible, allowing approximation with fewer number of wavelet coefficients [28]. On the other hand, the sampling frequency of monitoring devices (e.g., smart meter installed at the customer premises) is expected to rise from the necessity of real-time monitoring and

robust control of the future electric grid. In this scenario, we can exploit the correlated nature of distributed generation and load to efficiently aggregate power information using concepts from *Compressed Sensing*. A successful implementation of this technique is dependent upon the degree of sparsity of the information being collected. In Chapter 3, we discuss in detail about the application of compressed sensing in efficient aggregation of real and reactive power measurements both in space and time. Recently, authors in [29, 30] report application of compressed sensing in simultaneous reading of all smart meters in order to securely transmit the sensed data and to ensure identical delay in wireless transmission of all smart meter readings. The measurement compression is obtained using pseudorandom spreading sequence. In [29], compressed measurements are obtained in the presence of bounded noise whereas, the measurement noise is assumed to be normally distributed in [30]. In [31], an automated compressed meter reading scheme is proposed for a wireless home area network, where the smart meters are equipped with Zigbee standard communication devices. In all these works it is assumed that, the meters convey information only about the significant changes of the respective power levels. In this way sparsity is preserved to apply compressed sensing. However, in a practical scenario, the meters may send exact measurement information at regular time intervals. Therefore, the underlying correlation among the measurements both in space and time needs to be utilized in order to obtain approximate sparsity and hence the compressed sensing technique. In smart meter data collection perspective, this is still an open problem. Furthermore, the impact of correlated real and reactive power measurements over the voltage states needs to be investigated in order to assess the competence of compressed measurements in state estimation.

1.4.2 Question 2: State Estimation from Compressive Measurements

In an electric power grid, the state of the system is usually defined as system node voltage phasors or branch currents [32, 33]. These states form a system of nonlinear equations with

the typically available measurements of real and reactive power. These measurements may not be readily available from every customer. As a result, pseudo-measurements are synthesized using historical data and real-time measurements, so that an overdetermined system of equations can be formed [34, 35]. Such a system of nonlinear equations is then solved using weighted least square (WLS) method and its variants [33]. As discussed in Section 4.1, the application of compressed sensing causes a linear mapping to underdetermined system whereas, traditional cases are the direct mathematical models of corresponding physical system. A typical example is a set of parametrized nonlinear equations, arising from an attempt to solve a determined system in \mathbb{R}^M by continuation methods [36]. Another example is the method of finding an interior point of the polytope $\{\mathbf{z} \geq \mathbf{0}, \mathbf{A} \in \mathbb{R}^{M \times N} : \mathbf{A}\mathbf{z} = \mathbf{b}\}$, by using the non-linear transformation ($z_i = e^{x_i}, i = 1, \dots, N$) [37]. As a third example, we quote a recent article about the in-silico manipulation of biological signaling pathways, which is modeled as an underdetermined system of nonlinear equations [38]. In prior literature, Newton-Raphson method is reported to be the most suitable to solve these kinds of underdetermined nonlinear equations [37], which can give only the least square solution using pseudo-inverse of underlying Jacobian matrix. Furthermore, the compressed measurements obtained in [29, 30, 31] need to be exploited to estimate the states of a power system.

1.4.3 Question 3: Distributed Static Estimation

The word *decentralized* and *distributed* are used interchangeably in literature to describe the state estimation procedure by a set of communicable sensors distributed over the physical system. In this architecture, each sensor node observes only a distinct portion of underlying physical system and makes a local estimate of overall system states by exchanging information with neighboring nodes. It offers benefits relative to centralized and hierarchical data fusion architecture in terms of (1) real-time implementation, (2) configuration flexibility, and (3) communication bottlenecks [39]. In this regard, a brief survey of static state esti-

mators is given in [40], which also reports the weighted least square (WLS) method as the most popular one in this respect. The WLS method is used in various literature for global state estimation from decentralized observation model, specially for interconnected electric power transmission systems. The decentralization is generally obtained by decomposing an interconnected transmission system into a certain number of *nonoverlapping* subarea on a geographical basis [41]. In each of the subsystem, respective local estimates are obtained using Gauss-Newton iteration. The local estimates are then centrally coordinated to obtain overall system states subject to the boundary constraints of the interconnected system. A comprehensive survey of such multiarea scheme of state estimation is given in [42]. A hierarchical multilevel structure of state estimation based on factorized WLS method is described in [43]. Here, the authors define the distribution substations as the lowest level of the multilevel structure. In [44], authors use decentralized measurement model with the assumption that the global state may not be observable in some of the subareas. In each subarea, the global states are synchronously updated using the combined approach of global consensus and innovation. It should be noted that, the convergence to global consensus is not reported in [44], when local jacobian matrix violates the condition of full column rank in each of the subarea. Unlike [41], partial overlapping of local state vectors is considered in [45] and [46]. It is assumed that any subarea may have some state elements, which are shared with its neighbors. The sharing is due to the specific placement of measurement devices over the tielines of an interconnected system. In this scenario, the multi-area state estimation (MASE) is designed as a distributed optimization problem in each Gauss-Newton iteration [45]. As discussed in Section 5.1.1, this optimization problem represents a *partially* global consensus problem, that can be solved using alternating direction method of multipliers (ADMM) (See chapter 7 of [47] for details). In [46], authors implement R-SDP with the help of ADMM to solve the problem of MASE. However, two important assumptions are made in the works of [45] and [46]: (1) in each subarea, there exists at least one state element, unshared with any of its neighbors, and (2) if a state element is shared, it is only

between two neighboring subareas. As will be discussed in Chapter 5, these assumptions are restrictive and do not hold for the power distribution grid considered in the current research. In a nutshell, the prior efforts are based on either over-determined or fully determined measurement models. As a consequence, efficient factorization schemes [48] can be applied in WLS method as well as in partially global consensus problem. Furthermore, a scenario with more than two subareas sharing the same state element as well as the case when all local state elements are shared among neighbors is still an open challenge in distributed state estimation.

1.4.4 Questions 4 and 5: Distributed Dynamic Estimation

The current electric power grid can be tracked via a distributed network of communicable sensors typically over a large geographical region [49, 50]. The underlying communication and computational burden is considerably high with a centralized or hierarchical implementation of a dynamic state estimator, e.g., Kalman filter [39]. This issue is resolved through distributed implementation of Kalman filters that also offers benefits in terms of real-time implementation and configuration flexibility. Although, the fundamental concept is unchanged, the distributed Kalman filter has evolved through numerous versions. Among those, Kalman consensus filter (KCF) [51] and diffusion Kalman filter (DKF) [52] are worth mentioning. In both setups, the sensors collect detailed information about the measurements (i.e., sensed data and noisy observation space model) from neighbors. The fused information is then applied to the classical Kalman filtering algorithm. Thus, an intermediate estimate of the whole dynamics and the corresponding update in estimation error covariance matrix is obtained at each sensor. At the last stage, the sensors exchange information about the intermediate estimates with their neighbors. A correction is made to each sensor's estimate by either applying consensus (i.e., KCF), or through a weighted combination of the received neighborhood information (i.e., DKF). Finding the desired degree of consensus or the optimal weights for diffusion are the major issues in the design of such distributed Kalman

filters. For the Kalman consensus filter, the objective is to design a scalar consensus gain parameter. In this regard, an optimization problem is solved in [53] in order to find the desired consensus gain that ensures the convergence of the over all estimation error covariance to a known steady-state matrix. However, it is also mentioned that such steady-state matrix is hard to obtain in practical perspective and an approximate expression for the consensus parameter is derived. On the other hand, in [54] a Lyapunov stability analysis is carried out over the estimation error dynamics to find a range of real numbers within which the desired consensus parameter should be chosen. When it comes to implementing weighted diffusion of neighborhood estimates, the Laplacian [52], the Metropolis [55] and the nearest neighbor or uniform weighting rules [56] are the popular choices in designing the diffusion weights. The robustness of Kalman consensus filter is also investigated under the effect of lossy sensor network in [53] by incorporating a Bernoulli random variable in the consensus step. The effect is only illustrated through simulations necessitating a theoretical analysis that involves the consensus gain parameter and network reliability. The usage of mutually independent Bernoulli distributions is also reported in [57] to model the random presence of nonlinear dynamics as well as the quantization effect in the sensor communication. The filter designed is of diffusion characteristics and the corresponding weights are derived based on the average \mathcal{H}_∞ performance. On the other hand, relative variance and adaptive combination rule is proposed in [56] for stationary diffusive estimation of single parameter under noisy communication link. As a matter of fact, most of the advanced distributed Kalman filters are based on these two major approaches, although they bear widely varying application specific characteristics. We would like to refer [58] as a resource that summarizes the extensive research carried out in this arena. As mentioned in Section 1.2, the massive deployment of communicable sensors makes it quite challenging to regularly store and update the global state vector of a CPS-like large-scale system through a fully connected network. Consequently, the number of communication links, computational memory and hardware requirements increase with the increase in active sensor nodes. Thus, it may be impractical

to track the high dimensional state vector in its entirety at each communicable sensor. This constraint is overcome specifically for sparse large-scale linear systems [59, 60]. In this case, the corresponding transition of states can be reflected on (approximately) banded matrix to spatially decompose the overall dynamics among sensors even when local measurement space projects onto global states. This idea is further extended for system specific reduced order particle filtering [61] and distributed observer design for large-scale system partitioned into disjoint areas [62]. The key fact is, the observation space of each sensor is modified solely based on the characteristics of state dynamics.

On the contrary, in a practical physical system, overlaid with communication network, the observation space of a sensor may be coupled to a limited set of specific state elements. Some state elements may even be coupled to two or more sensors' observation space. Under these circumstances, each sensor may be relieved to track only the pertinent state elements. Feasibility of such a scenario is still unknown and demands extensive investigation.

1.4.5 **Question 6: Distributed and Decentralized Model Predictive Control**

Model predictive control (MPC) of a system relies on the future prediction of state trajectories so that the set of desired control inputs can fulfill system specific objectives under the given constraints. This is advantageous than finite horizon approximated linear quadratic regulator as it offers more stability in satisfying the control objective. However, as pointed out in [63, 64], the computational complexity of MPC is a major issue that makes its application in large-scale systems questionable in terms of efficiency, robustness as well as reliability. Such scalability issue can be overcome by delegating the control responsibility among a set of controllers spatially distributed across the cyber-physical system (CPS). In this regard, two basic approaches are typically used: (1) decentralized approach, where the set of controllers independently decides specific and non-overlapped control inputs without the requirement of communication, and (2) the distributed approach, where a more

reliable control decision is achieved through an iterative information exchange among the controllers over a communication network [65]. The characteristics of a large-scale system usually plays a major role in defining each controller’s observation and control space in a distributed control scheme. In [66], a sparse and large scale system is decomposed into multiple subsystems. The subsystems are mapped using binary matrices *exclusively* designed for specific dynamics irrespective of observation space. In absence of any negotiation, each entry of the control input set is chosen from individual subsystem. The decomposed local subsystem is also reported to include the effect of neighborhood control decisions. In [64], the neighboring control inputs appear as disturbance to the local controller of a radial network of water delivery canal system. Alternately, a weighted combination of neighborhood controls is considered in [67]. Here, a communication intensive repetitive procedure of control coordination is proposed. However, the typical effect of communication networks (delay, loss of packets etc.) are not considered. A successive loss of limited number of packets are considered in [63] to analyze the effect of lossy network over the closed loop stability of the global dynamical process. Consensus in control using linear quadratic Gaussian regulator is investigated in [68] and the stability criteria is derived via conditioning the spectral radius.

In recent years, distributed control is becoming popular in multi-agent operation of cyber-physical systems. Electric power transmission and distribution system can be considered a good candidate example. In this context, a graph theoretic approach is reported in [50] to maintain power balance at the prevalence of renewable energy sources. Authors in [69] use multi-objective genetic algorithm to minimize phase current unbalance in distribution network. More recently, the optimal power flow problem is distributed in [70] using the Lagrangian relaxation. A lower-upper-bound switching algorithm is designed to balance the power flow between the utility and microgrid enabled communities. In a broader perspective, the problem of optimal operation of electric power system can be addressed as synchronous optimization in a multi-objective environment. As a result, decision conflict is not uncommon in an agent based framework of such systems. In this regard, coalitional

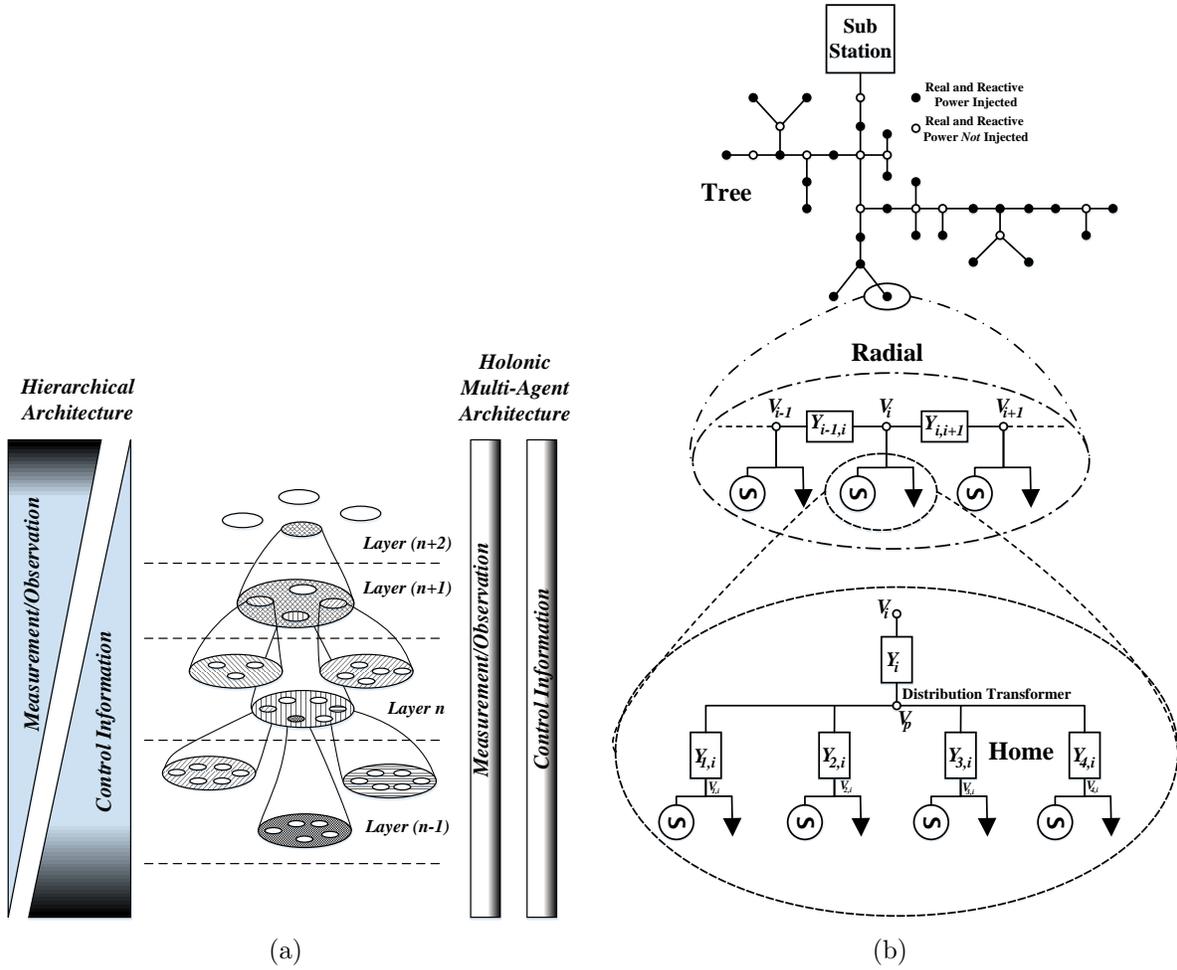


Figure 1.2: (a) *Holonic Multi-Agent System (HMAS)*, (b) *Depiction of Power Distribution Network as an HMAS*

game theory is used in [71] to combine different objective functions to ensure common control decisions. A recent review of the usage of consensus strategy in distributed multi-agent optimization problem shows the focus of existing research in individual agent's convergence to the average of the initial states [72]. Optimal convergence for general linear and nonlinear systems, however is still an open problem. Also, the distributed control framework still needs to be generalized based on deeper insights on system behavior as well as the limited observation spaces of the controllers [73].

1.5 Contributions of This Dissertation

In this dissertation, a holonic multi-agent framework is proposed, which enables distributed estimation and control of a CPS. A multi-agent system consists of virtual, intelligent and proactive agents, which are individually aware of the surrounding environment and able to interact with each other over a communication network. Through a collaborative approach, the multi-agent system enables the underlying physical system to be *artificially intelligent* to make decisions in various situations in order to reach a universal objective [74]. Some prior research in agent based modeling involves coordinated control of unmanned air vehicles [75]; reliable management of intelligent water distribution network [76], efficient energy management in commercial buildings [77], and intelligent traffic control systems [78]. When the agents are segmented in multiple layers and hierarchically organized, they form a holonic multi-agent system (HMAS). Such an architecture can reduce communication as well as computational burden and ensure synchronous tracking and control of underlying CPS. Fig.1.2(a) shows the basic HMAS model. A typical power distribution network is shown in three layers in Fig.1.2(b), namely the “tree” layer, “radial” layer and “home” layer. In this research, the spatially distributed sensors are defined as *agents*. Each of them can (1) collect measurements which relates to a subset of global states, (2) can make intermediate estimate of the local states and (3) can communicate with neighboring agents where the neighbors are defined according to the sharing of state elements. In this regard, the major contributions of this research is presented below, which forms the foundation of this dissertation,

- Questions 1 and 2: Correlated information aggregation and centralized state estimation.
 - Study and quantify the correlation structure induced on the voltages by the correlated generation from DGs. Results demonstrate that, the voltage phasors exhibit very strong correlation similar to that of distributed generation.
 - Demonstrate the feasibility of spatial, temporal and spatiotemporal compressed

sensing strategies on the reconstruction of power measurements. For the IEEE 34 node distribution test feeder, it is observed that, the correlated generation from DGs allow upto 50% reduction in all types of measurements while recovering the original data with 90% accuracy.

- Quantify the accuracy in estimating voltage states from compressed power readings. It is shown that, the voltage phasors of the grid can be estimated almost accurately by randomly measuring only half of the available apparent power information.
- Develop two (direct and indirect) methods of state estimation using compressed power measurements. The underlying computational complexities for the two methods are also discussed. It is shown that both the *indirect* and *direct* state estimation approach give similar performance with same complexity order and can estimate voltage states almost accurately by randomly projecting only half of the available power information.

These contributions are discussed in detail in Chapters 3 and 4 and also appear in the following articles:

[79]: S M Shafiul Alam, Bala Natarajan and Anil Pahwa, “Impact of Correlated Distributed Generation on Information Aggregation in Smart Grid”, *In Proceedings of 5th Annual Green Technologies Conference (IEEE GreenTech)*, April 3-5, 2013, Denver, Colorado, USA.

[80]: S M Shafiul Alam, Bala Natarajan, and Anil Pahwa, “Distribution Grid State Estimation from Compressed Measurements”, *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1631-1642, 2014.

- Question 3: Agent based static state estimation.

- The states of a physical system are estimated in a distributed fashion using agent based nonlinear measurement model. We develop a distributed local consensus based approach, especially when there is overlap of state elements among neighboring agents.
- The convergence rate of distributed consensus is analyzed. Based on this analysis, a general criteria is formulated in order to make system specific selection of coefficient values in the proposed method.

We discuss these contributions in Chapter 5 and they also appear in the following articles:

[81]: S M Shafiul Alam, Bala Natarajan, Anil Pahwa, and Sergio Curto, “Agent based State Estimation in Smart Distribution Grid”, *IEEE Latin America Transactions*, vol. 13, no. 2, pp. 496-502, 2015 .

[82]: Anil Pahwa, Scott A. DeLoach, Bala Natarajan, Sanjoy Das, Ahmad R. Malekpour, S M Shafiul Alam, and Denise M. Case, “Goal-Based Holonic Multi-Agent System for Operation of Power Distribution System”, *IEEE Transactions on Smart Grid (Special Issue on Cyber-Physical Systems and Security for Smart Grid)*, vol. 6, no. 5, pp. 2510-2518, 2015.

- Performance and robustness of consensus and diffusion approach in agent based dynamic state estimation.
 - In contrast to the traditional decomposition of global system dynamics (which is independent of measurement model), we introduce the use of binary projection matrices that reflect the local underdetermined observation space of individual agents. The agents can sense, compute as well as communicate with the agent neighbors (defined according to the sharing of state elements rather than the relative placement of physical devices).

- Develop agent based Kalman consensus filter (AKCF) and agent based diffusion Kalman filter (ADKF) where the observation space of each agent is (generally) underdetermined.
- Investigate the effect of communication over the performance of agent based tracking of the dynamical system.

These contributions are discussed in detail in Chapter 6 and in the following article:

[83]: S M Shafiu Alam, Bala Natarajan, and Anil Pahwa, “Distributed Agent Based Dynamic State Estimation over a Lossy Network”, *In Proceedings of 5th Int. Workshop on Networks of Cooperating Objects for Smart Cities (UBICITEC)*, April 14-17,2014, Berlin, Germany.

- Question 4: Design of stable and optimally weighted agent based Kalman consensus filter (AKCF).
 - Perform Lyapunov stability analysis from the AKCF estimation error dynamics.
 - Find the expression of optimal consensus weight in the form of $\epsilon \mathbf{A}$, where ϵ is a positive scalar value and \mathbf{A} is a system specific and timely updated matrix.
 - Find an upper bound of ϵ that ensures filter stability.
 - Incorporate random link failure among neighboring agents and find the theoretical bound on link failure rate till which guarantees estimation stability.

Chapter 7 discusses these contributions in detail and also appear in the following article:

[84]: S M Shafiu Alam, Bala Natarajan, and Anil Pahwa, “Agent based Optimally Weighted Kalman Consensus Filter over a Lossy Network”, *In Proceedings of 2015*

IEEE Conference on Global Communications (IEEE GLOBECOM), (accepted), December 6-10,2015, San Diego, CA, USA.

- Question 5: Development of a generalized optimally weighted and stable AKCF.
 - Introduce a new flexible policy for agents. Specifically, each agent is given the flexibility of random measurement and/or consensus. That is, agents can use either the local measurements or neighborhood information or both for estimating the states. This is a unique aspect of this research and completely different from the concept of *intermittent* measurements, which usually occurs due to faulty communication link among sensing and computing devices. This agent-level flexibility also allows the modeling of scenarios, when (1) the costs involved in sensing vs communication are significantly different; (2) there is partial/temporary malfunction of sensors that prevents its ability to take measurements but not impair its communication interface (and vice versa).
 - Define three cases of agent behavior based on the usage of measurement and consensus in state estimation. In the first case, each agent uses both the measurement and consensus in state estimation. In case 2, each agent is allowed to use either measurement or consensus. Finally, in the third case, each agent's decision of measurement usage is independent from that of participation in consensus. For each case, the filter stability is analyzed based on the mean behavior of Lyapunov energy function.
 - Derive bounds on the consensus level (in terms of eigenvalue ratios) that will guarantee convergence of the estimation process. For each case of agent behavior, three different scenario are considered while expressing these bounds. These scenario are characterized based on the relative eigenvalue distribution of the underlying symmetric positive definite matrices.

These contributions are discussed in detail in Chapter 8 and under review in the fol-

lowing article:

[85]: S M Shafiul Alam, Bala Natarajan, and Anil Pahwa, “On the Stability of Agent based Kalman Filters with Measurement and/or Consensus”, *IEEE Transactions on Control of Network Systems*(under review), 2015.

- Question 6: Agent based distributed model predictive control.
 - Propose a unique way of obtaining agent based small-scale model of global system dynamics. This down conversion depends upon agent specific observation space.
 - Perform Lyapunov function based stability analysis of the closed loop system resulting from agent based model predictive state feedback control.
 - Determine a reasonable control consensus weight in the form of $\nu \mathbf{A}$, where ν is a positive scalar value and \mathbf{A} is system specific and timely updated matrix.
 - Derive an upper bound of ν that ensures closed loop stability of global system.
 - Incorporate independent and identically distributed random link failure among neighboring agents and find theoretical bound on the link failure rate within which the global system will be stable.

We discuss these contributions in Chapter 9 and under review in the following article:

[86]: S M Shafiul Alam and Bala Natarajan, “Stability of Agent based Distributed Model Predictive Control over a Lossy Network”, *IEEE Transactions on Signal and Information Processing over Networks* (In Press), 2015.

1.6 Organization of This Dissertation

Chapter 2 provides the background on compressed sensing, static and dynamic estimation and control. Chapter 3 describes the application of spatial, temporal and spatiotemporal compressed sensing to aggregate correlated measurement information. Chapter 4 describes two methods of centralized estimation of states that nonlinearly relate to compressed measurements. In Chapter 5, an agent based static state estimation method is proposed along with the necessary condition of stability. The agent based approach is extended to dynamical process in Chapter 6 and the comparative performance of associated Kalman consensus and Diffusion Kalman filter both under perfect and lossy communication network is presented. Lyapunov energy based stability analysis is carried out in Chapter 7 that ensures a stable and optimally weighted deterministic AKCF under the effect of communication network. In Chapter 8, a flexible policy is introduced into AKCF formulation so that the agents have the freedom to collect local measurements and/or participate in consensus. The associated bounds on the degree of consensus is also derived from Lyapunov stability analysis of estimation error dynamics. Chapter 9 describes an agent based model predictive control scheme and analyze the closed loop state stability . Concluding remarks and future research directions are discussed in Chapter 10.

Chapter 2

Background

In this chapter, we discuss the concept of compressed sensing and describe the fundamentals of static and dynamic state estimation and control.

2.1 Compressed Sensing

The term *Compressed Sensing* refers to the direct acquisition of data in a compressed form, rather than first sampling at high rate and then compressing the sampled data (Fig.2.1). It allows exact or approximate reconstruction of a signal from a very small number of random projections of the signal itself, provided the signal under consideration is (approximately) sparse [87].

Fundamentally, it is possible with high probability to reconstruct data from comparatively very small number of random measurements of original dataset, provided that the original data is sparse itself or has approximate sparsity in a linear transformation basis. We briefly state the basic concept in the following theorem, the proof of which can be found in the seminal work by Candes and Tao [88]:

Theorem 2.1. *Let $\mathbf{w} \in \mathbb{R}^N$ be the original signal, compressible on a linear transformation*

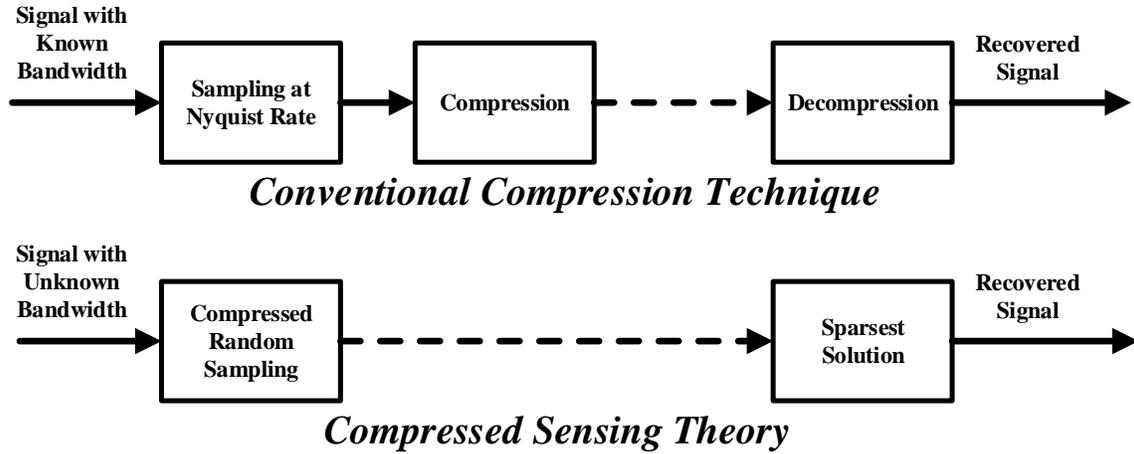


Figure 2.1: *Compressed Sensing*

basis Ψ such that,

$$\mathbf{w} = \Psi \mathbf{a}$$

where, \mathbf{a} has at most $K \ll N$ significant coefficients i.e., \mathbf{w} is K -sparse in sparsifying basis Ψ . Compressed measurements are achieved by taking $M \ll N$ random projections of \mathbf{w} ,

$$\mathbf{h} = \Phi \mathbf{w}; \quad \mathbf{h} \in \mathbb{R}^M, \quad \Phi \in \mathbb{R}^{M \times N}. \quad (2.1)$$

where, the entries of Φ are i.i.d. Gaussian random variable with mean 0 and variance $1/M$.

The original signal \mathbf{w} can be recovered by solving the following ℓ_1 minimization problem,

$$\mathbf{a}^* = \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{subject to } \mathbf{h} = \Phi \Psi \mathbf{z}$$

$$\mathbf{w}^* = \Psi \mathbf{a}^* \quad (2.2)$$

The result of the optimization problem in (2.2) provides an exact reconstruction with over-

whelming probability if there exists a $\delta \in (0, 1)$ such that,

$$(1 - \delta)\|\mathbf{z}\|_2^2 \leq \|\Phi\Psi\mathbf{z}\|_2^2 \leq (1 + \delta)\|\mathbf{z}\|_2^2,$$

holds for all K -sparse signal \mathbf{z} . This is called the Restricted Isometry Property (RIP) of order K . Correspondingly, the measurement dimension is bounded in following order [87],

$$\begin{aligned} M &= \mathcal{O}\left(\frac{1}{\delta^2}K \log\left(\frac{N}{K}\right)\right) \\ \Rightarrow \left(\frac{M}{N}\right) &= \mathcal{O}\left(\frac{1}{\delta^2}\frac{K}{N} \log\left(\frac{N}{K}\right)\right) \\ \Rightarrow CMR &= \mathcal{O}\left(\frac{1}{\delta^2}SR \log\left(\frac{1}{SR}\right)\right) \end{aligned} \tag{2.3}$$

where, CMR is compressed measurement ratio and SR is sparsity ratio.

The inherent correlations of a signal corresponds to low variability or some degree of smoothness across it, resulting in compressible coefficients in wavelet domain [27]. Hence, it is expected that, with high level of correlation, the signal will be more compressible, allowing approximation with fewer number of wavelet coefficients [28]. This will, in turn reduce the required number of random measurements.

Fig.2.2 illustrates the compressed sensing of a periodic signal. The original signal consists of 128 samples and exhibits sparsity in frequency domain, since only 4 out of 128 Fourier coefficients are nonzero. In this example, only 28 samples are aggregated through compressed sensing. As evident, the original signal can be successfully recovered from such aggregation due to the inherent sparse characteristics of the periodic signal.

The one-dimensional concept of compressed sensing can be extended to compressed imaging or 2D compressed sensing [89]:

Corollary 2.1. Let $\mathbf{W} \in \mathbb{R}^{N_{space} \times N_{time}}$ be the spatiotemporal data over N_{space} nodes and

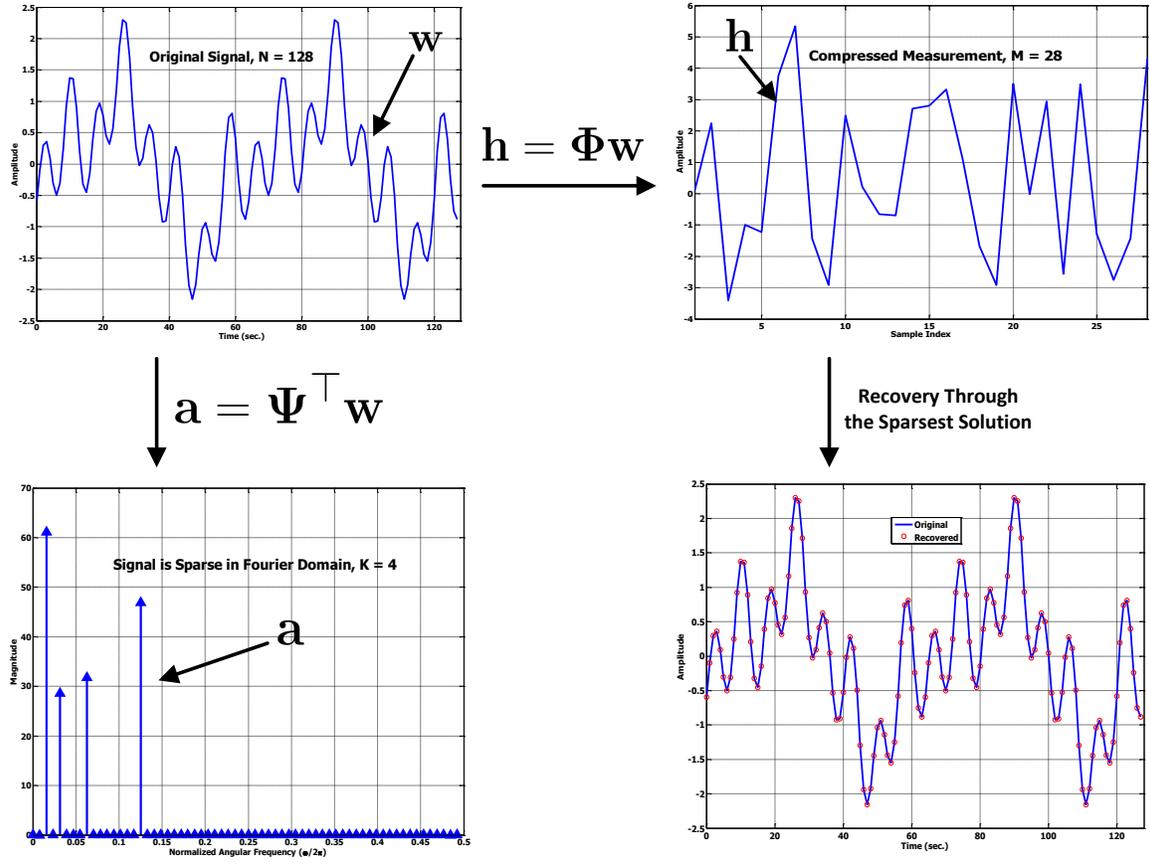


Figure 2.2: An Example of 1D Compressed Sensing

for N_{time} number of observations. \mathbf{W} is sparse in sparsifying basis $\Psi_{N_{space}}$ and $\Psi_{N_{time}}$ such that,

$$\mathbf{W} = \Psi_{N_{space}} \mathbf{A} \Psi_{N_{time}}^{\top}$$

The spatiotemporal compressed sensing of \mathbf{W} is,

$$\mathbf{H} = \Phi_{space} \mathbf{W} \Phi_{time}^{\top},$$

$$\Phi_{space} \in \mathbb{R}^{m_{space} \times N_{space}}, \quad \Phi_{time} \in \mathbb{R}^{m_{time} \times N_{time}}$$

$$m_{space} \ll N_{space} \text{ and } m_{time} \ll N_{time}$$

$$N = N_{space} \times N_{time} \text{ and } M = m_{space} \times m_{time} \quad (2.4)$$

where, the entries of Φ_{space} and Φ_{time} are *i.i.d.* Gaussian random variables with zero mean and respective variance of $1/m_{space}^2$ and $1/m_{time}^2$. The spatiotemporal data is recovered by solving the following ℓ_1 minimization problem,

$$\begin{aligned} \mathbf{a}^* &= \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1 \\ \text{subject to } \text{vec}(\mathbf{H}) &= (\Phi_{space} \otimes \Phi_{time}) (\Psi_{N_{space}} \otimes \Psi_{N_{time}}) \mathbf{z} \\ \text{vec}(\mathbf{W}^*) &= \Psi_{N_{space}} \otimes \Psi_{N_{time}} \mathbf{a}^* \end{aligned} \tag{2.5}$$

In this Corollary, $\text{vec}(\mathbf{W})$ represents the row-ordered vectorization of \mathbf{W} and \otimes represents the Kronecker product operator. We illustrate the compressed imaging in Fig.2.3, where a 64×64 image is compressively sampled to have a 50×50 image.

The concept of compressed sensing described in this section will be used in Chapters 3 and 4 in order to address the research questions 1 and 2, respectively.

2.2 State Estimation and Control

For any physical system, the term *state* summarizes the complete status of the underlying process at any given time [90]. Usually, not all of the system states can be directly obtained from the finite set observations made at any instant. As a consequence, the best possible values of these unobservable states are *estimated* from the available set of measurements/observations [91]. Primarily, the state estimation procedure can be divided into two major categories: (1) static, and (2) dynamic state estimation. Static estimation of states is possible as long as the rate of updating measurement sets is higher than the underlying system dynamics. In this scenario, the system remains in steady-state over the time horizon

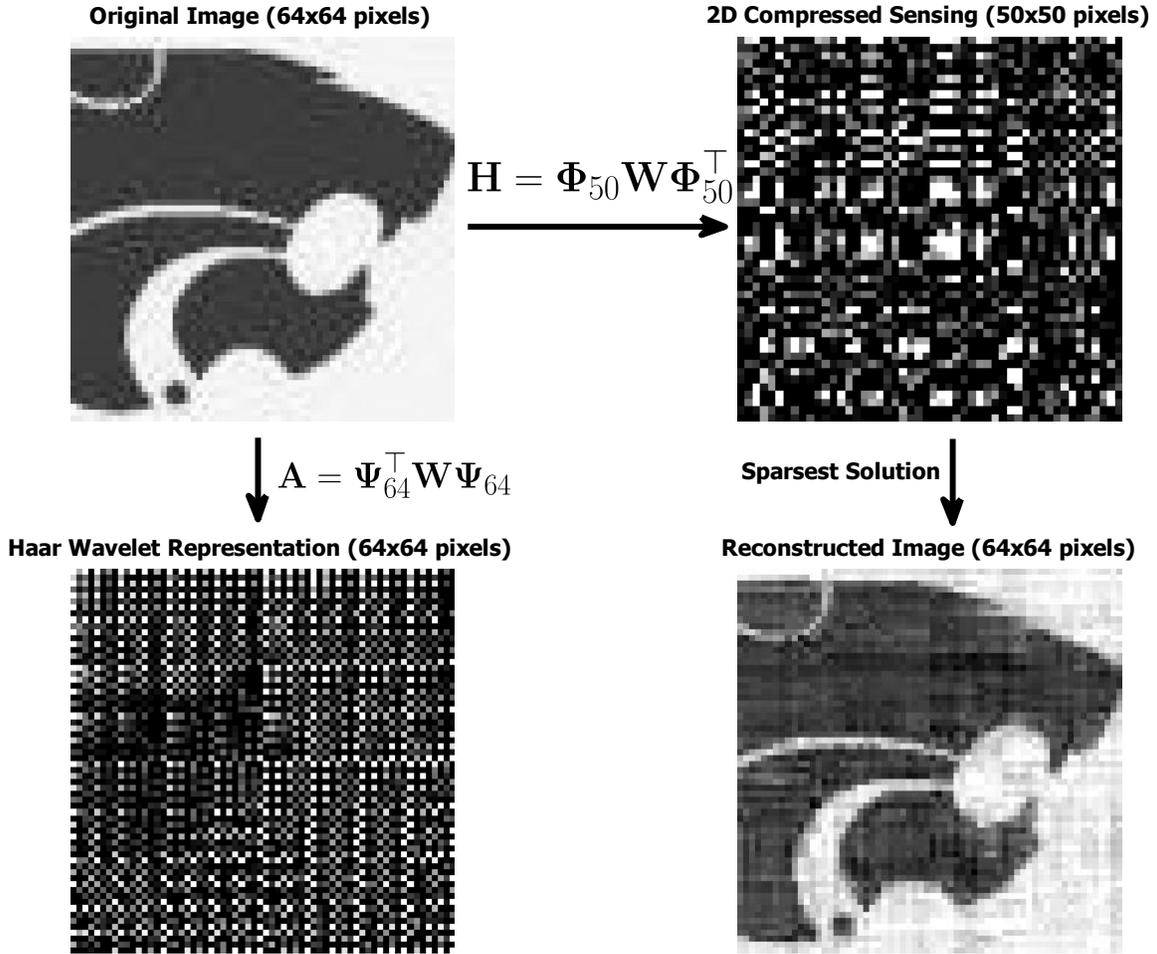


Figure 2.3: *An Example of 2D Compressed Sensing*

of state estimation. Dynamic estimation, on the other hand, takes into account the continuous evolution of the physical process so that the states can be recursively estimated from the periodically updated measurements. In the next two subsections, we briefly go over the system models and basic estimation procedures for the static and dynamic cases.

2.2.1 Static State Estimation

We assume a physical system that can be described by a state vector \mathbf{x} having L elements. In the most general sense, the measurements / observations may follow a nonlinear relationship

with \mathbf{x} ,

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) + \mathbf{w} \quad (2.6)$$

where, $\mathbf{y} \in \mathbb{R}^M$ is the vector of measurements and \mathbf{w} is a vector of independent and identically distributed (i.i.d.) Gaussian noise with zero mean and covariance \mathbf{C}_w . \mathbf{w} models the additive noise associated with the measurement procedure. The state elements are nonlinearly mapped to the measurement set by the vector of differentiable functions, $\mathbf{h} : \mathbb{R}^L \rightarrow \mathbb{R}^M$. Equation (2.6) thus represents the combined observation space of data collecting sensors. Usually, $M \gg L$ that makes equation (2.6) an over-determined system of nonlinear equations and weighted least square (WLS) is the most popular method of estimating \mathbf{x} from such system [40]. WLS finds the values of unknown state elements that minimize the squared residual error, weighted by noise variances. Mathematically, the optimum solution is,

$$\hat{\mathbf{x}} = \underset{\mathbf{z}}{\text{arg min}} [\mathbf{y} - \mathbf{h}(\mathbf{z})]^\top \mathbf{C}_w^{-1} [\mathbf{y} - \mathbf{h}(\mathbf{z})] \quad (2.7)$$

Due to nonlinear mapping, equation (2.7) represents a nonconvex problem. As a result, iterative Gauss-Newton algorithm is usually employed to obtain the solution of the problems of this kind. For this algorithm, the nonlinear measurement model of equation (2.6) is approximated upto 1st order of the corresponding Taylor series expansion. If the properties of state elements are unknown, the best linear unbiased estimate (BLUE) [92] is obtained based on this approximated linear model. Thus, at $(k + 1)^{\text{th}}$ iteration, the state elements are updated as follows,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + [\mathbf{H}(k)^\top \mathbf{C}_w^{-1} \mathbf{H}(k)]^{-1} \mathbf{H}(k)^\top \mathbf{C}_w^{-1} [\mathbf{y} - \mathbf{h}(\mathbf{x}^{(k)})] \quad (2.8)$$

where, $\mathbf{H}(k)$ is the Jacobian matrix such that, $[\mathbf{H}(k)]_{m,l} = \left. \frac{\partial h_m(\mathbf{x})}{\partial x_l} \right|_{\mathbf{x}=\mathbf{x}^{(k)}}$. The associated covariance of the BLUE estimate is $\mathbf{C}_x^{(k)} = [\mathbf{H}(k)^\top \mathbf{C}_w^{-1} \mathbf{H}(k)]^{-1}$.

On the other hand, if the state elements are known to be random with mean $\boldsymbol{\mu}_x$ and

covariance $\Sigma_{\mathbf{x}}$, the $(k + 1)^{th}$ iteration implements the linear minimum mean squared error (LMMSE) criterion through a Wiener-Hopf filter [93],

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Sigma_{\mathbf{x}}\mathbf{H}(k)^\top [\mathbf{H}(k)\Sigma_{\mathbf{x}}\mathbf{H}(k)^\top + \mathbf{C}_{\mathbf{w}}]^{-1}[\mathbf{y} - \mathbf{h}(\mathbf{x}^{(k)})] \quad (2.9)$$

The associated covariance of the LMMSE estimate is $\mathbf{C}_{\mathbf{x}}^{(k)} = \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x}}\mathbf{H}(k)^\top [\mathbf{H}(k)\Sigma_{\mathbf{x}}\mathbf{H}(k)^\top + \mathbf{C}_{\mathbf{w}}]^{-1}\mathbf{H}(k)\Sigma_{\mathbf{x}}$.

For over-determined systems, $\mathbf{H}(k)$ is full-column rank, which ensures observability. Intuitively, the rate of change of the objective function in equation (2.7) decreases as the estimate approaches the global minima of the objective function according to the recursion given in equations (2.8) and (2.9).

It is worthwhile to note that, for $\mathbf{h}(\cdot)$ being a quadratic form in equation (2.6), the outer product of state vector (i.e., $\mathbf{x}\mathbf{x}^\top$) constitutes a linear relationship with observation. Due to this linear mapping, relaxed semidefinite programming (R-SDP) can be used for state estimation [94].

The fundamental concept of static state estimation discussed in this section will be used in Chapter 5 to address the research question 3.

