Distributed Estimation of Dynamical Systems: A Structural Analysis

A dissertation submitted by

Mohammadreza Doostmohammadian

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

Tufts University

May 2015

Adviser: Prof. Usman A. Khan

Abstract

Distributed estimation is where a network of agents is tasked to estimate the state of a dynamical system. Agents only communicate over a *sparse* communication network. Recently, consensus-based estimation has been proposed as a distributed solution of this problem where the agents implement a large number of information exchanges between every two successive time-steps of the system dynamics. For optimal performance, this consensus-based estimator requires a consensus to be reached first. When the network is unable to implement a consensus due to, e.g., resource-constraints or faster system dynamics, distributed solutions have been proposed with single-time information exchanges. In this scenario both system dynamics and distributed estimator evolve at the same time-scale. This scenario requires the system to be observable at every estimator/agent, implying the new concept of *distributed observability*. Given this background, this thesis is devoted to (1) formulation of distributed observability in single-time scale estimation, (2) partitioning the necessary set of state measurements based on their role in distributed observability, and (3) characterization of necessary and sufficient connectivity of the underlying communication network topology among the agents. Employing structure-based generic methodology instead of algebraic approaches motivates application in power systems and social networks.

Acknowledgements

I would like to thank my adviser Professor Usman Khan for all his support during my PhD studies, as well as all his inspiring ideas and warm encouragement. He was a great mentor to me both in research and teaching. I appreciate all his useful comments and suggestions to hone my writing and perfect my presentation skills. I was very fortunate to have him as my adviser, as my mentor, and more important as my friend.

I am very thankful of my committee members, Professor Carlo Fischione from KTH, Professor Jason Rife from Tufts University, Mechanical Engineering Department, Professor Brian Tracy and Alexander Stankovic from Tufts University, Electrical Engineering Department. I appreciate all their useful suggestions and insightful comments to enhance the quality of this work. I am also grateful to my wonderful friends and colleagues at Tufts University, whose support and friendship always helped me overcome the challenges in my life.

Finally and most importantly, by dedicating this thesis to my family, I would like to thank them for all their endless love and support. My mother Hosni, the real source of love and energy, my father Aghajani the greatest support throughout my life, my lovely sisters Azam, Najme, Sorayya and my handsome brother Alireza who were always with me when I needed them. I was blessed to grow up in such a happy family with the most kind parents in the world.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Estimation of linear dynamical systems is a thriving field of research, pioneered by Kalman filtering. However, the complexity of system dynamics, high-dimensional state-space, and the diversity of available sensing methodologies mean that extensive computation is required to implement and analyze traditional estimation techniques. In practice, it is desirable to implement a scalable estimator that is robust to system perturbations and is computationally efficient. In this context, distributed (or networked) estimation provides scalable, efficient, and robust solution to instrument the sensing measurements all-together. This allows the system state to be estimated in a collaborative way without relying on a central computation entity. In a centralized scheme, each agent (or sensor) makes a *local* observation of the physical system, and sends this measurement to a central unit where an optimal estimator may be implemented. In contrast, in a distributed scenario, each agent makes a *local* estimate of the system state, shares its measurement and/or estimates with nearby agents, and combines the received information to improve its local estimate.

As an example, consider a scenario where a collection of agents (e.g. observers/sensors/robots) are assigned to estimate a system or a phenomenon of interest. Agents are distributed in the sense that each agent can only measure some of the states of the dynamical system. For example, a group of sensors spread geographically over a large region to monitor daily temperature evolution. The measurement data and dynamical models are further corrupted by noise and disturbances. Clearly, a centralized scheme may be impractical for a large-scale system. Therefore, in distributed fashion, the objective is to enable each agent to make decision on the *global state* relying only on its own measurement and the measurements of its immediate neighbors. Such a scheme is often referred to as networked estimation or distributed estimation where the term network implies that the information is restricted on a sparse network.

Consensus-based distributed estimation strategies have recently gained a lot of interest, where the main focus is to reduce the uncertainty of individual estimates by averaging on the measurements. Consensus protocols [1–4] define interaction rules among a network of agents to combine their information. These interactions primarily are defined over a graph (network), where the existence of an edge (communication link) between two nodes (agents) implies the flow of information from one agent to the other. Early work in [5–9] considers a two time-scale method, where consensus is implemented at a time-scale different than the system dynamics (see Fig. 1.1), where a large number ($\rightarrow \infty$) of data fusion iterations are implemented between every two successive time-steps, k and $k+1$, of the system dynamics. This approach requires communication over a much faster rate than the sampling of the dynamics, and thus, in general, becomes practically in-feasible when the underlying system is operating under power constraints and has restricted communication and computation budgets.

The key point is that in the two time-scale method (see Fig. 1.1–(Left)), the com-

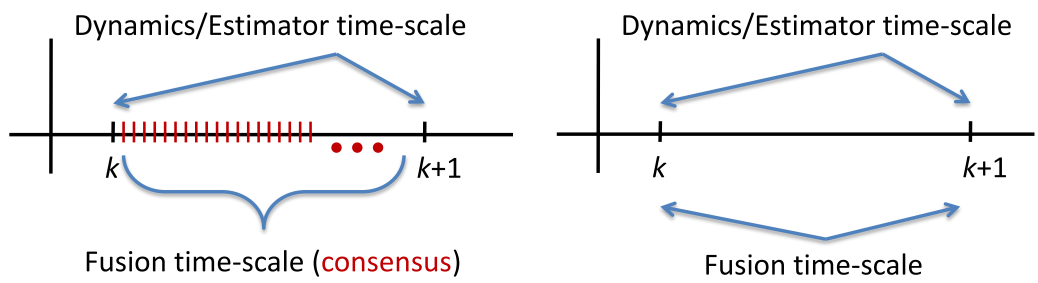

Figure 1.1: (Left) The traditional two time-scale consensus-based approach; each small dash represents one step of consensus/communication. (Right) single time-scale approach.

munication network becomes irrelevant due to more information exchanges among the individuals. This is because the information in a sparsely connected graph is equivalent to the information in a fully connected graph when a large number of information exchanges are carried out. For instance consider a network with diameter d, i.e. the maximum. This implies that more than d steps of communications and exchange of measurements between every two successive steps of system dynamics, conveys all taken measurements to each agent. Therefore, the performance and properties of the underlying estimator depends only on the data fusion principles among the agents. However, in the single time-scale scenario of Fig. 1.1–(Right), the underlying agent network remains sparse and an arbitrary communication network may not suffice to make the distributed estimation error stable (e.g., see [10,11]). This is where the concept of *system observability* plays the key role. Observability is a measure to quantify inference of internal (e.g. not accessible) states of the system based on the measurements of external (accessible) states. This is because states of a dynamical system are typically dependent and under influence of their *neighboring* states. If system is fully observable, it implies that a given set of measurements contain enough information to reconstruct the global state of the system [12–14].

In centralized estimation and observability, transmitting all measurement data makes the system globally observable to a central unit/processor assuming that the system is observable. Similar argument holds for multi-time scale distributed estimation (with number of consensus/communication step more than network diameter [9]). However, in the single-time scale estimator, at each step only *local* measurements are available to each agent. This local information may not contain necessary information to guarantee observability. In this context, the key problem is to *design* the structure of the multi-agent communication network according to the underlying fusion rules in order to *recover* the distributed observability, i.e. to make each estimator locally observable. A simple approach to solve this problem is to share all the necessary[1] measurements for observability at every step of system dynamics [15, 16]. The more challenging approach is to communicate *both measurements and predictions* to impose less communication among the agents. We show that the latter approach requires less connectivity in the agent network. A related study on this is carried out in [10], where a particular distributed estimator are shown to have bounded MSEE if the two-norm of system matrix is less than the Network Tracking Capacity (NTC). This quantity is a function of the communication network and system measurement model.

The next challenge that we address in this work, and in general, in real world systems, is the time varying nature and uncertainty of the system parameters. For example in power systems and social networks, the *structure* of the system remains time-invariant but the system parameters (values of electrical components) are subject to perturbations. This motivates us to implement methodologies that are indepen-

[1] In this work, necessary, critical, and crucial are interchangeably used for measurements/agents.

dent of exact parameter values of such systems that we refer to as Linear Structure-Invariant (LSI) systems. Our approach only relies on the underlying system structure, i.e., the zero and non-zero pattern of the system matrix. Such properties lead to robust observer design where the analysis is graph-theoretic rather than the traditional algebraic approach. This is the case when some physical quantities in the system change over time, as well as, in linearization of nonlinear models for which the parameters depend on the system operating point. In this sense, structured system theory is beneficial for analysis of system properties such as controllability and observability [15–21]. Using such graph-theoretic techniques, we tackle uncertainty issues in our analysis. It is noteworthy that the structurally-defined results based on the linearization is applicable for the structural observability of nonlinear models. Indeed, as mentioned in [20, 22], generic analysis holds for *smooth* nonlinear systems with fixed structured Jacobian representing LSI matrices . In other words, structural observability leads to observability over a continuum of system operating points in nonlinear cases.