2.2.2 Dynamic State Estimation and Control

The operating condition of a physical process—specifically the system stability—manifests itself through the temporal evolution of state trajectories. And a typical representation of this continuous-time state-space behavior is a set of 1st order lumped-parameter differential equations [90],

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \quad (2.10)$$

The corresponding observation space for the physical process can be defined as,

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}, \mathbf{u}, t) \quad (2.11)$$

Usually, the functions \mathbf{f} and \mathbf{h} are nonlinear and hence the 1st order Taylor series approximation leads to a linear continuous-time dynamical system,

$$\dot{\mathbf{x}} = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) \quad (2.12)$$

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) \quad (2.13)$$

Here, the state vector, $\mathbf{x} \in \mathbb{R}^n$, control or input vector, $\mathbf{u} \in \mathbb{R}^p$ and the measurement or output vector, $\mathbf{y} \in \mathbb{R}^m$. Accordingly, the dimensions of the matrices $\mathbf{F}(t)$, $\mathbf{G}(t)$, $\mathbf{H}(t)$ and $\mathbf{D}(t)$ are $n \times n$, $n \times p$, $m \times n$ and $m \times p$, respectively. A linear discrete-time dynamical system is derived from this model since in practical scenario, the sensors collect discrete samples of data as well as discrete control signals are sent to the actuators in regular time interval. Let us assume the sampling interval between the discrete time instances t_k and t_{k+1} be T , which is small enough to approximate the matrices of equations (2.12) and (2.13) as piecewise constants. In other words, $\mathbf{F}(t) = \mathbf{F}(t_k); t_k \leq t \leq t_{k+1}$ and $\mathbf{G}(t) = \mathbf{G}(t_k); t_k \leq t \leq t_{k+1}$. Thus, we obtain the following linear discrete-time dynamical system,

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{u}_k \quad (2.14)$$

Where, $\mathbf{F}_k = e^{\mathbf{F}(t_k)T}$ and $\mathbf{G}_k = \int_{t_k}^{t_{k+1}} e^{\mathbf{F}(t_k)(t_{k+1}-\tau)} d\tau \mathbf{G}(t_k)$. The discrete-time observation model is given by,

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{D}_k\mathbf{u}_k \quad (2.15)$$

In this dissertation, the linear discrete-time dynamical system will be used in formulating the agent based estimation and control of cyber-physical system. Now, the system states may be characterized as random. On the other hand, there may exist the effect of model prediction uncertainty and noisy measurements as well. In this setup, the most popular method in dynamic state estimation is to minimize the mean squared error in estimation. The minimum mean squared error (MMSE) criterion for dynamical system results in the

famous Kalman filter [95]. In the basic approach, the sensors communicate with a single fusion center either directly or hierarchically to send updated measurement information in timely manner. Based on the knowledge of previous state values and system dynamics, the fusion center makes the MMSE prediction of the states. Necessary corrections are made to the predicted states based on sensor measurements. The prediction uncertainty and noisy measurements are modeled through the incorporation of additive white Gaussian noise (AWGN) in equations (2.14) and (2.15). Therefore, the linear system dynamics in discrete-time is given by,

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{B}_k \mathbf{w}_k; k = 0, 1, 2, \dots \quad (2.16)$$

Here, the initial values of the state vector elements at $k = 0$ follow Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. This characterizes the random behavior of states. We assume that the eigenvalues of \mathbf{F}_k lie within a unit circle. The process noise $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$. The optimal control inputs for such stochastic process are obtained through the solution of a finite moving horizon linear quadratic optimization problem. The associated cost function at any time instance k is,

$$J(\mathcal{U}_k) = \mathbb{E} \left[\sum_{\tau=k}^{k+k_f-1} (\mathbf{x}_\tau^\top \mathbb{X} \mathbf{x}_\tau + \mathbf{u}_\tau^\top \mathbb{U} \mathbf{u}_\tau) + \mathbf{x}_{k+k_f}^\top \mathbb{X} \mathbf{x}_{k+k_f} \right] \quad (2.17)$$

Here, $\mathcal{U}_k = \{\mathbf{u}_k, \dots, \mathbf{u}_{k+k_f-1}\}$. $\mathbb{X} \succeq 0$, $\mathbb{U} \succ 0$ and k_f defines the length of moving horizon. Hence, the desired set of control inputs between k and $k + k_f - 1$ are,

$$\{\mathbf{u}_k^*, \dots, \mathbf{u}_{k+k_f-1}^*\} = \arg \min J(\mathcal{U}_k)$$

subject to

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{u}_k + \mathbf{B}_k \mathbf{w}_k$$

With the help of dynamic programming, one can obtain the control inputs in state feedback form. A model-predictive control (MPC) is achieved when only the first entry \mathbf{u}_k^* at the time instance k is applied. The global states at $(k + 1)$ is obtained from the model (2.16) and the finite time horizon is shifted one time step forward. The linear quadratic optimization is then repeated. The stochastic control can be achieved through state feedback in terms of estimated states multiplied by a feedback control gain matrix. The control matrix is the solution of the linear quadratic optimization problem, which can be calculated through the following steps of dynamic programming:

- Define time horizon

$$\tau = k + k_f : -1 : k + 1 \quad (2.18)$$

- Initialization

$$\mathbf{Y}_\tau = \mathbb{X} \quad (2.19)$$

- State feedback control gain

$$\mathbf{C}_{\tau-1} = -(\mathbf{U} + \mathbf{G}_\tau^\top \mathbf{Y}_\tau \mathbf{G}_\tau)^{-1} \mathbf{G}_\tau^\top \mathbf{Y}_\tau \mathbf{F}_\tau \quad (2.20)$$

- Update

$$\mathbf{Y}_{\tau-1} = \mathbf{F}_\tau^\top \mathbf{Y}_\tau (\mathbf{F}_\tau + \mathbf{G}_\tau \mathbf{C}_{\tau-1}) + \mathbb{X} \quad (2.21)$$

The dynamic state estimation is then carried out based on some set measurements collected at the fusion center. Now, the fusion center observation model is,

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (2.22)$$

Here, the measurement noise $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ and is independent of process noise. The state vector estimated at discrete time instant k is defined as,

$$\hat{\mathbf{x}}_{k|j} = \mathbb{E} [\mathbf{x}_k | \mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_j] \quad (2.23)$$

The corresponding error covariance matrix is,

$$\mathbf{M}_{k|j} = \mathbb{E} [(\mathbf{x}_k - \hat{\mathbf{x}}_{k|j})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|j})^\top] \quad (2.24)$$

Based on these definitions, we present the basic steps of a standard Kalman filter:

- Initialization:

$$\hat{\mathbf{x}}_{0|0} = \boldsymbol{\mu}, \mathbf{M}_{0|0} = \boldsymbol{\Sigma} \quad (2.25)$$

- Prediction:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{G}_{k-1} \mathbf{u}_{k-1} \quad (2.26)$$

- Predict error covariance:

$$\mathbf{M}_{k|k-1} = \mathbf{F}_{k-1} \mathbf{M}_{k-1|k-1} \mathbf{F}_{k-1}^\top + \mathbf{B}_{k-1} \mathbf{Q}_{k-1} \mathbf{B}_{k-1}^\top \quad (2.27)$$

- Kalman Gain:

$$\mathbf{K}_k = \mathbf{M}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{M}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \quad (2.28)$$

- Correction:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}) \quad (2.29)$$

- Correct Error Covariance:

$$\mathbf{M}_{k|k} = \mathbf{M}_{k|k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{M}_{k|k-1} \quad (2.30)$$

- State feedback control:

$$\mathbf{u}_k = \mathbf{C}_k \hat{\mathbf{x}}_{k|k} \tag{2.31}$$

Based on these basic Kalman filtering steps, we will develop the agent-based Kalman filters in Chapter 6 and derive the necessary conditions for stable estimation procedure in Chapters 7 and 8. Similarly, the basic state feedback control procedures discussed in this chapter will be used in Chapter 9 to design the agent based model predictive control.

Chapter 3

Impact of Correlated Distributed Generation on Information Aggregation in Smart Grid

Real-time control of a smart distribution grid with renewable energy based generators requires accurate state estimates, that are typically based on measurements aggregated from smart meters. However, the amount of data/measurements increases with the scale of the physical grid, posing a significant stress on both the communication infrastructure as well as data processing control centers. In this chapter, the effect of geographical footprint of distributed generation (DG) on the voltage states of a smart distribution system is investigated. We demonstrate that the strong coupling in the physical power system results in estimated voltage phasors exhibiting a correlation structure that allows for compression of measurements. Specifically, by exploiting principles of 1D and 2D compressed sensing, we illustrate the effectiveness of voltage estimation with significantly low number of random spatial, temporal as well as spatiotemporal power measurements. Results demonstrate the importance of accounting for correlation in information aggregation in smart grids.

3.1 Introduction

The electric power distribution network we know today is expected to transform to a “smart” network in the near future, enabling it to deliver affordable electric power with improved reliability, greater security and more efficiency [22]. A smart distribution grid is envisioned to contain distributed generations (e.g., solar, wind) along with advanced metering infrastructure (AMI), which will deploy communicable meters and actuators throughout the distribution grid to collect data and act on it (e.g., use it for state estimation and / or control). Consequently, an unusual increase of raw measurements is expected with increase in levels of DG penetration [19]. On the other hand, the sampling interval of commercially available smart meters can vary in the range of a few seconds to the order of minutes. Devices are currently being used to record solar generation with a temporal resolution of 1 second [96] and 15 seconds [18]. Reliable wireless communication of the resultant high volume of end-user information through advanced meters and actuators is an important issue in designing a real-time dynamic smart grid. It should be noted that, even under optimal circumstances, the effective throughput of underlying network is drastically reduced with the increase in sensors communicating at the same time [17]. As a consequence, the inherent grid structure should be efficiently utilized to reduce the number of active sensors/measurements while maintaining desired reliability and efficiency throughout the power system.

With the anticipated increase in distributed generators (DGs), measurements related to the power generated by the DGs are important for real time estimation and control of smart grid. DGs may be installed at the end-user premises of a power distribution grid [23]. The geographical proximity of these renewable energy sources causes the power generated from DGs to be correlated [24, 25]. This correlation structure induces sparsity in appropriate transformation basis [26, 27]. This in turn enables compressibility of spatial power measurements i.e., we can reduce the amount of aggregated power information at a time using concepts from *Compressed Sensing* introduced in Chapter 2. Temporal compressed sensing can be used at a node for efficient compression of current and stored load /gener-

ation data. The compressed information can then be sent to the utility/customer to study load/generation profile for that particular node. This also motivates the use of 2-dimensional spatiotemporal compressed sensing that enables information aggregation across both space and time. Recently, compressed sensing has been applied in simultaneous reading of all smart meters in order to securely transmit the sensed data and to ensure identical delay in wireless transmission of all smart meter readings [29, 30]. The compressed reading is obtained using pseudorandom spreading sequence. It is assumed that, the meters convey information only about the significant changes of the respective power levels. In [29], compressed measurements are obtained in presence of bounded noise whereas, the measurement noise is assumed to be normally distributed in [30]. In [31], an automated compressed meter reading scheme is proposed for a wireless home area network, where the smart meters are equipped with Zigbee standard communication devices. However, usage of such measurements for estimating states of corresponding power system is not reported in [29, 30, 31].

In this chapter, the correlation of voltage phasors is analyzed in the presence of correlated generation from DGs. Unlike prior works, correlation of distributed generation is considered in time as well as in space [79]. Approximate sparsity is obtained by transforming the correlated data into wavelet domain. Then, 1D (space, time) and 2D (space and time) compressed sensing is applied to aggregate the power readings from the distribution nodes. Finally, the voltage phasors are estimated from the power readings using power flow equations.

3.2 General System Model

We consider an N node single phase radial power distribution system as shown in Fig.3.1. The “grid” is designated as node 0 and all other nodes are denoted as $1, 2, \dots, N$. The electrical characteristics of all nodes may not be the same. Some nodes may act as a load only, whereas some others may contain photovoltaic (PV) panels or wind generators serving

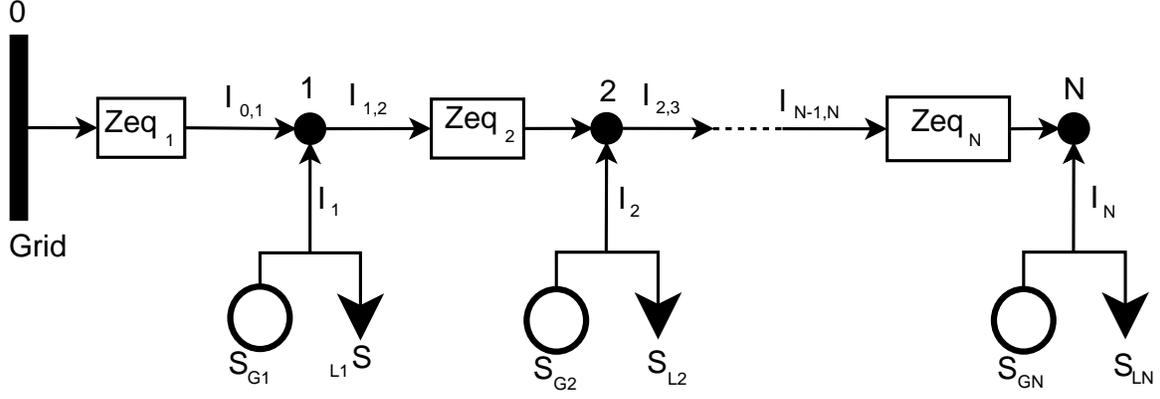


Figure 3.1: *Single Phase Distribution Network with Renewable Energy Source.*

as renewable energy sources. It is also possible for a node to have both loads and renewable energy sources. The electrical coupling among these nodes are based on Kirchoff's current and voltage laws. The inter-node series impedances and shunt admittances determine the coupling strength.

According to Kirchoff's current law, the i^{th} node injection current is given by,

$$I_i = \begin{bmatrix} -\frac{1}{Zeq_i} & \frac{1}{Zeq_i} + \frac{1}{Zeq_{i+1}} & -\frac{1}{Zeq_{i+1}} \end{bmatrix} \begin{bmatrix} V_{i-1} \\ V_i \\ V_{i+1} \end{bmatrix} \quad (3.1)$$

where, $\frac{1}{Zeq_i}$ is the single-phase equivalent admittance between node $i - 1$ and node i [97]. V_i is the i th node voltage. Using (3.1), the voltage-current relation for the whole network can be written in matrix-vector form,

$$\mathbf{I}_{(N+1) \times 1} = \mathbf{Y}_{(N+1) \times (N+1)} \mathbf{V}_{(N+1) \times 1} \quad (3.2)$$

A similar formulation for a three-phase system with each phase represented separately can be done to enhance applications of the concept presented in this paper.

The magnitudes and angles of voltage phasors at different nodes of the distribution net-

work constitute the state of the overall system. The estimation of these states plays an important role in voltage regulation and control of reactive power flows. However, measuring voltage phasors from every node of a distribution grid is not economically feasible. At present, energy usage of every end-user of conventional distribution grid is recorded for billing purposes only. With the evolution of the smart grid, these energy meters are expected to be upgraded so that the utility companies can monitor the usage of energy at regular intervals for improved estimation and controllability of the distribution system. Thus, the measurements of generation and load at every node may be readily available in smart distribution system.

Accurate estimation of voltage states is dependent upon the measurement of generation and load at different nodes of the smart distribution network. As shown in Fig.3.1, S_{Gi} and S_{Li} representing the respective apparent power generation and load at i th node, satisfy the following power flow equations:

$$\begin{aligned}
 P_i &= \text{real}(S_{Gi} - S_{Li}) = |V_i| \sum_{k=0}^N |V_k| (G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik}); \\
 Q_i &= \text{imag}(S_{Gi} - S_{Li}) = |V_i| \sum_{k=0}^N |V_k| (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik}) \quad (3.3)
 \end{aligned}$$

where,

$$[\mathbf{Y}]_{ik} = Y_{ik} = G_{ik} + jB_{ik};$$

$$V_i = |V_i| e^{j\theta_i}$$

and

$$\theta_{ik} = \theta_i - \theta_k.$$

In the above equation, P_i and Q_i represent the i th node injected real and reactive power, respectively. We assume that $|V|_0$ and θ_0 are known and define rest of the voltage phasors

as a composite vector:

$$\mathbf{x} = \left[\theta_1 \quad \cdots \quad \theta_N \quad |V|_1 \quad \cdots \quad |V|_N \right]^T$$

For the entire system, the injected powers are related to \mathbf{x} according to (3.3). This can be represented in the vector form:

$$\mathbf{y} = \left[P_1 \quad \cdots \quad P_N \quad Q_1 \quad \cdots \quad Q_N \right]^T = \mathbf{F}(\mathbf{x}) \quad (3.4)$$

Thus, voltage phasors i.e., \mathbf{x} can be estimated from the available measurements \mathbf{y} using Newton-Raphson iterative algorithm given a fully determined system [98]. An initial guess of all unknown voltage phasors is required in Newton-Raphson algorithm. Throughout this paper, *flat start* is used, in which the unknown voltage angles are set to zero and unknown voltage magnitudes are set to $1.0p.u.$ In the subsequent sections, this algorithm will be used to investigate the effect of spatial correlation of power generated on voltage phasor estimation.

3.3 Correlation Analysis in Distribution Grid

Since power and voltage phasors are nonlinearly related (3.3), it is important to analyze the inherent correlation among the distribution voltage phasors for a given set of spatially correlated power. As an illustrative example, we consider an exponential model for spatial correlation. That is, the correlation between the generated powers at two DGs spaced d miles apart corresponds to an exponentially decaying function of d , parameterized with θ_1 and θ_2 [99]:

$$r(d_{ij}) = \exp \left\{ - \left(\frac{d_{ij}}{\theta_1} \right)^{\theta_2} \right\}; \quad \theta_1 > 0, \quad \theta_2 \in (0, 2] \quad (3.5)$$

where, $r(d_{ij})$ is the coefficient of correlation and d_{ij} is the Euclidean distance between the i^{th} and j^{th} DG. This model becomes exponential, if $\theta_2 = 1$ and squared exponential for $\theta_2 = 2$.

We perform the correlation analysis of voltage phasors for two cases. In the first case, the generated power quantities are assumed only to be spatially correlated. And the loads are assumed to be completely random. The temporal dependence is considered for both load and spatially correlated generation in the second case. For this, we model the distributed generation and load as first order auto-regressive (AR(1)) process [100]. The generalized time-series representation of AR(1) apparent generation (S_g) and load (S_l) is:

$$S_{t+1} - D_{t+1} = \alpha(S_t - D_t) + W_t \quad (3.6)$$

where, S is common notation for S_g and S_l . α denotes the AR(1) coefficient. W_t is the corresponding noise component and t is the time in minute. D_t is the deterministic trend component of the time series, which is defined in per unit values[101]:

- for S_g : $0.35 + 0.025 \sin\left(\frac{6\pi t}{T} + 0.2\pi\right) + 0.1 \sin\left(\frac{\pi t}{T}\right) + 0.025 \sin\left(\frac{6\pi t}{T} + 0.3\pi\right)$;
- for $P_l = \text{real}(S_l)$: $0.6 + 0.05 \sin\left(\frac{6\pi t}{T} + 0.1\pi\right) + 0.01 \sin\left(\frac{\pi t}{T}\right)$ and
- for $Q_l = \text{imag}(S_l)$: $0.3 + 0.05 \sin\left(\frac{6\pi t}{T} + 0.1\pi\right) + 0.01 \sin\left(\frac{\pi t}{T}\right)$

Here, T is the number of equidistant samples taken over 1440 minutes (i.e. 24 hours). These functions are chosen for illustrations. However, other functions that can capture the periodic variation in these variables can be used.

A single-phase equivalent test system based on the IEEE 34 node distribution test feeder [102] is used in this paper to perform the proposed correlation analysis. The base kVA and base kV is assumed to be 100 and 24.9, respectively. As shown in Fig.3.2, node 800 is the distribution grid and there is a transformer between node 832 and node 888. In our analysis we replace the transformer with an equivalent circuit and merge node 888 to node 832. We mark the grid node i.e., node 800 as node 0. Thus, we have 32 nodes. We assume that each

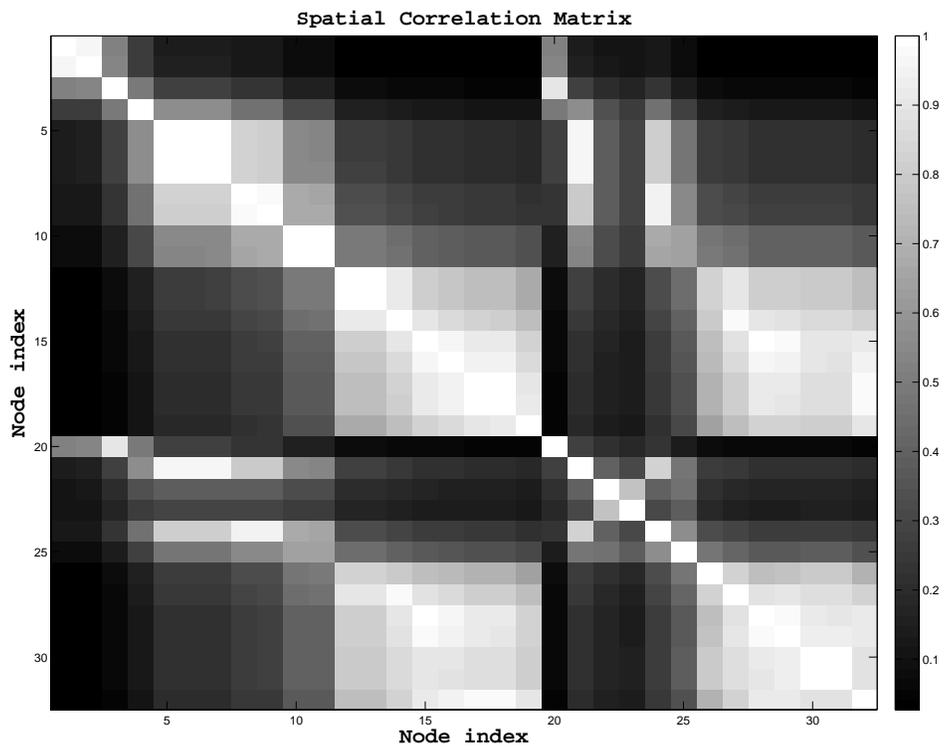


Figure 3.3: *Spatial Correlation Matrix of Distributed Generation based on (3.5)*

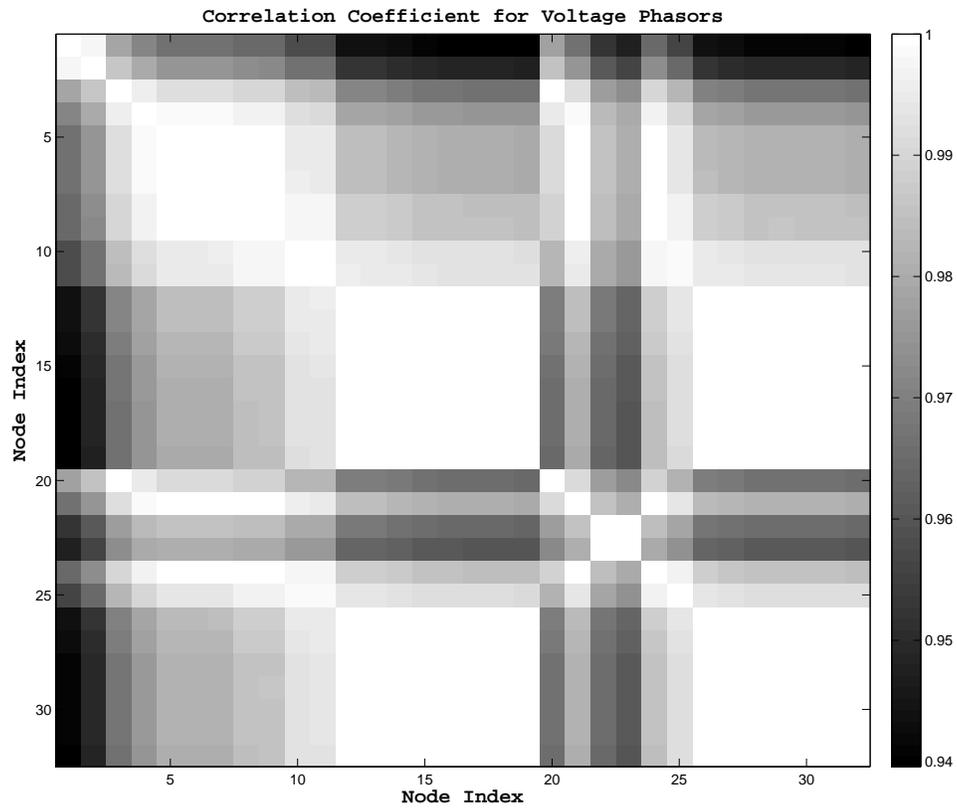


Figure 3.4: *Correlation of Voltage Phasors for Spatially Correlated Generation and Random Load*

load and generation. Both the magnitudes and angles of voltage phasors exhibit similar correlation structure [79]. The high correlation of voltage phasors follows the electrical coupling among distribution nodes of a radial system. An increase of load at any node will cause excessive drawing of current and consequently all the nodes along the path of this current will experience a drop in voltage magnitude and phase shift in voltage angles. Conversely, a decrease of load or increase of distributed generation at a particular node will contribute to the rise of voltage levels of all other nodes. The degree of voltage rise or drop among the nodes follow the voltage-current relation in (3.2). Thus the admittance matrix has a direct impact on having a unique spatial correlation structure of voltage phasors for a particular distribution system.

3.4 Information Aggregation in Distribution Grid

As mentioned earlier, conventional state estimation method requires all available power measurements. Thus for a practical distribution system, large number of measurements have to be communicated and aggregated for estimation. This would require a communication link with large bandwidth and high reliability. Henceforth, it is difficult to say whether the design and maintenance of a communication link with high data overload and sophisticated error correction scheme would be economically feasible or not. This drawback can be alleviated in one way by attempting to reduce the number of measurements transmitted to the estimator. We exploit the compressibility of correlated data to reduce the number of measurements using theory of *Compressive Sensing* [87]. In Section 2.1, we discuss the fundamentals of this concept and also describe the conditions for successful recovery of original data. To better explain the role of compressed sensing method in the context of information exchange we present one possible architecture. Let us consider our model of information

aggregation as depicted in equation (2.1),

$$\mathbf{h} = \Phi \mathbf{w} = \sum_{i=1}^N w_i \phi_i \quad (3.7)$$

where, each entry of the vector \mathbf{w} is an observation value or measurement obtained in the respective sensor. For further analysis, these measurements are required to be sent to a fusion center. The i^{th} sensor encrypts its measurement using the respective M -dimensional column vector ϕ_i . The sensor at the farthest location need not send its information directly to the fusion center. It can communicate with a sensor nearest to it. Upon receiving the encrypted message, the neighboring sensor can aggregate the received measurement with its own encrypted measurement and send it to its neighbor. This hierarchical process of local information exchange is repeated sequentially all the way to the fusion center. As illustrated in Fig. 3.5, the memory requirement at each sensor is always limited to $M = m_{space}$ and does not grow with the increase in number of aggregated measurements. Furthermore, the communication range for all sensors (even for the farthest one) is uniform and moderate since not every sensor has to directly link to the fusion center. Upon receiving the compressed information, the fusion center can make a *one-time recovery of the entire set of measurements* as long as the actual information is at least approximately sparse in some linear transformation basis. Thus, the memory usage and processing time can be considerably reduced especially when the number of active sensors $N = N_{space}$ is excessively large. This will, in turn, enhance the capability of real-time monitoring and control of the whole system.

In this study, we only know that the signal under consideration is approximately sparse in wavelet domain. Mathematically, in equation (2.3), we don't have the exact knowledge of SR , rather an overall effect of compressed sensing is obtained by varying the CMR from as low as 10% to 100% and observing the reconstruction error at CMR level.

Based on the discussion above, we apply Theorem 2.1, where \mathbf{w} stands for the power

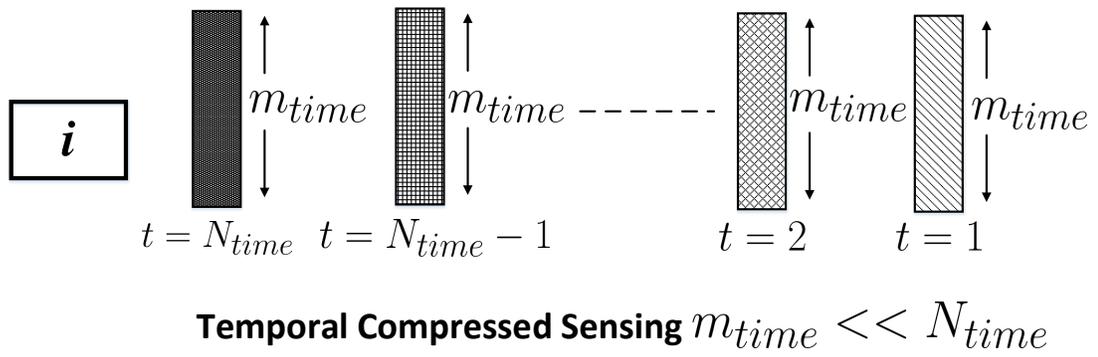
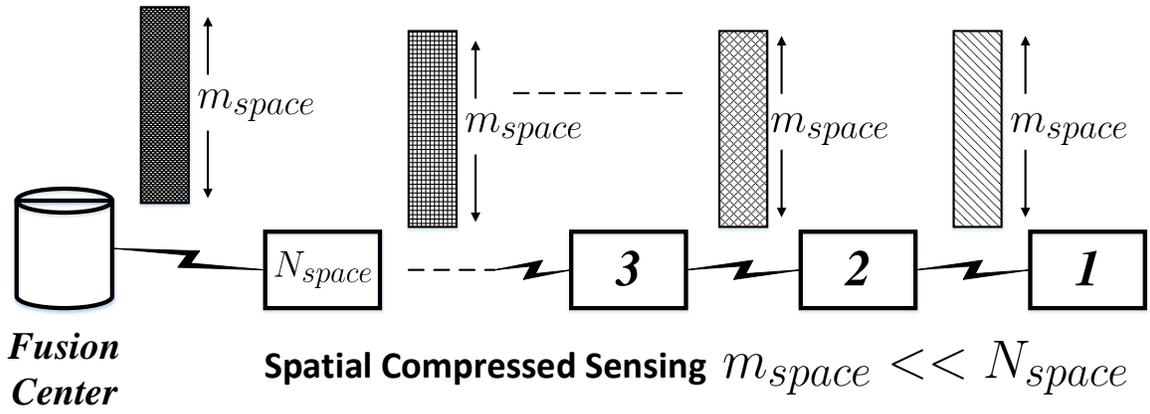


Figure 3.5: *Illustration of Spatial and Temporal Compressed Sensing*

generated from DGs and Ψ is wavelet transform basis. The random measurements of power can be taken across space (spatial compressed sensing) or across time (temporal compressed sensing). The compressions thus achieved are illustrated in Fig. 3.5. Alternately, we can randomize the power measurements across both space and time (spatiotemporal compressed sensing) applying the concept of compressed imaging or 2D compressed sensing [89].

It should be noted that, equation (2.3) gives a sparsity level based lower limit of reducing compressed measurement, irrespective of the transformation basis. However, the numerical process of recovering original signal or image from compressed measurement does not require specific information about the sparsity level. Based on this observation, we explore the effect of compressed measurement dimension on the accuracy of recovering signal or image which are known to be approximately sparse in wavelet domain.

We investigate the effect of CMR on the recovery performance of different compressed sensing phenomena for the IEEE 34 node distribution test feeder as used in Section 3.3. We consider spatial, temporal and spatiotemporal compressed sensing. For each of the cases, we take different number of random measurements and use (2.2) and (2.5) to reconstruct the original signals. The ℓ_1 minimization problems in (2.2) and (2.5) are solved using `cvx` ([103], [104]). In our analysis, we use ‘‘Haar’’ mother wavelet to define the sparsifying basis Ψ [105].

3.5 Results

We investigate the performance of the three compressed sensing phenomena by performing 1000 Monte Carlo simulation at different CMRs. For the purpose of comparison, the measurements of real generated power P_g are used in this simulation. The performance variability of spatial, temporal and spatiotemporal compressed sensing is compared using a percentage ratio, called *Integrated Normalized Absolute Error (INAE)*. For 1-D (spatial or

temporal) case, we define *INAE* as,

$$INAE = \frac{\sum_{j=1}^N |x_j - x_j^*|}{\sum_{j=1}^N |x_j|} \times 100 \quad (3.8)$$

where, $x_j = [\mathbf{x}]_j$ is the real power generated at j^{th} DG (spatial) or at the j^{th} time instance (temporal). The lower *INAE* is, the better is the reconstruction performance. For the spatiotemporal case, *INAE* corresponds to

$$INAE = \frac{\sum_{j=1}^{N_{space}} \sum_{k=1}^{N_{time}} |X_{jk} - X_{jk}^*|}{\sum_{j=1}^{N_{space}} \sum_{k=1}^{N_{time}} |X_{jk}|} \times 100 \quad (3.9)$$

where, $X_{jk} = [\mathbf{X}]_{jk}$ is the real power generated at j^{th} DG, which is measured at the k^{th} time instance. Using (3.8) and (3.9), we show the average variation of *INAEs* for the three types of compressed sensing in Fig.3.6. Here, we observe that, the accuracy in reconstruction corresponds to the respective level of sparsity obtained in different schemes of compressed sensing. However, for all the three cases, the reconstruction error can be kept within 10% if the number of random measurements is at least 50% of the original data dimension [79]. Bearing this in mind, we show example reconstructions from spatial and temporal compressed measurements in Fig.3.5 and Fig.3.8, respectively.

We extend our analysis by observing the effect of compressed power measurements on the estimation of voltage phasors. For this purpose, equation (3.4) is used to estimate voltage magnitudes and angles using the injected apparent power readings. Fig. 3.9 shows the estimated voltages for all the nodes of the distribution grid. The “+” sign denotes the voltage phasors estimated from direct measurements of injected power. The “o” sign indicates the estimated results, when the injected power information are recovered from 50% compressed measurements. It is interesting to see that, the voltage phasors estimated from compressed measurements are almost identical to those estimated from direct measurements. Thus, only half of the available smart meter readings, which are randomly selected, need to be aggregated via a communication link for reasonable accuracy in voltage estimation.

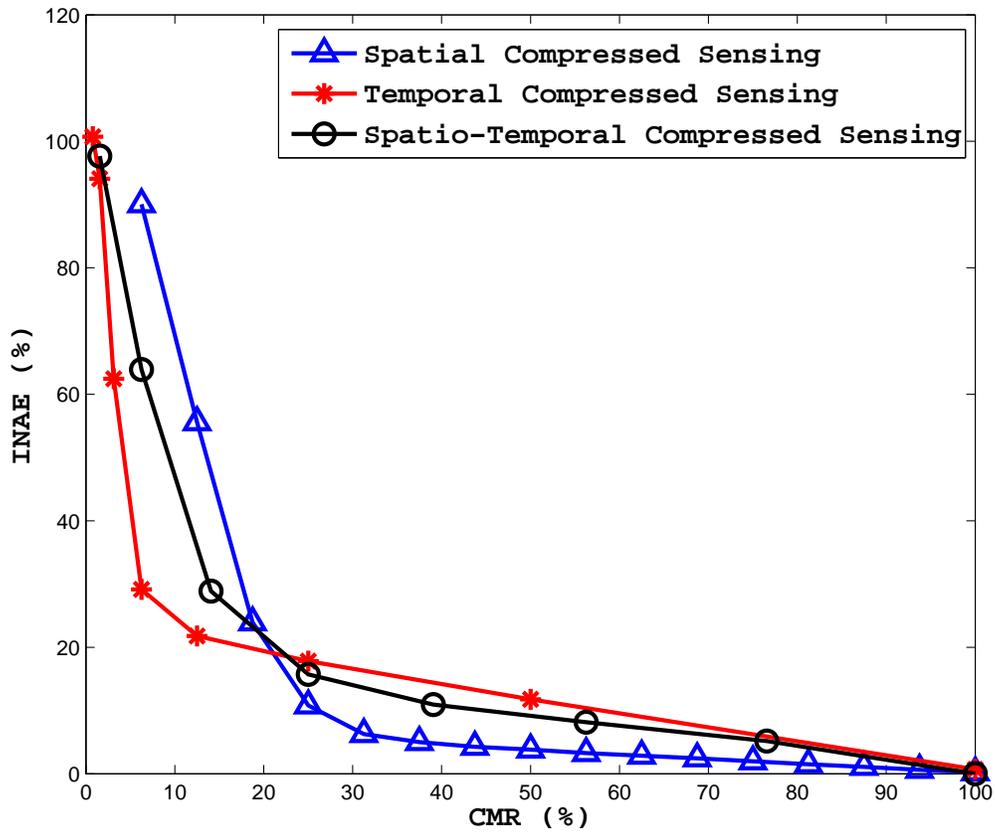


Figure 3.6: *Variation of INAE for Spatial, Temporal and Spatiotemporal Compressed Sensing*

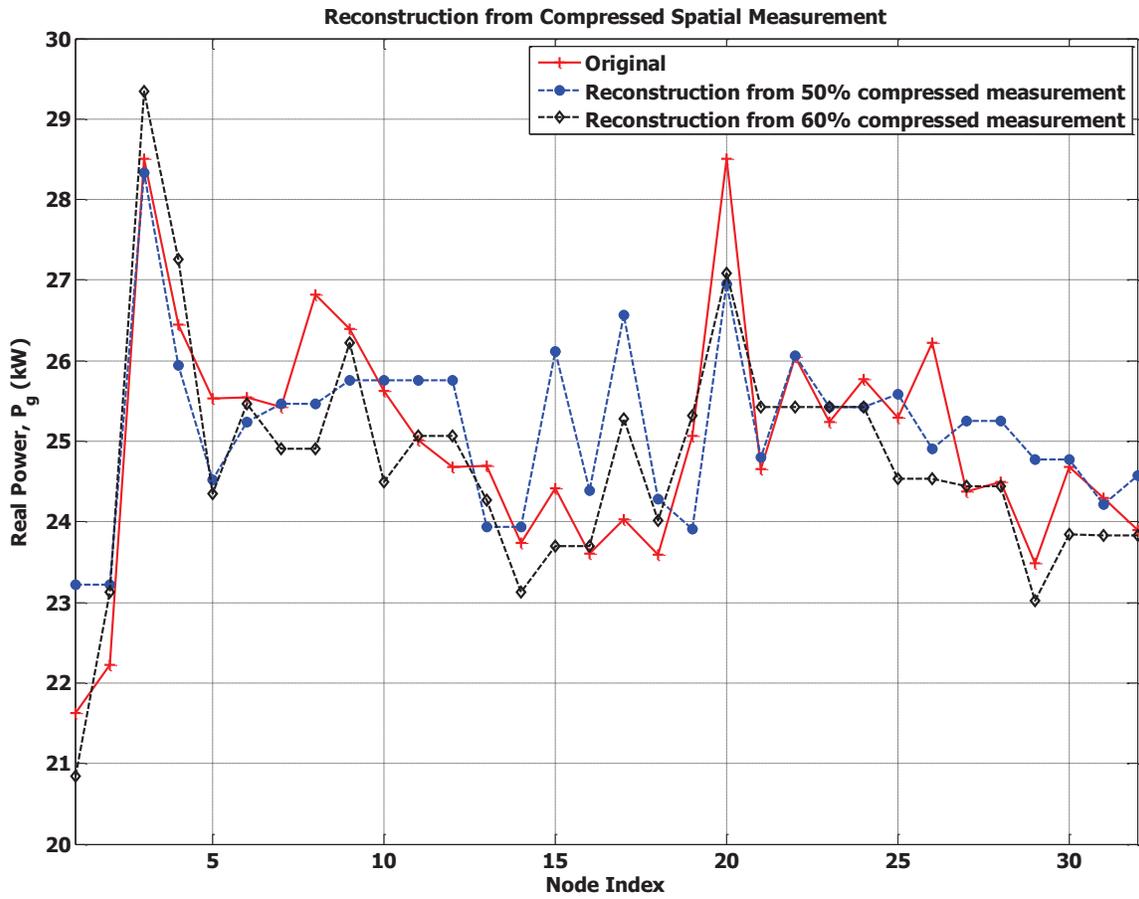


Figure 3.7: Reconstruction from Spatially Compressed Real Power Measurements

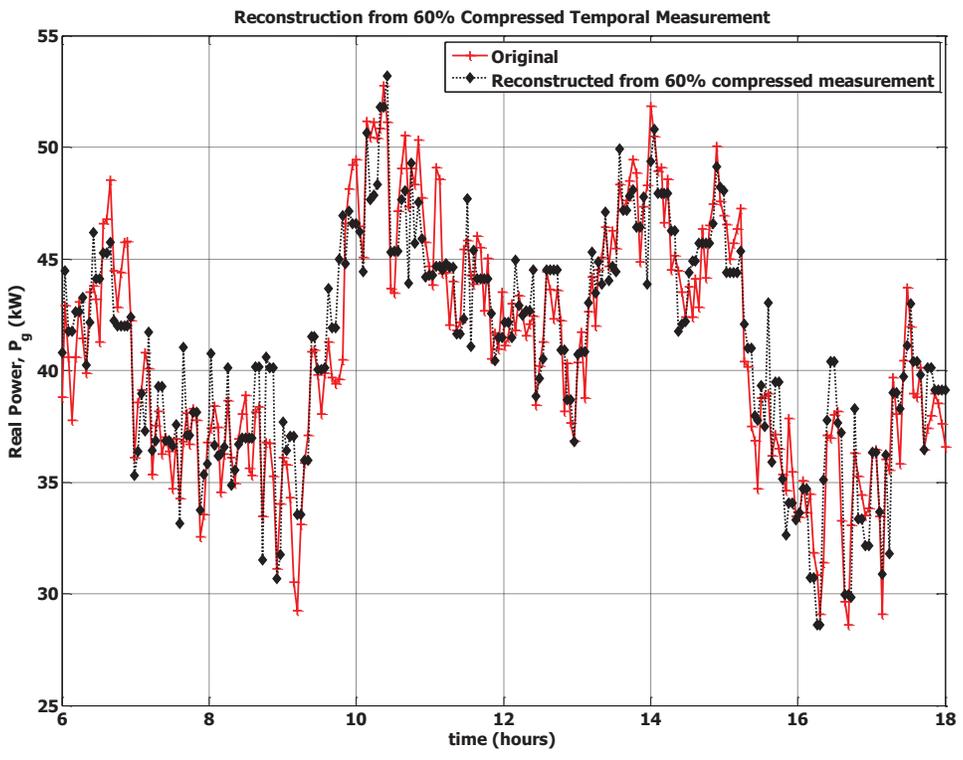
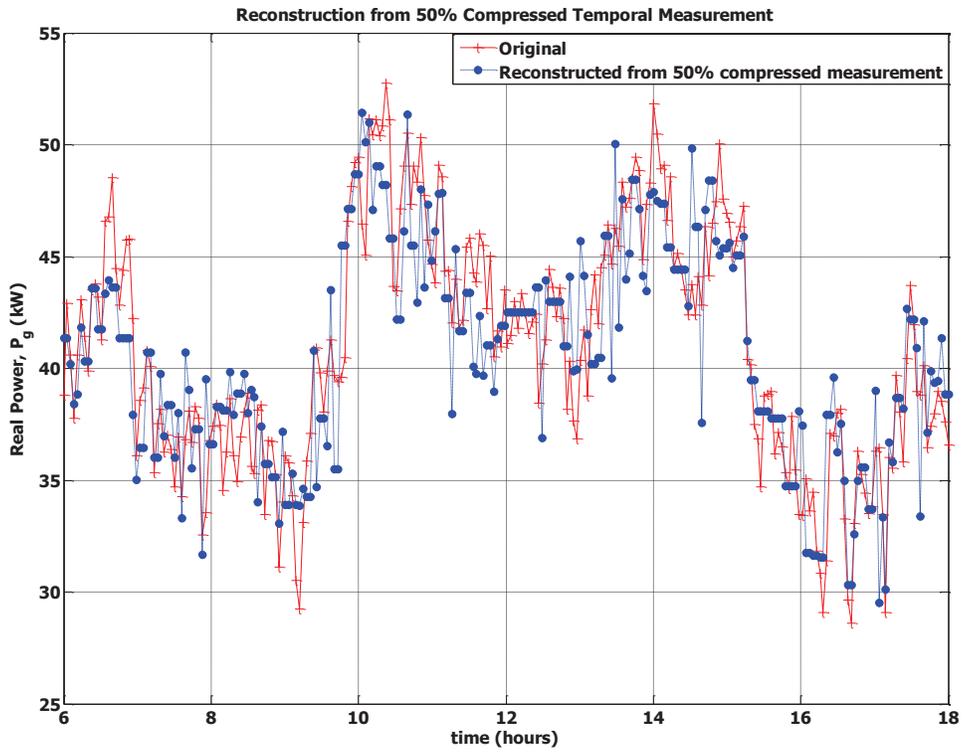


Figure 3.8: *Reconstruction from Temporally Compressed Real Power Measurements*

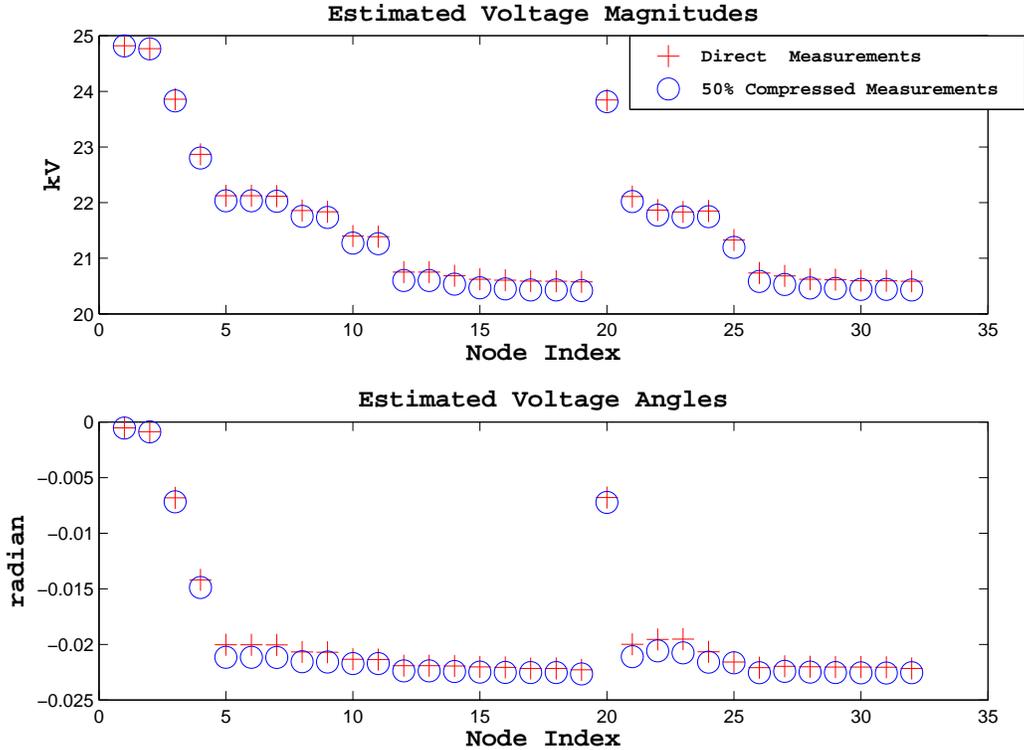


Figure 3.9: Voltage Phasors Estimated from Direct and Compressed Power Measurements

3.6 Conclusions

In this chapter, the effect of correlated distributed generation on the voltage phasors in a smart distribution grid is analyzed. It is observed that the estimated voltage phasors exhibit strong correlation. This observation allows us to employ compressed sensing in aggregating information from voltage/power sensors, which are expected to be used in future smart grid. The correlated nature of distributed generation also permits significant reduction of required power measurements in spatial, temporal as well as spatiotemporal domain. A comparative study shows that, the voltage phasors can be estimated almost accurately with only a random half of the available measurements. Thus, compressed reading of power as well as voltages is expected to have considerable influence in the measurement and sensing technologies associated with smart grid. Our next step is to circumvent the need of recon-

structuring power measurements for voltage estimation i.e., we aim to combine compressed sensing and power flow equations in order to develop a direct method to estimate voltage states. This challenge is addressed in Chapter 4.

Chapter 4

Distribution Grid State Estimation from Compressed Measurements

In Chapter 3 the potential application of 1D and 2D compressed sensing in correlated measurement aggregation is illustrated through significant reduction in the requirement of raw data collection. Starting from these compressed power measurements, we develop two approaches, an *indirect* and *direct* method for state estimation. We illustrate the effectiveness of voltage estimation with significantly low number of random spatial, temporal as well as spatiotemporal power measurements using the IEEE 34 node distribution test feeder and a larger 100 node radial distribution system. Results show similar performance for both methods at all levels of compression. It is observed that, even with only 50% compressed power measurements, both methods estimate the states of the test feeder with high level of accuracy.

4.1 Introduction

Real-time monitoring and active control of smart distribution grid is contingent upon the knowledge of existing system states, network topology and timely update of any change in

the network architecture. In this regard, a brief survey of the state-of-the-art estimators is given in [40]. A multi-microgrid state estimation is proposed in [106], which incorporates the network topology identification as well as is robust to islanded mode of operation. In [107], an event-triggered recursive Bayesian estimator is developed for topology identification. The change in network topology is identified in [108] by performing normalized residual test to detect bad status of switching devices.

The states of a power system are usually defined in terms of magnitudes and angles of either system node voltages or branch currents [32],[33]. These states form a system of nonlinear equations with the available information of real and reactive power,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) \tag{4.1}$$

where, $\mathbf{F}: \mathbb{R}^L \rightarrow \mathbb{R}^N$ is differentiable and non-linearly maps the L -dimensional state vector \mathbf{x} to the N -dimensional power vector \mathbf{y} . Ideally, Newton-Raphson method can be used to get exact solution of the system of equations (4.1), if it is fully determined (N equals L)[98]. However, in a conventional power distribution system, real and reactive power measurements are not readily available from each customer. Furthermore, a distribution network consists of loads characterized in several categories such as residential, commercial etc. Typically, the load demands of similar class of customers (e.g., residential, commercial) is expected to be highly correlated. As a result, pseudo-measurements are synthesized using historical data and real-time measurements, so that an overdetermined system of equations can be formed (i.e., $N > L$) [34, 35]. Such a system of nonlinear equations is then solved using weighted least square (WLS) method (Section 2.2.1) and its variants [33]. It is worthwhile to mention that, use of correlated load information in generating pseudo-measurements reduces the variability in statistical distribution of estimated power system states [109, 110, 111].