1.1 Related Work

A variety of solutions exists for distributed estimation pioneered by the earlier work [23, 24], and references therein on Parallel Kalman Filtering architectures for *all-to-all* connected networks, to more recent consensus-based protocols [5]. The latter was also referred to as a *two time-scale* approach, Fig. 1.1–(a). In contrast to the two time-scale distributed estimation, recently Refs. [9, 10, 25–32] study the behavior of distributed estimators when communication and dynamics time-scale are the same, as shown in Fig. 1.1–(b). Clearly, this method is practically feasible for real-time applications and computationally efficient as compared to the two time-scale approach .

Observability analysis of the centralized estimation is primarily introduced in [12, 33]. Recently, the works by [17, 34] applied a structural approach for observability analysis. Based on the same approach, [15,19,20] find the critical set of measurements for centralized observability. The same observability analysis works for the two time-scale estimators as all measurements are accessible between every two-steps of system dynamics. However, the observability analysis for single time-scale estimators is more challenging, as discussed in [27, 30] for particular distributed single-time estimators.

The literature can also be classified into static and dynamic estimation. In static estimation [25, 26, 29, 35, 36] or quasi-static estimation [32], the target state to be estimated does not change over time. A similar case is when system state is relatively stationary (quasi-stationary) over a period to allow the adaptive algorithm to converge [26], i.e. the evolution-rate of the target state is not too fast. On the other hand, dynamic estimation [5–9, 23, 37–39] takes the (non-stationary) time-evolution of the system into account. Typically, these works make a pre-assumption on the communication network to be (strongly) connected [6, 11, 26, 32, 35, 36, 39] or for it to include a cycle path connecting through all agents [25, 29].

Chapter 2 of this thesis provides more review of the literature, and further, preliminary concepts prerequisite for the rest of the thesis. The same chapter includes consensus protocols and average consensus for information processing, structured system theory and generic analysis, and specifically, generic observability, structural rank, accessibility and matching conditions. It further reviews background on centralized estimation problem and distributed methods.

1.2 Problem Notation and Formulation

Consider distributed estimation of a discrete-time linear dynamics of the form:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{v}_k, \tag{1.1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector[2], $A = \{a_{ij}\} \in \mathbb{R}^{n\times n}$ is the system matrix, and $\mathbf{v}_k \sim \mathcal{N}(0, V)$ is the system noise.

$$\mathbf{x}_k = \begin{pmatrix} x_k^1 \\ \vdots \\ x_k^n \end{pmatrix}, \quad \mathbf{v}_k = \begin{pmatrix} v_k^1 \\ \vdots \\ v_k^n \end{pmatrix}, \tag{1.2}$$

For estimation, we assume N measurements of system states are taken as:

$$\mathbf{y}_k^i = H_i\mathbf{x}_k + \mathbf{r}_k^i, \tag{1.3}$$

where $\mathbf{y}_k^i \in \mathbb{R}^{p_i}$, $i \in \{1, ..., N\}$ is the output vector[3] at agent i, $\mathbf{r}_k^i \sim \mathcal{N}(0, R_i)$ is the output noise, and H_i is the output matrix at agent i. With this notation, we can write the global observation model as:

$$\mathbf{y}_k = H\mathbf{x}_k + \mathbf{r}_k, \tag{1.4}$$

[2]As a general notation notice, we use *boldface* letters for *vectors* and plain italic letters for *scalar* variables, and capital italic for *matrices*.

[3]In this thesis, without loss of generality, we assume $p_i = 1$. Simply, if $p_i > 1$ we may consider p_i observation $\mathbf{y}^i \in \mathbb{R}$, $i \in \{1, ..., p_i\}$

where $\mathbf{y}_k$, $H = \{h_{ij}\}$, and $\mathbf{r}_k$ are collections of the local variables.

$$\mathbf{y}_k = \begin{pmatrix} \mathbf{y}_k^1 \\ \vdots \\ \mathbf{y}_k^N \end{pmatrix}, \quad H = \begin{pmatrix} H_1 \\ \vdots \\ H_N \end{pmatrix}, \quad \mathbf{r}_k = \begin{pmatrix} \mathbf{r}_k^1 \\ \vdots \\ \mathbf{r}_k^N \end{pmatrix}, \quad R = \begin{pmatrix} R_1 & & 0 \\ & \ddots & \\ 0 & & R_N \end{pmatrix}. \tag{1.5}$$

Let $\widehat{\mathbf{x}}^c_{k|k}$ represent the state of the Centralized Kalman Filter (CKF) estimator at time k given all the observations, y_k, up to time k. It can be shown that the estimation error (including the transient error) is given by

$$\widehat{\mathbf{e}}^c_{k|k} \triangleq \mathbf{x}_k - \widehat{\mathbf{x}}^c_{k|k}, \tag{1.6}$$

$$= (A - K^c_k HA)\widehat{\mathbf{e}}^c_{k-1|k-1} + \eta_k, \tag{1.7}$$

where K^c_k is the appropriate centralized Kalman gain and η_k represents the noise terms independent of $\widehat{\mathbf{e}}^c_{k|k}$. The estimation error, $\mathbf{e}_{k|k}$, in (1.6) is stable if and only if the system is (A, H)-observable.

We now state in detail the problems considered in this thesis. *The first problem* is to find the measurement matrix H (or the set of states to be measured denoted as $\mathcal{Y}$) such that the pair (A, H) is observable. Our contribution is to further partition these critical measurements based on their specific role in system observability.[4] An extension of this problem is to find the set of *equivalent* measurements for observability. The concept of (observational) equivalency implies that two different measurements,

[4]In *Chapter 3* we discuss these measurements in details. We classify and name these measurements and their assigned observer as Type-α and Type-β.

for example,

$$\begin{cases} \mathbf{y}_i = H_i\mathbf{x} + \mathbf{r}_i, \\ \mathbf{y}_j = H_j\mathbf{x} + \mathbf{r}_j. \end{cases} \tag{1.8}$$

play the same role for observability. Mathematically, it can be formally characterized as,

$$(A, H_i)\text{-observability} \Longleftrightarrow (A, H_j)\text{-observability} \Longleftrightarrow \left(A, \begin{pmatrix} H_i \\ H_j \end{pmatrix}\right)\text{-observability} \tag{1.9}$$

We find observational equivalence sets for different measurement types. Indeed, for each measurement $\mathbf{y}_i$, we find the *set* of measurements equivalent to $\mathbf{y}_i$ such that only one measurement in this set is required to ensure observability.

Next, for distributed estimation, we assume that the agents communicate over a communication network, $\mathcal{G}_{net} = (\mathcal{V}_{net}, \mathcal{E}_{net})$, where $\mathcal{V}_{net}$ includes the nodes representing the agents and $\mathcal{E}_{net}$ includes the edges; an edge $j \rightarrow i$ represents a communication link from agent j to agent i. The adjacency matrix of this network is defined by matrices $W = \{w_{ij}\}$ and $U = \{u_{ij}\}$. For example, for W we have,

$$\begin{cases} w_{ij} \neq 0, & j \rightarrow i, \\ 0, & \text{otherwise}. \end{cases} \tag{1.10}$$

The neighborhood at agent i is defined as

$$\mathcal{N}(i) = \{i\} \cup \{j \mid j \rightarrow i\}. \tag{1.11}$$

An edge (link) from node j to node i implies that agent i transmits information to agent j. In this sense, $\mathcal{N}(i)$ includes all the agents that *send* information to agent i. The neighborhood is defined based on adjacency matrices W and U. We discus these matrices in detail in *Chapters 4 and 5*. Clearly, due to the interaction, each agent i now estimates the state, $\mathbf{x}_k$, with its own information, including $\mathbf{y}_k^i$, and with its neighboring information, including observations $\mathbf{y}_k^j, j \in \mathcal{N}_i$. Now, let $\widehat{\mathbf{x}}_{k|k}^i$, be the estimate of the state, $\mathbf{x}_k$, by all information available to agent i up to time k through the interaction graph, $\mathcal{G}_{net}$. Define the local error at agent i as,

$$\mathbf{e}_{k|k}^i = \mathbf{x}_{k|k} - \widehat{\mathbf{x}}_{k|k}^i, \tag{1.12}$$

Concatenating the estimates at all agents, the *global* state estimate and error in the agent network is given by,

$$\underline{\widehat{\mathbf{x}}}_{k|k} \triangleq \begin{pmatrix} \widehat{\mathbf{x}}_{k|k}^1 \\ \widehat{\mathbf{x}}_{k|k}^2 \\ \vdots \\ \widehat{\mathbf{x}}_{k|k}^N \end{pmatrix}, \quad \mathbf{e}_{k|k} = \begin{pmatrix} \widehat{\mathbf{e}}_{k|k}^1 \\ \widehat{\mathbf{e}}_{k|k}^2 \\ \vdots \\ \widehat{\mathbf{e}}_{k|k}^N \end{pmatrix}. \tag{1.13}$$