In order to estimate the states of power distribution system with compressed power measurements we need to find a sparse solution to equation (4.1). In the context of compressed sensing (discussed in Section 3.4), an M dimensional compressed measurement vector is

obtained by taking $M \ll N$ random linear projections of available N original readings for a fully determined nonlinear system. Mathematically, this process is performed by multiplying both sides of equation (4.1) by a random projection matrix $\Phi: \mathbb{R}^N \rightarrow \mathbb{R}^M$,

$$\begin{aligned} \mathbf{h} &= \Phi \mathbf{y} \\ \Rightarrow \mathbf{h} &= \Phi \mathbf{F}(\mathbf{x}) \end{aligned} \tag{4.2}$$

where,

$$\mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^N; \text{ (} N \text{ equals } L \text{)}$$

and

$$\Phi \mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^M; M \ll N$$

Thus, a fully determined system of nonlinear equations is *linearly mapped* into an underdetermined system. It should be noted that, conventional underdetermined nonlinear systems are the direct mathematical models of corresponding physical system. A typical example is a set of parameterized nonlinear equations, arising from an attempt to solve a determined system in \mathbb{R}^M by continuation methods [36]. Another example is the method of finding an interior point of the polytope $\{\mathbf{z} \geq \mathbf{0}, \mathbf{A} \in \mathbb{R}^{M \times N}: \mathbf{A}\mathbf{z} = \mathbf{b}\}$, by using the non-linear transformation ($z_i = e^{x_i}, i = 1, \dots, N$) [37]. As a third example, we quote a recent article about the in-silico manipulation of biological signaling pathways, which is modeled as an underdetermined system of nonlinear equations [38]. In prior literature, Newton-Raphson method is reported to be the most suitable to solve these kinds of underdetermined nonlinear equations [37]. However, conventional Newton-Raphson algorithm can give only the least square solution with the use of pseudo-inverse of underlying Jacobian matrix. On the other hand, search for the sparse solution is expected to be the appropriate

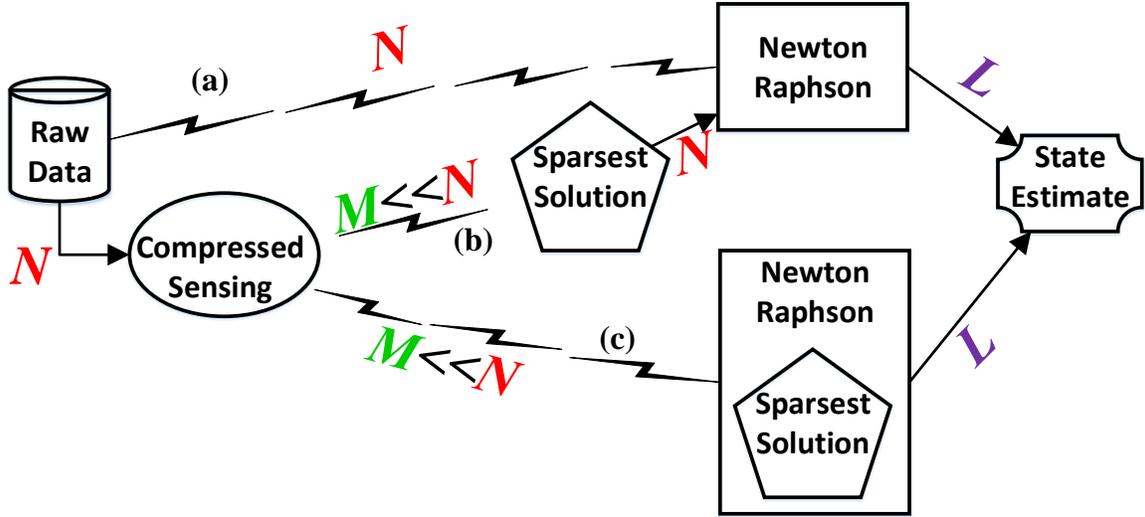


Figure 4.1: *Centralized State Estimation: (a) Conventional, (b) Indirect Method, (c) Direct Method*

one for an underdetermined nonlinear system as in equation (4.2). Therefore, an alternate approach is required to solve such a system considering the sparsity of aggregated power information.

In this chapter, our main objective is to develop algorithms to estimate states from the compressed measurements, given that the system is non-linear. The underlying network topology is assumed to be known a priori and fixed over the time duration of state estimation. Two methods of voltage phasor estimation from the compressed power readings are developed. In the first method, referred to as the *indirect* method, all power values are reconstructed from compressed measurements and then fed as input to a Newton-Raphson algorithm to estimate voltage states.

In the second method, referred to as the *direct* method, we avoid the reconstruction procedure and the compressed measurements are directly used in a modified Newton-Raphson algorithm to estimate the states [80]. At every iteration of modified Newton-Raphson algorithm, the linear system involving Jacobian matrix and random projection matrix is solved in two steps. At the first step, sparsest solution of the mismatch vector is obtained considering only the projection matrix. The existence of sparsity for mismatch vector is also

discussed, which authenticates the sparse solution. The mismatch vector just obtained and the Jacobian matrix are then used to get the least square solution of the state-update vectors. Fig. 4.1 shows the contrast of the proposed methods with the conventional centralized state estimation.

Performance of these methods are obtained for different levels of compression based on the IEEE 34 node distribution test feeder [102] and a larger radial network of 100 nodes.

4.2 State Estimation from Compressed Measurements

According to Theorem 2.1, a random projection matrix $\Phi_{projection} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ is used to make individual compressive measurements of real and reactive powers for an N -node power distribution system:

$$\begin{aligned} \mathbf{h}_P &= \Phi_{projection} \begin{bmatrix} P_1 & \cdots & P_N \end{bmatrix}^T ; \\ \mathbf{h}_Q &= \Phi_{projection} \begin{bmatrix} Q_1 & \cdots & Q_N \end{bmatrix}^T \end{aligned} \quad (4.3)$$

Because of the inherent correlation structure, the real and reactive injected powers are approximately sparse on respective linear transformation basis Ψ_P and Ψ_Q . According to equations (3.4) and (4.3), the following relation among \mathbf{y} , \mathbf{h}_P and \mathbf{h}_Q are obtained:

$$\begin{aligned} \mathbf{h} &= \begin{bmatrix} \mathbf{h}_P^T & \mathbf{h}_Q^T \end{bmatrix}^T \\ \Rightarrow \mathbf{h} &= \Phi \mathbf{y} \\ \Rightarrow \mathbf{h} &= \Phi \Psi \mathbf{a} \end{aligned} \quad (4.4)$$

Where,

$$\Phi = \begin{bmatrix} \Phi_{projection} & \mathbf{O}_{M \times N} \\ \mathbf{O}_{M \times N} & \Phi_{projection} \end{bmatrix}$$

And,

$$\Psi = \begin{bmatrix} \Psi_P & \mathbf{O}_{M \times N} \\ \mathbf{O}_{M \times N} & \Psi_Q \end{bmatrix}$$

We incorporate these compressed measurements in voltage phasor estimation using two different methods:

4.2.1 Indirect Method

This method, consists of two steps. At first, all the injected power values are reconstructed from the compressed measurements. In the second step, Newton-Raphson method uses these reconstructed values to estimate voltage phasors. The overall method constitutes Algorithm 1, which returns the estimated states, when the ℓ_∞ norm of mismatch vector \mathbf{b} reaches the tolerance level ϵ .

Algorithm 1 Indirect Method

```

1: procedure RECONSTRUCT( $\Phi, \Psi, \mathbf{h}$ )
2:    $\mathbf{a}^* \leftarrow \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1$  subject to  $\mathbf{h} = \Phi \Psi \mathbf{z}$ 
3:    $\mathbf{y}^* \leftarrow \Phi \Psi \mathbf{a}^*$ 
4: end procedure
5: Initialize  $\mathbf{x}^{(0)}$  ▷ Flat Start
6: procedure NEWTON-RAPHSON( $\mathbf{x}^{(0)}, \mathbf{y}^*, \epsilon$ )
7:    $\mathbf{x} \leftarrow \mathbf{x}^{(0)}$  ▷ Take State Values
8:    $\mathbf{b} \leftarrow \mathbf{y}^* - \mathbf{F}(\mathbf{x})$  ▷ Calculate Initial Mismatch
9:   while  $\|\mathbf{b}\|_\infty > \epsilon$  do
10:     $\mathbf{J} \leftarrow \text{Jacobian}(\mathbf{x})$  ▷ Calculate Jacobian Matrix
11:    Solve:  $\mathbf{J} \Delta \mathbf{x} = \mathbf{b}$  for  $\Delta \mathbf{x}$ 
12:     $\mathbf{x} \leftarrow \mathbf{x} + \Delta \mathbf{x}$  ▷ Update  $\mathbf{x}$ 
13:     $\mathbf{b} \leftarrow \mathbf{y}^* - \mathbf{F}(\mathbf{x})$  ▷ Calculate Mismatch
14:  end while
15:  Return  $\mathbf{x}$ 
16: end procedure

```

4.2.2 Direct Method

In the *indirect* method, all the power values are required to be reconstructed from compressed power measurements prior to state estimation. The *direct* method uses compressed power measurements within the Newton-Raphson iteration itself, thus avoiding the unnecessary reconstruction of all power values. The associated procedure results in Algorithm 2 [80]. Unlike Algorithm 1, here we are required to solve an underdetermined system of nonlinear equations as evident from equations (3.4) and (4.4),

$$\mathbf{h} = \Phi \mathbf{F}(\mathbf{x}) \quad (4.5)$$

As a result, the mismatch vector and system of linear equations to be solved at the v th iteration would be modified as: Mismatch vector:

$$\begin{aligned} \mathbf{b}^{(v)} &= \mathbf{y} - \mathbf{F}(\mathbf{x}^{(v)}) \\ \Rightarrow \mathbf{c}^{(v)} &= \mathbf{h} - \Phi \mathbf{F}(\mathbf{x}^{(v)}) \end{aligned} \quad (4.6)$$

where,

$$\mathbf{c}^{(v)} = \Phi \mathbf{b}^{(v)}$$

System of linear equations:

$$\begin{aligned} \mathbf{J}^{(v)} \Delta \mathbf{x}^{(v)} &= \mathbf{b}^{(v)} \\ \Rightarrow \Phi \mathbf{J}^{(v)} \Delta \mathbf{x}^{(v)} &= \mathbf{c}^{(v)} \end{aligned} \quad (4.7)$$

Conventional least-square method will not give accurate solution of equation (4.7), since this system is the result of random mapping of a fully determined system. One possibility is to

search for the sparsest solution [38]. However, the recovery of a sparse solution is dependent upon the coherence of $\Phi\mathbf{J}^{(v)}$. Theorem 4.1 (proof can be found in [112]) specifies the strict upper bound on coherence to ensure at most one K -sparse solution [87].

Theorem 4.1. *For any matrix $\mathbf{D} \in \mathbb{R}^{M \times N}$, $\mu(\mathbf{D}) \in \left[\sqrt{\frac{N-M}{M(N-1)}}, 1 \right]$ and if*

$$\mu(\mathbf{D}) < \frac{1}{2K-1}$$

there exist at most one K -sparse signal \mathbf{w} for every measurement $\mathbf{h} \in \mathbb{R}^M$ such that $\mathbf{h} = \mathbf{D}\mathbf{w}$. Where, $\mu(\mathbf{D})$ is the coherence of \mathbf{D} and it is the maximum absolute off-diagonal entry of the normalized version of Gram matrix $\mathbf{D}^T\mathbf{D}$.

The Gram matrix for equation (4.7) is $\mathbf{J}^{(v)T}\Phi^T\Phi\mathbf{J}^{(v)}$. For the moment, we assume Φ be so chosen that $\Phi^T\Phi$ becomes an identity matrix and hence the coherence is dependent only on the Gram matrix $\mathbf{J}^{(v)T}\mathbf{J}^{(v)}$. As discussed earlier, for any two different nodes, the rate of change in injected power is affected by the electrical coupling in between them. So, two distribution nodes, directly connected and at sufficiently close geographical proximity can have the same rate of change in injected power with respect to voltage magnitudes as well as voltage angles. So, the off-diagonal entries of normalized version of $\mathbf{J}^{(v)T}\mathbf{J}^{(v)}$ can be unity or very close to unity. Fig. 4.2 shows such an instance for the distribution system model described in Section III. As a result, sparse and unique recovery of $\Delta\mathbf{x}^{(v)}$ cannot be guaranteed from equation (4.7). This lack of incoherence can be resolved if we consider the term $\mathbf{J}^{(v)}\Delta\mathbf{x}^{(v)}$ as a single variable \mathbf{w} and solve equation (4.7) in following two steps,

Solve for \mathbf{w} from:

$$\Phi\mathbf{w} = \mathbf{c}^{(v)} \quad (4.8)$$

Solve for $\Delta\mathbf{x}^{(v)}$ from:

$$\mathbf{J}^{(v)}\Delta\mathbf{x}^{(v)} = \mathbf{w} \quad (4.9)$$

Equation (4.9) represents a fully determined system and its solution is trivial. On the other

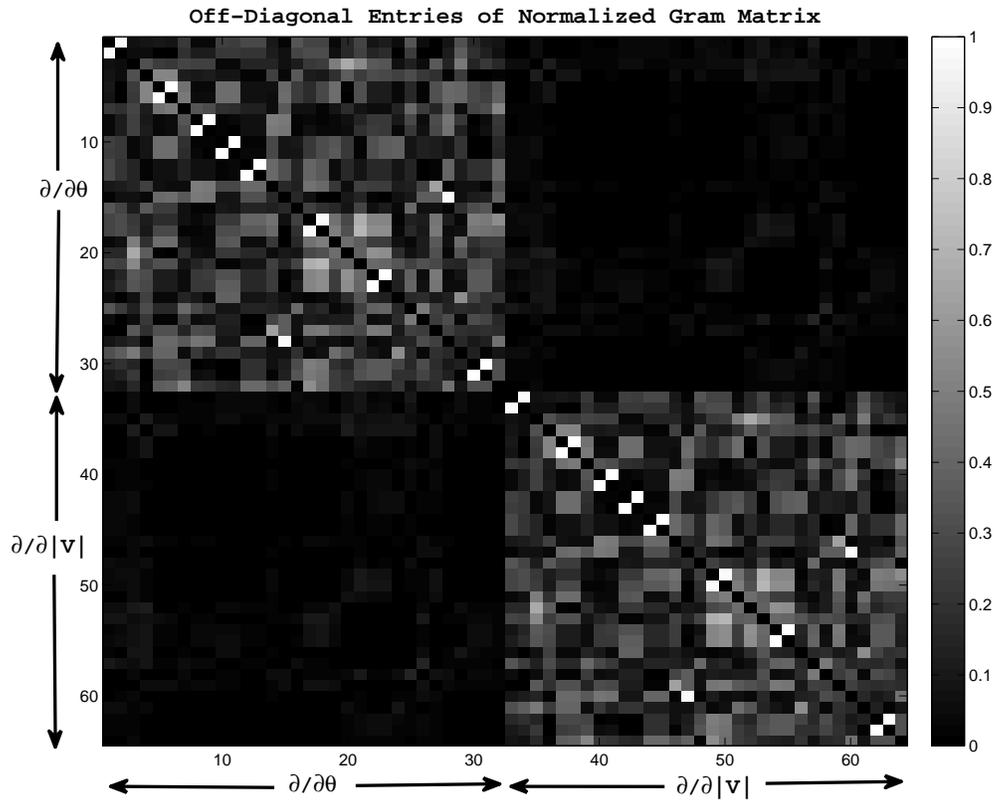


Figure 4.2: Off-diagonal Entries of Normalized $\mathbf{J}^{(0)T} \Phi^T \Phi \mathbf{J}^{(0)}$

hand, we obtain a unique solution of equation (4.8) if the intermediate variable \mathbf{w} is sparse in appropriate domain. Comparing equation (4.6) and (4.8) it is observed that, at the v th iteration, the sparsity of \mathbf{w} corresponds to the sparsity of original mismatch vector $\mathbf{b}^{(v)}$. Thus, unique sparse solution for \mathbf{w} is guaranteed if the mismatch vector entries of same kind (i.e., real or reactive) are correlated. We investigate this by obtaining spatial correlation for the 0^{th} mismatch vector, which is shown in Fig. 4.3. As expected, the real power mismatch is uncorrelated to reactive power mismatch. However, comparing Fig. 4.3 and Fig. 3.3, it is observed that the auto-correlation pattern of both type of mismatch is almost same to the correlation matrix defined in Section III. As a consequence, at every iteration, Theorem 2.1 can be applied to solve equation (4.8) and this is incorporated in steps 6 – 9 of Algorithm 2.

Algorithm 2 Direct Method

```

1: Initialize  $\mathbf{x}^{(0)}$  ▷ Flat Start
2: procedure COMPRESSED-NEWTON-RAPHSON( $\mathbf{x}^{(0)}, \mathbf{h}, \Phi, \Psi, \epsilon$ )
3:    $\mathbf{x} \leftarrow \mathbf{x}^{(0)}$  ▷ Take State Values
4:    $\mathbf{c} \leftarrow \mathbf{h} - \Phi\mathbf{F}(\mathbf{x})$  ▷ Calculate Initial Mismatch
5:   while  $\|\mathbf{c}\|_{\infty} > \epsilon$  do
6:      $\mathbf{a}^* \leftarrow \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1$  subject to  $\mathbf{c} = \Phi\Psi\mathbf{z}$ 
7:      $\mathbf{w}^* \leftarrow \Phi\Psi\mathbf{a}^*$ 
8:      $\mathbf{J} \leftarrow \text{Jacobian}(\mathbf{x})$  ▷ Calculate Jacobian Matrix
9:     Solve:  $\mathbf{J}\Delta\mathbf{x} = \mathbf{w}^*$  for  $\Delta\mathbf{x}$ 
10:     $\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x}$  ▷ Update  $\mathbf{x}$ 
11:     $\mathbf{c} \leftarrow \mathbf{h} - \Phi\mathbf{F}(\mathbf{x})$  ▷ Calculate Mismatch
12:  end while
13:  Return  $\mathbf{x}$ 
14: end procedure

```

4.2.3 Computational Complexity

The complexity order for the two methods can be obtained by observing Algorithms 1 and 2. The *indirect* method involves an ℓ_1 minimization problem outside the iterative loop, whereas, the *direct* method has ℓ_1 minimization problem within the iterative loop. On the

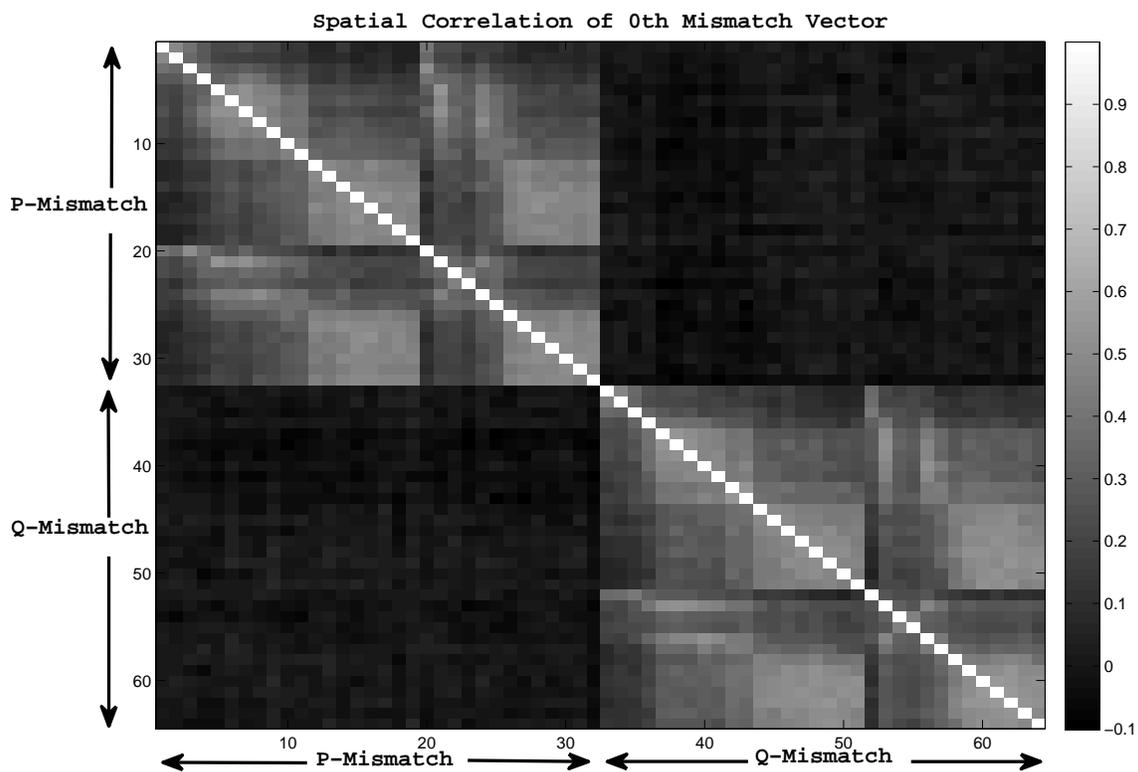


Figure 4.3: *Spatial Correlation of $\mathbf{b}^{(0)}$*

other hand, both methods have a system of equations with square jacobian matrix \mathbf{J} within the iterative loop.

For a typical power distribution system, the jacobian matrix \mathbf{J} is usually sparse, since relative changes in injected power at a particular distribution node is only affected by the directly coupled neighboring nodes [98]. As a result, the computational complexity of solving a fully determined linear system as in step 11 of Algorithm 1 and step 9 of Algorithm 2 can be reduced to $\mathcal{O}(N)$ using various efficient factorization techniques [48]. On the other hand, since ℓ_1 norm is convex, computational complexity of implementing the ℓ_1 minimization as a linear programming problem is $\mathcal{O}(N \log_2(N))$ [87].

Based on the above discussion, we conclude that both the *indirect* and *direct* method have complexity order $\mathcal{O}(N \log_2(N))$. In the next section, we apply Algorithms 1 and 2 to the system model described in Section III and discuss their relative performance.

4.3 Results

For voltage phasor estimation, we apply *indirect* and *direct* method over the compressed power measurements. For both algorithms, the tolerance level ϵ is kept at 10^{-9} . That is, the algorithm converges when the maximum absolute value of the elements in the mismatch vector is $\leq 10^{-9}$. Fig.4.4 shows the estimated voltages for all the nodes of the distribution grid. The “o” sign denotes the actual voltage phasors. With $CMR = 50\%$, the “+” and “ Δ ” signs indicate voltage phasors estimated from *indirect* and *direct* method, respectively. It is observed that, both *indirect* and *indirect* method can estimate the states almost accurately from only 50% compressed measurements [80]. Also, the deviation of the actual state values from the estimated ones is almost constant for bus 12 to bus 19 and for bus 26 to bus 32. This is a consequence of the ℓ_1 -minimization inherent in the proposed approaches. The solution of ℓ_1 -minimization is biased, because it tends to under-estimate the value of true coefficients by having no prior knowledge of where the nonzero elements of the sparse vector are located.

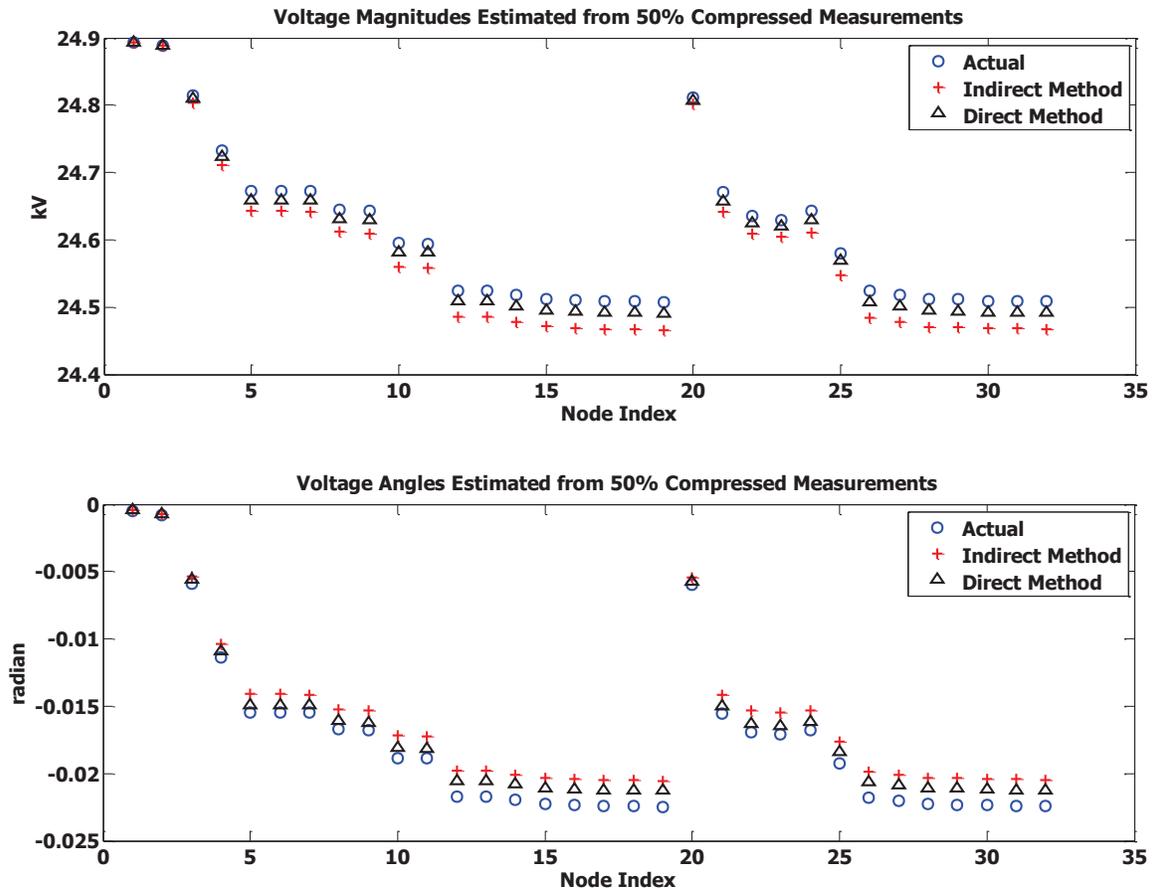


Figure 4.4: Voltage Phasors Estimated from Direct and Indirect Method

Column 1 Heading	Direct Method	Indirect Method
34 Node Test Feeder	0.3936 sec.	0.1363 sec.
100 Node Test Feeder	6.3494 sec.	1.6369 sec.

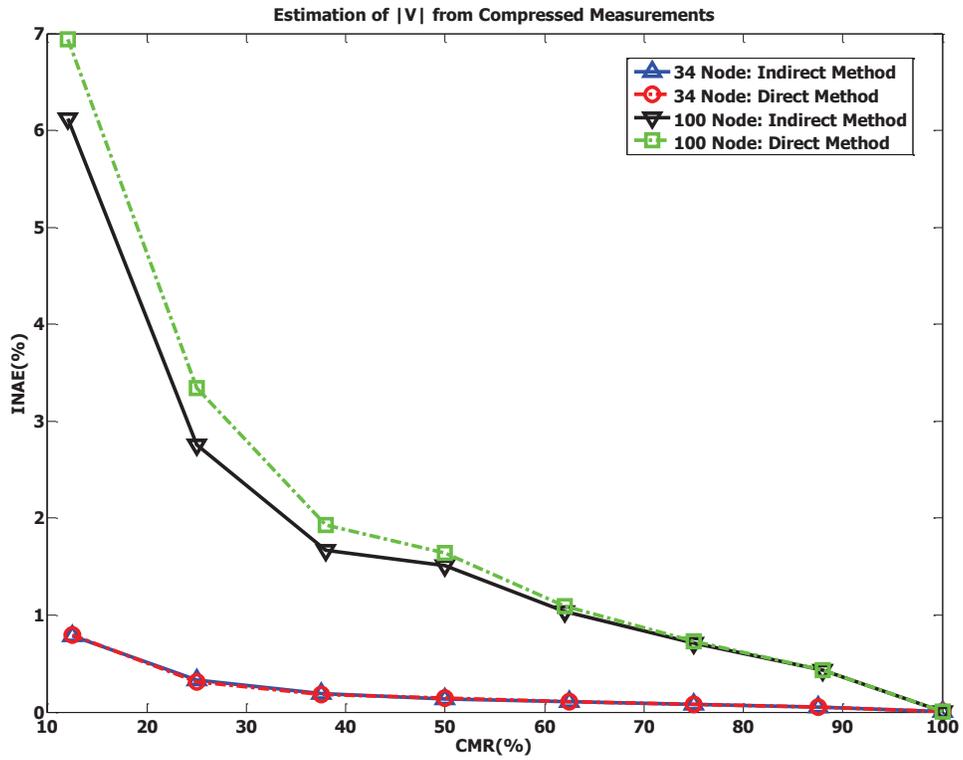
Table 4.1: CPU time for the proposed estimators

It is established that the bias incurred due to ℓ_1 -minimization is within a logarithmic factor of the ideal mean squared error (Theorem 1.1 of [113]). To illustrate the scalability of our analysis, the proposed method is applied to a larger 100 node radial distribution test feeder. The test feeder is a modified version of IEEE 123 node distribution test feeder with the extreme case of DGs being installed at every node. The test systems are simulated on intel[®] CORE[™] i7 PC, with clock speed 3.4 GHz and 16.0 GB RAM. MATLAB[®] for Windows is used as the simulation platform. Table 4.1 shows the CPU time required for state estimation for the two test systems using the proposed estimation algorithms. For both cases, CMR is maintained at 50%. It should be noted that, both methods converge to solution within 4 and 6 iterations for the 34 node and 100 node systems, respectively.

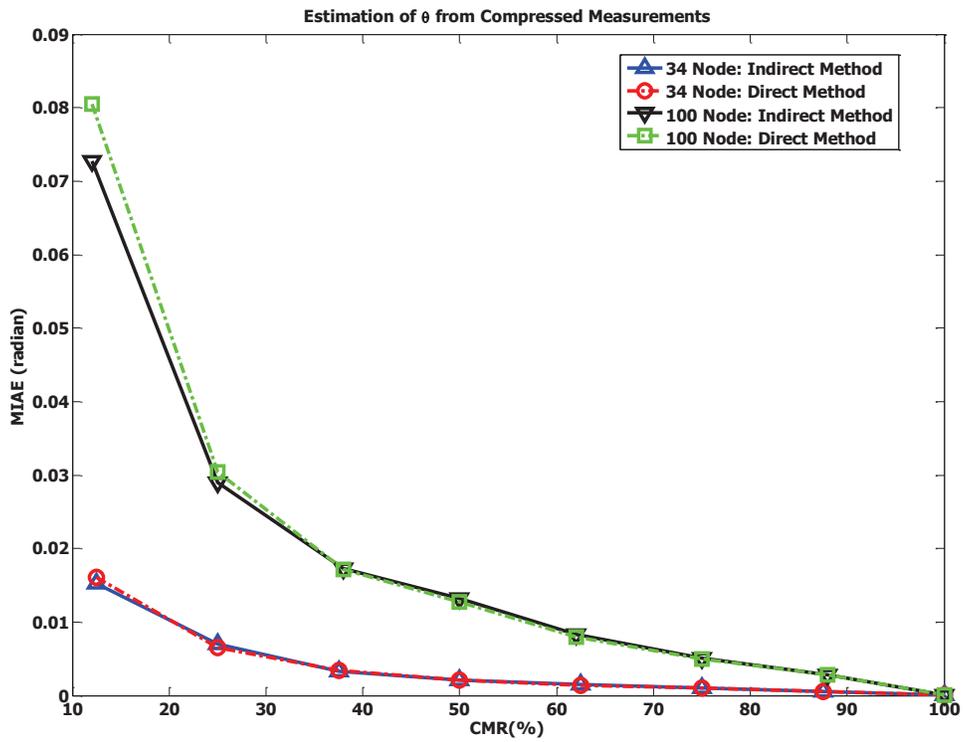
The two methods of state estimation are also investigated at different CMRs and corresponding INAEs are obtained for the voltage magnitudes. The INAEs for voltage magnitudes are calculated using equation (3.8), where x_j represents the voltage magnitude of j th node. To compare performance of the two methods in estimating voltage angles, *Mean Integrated Absolute Error (MIAE)* is defined as follows,

$$MIAE = \frac{\sum_{j=1}^N |\theta_j - \theta_j^*|}{N} \quad (4.10)$$

The overall performance for both test systems are shown in Fig. 4.5. It is observed that, both the *indirect* and *direct* methods give almost the same performance at all CMRs. At CMRs $\geq 50\%$, the error in estimating voltage magnitudes is $< 2\%$ for both test systems and corresponding error in voltage angles is < 0.015 radian [80]. Intuitively, it is possible to estimate the states of a smart distribution system without even recovering power information of



(a)



(b)

Figure 4.5: Variation of INAE and MIAE for Indirect and Direct Methods

all nodes and with no increase in computational complexity. These methods, therefore, hold promising application in distribution system state estimation with a significant reduction in storage / memory requirement.

4.4 Conclusions

In this chapter, two novel methods: *indirect* and *direct* method of state estimation from compressed power measurements are developed, based on the well-known Newton-Raphson algorithm. It is observed that, both methods give similar performance in distribution system state estimation at all level of compression and maintain same complexity order. Even with only 50% compressed measurements both methods provide accurate estimation of voltage states. Thus, the proposed methods of state estimation from compressed measurements can play a significant role in the perspective of centralized information processing for any CPS including the smart distribution grid. In the upcoming chapters, we investigate the decentralized approach, which can also reduce the information burden while distributing the computational task among the communicable sensors. As pointed in Section 1.2, this will help making the CPS more robust to the faulty communication network.

Chapter 5

Agent based State Estimation in Smart Distribution Grid

A novel agent based static state estimation strategy for a specific class of physical systems is proposed. Physical systems, (e.g. smart distribution grid) which are well modeled using decentralized measurements and distributed state-space formulations are considered. In such systems, sensor nodes acting as agents estimate only a subset of states, instead of evaluating local estimates of global states. In general, for each agent, the measurement model reduces to an underdetermined nonlinear system and in many cases, the state elements associated with an agent may overlap with neighboring agents. A classic example of such a physical system is a radial power distribution grid. We propose a state estimation strategy, which effectively integrates the principles of local consensus and least squares solution and illustrate its potency using the power distribution grid. We also present rigorous analysis of convergence of the proposed approach to motivate its application to other multi-agent systems.

5.1 Introduction

The future electric power distribution network is envisioned to be *smart* and *intelligent* using modern information and communication technology. The future smart distribution grid is characterized by the ability of end-user to participate in system monitoring, operation and control through two-way communication with electrically coupled neighbors and exchanging relevant encrypted information [22]. In a sense, each end-user will act as an *agent* by being proactive to environments and neighbors. It will ensure affordable, secured and reliable electric power delivery. Decentralized estimation of states is one of the important aspects of an agent based system, that offers benefits relative to centralized and hierarchical data fusion architecture in terms of (1) real-time implementation, (2) configuration flexibility, and (3) communication bottlenecks [39]. In a typical decentralized architecture, each sensor node observes only a distinct portion of underlying physical system and makes a local estimate of overall system states by exchanging information with neighboring nodes. However, full connectivity of sensor nodes is required to make the local estimate equivalent to that obtained using a centralized systems. With the increase in sensor nodes, the number of communication links, computational memory and hardware requirements increase excessively in a fully connected decentralized multisensor system. This problem can be resolved by introducing sensor specific distributed state-space model (whenever applicable), so that each sensor estimates only a relevant subset of the overall physical system. Thus, the communication and computational redundancy inherent to decentralized observation model can be reduced. In this way, the sensors are acting as agents and the underlying system can be referred to as a multi-agent system. In this chapter, we formulate an agent based scenario for estimating the voltage states of a typical radial distribution network.

As discussed in Section 2.2.1, static estimation of states from nonlinear observations is typically performed through Gauss-Newton iteration constituting the weighted least square (WLS) method. In the next subsection, we briefly describe the application of WLS and its variations for decentralized static state estimation in power system.

5.1.1 Related Work

The WLS method is used in various literature for global state estimation from decentralized observation model, specially for interconnected electric power transmission systems. The decentralization is generally obtained by decomposing an interconnected transmission system into a certain number of *nonoverlapping* subarea on a geographical basis [41]. Let us assume that the physical system is decomposed into N subareas. And the global state vector \mathbf{x} is decomposed into N nonoverlapped local state vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$. The set of measurement / observations made in the i^{th} subarea constitutes the following measurement model,

$$\mathbf{y}_i = \mathbf{h}_i(\mathbf{x}_i) + \mathbf{w}_i; \quad i = 1, 2, \dots, N \quad (5.1)$$

where, $\mathbf{y}_i \in \mathbb{R}^{M_i}$ is the measurement vector for i^{th} subarea, $\mathbf{x}_i \in \mathbb{R}^{L_i}$ is the state vector having only local state elements of i^{th} subarea and $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_i)$. The global state vector is related to inter-area boundary measurements as,

$$\mathbf{y}_{boundary} = \mathbf{h}_{boundary}(\mathbf{x}) + \mathbf{w}_{boundary} \quad (5.2)$$

In each of the subsystem, respective local estimates are obtained using equation (2.8). The local estimates are then centrally coordinated to obtain overall system states subject to the constraints imposed by equation (5.2). A comprehensive survey of such multiarea scheme of state estimation is given in [42]. A hierarchical multilevel structure of state estimation based on factorized WLS method is described in [43]. Here, the authors define the distribution substations as the lowest level of the multilevel structure. In [44], authors use decentralized measurement model with the assumption that the global state may not be observable in some of the subareas. In each subarea, the global states are synchronously updated using the combined approach of global consensus and innovation. Mathematically, at $(k + 1)^{th}$ iteration, state update at subarea i is given as,

$$\begin{aligned} \mathbf{x}_i^{(k+1)} &= \mathbf{x}_i^{(k)} + \alpha(k) \mathbf{H}_i(k)^\top \mathbf{C}_i^{-1} [\mathbf{y}_i - \mathbf{H}_i(k) \mathbf{x}_i^{(k)}] \\ &\quad - \beta(k) \sum_{j \in \mathcal{S}_i} [\mathbf{x}_i^{(k)} - \mathbf{x}_j^{(k)}] \end{aligned} \quad (5.3)$$

where, \mathbf{x}_i represents the global state maintained and updated in each iteration by subarea i and \mathcal{S}_i is the corresponding set of neighbors. The local jacobian matrix $\mathbf{H}_i(k)$ in equation (5.3) is updated after a fixed number of iterations to reduce computational requirements. It should be noted that, the convergence to global consensus is not reported in [44], when $\mathbf{H}_i(k)$ violates the condition of full column rank in each of the subarea i . Unlike [41], partial overlapping of local state vectors is considered in [45] and [46]. It is assumed that any subarea may have some state elements, which are shared with its neighbors. The sharing is due to the specific placement of measurement devices over the tielines of an interconnected system. In this scenario, the multi-area state estimation (MASE) is designed as a distributed optimization problem in each Gauss-Newton iteration [45]. At $(k + 1)^{th}$ iteration,

$$\min \sum_{i=1}^N \frac{1}{2} \|\mathbf{y}_i - \mathbf{H}_i(k) \mathbf{x}_i^{(k+1)}\|^2 \quad (5.4)$$

subject to $\mathbf{x}_i^{(k+1)}[j] = \mathbf{x}_{ij}^{(k+1)}, \forall j \in \mathcal{S}_i; i = 1, 2, \dots, N$. Here, \mathbf{x}_i represents the state elements local to subarea i . $\mathbf{x}_i[j]$ represents a vector of state elements for the i^{th} subarea, which are shared with its neighbor j . The auxiliary variables \mathbf{x}_{ij} are associated with the tieline measurements. Thus, at every iteration, equation (5.4) represents a *partially* global consensus problem, that can be solved using alternating direction method of multipliers (ADMM) (See chapter 7 of [47] for details). In [46], authors implement R-SDP with the help of ADMM to solve the problem of MASE. However, two important assumptions are made in the works of [45] and [46]: (1) in each subarea, there exists at least one state element, unshared with

any of its neighbors, and (2) if a state element is shared, it is only between two neighboring subareas.

In summary, the aforementioned prior efforts are based on either over-determined or fully determined measurement models. As a consequence, efficient factorization schemes [48] can be applied in equations (2.8) and (5.4). Furthermore, a scenario of more than two subareas sharing the same state element as well as the case when all local state elements are shared among neighbors is still an open challenge in distributed state estimation.

In this chapter, we propose an agent based static estimation method, specifically for a radial network of electric power distribution. Each node in the network is considered an agent that has access to a distinct set of measurements. A local nonlinear measurement model is developed for every agent, which is based on the physical coupling with its neighbors. Using this measurement model each agent attempts to estimate local states, when (1) more than two agents share a particular state element and (2) each state element is shared at least between two agents [81, 114].

5.2 System Model

The proposed methodology is based on mathematical modeling of the underlying physical system in multi-agent framework. We consider a physical system that is monitored by a set of distributed sensors. These sensors acquire a distinct set of measurements and can communicate with other neighbors. Thus, the physical system can be visualized as a multi-agent network, where each sensor acts like an *agent*. Additionally, the neighbors of an agent are defined based on existing physical connections rather than geographical proximity.

Let, \mathcal{S}_i denotes the set of neighbors of agent i . The measurement model for agent i is,

$$\mathbf{y}_i = \mathbf{h}_i(\mathbf{x}_i) + \mathbf{w}_i; i = 1, 2, 3, \dots, N \quad (5.5)$$

where, $\mathbf{y}_i \in \mathbb{R}^{M_i}$ is the set of measurements taken by agent i and $\mathbf{x}_i \in \mathbb{R}^{L_i}$ represents the

associated local state vector. Similar to equation (5.1), \mathbf{w}_i is the Gaussian noise vector for i^{th} agent with $\mathbf{0}$ mean and covariance \mathbf{C}_i . Since, each agent is now an estimator, it is likely that $M_i \leq L_i, \forall i$. Thus, if Gauss-Newton iteration is used, the underlying Jacobian matrix will be full row rank, representing an underdetermined system of equations. Furthermore, each sensor / measurement unit is now acting itself as an agent / local estimator. As a consequence, each state element associated with an agent can be shared with its neighbors leaving the set of nonoverlapped state elements empty.

As an illustrative example, let us consider an N -node radial power distribution network as shown in Fig. 3.1. Node 0 represents the substation grid. According to Kirchhoff's current law, the i^{th} node injection current is given by,

$$I_i = \begin{bmatrix} -Y_{i-1,i} & Y_{i-1,i} + Y_{i,i+1} & -Y_{i,i+1} \end{bmatrix} \begin{bmatrix} V_{i-1} \\ V_i \\ V_{i+1} \end{bmatrix} \quad (5.6)$$

where, $Y_{i-1,i} = \frac{1}{Z_{eq_i}}$ is the single-phase equivalent admittance between node $i-1$ and node i [97]. V_i (or $|V_i|\angle\theta_i$) is the i^{th} node voltage phasor. Since, the network is limited to N nodes, for consistency it is assumed that $Y_{N,N+1} = 0$ and $|V_{N+1}|\angle\theta_{N+1} = 0$ as well.

The voltage phasors at different nodes of the network constitute the state of the overall system. For the purpose of analyzing the network, Node 0 is taken as the reference bus with $|V_0|\angle\theta_0 = 1p.u.\angle 0$.

As shown in Fig. 3.1, S_{G_i} and S_{L_i} represents the respective apparent power generation and load at i^{th} node. The corresponding real injected power, $P_i = real(S_{G_i} - S_{L_i})$ and reactive injected power, $Q_i = imag(S_{G_i} - S_{L_i})$. The injected powers satisfy the following power flow equations throughout the network:

$$P_i = |V_i|^2(G_{i-1,i} + G_{i,i+1}) - |V_i| \sum_{j=i-1,i+1} |V_j| (G_{i,j} \cos \theta_{i,j} + B_{i,j} \sin \theta_{i,j});$$

$$Q_i = -|V_i|^2(B_{i-1,i} + B_{i,i+1}) - |V_i| \sum_{j=i-1,i+1} |V_j| (G_{i,j} \sin \theta_{i,j} - B_{i,j} \cos \theta_{i,j}) \quad (5.7)$$

where, $G_{i,j} = \text{real}(Y_{i,j})$; $B_{i,j} = \text{imag}(Y_{i,j})$ and $\theta_{i,j} = \theta_i - \theta_j$. From these equations, we have the following noiseless measurement model for each agent

$$\mathbf{y}_i = \begin{bmatrix} P_i \\ Q_i \end{bmatrix} = \mathbf{h}_i(\mathbf{x}_i) \quad (5.8)$$

where, $\mathbf{h}_i : \mathbb{R}^6 \rightarrow \mathbb{R}^2$; $1 \leq i \leq N$, and the associated state vector of agent i ,

$$\mathbf{x}_i = \left[\theta_{i-1}, |V_{i-1}|, \theta_i, |V_i|, \theta_{i+1}, |V_{i+1}| \right]^\top ; \forall i \quad (5.9)$$

Thus, equation (5.8) represents an underdetermined nonlinear system for each agent. The local state vectors as defined in equation (5.9) can be accumulated in a single matrix \mathbf{X} such that agent i 's state vector is in the i^{th} column of \mathbf{X} . The matrix takes the following form:

$$\begin{bmatrix} 0, & \theta_1, & \dots, & \theta_{i-2}, & \theta_{i-1}, & \theta_i, & \dots, & \theta_{N-1} \\ 1, & |V_1|, & \dots, & |V_{i-2}|, & |V_{i-1}|, & |V_i|, & \dots, & |V_{N-1}| \\ \theta_1, & \theta_2, & \dots, & \theta_{i-1}, & \theta_i, & \theta_{i+1}, & \dots, & \theta_N \\ |V_1|, & |V_2|, & \dots, & |V_{i-1}|, & |V_i|, & |V_{i+1}|, & \dots, & |V_N| \\ \theta_2, & \theta_3, & \dots, & \theta_i, & \theta_{i+1}, & \theta_{i+2}, & \dots, & 0 \\ |V_2|, & |V_3|, & \dots, & |V_i|, & |V_{i+1}|, & |V_{i+2}|, & \dots, & 0 \end{bmatrix} \quad (5.10)$$

It can be observed from equation (5.10) that, the state vector for each agent consists not only of its own voltage and angle states but also includes four shared elements from its two neighbors. Furthermore, for any agent, no state element is left unshared as illustrated in the Venn diagram (Fig. 5.1). This scenario is a complete violation of the assumptions used in [45], and [46].

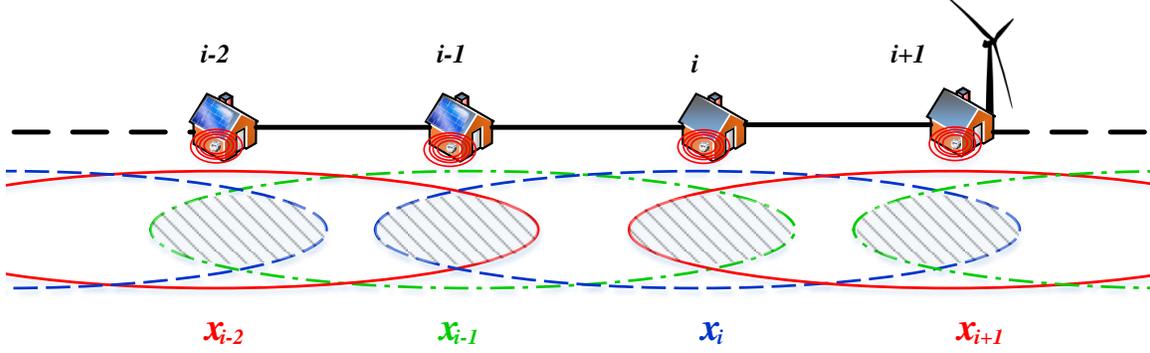


Figure 5.1: Venn Diagram of State Elements for a Radial Power Distribution Network.

5.3 Proposed Method

We wish to develop an agent-wise local consensus procedure, when (1) each agent's observation space is nonlinear and underdetermined, and (2) each state element of the system is shared at least between two agents. For any multi-agent system, the effect of neighbors can be incorporated in local consensus with the help of *selection* matrices [114, 81]. These matrices project the neighboring agents' state vector with correct sign and orientation onto the local state vector to accomplish distributed local consensus.

As shown in equation (5.10), agent $(i - 1)$ and $(i + 1)$ are in the set \mathcal{S}_i , located at the trailing and leading edge of the i^{th} agent, respectively. To compactly describe the local consensus, three selection matrices are defined in accordance with equation (5.10),

$$\mathbf{P}_{lead} = \begin{bmatrix} -\mathbf{e}_3, -\mathbf{e}_4, -\mathbf{e}_5, -\mathbf{e}_6, \mathbf{0}, \mathbf{0} \end{bmatrix}^\top; \mathbf{P}_{loc} = \begin{bmatrix} \mathbf{e}_1, \mathbf{e}_2, 2\mathbf{e}_3, 2\mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6 \end{bmatrix}^\top; \text{ and}$$

$$\mathbf{P}_{trail} = \begin{bmatrix} \mathbf{0}, \mathbf{0}, -\mathbf{e}_1, -\mathbf{e}_2, -\mathbf{e}_3, -\mathbf{e}_4 \end{bmatrix}^\top \quad (5.11)$$

where, \mathbf{e}_n is a column vector with 1 as the n^{th} entry and zeros elsewhere. The dimension and entries of these selection matrices depend upon the dimension of the state vectors as well as the orientation of corresponding state elements. Therefore, these matrices are expected to be fixed for a particular multi-agent system. Furthermore, \mathbf{P}_{loc} is always symmetric positive definite, and in this particular case, $\mathbf{P}_{lead} = \mathbf{P}_{trail}^\top$. Using the selection matrices

and measurement model, we can write the state updates for each agent in the following way [114, 81],

$$\begin{aligned} \mathbf{x}_i^{(k+1)} = & \mathbf{x}_i^{(k)} + \alpha_i(k)\mathbf{H}_i(k)^\dagger \mathbf{y}_i - \alpha_i(k)\mathbf{G}_i(k)\mathbf{x}_i^{(k)} \\ & - \beta(k)\left(\mathbf{P}_{lead}\mathbf{x}_{i-1}^{(k)} + \mathbf{P}_{loc}\mathbf{x}_i^{(k)} + \mathbf{P}_{trail}\mathbf{x}_{i+1}^{(k)}\right) \end{aligned} \quad (5.12)$$

where, $\alpha_i(k), \beta(k) > 0$ are the time varying coefficients and $\mathbf{G}_i(k) = \mathbf{H}_i(k)^\dagger \mathbf{H}_i(k)$. All quantities having subscript “ i ” correspond to the i^{th} agent and k indicates the iteration index. $\mathbf{H}_i(k)^\dagger$ represents the pseudo-inverse for a full row rank local Jacobian matrix. Mathematically, $\mathbf{H}_i(k)^\dagger = \mathbf{H}_i(k)^\top \left(\mathbf{H}_i(k)\mathbf{H}_i(k)^\top\right)^{-1}$. Hence, by definition, $\mathbf{G}_i(k)$ is symmetric positive definite.

It can be observed from equation (5.12) that, the system observability depends upon the measurements acquired by each agent. If the agents are allowed to share measurements among neighbors, each agent is expected to have accurate estimation with faster convergence. However, issues like security and privacy has to be considered in defining the type of information the neighboring agents can share. Furthermore, the choice of values for $\alpha_i(k), \beta(k)$ is another important factor that determines the rate of convergence of the proposed method. In the following subsection we attempt to explore the theoretical analysis for convergence rate of equation (5.12).

5.3.1 Convergence Analysis

Before going to detailed convergence analysis, we briefly review the concept of *graph*. A networked physical system can be represented as a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of vertices or nodes, whose interconnections are defined as a set of undirected edges \mathcal{E} . An alternate representation of \mathcal{G} is a symmetric *adjacency* matrix \mathbf{A} . The $\{i, j\}^{th}$ entry of this

matrix is,

$$[\mathbf{A}]_{i,j} = [\mathbf{A}]_{j,i} = \begin{cases} 1 & \{i,j\} \in \mathcal{E} \\ 0 & \{i,j\} \notin \mathcal{E} ; \forall i,j \in \mathcal{V} \\ 0 & i = j \end{cases} \quad (5.13)$$

The above definition enables the decomposition, $\mathbf{A} = \mathbf{B} + \mathbf{F}$ and $\mathbf{B} = \mathbf{F}^\top$ where, \mathbf{B} and \mathbf{F} are strictly lower and upper triangular matrices, respectively.

In order to represent the iteration in equation (5.12) in compact form, we stack the state vectors of all agents in a single long state vector,

$$\mathbf{x} = \left[\mathbf{x}_1^\top, \mathbf{x}_2^\top, \dots, \mathbf{x}_N^\top \right]^\top \quad (5.14)$$

For simplicity, let's assume that the state vector dimension for each agent is L and that for measurement set is M . As a consequence, the accumulated state updates for all agents can be represented by the following recursive relation,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbb{H}(k)\mathbf{y} - \left(\mathbb{G}(k) + \beta(k)\mathbb{D} \right) \mathbf{x}^{(k)} \quad (5.15)$$

where, $\mathbf{y} = \left[\mathbf{y}_1^\top, \mathbf{y}_2^\top, \dots, \mathbf{y}_N^\top \right]^\top$; $\mathbb{D} = (\mathbf{B} \otimes \mathbf{P}_{lead}) + (\mathbf{F} \otimes \mathbf{P}_{trail}) + (\mathbf{I}_N \otimes \mathbf{P}_{loc})$;

$\mathbb{H}(k) = \text{Blockdiag} \left[\alpha_1(k)\mathbf{H}_1(k)^\dagger, \alpha_2(k)\mathbf{H}_2(k)^\dagger, \dots, \alpha_N(k)\mathbf{H}_N(k)^\dagger \right]$; and

$\mathbb{G}(k) = \text{Blockdiag} \left[\alpha_1(k)\mathbf{G}_1(k), \alpha_2(k)\mathbf{G}_2(k), \dots, \alpha_N(k)\mathbf{G}_N(k) \right]$. \mathbf{I}_N is the $N \times N$ identity matrix. The symbol \otimes represents Kronecker product.

Lemma 5.1. *For radial network, \mathbb{D} is symmetric positive semidefinite.*

Proof. For the state matrix in equation (5.10), $\mathbf{P}_{lead} = \mathbf{P}_{trail}^\top$ and $\mathbf{B} = \mathbf{F}^\top$. Therefore, $(\mathbf{B} \otimes \mathbf{P}_{lead}) = (\mathbf{F}^\top \otimes \mathbf{P}_{trail}^\top)$. Using the property of Kronecker product, $(\mathbf{B} \otimes \mathbf{P}_{lead}) = (\mathbf{F} \otimes \mathbf{P}_{trail})^\top$ and, $(\mathbf{I}_N \otimes \mathbf{P}_{loc})^\top = (\mathbf{I}_N \otimes \mathbf{P}_{loc})$. Hence, $\mathbb{D} = (\mathbf{F} \otimes \mathbf{P}_{trail})^\top + (\mathbf{F} \otimes \mathbf{P}_{trail}) + (\mathbf{I}_N \otimes \mathbf{P}_{loc})$ or, $\mathbb{D} = \mathbf{A} + (\mathbf{I}_N \otimes \mathbf{P}_{loc})$ where, $\mathbf{A} = (\mathbf{F} \otimes \mathbf{P}_{trail})^\top + (\mathbf{F} \otimes \mathbf{P}_{trail})$. Furthermore, $(\mathbf{I}_N \otimes \mathbf{P}_{loc})$ is a

diagonal matrix with non-zero elements, making it symmetric positive definite. However, the strict upper triangularity of \mathbf{F} makes the diagonal elements of \mathbf{A} zero. Thus, \mathbf{A} is symmetric indefinite matrix. As a combined effect, \mathbf{D} is symmetric positive semidefinite. \square

From equation (5.15) it is observed that, convergence will be achieved if the following two limiting conditions are fulfilled [81, 114],

- Limiting condition of consensus

$$\lim_{k \rightarrow \infty} \beta(k) \mathbf{D} \mathbf{x}^{(k)} = \mathbf{0} \quad (5.16)$$

- Limiting condition of zero mismatch

$$\lim_{k \rightarrow \infty} \mathbf{H}(k) \mathbf{y} = \lim_{k \rightarrow \infty} \mathbf{G}(k) \mathbf{x}^{(k)} \quad (5.17)$$

Given the initial estimate $\mathbf{x}^{(0)}$, the $(k + 1)^{th}$ update can also be represented in the following way,

$$\mathbf{x}^{(k+1)} = \left\{ \mathbf{H}(k) + \sum_{n=0}^{k-1} \left(\prod_{m=k}^{n+1} \mathbf{P}(m) \right) \mathbf{H}(n) \right\} \mathbf{y} + \prod_{m=k}^0 \mathbf{P}(m) \mathbf{x}^{(0)} \quad (5.18)$$

where, $\mathbf{P}(m) = \mathbf{I}_{NL} - \mathbf{G}(m) - \beta(m) \mathbf{D}$.

From equations (5.17) and (5.18), we see that if,

$$\lim_{k \rightarrow \infty} \prod_{m=k}^0 \mathbf{P}(m) = \mathbf{0} \quad (5.19)$$

then,

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k+1)} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \lim_{k \rightarrow \infty} \mathbf{G}(k) \mathbf{x}^{(k)} \quad (5.20)$$

Therefore, if λ_{max} denotes the maximum absolute eigenvalue, equation (5.19) will be satisfied if and only if, $\lambda_{max}(\mathbf{P}(k)) < 1$ [115]. Using Lemma 5.1 and properties of positive definiteness, it implies that, $0 < \lambda_{max}(\mathbf{G}(k) + \beta(k)\mathbf{D}) < 1$. Applying equal weights to both matrices, the inequality breaks down as,

$$\lambda_{max}(\mathbf{G}(k)) < 1/2; \lambda_{max}(\beta(k)\mathbf{D}) < 1/2$$

Equivalently, the time varying coefficients are bounded according to following inequalities ,

$$\alpha_i(k) < \frac{1}{2\lambda_{max}(\mathbf{G}_i(k))}; \beta(k) < \frac{1}{2\lambda_{max}(\mathbf{D})}, \forall i, \forall k \quad (5.21)$$

The time varying coefficients are also expected to satisfy the limiting conditions of equations (5.16) and (5.20). Based on these observations, we infer the following expressions of the coefficients [81, 114],

$$\alpha_i(k) = \frac{1}{a\lambda_{max}(\mathbf{G}_i(k))(1+k)^{\tau_1}}; a > 2, \tau_1 > 0 \quad (5.22)$$

$$\beta(k) = \frac{1}{b\lambda_{max}(\mathbf{D})(1+k)^{\tau_2}}; b > 2, \tau_2 > 0 \quad (5.23)$$

5.4 Simulation and Results

The proposed method is simulated using the following model specifications:

- Number of nodes excluding substation grid, $N = 10$
- Base kV = 24.9; Base kVA = 100
- Inter node distance starting from grid:
 $\{2580, 1730, 32230, 37500, 29730, 10, 310, 10210, 840, 20440\}$ ft.
- Inter node series impedance per mile:

- From grid to 5th node: $1.3368 + j1.3343$
- From 5th to 10th node: $1.9300 + j1.4115$; $j = \sqrt{-1}$
- Initial estimate for each agent: Voltage magnitudes are initiated as uniform random values between 0.95 *p.u.* and 1.05 *p.u.*. Similarly, voltage angles are initiated between $-\pi$ *radian* and π *radian*.
- For radial structure and selection matrices of equation (5.11), $\lambda_{max}(\mathbf{D}) = 3$.
- Empirical values for time varying coefficients
 - $\tau_1 = 2, \tau_2 = 1$
 - Effect of parameter a : $a = \{1, 4, 16, 64\}, b = 3$.
 - Effect of parameter b : $a = 4, b = \{1, 2, 3\}$
- Maximum number of iterations = 100

The proposed method is applied to obtain the agent based static estimation of voltage states. For comparison, centralized estimation of states are also obtained. Convergence analysis is performed by calculating the following two metrics at each iterations,

$$INAE(k) = \frac{\sum_{i=1}^N ||V_i^{(k)} - |V_i|^{central}|}{\sum_{i=1}^N |V_i^{(k)}|} \times 100 \quad (5.24)$$

$$MIAE(k) = \frac{1}{N} \sum_{i=1}^N |\theta_i^{(k)} - \theta_i^{central}| \quad (5.25)$$

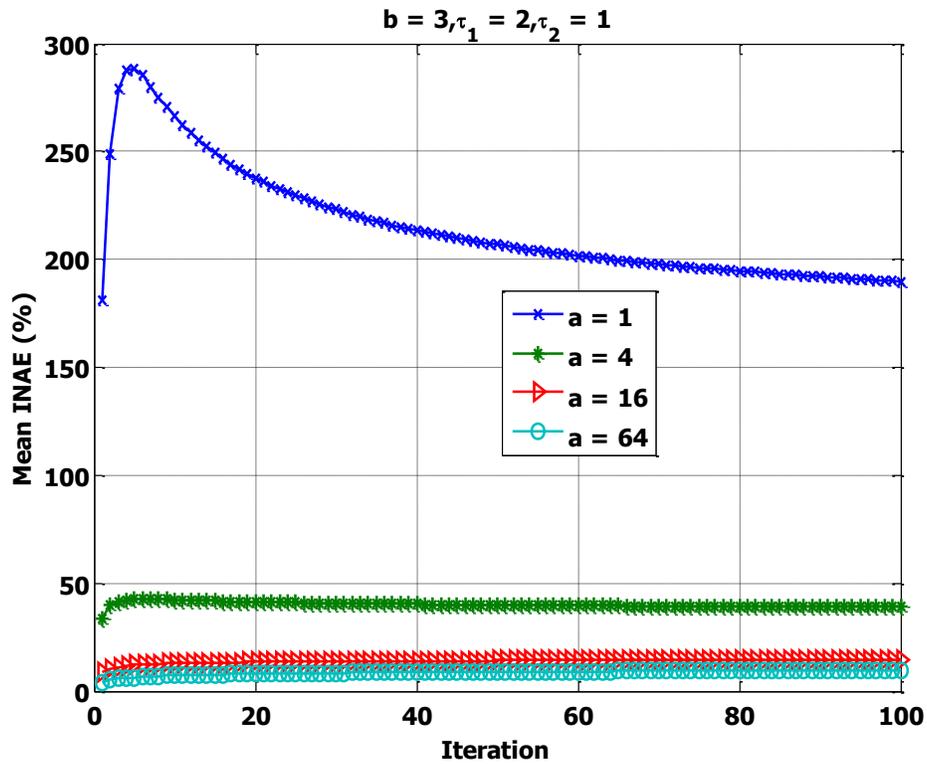
where, *INAE* and *MIAE* follows from equations (3.8) and (4.10), respectively. k represents the iteration index. The superscript “*central*” indicates the voltage phasors estimated using conventional WLS method as in equation (2.8). *INAE* and *MIAE* represents the average estimation error in voltage magnitudes and angles, respectively. The entire process of distributed estimation is repeated over 1000 independent Monte Carlo trials. At each

iteration, the mean values of $INAE$ and $MIAE$ are obtained from these trials. Figures 5.2 and 5.3 illustrate respective effects of parameter a and b over the mean convergence of error in voltage phasors.

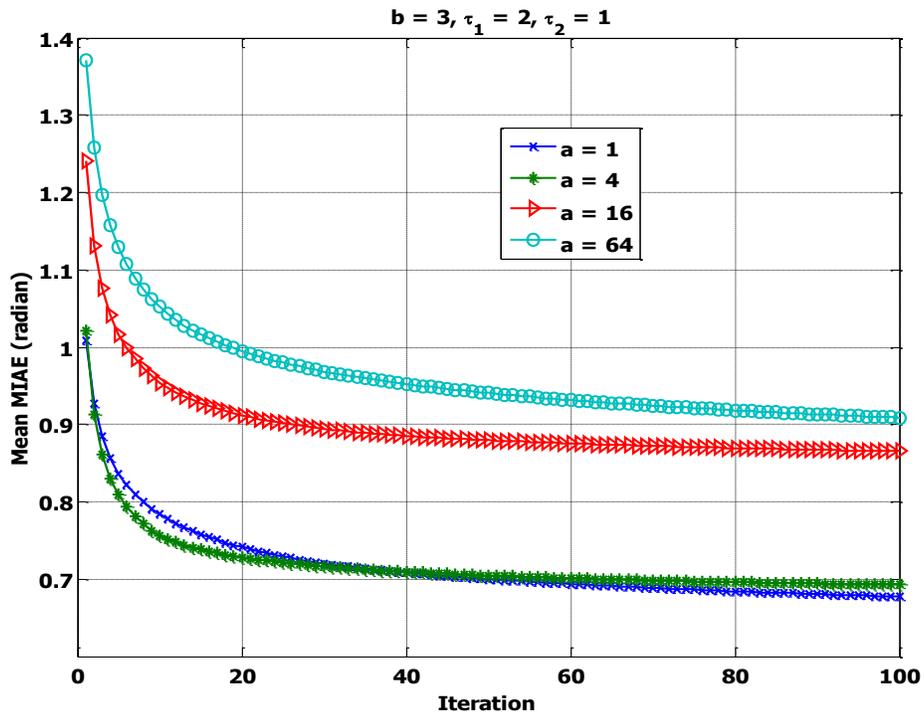
From Fig.5.2 it can be noticed that, with the increase of a , voltage magnitude estimation error ($INAE$) decreases, while phase angle estimation error ($MIAE$) increases with the former being more significant. As shown in Fig.5.3, parameter b has similar effect over $MIAE$. However, $INAE$ seems to increase with the increase of b . When $a < 2$, $INAE$ is high enough to conclude as adequately convergent over iterations. Similarly, $INAE$ is divergent over iterations when $b < 2$. Thus, it is observed that, the conditions of convergence are applied over the voltage magnitude estimation errors whereas, the phase angle estimation errors converge asymptotically, with minimal effect of a and b . Henceforth, for satisfactory convergence, the parameters a, b and exponents τ_1, τ_2 should be selected such that, (1) $a > b > 2$ and (2) $\tau_1 > \tau_2$.

5.5 Conclusions

In this chapter, we propose a distributed state model and an agent based static state estimation method for smart distribution grid. We specially consider the case when for each agent, the local measurement model is underdetermined and all state elements for a particular agent is completely shared with its neighbors. An agent-wise distributed local consensus procedure is developed, which at every iteration incorporates the least square mismatch. Simulation results on a radial distribution grid show that the proposed method can give satisfactory convergence based on the appropriate selection of constants and exponents. An in-depth analysis of convergence gives an optimum selection criteria of these constants ensuring universal application in multi-agent systems. In the next chapter, we generalize the multi-agent formulation for dynamic state estimation and study the relative performance of Kalman consensus filter and diffusion Kalman filter.

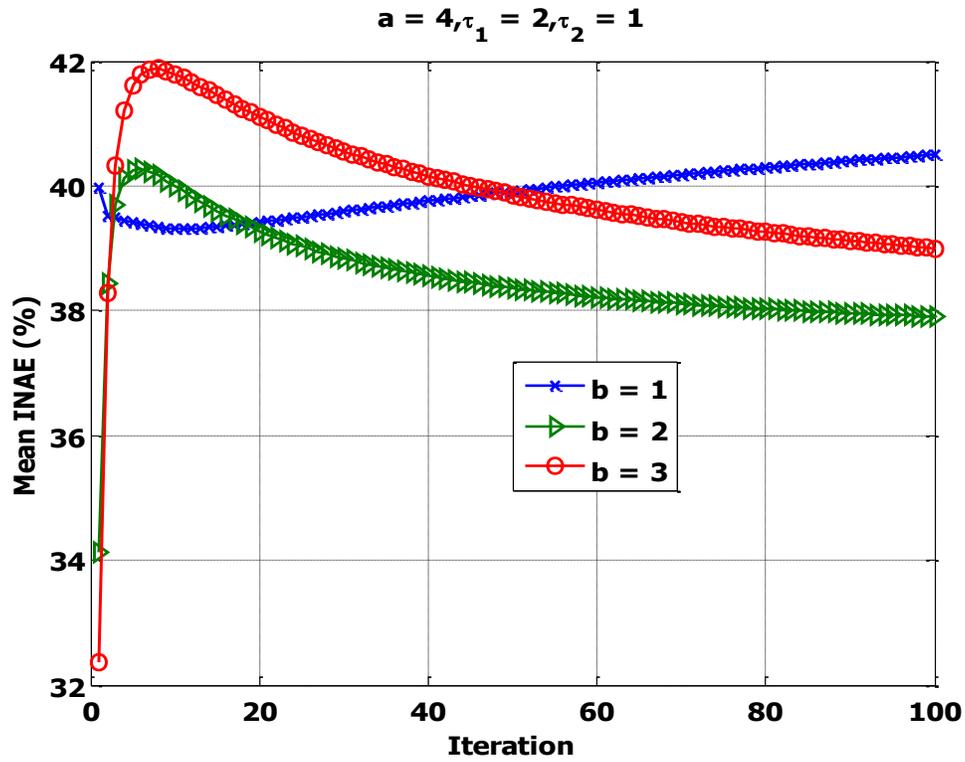


(a)

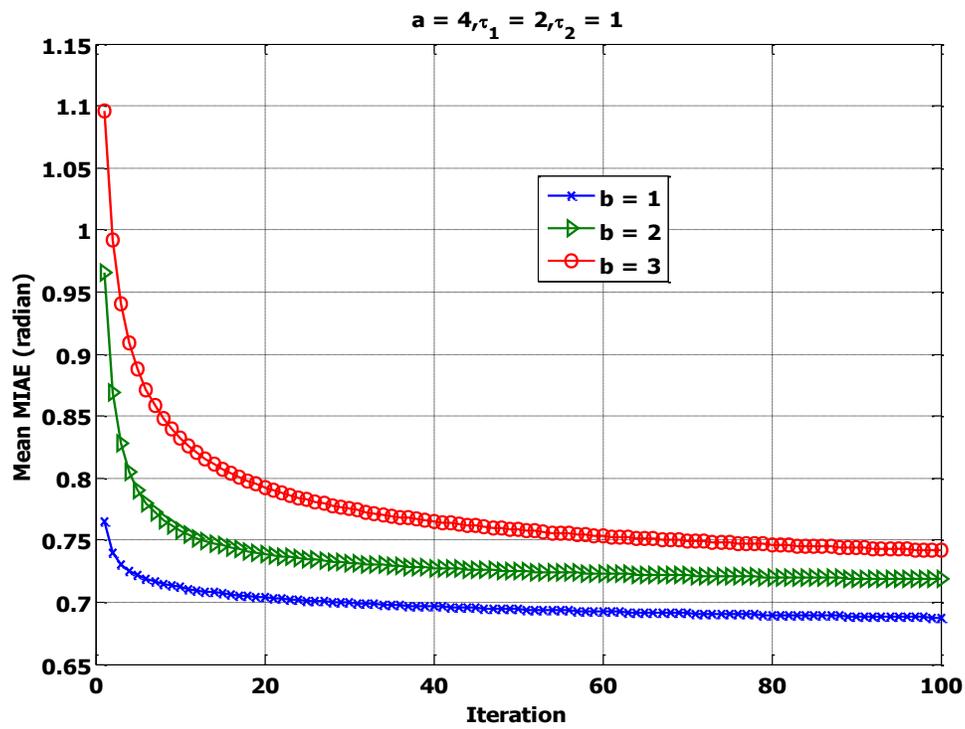


(b)

Figure 5.2: Effect of a : (a) Mean Convergence of Error in Voltage Magnitudes. (b) Mean Convergence of Error in Voltage Angles.



(a)



(b)

Figure 5.3: Effect of b : (a) Mean Convergence of Error in Voltage Magnitudes. (b) Mean Convergence of Error in Voltage Angles.

Chapter 6

Distributed Agent based Dynamic State Estimation over a Lossy Network

In this chapter, a novel distributed agent based dynamical system estimation strategy is proposed. We consider a spatially distributed system where each agent has a local observation space and is interested in a specific set of system state elements. The agents have the ability of two-way communication with its neighbors (i.e., agents who share at least one state element). At a particular time instant, each agent predicts its state and makes intermediate correction based on its local measurements. Information about the corrected state elements are then exchanged among the neighboring agents. Based on the final processing of these exchanged information, an agent based Kalman consensus Filter (AKCF) and uniform weighting based diffusion Kalman filter (ADKF) are proposed in the light of well-established theory of distributed Kalman filtering. Two different systems are simulated using the proposed filters. The effect of communication is also investigated by introducing random failures in the communication link among neighboring agents. It is observed that the mean square deviation (MSD) of AKCF is lower than that of ADKF for the scenarios

considered. Additionally, the results also demonstrate that the AKCF is more robust to communication link failures than the ADKF.

6.1 Introduction

Through the last five decades, Kalman filter [95] has been one tool of choice for real-time estimation and tracking of dynamical processes. In a classical filtering approach, the sensors communicate with a single fusion center either directly or hierarchically to send updated measurement information in timely manner. Based on the knowledge of previous state values and overall system dynamics, the fusion center makes a minimum mean squared error (MMSE) prediction of the states. Necessary corrections are made to the predicted states based on sensor measurements. Now-a-days, physical systems, like the electric power grid can be tracked via a distributed network of communicable sensors typically over a large geographical region [49, 50]. The underlying communication and computational burden is considerably high with a centralized or hierarchical implementation of a Kalman filter [39]. This issue is resolved through distributed implementation of Kalman filters that also offers benefits in terms of real-time implementation and configuration flexibility. In this scenario, the sensors have additional responsibility of implementing a local Kalman filter and inter-sensor communication. The objective of each sensor is to have updated status of the overall system through local prediction and necessary correction based on the type of information (measurement and/or predicted state values) exchanged among the neighbors. Although, the fundamental concept is unchanged, the distributed Kalman filter has evolved through numerous versions. Among those, Kalman consensus filter (KCF) [51] and diffusion Kalman filter (DKF) [52] are worth mentioning. In both setups, the sensors collect detailed information about the measurements (i.e., sensed data and noisy observation space model) from neighbors. The fused information is then applied to the classical Kalman filtering algorithm. Thus, an intermediate estimate of the whole dynamics and the corresponding update in es-

estimation error covariance matrix is obtained at each sensor. At the last stage, the sensors exchange information about the intermediate estimates with their neighbors. A correction is made to each sensor's estimate by either applying consensus (i.e., KCF), or through a weighted combination of the received neighborhood information (i.e., DKF). Finding the desired degree of consensus or the optimal weights for diffusion are the major issues in the design of such distributed Kalman filters. For the Kalman consensus filter, the objective is to design a scalar consensus gain parameter. In this regard, an optimization problem is solved in [53] in order to find the desired consensus gain that ensures the convergence of the over all estimation error covariance to a known steady-state matrix. However, it is also mentioned that such steady-state matrix is hard to obtain in practical perspective and an approximate expression for the consensus parameter is derived. On the other hand, in [54] a Lyapunov stability analysis is carried out over the estimation error dynamics to find a range of real numbers within which the desired consensus parameter should be chosen. When it comes to implementing weighted diffusion of neighborhood estimates, the Laplacian [52], the Metropolis [55] and the nearest neighbor or uniform weighting rules [56] are the popular choices in designing the diffusion weights.

It should be noted that, besides the system dynamics and local observation space, the topology of active sensor network as well as the inter-sensor communication reliability play vital roles in successful implementation of distributed Kalman filters. In this regard, Kalman consensus filter is characterized under the effect of lossy sensor network in [53] by incorporating a Bernoulli random variable in the consensus step. The effect is only illustrated through simulations necessitating a theoretical analysis that involves the consensus gain parameter and network reliability. The usage of mutually independent Bernoulli distributions is also reported in [57] to model the random presence of nonlinear dynamics as well as the quantization effect in the sensor communication. The filter designed is of diffusion characteristics and the corresponding weights are derived based on the average \mathcal{H}_∞ performance. On the other hand, relative variance and adaptive combination rule is proposed in [56] for

stationary diffusive estimation of single parameter under noisy communication link.

As a matter of fact, most of the advanced distributed Kalman filters are based on these two major approaches, although they bear widely varying application specific characteristics. We would like to refer [58] as a resource that summarizes the extensive research carried out in this arena.

Dynamic state estimation in a large scale cyber-physical system (CPS) (e.g., Smart Grid) presents some unique challenges namely, (1) dimension of the state vector is quite large, (2) sensors are spatially distributed over a large area. Consequently, sensors are required to be fully connected in order to regularly store and update the global state vector and estimation error covariance matrix. Additionally, the number of communication links, computational memory and hardware requirements increase with the increase in active sensor nodes. Thus, it may be impractical to track the high dimensional state vector in its entirety at each communicable sensor. This constraint can be overcome specifically for sparse large-scale linear systems [59, 60]. In this case, the corresponding transition of states can be reflected on (approximately) banded matrix to spatially decompose the overall dynamics among sensors even when local measurement space projects onto global states. This idea is further extended for system specific reduced order particle filtering [61] and distributed observer design for large-scale system partitioned into disjoint areas [62]. The key fact is, the observation space of each sensor is modified solely based on the characteristics of state dynamics.

On the contrary, in a practical physical system, overlaid with communication network, the observation space of a sensor may be coupled to a limited set of specific state elements. Some state elements may even be coupled to two or more sensors' observation space. Under these circumstances, each sensor may be relieved to track only the pertinent state elements. In this way, the sensors are acting as agents and the underlying system can be referred to as a multi-agent system.

In this chapter, an agent based general formulation of KCF and DKF is proposed. One is called agent based KCF (AKCF) and the other one is agent based DKF (ADKF). Each

agent has access to a distinct set of measurements, that are coupled to a subset of global state elements. A set of binary projection matrices are defined based on the distribution of system state elements over the observation space of the agents. These matrices map the overall system dynamics to agent-specific state-space model and also define the set of neighbors of a particular agent. AKCF and ADKF are developed by proper inclusion of these projection matrices in the basic filtering steps. The application of proposed filters are illustrated with two custom built 3-agent systems to make a comparative performance analysis. The effect of losses in the inter-agent information exchange is also investigated by allowing random and independent failure of the existing communication links.

6.2 System Model

We consider a system whose dynamics can be modeled in discrete time as 1st order Gauss-Markov Process, i.e.,

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{w}_{t-1}; t = 0, 1, 2, \dots \quad (6.1)$$

where, the overall system state is represented by the n -dimensional state vector \mathbf{x}_t at time instant t . The initial values of the state vector elements at $t = -1$ follow Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Unlike [59, 60, 61, 62], the state transition matrix $\mathbf{F} \in \mathbb{R}^{n \times n}$ is a general square matrix with eigenvalues lying within a unit circle. The process noise $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. The underlying physical system is observed by N agents. The linear observation model for the k^{th} agent is,

$$\mathbf{y}_{t,k} = \mathbf{H}_k \mathbf{x}_{t,k} + \mathbf{v}_{t,k}; k = 1, 2, \dots, N \quad (6.2)$$

where, the n_k -dimensional vector $\mathbf{x}_{t,k}$ is Agent k 's local state vector – a subset of \mathbf{x}_t . The observation matrix $\mathbf{H}_k \in \mathbb{R}^{m_k \times n_k}$ is full row rank (i.e., $m_k \leq n_k$). The measurement noise $\mathbf{v}_{t,k} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ and is independent of process noise. Unlike conventional distributed

Kalman filtering, a particular agent k attempts to estimate only the local state vector $\mathbf{x}_{t,k}$ instead of \mathbf{x}_t . Consequently, an agent-specific local system dynamics is associated with the filtering problem. In this regard, we introduce a binary projection matrix, which extracts out the state elements from the global set of states and rearrange them according to the observation space of a particular agent. For the k^{th} agent, we name it \mathbf{T}_k such that,

$$\mathbf{x}_{t,k} = \mathbf{T}_k \mathbf{x}_t; k = 1, 2, \dots, N. \quad (6.3)$$

The projection matrix \mathbf{T}_k is an $n_k \times n$ matrix and rank deficient ($n_k < n$). Therefore, exact recovery of the global set of states \mathbf{x}_t from individual agent's local state vector $\mathbf{x}_{t,k}$ is not possible and also not required in current scenario. We define it as binary projection matrix since intuitively the matrix is projecting the n -dimensional state vector onto n_k -dimension. Here, by projection, we imply the extraction of one state element at a time instead of creating a new one from some weighted combinations. Such binary matrices are first reported in [66] to define multiple subsystems for decentralized model predictive control. Unlike the proposed approach, the decomposition is performed solely based on the sparse or block diagonal characteristics of the state transition matrix. Nevertheless, the projection matrix also follows Lemma 6.1.

Lemma 6.1. *For the binary projection matrix \mathbf{T}_k of dimension $n_k \times n$ ($n_k < n$), $\mathbf{T}_k \mathbf{T}_k^\top = \mathbf{I}_{n_k}$. Here, \mathbf{I}_{n_k} is an identity matrix of dimension n_k .*

Proof. An alternate representation of the projection matrix is, $\mathbf{T}_k = \begin{bmatrix} \mathbf{I}_{n_k} & \mathbf{O} \end{bmatrix} \mathcal{P}$. Here, \mathbf{O} is $n_k \times (n - n_k)$ matrix of zeros and \mathcal{P} is an $n \times n$ permutation matrix. Using the properties of permutation matrices [116],

$$\mathbf{T}_k \mathbf{T}_k^\top = \begin{bmatrix} \mathbf{I}_{n_k} & \mathbf{O} \end{bmatrix} \mathcal{P} \mathcal{P}^\top \begin{bmatrix} \mathbf{I}_{n_k}^\top \\ \mathbf{O}^\top \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_k} & \mathbf{O} \end{bmatrix} \mathbf{I}_n \begin{bmatrix} \mathbf{I}_{n_k}^\top \\ \mathbf{O}^\top \end{bmatrix} = \mathbf{I}_{n_k}$$

□

Using the projection matrix \mathbf{T}_k , the system dynamics of k^{th} agent can be mapped from equation (6.1),

$$\mathbf{x}_{t,k} = \mathbf{T}_k \mathbf{F} \mathbf{x}_{t-1} + \mathbf{w}_{t-1,k} \quad (6.4)$$

where, $\mathbf{w}_{t-1,k} = \mathbf{T}_k \mathbf{w}_{t-1}$. Therefore, $\mathbf{w}_{t-1,k} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ with $\mathbf{Q}_k = \mathbf{T}_k \mathbf{Q} \mathbf{T}_k^\top$. It should be noted that, the local system dynamics for k^{th} agent should reflect the corresponding observation model of equation (6.2). As a consequence, the desired 1st order Gauss-Markov process of k^{th} agent can be expressed as,

$$\mathbf{x}_{t,k} = \mathbf{F}_k \mathbf{x}_{t-1,k} + \mathbf{w}_{t-1,k}, \quad (6.5)$$

The above model conforms with equation (6.4) if and only if $\mathbf{F}_k \mathbf{T}_k = \mathbf{T}_k \mathbf{F}$. Or, using Lemma 6.1, $\mathbf{F}_k = \mathbf{T}_k \mathbf{F} \mathbf{T}_k^\top$. It is important to note that, this irreversible mapping is intended only to design the agent based Kalman filter, and subsequent derivation of theoretical bounds for filter stability. Later in Chapters 7 and 8, a case study is carried out to verify these bounds, where equations (6.1) (global system dynamics) and (6.3) (local state mapping) serve as true state vectors for individual agents.

The static set of *physical* neighbors for the k^{th} agent is defined based on the overlap/sharing of state elements. Mathematically,

$$\mathcal{S}_k = \{i : \mathbf{P}_{i,k} \mathbf{x}_{t,i} \text{ projects onto } \mathbf{L}_{i,k} \mathbf{x}_{t,k}, \forall t\} \quad (6.6)$$

Here, $\mathbf{P}_{i,k}$ and $\mathbf{L}_{i,k}$ are $n_k \times n_i$ and $n_k \times n_k$ binary projection matrices, respectively. When agent i sends state information to agent k , the projection matrix $\mathbf{P}_{i,k}$ is multiplied to the sender agent's vector of shared states whereas $\mathbf{L}_{i,k}$ is used with receiver agent's shared state vector. By default, $\mathbf{P}_{k,k} = \mathbf{L}_{k,k} = \mathbf{I}_{n_k}$.

6.3 Proposed Method

The agent based dynamic state estimation procedure is developed with minimum mean square error (MSE) as the metric of interest. The state vector estimated by the k^{th} agent at discrete time instant i is defined as,

$$\hat{\mathbf{x}}_{i,k|j} = \mathbb{E}[\mathbf{x}_{i,k} | \mathbf{y}_{0,k}, \mathbf{y}_{1,k}, \dots, \mathbf{y}_{j,k}] \quad (6.7)$$

The corresponding error covariance matrix is,

$$\mathbf{M}_{i,k|j} = \mathbb{E}[(\mathbf{x}_{i,k} - \hat{\mathbf{x}}_{i,k|j})(\mathbf{x}_{i,k} - \hat{\mathbf{x}}_{i,k|j})^\top] \quad (6.8)$$

The first five steps of estimation are performed according to traditional Kalman filtering, which are represented according to the definitions given in equations (6.7) and (6.8). For the k^{th} agent,

- Initialization:

$$\hat{\mathbf{x}}_{-1,k|-1} = \boldsymbol{\mu}_k, \mathbf{M}_{-1,k|-1} = \boldsymbol{\Sigma}_k; \quad (6.9)$$

where $\boldsymbol{\mu}_k = \mathbf{T}_k \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_k = \mathbf{T}_k \boldsymbol{\Sigma} \mathbf{T}_k^\top$.

- Prediction:

$$\hat{\mathbf{x}}_{t,k|t-1} = \mathbf{F}_k \hat{\mathbf{x}}_{t-1,k|t-1} \quad (6.10)$$

- Update Error Covariance:

$$\mathbf{M}_{t,k|t-1} = \mathbf{F}_k \mathbf{M}_{t-1,k|t-1} \mathbf{F}_k^\top + \mathbf{Q}_k \quad (6.11)$$

For this process to be stable the spectral radius of \mathbf{F}_k has to be less than 1. The next two steps are based on k^{th} agent's observation model.

- Minimum MSE:

$$\mathbf{M}_{t,k|t} = \left(\mathbf{M}_{t,k|t-1}^{-1} + \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \right)^{-1} \quad (6.12)$$

- Intermediate Correction:

$$\hat{\mathbf{b}}_{t,k} = \hat{\mathbf{x}}_{t,k|t-1} + \mathbf{M}_{t,k|t} \mathbf{H}_k^\top \mathbf{R}_k^{-1} (\mathbf{y}_{t,k} - \mathbf{H}_k \hat{\mathbf{x}}_{t,k|t-1}) \quad (6.13)$$

At the last step of estimation, $\hat{\mathbf{b}}_{t,k}$ is used along with the exchanged information from the neighbors to arrive at the estimate of individual agents' local state values. This step can be performed either using consensus [54] or uniformly diffusing the exchanged information [52]. These approaches are called agent based Kalman consensus filter (AKCF) and Diffusion Kalman filter (ADKF), respectively [83]. The mathematical representation of this step of information exchange is given below,

- Final Correction: AKCF

$$\hat{\mathbf{x}}_{t,k|t} = \hat{\mathbf{b}}_{t,k} + \epsilon \mathbf{M}_{t,k|t} \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \hat{\mathbf{x}}_{t,k|t-1}) \quad (6.14)$$

where, $0 < \epsilon \leq 1$. Larger value of ϵ allows greater contribution of consensus and vice versa.

- Final Correction: ADKF

$$\hat{\mathbf{x}}_{t,k|t} = \mathbf{D}_k \sum_{i \in \mathcal{S}_k \cup \{k\}} \mathbf{P}_{i,k} \hat{\mathbf{b}}_{t,i} \quad (6.15)$$

where, $\mathbf{D}_k = \text{diag}(d_k[1], d_k[2], \dots, d_k[n_k])$. For uniform weighting,

$$d_k[j] = \frac{1}{\text{Number of Agents that share the state } \mathbf{x}_k[j]} \quad (6.16)$$

Here, $\mathbf{x}_k[j]$ represents the j^{th} element of agent k 's local state vector.

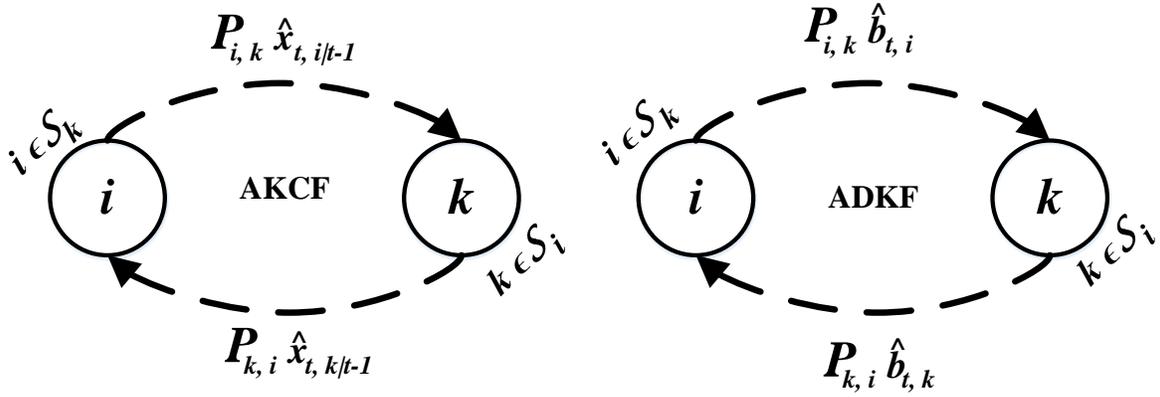


Figure 6.1: Inter-agent Information Exchange

Furthermore, unlike conventional algorithms of distributed Kalman filter [], the *agents are not required to aggregate information about their neighbors' measurements or individual error covariance matrices.*

6.4 Effect of Lossy Communication Network

In the proposed method of agent based filtering, it is evident that only the information about relevant state elements are being exchanged among neighbors. This is illustrated in Fig. 6.1. In AKCF, the k^{th} agent exchanges information about the predicted state elements obtained in equation (6.10) with its neighbors. Whereas, in ADKF, the intermediate correction vector obtained in equation (6.13) is exchanged. Therefore, the inter-agent two way information exchange plays an important role in agent based Kalman filtering and can be hampered if the underlying communication link fails. These circumstances can be simulated by introducing random link failures (RLF). Mathematically, effect of RLF can be analyzed by inserting Bernoulli random variables $\zeta_{i,k}(t)$ in equations (6.14) and (6.15). These random variables have the following probability mass functions,

$$\zeta_{i,k}(t) = \begin{cases} 0 & \text{with Prob. } \rho_{i,k} \\ 1 & \text{with Prob. } 1 - \rho_{i,k} \end{cases}; \forall i \in \mathcal{S}_k, \forall k. \quad (6.17)$$

Here, $\rho_{i,k}$ represents the probability of failure to send information from agent i to a neighbor agent k . From the properties of Bernoulli random variables, $\mathbb{E}[\zeta_{i,k}(t)] = 1 - \rho_{i,k}$. Consequently, the final correction step in AKCF becomes,

$$\hat{\mathbf{x}}_{t,k|t} = \hat{\mathbf{b}}_{t,k} + \epsilon \mathbf{M}_{t,k|t} \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{P}_{i,k} \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \hat{\mathbf{x}}_{t,k|t-1}) \quad (6.18)$$

It is evident that only the consensus part of equation (6.18) is affected by communication. On the contrary, the final correction step of ADKF is modified as follows,

$$\hat{\mathbf{x}}_{t,k|t} = \mathbf{D}_{t,k} \sum_{i \in \mathcal{S}_k \cup \{k\}} \mathbf{P}_{i,k} \hat{\mathbf{b}}_{t,i} \quad (6.19)$$

where, $\mathbf{D}_{t,k} = \text{diag}(d_{t,k}[1], d_{t,k}[2], \dots, d_{t,k}[n_k])$. And for uniform weighting,

$$d_{t,k}[j] = \frac{1}{c_{t,k}[j]} \quad (6.20)$$

where, $c_{t,k}[j]$ represents the number of *Successfully Received* estimates for $\mathbf{x}_k[j]$ at discrete time instant t .

The following assumptions are made in analyzing the lossy network effect over the agent based estimation problem.

Assumption 6.1. *The random events $\zeta_{i,k}(t)$ and $\zeta_{k,i}(t)$ are independent of each other $\forall i \in \mathcal{S}_k; \forall k$.*

Assumption 6.2. *The random events $\zeta_{i,k}(t)$ and $\zeta_{j,l}(t)$ are independent of each other $\forall i \neq j; \forall k \neq l$.*

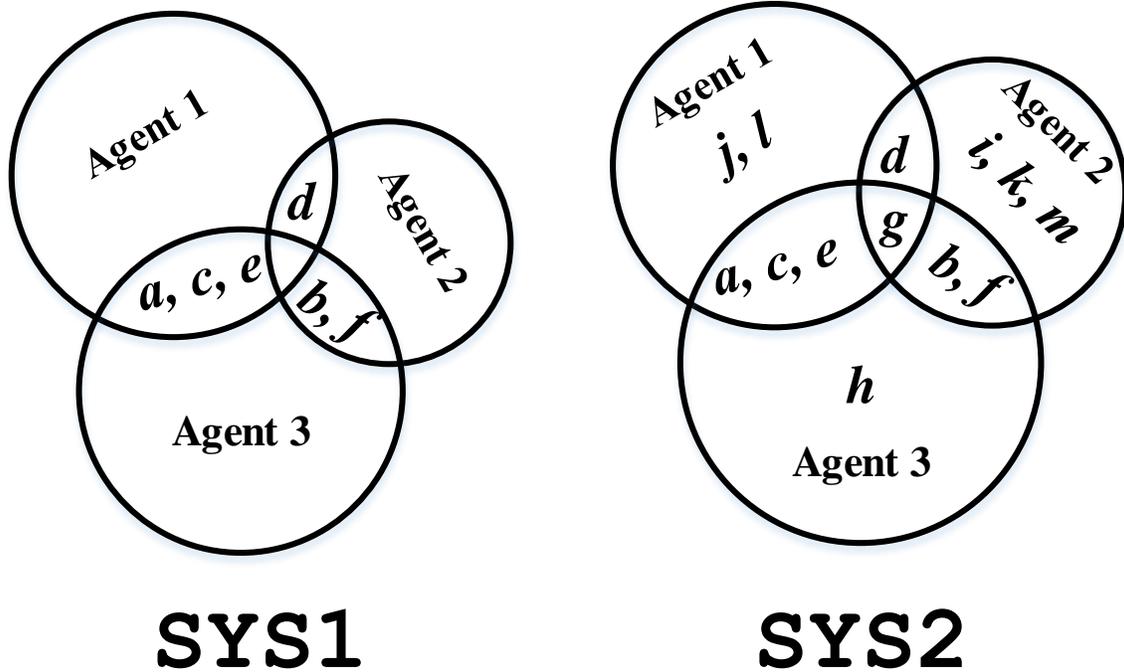


Figure 6.2: Venn diagrams for *SYS1* and *SYS2*

6.5 Simulation and Results

We investigate the performance of the proposed filters for two 3-agent systems. The set of global state elements for the 1st system *SYS1* is $\{a, b, c, d, e, f\}$ and that for the 2nd system *SYS2* is $\{a, b, c, d, e, f, g, h, i, j, k, l, m\}$. The Venn diagrams in Fig. 6.2 show the agent-wise distribution of the state elements for the two systems. It should be noted that in *SYS2*, there exist some state elements *strictly* local to the agents, whereas, each of the state elements is shared between two agents in *SYS1*. The rest of the parameters of *SYS1* and *SYS2* are given in Appendix A and B, respectively [83]. The proposed filters are applied to these systems as illustrative examples.

6.5.1 Figure of Merit

To investigate the performance of AKCF and ADKF, an estimation error vector associated with each agent is calculated. The estimation error of k^{th} agent at discrete time instant t

is, $\boldsymbol{\eta}_{t,k} = \hat{\mathbf{x}}_{t,k|t} - \mathbf{x}_{t,k}$. A global estimation error vector is formed by stacking $\boldsymbol{\eta}_{t,k}$ from all agents,

$$\boldsymbol{\eta}_t = \left[\boldsymbol{\eta}_{t,1}^\top \cdots \boldsymbol{\eta}_{t,N}^\top \right]^\top$$

This vector is used to define the mean square deviation (MSD) as follows,

$$\mathbf{MSD}_t = \text{trace} \left(\mathbb{E} \left[\boldsymbol{\eta}_t \boldsymbol{\eta}_t^\top \right] \right) \quad (6.21)$$

MSD is used as the figure of merit to compare the performance of AKCF and ADKF. The lower the MSD is, the better. In the upcoming subsections we present the performance of the proposed filters in terms of MSD obtained from simulations.

6.5.2 Case Study: Perfect Communication

In this scenario, all the inter-agent communication links are assumed to be working perfectly. The proposed ADKF is applied to **SYS1** with uniform weighting rule (equation (6.16)). AKCF is applied to **SYS1** with different values of ϵ . MSD is calculated from 1000 independent Monte Carlo trials at each time step. The comparative performance of ADKF and AKCF for **SYS1** is shown in Fig. 6.3. In the same way, AKCF and ADKF is applied to **SYS2** and the performance is summarized in Fig. 6.4. The MSD values from Fig. 6.3 and Fig. 6.4 differ from the classical distributed filtering approach [52] by showing better performance of AKCF, irrespective of the selection of ϵ within the prescribed range [83]. In particular, the best performance of AKCF is obtained when ϵ is selected to be 0.1 and 0.01 for **SYS1** and **SYS2**, respectively. The optimum values of ϵ thus obtained is used in AKCF for the following case study.