The second problem is to mathematically formulate distributed observability from first principles. We prove that distributed observability is characterized as observability of the pair:

$$(W \otimes A, D_H) \tag{1.14}$$

The matrix D_H is a block-diagonal matrix defined as,

$$D_H = \begin{pmatrix} \sum_{j\in\mathcal{N}(1)} H_j^T H_j & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sum_{j\in\mathcal{N}(N)} H_j^T H_j \end{pmatrix}. \tag{1.15}$$

where the neighborhood is defined based on based on the adjacency of U matrix. Next, we provide a Networked Kalman-type Estimator (NKE):

$$\widehat{\mathbf{x}}^i_{k|k-1} = \sum_{j\in\mathcal{N}(i)} w_{ij} A \widehat{\mathbf{x}}^j_{k-1|k-1}, \tag{1.16}$$

$$\widehat{\mathbf{x}}^i_{k|k} = \widehat{\mathbf{x}}^i_{k|k-1} + K^i_k \sum_{j\in\mathcal{N}(i)} H_j^T \left(y^j_k - H_j \widehat{\mathbf{x}}^i_{k|k-1} \right), \tag{1.17}$$

We show that the estimator error of the NKE estimator evolves as follows:

$$\mathbf{e}_k = (W \otimes A - K_k D_H (W \otimes A))\mathbf{e}_{k-1} + \mathbf{q}_k, \tag{1.18}$$

where K_k is the block diagonal gain matrix and q_k contains noise terms independent of the $\mathbf{e}_{k|k}$. We verify that the above error evolution is steady-state stable if and only if $(W \otimes A, D_H)$ is observable.

The third problem is to define necessary connectivity of the multi-agent network $\mathcal{G}_W$ to ensure $(W \otimes A, D_H)$ observability. Notice that the structure of the network is tied with the adjacency matrix W and U. In the sufficiency case, we design $\mathcal{G}_{net}$, for a given set of measurements $\mathcal{Y}$ satisfying (A, H)-observability. Since we aim to solve the problem generically, the observability analysis is performed on system graph. We provide combinatorial graphical algorithms of *polynomial order* to check for system

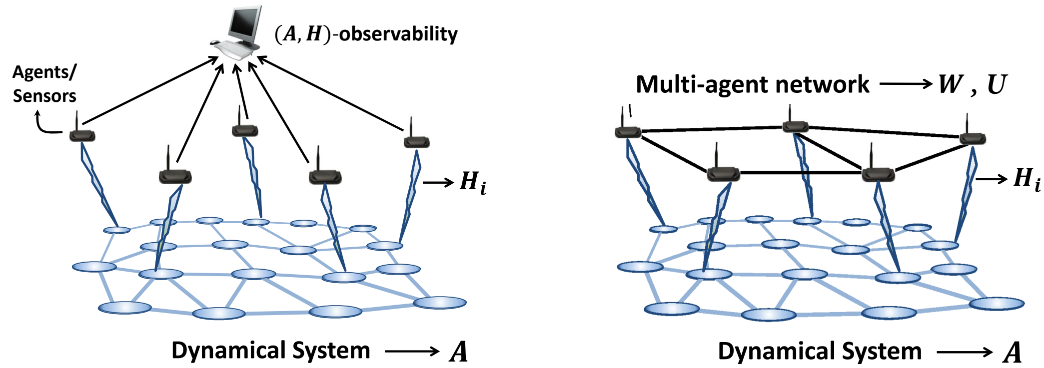

Figure 1.2: (Left) Centralized estimation vs. (Right) Distributed estimation.

observability, to find the necessary set of measurements and observationally equivalent sets. In the same way, we aim to graphically design the *structure* of the agent network to ensure distributed observability. This mathematical model can be implemented, for example, by a wireless communication network where the sampling time is long enough as compared to, for example, coherence time of the wireless communication channel [32]. This point could be considered another drawback for the multi time scale estimation where high rate of communication potentially encounters more packet loss. To avoid this, reliable communication channels with smaller coherence time is required in multi time scale scenario.

We illustrate the distributed and centralized estimation problem in Fig. 1.2. It should be mentioned that in this thesis we design the communication network such that all agents can track the *global* state of the dynamical system. Therefore the observability condition could not be checked locally as we assume that the system is not observable at any agent or in its direct neighborhood. In terms of implementation, global inference/tracking requires the knowledge of the entire system matrix, while the measurement matrix is distributed among the agents. This is for any distributed estimation protocol claiming global state inference. The other scenario could be to estimate a local portion of system states. For example, in a spatially distributed system, each agent might be interested to only monitor its neighboring area instead of the entire system. In partial inference, the observability condition might be checked locally at each agent.

1.2.1 Notations

In this section we list all the notations in Table 1.1. Some of these notations may be introduced in the next chapters.

1.2.2 Assumptions

In the rest of this thesis we make these general assumptions:

(i) The communication between the agents is stable (static), i.e., the network topology is time-invariant;

(ii) We impose *no hierarchy* in the multi-agent network, i.e., we assume that the processing/communication capabilities of all agents are the same;

(iii) The system is globally (A, H)-observable;

(iv) For every agent, i, the pairs, (A, H_i) or $(A, \sum_{j \in \mathcal{N}(i)} H_j^T H_j)$, are not necessarily observable.