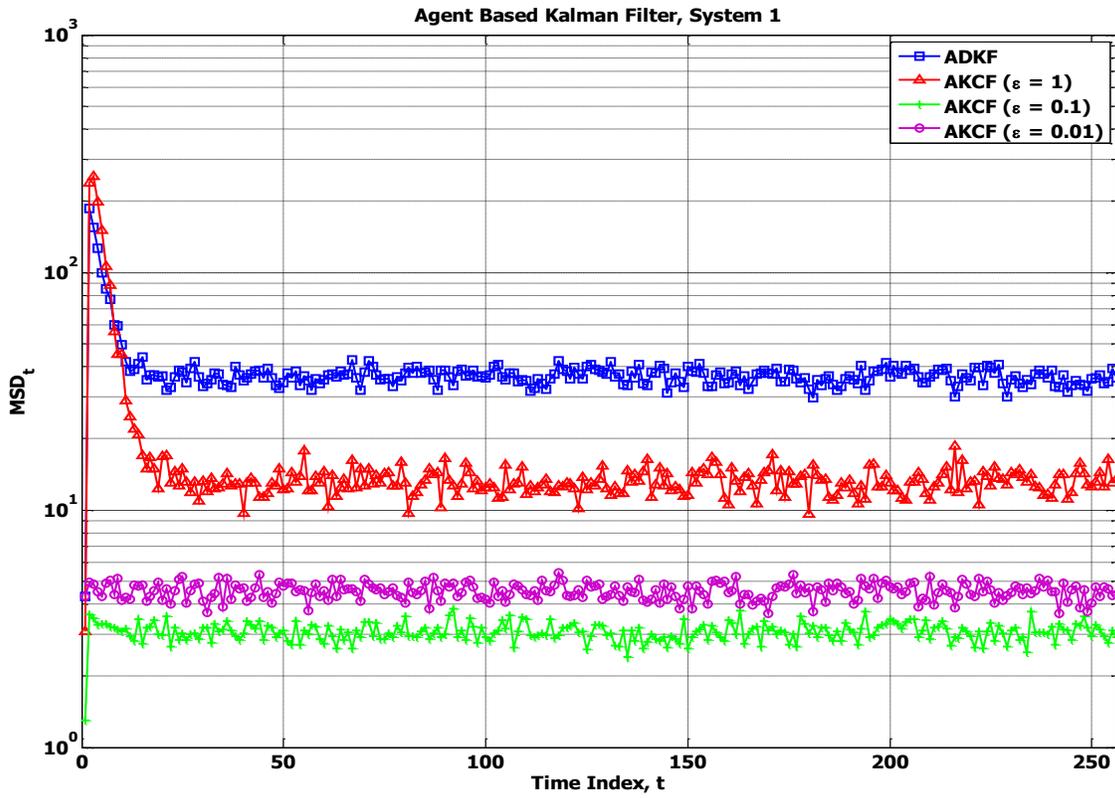


Figure 6.3: *AKCF versus ADKF for SYS1*

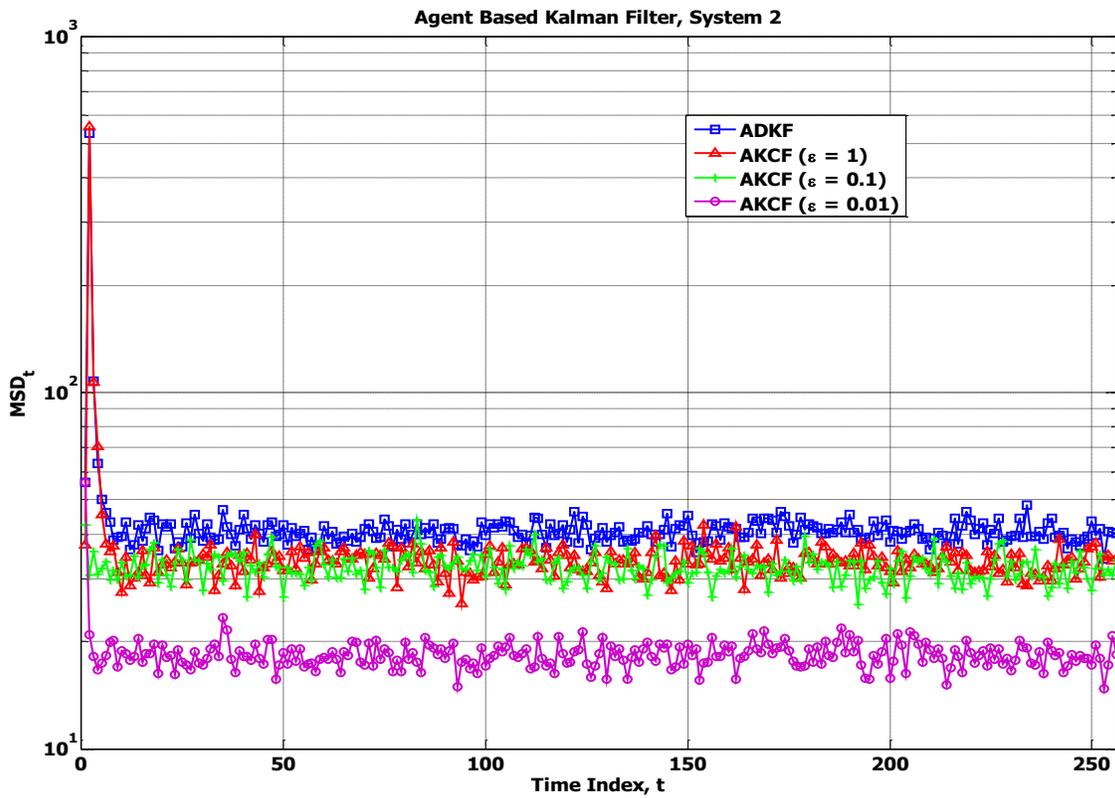


Figure 6.4: *AKCF versus ADKF for SYS2*

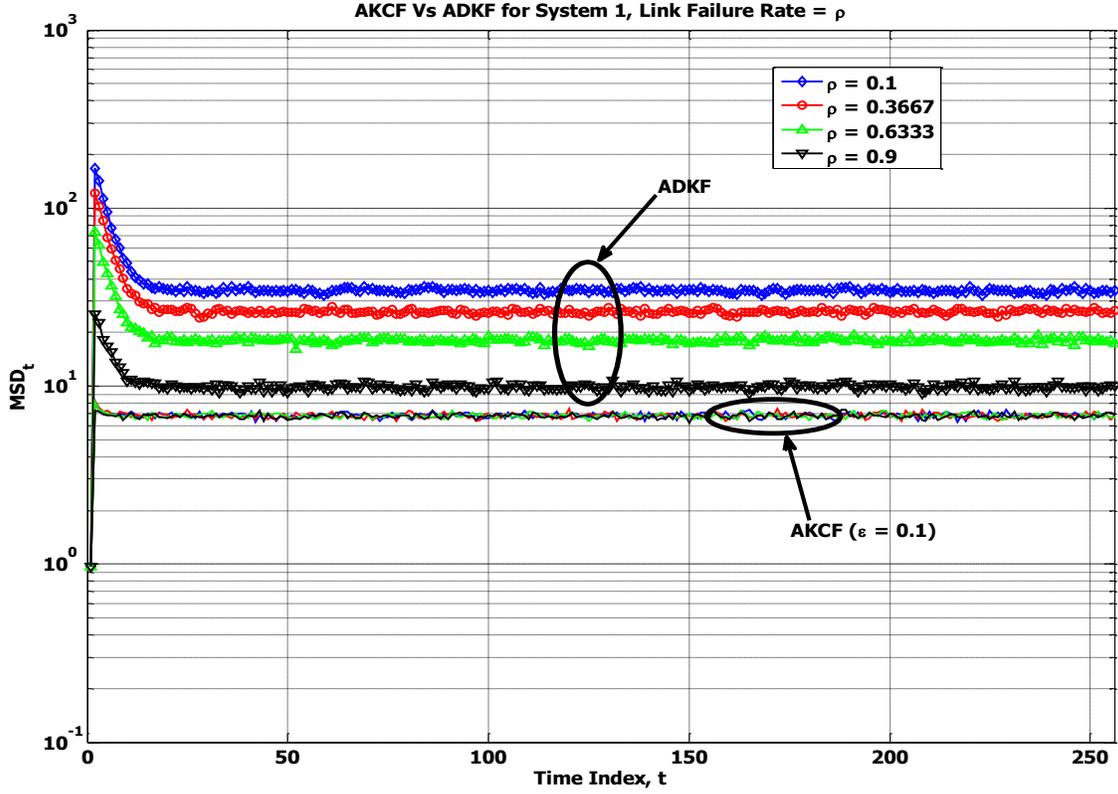


Figure 6.5: *Effect of communication for SYS1*

6.5.3 Case Study: Random Link Failure

The effect of inter-agent communication is investigated for SYS1 and SYS2 by using equations (6.18-6.20) at the information exchange steps of the proposed filters. For simplicity, the probability of link failure, $\rho_{i,k} = \rho; \forall i, k, i \neq k$. The MSD is obtained at different link failure rates, which is illustrated in Fig. 6.5 for SYS1. Based on the previous case study, the value of ϵ in AKCF is 0.1. The whole procedure is repeated for SYS2 with the corresponding value of ϵ in AKCF being 0.01. Fig. 6.6 shows the relative performance for SYS2 affected by imperfect communication. The effect of faulty inter-agent communication link is insignificant for AKCF as evident from Fig. 6.5 and Fig. 6.6. This is because of relatively small values of ϵ chosen for the two systems. On the other hand, a high link failure rate results in less contribution from neighboring agents in the final correction step of ADKF. It is interesting to see that ADKF performs better when the communication link is highly unreliable. While

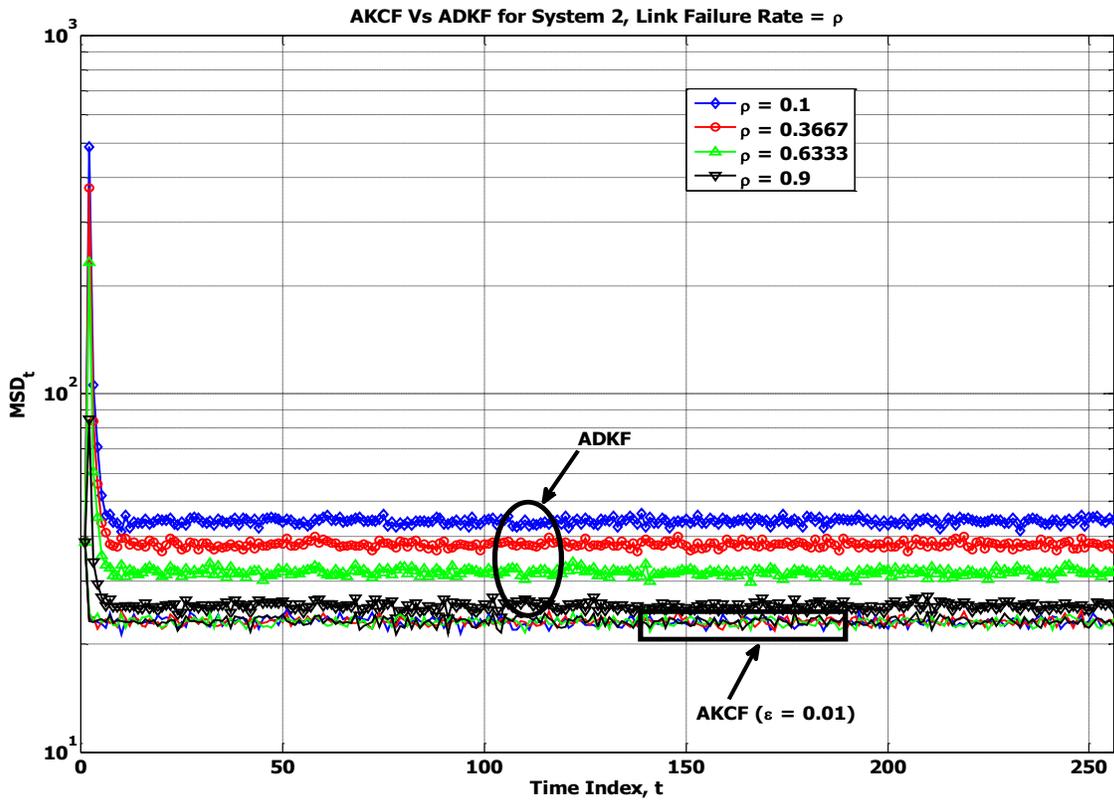


Figure 6.6: *Effect of communication for SYS2*

this may appear counter intuitive, it is in fact a direct consequence of the system parameter choice. These observations suggest that, for the two systems considered in our simulations, the underlying system states are more dependent upon the agent-wise observation space as compared to the system dynamics itself. Nevertheless, the steady-state MSD values are smaller for AKCF, irrespective of the choice of ϵ value as well as the condition of inter-agent communication link and is more robust than ADKF [83].

6.6 Conclusions

An agent based Kalman consensus and diffusion Kalman filter is proposed. Each agent is interested in a distinct subset of state elements and is able to communicate to its neighbors who share at least one state element. The proposed filtering procedures are applied to two multi-agent systems to compare their performance. The effect of communication is also observed for the agent based Kalman filters. It is observed that AKCF performs better than ADKF for both systems even under random failure of inter-agent communication link. In the next chapter we propose a strategy of choosing optimal consensus weights in AKCF formulation both under perfect and lossy communication network.

Chapter 7

Agent based Optimally Weighted Kalman Consensus Filter over a Lossy Network

7.1 Introduction

The primary study of agent based Kalman filtering in Chapter 6 indicates the superior and more robust characteristics of the consensus approach in inter-agent information exchange. Consequently, this chapter further investigates the AKCF performance by incorporating a varying consensus weight. In addition to that individual agent's state elements are clearly distinguished as shared and unshared ones. The optimal weighting of consensus is derived through Lyapunov function based stability analysis of corresponding estimation error. The effect of communication is also investigated by introducing random failures in the communication link among neighboring agents. The corresponding bounds on the degree of consensus and inter-agent link failure rate are also derived for stable implementation of the agent based dynamic estimation. The proposed filter is applied to a custom built 2-agent system to conform the desired optimal limit for the degree of consensus based information

exchange both under perfect and lossy communication network.

7.2 Shared and Unshared State Elements in Agent based model

Given the agent based observation model 6.2 and state dynamics 6.5, let us redefine the static set of *physical* neighbors for the k^{th} agent based on the overlap/sharing of state elements. Mathematically,

$$\mathcal{S}_k = \{i : \mathbf{P}_{i,k} \mathbf{S}_i \mathbf{x}_{t,i} \text{ projects onto } \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{x}_{t,k}, \forall t\}. \quad (7.1)$$

Here, $\mathbf{P}_{i,k}$ and $\mathbf{L}_{i,k}$ are $n_k^s \times n_i^s$ and $n_k^s \times n_k^s$ binary projection matrices, respectively. n_k^s is the number of state elements shared between agent k and its neighbors. Similar definition applies for n_i^s . When agent i sends state information to agent k , the projection matrix $\mathbf{P}_{i,k}$ is multiplied to the sender agent's vector of shared states whereas $\mathbf{L}_{i,k}$ is used with receiver agent's shared state vector. By default, $\mathbf{P}_{k,k} = \mathbf{L}_{k,k} = \mathbf{I}_{n_k^s}$. The other binary matrices ($\mathbf{S}_i, \mathbf{S}_k$) are required to extract the shared state elements for the respective agents. In addition to that, a reordering binary matrix \mathbf{O}_k^s is used to assign the shared state elements to their original position in the k^{th} agent's state vector with the entries of unshared elements kept zero. Similar operation of extracting and reordering can be defined for completely unshared state elements for any agent. Fig. 7.1 illustrates these operations for shared and unshared state elements of agent k . The decomposition into shared and unshared portion and recombining them should constitute the original state vector for the k^{th} agent. Thus, mathematically, it is required that [84],

$$\Psi_k \mathbf{x}_{t,k} = \mathbf{O}_k^s \mathbf{S}_k \mathbf{x}_{t,k} + \mathbf{O}_k^u \mathbf{U}_k \mathbf{x}_{t,k} = \mathbf{x}_{t,k}$$

To clarify these matrices we present an example of 3-agent system. The agent-wise distribution of global state elements are obtained through their respective observation spaces.

Fig.7.2 shows the corresponding Venn diagram. The mapping of local state vector from the global one are obtained using the following matrices:

$$\mathbf{T}_1 = \begin{bmatrix} 0,1,0,0,0 \\ 0,0,1,0,0 \\ 0,0,0,0,1 \end{bmatrix}; \mathbf{T}_2 = \begin{bmatrix} 1,0,0,0,0 \\ 0,0,0,1,0 \\ 0,0,0,0,1 \end{bmatrix}; \mathbf{T}_3 = \begin{bmatrix} 0,0,1,0,0 \\ 0,0,0,1,0 \\ 0,0,0,0,1 \end{bmatrix}$$

Also, from Venn diagram, the set of neighbors, $\mathcal{S}_1 = \{2, 3\}$; $\mathcal{S}_2 = \{1, 3\}$ and $\mathcal{S}_3 = \{1, 2\}$. The corresponding \mathbf{S} , \mathbf{P} and \mathbf{L} matrices are as follows:

$$\mathbf{S}_1 = \begin{bmatrix} 0,1,0 \\ 0,0,1 \end{bmatrix}; \mathbf{S}_2 = \mathbf{S}_1; \mathbf{S}_3 = \mathbf{I}_3.$$

- Agent 2 \rightarrow Agent 1 \leftarrow Agent 3,

$$\mathbf{P}_{2,1} = \begin{bmatrix} 0,0 \\ 0,1 \end{bmatrix}; \mathbf{L}_{2,1} = \mathbf{P}_{2,1}; \mathbf{P}_{3,1} = \begin{bmatrix} 1,0,0 \\ 0,0,1 \end{bmatrix}; \mathbf{L}_{3,1} = \mathbf{I}_2.$$

- Agent 1 \rightarrow Agent 2 \leftarrow Agent 3,

$$\mathbf{P}_{1,2} = \mathbf{P}_{2,1}; \mathbf{L}_{1,2} = \mathbf{P}_{1,2}; \mathbf{P}_{3,2} = \begin{bmatrix} 0,1,0 \\ 0,0,1 \end{bmatrix}; \mathbf{L}_{3,2} = \mathbf{I}_2.$$

- Agent 1 \rightarrow Agent 3 \leftarrow Agent 2,

$$\mathbf{P}_{1,3} = \begin{bmatrix} 1,0 \\ 0,0 \\ 0,1 \end{bmatrix}; \mathbf{L}_{1,3} = \begin{bmatrix} 1,0,0 \\ 0,0,0 \\ 0,0,1 \end{bmatrix}; \mathbf{P}_{2,3} = \begin{bmatrix} 0,0 \\ 1,0 \\ 0,1 \end{bmatrix}; \mathbf{L}_{2,3} = \begin{bmatrix} 0,0,0 \\ 0,1,0 \\ 0,0,1 \end{bmatrix}.$$

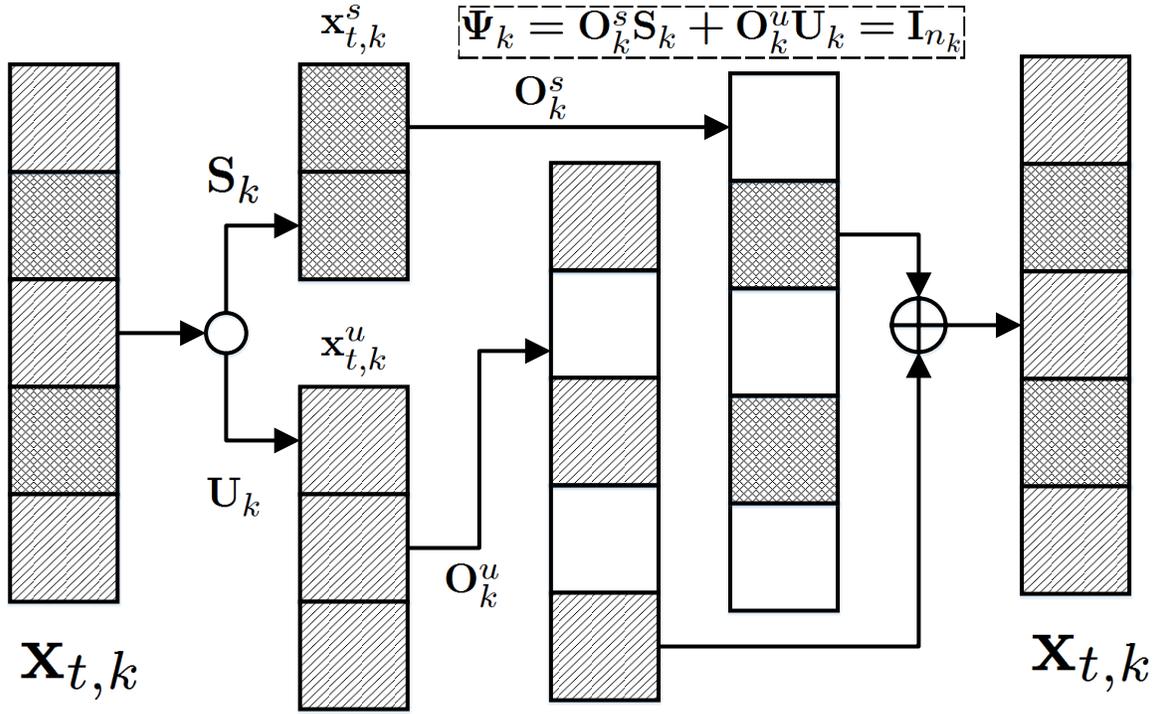


Figure 7.1: *Extracting and Reordering of Shared and Unshared State Elements.*

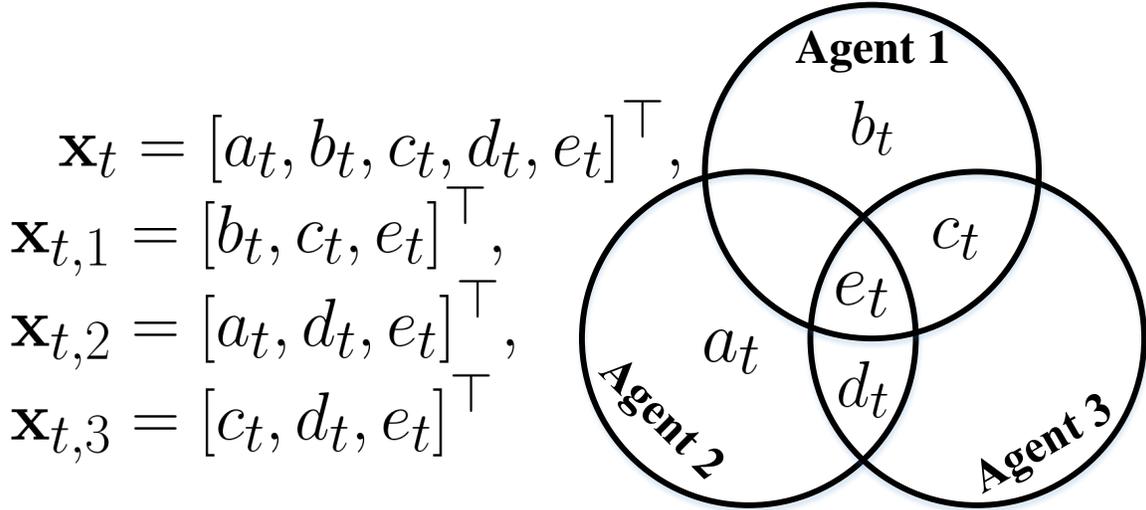


Figure 7.2: *Venn Diagram of an Example 3-Agent System.*

As can be seen from preceding definitions and example, the proposed concept of projection matrices is quite different from the well known successive orthogonalization procedure by Hassan et al. [117]. In that work, the authors talk about N interconnected linear dynamical subsystems where, the dynamics of a particular subsystem is driven by the *weighted combination of other subsystem dynamics*. This weighted aggregation also reflects the interconnected characteristics of the overall process. In addition to that, the state for any subsystem is being estimated by successively orthogonalizing the estimation error against a Hilbert space, which consists of *measurements from all the subsystems*. Accordingly, we would like to clarify the difference with the current approach in three fundamental ways. First, the single large-scale dynamical process is decomposed exclusively based on the individual agent’s measurement space. The weighted combination is absent here and the agent-wise decomposed dynamics does not depend upon that of other agents. Secondly, not all the agents are interconnected rather the neighbors are defined based on the sharing of same state elements. Third, as will be described in the upcoming AKCF algorithm, the intermediate state estimate for each agent is based on the orthogonalization procedure with respect to its own set of measurements and the final estimate is obtained through *neighborhood consensus* of these estimates. Thus, the strategy proposed in this research can efficiently handle large-scale problems while preserving data privacy.

7.3 AKCF with Adjustable Consensus Weights

The AKCF algorithm proposed in Section 6.3 is modified by incorporating the concept of shared and unshared state elements. Algorithm 3 summarizes the procedure.

At the initialization step, $\boldsymbol{\mu}_k = \mathbf{T}_k \boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_k = \mathbf{T}_k \boldsymbol{\Sigma} \mathbf{T}_k^\top$. Steps 2 to 6 are performed according to traditional Kalman filtering. The standard form calculation of Kalman gain (step 4) is feasible since each agent k has to bear $\mathcal{O}(n_k^3)$ complexity as compared to $\mathcal{O}(n^3)$ in conventional distributed Kalman filtering. At the 7th step of estimation, correction is made

Algorithm 3 AKCF

- 1: $\hat{\mathbf{x}}_{-1,k|-1} = \boldsymbol{\mu}_k, \mathbf{M}_{-1,k|-1} = \boldsymbol{\Sigma}_k$ ▷ Initialization
 - 2: $\hat{\mathbf{x}}_{t,k|t-1} = \mathbf{F}_k \hat{\mathbf{x}}_{t-1,k|t-1}$ ▷ Predict State
 - 3: $\mathbf{M}_{t,k|t-1} = \mathbf{F}_k \mathbf{M}_{t-1,k|t-1} \mathbf{F}_k^\top + \mathbf{Q}_k$ ▷ Update Error Covariance
 - 4: $\mathbf{K}_{t,k}^f = \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1}$ ▷ Kalman Gain
 - 5: $\mathbf{M}_{t,k|t} = \left(\mathbf{I}_{n_k} - \mathbf{K}_{t,k}^f \mathbf{H}_k \right) \mathbf{M}_{t,k|t-1}$ ▷ Correct Error Covariance
 - 6: $\hat{\mathbf{b}}_{t,k} = \hat{\mathbf{x}}_{t,k|t-1} + \mathbf{K}_{t,k}^f (\mathbf{y}_{t,k} - \mathbf{H}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Intermediate Correction
 - 7: $\hat{\mathbf{x}}_{t,k|t}^s = \mathbf{S}_k \hat{\mathbf{b}}_{t,k}$
 $+ \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \mathbf{S}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Shared Element Correction through Inter-Agent Information Exchange
 - 8: $\hat{\mathbf{x}}_{t,k|t} = \mathbf{O}_k^s \hat{\mathbf{x}}_{t,k|t}^s + \mathbf{O}_k^u \mathbf{U}_k \hat{\mathbf{b}}_{t,k}$ ▷ Combining the Shared and Unshared Parts
-

to individual agents' shared state elements by using $\hat{\mathbf{b}}_{t,k}$ and exchanged information from the neighbors. Finally, the estimates of shared and unshared state elements are combined in a single vector for each agent. The superscript “ f ” is used to indicate the Kalman filter gain and consensus weight for filtering. It should be noted that, the fixed consensus weight $\mathbf{M}_{t,k|t}$ used in Section 6.3 is changed to a general weight matrix $\mathbf{W}_{t,k|t}^f$. In the upcoming sections we discuss the probable choice of this weight matrix for a stable filtering operation.

7.4 Main Results

We present the primary outcome of this research starting with an important lemma, which will be used in upcoming theorems and subsequent sections.

Lemma 7.1. *Let's assume \mathbf{G} and \mathbf{L} are two symmetric positive definite matrices. The maximum and minimum eigenvalues of \mathbf{G} are g_{max} and g_{min} , respectively. Similarly, l_{max} and l_{min} denote the respective maximum and minimum eigenvalues of \mathbf{L} . Let $\epsilon \in \mathbb{R}$. The positive definiteness of the symmetric matrix quantity $(\mathbf{G} - \epsilon^2 \mathbf{L})$ is guaranteed, if $|\epsilon|$ is bounded as follows [84],*

Proof. Let, $\mathbf{B} = \mathbf{G} - \epsilon^2 \mathbf{L}$. Therefore, \mathbf{B} will be positive definite if and only if for all non zero vectors \mathbf{u} , the quadratic form, $\mathbf{u}^\top \mathbf{B} \mathbf{u} > 0$. The primary decomposition of this form

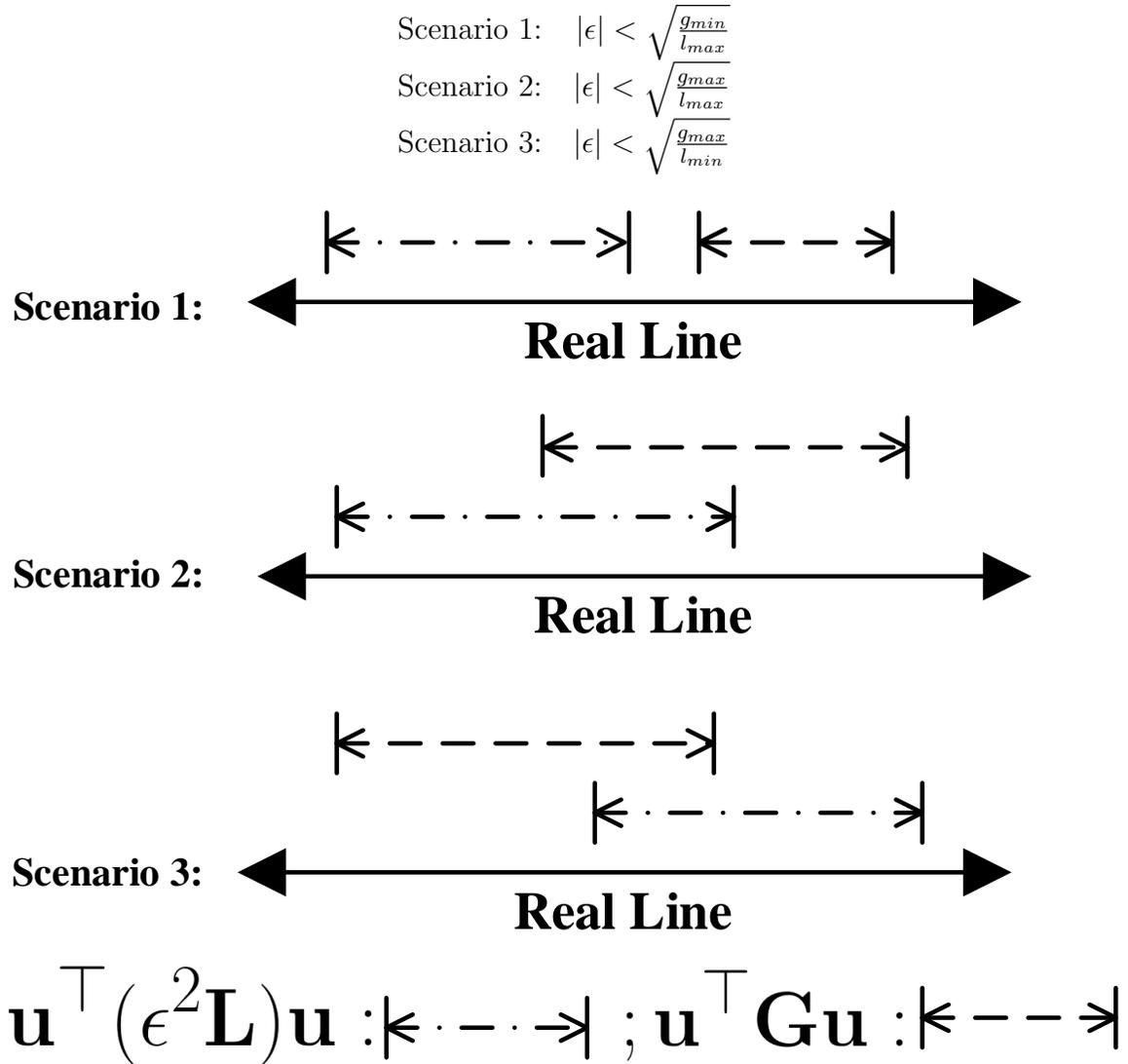


Figure 7.3: *Relative Positioning of the Range of Quadratic Values.*

reveals the required inequality, $\mathbf{u}^\top (\epsilon^2 \mathbf{L}) \mathbf{u} < \mathbf{u}^\top \mathbf{G} \mathbf{u}$. Because of the symmetric matrices, the quadratic values $\mathbf{u}^\top (\cdot) \mathbf{u}$ lie on the real line whose range is defined according to Rayleigh-Ritz inequality. Thus, for the two sides of the above equation we have the following range of real quadratic values, $g_{min} \mathbf{u}^\top \mathbf{u} \leq \mathbf{u}^\top \mathbf{G} \mathbf{u} \leq g_{max} \mathbf{u}^\top \mathbf{u}$, and $\epsilon^2 l_{min} \mathbf{u}^\top \mathbf{u} \leq \mathbf{u}^\top (\epsilon^2 \mathbf{L}) \mathbf{u} \leq \epsilon^2 l_{max} \mathbf{u}^\top \mathbf{u}$.

The relative positioning of the two ranges of quadratic values over the real line results in three possible scenarios as illustrated in Fig.7.3. These scenarios satisfy the condition of positive definiteness when,

$$\begin{aligned}
&\text{in scenario 1: } \epsilon^2 l_{max} \mathbf{u}^\top \mathbf{u} < g_{min} \mathbf{u}^\top \mathbf{u} \\
&\text{in scenario 2: } \epsilon^2 l_{max} \mathbf{u}^\top \mathbf{u} < g_{max} \mathbf{u}^\top \mathbf{u} \\
&\text{and in scenario 3: } \epsilon^2 l_{min} \mathbf{u}^\top \mathbf{u} < g_{max} \mathbf{u}^\top \mathbf{u}
\end{aligned}$$

Rearranging the inequalities and taking square roots on both sides complete the proof. \square

We also represent the sufficient condition in the following corollary of Lemma 7.1 considering the stochastic nature of symmetric positive definite matrices and following their mean behavior.

Corollary 7.1. *Let's assume $\underline{\mathbf{G}}$ and $\underline{\mathbf{L}}$ are two symmetric positive definite matrices with random elements that are independent of each other. The expectation of positive definiteness of the matrix $(\underline{\mathbf{G}} - \epsilon^2 \underline{\mathbf{L}})$ is guaranteed if ϵ is bounded as follows [85],*

$$\begin{aligned}
\text{Scenario 1: } & |\epsilon| < \sqrt{\mathbb{E}[\lambda_{min}(\underline{\mathbf{G}})]/\mathbb{E}[\lambda_{max}(\underline{\mathbf{L}})]} \\
\text{Scenario 2: } & |\epsilon| < \sqrt{\lambda_{max}(\mathbb{E}[\underline{\mathbf{G}}])/\mathbb{E}[\lambda_{max}(\underline{\mathbf{L}})]} \\
\text{Scenario 3: } & |\epsilon| < \sqrt{\lambda_{max}(\mathbb{E}[\underline{\mathbf{G}}])/\lambda_{min}(\mathbb{E}[\underline{\mathbf{L}}])}
\end{aligned}$$

Proof. The positive definiteness is guaranteed in Lemma 7.1 by expressing the bounds in terms of the eigenvalue ratios. The random nature of matrices under consideration results in random ratios of eigenvalues. Therefore, we are interested in expected eigenvalue ratios. Mathematically, our goal is to modify the general inequality $\epsilon^2 < \mathbb{E}[\lambda(\underline{\mathbf{G}})/\lambda(\underline{\mathbf{L}})]$ based on the three scenarios mentioned in the proof of Lemma 7.1. Since, the random matrices are independent, their eigenvalues (which are real and positive because of symmetry and positive definiteness) are also independent. Therefore,

$$\mathbb{E}[\lambda(\underline{\mathbf{G}})/\lambda(\underline{\mathbf{L}})] = \mathbb{E}[\lambda(\underline{\mathbf{G}})]\mathbb{E}[1/\lambda(\underline{\mathbf{L}})] > \mathbb{E}[\lambda(\underline{\mathbf{G}})]/\mathbb{E}[\lambda(\underline{\mathbf{L}})]$$

The above inequality is the direct consequence of the relationship between arithmetic and harmonic means of positive datasets. Hence, we focus on more strict inequality $\epsilon^2 <$

$\mathbb{E}[\lambda(\underline{\mathbf{G}})]/\mathbb{E}[\lambda(\underline{\mathbf{L}})]$. In addition to that, the Jensen's inequality is used to define the following properties of maximum and minimum eigenvalues of a random matrix $\underline{\mathbf{X}}$.

- Convexity of maximum eigenvalue: $\lambda_{max}(\mathbb{E}[\underline{\mathbf{X}}]) \leq \mathbb{E}[\lambda_{max}(\underline{\mathbf{X}})]$
- Concavity of minimum eigenvalue: $\lambda_{min}(\mathbb{E}[\underline{\mathbf{X}}]) \geq \mathbb{E}[\lambda_{min}(\underline{\mathbf{X}})]$

Using the above properties, we can express the sufficient bounds on ϵ^2 under the three scenarios.

$$\begin{aligned} \text{Scenario 1: } & \epsilon^2 < \mathbb{E}[\lambda_{min}(\underline{\mathbf{G}})]/\mathbb{E}[\lambda_{max}(\underline{\mathbf{L}})] \\ \text{Scenario 2: } & \epsilon^2 < \lambda_{max}(\mathbb{E}[\underline{\mathbf{G}}])/\mathbb{E}[\lambda_{max}(\underline{\mathbf{L}})] \\ \text{Scenario 3: } & \epsilon^2 < \lambda_{max}(\mathbb{E}[\underline{\mathbf{G}}])/\lambda_{min}(\mathbb{E}[\underline{\mathbf{L}}]) \end{aligned}$$

Taking square root on both sides of these inequalities complete the proof. \square

In addition to this, we define the following matrices according to AKCF steps in Algorithm 3,

Definition 7.1. $\mathcal{C}_{t,k} = (\mathbf{I}_{n_k} - \mathbf{K}_{t,k}^f \mathbf{H}_k) \mathbf{F}_k$; $\mathcal{D}_{t,k} = \mathcal{C}_{t,k}^{-1} \mathbf{M}_{t,k|t} (\mathcal{C}_{t,k}^{-1})^\top$;
 $\mathbf{G}_{t,k} = \mathbf{M}_{t-1,k|t-1}^{-1} - \mathcal{D}_{t,k}^{-1}$; and,

$$[\mathbf{A}^F]_{(k,l)^{th} \text{ block}} = \begin{cases} -\mathbf{O}_k^s \mathbf{P}_{l,k} \mathbf{S}_l \mathbf{F}_l; & l \in \mathcal{S}_k \\ \sum_{i \in \mathcal{S}_k} \mathbf{O}_k^s \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k; & l = k \\ \mathbf{0}_{n_k \times n_l}; & \text{otherwise.} \end{cases} \quad (7.2)$$

The following theorem gives an optimal expression of the consensus weight $\mathbf{W}_{t,k}^f$ in association with a tuning parameter ϵ in order to set the degree of participation in consensus.

Theorem 7.1. *The error dynamics of the Kalman consensus filter for multi-agent system described in Algorithm 3 is globally asymptotically stable if, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top \mathbf{O}_k^s$; $\forall k, \forall t$; and the sufficient conditions on the level of consensus to guarantee stability are [84],*

$$\begin{aligned}
\text{Scenario 1: } & |\epsilon| < \sqrt{\min_k \lambda_{\min}(\mathbf{G}_{t,k}) / \lambda_{\max}(\mathbf{A}^{F\top} \mathbf{D}_t \mathbf{A}^F)} \\
\text{Scenario 2: } & |\epsilon| < \sqrt{\max_k \lambda_{\max}(\mathbf{G}_{t,k}) / \lambda_{\max}(\mathbf{A}^{F\top} \mathbf{D}_t \mathbf{A}^F)} \\
\text{Scenario 3: } & |\epsilon| < \sqrt{\max_k \lambda_{\max}(\mathbf{G}_{t,k}) / \lambda_{\min}(\mathbf{A}^{F\top} \mathbf{D}_t \mathbf{A}^F)}
\end{aligned}$$

Proof. To investigate the error dynamics, an estimation error vector associated with each agent is defined in two steps. At first, error in estimating the shared states and those of unshared ones are defined. For shared states, the error is, $\boldsymbol{\eta}_{t,k}^s = \hat{\mathbf{x}}_{t,k|t}^s - \mathbf{S}_k \mathbf{x}_{t,k}$. And for the unshared elements, $\boldsymbol{\eta}_{t,k}^u = \mathbf{U}_k (\hat{\mathbf{b}}_{t,k} - \mathbf{x}_{t,k})$. Here, $\mathbf{x}_{t,k}$ is the true state vector of the k^{th} agent at discrete time instant t . At the second step, the estimation error of k^{th} agent is defined as,

$$\boldsymbol{\eta}_{t,k} = \mathbf{O}_k^s \boldsymbol{\eta}_{t,k}^s + \mathbf{O}_k^u \boldsymbol{\eta}_{t,k}^u \quad (7.3)$$

Therefore, the error dynamics for the k^{th} agent can be derived from equations (6.2), (6.5), steps 2,6,7 and 8 of Algorithm 3,

$$\begin{aligned}
\boldsymbol{\eta}_{t,k} &= \mathcal{C}_{t,k} \boldsymbol{\eta}_{t-1,k} + \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} + \mathbf{K}_{t,k} \mathbf{v}_{t,k} - (\mathbf{I}_{n_k} - \mathbf{K}_{t,k} \mathbf{H}_k) \mathbf{w}_{t-1,k} \\
&\quad - \mathbf{O}_k^s \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{w}_{t-1,i} - \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{w}_{t-1,k})
\end{aligned} \quad (7.4)$$

where, $\mathbf{z}_{t-1,k} = \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{F}_i \boldsymbol{\eta}_{t-1,i} - \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k \boldsymbol{\eta}_{t-1,k})$ and $\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{x}_{t,i} = \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{x}_{t,k}, \forall i \in \mathcal{S}_k, \forall k, \forall t$. From equation (7.4) it can be observed that, the measurement and process noise terms (i.e., \mathbf{v} and \mathbf{w}) are random inputs characterized by zero mean Gaussian probability distribution. Therefore, for k^{th} agent, the stability of equation (7.4) is governed by the following homogeneous equation [118],

$$\boldsymbol{\eta}_{t,k} = \mathcal{C}_{t,k} \boldsymbol{\eta}_{t-1,k} + \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} \quad (7.5)$$

The optimum choice for $\mathbf{W}_{t,k}^f$ is derived using the Lyapunov stability criteria [54, 118]. The

candidate Lyapunov function for the k^{th} agent, $V_k(t) = \boldsymbol{\eta}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \boldsymbol{\eta}_{t,k}$. Hence, the change in $V_k(t)$ can be written as,

$$\begin{aligned} \delta V_k &= V_k(t) - V_k(t-1) \\ &= -\boldsymbol{\eta}_{t-1,k}^\top \mathbf{G}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\boldsymbol{\eta}_{t-1,k}^\top \left(\mathcal{C}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \right) \mathbf{z}_{t-1,k} \\ &\quad + \mathbf{z}_{t-1,k}^\top \left(\mathbf{O}_k^s \mathbf{W}_{t,k}^f \right)^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} \end{aligned} \quad (7.6)$$

In the above equation, the positive definiteness of $\mathbf{G}_{t,k}$ is ensured by the recursive property of Kalman filter and it also justifies the current choice of Lyapunov function. We add a related explanation in Appendix C and it complements the use of similar Lyapunov function in [119], too. Now, second term matrix quantity of the above equation can be reduced as follows,

$$\begin{aligned} \mathcal{C}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f &= \epsilon \mathbf{O}_k^s; \epsilon \in \mathbb{R} \\ \Rightarrow \mathbf{O}_k^s \mathbf{W}_{t,k}^f &= \epsilon \mathbf{M}_{t,k|t} \left(\mathcal{C}_{t,k}^\top \right)^{-1} \mathbf{O}_k^s \end{aligned} \quad (7.7)$$

Consequently, the third term reduces to $\epsilon^2 (\mathbf{O}_k^s \mathbf{z}_{t-1,k})^\top \mathcal{D}_{t,k} (\mathbf{O}_k^s \mathbf{z}_{t-1,k})$. Since, \mathbf{O}_k^s is at least a column rank matrix (i.e., $n_k \geq n_k^s$), we can write, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t} \left(\mathcal{C}_{t,k}^\top \right)^{-1} \mathbf{O}_k^s$. Using step 5 of Algorithm 3 and Definition 7.1,

$$\begin{aligned} \mathbf{M}_{t,k|t} &= \left(\mathbf{I}_{n_k} - \mathbf{K}_{t,k}^f \mathbf{H}_k \right) \mathbf{M}_{t,k|t-1} \\ \Rightarrow \mathbf{M}_{t,k|t} &= \mathbf{M}_{t,k|t-1} \left(\mathbf{I}_{n_k} - \mathbf{K}_{t,k}^f \mathbf{H}_k \right)^\top \\ \Rightarrow \mathbf{M}_{t,k|t} &= \mathbf{M}_{t,k|t-1} \left(\mathbf{F}_k^{-1} \right)^\top \mathcal{C}_{t,k}^\top \end{aligned} \quad (7.8)$$

Therefore, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} \left(\mathbf{F}_k^{-1} \right)^\top \mathbf{O}_k^s$. With this consensus weight, we get the change in Lyapunov function for the multi-agent system by accumulating δV_k for all agents,

$\delta V = \sum_{k=1}^N \delta V_k$ [54]. Therefore,

$$\begin{aligned}
\delta V &= - \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \mathbf{G}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\epsilon \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top (\mathbf{O}_k^s \mathbf{z}_{t-1,k}) \\
&\quad + \epsilon^2 \sum_{k=1}^N (\mathbf{O}_k^s \mathbf{z}_{t-1,k})^\top \mathcal{D}_{t,k} (\mathbf{O}_k^s \mathbf{z}_{t-1,k}) \\
&= -\boldsymbol{\eta}_{t-1}^\top \mathbf{G}_t \boldsymbol{\eta}_{t-1} - 2\epsilon \boldsymbol{\eta}_{t-1}^\top \mathbf{A}^F \boldsymbol{\eta}_{t-1} \\
&\quad + \epsilon^2 \boldsymbol{\eta}_{t-1}^\top \mathbf{A}^{F^\top} \mathbf{D}_t \mathbf{A}^F \boldsymbol{\eta}_{t-1}
\end{aligned} \tag{7.9}$$

where, $\boldsymbol{\eta}_t = [\boldsymbol{\eta}_{t,1}^\top \cdots \boldsymbol{\eta}_{t,N}^\top]^\top$, $\mathbf{D}_t = \text{Blockdiag} [\mathcal{D}_{t,1} \cdots \mathcal{D}_{t,N}]$ and $\mathbf{G}_t = \text{Blockdiag} [\mathbf{G}_{t,1} \cdots \mathbf{G}_{t,N}]$. From the properties of block-diagonal matrices, $\lambda_{\min}(\mathbf{G}_t) = \min_k \lambda_{\min}(\mathbf{G}_{t,k})$. Hence, the error dynamics of the multi-agent system will be globally asymptotically stable if and only if $\delta V < 0$ i.e., the matrix $(\mathbf{G}_t - \epsilon^2 \mathbf{A}^{F^\top} \mathbf{D}_t \mathbf{A}^F)$ is positive definite. Using Lemma 7.1, this condition gives the bounds on $|\epsilon|$ in three scenarios. □

Theorem 7.1 is system specific as can be seen from the three possible upper bounds in the degree of consensus participation. In proving this theorem, we maintain consistency in the definition of Lyapunov function at each discrete time instances. However, following the analysis of [54], the respective Lyapunov functions would be $V_k(t) = \boldsymbol{\eta}_{t,k}^\top (\mathbf{M}_{t,k}^+)^{-1} \boldsymbol{\eta}_{t,k}$ and $V_k(t-1) = \boldsymbol{\eta}_{t-1,k}^\top \mathbf{M}_{t,k|t}^{-1} \boldsymbol{\eta}_{t-1,k}$. Here, the equivalent expression of $\mathbf{M}_{t,k}^+$ would be,

$$\mathbf{M}_{t,k}^+ = \mathbf{M}_{t,k|t} \left(\mathbf{M}_{t,k|t-1}^{-1} \mathbf{M}_{t+1,k|t} \mathbf{M}_{t,k|t-1}^{-1} + \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \right) \mathbf{M}_{t,k|t}$$

It should be noted that the negativity of δV can also be ensured by making a least square estimate of $\mathbf{W}_{t,k}^f$. However, the existence of such weight matrix is dependent upon the rank property of a system specific composite matrix. Furthermore, the degree of consensus cannot be adjusted for stable filtering operation. The associated derivation is given in Appendix D.

7.5 Stability under Lossy Communication

From Algorithm 3 it is observed that only the information about relevant state elements are being exchanged among neighbors at the final correction steps. Therefore, the inter-agent two way information exchange plays an important role in agent based Kalman filtering and can be hampered if the underlying communication link fails.

These circumstances can be simulated by introducing random link failures (RLF). Mathematically, effect of RLF can be analyzed by inserting Bernoulli random variables $\zeta_{i,k}(t)$ in step 7 of Algorithm 3. The associated probability mass functions are defined in equation (6.17) and assumptions 6.1,6.2 are also followed in this analysis.

With the insertion of Bernoulli random variable, the step 7 in Algorithm 1 becomes,

$$\hat{\mathbf{x}}_{t,k|t}^s = \mathbf{S}_k \hat{\mathbf{b}}_{t,k} + \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{P}_{i,k} \mathbf{S}_i \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \mathbf{S}_k \hat{\mathbf{x}}_{t,k|t-1}) \quad (7.10)$$

Thus, it is evident that only the consensus part of equation (7.10) is affected by communication network. Additionally, the introduction of Bernoulli random variable $\zeta_{i,k}(t)$ randomizes equation (7.2) in the following way,

$$[\mathbf{A}_t^F]_{(k,l)^{th} block} = \begin{cases} -\zeta_{l,k}(t) \mathbf{O}_k^s \mathbf{P}_{l,k} \mathbf{S}_l \mathbf{F}_l; & l \in \mathcal{S}_k \\ \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) \mathbf{O}_k^s \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k; & l = k \\ \mathbf{0}_{n_k \times n_l}; & otherwise. \end{cases} \quad (7.11)$$

Now, we represent a corollary of Theorem 7.1 for the optimal choice of $\mathbf{W}_{t,k}^f$ under lossy communication network.

Corollary 7.2. *If the probability of failure for any neighboring agent pair (i, k) in an N -agent system follows assumptions 1,2 and equals to ρ , the stochastic error dynamics of the associated Kalman Consensus filter is globally asymptotically stable if, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top \mathbf{O}_k^s$; $\forall k, \forall t$; and the sufficient conditions on the level of consensus to guarantee stability are [84],*

$$\begin{aligned}
\text{Scenario 1:} \quad & |\epsilon| < \sqrt{\frac{\min_k \lambda_{\min}(\mathbf{G}_{t,k})}{\mathbb{E}[\lambda_{\max}(\mathbf{A}_t^{F\top} \mathbf{D}_t \mathbf{A}_t^F)]}} \\
\text{Scenario 2:} \quad & |\epsilon| < \sqrt{\frac{\max_k \lambda_{\max}(\mathbf{G}_{t,k})}{\mathbb{E}[\lambda_{\max}(\mathbf{A}_t^{F\top} \mathbf{D}_t \mathbf{A}_t^F)]}} \\
\text{Scenario 3:} \quad & |\epsilon|(1 - \rho) < \sqrt{\frac{\max_k \lambda_{\max}(\mathbf{G}_{t,k})}{\lambda_{\min}(\mathbf{A}^{F\top} \mathbf{D}_t \mathbf{A}^F)}}
\end{aligned}$$

Proof. From equations (6.2), (6.5) and (7.10) and steps 2,6 of Algorithm 3, we express the homogeneous equation governing the error dynamics of the k^{th} agent,

$$\boldsymbol{\eta}_{t,k} = \mathcal{C}_{t,k} \boldsymbol{\eta}_{t-1,k} + \mathbf{O}_k^s \mathbf{W}_{t,k}^f \underline{\mathbf{z}}_{t-1,k} \quad (7.12)$$

where, $\underline{\mathbf{z}}_{t-1,k} = \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{F}_i \boldsymbol{\eta}_{t-1,i} - \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k \boldsymbol{\eta}_{t-1,k})$ and $\boldsymbol{\eta}_{t,k}$ is defined in equation (7.3). The stochastic nature introduced by RLF is reflected through an *underline* embedded below the \mathbf{z} vector. Furthermore, if every agent pair (i, k) within the neighborhood fails to communicate with the same probability ρ , then $\rho_{i,k} = \rho, \forall i \in \mathcal{S}_k; \forall k$. Hence, from equation (7.11), $\mathbb{E}(\mathbf{A}_t^F) = (1 - \rho) \mathbf{A}^F$. The corresponding change in k^{th} agent's stochastic Lyapunov function is given by,

$$\begin{aligned}
\delta \underline{V}_k &= \underline{V}_k(t) - \underline{V}_k(t-1) \\
&= -\boldsymbol{\eta}_{t-1,k}^\top \mathbf{G}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\boldsymbol{\eta}_{t-1,k}^\top \left(\mathcal{C}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \right) \underline{\mathbf{z}}_{t-1,k} \\
&\quad + \underline{\mathbf{z}}_{t-1,k}^\top \left(\mathbf{O}_k^s \mathbf{W}_{t,k}^f \right)^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \underline{\mathbf{z}}_{t-1,k}
\end{aligned} \quad (7.13)$$

The above equation have similar form as in equation (7.6). Hence, $\mathbf{W}_{t,k}^f$ is the same as in Theorem 7.1. Consequently, the third term reduces to $\epsilon^2 (\mathbf{O}_k^s \underline{\mathbf{z}}_{t-1,k})^\top \mathcal{D}_{t,k} (\mathbf{O}_k^s \underline{\mathbf{z}}_{t-1,k})$. It is observed that the Lyapunov function becomes stochastic because of $\underline{\mathbf{z}}_{t-1,k}$. The collected

stochastic Lyapunov dynamics for the multi-agent system, $\delta \underline{V} = \sum_{k=1}^N \delta \underline{V}_k$, or, equivalently,

$$\begin{aligned}
\delta \underline{V} &= - \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \mathbf{G}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\epsilon \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \mathbf{O}_k^s \mathbf{z}_{t-1,k} \\
&\quad + \epsilon^2 \sum_{k=1}^N (\mathbf{O}_k^s \mathbf{z}_{t-1,k})^\top \mathcal{D}_{t,k} \mathbf{O}_k^s \mathbf{z}_{t-1,k} \\
&= - \boldsymbol{\eta}_{t-1}^\top \mathbf{G}_t \boldsymbol{\eta}_{t-1} - 2\epsilon \boldsymbol{\eta}_{t-1}^\top \mathbf{A}_t^F \boldsymbol{\eta}_{t-1} + \epsilon^2 \boldsymbol{\eta}_{t-1}^\top \underline{\mathbf{L}}_t \boldsymbol{\eta}_{t-1}
\end{aligned} \tag{7.14}$$

where, $\underline{\mathbf{L}}_t = \mathbf{A}_t^F{}^\top \mathbf{D}_t \mathbf{A}_t^F$ and \mathbf{G}_t , \mathbf{D}_t are the block-diagonal matrices as defined in previous theorem. Thus, the stochastic error dynamics of the multi-agent system will be globally asymptotically stable if and only if $\delta \underline{V} < 0$, i.e., the random matrix $(\mathbf{G}_t - \epsilon^2 \underline{\mathbf{L}}_t)$ is positive definite. Using Corollary 7.1, this condition gives the bounds on $|\epsilon|$ in three scenarios. \square

Intuitively, more consideration is expected in agent-wise consensus as the inter-agent communication link continues to become more and more prone to failure. This can be observed from Corollary 7.2 through the proportional relationship between the degree of participation in consensus “ ϵ ” and the rate of link failure “ ρ ”.

7.6 Case Study

The optimal AKCF is applied to a custom 2-agent system “SYS3” as illustrative example. The set of state elements for the 1st agent is $\{a, b\}$ and that for the 2nd agent $\{b, c\}$. It is clear that the state element b is shared between the two agents and rest of the state elements are strictly local to the respective agents. The system parameters described in Sections 6.2, 7.2 and corresponding extraction, reordering and projection matrices are given in Appendix E. The parameters are so chosen that SYS3 follows Scenario 2 of Lemma 7.1. Table 7.1 lists the corresponding steady-state upper bounds of $|\epsilon|$ under perfect and lossy communication network following Theorem 7.1 and Corollary 7.2, respectively.

The stability performance of the proposed optimal AKCF is investigated by running

	ρ	$ \epsilon $ Upper bound
Perfect Network	0	0.3849
Lossy Network	0.2	0.4103
	0.4	0.4410
	0.6	0.4791
	0.8	0.5279

Table 7.1: Steady State Bounds for SYS3

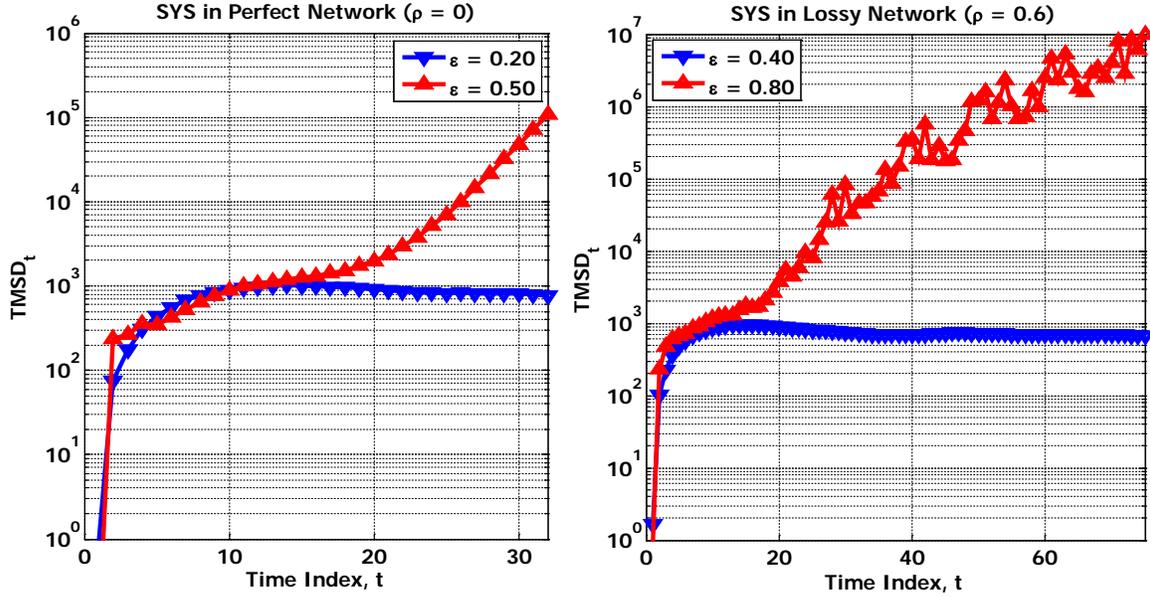


Figure 7.4: AKCF Stability for SYS3 in Perfect and Lossy Network.

1000 independent Monte Carlo trials over SYS3 at each time step. These trials simulate the total mean squared deviation, (TMSD), which has the following definition,

$$\mathbf{TMSD}_t = \sum_{k=1}^2 \mathbb{E} [\boldsymbol{\eta}_{t,k}^\top \boldsymbol{\eta}_{t,k}] \quad (7.15)$$

where, the estimation error, $\boldsymbol{\eta}_{t,k} = \hat{\mathbf{x}}_{t,k|t} - \mathbf{T}_k \mathbf{x}_t$, the true states \mathbf{x}_t being obtained according to equation (6.1). TMSD is used as an indication of AKCF stability under different degree of consensus and communication reliability. The filter is stable if TMSD values converge over time. In Fig. 7.4 this metric shows the stability performance under perfect and lossy

network. Consistent with Theorem 7.1, it is observed that the filter becomes unstable when $\epsilon > 0.3849$ under perfect communication. Additionally, as defined in Corollary 7.2, stability is preserved under lossy network as long as the the degree of participation in consensus does not exceed 0.4791 given the rate of link failure is 6 out of 10.

7.7 Conclusions and Future Work

An agent based optimally weighted Kalman consensus filter is proposed. Through Lyapunov stability analysis, the optimal degree of participation in consensus is derived under both perfect and lossy communication network. In the next chapter we investigate the impact on stability when agents are given the option to either aggregate their measurements and/or exchange information.

Chapter 8

Agent based Kalman Filtering - Measurement vs Consensus?

This chapter discusses a generalized architecture of AKCF where the agents have a choice in taking measurements and/or participating in information exchange. Within this framework, each agent is characterized under three different cases - *deterministic*, *mutual* and *independent*. Lyapunov energy functions are obtained from the corresponding estimation errors and recursive error covariance matrices. Stability analysis is performed based on the mean behavior of these functions in order to derive the bounds on the consensus level. In each case, the consensus bounds are expressed under three different scenarios of relative eigenvalue distributions. The proposed filtering method is applied to a 10-agent system with radial topology. It is observed that the filter becomes unstable in the event of violating the derived bounds confirming the theoretical results.

8.1 Introduction

In Chapter 7 we have analyzed an agent based Kalman consensus filter (AKCF). Through Lyapunov stability analysis some bounds on the degree of consensus is obtained for both

perfect and lossy inter-agent communication network. In this chapter, we further improve the optimally weighted AKCF as a generalized one, while the agents are introduced to a new flexible policy. Specifically, each agent is given the flexibility of random measurement and/or consensus. That is, agents can use either the local measurements or neighborhood information or both for estimating the states. This is a unique aspect of this research and completely different from the concept of *intermittent* measurements, which usually occurs due to faulty communication link among sensing and computing devices. This agent-level flexibility also allows the modeling of scenarios, when (1) the costs involved in sensing vs communication are significantly different; (2) there is partial/temporary malfunction of sensors that prevents its ability to take measurements but not impair its communication interface (and vice versa). Based on the usage of measurement and consensus in state estimation, three cases are defined. In the first case, which is discussed in Chapter 7, each agent uses both the measurement and consensus in state estimation. In case 2, each agent is allowed to use either measurement or consensus. Finally, in the third case, each agent's decision of measurement usage is independent from that of participation in consensus. The filter stability for case 2 and 3 are analyzed in this chapter based on the mean behavior of Lyapunov energy function. The corresponding bounds on the consensus level (in terms of eigenvalue ratios) are derived that will guarantee convergence of the estimation process. For each case of agent behavior, three different scenario are considered while expressing these bounds. These scenario are characterized based on the relative eigenvalue distribution of the underlying symmetric positive definite matrices.

8.2 Incorporation of Flexible Policy in Basic AKCF

The fundamental fileting procedure in agent based architecture is presented in Algorithm 3. It can be observed that, at every time instance, each agent is required to collect measurement (steps 5 and 6) as well as participate in consensus (step 7). We can introduce some

flexibility into the filter by enabling the agents to decide whether to take measurement and/or consensus into account. One way to achieve this is to incorporate some randomness in measurement usage and consensus participation. As a consequence, we expect a reduction in the total number of active inter-agent links as well as the amount of measurement data aggregated by each agent.

8.2.1 Random Measurement

The concept of random measurement collection can be implemented by introducing a Bernoulli random variable $\phi_k(t)$ in steps 5 and 6 of Algorithm 3,

- Minimum MSE with random measurement:

$$\underline{\mathbf{M}}_{t,k|t} = \left(\mathbf{I}_{n_k} - \phi_k(t) \mathbf{K}_{t,k}^f \mathbf{H}_k \right) \mathbf{M}_{t,k|t-1} \quad (8.1)$$

- Intermediate correction with random measurement:

$$\hat{\mathbf{b}}_{t,k} = \hat{\mathbf{x}}_{t,k|t-1} + \phi_k(t) \mathbf{K}_{t,k}^f (\mathbf{y}_{t,k} - \mathbf{H}_k \hat{\mathbf{x}}_{t,k|t-1}) \quad (8.2)$$

where, the probability mass functions are as follows,

$$\phi_k(t) = \begin{cases} 1 & \text{with Prob. } \rho_\phi \\ 0 & \text{with Prob. } 1 - \rho_\phi \end{cases} ; \forall k. \quad (8.3)$$

Let, the expected number of agents collecting measurements be m . Then for N -agent system, $\rho_\phi = m/N$. Also, $\mathbb{E}[\phi_k^2(t)] = \rho_\phi$. Intuitively, m out of N agents are expected to use their collected measurements for estimation correction and subsequent error covariance update. The inclusion of $\phi_k(t)$ makes steps 5 and 6 of Algorithm 3 stochastic in nature. Although this characteristic is similar to the case of *intermittent* measurements, there exists a fundamental difference between these two scenarios. In conventional approach of dynamic state estimation, the sensors just have to collect measurements and send this to the fusion center

or neighbors through a lossy communication channel. This results in intermittent reception of observations and manifests itself through the stochastic error covariance update and correction from random measurement [120]. It is worth to mention that very recently the effect of such lossy communication channel over measurement reception is introduced in traditional distributed Kalman consensus filter [119]. It is shown that the consensus feature performs better than a typical Kalman filter in that scenario. In [121, 8], however, the measurement intermittency is reflected through time-varying observation model in the distributed implementation of Kalman Consensus filter. In this scenario, the communication channel is assumed to follow 1st order Markov chain. On the other hand, in agent based formulation, the sensing device itself has the ability to (1) collect measurements, (2) perform computation and (3) communicate with neighbors. Consequently, the communication channel would affect the inter-agent information exchange rather than the intermittent measurement-based dynamic estimation. We address the case of lossy communication network for AKCF in Section 7.5[84]. A Bernoulli random variable is used to model the random failure of neighborhood communication links. A bound on the link failure rate is derived that ensures the filter stability.

8.2.2 Random Consensus

The random consensus is modeled by incorporating another Bernoulli random variable $\gamma_k(t)$ on the right hand side in step 7 of Algorithm 3,

$$\hat{\mathbf{x}}_{t,k|t}^s = \mathbf{S}_k \hat{\mathbf{b}}_{t,k} + \gamma_k(t) \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \mathbf{S}_k \hat{\mathbf{x}}_{t,k|t-1}) \quad (8.4)$$

The corresponding probability mass functions are,

$$\gamma_k(t) = \begin{cases} 1 & \text{with Prob. } \rho_\gamma \\ 0 & \text{with Prob. } 1 - \rho_\gamma \end{cases} ; \forall k. \quad (8.5)$$

If r agents are expected to participate in consensus, then $\rho_\gamma = r/N$, for an N -agent system. Also, $\mathbb{E}[\gamma_k^2(t)] = \rho_\gamma$. Based on these discussions, the AKCF with flexible policy is presented in Algorithm 4

Algorithm 4 Generalized AKCF

- 1: $\hat{\mathbf{x}}_{-1,k|-1} = \boldsymbol{\mu}_k, \mathbf{M}_{-1,k|-1} = \boldsymbol{\Sigma}_k$ ▷ Initialization
 - 2: $\hat{\mathbf{x}}_{t,k|t-1} = \mathbf{F}_k \hat{\mathbf{x}}_{t-1,k|t-1}$ ▷ Predict State
 - 3: $\mathbf{M}_{t,k|t-1} = \mathbf{F}_k \mathbf{M}_{t-1,k|t-1} \mathbf{F}_k^\top + \mathbf{Q}_k$ ▷ Update Error Covariance
 - 4: $\mathbf{K}_{t,k}^f = \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1}$ ▷ Kalman Gain
 - 5: $\underline{\mathbf{M}}_{t,k|t} = (\mathbf{I}_{n_k} - \phi_k(t) \mathbf{K}_{t,k}^f \mathbf{H}_k) \mathbf{M}_{t,k|t-1}$ ▷ Correct Error Covariance if $\phi_k(t) = 1$
 - 6: $\hat{\mathbf{b}}_{t,k} = \hat{\mathbf{x}}_{t,k|t-1} + \phi_k(t) \mathbf{K}_{t,k}^f (\mathbf{y}_{t,k} - \mathbf{H}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Intermediate Correction if $\phi_k(t) = 1$
 - 7: $\hat{\mathbf{x}}_{t,k|t}^s = \mathbf{S}_k \hat{\mathbf{b}}_{t,k}$
 $+ \gamma_k(t) \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \mathbf{S}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Shared Element Correction through Inter-Agent Information Exchange if $\gamma_k(t) = 1$
 - 8: $\hat{\mathbf{x}}_{t,k|t} = \mathbf{O}_k^s \hat{\mathbf{x}}_{t,k|t}^s + \mathbf{O}_k^u \mathbf{U}_k \hat{\mathbf{b}}_{t,k}$ ▷ Combining the Shared and Unshared Parts
-

Given our interest in a flexible policy for agents, we can think of three possible cases, [85]:

- Case 1: Deterministic \Rightarrow Each agent collects measurement as well as exchanges information with neighbors. There is no reduction in measurement and communication, i.e., $m = N$ and $r = N$. This case is discussed in detail in Chapter 7
- Case 2: Mutually exclusive \Rightarrow Each agent either collects measurement or exchanges information with neighbors, i.e., $\phi_k(t) + \gamma_k(t) = 1$ and it is expected that, $m + r = N$.
- Case 3: Independent \Rightarrow The events of measurement collection and information exchange are independent, i.e., $0 \leq \phi_k(t) + \gamma_k(t) \leq 2$ and consequently the inequality $0 \leq m + r \leq 2N$ is expected.

These cases constitute the generalized AKCF formulation.

8.3 AKCF Stability in Case 2 and Case 3

We investigate Case 2 and Case 3 through the stochastic nature of equations (8.1-8.4). In this regard, the following matrices are defined with an *underline* to indicate their stochasticity.

Definition 8.1. $\underline{\mathcal{C}}_{t,k} = (\mathbf{I}_{n_k} - \phi_k(t)\mathbf{K}_{t,k}\mathbf{H}_k) \mathbf{F}_k$; $\underline{\mathcal{D}}_{t,k} = \underline{\mathcal{C}}_{t,k}^{-1}\underline{\mathbf{M}}_{t,k|t} (\underline{\mathcal{C}}_{t,k}^{-1})^\top$;
 $\underline{\mathbf{G}}_{t,k} = \underline{\mathbf{M}}_{t-1,k|t-1}^{-1} - \underline{\mathcal{D}}_{t,k}^{-1}$; $\mathcal{B}_{t,k} = \mathbf{F}_k^{-1}\mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top$; $\mathcal{L}_{t,k} = \rho_\phi \mathcal{D}_{t,k} + (1 - \rho_\phi)\mathcal{B}_{t,k}$; and
 $\mathbf{J}_{t,k} = \rho_\phi(\mathbf{M}_{t-1,k|t-1}^{-1} - \mathcal{D}_{t,k}^{-1}) + (1 - \rho_\phi)(\mathbf{M}_{t-1,k|t-2}^{-1} - \mathcal{B}_{t,k}^{-1})$.

The matrix $\mathcal{D}_{t,k}$ is the same as in Definition 7.1. Since, $\mathcal{B}_{t,k}$ and $\mathcal{D}_{t,k}$ are symmetric matrices, $\mathcal{L}_{t,k}$ and $\mathbf{J}_{t,k}$ are also symmetric. Furthermore, the incorporation of error covariance matrices ensure the positive definiteness of these new definitions. A closer look reveals that incorporation of random variable $\phi_k(t)$ results in stochastic behavior of all the relevant matrices. Hence, we investigate the randomness of $\underline{\mathbf{M}}_{t,k|t}$, $\underline{\mathcal{C}}_{t,k}$ and its inverse. By definitions 7.1 and 8.1, $\underline{\mathbf{M}}_{t,k|t} = \phi_k(t)\mathbf{M}_{t,k|t} + (1 - \phi_k(t))\mathbf{M}_{t,k|t-1}$; $\forall k$; and

$$\underline{\mathcal{C}}_{t,k} = \phi_k(t)\mathcal{C}_{t,k} + (1 - \phi_k(t))\mathbf{F}_k; \underline{\mathcal{C}}_{t,k}^{-1} = \phi_k(t)\mathcal{C}_{t,k}^{-1} + (1 - \phi_k(t))\mathbf{F}_k^{-1}; \forall k$$

Therefore, $\mathbb{E}(\underline{\mathbf{M}}_{t,k|t}) = \rho_\phi\mathbf{M}_{t,k|t} + (1 - \rho_\phi)\mathbf{M}_{t,k|t-1}$; $\mathbb{E}(\underline{\mathcal{C}}_{t,k}) = \rho_\phi\mathcal{C}_{t,k} + (1 - \rho_\phi)\mathbf{F}_k$; $\mathbb{E}(\underline{\mathbf{G}}_{t,k}) = \mathbf{J}_{t,k}$; $\mathbb{E}(\underline{\mathcal{C}}_{t,k}^{-1}) = \rho_\phi\mathcal{C}_{t,k}^{-1} + (1 - \rho_\phi)\mathbf{F}_k^{-1}$; and $\mathbb{E}[\underline{\mathcal{D}}_{t,k}] = \mathcal{L}_{t,k}$.

Now, we express the optimal choice of $\mathbf{W}_{t,k}^f$ and the corresponding bounds on consensus using Corollary 7.1.

Theorem 8.1. *For an N -agent system, assume r agents are expected to participate in consensus. Then the stochastic error dynamics of the agent based Kalman consensus filter is globally asymptotically stable if, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top \mathbf{O}_k^s$; $\forall k, \forall t$; and the sufficient conditions on the level of consensus to guarantee stability are [85],*

Proof. Following the proof of Theorem 7.1, the estimation error vector associated with the k^{th} agent at discrete time instant t is, $\boldsymbol{\eta}_{t,k} = \mathbf{O}_k^s \boldsymbol{\eta}_{t,k}^s + \mathbf{O}_k^u \boldsymbol{\eta}_{t,k}^u$. Here, for shared states, the error is, $\boldsymbol{\eta}_{t,k}^s = \hat{\mathbf{x}}_{t,k|t}^s - \mathbf{S}_k \mathbf{x}_{t,k}$. And for the unshared elements, $\boldsymbol{\eta}_{t,k}^u = \mathbf{U}_k (\hat{\mathbf{b}}_{t,k} - \mathbf{x}_{t,k})$.

$$\begin{aligned}
\text{Scenario 1:} \quad & |\epsilon| < \sqrt{\frac{\mathbb{E}[\lambda_{\min}(\mathbf{G}_t)]}{\mathbb{E}[\lambda_{\max}(\mathbf{A}^{F^\top} \mathbf{D}_t \mathbf{A}^F)]}} \\
\text{Scenario 2:} \quad & |\epsilon| < \sqrt{\frac{\max_k \lambda_{\max}(\mathbf{J}_{t,k})}{\mathbb{E}[\lambda_{\max}(\mathbf{A}^{F^\top} \mathbf{D}_t \mathbf{A}^F)]}} \\
\text{Scenario 3 (Case 2):} \quad & |\epsilon| < \sqrt{\frac{N}{N-m}} \sqrt{\frac{\max_k \lambda_{\max}(\mathbf{J}_{t,k})}{\lambda_{\min}(\mathbf{A}^{F^\top} \mathbf{B}_t \mathbf{A}^F)}} \\
\text{Scenario 3 (Case 3):} \quad & |\epsilon| < \sqrt{\frac{N}{r}} \sqrt{\frac{\max_k \lambda_{\max}(\mathbf{J}_{t,k})}{\lambda_{\min}(\mathbf{A}^{F^\top} \mathbf{L}_t \mathbf{A}^F)}}
\end{aligned}$$

Therefore, the error dynamics for the k^{th} agent can be derived from equations (6.2), (6.5), (8.1-8.5) and step 2 of Algorithm 4,

$$\begin{aligned}
\boldsymbol{\eta}_{t,k} &= \underline{\mathcal{C}}_{t,k} \boldsymbol{\eta}_{t-1,k} + \gamma_k(t) \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} \\
&\quad + \phi_k(t) \mathbf{K}_{t,k} \mathbf{v}_{t,k} - (\mathbf{I}_{n_k} - \phi_k(t) \mathbf{K}_{t,k} \mathbf{H}_k) \mathbf{w}_{t-1,k} \\
&\quad - \gamma_k(t) \mathbf{O}_k^s \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{w}_{t-1,i} - \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{w}_{t-1,i})
\end{aligned}$$

where, $\mathbf{z}_{t-1,k} = \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{F}_i \boldsymbol{\eta}_{t-1,i} - \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k \boldsymbol{\eta}_{t-1,i})$ and $\mathbf{P}_{i,k} \mathbf{S}_i \mathbf{x}_{t,i} = \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{x}_{t,k}, \forall i \in \mathcal{S}_k, \forall k, \forall t$. From the above equation it can be observed that, the random noise terms (i.e., \mathbf{v} and \mathbf{w}) are zero mean Gaussian distributed. Therefore, the following homogeneous equation governs the stability of the error dynamics [118],

$$\boldsymbol{\eta}_{t,k} = \underline{\mathcal{C}}_{t,k} \boldsymbol{\eta}_{t-1,k} + \gamma_k(t) \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} \tag{8.6}$$

Here, we choose $\underline{V}_k(t) = \boldsymbol{\eta}_{t,k}^\top \underline{\mathbf{M}}_{t,k}^{-1} \boldsymbol{\eta}_{t,k}$ as the candidate Lyapunov function. The corresponding change in $\underline{V}_k(t)$ can be written as,

$$\begin{aligned}
\delta \underline{V}_k &= \underline{V}_k(t) - \underline{V}_k(t-1) \\
&= -\boldsymbol{\eta}_{t-1,k}^\top \underline{\mathbf{G}}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\gamma_k(t) \boldsymbol{\eta}_{t-1,k}^\top \left(\underline{\mathcal{C}}_{t,k}^\top \underline{\mathbf{M}}_{t,k}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \right) \mathbf{z}_{t-1,k} \\
&\quad + \gamma_k^2(t) \mathbf{z}_{t-1,k}^\top \left(\mathbf{O}_k^s \mathbf{W}_{t,k}^f \right)^\top \underline{\mathbf{M}}_{t,k}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k}
\end{aligned}$$

The second term matrix quantity of the above equation can be reduced as follows,

$$\begin{aligned}\underline{\mathcal{C}}_{t,k}^\top \underline{\mathbf{M}}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f &= \epsilon \mathbf{O}_k^s; \epsilon \in \mathbb{R} \\ \Rightarrow \mathbf{O}_k^s \mathbf{W}_{t,k}^f &= \epsilon \underline{\mathbf{M}}_{t,k|t} (\underline{\mathcal{C}}_{t,k}^\top)^{-1} \mathbf{O}_k^s\end{aligned}\quad (8.7)$$

Consequently, the third term reduces to $\epsilon^2 (\mathbf{O}_k^s \mathbf{z}_{t-1,k})^\top \underline{\mathcal{D}}_{t,k} (\mathbf{O}_k^s \mathbf{z}_{t-1,k})$. Since, \mathbf{O}_k^s is at least a column rank matrix (i.e., $n_k \geq n_k^s$), we can write, $\mathbf{W}_{t,k}^f = \epsilon (\mathbf{O}_k^s)^\dagger \underline{\mathbf{M}}_{t,k|t} (\underline{\mathcal{C}}_{t,k}^\top)^{-1} \mathbf{O}_k^s$. Now, because of the stochastic nature of $\underline{\mathbf{M}}_{t,k|t}$ and $\underline{\mathcal{C}}_{t,k}$, $\mathbf{W}_{t,k}^f$ can have the following expressions,

$$\mathbf{W}_{t,k}^f = \begin{cases} \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top \mathbf{O}_k^s; \phi_k(t) = 0 \\ \epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t} (\underline{\mathcal{C}}_{t,k}^\top)^{-1} \mathbf{O}_k^s; \phi_k(t) = 1 \end{cases}\quad (8.8)$$

It can be observed that, when $\phi_k(t) = 1$, $\mathbf{W}_{t,k}^f$ equals to $\epsilon (\mathbf{O}_k^s)^\dagger \mathbf{M}_{t,k|t-1} (\mathbf{F}_k^{-1})^\top \mathbf{O}_k^s$ according to equation (7.8). So the final expression of $\mathbf{W}_{t,k}^f$ is independent of the flexible policy considered in this research. In addition to that the Lyapunov function becomes stochastic because of $\gamma_k(t)$ and $\underline{\mathcal{C}}_{t,k}$. The collected stochastic Lyapunov dynamics for the multi-agent system, $\delta \underline{V} = \sum_{k=1}^N \delta \underline{V}_k$, or, equivalently,

$$\begin{aligned}\delta \underline{V} &= - \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \underline{\mathbf{G}}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2\epsilon \sum_{k=1}^N \gamma_k(t) \boldsymbol{\eta}_{t-1,k}^\top \mathbf{O}_k^s \mathbf{z}_{t-1,k} \\ &\quad + \epsilon^2 \sum_{k=1}^N (\mathbf{O}_k^s \mathbf{z}_{t-1,k})^\top \gamma_k^2(t) \underline{\mathcal{D}}_{t,k} \mathbf{O}_k^s \mathbf{z}_{t-1,k} \\ &= -\boldsymbol{\eta}_{t-1}^\top \underline{\mathbf{G}}_t \boldsymbol{\eta}_{t-1} - 2\epsilon \boldsymbol{\eta}_{t-1}^\top \mathbf{A}_t^F \boldsymbol{\eta}_{t-1} + \epsilon^2 \boldsymbol{\eta}_{t-1}^\top \mathbf{A}^{F\top} \underline{\mathbf{D}}_t \mathbf{A}^F \boldsymbol{\eta}_{t-1}\end{aligned}\quad (8.9)$$

where, $\underline{\mathbf{G}}_t = \text{Blockdiag} \left[\underline{\mathbf{G}}_{t,1} \quad \cdots \quad \underline{\mathbf{G}}_{t,N} \right]$; $\underline{\mathbf{D}}_t = \text{Blockdiag} \left[\gamma_1^2(t) \underline{\mathcal{D}}_{t,1} \quad \cdots \quad \gamma_N^2(t) \underline{\mathcal{D}}_{t,N} \right]$ and $[\mathbf{A}_t^F]_{(k,l)\text{th block}} = \gamma_k(t) [\mathbf{A}^F]_{(k,l)\text{th block}}$.

Thus, the stochastic error dynamics of the multi-agent system will be globally asymptotically stable if and only if $\delta \underline{V} < 0$, i.e., the random matrix $\left(\underline{\mathbf{G}}_t - \epsilon^2 \mathbf{A}^{F\top} \underline{\mathbf{D}}_t \mathbf{A}^F \right)$ is positive definite. Now, we investigate the properties of $\underline{\mathbf{G}}_t$ and $\underline{\mathbf{D}}_t$. Based on the definitions, both

these matrices are symmetric positive definite. Using the expectation operator, $\mathbb{E}[\underline{\mathbf{G}}_t] = \mathbf{J}_t$, where, $\mathbf{J}_t = \text{Blockdiag} \left[\mathbf{J}_{t,1} \quad \dots \quad \mathbf{J}_{t,N} \right]$. And,

$$\mathbb{E}[\underline{\mathbf{D}}_t] = \begin{cases} \left(\frac{N-m}{N}\right)\mathbf{B}_t & ; \text{Case 2} \\ \frac{r}{N}\mathbf{L}_t & ; \text{Case 3} \end{cases} \quad (8.10)$$

Here, $\mathbf{B}_t = \text{Blockdiag} \left[\mathbf{B}_{t,1} \quad \dots \quad \mathbf{B}_{t,N} \right]$ and $\mathbf{L}_t = \text{Blockdiag} \left[\mathcal{L}_{t,1} \quad \dots \quad \mathcal{L}_{t,N} \right]$. The Case 2 expression of $\mathbb{E}[\underline{\mathbf{D}}_t]$ is explained in Appendix F. Therefore, using the properties of block-diagonal matrices and Corollary 7.1, the bounds on $|\epsilon|$ are obtained. \square

From Theorem 8.1 it can be observed that increase of $|\epsilon|$ allows the reduction in the expected number of consensus agents as long as the product $|\epsilon|\sqrt{r}$ is bounded by the specified system parameters. Furthermore, an upper bound for ϵ is achieved when r equals $N - m$.

8.4 Simulation and Results

The generalized AKCF is applied to a 10-agent system with a radial topology. Each agent has access to two measurements. The 1st and 2nd agent's measurements are linked to 4 state elements whereas, the rest of the agents can track 6 state elements each. In the radial structure, each agent shares its state elements with two nearest leading and two nearest trailing neighbors. Additionally, no state elements are kept strictly local to any of the agent's state-space. This model resembles a radial network of power system [114]. Assuming agents follow the mutually exclusive policy of case 2, the parameters of the multi-agent system are so chosen that the upperbound on the degree of consensus (called ϵ^* hereafter) is determined from Scenario 1 of Theorem 8.1. If agents follow the independent policy of case 3, the system parameters are chosen such that Scenario 2 of Theorem 8.1 is applicable. Fig. 8.1 shows ϵ^* as a function of m and r for case 2 and case 3, respectively. In both cases, 1000 independent Monte Carlo trials are run to obtain these system properties. An increase in ϵ^* is observed

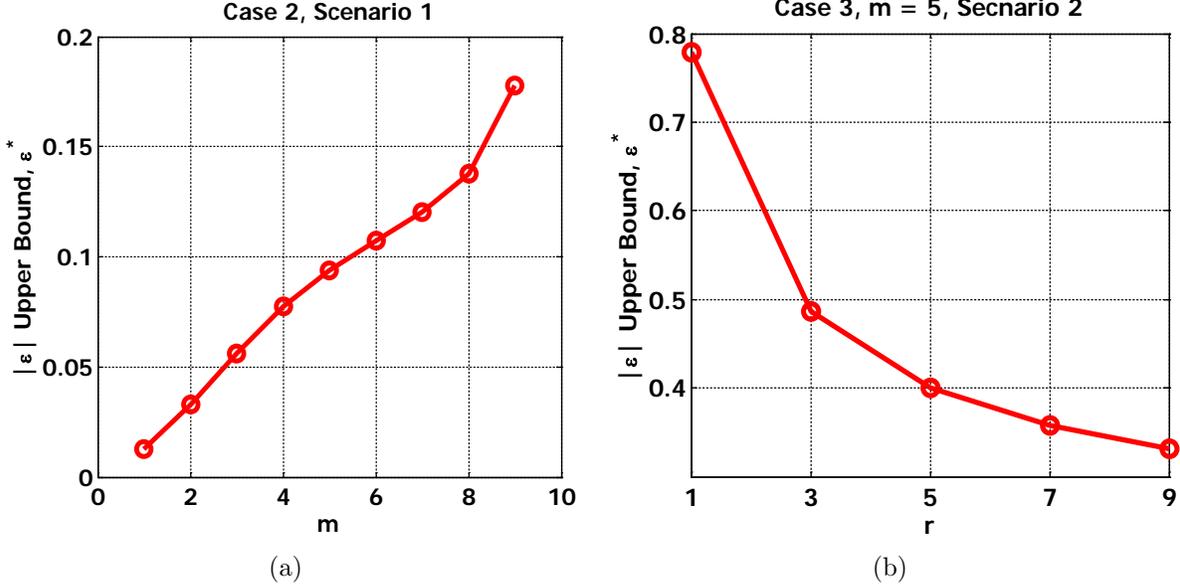


Figure 8.1: ϵ Upper bounds (ϵ^*) for 10 agent system. (a) $m = 10 - r$ agents are expected to use local measurements only and don't exchange information with neighbors. (b) $m = 5$ agents are expected to use local measurements while $1 \leq r \leq 10$ agents are expected to exchange information with neighbors.

in Fig.8.1 (a) with the increase of m (hence reduction in r). On the other hand, the upper bound decreases with r in Fig.8.1 (b) [85]. These responses indicate that as more agents tend to participate in information exchange, the degree of consensus reduces accordingly. Intuitively, such characteristics can also be observed in case 1 under the effect of lossy communication network. In this scenario, if the consensus parameter ϵ is kept fixed, the random failure of agent-to-agent links cause expected reduction in the information exchange and thus balances the neighborhood contribution to the state estimation. This realization is theorized and verified through simulation in Chapter 7 [84].

Next, we investigate the stability performance of the proposed AKCF under cases 2 and 3. The corresponding upper bounds (ϵ^*) are obtained from Fig.8.1 (a) and (b), respectively. As a metric of stability we use the total mean squared deviation, (TMSD), which has the following definition,

$$\mathbf{TMSD}_t = \sum_{k=1}^{10} \mathbb{E} [\boldsymbol{\eta}_{t,k}^\top \boldsymbol{\eta}_{t,k}] \quad (8.11)$$

where, the estimation error, $\boldsymbol{\eta}_{t,k} = \hat{\mathbf{x}}_{t,k|t} - \mathbf{T}_k \mathbf{x}_t$, the true states \mathbf{x}_t being obtained according to equation (6.1). The expectation operation is simulated by taking average over 1000 independent Monte Carlo trials at each time step. TMSD is used as an indication of AKCF stability under different degree of consensus. The filter is stable if TMSD values converge over time. Fig. 8.2 shows the stability performance when the AKCF operates in *mutual* manner with $m \in \{3, 6, 9\}$. The *independent* case is shown in Fig.8.3 with $m = 5$ and $r \in \{1, 5, 9\}$. It is evident from these figures that the filter becomes unstable whenever the consensus parameter ϵ is chosen well above the prescribed upper bound. However, the TMSD values in Fig. 8.3 (a) show a temporary overshoot while converging over time (i.e., ϵ is chosen below ϵ^*). Such behavior is not surprising as the filtering process itself is stochastic in nature through the inclusion of randomness in measurement as well as consensus based state estimation [85]. It is important to remember that, Theorem 8.1 provides the *sufficient* condition and not the necessary that guarantees convergence.

8.5 Conclusions

A generalized version of AKCF (agent based distributed Kalman consensus filter) is proposed in this chapter. The agents have the flexibility to independently or mutually decide between measurement collection and consensus participation in order to estimate the local states of a dynamical system. The underlying error dynamics are characterized and used for the Lyapunov stability analysis. Bounds on the degree of consensus is derived under three scenarios. The proposed filtering is applied to a custom-built 10 agent dynamical system. It is found that the filter becomes unstable for specific system characteristics and agent behavior when the degree of consensus exceeds the corresponding bound. In the next chapter, this concept of agent based design will be extended to distributed control, where the agents can take independent control decisions based on the local estimates. An efficient and rational arbitration strategy will also be implemented to handle the conflicting (if any)

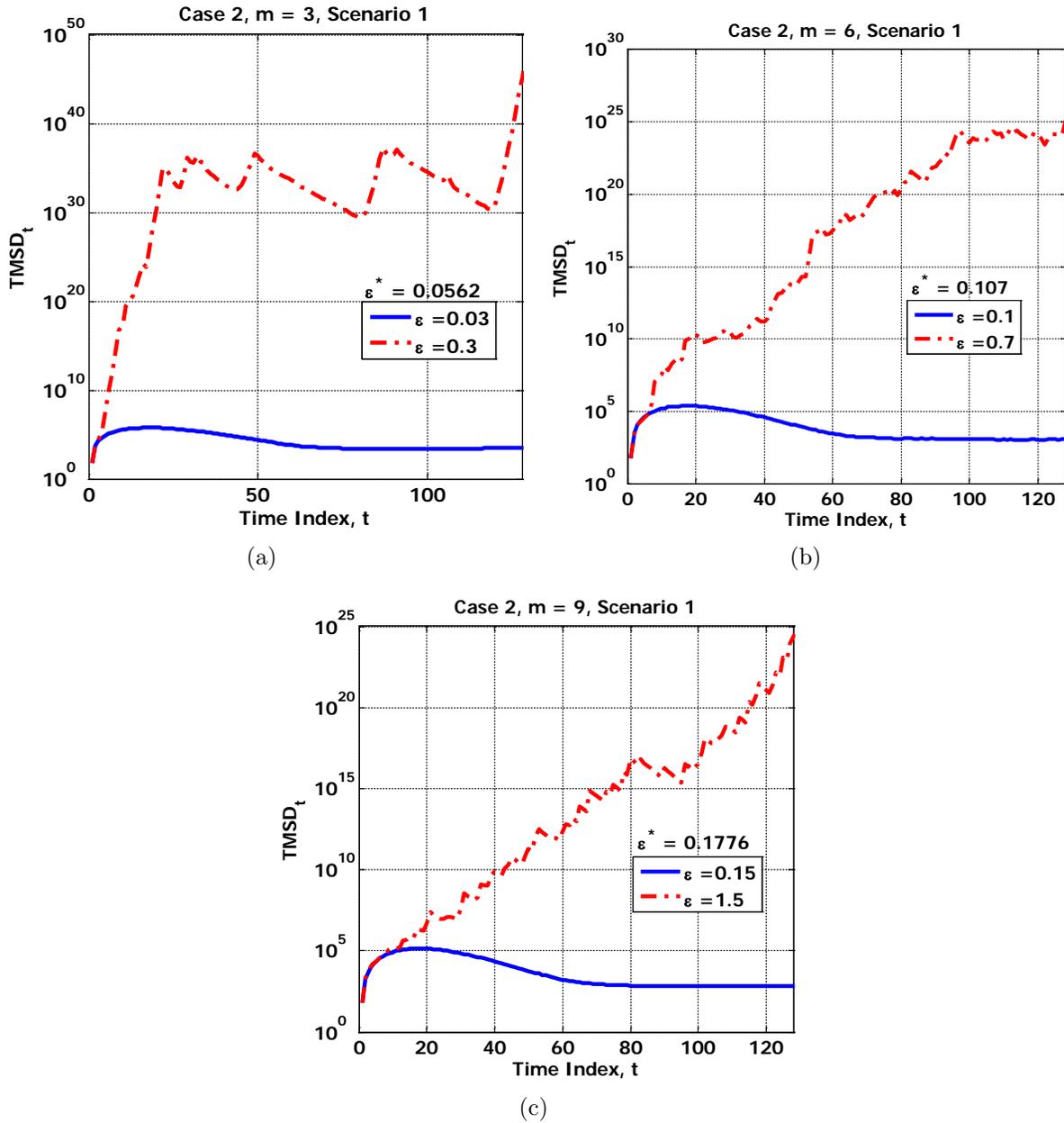


Figure 8.2: AKCF stability performance of 10 agent system under case 2 (mutual events). The upper bound ϵ^* is mentioned in respective plots.

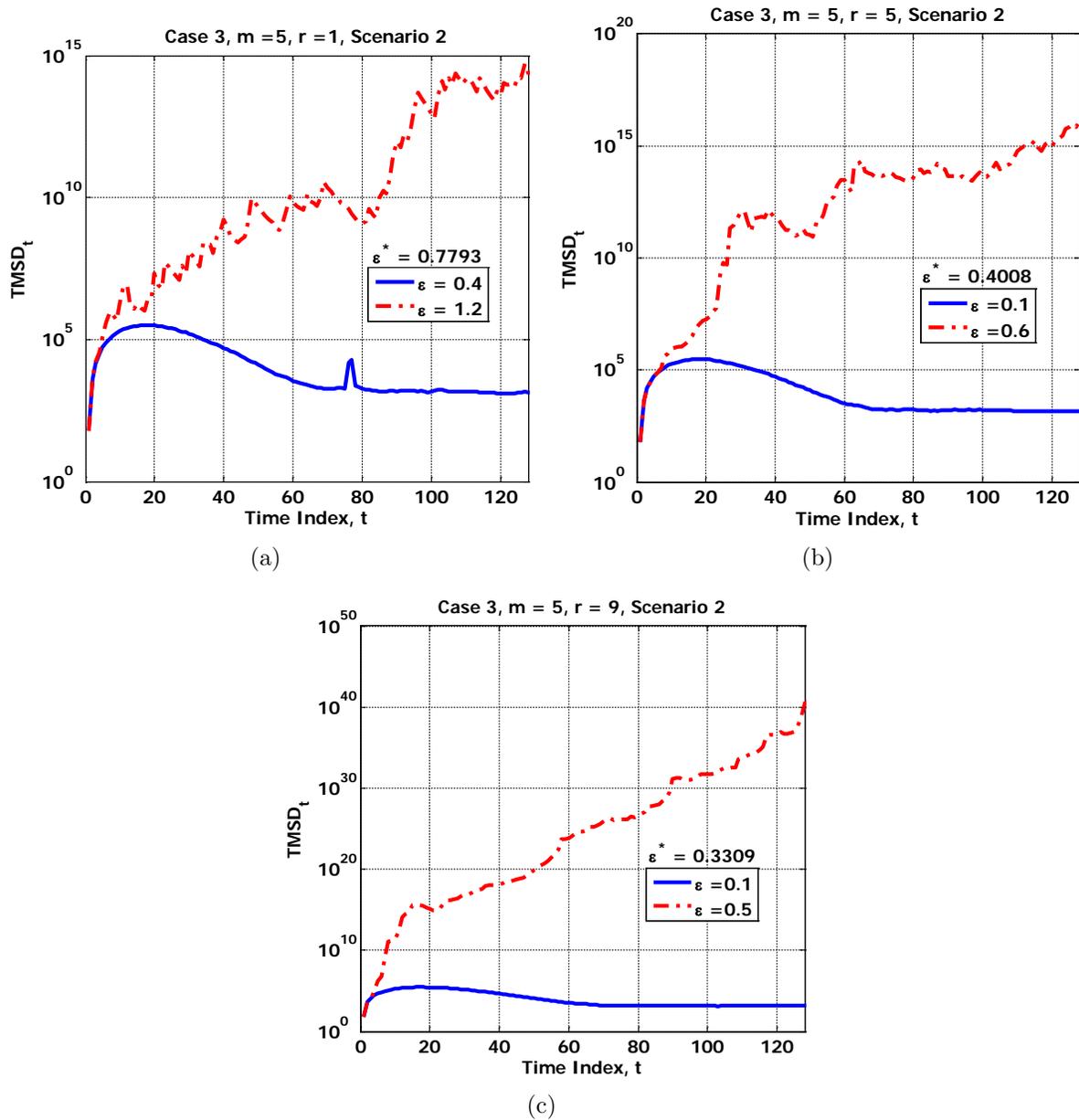


Figure 8.3: AKCF stability performance of 10 agent system under case 3 (independent events). The upper bound ϵ^* is mentioned in respective plots.

control decisions, while fulfilling agent-specific control objectives.