(v) The system is *dynamic* and in general it might be unstable, i.e. its spectral radius is greater than 1 ($\rho(A) > 1$).

(vi) We assume no packet loss in communication channels.

Table 1.1: List of Notations.

$\mathbf{x}$	System state vector	k	Sampling-time/Time-step
n	Number of states	N	Number of measurements
A	System matrix	$\mathcal{A}$	Structured system matrix
$\mathbf{v}$	System noise	$\mathbf{y}$	Measurement vector
H	Measurement matrix	$\mathcal{H}$	Structured measurement matrix
$\otimes$	Kronecker matrix product	$\mathbf{r}$	Measurement noise
K	Gain matrix	$\underline{K}$	Block-diagonal gain matrix
$\widehat{\mathbf{x}}$	State estimate	$\widehat{\mathbf{e}}$	Estimation error
$\mathcal{G}$	Graph	$\mathcal{E}$	Set of edges
$\mathcal{V}$	Set of nodes	$\mathcal{G}_{net}$	Network of agents/observers
$\mathcal{X}$	Set of state nodes	$\mathcal{Y}$	Set of measurement nodes
$\mathcal{G}_{\text{sys}}$	System digraph	$\mathcal{G}_{Dist}$	Distributed system digraph
$\longrightarrow$	Direct link in graph	$\xrightarrow{\text{path}}$	Path in graph
W	Prediction-fusion matrix	U	Measurement-fusion matrix
$\widehat{\mathbf{x}}^i_{k\|k}$	Estimate of agent i at time k given information by time k	$\mathcal{N}(i)$	Neighborhood of agent i
D_H	Global measurement matrix	$\overline{D}_H$	Global measurement matrix (no information fusion)
$\mathcal{O}$	Observability Gramian	Γ_A	Bipartite graph
$\mathcal{V}^-$	Set of end nodes in Γ_A	$\mathcal{V}^+$	Set of start nodes in Γ_A
$\mathcal{M}$	Maximal matching	$\Gamma_A^{\mathcal{M}}$	Auxiliary graph
$\delta\mathcal{M}^+$	Unmatched nodes in $\mathcal{V}^+$	$\mathcal{C}$	Contraction set
$\mathcal{S}$	Set of states in SCC	$\mathcal{S}^p$	Parent SCC
$\mathcal{S}^c$	Child SCC	$\mathcal{S}^{\circlearrowleft}$	Matched SCC
α,β,γ	Measurement/Agent types	$\mathcal{N}_\alpha$	Neighboring α agents
$\mathcal{N}_\beta$	Neighboring β agents	$\mathcal{G}_\alpha$	Graph of α agents
$\mathcal{G}_\beta$	SC Graph of β agents	$\mathcal{G}^*_\alpha$	Graph of β agents (Type II)
$\mathcal{G}_0$	Graph of self-cycles	$\mathbf{z}$	Global measurement
$\underline{\mathbf{x}}$	Global state	$\rho(.)$	Spectral radius of matrix
$\mathbb{N}(.)$	Gaussian distribution	H_γ	Matrix of non-critical measurements
H_α	Matrix of α measurements	H_β	Matrix of β measurements
n_α	Number of α states	n_β	Number of β states

Assumption (ii) increases the reliability of node/link failure by avoiding agent hierarchy. Assumption (iii) is a typical assumption in distributed estimation implying the observability of centralized estimator; without this, no estimation scheme works. Assumption (iv), in practice, makes the networked estimation problem more challenging. This is where this work becomes significantly different from current approaches. Assumption (v) is another challenge in distributed estimation. The reason is that for stable systems, the stability analysis of the estimation error is not related, since all states and estimates eventually reach the stability (boundary). In other words, for stable systems the MSEE converges to stability boundary for any estimation protocol. In terms of implementation, assumption (vi) is practically feasible where the dynamic sampling is long as compared to, for example, coherence time of the wireless communication channel. This allows sufficient time to re-transmit the wrong packet till one received successfully.