Chapter 9

Stability of Agent based Distributed Model Predictive Control over a Lossy Network

In this chapter, an agent based control formulation of a large scale cyber-physical system is proposed. Each agent can partially observe a part of the global dynamical process and estimate the associated local states through a combination of traditional Kalman filtering algorithm and consensus. The local estimates are then used for state feedback control. The optimal feedback gain of individual agents are obtained through dynamically solving a moving horizon linear quadratic optimization problem. The agents also exchange information among neighbors in order to design the feedback gain matrix. Finally, a control decision incurring the least cost among all agents is applied to the global system in a distributed manner. A Lyapunov function based stability analysis is performed in order to obtain a bound over the degree of agent negotiation in designing the control decision. Besides, the effect of lossy communication network in control design and henceforth in global system stability is also investigated and corresponding bound in control consensus is obtained. The theoretical results are verified through simulating a 10-agent dynamical process having a

radial topology.

9.1 Introduction

The application of model predictive control has been popular in various complex multivariate systems like energy [122], water [64], chemical processes [73], traffic networks [123] and even supply chains [124]. As the name implies, model predictive control (MPC) relies on the future prediction of state trajectories so that the set of desired control inputs can fulfill system specific objectives under the given constraints. The set of control inputs are designed for a finite length of future time instances. The control signal of only the immediate future is applied while the whole design procedure is repeated with the finite time horizon advanced by one step. Thus, MPC is advantageous than finite horizon approximated linear quadratic regulator as it offers more stability in satisfying the control objective. However, as pointed out in [63, 64], the computational complexity of MPC is the major issue that makes its application in large-scale systems questionable in terms of efficiency, robustness as well as reliability. Such scalability issue can be overcome by delegating the control responsibility among a set of controllers spatially distributed across the cyber-physical system (CPS). In this regard, two basic approaches are typically used: (1) decentralized approach, where the set of controllers independently decides specific and non-overlapped control inputs without the requirement of communication, and (2) the distributed approach, where a more reliable control decision is achieved through an iterative information exchange among the controllers over a communication network [65]. The characteristics of a large-scale system usually plays a major role in defining each controller's observation and control space in a distributed control scheme. In [66], a sparse and large scale system is decomposed into multiple subsystems. The subsystems are mapped using binary matrices *exclusively* designed for specific dynamics irrespective of observation space. In absence of any negotiation, each entry of the control input set is chosen from individual subsystem. The decomposed local

subsystem is also reported to include the effect of neighborhood control decisions. In [64], the neighboring control inputs appear as disturbance to the local controller of a radial network of water delivery canal system. Alternately, a weighted combination of neighborhood controls is considered in [67]. Here, a communication intensive repetitive procedure of control coordination is proposed. However, the typical effect of communication networks (delay, loss of packets etc.) are not considered. A successive loss of limited number of packets are considered in [63] to analyze the effect of lossy network over the closed loop stability of the global dynamical process. Consensus in control using linear quadratic Gaussian regulator is investigated in [68] and the stability criteria is derived via conditioning the spectral radius.

In recent years, distributed control is becoming popular in multi-agent operation of cyber-physical systems. Electric power transmission and distribution system can be considered a good candidate example. In this context, a graph theoretic approach is reported in [50] to maintain power balance at the prevalence of renewable energy sources. Authors in [69] use multi-objective genetic algorithm to minimize phase current unbalance in distribution network. More recently, the optimal power flow problem is distributed in [70] using the Lagrangian relaxation. A lower-upper-bound switching algorithm is designed to balance the power flow between the utility and microgrid enabled communities. In a broader perspective, the problem of optimal operation of electric power system can be addressed as synchronous optimization in a multi-objective environment. As a result, decision conflict is not uncommon in an agent based framework of such systems. In this regard, coalitional game theory is used in [71] to combine different objective functions to ensure common control decisions. A recent review of the usage of consensus strategy in distributed multi-agent optimization problem shows the focus of existing research in individual agent's convergence to the average of the initial states [72]. Optimal convergence for general linear and nonlinear systems, however is still an open problem. Also, the distributed control framework still needs to be generalized based on deeper insights on system behavior as well as the limited observation spaces of the controllers [73].

In this paper, we take a different perspective on agent based distributed control of a large CPS. We assume that each agent can observe only a part of the global dynamical process. For each agent, the global dynamics is scaled down to agent-wise local system model according to respective measurement space. Irrespective of physical placements, the agents identify neighbors based on the sharing of state elements. This enables to overcome the limitation of scalability in model predictive control. The control objective functions of individual neighboring agents won't be combined rather the agents participate in respective control decision adjustment through a consensus strategy. We investigate the agent based MPC performance under perfect and lossy communication network. Through a collective measure of agent-wise Lyapunov functions we present three sufficient conditions that will ensure global system stability under different network scenario.

9.2 System Model

We consider a linear time-invariant system whose dynamics can be modeled in discrete time as 1st order Gauss-Markov Process, i.e.,

$$\mathbf{x}_{t+1} = \mathbf{F}\mathbf{x}_t + \mathbf{G}\mathbf{u}_t + \mathbf{w}_t; t = 0, 1, 2, \dots \quad (9.1)$$

where, the overall system state is represented by the n -dimensional state vector \mathbf{x}_t at time instant t . The initial values of the state vector elements at $t = 0$ follow Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Unlike [66, 59, 60, 61, 62], the state transition matrix $\mathbf{F} \in \mathbb{R}^{n \times n}$ is a general square matrix with eigenvalues lying within a unit circle. The process noise is $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$; the control input is $\mathbf{u}_t \in \mathbb{R}^p$, and $\mathbf{G} \in \mathbb{R}^{n \times p}$. The optimal control inputs for such stochastic process are obtained through the solution of a finite moving horizon linear quadratic optimization problem. The associated cost function at any time

instance t is,

$$J(\mathcal{U}_t) = \mathbb{E} \left[\sum_{\tau=t}^{t+t_f-1} (\mathbf{x}_\tau^\top \mathbb{X} \mathbf{x}_\tau + \mathbf{u}_\tau^\top \mathbb{U} \mathbf{u}_\tau) + \mathbf{x}_{t+t_f}^\top \mathbb{X} \mathbf{x}_{t+t_f} \right] \quad (9.2)$$

Here, $\mathcal{U}_t = \{\mathbf{u}_t, \dots, \mathbf{u}_{t+t_f-1}\}$. $\mathbb{X} \succeq 0$, $\mathbb{U} \succ 0$ and t_f defines the length of moving horizon.

Hence, the desired set of control inputs between t and $t + t_f - 1$ are,

$$\{\mathbf{u}_t^*, \dots, \mathbf{u}_{t+t_f-1}^*\} = \arg \min J(\mathcal{U}_t)$$

subject to

$$\mathbf{x}_{t+1} = \mathbf{F} \mathbf{x}_t + \mathbf{G} \mathbf{u}_t + \mathbf{w}_t$$

With the help of dynamic programming, one can obtain the control inputs in state feedback form. A model-predictive control is achieved when only the first entry \mathbf{u}_t^* at the time instance t is applied. The global states at $(t + 1)$ is obtained from the model (9.1) and the finite time horizon is shifted one time step forward. The linear quadratic optimization is then repeated. To achieve stochastic control through state feedback, estimate of system states has to be obtained from some set of observations. We introduce the multi-agent architecture by letting N agents to observe the entire system. The observation of the k^{th} agent follows the additive noise model of equation (6.2), where the noise is Gaussian with zero mean and covariance \mathbf{R}_k . This model reflects individual agent's attempt to estimate only the local state vector $\mathbf{x}_{t,k}$ instead of \mathbf{x}_t . Consequently, an agent-specific local system dynamics is associated with the filtering problem. In this regard, we utilize the binary projection matrix of equation (6.3), which projects the n -dimensional state vector onto n_k -dimension. Here, by projection, we imply the extraction of one state element at a time instead of creating a new one from some weighted combinations. Such binary matrices are first reported in [66] to define multiple subsystems for decentralized model predictive control. Unlike the proposed approach, the decomposition in [66] is performed solely based on the sparse or block diagonal characteristics of the state transition matrix. Nevertheless, the projection matrix also follows Lemma 6.1.

Using the projection matrix \mathbf{T}_k , the system dynamics of k^{th} agent can be mapped from equation (9.1),

$$\mathbf{x}_{t+1,k} = \mathbf{T}_k \mathbf{F} \mathbf{x}_t + \mathbf{G}_k \mathbf{u}_{t,k} + \mathbf{w}_{t,k} \quad (9.3)$$

where, $\mathbf{G}_k = \mathbf{T}_k \mathbf{G}$ and $\mathbf{w}_{t,k} = \mathbf{T}_k \mathbf{w}_t$. Therefore, $\mathbf{w}_{t,k} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ with $\mathbf{Q}_k = \mathbf{T}_k \mathbf{Q} \mathbf{T}_k^\top$. It should be noted that, the local system dynamics for k^{th} agent should reflect the corresponding observation model of equation (6.2). As a consequence, the desired 1st order Gauss-Markov process of k^{th} agent can be expressed as,

$$\mathbf{x}_{t+1,k} = \mathbf{F}_k \mathbf{x}_{t,k} + \mathbf{G}_k \mathbf{u}_{t,k} + \mathbf{w}_{t,k}, \quad (9.4)$$

The above model conforms with equation (9.3) if and only if $\mathbf{F}_k \mathbf{T}_k = \mathbf{T}_k \mathbf{F}$. Or, using Lemma 6.1, $\mathbf{F}_k = \mathbf{T}_k \mathbf{F} \mathbf{T}_k^\top$. It is important to note that, this irreversible mapping is intended only to design the agent based model predictive control (AMPC) (Algorithm 6), and subsequent derivation of theoretical bounds for system stability. Later in Section 9.5, a case study is carried out to verify these bounds, where the system stability is illustrated through a quadratic value calculated from global state vector \mathbf{x}_t .

The static set of *physical* neighbors for the k^{th} agent is defined according to equation (7.1) and follows the subsequent arguments of Section 7.2.

In multi-agent system, every agent has four major responsibilities, (1) sensing, (2) computing (3) networking and (4) controlling. At any time instant, the desired control signal is goal specific and depends upon the current system dynamics [122]. As can be seen from equation (9.4), each agent is responsible for making the same set of control decisions. As a consequence, agents will have to negotiate their control decisions among neighbors, while reaching agent-specific goal as close as possible. Bearing this in mind, we define the following

distributed optimization problem for the k^{th} agent,

$$\begin{aligned} \{\mathbf{u}_{t,k}^*, \dots, \mathbf{u}_{t+t_f-1,k}^*\} &= \arg \min J(\mathcal{U}_{t,k}) \\ \text{subject to} \\ \mathbf{x}_{t+1,k} &= \mathbf{F}_k \mathbf{x}_{t,k} + \mathbf{G}_k \mathbf{u}_{t,k} + \mathbf{w}_{t,k} \end{aligned}$$

here, $\mathcal{U}_{t,k} = \{\mathbf{u}_{t,k}, \dots, \mathbf{u}_{t+t_f-1,k}\}$; $\mathbb{X}_k = \mathbf{T}_k \mathbb{X} \mathbf{T}_k^\top$. And,

$$\begin{aligned} J(\mathcal{U}_{t,k}) &= \mathbb{E} \left[\sum_{\tau=t}^{t+t_f-1} (\mathbf{x}_{\tau,k}^\top \mathbb{X}_k \mathbf{x}_{\tau,k} + \mathbf{u}_{\tau,k}^\top \mathbf{U} \mathbf{u}_{\tau,k}) \right] \\ &+ \mathbb{E} \left[\mathbf{x}_{t+t_f,k}^\top \mathbb{X}_k \mathbf{x}_{t+t_f,k} \right] \end{aligned} \quad (9.5)$$

From a model predictive control perspective, the first element of the set $\mathcal{U}_{t,k}$, $\mathbf{u}_{t,k}^*$ represents the control decision of the k^{th} agent at time t . It can be observed that the agents are independent in control decisions rather than cooperative [125] as long as a conflict of interest does not occur. In this scenario, the fundamental question arise about agents' behavior related to conflicting (if any) control decisions. Our endeavor is to model a consensus strategy in agent-wise control decisions.

As a first step, the agents solve their individual linear quadratic optimization problem in order to obtain the state feedback gain. Algorithm 5 depicts calculation of the corresponding control gain matrix. The superscript “c” is used to denote control gain. The algorithm is

Algorithm 5 State Feedback Gain

- | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| 1: $\tau = t + t_f : -1 : t + 1$ | ▷ Define Time Horizon |
| 2: $\mathbf{\Upsilon}_{\tau,k} = \mathbb{X}_k$ | ▷ Initialization |
| 3: $\mathbf{K}_{\tau-1,k}^c = -(\mathbf{U} + \mathbf{G}_k^\top \mathbf{\Upsilon}_{\tau,k} \mathbf{G}_k)^{-1} \mathbf{G}_k^\top \mathbf{\Upsilon}_{\tau,k} \mathbf{F}_k$ | ▷ Control Gain |
| 4: $\mathbf{\Upsilon}_{\tau-1,k} = \mathbf{F}_k^\top \mathbf{\Upsilon}_{\tau,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{\tau-1,k}^c) + \mathbb{X}_k$ | ▷ Update |
-

based on the following Hamilton-Jacobi inequality:

$$\mathbf{\Upsilon}_{\tau-1,k} - \mathbf{F}_k^\top \mathbf{\Upsilon}_{\tau,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{\tau-1,k}^c) \geq \mathbb{X}_k \quad (9.6)$$

Each agent uses the control gain $\mathbf{K}_{t,k}^c$ to design a state feedback control. The states are being estimated using the agent based Kalman consensus filter embedded with a flexible policy of measurement and/or consensus [85]. When the agents are aware about the local states through this MMSE criterion, they make a second phase of neighborhood information exchange to make agent based control decisions. Individual control decisions are then sent back to the original system and control input with the lowest cost (i.e., $\mathbf{u}_{t,k}^\top \mathbf{U} \mathbf{u}_{t,k}$) is implemented. These whole procedure is represented as Algorithm 6.

Algorithm 6 AMPC

- 1: $\hat{\mathbf{x}}_{0,k|0} = \boldsymbol{\mu}_k, \mathbf{M}_{0,k|0} = \boldsymbol{\Sigma}_k$ ▷ Initialization
 - 2: $\hat{\mathbf{x}}_{t,k|t-1} = \mathbf{F}_k \hat{\mathbf{x}}_{t-1,k|t-1} + \mathbf{G}_k \mathbf{u}_{t-1,k}$ ▷ Prediction
 - 3: $\mathbf{M}_{t,k|t-1} = \mathbf{F}_k \mathbf{M}_{t-1,k|t-1} \mathbf{F}_k^\top + \mathbf{Q}_k$ ▷ Predict Error Covariance
 - 4: $\mathbf{K}_{t,k}^f = \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{M}_{t,k|t-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1}$ ▷ Kalman Gain
 - 5: $\underline{\mathbf{M}}_{t,k|t} = (\mathbf{I}_{n_k} - \phi_k(t) \mathbf{K}_{t,k}^f \mathbf{H}_k) \mathbf{M}_{t,k|t-1}$ ▷ Correct Error Covariance if $\phi_k(t) = 1$
 - 6: $\hat{\mathbf{b}}_{t,k} = \hat{\mathbf{x}}_{t,k|t-1} + \phi_k(t) \mathbf{K}_{t,k}^f (\mathbf{y}_{t,k} - \mathbf{H}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Intermediate Correction if $\phi_k(t) = 1$
 - 7: $\hat{\mathbf{x}}_{t,k|t}^s = \mathbf{S}_k \hat{\mathbf{b}}_{t,k}$
 $+ \gamma_k(t) \mathbf{W}_{t,k}^f \sum_{i \in \mathcal{S}_k} (\mathbf{P}_{i,k} \mathbf{S}_i \hat{\mathbf{x}}_{t,i|t-1} - \mathbf{L}_{i,k} \mathbf{S}_k \hat{\mathbf{x}}_{t,k|t-1})$ ▷ Shared Element Correction through Inter-Agent Information Exchange if $\gamma_k(t) = 1$
 - 8: $\hat{\mathbf{x}}_{t,k|t} = \mathbf{O}_k^s \hat{\mathbf{x}}_{t,k|t}^s + \mathbf{O}_k^u \mathbf{U}_k \hat{\mathbf{b}}_{t,k}$ ▷ Combining the Shared and Unshared Parts
 - 9: $\mathbf{u}_{t,k} = \mathbf{K}_{t,k}^c \hat{\mathbf{x}}_{t,k|t} + \mathbf{W}_{t,k}^c \sum_{i \in \mathcal{S}_k} (\mathbf{K}_{t,i}^c \hat{\mathbf{x}}_{t,i|t} - \mathbf{K}_{t,k}^c \hat{\mathbf{x}}_{t,k|t})$ ▷ Consensus in Control
 - 10: $\mathbf{u}_t = \arg \min_{\mathbf{u}_{t,k}, k=1, \dots, N} (\mathbf{u}_{t,k}^\top \mathbf{U} \mathbf{u}_{t,k})$ ▷ Desired Control for Global System
-

In the above algorithm, the flexible policies are modeled through Bernoulli random variable $\phi_k(t)$ and $\gamma_k(t)$. In Chapter 8 they have been introduced in equations 8.3 and 8.5, respectively. The state estimation consensus weights (i.e., $\mathbf{W}_{t,k}^f$) are chosen according to Theorem 8.1 [85]. How does the choice of $\mathbf{W}_{t,k}^c$ then impact the stability of our feedback control system? The results in the next section help address this fundamental question.

9.3 Main Results

We present the primary outcome of this research starting with the definitions of following matrices [86]. These definitions are based on the AMPC steps in Algorithm 6.

Definition 9.1. $\mathbf{B}_{t,k} = \Upsilon_{t,k} - (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \Upsilon_{t+1,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)$; $\mathbf{C}_t = \Xi_t^{-1} \Upsilon_{t+1}^{-1} (\Xi_t^{-1})^\top$;
 $\mathbb{B}_t = \text{Blockdiag}[\mathbf{B}_{t,k}]$; and,

$$[\mathbb{K}_t]_{(k,l)^{th} \text{ block}} = \begin{cases} -(\mathbf{K}_{t,k}^c)^\top \mathbf{K}_{t,l}^c; & l \in \mathcal{S}_k \\ |\mathcal{S}_k| (\mathbf{K}_{t,k}^c)^\top \mathbf{K}_{t,k}^c; & l = k \\ \mathbf{0}_{n_k \times n_l}; & \text{otherwise.} \end{cases} \quad (9.7)$$

Using these definitions, we state Theorem 9.1 [86], which gives an optimal expression of the consensus weight $\mathbf{W}_{t,k}^c$ in association with a tuning parameter ν in order to set the degree of participation in control consensus.

Theorem 9.1. *The closed loop system dynamics obtained through the state feedback control for multi-agent system described in Algorithm 6 is globally asymptotically stable if, $\mathbf{W}_{t,k}^c = \nu \mathbf{G}_k^\dagger \Upsilon_{t+1,k}^{-1} [\mathbf{F}_k^\top + (\mathbf{K}_{t,k}^c)^\top \mathbf{G}_k^\top]^{-1} (\mathbf{K}_{t,k}^c)^\top$; $\forall k, \forall t$; and the sufficient conditions on the level of consensus to guarantee stability are,*

$$\begin{aligned} \text{Scenario 1:} \quad & |\nu| < \sqrt{\min_k \lambda_{\min}(\mathbf{B}_{t,k}) / \lambda_{\max}(\mathbf{K}_t^\top \mathbf{C}_t \mathbf{K}_t)} \\ \text{Scenario 2:} \quad & |\nu| < \sqrt{\max_k \lambda_{\max}(\mathbf{B}_{t,k}) / \lambda_{\max}(\mathbf{K}_t^\top \mathbf{C}_t \mathbf{K}_t)} \\ \text{Scenario 3:} \quad & |\nu| < \sqrt{\max_k \lambda_{\max}(\mathbf{B}_{t,k}) / \lambda_{\min}(\mathbf{K}_t^\top \mathbf{C}_t \mathbf{K}_t)} \end{aligned}$$

Proof. Following the proof of Theorem 7.1, the estimation error vector associated with the k^{th} agent at discrete time instant t is, $\boldsymbol{\eta}_{t,k} = \mathbf{O}_k^s \boldsymbol{\eta}_{t,k}^s + \mathbf{O}_k^u \boldsymbol{\eta}_{t,k}^u$. Here, for shared states, the error is, $\boldsymbol{\eta}_{t,k}^s = \hat{\mathbf{x}}_{t,k|t}^s - \mathbf{S}_k \mathbf{x}_{t,k}$. And for the unshared elements, $\boldsymbol{\eta}_{t,k}^u = \mathbf{U}_k (\hat{\mathbf{b}}_{t,k} - \mathbf{x}_{t,k})$. Therefore, the closed loop dynamics for the k^{th} agent can be derived from this definition of estimation error, equation (9.4) and step 8 of Algorithm 6,

$$\begin{aligned}
\mathbf{x}_{t+1,k} &= (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) \mathbf{x}_{t,k} + \mathbf{G}_k \mathbf{K}_{t,k}^c \boldsymbol{\eta}_{t,k} + \mathbf{w}_{t,k} \\
&\quad + \mathbf{G}_k \mathbf{W}_{t,k}^c \sum_{i \in \mathcal{S}_k} (\mathbf{K}_{t,i}^c \mathbf{x}_{t,i} - \mathbf{K}_{t,k}^c \mathbf{x}_{t,k}) \\
&\quad + \mathbf{G}_k \mathbf{W}_{t,k}^c \sum_{i \in \mathcal{S}_k} (\mathbf{K}_{t,i}^c \boldsymbol{\eta}_{t,i} - \mathbf{K}_{t,k}^c \boldsymbol{\eta}_{t,k})
\end{aligned} \tag{9.8}$$

From the above equation it can be observed that, the estimation error terms are not recursive in nature. Therefore, for k^{th} agent, the stability of the corresponding closed loop state dynamics is governed by the following homogeneous equation [118],

$$\mathbf{x}_{t+1,k} = (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) \mathbf{x}_{t,k} + \mathbf{G}_k \mathbf{W}_{t,k}^c \boldsymbol{\xi}_{t,k} \tag{9.9}$$

where, $\boldsymbol{\xi}_{t,k} = \sum_{i \in \mathcal{S}_k} (\mathbf{K}_{t,i}^c \mathbf{x}_{t,i} - \mathbf{K}_{t,k}^c \mathbf{x}_{t,k})$. Consider, the candidate Lyapunov function, $V_k(t) = \mathbf{x}_{t,k}^\top \boldsymbol{\Upsilon}_{t,k} \mathbf{x}_{t,k}$. The change in Lyapunov function,

$$\begin{aligned}
\delta V_k &= V_k(t+1) - V_k(t) \\
&= -\mathbf{x}_{t,k}^\top \mathbf{B}_{t,k} \mathbf{x}_{t,k} \\
&\quad + 2\mathbf{x}_{t,k}^\top (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \boldsymbol{\xi}_{t,k} \\
&\quad + \boldsymbol{\xi}_{t,k}^\top (\mathbf{G}_k \mathbf{W}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \boldsymbol{\xi}_{t,k}
\end{aligned} \tag{9.10}$$

In the above equation, the positive definiteness of $\mathbf{B}_{t,k}$ is ensured by the Hamilton-Jacobi inequality and it also justifies the current choice of Lyapunov function. We add a related explanation in Appendix G. From the second term of the above equation, $\mathbf{W}_{t,k}^c$ can be chosen such that, $(\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c = \nu (\mathbf{K}_{t,k}^c)^\top$; $\nu \in \mathbb{R}$. That is, $\mathbf{W}_{t,k}^c = \nu \mathbf{G}_k^\dagger \boldsymbol{\Upsilon}_{t+1,k}^{-1} [\mathbf{F}_k^\top + (\mathbf{K}_{t,k}^c)^\top \mathbf{G}_k^\top]^{-1} (\mathbf{K}_{t,k}^c)^\top$. Here, the superscript ‘‘c’’ is used to denote the weight matrix for control. We get the change in Lyapunov function for the multi-agent system by accumulating δV_k for all agents, $\delta V = \sum_{k=1}^N \delta V_k$. Therefore,

$$\delta V = -\mathbf{a}_t^\top \mathbf{B}_t \mathbf{a}_t + \nu^2 \mathbf{a}_t^\top \mathbf{K}_t^\top \mathbf{C}_t \mathbf{K}_t \mathbf{a}_t - 2\nu \mathbf{a}_t^\top \mathbf{K}_t \mathbf{a}_t \tag{9.11}$$

where, $\mathbf{a}_t = [\mathbf{x}_{t,1}^\top \cdots \mathbf{x}_{t,N}^\top]^\top$. Therefore, the closed loop system dynamics of the multi-agent system will be globally asymptotically stable if and only if $\delta V < 0$ i.e., the matrix $(\mathbf{B}_t - \nu^2 \mathbf{K}_t^\top \mathbf{C}_t \mathbf{K}_t)$ is positive definite. Using Lemma 7.1, this condition gives the bounds on $|\nu|$ in three scenarios. \square

It should be noted that the negativity of δV can also be ensured by making a least square estimate of $\mathbf{W}_{t,k}^c$. However, the existence of such weight matrix is dependent upon the rank property of a system specific composite matrix. Furthermore, the degree of consensus cannot be adjusted for stable state feedback control. The associated derivation is given in Appendix H.

9.4 Effect of Lossy Network on Agent based control

In the proposed method of agent based control, it is evident from Algorithm 6 that the agents exchange control information only among neighbors. Therefore, the inter-agent two way information exchange plays an important role in agent based model predictive control and can be hampered if the underlying communication link fails. These circumstances can be simulated by introducing random link failures (RLF) at every link of inter-agent communication. Mathematically, effect of RLF can be analyzed by inserting Bernoulli random variables $\zeta_{i,k}(t)$ in step 9 of Algorithm 6. The associated probability mass functions are defined in equation (6.17) and assumptions 6.1,6.2 are also followed in this analysis. Consequently, step 9 in Algorithm 6 becomes,

$$\mathbf{u}_{t,k} = \mathbf{K}_{t,k}^c \hat{\mathbf{x}}_{t,k|t} + \mathbf{W}_{t,k}^c \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{K}_{t,i}^c \hat{\mathbf{x}}_{t,i|t} - \mathbf{K}_{t,k}^c \hat{\mathbf{x}}_{t,k|t}) \quad (9.12)$$

Thus, it is evident that only the consensus part of equation (9.12) is affected by communication network. Additionally, the introduction of Bernoulli random variable $\zeta_{i,k}(t)$

randomizes equation (9.7) in the following way,

$$[\underline{\mathbf{K}}_t]_{(k,l)^{th}block} = \begin{cases} -\zeta_{l,k}(t) (\mathbf{K}_{t,k}^c)^\top \mathbf{K}_{t,l}^c; & l \in \mathcal{S}_k \\ \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{K}_{t,k}^c)^\top \mathbf{K}_{t,i}^c; & l = k \\ \mathbf{0}_{n_k \times n_l}; & otherwise. \end{cases} \quad (9.13)$$

Now, we represent a corollary of Theorem 9.1 for the optimal choice of $\mathbf{W}_{t,k}^c$ under lossy communication network.

Corollary 9.1. *If the probability of failure for any neighboring agent pair (i, k) in an N -agent system follows assumptions 1,2 and equals to ρ , the stochastic closed loop dynamics of the multi-agent system is globally asymptotically stable if,*

$\mathbf{W}_{t,k}^c = \nu \mathbf{G}_k^\dagger \Upsilon_{t+1,k}^{-1} [\mathbf{F}_k^\top + (\mathbf{K}_{t,k}^c)^\top \mathbf{G}_k^\top]^{-1} (\mathbf{K}_{t,k}^c)^\top; \forall k, \forall t;$ and the sufficient conditions on the level of consensus to guarantee stability are,

$$\begin{aligned} \text{Scenario 1:} \quad & |\nu| < \sqrt{\min_k \lambda_{\min}(\mathbf{B}_{t,k}) / \mathbb{E} [\lambda_{\max}(\underline{\mathbf{K}}_t^\top \mathbf{C}_t \underline{\mathbf{K}}_t)]} \\ \text{Scenario 2:} \quad & |\nu| < \sqrt{\max_k \lambda_{\max}(\mathbf{B}_{t,k}) / \mathbb{E} [\lambda_{\max}(\underline{\mathbf{K}}_t^\top \mathbf{C}_t \underline{\mathbf{K}}_t)]} \\ \text{Scenario 3:} \quad & |\nu|(1 - \rho) < \sqrt{\max_k \lambda_{\max}(\mathbf{B}_{t,k}) / \lambda_{\min}(\underline{\mathbf{K}}_t^\top \mathbf{C}_t \underline{\mathbf{K}}_t)} \end{aligned}$$

Proof. We use the same definitions of estimation errors as used in the proof of Theorem 9.1. Therefore, using equations (9.4), and (9.12), we express the homogeneous equation governing the state dynamics of the k^{th} agent,

$$\mathbf{x}_{t+1,k} = (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) \mathbf{x}_{t,k} + \mathbf{G}_k \mathbf{W}_{t,k}^c \underline{\boldsymbol{\xi}}_{t,k} \quad (9.14)$$

Where, $\underline{\boldsymbol{\xi}}_{t,k} = \sum_{i \in \mathcal{S}_k} \zeta_{i,k}(t) (\mathbf{K}_{t,i}^c \mathbf{x}_{t,i} - \mathbf{K}_{t,k}^c \mathbf{x}_{t,k})$. The stochastic nature introduced by RLF is reflected through an *underline* embedded below the $\boldsymbol{\xi}$ vector. Furthermore, if every agent pair (i, k) within the neighborhood fails to communicate with the same probability ρ , then $\rho_{i,k} = \rho, \forall i \in \mathcal{S}_k; \forall k$. Hence, from equation (9.13), $\mathbb{E}(\underline{\mathbf{K}}_t) = (1 - \rho) \mathbf{K}_t$. The candidate

stochastic Lyapunov function is, $\underline{V}_k(t) = \mathbf{x}_{t,k}^\top \mathbf{\Upsilon}_{t,k} \mathbf{x}_{t,k}$. Hence, the change in stochastic Lyapunov function,

$$\begin{aligned}
\delta \underline{V}_k &= \underline{V}_k(t+1) - \underline{V}_k(t) \\
&= -\mathbf{x}_{t,k}^\top \mathbf{B}_{t,k} \mathbf{x}_{t,k} \\
&\quad + 2\mathbf{x}_{t,k}^\top (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \mathbf{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \underline{\boldsymbol{\xi}}_{t,k} \\
&\quad + \underline{\boldsymbol{\xi}}_{t,k}^\top (\mathbf{G}_k \mathbf{W}_{t,k}^c)^\top \mathbf{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \underline{\boldsymbol{\xi}}_{t,k}
\end{aligned} \tag{9.15}$$

With the same choice of $\mathbf{W}_{t,k}^c$ as in Theorem 9.1 and using equation (9.13), the collective Lyapunov dynamics is,

$$\delta \underline{V} = -\mathbf{a}_t^\top \mathbf{B}_t \mathbf{a}_t + \nu^2 \mathbf{a}_t^\top \underline{\mathbf{K}}_t^\top \mathbf{C}_t \underline{\mathbf{K}}_t \mathbf{a}_t - 2\nu \mathbf{a}_t^\top \underline{\mathbf{K}}_t \mathbf{a}_t \tag{9.16}$$

where, \mathbf{B}_t and \mathbf{C}_t are the block-diagonal matrices as defined in previous theorem. \mathbf{a}_t represents the accumulation of local state vectors of individual agents. Thus, the stochastic closed-loop dynamics of the multi-agent system will be globally asymptotically stable if and only if $\delta \underline{V} < 0$, i.e., the random matrix $(\mathbf{B}_t - \nu^2 \underline{\mathbf{K}}_t^\top \mathbf{C}_t \underline{\mathbf{K}}_t)$ is positive definite. Therefore, using the properties of block-diagonal matrices and Corollary 7.1, the bounds on $|\nu|$ are obtained. \square

9.5 Simulation and Results

The proposed agent based model predictive control is applied to a 10-agent system with a radial topology. Each agent has access to two measurements. The 1st and 2nd agent's measurements are linked to 4 state elements whereas, the rest of the agents can track 6 state elements each. In the radial structure, each agent shares its state elements with two nearest leading and two nearest trailing neighbors. Additionally, no state elements are kept strictly local to any of the agent's state-space. This model resembles a radial network of

power system [114], [82]. For state estimation at any time instant, we assume that each agent is allowed either to use the local measurement or strictly participate in consensus. Under this assumption, the simulation is set up such that 6 agents are expected to use only the local measurements while the rest 4 agents are expected just to exchange information with neighbors. This constitutes the mutually exclusive event, which is defined as Case 2 in Chapter 8. The corresponding parameters are so chosen that the multi-agent system follows scenario 1 of Corollary 7.1 and the upperbound on the degree of estimation consensus is 0.107 [85]. On the other hand, the control parameters are so chosen that the multi-agent system follows Scenario 2 of Lemma 7.1. Table 9.1 lists the corresponding steady-state upper bounds of $|\nu|$ under perfect and lossy communication network following Theorem 9.1 and Corollary 9.1, respectively.

	ρ	$ \nu $ Upper bound
Perfect Network	0	0.02
Lossy Network	0.2	0.0224
	0.4	0.0263
	0.6	0.0332
	0.8	0.049

Table 9.1: *Steady State Bounds*

The continuous increase of $|\nu|$ upper bound from top to bottom in Table 9.1 indicates the necessity of more and more degree of control consensus as the agent communication links become less reliable. Next, we investigate the closed loop stability performance of the proposed AMPC. As a metric of stability we use the following,

$$\mathbf{J}_t = \mathbb{E} [\mathbf{x}_t^\top \mathbf{X} \mathbf{x}_t] \quad (9.17)$$

At each time step, \mathbf{J}_t is simulated through 1000 independent Monte Carlo trials over the multi-agent system. Fig. 9.1(a) shows the stability performance for perfect communication. Fig. 9.1(b) and Fig. 9.1(c) show the closed loop stability performance of the global dynam-

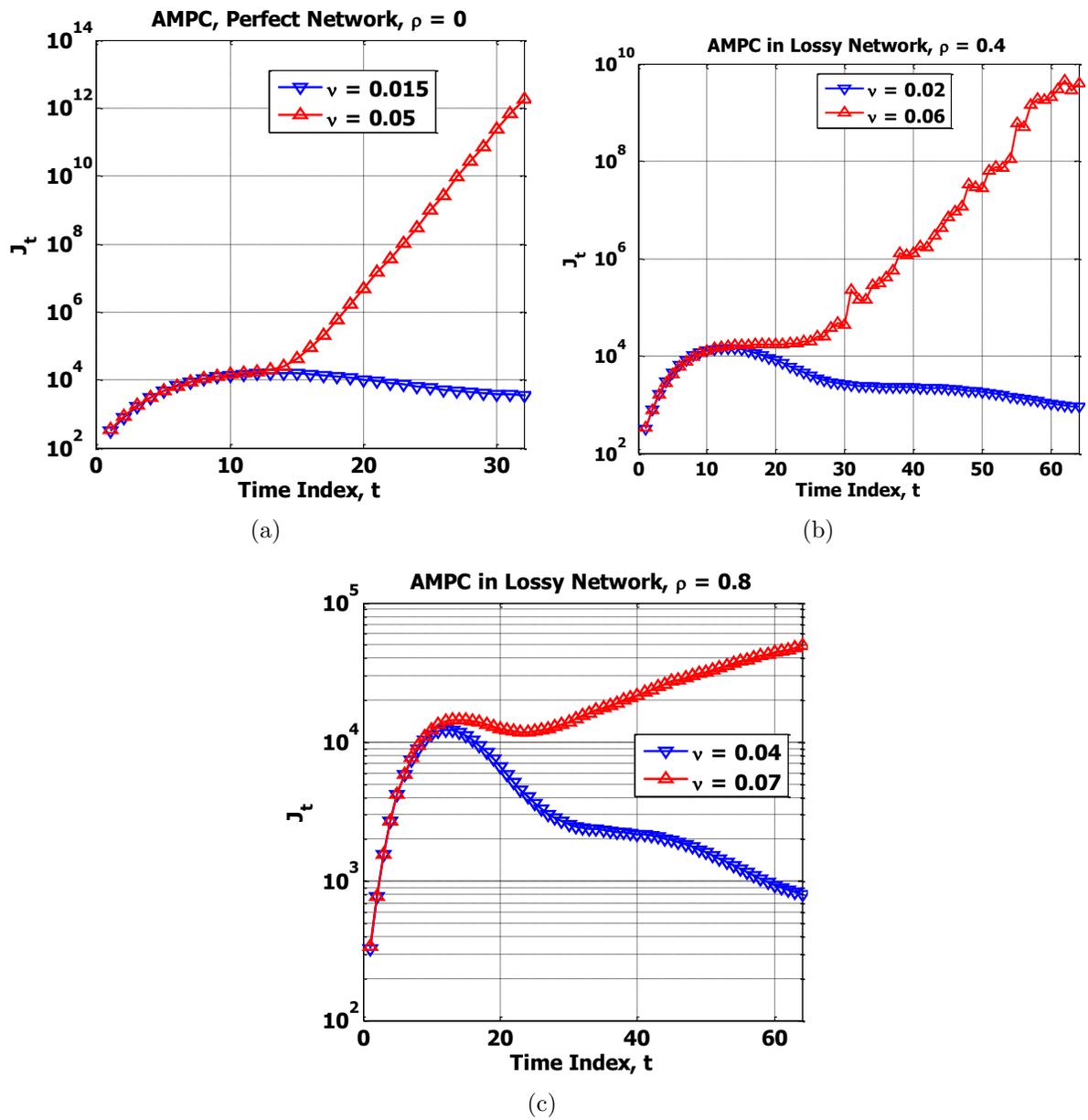


Figure 9.1: Agent based MPC Performance. (a) Perfect Network, (b) Lossy Network with Failure Probability = 0.4, (c) Lossy Network with Failure Probability = 0.8

ics when the communication between two neighboring agents fails 40% and 80% of time, respectively. For each of these cases, two values of ν is chosen - one following the bound in Table 9.1 and other violating them. It is evident from Fig. 9.1 that, the global system becomes unstable as soon as the bound on the degree of control consensus is violated under both perfect and lossy communication network.

9.6 Conclusions

In this chapter, an agent based distributed model predictive control is proposed for large-scale cyber-physical system. For each agent a local model of system dynamics is derived based on the respective observation space. Using this local model, each agent solves a linear quadratic cost minimization problem in order to derive optimal state feedback gain matrix. Through neighborhood agent consensus the final control decisions are obtained and the one with minimum cost is applied to the actual process. The bounds on the underlying degree of consensus is obtained through Lyapunov function based closed loop stability analysis under both perfect and lossy communication network. Simulation results confirm the findings by illustrating stable and unstable system behavior while the degree of control consensus obeys and violates the theoretical bounds, respectively.

Chapter 10

Conclusion and Future Work

This chapter concludes the dissertation with a summary of research results and future research directions.

10.1 Summary

In this dissertation, a compressed sensing based information aggregation and techniques for agent based distributed information processing are proposed, which address the scalability and complexity issues of a large-scale cyber-physical system. Specifically, six fundamental research questions are identified that represent the challenges of sensing massive volume of raw data, making fast and concurrent control decisions and most importantly ensuring stable operation under a lossy communication network. We summarize the key research contributions of this dissertation below:

- As an example CPS, we consider the smart grid and investigate the effect of correlated distributed generation on the voltage phasors. The estimated voltage states exhibit strong correlation that encourages the application of compressed sensing in aggregating information from the spatially deployed voltage/power sensors.
- The correlated nature of distributed generation enables the spatial, temporal and

spatiotemporal compressed sensing and it is found that, such aggregated information can be used to estimate the power system states (i.e., the voltage phasors) accurately without recovering all the measurements.

- It is also verified that the direct usage of compressed measurements within the Newton-Raphson iteration can offer excellent state estimation performance while having the same computational complexity as the indirect method.
- We decentralize the static state estimation procedure for a radial power distribution network by defining network specific binary selection matrices and enabling each sensor to compute and interact with other sensors. We refer to each sensor as an agent and address two unique aspects of radial topology commonly encountered in a distribution grid: (1) each agent shares state elements with two other agents in immediate vicinity, and (2) no state elements are left unshared. Through an extensive convergence analysis, we derive the necessary bounds for such agent based static state estimation and verify these bounds via a simulation of a 10-node system.
- We extend the agent based formulation to the 1st order Gauss-Markov process by introducing binary projection matrices that reflect individual agents' perspective on the global system dynamics. In this setup, for two custom built multi-agent systems, the agent based Kalman consensus filter (AKCF) performs better than the agent based uniformly weighted diffusion Kalman filter (ADKF) and is more robust to communication network impairments.
- We perform the Lyapunov stability analysis for the AKCF estimation error dynamics, which gives the optimal consensus weight as well as the upper bound on the degree of consensus for stable filtering operation. The analysis is further extended to incorporate the effect of lossy communication network resulting in a bound on the link failure rate that guarantees system stability. The theoretical bounds are verified via a custom built 2-agent system.

- The AKCF formulation is also generalized by introducing a flexible policy where the agents have the freedom of either collecting measurement and/or participating in consensus. Within this framework, the agents are characterized under three different cases - *deterministic*, *mutual* and *independent*. In the first case, each agent has both tasks of collecting measurements and engaging in consensus. Whereas, in the second case, an agent can choose between the two tasks at any time instance. An agent may also independently decide whether or not to take measurements or participate in consensus or even avoid both, which constitutes the third case. Through the stability analysis of the associated stochastic Lyapunov energy functions, we derive the necessary consensus bound in each case under three different scenarios of relative eigenvalue distributions. We verify the bounds and the filter stability for a custom built 10-agent system with radial topology.
- Finally, we develop an agent based distributed model predictive control, where each agent designs the state feedback control decisions, interacts with respective neighbors and control signal with the least cost is sent to the physical system. Through Lyapunov stability analysis, the corresponding bound over the control consensus is derived that ensures the closed loop system stability. The control strategy is verified by simulating a radial 10-agent system, where the system states exhibit instability once the theoretical bounds are violated.

Based on the research accomplished in this dissertation, some future research directions are highlighted in the next section.

10.2 Future Work

In this research, we do not consider the effect of bad or noisy data while implementing the centralized state estimation from compressed measurements. Secondly, we apply the agent based dynamic state estimation and control techniques to a linear system. Thirdly, we

assume synchronous communication among neighboring agents, i.e., equal packet delay is considered in neighborhood communication. Only random and independent failure of communication links is considered to analyze the stability of the proposed estimation technique. Overall, the research is carried out for a single layer of agents out of the holonic structure. As a consequence, we summarize the following potential research directions:

- In the centralized estimation of static states from compressed measurements, the effect of topology changes, noise and bad data can be considered. Since bad data is expected to be sparse (i.e., infrequent in space and / or time), we can filter out the unaffected measurements through a weighted combination of ℓ_1 -minimization and least square solution as presented in [126].
- The agent based dynamic state estimation technique can be applied to the general case of nonlinear systems. Usually, nonlinear systems are tracked using extended or unscented Kalman filters. The distributed implementation of these filters is reported in [127] with the use of diffusion algorithm. Authors in [128] include a consensus part in the unscented Kalman filter and apply it to a jump Markov nonlinear systems. In recent work, unscented Kalman filter is applied in power system dynamic state estimation [129]. In addition to that, applicability of the proposed approach can be investigated for multi-agent formulation of stochastic hybrid systems [130].
- Effect of correlated link failure as well as packet delay can be investigated for the agent based estimation and control. Recently, correlated link failure is considered in [20] in order to derive the critical link failure rate above which the centralized Kalman filter becomes unstable for a spatially distributed linear dynamical process. In [131], the effect of randomly varying packet delay is considered to derive the asymptotic stability condition of \mathcal{H}_∞ filter.
- The effect of random switching topologies can be investigated for the proposed agent based estimation and control approach. Such effect is recently investigated in [132] for

multi-agent consensus problem with Markovian switching topologies.

- The detection of faulty as well as malicious agents in the proposed architecture can be another potential research thrust. Faulty behavior of agents are generally the outcome of malfunctioning of the physical device (e.g., the sensing, computation or the communication part of the agent may be out of order or even the energy source for the agent can be depleted). One approach in detecting such faulty agent can be the assessment of the excessive delay in communication with the neighbors. On the other hand, an agent can act as an “attacker” by exhibiting abnormal behavior through delivering unexpected measurement or by giving harmful or impractical control decisions [133]. Existence of such agents in cyber-physical system necessitates a robust and vigilant multi-agent framework.
- The single layer multi-agent system analysis conducted in this research can be extended to the multi-layer holonic structure. As seen in Fig.1.2(a), the holonic model of multi-agent system is a hierarchical structure of multiple (heterogeneous) layers, where each layer is formed by a number of homogeneous agents. It is evident that we want to keep a uniform flow of information and control decisions across multiple layers instead of overburdening either end of the hierarchical structure. As a consequence, the following questions arise,
 - What is the minimum information that should flow from layer “ $(n - 1)$ ” to layer “ (n) ”?
 - What control decisions must be passed from layer “ (n) ” to “ $(n - 1)$ ”?

The answer to these questions is system specific and will also determine the inter-layer channel capacity, minimum allowable latency and also the critical rate above which the inter-layer communication should not fail.

Bibliography

- [1] S. Deshmukh, “Estimation & control in spatially distributed cyber physical systems,” Ph.D. dissertation, Kansas State University, 2013.
- [2] “ECE colloquium series: The challenges of cyber-physical systems,” 2014. [Online]. Available: <https://youtu.be/7pLoepM6i2s>
- [3] V. Gunes, S. Peter, T. Givargis, and F. Vahid, “A survey on concepts, applications, and challenges in cyber-physical systems,” *Internet and Information Systems, KSII Transactions on*, vol. 8, no. 12, pp. 4242–4268, 2014. [Online]. Available: http://cps.ics.uci.edu/publication/cps_survey/
- [4] E. A. Lee. (2012) Cyber-Physical Systems - a Concept Map. [Online]. Available: <http://cyberphysicalsystems.org/>
- [5] “CPS Vision Statement: National Science Foundation,” 2015. [Online]. Available: [https://www.nitrd.gov/nitrdgroups/images/6/6a/Cyber_Physical_Systems_\(CPS\)_Vision_Statement.pdf](https://www.nitrd.gov/nitrdgroups/images/6/6a/Cyber_Physical_Systems_(CPS)_Vision_Statement.pdf)
- [6] E. A. Lee and S. A. Seshia, *Introduction to Embedded Systems - A Cyber-Physical Systems Approach*, 1st ed. LeeSeshia.org, 2014.
- [7] N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, 2nd ed. The MIT Press, 1961.
- [8] D. Y. Kim and M. Jeon, “Robust distributed Kalman filter for wireless sensor networks with uncertain communication channels,” *Mathematical Problems in Engineering*, vol. 2012, 2012, article id 238597.

- [9] T. Sanislav and L. Miclea, “Cyber-physical systems - concept, challenges and research areas,” *Journal of Control Engineering and Applied Informatics*, vol. 14, no. 2, pp. 28–33, 2012.
- [10] S. Karnouskos, “Cyber-physical systems in the smartgrid,” in *Industrial Informatics (INDIN), 2011 9th IEEE International Conference on*, July 2011, pp. 20–23.
- [11] “Usdot intelligent transportation systems joint program office,” 2015. [Online]. Available: <http://www.its.dot.gov/>
- [12] S. A. Haque, S. M. Aziz, and M. Rahman, “Review of cyber-physical system in health-care,” *International Journal of Distributed Sensor Networks*, 2014, article id 217415.
- [13] “Climate-smart agriculture,” 2015. [Online]. Available: <http://www.fao.org/climate-smart-agriculture/en/>
- [14] F. Mehdipour, “Smart field monitoring: An application of cyber-physical systems in agriculture (work in progress),” in *Advanced Applied Informatics (IIAIAAI), 2014 IIAI 3rd International Conference on*, Aug 2014, pp. 181–184.
- [15] P. Derler, E. Lee, and A.-S. Vincentelli, “Modeling cyber-physical systems,” *Proceedings of the IEEE*, vol. 100, no. 1, pp. 13–28, Jan 2012.
- [16] A. B. Sharma, F. Ivančić, A. Niculescu-Mizil, H. Chen, and G. Jiang, “Modeling and analytics for cyber-physical systems in the age of big data,” *SIGMETRICS Perform. Eval. Rev.*, vol. 41, no. 4, pp. 74–77, Apr. 2014.
- [17] P. Gupta and P. Kumar, “The capacity of wireless networks,” *Information Theory, IEEE Transactions on*, vol. 46, no. 2, pp. 388–404, mar 2000.
- [18] “Pecan street research institute,” 2013. [Online]. Available: <http://www.pecanstreet.org/>

- [19] K. Nagasawa, C. R. Upshaw, J. D. Rhodes, C. L. Holcomb, D. A. Walling, and M. E. Webber, "Data management for a large-scale smart grid demonstration project in Austin, Texas," in *Energy Sustainability, ASME 2012 6th International Conference on*, 2012, pp. 1027–1031.
- [20] S. Deshmukh, B. Natarajan, and A. Pahwa, "State estimation over a lossy network in spatially distributed cyber-physical systems," *Signal Processing, IEEE Transactions on*, vol. 62, no. 15, Aug 2014.
- [21] M. Amin, "Security challenges for the electricity infrastructure," *Computer*, vol. 35, no. 4, pp. 8–10, Apr 2002.
- [22] "Smart Grid: An Introduction," 2008. [Online]. Available: <http://www.smartgrid.gov>
- [23] P. Barker and R. De Mello, "Determining the impact of distributed generation on power systems: Part 1 - radial distribution systems," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 3, 2000, pp. 1645 –1656.
- [24] C. N. Long and T. P. Ackerman, "Surface measurements of solar irradiance: A study of the spatial correlation between simultaneous measurements at separated sites," *J. Appl. Meteor.*, vol. 34, no. 5, p. 1039 1046, may 1995.
- [25] J. Bebic, R. Walling, K. O'Brien, and B. Kroposki, "The Sun also rises," *Power and Energy Magazine, IEEE*, vol. 7, no. 3, pp. 45 –54, may-june 2009.
- [26] A. Ciancio, S. Pattem, A. Ortega, and B. Krishnamachari, "Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm," in *Proceedings of the 5th international conference on Information processing in sensor networks*, ser. IPSN '06. New York, NY, USA: ACM, 2006, pp. 309–316.
- [27] S. Mallat, *A Wavelet Tour of Signal Processing*. Elsevier Inc., 2009.

- [28] R. A. DeVore, “Nonlinear approximation,” *Acta Numerica*, vol. 1998, pp. 51–150, 1998.
- [29] H. Li, R. Mao, L. Lai, and R. Qiu, “Compressed meter reading for delay-sensitive and secure load report in smart grid,” in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, oct. 2010, pp. 114 –119.
- [30] R. Ranganathan, R. Qiu, and Z. Hu, “Cognitive radio for smart grid: Theory, algorithms, and security,” *International Journal of Digital Multimedia Broadcasting*, vol. 2011, 2011, article id 502087.
- [31] T. Islam and I. Koo, “Compressed sensing-based data gathering in wireless home area network for smart grid,” in *Informatics, Electronics Vision (ICIEV), 2012 International Conference on*, may 2012, pp. 82 –86.
- [32] H. Wang and N. Schulz, “A revised branch current-based distribution system state estimation algorithm and meter placement impact,” *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 207 – 213, feb. 2004.
- [33] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. New York, USA: Marcel Dekker, Inc., 2004.
- [34] E. Manitsas, R. Singh, B. Pal, and G. Strbac, “Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling,” *Power Systems, IEEE Transactions on*, vol. 27, no. 4, pp. 1888 –1896, nov. 2012.
- [35] R. Singh, B. Pal, and R. Jabr, “Distribution system state estimation through gaussian mixture model of the load as pseudo-measurement,” *Generation, Transmission Distribution, IET*, vol. 4, no. 1, pp. 50 –59, january 2010.
- [36] E. L. Allgower and K. Georg, *Numerical Continuation Methods: An Introduction*. Springer Series in Computational Mathematics, 1990, vol. 13.

- [37] J. M. Martínez, “Quasi-Newton methods for solving underdetermined nonlinear simultaneous equations,” *Computational and Applied Mathematics, Journal of*, vol. 34, no. 2, pp. 171–190, April 1991.
- [38] P. Kuegler, “A sparse update method for solving underdetermined systems of nonlinear equations applied to the manipulation of biological signaling pathways,” *SIAM J. Appl. Math.*, vol. 72, no. 4, pp. 982–1001, 2012.
- [39] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*, 1st ed. Florida, USA: CRC Press LLC, 1998.
- [40] Y.-F. Huang, S. Werner, J. Huang, N. Kashyap, and V. Gupta, “State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid,” *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 33–43, 2012.
- [41] G. Korres, “A distributed multiarea state estimation,” *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 73–84, 2011.
- [42] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. V. Cutsem, “A taxonomy of multi-area state estimation methods,” *Electric Power Systems Research*, vol. 81, no. 4, pp. 1060 – 1069, 2011.
- [43] A. Gómez-Expósito, A. Abur, A. de la Villa Jaén, and C. Gómez-Quiles, “A multilevel state estimation paradigm for smart grids,” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 952–976, 2011.
- [44] L. Xie, D.-H. Choi, S. Kar, and H. Poor, “Fully distributed state estimation for wide-area monitoring systems,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 3, pp. 1154–1169, 2012.
- [45] V. Kekatos and G. Giannakis, “Distributed robust power system state estimation,” *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 1617–1626, 2013.

- [46] H. Zhu and G. Giannakis, “Multi-area state estimation using distributed sdp for nonlinear power systems,” in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*, 2012, pp. 623–628.
- [47] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [48] J. Gilbert, C. Moler, and R. Schreiber, “Sparse matrices in matlab: Design and implementation,” *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 333–356, 1992.
- [49] H. Farag, E. El-Saadany, and R. Seethapathy, “A two ways communication-based distributed control for voltage regulation in smart distribution feeders,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 1, pp. 271–281, 2012.
- [50] C.-H. Lo and N. Ansari, “Decentralized controls and communications for autonomous distribution networks in smart grid,” *Smart Grid, IEEE Transactions on*, vol. 4, no. 1, pp. 66–77, 2013.
- [51] R. Olfati-Saber, “Distributed Kalman filtering for sensor networks,” in *Decision and Control, 2007 46th IEEE Conference on*, 2007, pp. 5492–5498.
- [52] F. Cattivelli and A. Sayed, “Diffusion strategies for distributed Kalman filtering and smoothing,” *Automatic Control, IEEE Transactions on*, vol. 55, no. 9, pp. 2069–2084, 2010.
- [53] W. Yang, X. Wang, and H. Shi, “Optimal consensus-based distributed estimation with intermittent communication,” *International Journal of Systems Science*, vol. 42, no. 9, pp. 1521–1529, 2011.

- [54] R. Olfati-Saber, “Kalman-consensus filter : Optimality, stability, and performance,” in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, 2009, pp. 7036–7042.
- [55] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.
- [56] X. Zhao, S.-Y. Tu, and A. Sayed, “Diffusion adaptation over networks under imperfect information exchange and non-stationary data,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 7, pp. 3460–3475, 2012.
- [57] H. Dong, Z. Wang, and H. Gao, “Distributed filtering for a class of time-varying systems over sensor networks with quantization errors and successive packet dropouts,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 6, pp. 3164–3173, 2012.
- [58] M. Mahmoud and H. Khalid, “Distributed Kalman filtering: a bibliographic review,” *Control Theory Applications, IET*, vol. 7, no. 4, pp. 483–501, 2013.
- [59] A. Alessio and A. Bemporad, “Decentralized model predictive control of constrained linear systems,” in *European Control Conference. EUCA*, 2007, pp. 2813–2818.
- [60] U. Khan and J. Moura, “Distributing the Kalman filter for large-scale systems,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 10, pp. 4919–4935, 2008.
- [61] A. Mohammadi and A. Asif, “Distributed state estimation for large-scale nonlinear systems: A reduced order particle filter implementation,” in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, Aug 2012, pp. 249–252.

- [62] F. Dorfler, F. Pasqualetti, and F. Bullo, “Continuous-time distributed observers with discrete communication,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 2, pp. 296–304, April 2013.
- [63] A. Alessio, D. Barcelli, and A. Bemporad, “Decentralized model predictive control of dynamically coupled linear systems,” *Journal of Process Control*, vol. 21, no. 5, pp. 705 – 714, 2011, special Issue on Hierarchical and Distributed Model Predictive Control.
- [64] J. Lemos, L. Pinto, L. Rato, and M. Rijo, “Distributed lqg control of a water delivery canal with feedforward from measured consumptions,” in *Control Automation (MED), 2012 20th Mediterranean Conference on*, July 2012, pp. 722–727.
- [65] A. Bemporad and D. Barcelli, “Decentralized model predictive control,” in *Networked Control Systems*, ser. Lecture Notes in Control and Information Sciences, A. Bemporad, M. Heemels, and M. Johansson, Eds. Springer London, 2010, vol. 406, pp. 149–178.
- [66] A. Alessio and A. Bemporad, “Decentralized model predictive control of constrained linear systems,” in *Control Conference (ECC), 2007 European*, July 2007, pp. 2813–2818.
- [67] J. Maestre, D. M. de la Pea, E. Camacho, and T. Alamo, “Distributed model predictive control based on agent negotiation,” *Journal of Process Control*, vol. 21, no. 5, pp. 685 – 697, 2011, special Issue on Hierarchical and Distributed Model Predictive Control.
- [68] W. Yang and H. Shi, “Distributed estimation based on lqg control over homogeneous sensor networks,” *International Journal of Control, Automation and Systems*, vol. 10, no. 6, pp. 1173–1181, 2012.

- [69] H. Mostafa, R. El-Shatshat, and M. Salama, “Multi-objective optimization for the operation of an electric distribution system with a large number of single phase solar generators,” *Smart Grid, IEEE Transactions on*, vol. 4, no. 2, pp. 1038–1047, 2013.
- [70] V. R. Disfani, L. Fan, L. Piyasinghe, and Z. Miao, “Multi-agent control of community and utility using lagrangian relaxation based dual decomposition,” *Electric Power Systems Research*, vol. 110, no. 0, pp. 45 – 54, 2014.
- [71] P. Nguyen, W. Kling, and P. Ribeiro, “A game theory strategy to integrate distributed agent-based functions in smart grids,” *Smart Grid, IEEE Transactions on*, vol. 4, no. 1, pp. 568–576, 2013.
- [72] Y. Cao, W. Yu, W. Ren, and G. Chen, “An overview of recent progress in the study of distributed multi-agent coordination,” *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 1, pp. 427–438, 2013.
- [73] P. D. Christofides, R. Scattolini, D. M. noz de la Peña, and J. Liu, “Distributed model predictive control: A tutorial review and future research directions,” *Computers & Chemical Engineering*, vol. 51, pp. 21–41, 2013.
- [74] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall, 2009.
- [75] J. Sztipanovits, X. Koutsoukos, G. Karsai, N. Kottenstette, P. Antsaklis, V. Gupta, B. Goodwine, J. Baras, and S. Wang, “Toward a science of cyber-physical system integration,” *Proceedings of the IEEE*, vol. 100, no. 1, pp. 29–44, Jan 2012.
- [76] J. Lin, “Agent-based analysis and mitigation of failure for cyber-physical systems,” Ph.D. dissertation, Missouri University of Science and Technology, 2011.

- [77] M. Simoes and S. Bhattarai, “Improving energy efficiency of cyber physical systems using multi-agent based control,” in *Industry Applications Society Annual Meeting (IAS), 2012 IEEE*, Oct 2012, pp. 1–7.
- [78] C. Ego, T. Hauptenthal, T. Schöler, and F. Weichenmeier, “Agent-based traffic control a cyber-physical system approach,” in *Mobil.TUM 2013, Mobility and Transport ITS for Connected Mobility International Scientific Conference on*, June 2013, pp. 1–6.
- [79] S. Alam, B. Natarajan, and A. Pahwa, “Impact of correlated distributed generation on information aggregation in smart grid,” in *Green Technologies Conference, 2013 IEEE*, April 2013, pp. 160–166.
- [80] —, “Distribution grid state estimation from compressed measurements,” *Smart Grid, IEEE Transactions on*, vol. 5, no. 4, pp. 1631–1642, July 2014.
- [81] S. Alam, B. Natarajan, A. Pahwa, and S. Curto, “Agent based state estimation in smart distribution grid,” *Latin America Transactions, IEEE (Revista IEEE America Latina)*, vol. 13, no. 2, pp. 496–502, Feb 2015.
- [82] A. Pahwa, S. DeLoach, B. Natarajan, S. Das, A. Malekpour, S. M. S. Alam, and D. Case, “Goal-based holonic multiagent system for operation of power distribution systems,” *Smart Grid, IEEE Transactions on*, vol. 6, no. 5, pp. 2510–2518, September 2015.
- [83] S. M. S. Alam, B. Natarajan, and A. Pahwa, “Distributed agent-based dynamic state estimation over a lossy network,” in *5th Int. Workshop on Networks of Cooperating Objects for Smart Cities (UBICITEC 2014)*, 2014.
- [84] —, “Agent based optimally weighted Kalman consensus filter over a lossy network,” in *Global Communications, 2015 IEEE Conference on (accepted)*, December 2015.

- [85] —, “Agent based Kalman filtering - measurement vs consensus?” *Control of Network Systems, IEEE Transactions on (under review)*, 2015.
- [86] S. M. S. Alam and B. Natarajan, “Stability of agent based distributed model predictive control over a lossy network,” *Signal and Information Processing over Networks, IEEE Transactions on (In Press)*, 2015.
- [87] M. Duarte and Y. Eldar, “Structured compressed sensing: From theory to applications,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4053–4085, sept. 2011.
- [88] E. Candes and T. Tao, “Decoding by linear programming,” *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, dec. 2005.
- [89] Y. Rivenson and A. Stern, “Compressed imaging with a separable sensing operator,” *Signal Processing Letters, IEEE*, vol. 16, no. 6, pp. 449–452, june 2009.
- [90] W. L. Brogan, *Modern Control Theory*, 3rd ed. New Jersey: Prentice Hall, Inc., 1991.
- [91] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed., ser. Springer Texts in Electrical Engineering. New York, USA: Springer New York, 1994.
- [92] C. R. Henderson, “Best linear unbiased estimation and prediction under a selection model,” *Biometrics*, vol. 31, no. 2, pp. pp. 423–447, 1975.
- [93] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Prentice-Hall, Inc., 2000.
- [94] H. Zhu and G. Giannakis, “Estimating the state of ac power systems using semidefinite programming,” in *North American Power Symposium (NAPS), 2011*, 2011, pp. 1–7.
- [95] R. E. Kalman *et al.*, “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

- [96] “egauge systems llc,” 2013. [Online]. Available: <http://egauge.net/>
- [97] S. Deshmukh, B. Natarajan, and A. Pahwa, “Voltage/var control in distribution networks via reactive power injection through distributed generators,” *Smart Grid, IEEE Transactions on*, vol. 3, no. 3, pp. 1226–1234, sept. 2012.
- [98] A. R. Bergen and V. Vittal, *Power System Analysis*, 2nd ed. New Jersey, USA: Prentice Hall, 2000.
- [99] J. O. Berger, V. de Oliveira, and B. Sanso, “Objective bayesian analysis of spatially correlated data,” *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1361–1374, Dec. 2001.
- [100] P. Bacher, H. Madsen, and H. A. Nielsen, “Online short-term solar power forecasting,” *Solar Energy*, vol. 83, no. 10, pp. 1772–1783, 2009.
- [101] S. Deshmukh, B. Natarajan, and A. Pahwa, “Stochastic state estimation for smart grids in the presence of intermittent measurements,” in *Proceedings of the 4th IEEE Latin-American Conference on Communications*, ser. LATINCOM’12, nov. 2012.
- [102] “Distribution test feeder,” 2012. [Online]. Available: <http://ewh.ieee.org/soc/pes/dsacom/testfeeders/index.html>
- [103] “CVX: Matlab software for disciplined convex programming, version 2.0 beta,” sep. 2012. [Online]. Available: <http://cvxr.com/cvx>
- [104] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.
- [105] B. Vidakovic, *Statistical Modeling by Wavelets*. John Wiley & Sons, Inc., 1999.

- [106] G. N. Korres, N. D. Hatziargyriou, and P. J. Katsikas, "State estimation in multi-microgrids," *European Transactions on Electrical Power*, vol. 21, no. 2, pp. 1178–1199, 2011.
- [107] J. Huang, V. Gupta, and Y.-F. Huang, "Electric grid state estimators for distribution systems with microgrids," in *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*, 2012, pp. 1–6.
- [108] G. Korres and N. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *Power and Energy Society General Meeting, 2012 IEEE*, 2012, pp. 1–8.
- [109] A. Arefi and M.-R. Haghifam, "State estimation in smart power grids," in *Smart Power Grids 2011*, ser. Power Systems, A. Keyhani and M. Marwali, Eds. Springer Berlin Heidelberg, 2012, pp. 439–478.
- [110] E. Caro, A. Conejo, and R. Minguez, "Power system state estimation considering measurement dependencies," *Power Systems, IEEE Transactions on*, vol. 24, no. 4, pp. 1875–1885, 2009.
- [111] G. Valverde, A. Saric, and V. Terzija, "Stochastic monitoring of distribution networks including correlated input variables," *Power Systems, IEEE Transactions on*, vol. 28, no. 1, pp. 246–255, 2013.
- [112] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [113] E. Candes and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n ," *The Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.

- [114] S. M. S. Alam, B. Natarajan, and A. Pahwa, “Agent based state estimation in smart distribution grid,” in *Communications (LATINCOM), 2013 IEEE Latin-America Conference on*, November 2013, pp. 1–6.
- [115] I. Daubechies and J. C. Lagarias, “Sets of matrices all infinite products of which converge,” *Linear Algebra and its Applications*, vol. 161, no. 0, pp. 227 – 263, 1992.
- [116] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. New York, USA: Cambridge University Press, 1994.
- [117] M. F. Hassan, G. Salut, G. Singh Madan, and A. Titli, “A decentralized computational algorithm for the global Kalman filter,” *Automatic Control, IEEE Transactions on*, vol. 23, no. 2, pp. 262–268, Apr 1978.
- [118] R. F. Stengel, *Optimal Control and Estimation*. New York, USA: Dover Publications, Inc., 1994.
- [119] W. Li, Y. Jia, and J. Du, “Distributed Kalman consensus filter with intermittent observations,” *Journal of the Franklin Institute*, no. 0, pp. –, 2015.
- [120] Y. Mo and B. Sinopoli, “Kalman filtering with intermittent observations: Tail distribution and critical value,” *Automatic Control, IEEE Transactions on*, vol. 57, no. 3, pp. 677–689, March 2012.
- [121] D. Y. Kim, J. hong Yoon, Y.-H. Kim, and V. Shin, “Distributed information fusion filter with intermittent observations,” in *Information Fusion (FUSION), 2010 13th Conference on*, July 2010, pp. 1–7.
- [122] S. Deshmukh, B. Natarajan, and A. Pahwa, “State estimation and voltage/var control in distribution network with intermittent measurements,” *Smart Grid, IEEE Transactions on*, vol. 5, no. 1, pp. 200–209, Jan 2014.

- [123] T. Tettamanti, T. Luspay, B. Kulcsar, T. Peni, and I. Varga, “Robust control for urban road traffic networks,” *Intelligent Transportation Systems, IEEE Transactions on*, vol. 15, no. 1, pp. 385–398, Feb 2014.
- [124] K. Subramanian, J. B. Rawlings, and C. T. Maravelias, “Economic model predictive control for inventory management in supply chains,” *Computers & Chemical Engineering*, vol. 64, pp. 71–80, 2014.
- [125] M. Rabbat and R. Nowak, “Distributed optimization in sensor networks,” in *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, April 2004, pp. 20–27.
- [126] B. Moore and B. Natarajan, “A general framework for robust compressive sensing based nonlinear regression,” in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2012 IEEE 7th*, 2012, pp. 225–228.
- [127] F. Cattivelli and A. Sayed, “Distributed nonlinear kalman filtering with applications to wireless localization,” in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 3522–3525.
- [128] W. Li and Y. Jia, “Consensus-based distributed multiple model ukf for jump markov nonlinear systems,” *Automatic Control, IEEE Transactions on*, vol. 57, no. 1, pp. 227–233, Jan 2012.
- [129] A. Singh and B. Pal, “Decentralized dynamic state estimation in power systems using unscented transformation,” *Power Systems, IEEE Transactions on*, vol. 29, no. 2, pp. 794–804, March 2014.
- [130] J. Hespanha, “Modelling and analysis of stochastic hybrid systems,” *IEE Proceedings - Control Theory and Applications*, vol. 153, pp. 520–535(15), September 2006.

- [131] B. Shen, Z. Wang, H. Shu, and G. Wei, “ H_∞ filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays,” *Automatica*, vol. 45, no. 4, pp. 1032 – 1037, 2009.
- [132] K. You, Z. Li, and L. Xie, “Consensus condition for linear multi-agent systems over randomly switching topologies,” *Automatica*, vol. 49, no. 10, pp. 3125 – 3132, 2013.
- [133] A. Teixeira, K. C. Sou, H. Sandberg, and K. Johansson, “Secure control systems: A quantitative risk management approach,” *Control Systems, IEEE*, vol. 35, no. 1, pp. 24–45, Feb 2015.

Appendix A

SYS1 Parameters

- State initialization: $\boldsymbol{\mu} = \begin{bmatrix} 10 & 5 & 20 & 2.5 & 40 & 1.25 \end{bmatrix}^\top$; $\boldsymbol{\Sigma} = \text{diag}(.8, .2, .5, 1.3, .1, .3)$;

- Global state transition matrix: $\mathbf{F} =$

$$\begin{bmatrix} 0.3153 & 0.0090 & 0.0541 & 0.2342 & 0.1712 & 0.2162 \\ 0.0270 & 0.2883 & 0.0631 & 0.1892 & 0.2072 & 0.2252 \\ 0.2793 & 0.0811 & 0.0180 & 0.1982 & 0.2432 & 0.1802 \\ 0.0721 & 0.2523 & 0.2973 & 0.1532 & 0.0901 & 0.1351 \\ 0.2703 & 0.0450 & 0.3063 & 0.1081 & 0.1261 & 0.1441 \\ 0.0360 & 0.3243 & 0.2613 & 0.1171 & 0.1622 & 0.0991 \end{bmatrix};$$

- Process noise: $\mathbf{Q} = \text{diag}(1.8, .9, 2.7, 3.6, 1.0, .5)$;

- Agent-wise observation model:

$$\mathbf{H}_1 = \begin{bmatrix} 13.3758 & 3.2277 & 5.2820 & 18.3034 \\ 6.7976 & 0.3931 & 4.9316 & 4.1200 \\ 8.1333 & 6.3871 & 13.2687 & 13.6812 \end{bmatrix};$$

$$\mathbf{R}_1 = \text{diag}(0.9217, 0.3057, 0.7316);$$

$$\mathbf{H}_2 = \begin{bmatrix} 41.4826 & 10.9732 & 27.7157 \\ 22.6780 & 28.1071 & 39.7014 \end{bmatrix};$$

$$\mathbf{R}_2 = \text{diag}(0.0019, 0.2653);$$

$$\mathbf{H}_3 = \begin{bmatrix} 41.6863 & 49.7289 & 39.0787 & 26.8042 & 4.2471 \\ 13.5406 & 48.6878 & 63.23 & 7.3416 & 14.5855 \\ 15.2408 & 46.3493 & 2.1896 & 73.8971 & 73.7868 \\ 68.2435 & 38.4299 & 46.5694 & 43.9841 & 10.4298 \end{bmatrix};$$

$$\mathbf{R}_3 = \text{diag}(0.3703, 0.3765, 0.2747, 0.3410).$$

Appendix B

SYS2 Parameters

- Dynamics of SYS2:

$$\boldsymbol{\mu} = \left[10 \ 5 \ 20 \ 2.5 \ 40 \ 1.25 \ 4 \ 3.7 \ 25.8 \ 21.4 \ 76 \ 51.6 \ 88.9 \right]^T ;$$

$$\boldsymbol{\Sigma} = \text{diag}(.8, .2, .5, 1.3, .1, .3, .4, 1, .7, 1.2, .9, 3.9, 5.7);$$

$$\mathbf{Q} = \text{diag}(1.8, 0.9, 2.7, 3.6, 1, 0.5, 0.1, 4.5, 2, 8, 5, 1.5, 0.3);$$

$$\mathbf{F} = \begin{bmatrix} 0.0842 & 0.0977 & 0.1113 & 0.1249 & 0.1385 & 0.1520 & 0.0009 & 0.0145 & 0.0281 & 0.0416 & 0.0552 & 0.0688 \\ 0.0968 & 0.1104 & 0.1240 & 0.1376 & 0.1511 & 0.0118 & 0.0136 & 0.0271 & 0.0407 & 0.0543 & 0.0679 & 0.0814 \\ 0.1095 & 0.1231 & 0.1367 & 0.1502 & 0.0109 & 0.0127 & 0.0262 & 0.0398 & 0.0534 & 0.0670 & 0.0805 & 0.0941 \\ 0.1222 & 0.1357 & 0.1493 & 0.0100 & 0.0235 & 0.0253 & 0.0389 & 0.0525 & 0.0661 & 0.0796 & 0.0932 & 0.0950 \\ 0.1348 & 0.1484 & 0.0090 & 0.0226 & 0.0244 & 0.0380 & 0.0516 & 0.0652 & 0.0787 & 0.0923 & 0.1059 & 0.1077 \\ 0.1475 & 0.0081 & 0.0217 & 0.0353 & 0.0371 & 0.0507 & 0.0643 & 0.0778 & 0.0914 & 0.1050 & 0.1068 & 0.1204 \\ 0.0072 & 0.0208 & 0.0344 & 0.0362 & 0.0498 & 0.0633 & 0.0769 & 0.0905 & 0.1041 & 0.1176 & 0.1195 & 0.1330 \\ 0.0199 & 0.0335 & 0.0471 & 0.0489 & 0.0624 & 0.0760 & 0.0896 & 0.1032 & 0.1167 & 0.1186 & 0.1321 & 0.1457 \\ 0.0326 & 0.0462 & 0.0480 & 0.0615 & 0.0751 & 0.0887 & 0.1023 & 0.1158 & 0.1294 & 0.1312 & 0.1448 & 0.0054 \\ 0.0452 & 0.0588 & 0.0606 & 0.0742 & 0.0878 & 0.1014 & 0.1149 & 0.1285 & 0.1303 & 0.1439 & 0.0045 & 0.0181 \\ 0.0579 & 0.0597 & 0.0733 & 0.0869 & 0.1005 & 0.1140 & 0.1276 & 0.1412 & 0.1430 & 0.0036 & 0.0172 & 0.0308 \\ 0.0706 & 0.0724 & 0.0860 & 0.0995 & 0.1131 & 0.1267 & 0.1403 & 0.1421 & 0.0027 & 0.0163 & 0.0299 & 0.0434 \\ 0.0715 & 0.0851 & 0.0986 & 0.1122 & 0.1258 & 0.1394 & 0.1529 & 0.0018 & 0.0154 & 0.0290 & 0.0425 & 0.0561 \end{bmatrix}$$

- Agent-wise observation model:

$$\mathbf{H}_1 = \begin{bmatrix} 5.03 & 1.9 & 19.88 & 15.16 & 5.23 & 2.1 & 18.42 \\ 14.55 & 4.19 & 2.18 & 8.78 & 8.23 & 4.46 & 12.04 \\ 4.82 & 7.87 & 17.23 & 3.49 & 13.23 & 17.94 & 10.37 \\ 6.91 & 2.91 & 4.45 & 7.97 & 4.73 & 12.60 & 7.36 \end{bmatrix};$$

$$\mathbf{R}_1 = \text{diag}(0.84, 0.16, 0.29, 0.37);$$

$$\mathbf{H}_2 = \begin{bmatrix} 15.45 & 25.53 & 6.17 & 46.70 & 16.23 & 41.32 & 32.43 \\ 5.06 & 39.37 & 31.16 & 0.27 & 32.76 & 8.01 & 20.26 \\ 39.33 & 18.28 & 44.17 & 27.31 & 6.27 & 27.23 & 13.33 \\ 42.76 & 8.67 & 35.13 & 49.78 & 44.14 & 15.19 & 22.33 \\ 17.46 & 39.27 & 48.45 & 31.98 & 15.56 & 6.85 & 44.79 \\ 9.14 & 27.59 & 2.34 & 34.65 & 34.39 & 42.29 & 30.44 \end{bmatrix};$$

$$\mathbf{R}_2 = \text{diag}(0.47, 0.23, 0.82, 0.32, 0.58, 0.42);$$

$$\mathbf{H}_3 = \begin{bmatrix} 8.19 & 75.38 & 68.21 & 77.34 & 13.89 & 58.98 & 73.77 \\ 15.3 & 38.88 & 68.12 & 45.0 & 56.2 & 65.84 & 17.77 \\ 18.96 & 20.23 & 66.23 & 61.85 & 51.02 & 18.76 & 58.98 \end{bmatrix};$$

$$\mathbf{R}_3 = \text{diag}(0.45, 0.81, 0.53).$$

Appendix C

Positive Definiteness of $\mathbf{G}_{t,k}$

From the definition of $\mathcal{D}_{t,k}$, $\mathbf{G}_{t,k} = \mathbf{M}_{t-1,k|t-1}^{-1} - \mathcal{C}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \mathcal{C}_{t,k}$. Using the definition of $\mathcal{C}_{t,k}$ and then step 5 of Algorithm 3, $\mathbf{G}_{t,k} = \mathbf{M}_{t-1,k|t-1}^{-1} - \mathbf{F}_k^\top \mathbf{M}_{t,k|t-1}^{-1} \mathbf{M}_{t,k|t} \mathbf{M}_{t,k|t-1}^{-1} \mathbf{F}_k$. Now, in the information form of Kalman filter, $\mathbf{M}_{t,k|t}^{-1} = \mathbf{M}_{t,k|t-1}^{-1} + \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k$. Therefore,

$$\begin{aligned}
 \mathbf{G}_{t,k} &= \mathbf{M}_{t-1,k|t-1}^{-1} \\
 &\quad - \mathbf{F}_k^\top \left(\mathbf{M}_{t,k|t}^{-1} - \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \right) \mathbf{M}_{t,k|t} \mathbf{M}_{t,k|t-1}^{-1} \mathbf{F}_k \\
 &= \mathbf{M}_{t-1,k|t-1}^{-1} - \mathbf{F}_k^\top \mathbf{M}_{t,k|t-1}^{-1} \mathbf{F}_k \\
 &\quad + \left(\mathbf{M}_{t,k|t} \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{F}_k \right)^\top \mathbf{M}_{t,k|t-1}^{-1} \mathbf{F}_k
 \end{aligned} \tag{C.1}$$

From step 3 of Algorithm 3, the alternate expression of $\mathbf{F}_k^\top \mathbf{M}_{t,k|t-1}^{-1} \mathbf{F}_k$ is,

$$\begin{aligned}
 &= \mathbf{F}_k^\top \left\{ \mathbf{F}_k \mathbf{M}_{t-1,k|t-1} \mathbf{F}_k^\top + \mathbf{Q}_k \right\}^{-1} \mathbf{F}_k \\
 &= \left\{ \mathbf{F}_k^{-1} \mathbf{Q}_k (\mathbf{F}_k^{-1})^\top + \mathbf{M}_{t-1,k|t-1} \right\}^{-1} \\
 &= \mathbf{M}_{t-1,k|t-1}^{-1} - \mathbf{M}_{t-1,k|t-1}^{-1} \mathbf{F}_k^{-1} \mathcal{Y}_{t,k}^{-1} (\mathbf{F}_k^{-1})^\top \mathbf{M}_{t-1,k|t-1}^{-1}
 \end{aligned} \tag{C.2}$$

where, $\mathcal{Y}_{t,k} = \mathbf{Q}_k^{-1} + (\mathbf{F}_k^{-1})^\top \mathbf{M}_{t-1,k|t-1}^{-1} \mathbf{F}_k^{-1}$. Combining the above two equations,

$$\begin{aligned} \mathbf{G}_{t,k} &= \mathbf{M}_{t-1,k|t-1}^{-1} (\mathbf{F}_k^\top \mathcal{Y}_{t,k} \mathbf{F}_k)^{-1} \mathbf{M}_{t-1,k|t-1}^{-1} \\ &\quad + \mathbf{F}_k^\top \left(\{ \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \} \mathbf{M}_{t,k|t} \mathbf{M}_{t,k|t-1}^{-1} \right) \mathbf{F}_k \end{aligned} \quad (\text{C.3})$$

The final expression ensures the positive definiteness of $\mathbf{G}_{t,k}$.

Appendix D

Least Square Estimate of AKCF

Weights

We start with the definitions of following block matrices,

- A block matrix \mathbf{A}^f ,

$$[\mathbf{A}^f]_{(k,l)^{th}block} = \begin{cases} -\mathbf{P}_{l,k} \mathbf{S}_l \mathbf{F}_l; & l \in \mathcal{S}_k \\ \sum_{i \in \mathcal{S}_k} \mathbf{L}_{i,k} \mathbf{S}_k \mathbf{F}_k; & l = k \\ \mathbf{0}_{n_k^s \times n_l}; & otherwise. \end{cases} \quad (\text{D.1})$$

- Block diagonal matrices, $\mathbf{C}_t = \text{Blockdiag} [\mathbf{C}_{t,1} \cdots \mathbf{C}_{t,N}]$;
 $\mathbf{M}_t^{-1} = \text{Blockdiag} [\mathbf{M}_{t,1|t}^{-1} \cdots \mathbf{M}_{t,N|t}^{-1}]$; $\mathbf{O}^s = \text{Blockdiag} [\mathbf{O}_1^s \cdots \mathbf{O}_N^s]$ and
 $\mathbf{W}_t^f = \text{Blockdiag} [\mathbf{W}_{t,1}^f \cdots \mathbf{W}_{t,N}^f]$.

All the inner matrices of the above block structures are according to Chapter 7. Following equation 7.6, we get the change in Lyapunov function for the multi-agent system by

accumulating δV_k for all agents, $\delta V = \sum_{k=1}^N \delta V_k$. Therefore,

$$\begin{aligned}
\delta V &= - \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \mathbf{G}_{t,k} \boldsymbol{\eta}_{t-1,k} + 2 \sum_{k=1}^N \boldsymbol{\eta}_{t-1,k}^\top \left(\mathbf{C}_{t,k}^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \right) \mathbf{z}_{t-1,k} \\
&\quad + \sum_{k=1}^N \mathbf{z}_{t-1,k}^\top \left(\mathbf{O}_k^s \mathbf{W}_{t,k}^f \right)^\top \mathbf{M}_{t,k|t}^{-1} \mathbf{O}_k^s \mathbf{W}_{t,k}^f \mathbf{z}_{t-1,k} \\
&= - \boldsymbol{\eta}_{t-1}^\top \mathbf{G}_t \boldsymbol{\eta}_{t-1} - 2 \boldsymbol{\eta}_{t-1}^\top \mathbf{C}_t^\top \mathbf{M}_t^{-1} \mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f \boldsymbol{\eta}_{t-1} \\
&\quad + \boldsymbol{\eta}_{t-1}^\top \left(\mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f \right)^\top \mathbf{M}_t^{-1} \mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f \boldsymbol{\eta}_{t-1} \\
&= - \boldsymbol{\eta}_{t-1}^\top \mathbf{G}_t \boldsymbol{\eta}_{t-1} + \boldsymbol{\eta}_{t-1}^\top \left(\mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f - 2 \mathbf{C}_t \right)^\top \mathbf{M}_t^{-1} \mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f \boldsymbol{\eta}_{t-1} \tag{D.2}
\end{aligned}$$

Therefore, $\delta V < 0$ when, $\mathbf{O}^s \mathbf{W}_t^f \mathbf{A}^f = 2 \mathbf{C}_t$. Using the block matrix properties,

$$\mathbf{O}_k^s \mathbf{W}_{t,k}^f [\mathbf{A}^f]_{(k,l)^{th} block} = \begin{cases} 2 \mathbf{C}_{t,k}; & l = k \\ \mathbf{0}_{n_k \times n_l}; & l \neq k, l \in \mathcal{S}_k. \end{cases} \tag{D.3}$$

For all the neighbors of the k^{th} agent, the above relation can be written as,

$$\mathbf{O}_k^s \mathbf{W}_{t,k}^f \left[[\mathbf{A}^f]_{(k,k)^{th} block}, \dots, [\mathbf{A}^f]_{(k,l)^{th} block}, \dots \right] = 2 [\mathbf{C}_{t,k}, \dots, \mathbf{0}_{n_k \times n_l}, \dots]$$

Taking transpose,

$$\begin{bmatrix} [\mathbf{A}^f]_{(k,k)^{th} block}^\top \\ \cdot \\ \cdot \\ [\mathbf{A}^f]_{(k,l)^{th} block}^\top \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \left(\mathbf{O}_k^s \mathbf{W}_{t,k}^f \right)^\top = 2 \begin{bmatrix} \mathbf{C}_{t,k}^\top \\ \cdot \\ \cdot \\ \mathbf{0}_{n_l \times n_k} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{D.4}$$

Now, the matrix \mathbf{O}_k^s operates on $\mathbf{W}_{t,k}^f$ and creates zero rows at corresponding indices. Therefore, $(\mathbf{O}_k^s \mathbf{W}_{t,k}^f)^\top$ will have the columns of $(\mathbf{W}_{t,k}^f)^\top$ along with some zero columns. From the prior knowledge of \mathbf{O}_k^s we can find the indices i_k of nonzero columns such that $1 \leq i_k \leq n_k$ and $(\mathbf{O}_k^s \mathbf{W}_{t,k}^f)^\top[:, i_k] \neq \mathbf{0}_{n_k^s \times 1}$. These non-zero columns will constitute the transpose of desired weight matrix $\mathbf{W}_{t,k}^f$. Therefore,

$$(\mathbf{O}_k^s \mathbf{W}_{t,k}^f)^\top[:, i_k] = 2 \left(\sum_{l \in S_k} [\mathbb{A}^f]_{(k,l)th\ block} [\mathbb{A}^f]_{(k,l)th\ block}^\top \right)^{-1} [\mathbb{A}^f]_{(k,k)th\ block} \mathcal{C}_{t,k}^\top[:, i_k] \quad (\text{D.5})$$

Rearranging these columns as rows gives us $\mathbf{W}_{t,k}^f$. The solution can also be written as,

$$\mathbf{W}_{t,k}^f = 2 (\mathbf{O}_k^s)^\dagger \mathcal{C}_{t,k} [\mathbb{A}^f]_{(k,k)th\ block}^\top \left[\left(\sum_{l \in S_k} [\mathbb{A}^f]_{(k,l)th\ block} [\mathbb{A}^f]_{(k,l)th\ block}^\top \right)^{-1} \right]^\top \quad (\text{D.6})$$

The inverse exists if $\text{rank} \left(\sum_{l \in S_k} [\mathbb{A}^f]_{(k,l)th\ block} [\mathbb{A}^f]_{(k,l)th\ block}^\top \right) = n_k^s$.

Appendix E

SYS3 Parameters

- Global state transition matrix:

$$\mathbf{F} = \begin{bmatrix} 0.95, & 0, & 0 \\ 1, & 0.9, & 0 \\ 1, & 1, & 0.8 \end{bmatrix}$$

- Process noise: $\mathbf{Q} = \text{diag}(1.8, 0.9, 0.5)$.

- State initialization: $\boldsymbol{\mu} = [10, 5, 8]^\top$; $\boldsymbol{\Sigma} = \text{diag}(0.8, 0.2, 0.5)$.

- Agent 1:

$$\mathbf{H}_1 = [2, 0]; \mathbf{R}_1 = 0.0648; \mathbf{S}_1 = [0, 1]; \mathbf{O}_1^s = \mathbf{S}_1^\top; \mathbf{U}_1 = [1, 0]; \mathbf{O}_1^u = \mathbf{U}_1^\top;$$

$$\mathbf{P}_{2,1} = 1; \mathbf{L}_{2,1} = 1;$$

$$\mathbf{T}_1 = \begin{bmatrix} 1, & 0, & 0 \\ 0, & 1, & 0 \end{bmatrix}$$

- Agent 2: $\mathbf{H}_2 = [3, 0]; \mathbf{R}_2 = 0.05; \mathbf{S}_2 = \mathbf{U}_1; \mathbf{O}_2^s = \mathbf{O}_1^u; \mathbf{U}_2 = \mathbf{S}_1; \mathbf{O}_2^u = \mathbf{O}_1^s; \mathbf{P}_{1,2} = 1;$
 $\mathbf{L}_{1,2} = 1;$

$$\mathbf{T}_2 = \begin{bmatrix} 0, & 1, & 0 \\ 0, & 0, & 1 \end{bmatrix}$$

Appendix F

Expected Value in Equation (8.10)

From definition 8.1, $\underline{\mathcal{D}}_{t,k} = \phi_k(t)\mathcal{D}_{t,k} + (1 - \phi_k(t))\mathcal{B}_{t,k}$. For mutually exclusive events, $\gamma_k(t) = 1 - \phi_k(t)$. Therefore, $\gamma_k^2(t)\underline{\mathcal{D}}_{t,k} = \{\phi_k(t) - 2\phi_k^2(t) + \phi_k^3(t)\}\mathcal{D}_{t,k} + \{1 + 3\phi_k^2(t) - 3\phi_k(t) - \phi_k^3(t)\}\mathcal{B}_{t,k}$.

In other words,

$$\gamma_k^2(t)\underline{\mathcal{D}}_{t,k} = \begin{cases} \mathbf{0} & ; \phi_k(t) = 1 \\ \mathcal{B}_{t,k} & ; \phi_k(t) = 0 \end{cases} ; \forall k.$$

Hence, $\mathbb{E}[\gamma_k^2(t)\underline{\mathcal{D}}_{t,k}] = (1 - \rho_\phi)\mathcal{B}_{t,k}; \forall k$.

Appendix G

Positive Definiteness of $\mathbf{B}_{t,k}$

From Definition 9.1

$$\begin{aligned} \mathbf{B}_{t,k} &= \boldsymbol{\Upsilon}_{t,k} - \mathbf{F}_k^\top \boldsymbol{\Upsilon}_{t+1,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) \\ &\quad - (\mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) \end{aligned} \quad (\text{G.1})$$

From step 3 of Algorithm 5, $(\mathbf{U} + \mathbf{G}_k^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k) \mathbf{K}_{t,k}^c = -\mathbf{G}_k^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{F}_k$. Therefore, $(\mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) = -(\mathbf{K}_{t,k}^c)^\top \mathbf{U} \mathbf{K}_{t,k}^c$. Using these relations and Hamilton-Jacobi inequality, from equation (G.1),

$$\mathbf{B}_{t,k} \geq (\mathbf{K}_{t,k}^c)^\top \mathbf{U} \mathbf{K}_{t,k}^c + \mathbb{X}_k \quad (\text{G.2})$$

Hence, the positive definiteness of $\mathbf{B}_{t,k}$ is ensured.

Appendix H

Least Square Estimate of AMPC Weights

We start with the definitions of following block matrices,

- A block matrix \mathbf{K}_t^c ,

$$[\mathbf{K}_t^c]_{(k,l)^{th}block} = \begin{cases} -\mathbf{K}_{t,l}^c; & l \in \mathcal{S}_k \\ |\mathcal{S}_k| \mathbf{K}_{t,k}^c; & l = k \\ \mathbf{0}_{p \times n_l}; & otherwise. \end{cases} \quad (\text{H.1})$$

- Block diagonal matrices,

$$\mathbf{G} = \text{Blockdiag} [\mathbf{G}_k]; \quad \mathbf{\Xi}_t = \text{Blockdiag} [\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c];$$

$$\mathbf{\Upsilon}_{t+1} = \text{Blockdiag} [\mathbf{\Upsilon}_{t+1,k}]; \quad \text{and, } \mathbf{W}_t^c = \text{Blockdiag} [\mathbf{W}_{t,1}^c \cdots \mathbf{W}_{t,N}^c].$$

All the inner matrices of the above block structures are according to Chapter 9. Following equation 9.10, we get the change in Lyapunov function for the multi-agent system by accumulating δV_k for all agents, $\delta V = \sum_{k=1}^N \delta V_k$. Therefore,

$$\begin{aligned}
\delta V &= \sum_{k=1}^N \delta V_k \\
&= - \sum_{k=1}^N \mathbf{x}_{t,k}^\top \mathbf{B}_{t,k} \mathbf{x}_{t,k} \\
&\quad + 2 \sum_{k=1}^N \mathbf{x}_{t,k}^\top (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \boldsymbol{\xi}_{t,k} \\
&\quad + \sum_{k=1}^N \boldsymbol{\xi}_{t,k}^\top (\mathbf{G}_k \mathbf{W}_{t,k}^c)^\top \boldsymbol{\Upsilon}_{t+1,k} \mathbf{G}_k \mathbf{W}_{t,k}^c \boldsymbol{\xi}_{t,k} \\
&= - \mathbf{x}_t^\top \mathbf{B}_t \mathbf{x}_t - 2 \mathbf{x}_t^\top \boldsymbol{\Xi}_t^\top \boldsymbol{\Upsilon}_{t+1} \mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c \mathbf{x}_t \\
&\quad + \mathbf{x}_t^\top (\mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c)^\top \boldsymbol{\Upsilon}_{t+1} \mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c \mathbf{x}_t \\
&= - \mathbf{x}_t^\top \mathbf{B}_t \mathbf{x}_t \\
&\quad + \mathbf{x}_t^\top \left[(\mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c)^\top - 2 \boldsymbol{\Xi}_t^\top \right] \boldsymbol{\Upsilon}_{t+1} \mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c \mathbf{x}_t
\end{aligned} \tag{H.2}$$

Therefore, $\delta V < 0$ when, $\mathbf{G} \mathbf{W}_t^c \mathbf{K}_t^c = 2 \boldsymbol{\Xi}_t$. Using the block matrix properties,

$$\mathbf{G}_k \mathbf{W}_{t,k}^c [\mathbf{K}_t^c]_{(k,l)^{th} block} = \begin{cases} 2(\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c); & l = k \\ \mathbf{0}_{n_k \times n_l}; & l \neq k, l \in \mathcal{S}_k. \end{cases} \tag{H.3}$$

For all the neighbors of the k^{th} agent, the above relation can be written as follows,

$$\mathbf{G}_k \mathbf{W}_{t,k}^c \left[[\mathbf{K}_t^c]_{(k,k)^{th} block}, \dots, [\mathbf{K}_t^c]_{(k,l)^{th} block}, \dots \right] = 2 \left[(\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c), \dots, \mathbf{0}_{n_k \times n_l}, \dots \right]$$

The least-square solution is then given by,

$$\mathbf{W}_{t,k}^c = 2 (\mathbf{G}_k)^\dagger (\mathbf{F}_k + \mathbf{G}_k \mathbf{K}_{t,k}^c) [\mathbf{K}_t^c]_{(k,k)^{th} block}^\top \left[\left(\sum_{l \in \mathcal{S}_k} [\mathbf{K}_t^c]_{(k,l)^{th} block} [\mathbf{K}_t^c]_{(k,l)^{th} block}^\top \right)^{-1} \right]^\top \tag{H.4}$$

We get this solution with assumptions (1) $\text{rank} \left(\sum_{l \in \mathcal{S}_k} [\mathbb{K}_t^c]_{(k,l)^{\text{th block}}} [\mathbb{K}_t^c]_{(k,l)^{\text{th block}}}^\top \right) = p$, and (2) $p \leq n_k, \forall k$.