1.3 Contributions and Thesis Organization

We summarize the contributions of this thesis in Fig.1.3, and in more details in the following:

- **In Chapter 3, we define measurement partitioning and observational equivalence in system estimation.** Using the related graph-theoretic concepts, we derive the necessary set of states required to ensure LTI state-space observability (in both distributed and centralized cases). Further, we show that the set of crucial measurements required for *centralized observability* can be subdivided into two types, based on their role in *distributed observability.* This

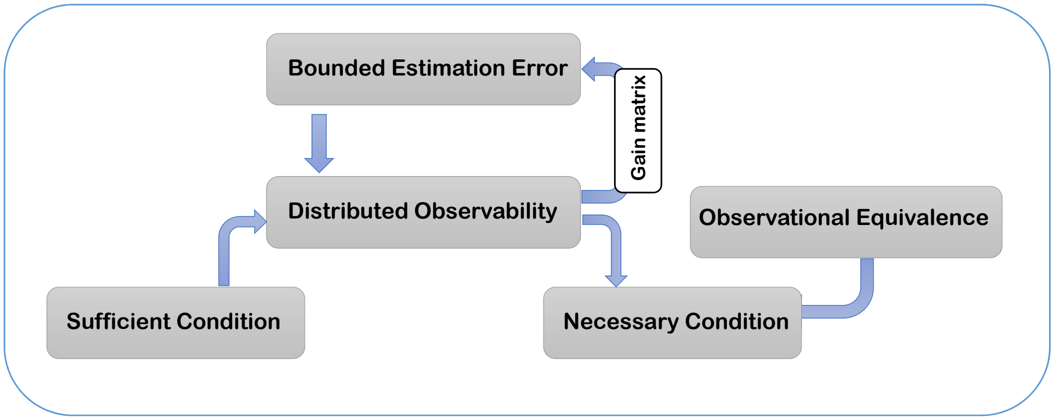

Figure 1.3: Thesis contribution.

partitioning is driven by both graphical and algebraic methods used to define the corresponding measurements. Using both graphical and algebraic methods, we find the observational equivalent sets. These equivalence set of measurements are applicable to recover for the loss of observability.

- **In Chapter 4, we mathematically derive distributed observability from the first principles.** In particular, we extend the centralized estimation setup to a distributed framework where in addition to the state and sensing, we also have communication among the agents. Combining the estimates at all of the sensors, we arrive at the networked estimator, which subsequently results into the corresponding networked dynamics. We show that the networked dynamics are not just a mere extension of the original dynamics repeated (block-diagonally) to accommodate for each sensor, but belongs to a large class of systems that naturally defines the allowable collaboration among the agents. Further, we cast the NKE estimator with two recursive measurement-fusion and prediction-fusion levels, where adding the prediction step is a distinction from the existing works. We verify that the error stability of the NKE estimator leads to the characterized distributed observability.

- **In Chapter 5, we derive the necessary and sufficient conditions for distributed observability and we *design* the multi-agent network accordingly.** We define two graph topology determining the connectivity of different types of agents and their role to recover observability. The communication network among the agents $\mathcal{G}_{net}$ is the union of these two graphs. The proposed condition on network sparsity is in contrast with the current densely wired networks in the literature [16, 40]. Particularly, for NKE estimator, we perform measurement- and prediction-fusion over different graphs. Moreover, we provide a system classification based on the system rank, and we compare the network connectivity of the distributed observer for (structurally) full rank and rank deficient systems.

- **In Chapter 6, we propose application of observational equivalency to recover the observability of power systems. Further, we show applicability of the distributed estimation/observability setup in social network inference.** The graph-theoretic approach in this work gives scalable and computationally efficient algorithms for distributed estimation over large-scale social systems, while the time-varying but structured nature of power systems further motivates application of our robust observability analysis.

In each chapter, we give graphical examples and/or simulations to illustrate the problem and support the results. *Chapter 7* of this thesis concludes the results and restates the contributions. In the appendix, we provide combinatorial algorithms (of polynomial order) for graph-theoretic analysis of this work. Also, the *block-diagonal feedback gain matrix* to stabilize the MSEE of the NKE (see Fig.1.3) is given in the appendix. The results of this thesis are published in the peer-reviewed conferences and journals [41–49].

Chapter 2

Background

In this chapter, we provide the general background and preliminaries on graph theoretic observability and distributed estimation protocols. Recall from Chapter 1 that we are interested to estimate discrete-time linear (structured) time-invariant (DT LTI) dynamical systems in the form:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{v}_k, \tag{2.1}$$

$$\mathbf{y}_k^i = H_i\mathbf{x}_k + \mathbf{r}_k^i, \tag{2.2}$$

First we state the centralized case where all the measurements are conveyed to a central processor. We state the Centralized Kalman Filter (CKF), its tracking error dynamics, and the general requirement for it to be bounded at steady state. In this regard, the concept of observability and different approaches to check for it are discussed. In particular, structured system theory and related basic graph notions are studied in this chapter. These concepts establishes the basis of the graph theoretic approach in this work; more advanced graph notions are postponed to Chapter 3.

Next, we review the common distributed estimation protocols in the literature. These notions along with other related preliminaries, such as Kronecker product of matrices, are employed in the next chapters to derive the main results of the thesis.

2.1 Introduction to graph theory

Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, \ldots, y_p\}$ denote the set of states and measurements, respectively. The *system digraph* is a directed graph defined as $\mathcal{G}_{\text{sys}} = (\mathcal{V}_{\text{sys}}, \mathcal{E}_{\text{sys}})$, where $\mathcal{V}_{\text{sys}} = \mathcal{X} \cup \mathcal{Y}$ is the set of nodes and $\mathcal{E}_{\text{sys}}$ is the set of edges; this digraph is induced by the structure of the system and measurement matrices, $A = \{a_{ij}\}$, and $H = \{h_{ij}\}$. An edge, $x_j \rightarrow x_i$, in $\mathcal{E}_{\text{sys}}$ exists from x_j to x_i if $a_{ij} \neq 0$. Similarly, an edge, $x_j \rightarrow y_i$, in $\mathcal{E}_{\text{sys}}$ exists from x_j to y_i if $h_{ij} \neq 0$. A *path* from x_j to x_i (or y_i) is a sequences of nodes originating from x_j and terminating to x_i (or y_i) with each subsequent edge in $\mathcal{E}_{\text{sys}}$. Denote such a path as $x_j \xrightarrow{\text{path}} x_i$. A path is called $\mathcal{Y}$*-connected*, denoted by $\xrightarrow{\text{path}} \mathcal{Y}$, if it terminates in a measurement/output node (i.e., in a measured state node). Here, we assume that each node is included in a path only once (a *simple* path). A *cycle* is a path where the originating and terminating nodes are the same. A cycle family is a group of cycles which are mutually disjoint, i.e. they don't share any node. Similarly, a path and a cycle are disjoint if they do not share any node.

Here, we provide some useful graph properties. A directed graph $\mathcal{G}_{\text{sys}}$ is Strongly-Connected (SC) if every two nodes are connected by a path, i.e., $x_i \xrightarrow{\text{path}} x_j$ for every $x_i, x_j \in \mathcal{X}$. A non-SC graph is called *acyclic*. In an acyclic digraph define Strongly Connected Components (SCCs), denoted as $\mathcal{S}_i$, as its maximal strongly connected partitions or sub-graphs. A directed tree is a directed graph where every node has exactly one incoming link, except the root of the tree (also known as the leader) which

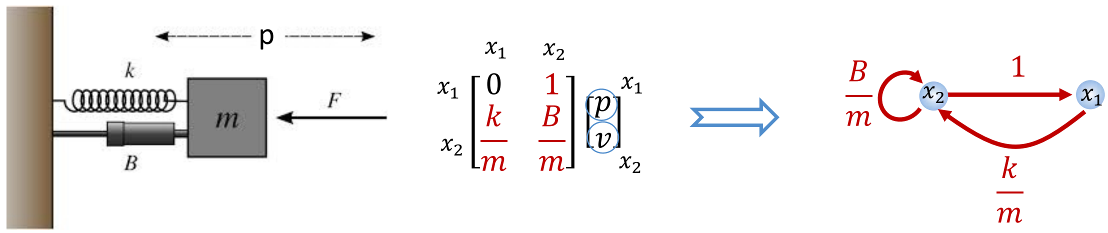

Figure 2.1: A dynamical system, the state-space representation and digraph representation. The cycle in graph represents the dynamical mode of the system.

has no incoming link. A graph contains a *spanning tree* if a subset of edges form a directed tree that connects (spans) all nodes. A dynamical system of states can be represented as a graph. Each node represents a dynamic state and the edges represent the interaction of different states governed by a dynamic equation, see Fig. 2.1. For example, representing two system states x_i and x_j as two nodes in the graph, an edge from x_i to x_j implies $\dot{x}_j = f(x_i)$ in continuous time model or $x_j(k+1) = f(x_i(k))$ in discrete time model where $f(.)$ represents a smooth function. The discussed graph features each imply a property of dynamic system. For example, an SC graph implies an irreducible system matrix and a cycle in the system digraph represents a dynamical mode of the system. We refer interested readers to [50, 51] for more detailed explanation on graph concepts.

2.2 Kronecker Product of Matrices

The Kronecker product of matrices are known to have implications in network theory and estimation. In [52], authors show that the graph associated to Kronecker product of matrices, termed as *Kronecker graphs*, mimics the structural properties of large real networks. Such recursively generated graphs are statistically proved to have properties, e.g. diameter and degree distribution, matching those of real networks. Further, Spectral analysis of Kronecker matrices with an application in LTI observability is primarily stated in [53]. In this thesis, we apply this concept to derive the *distributed* observability condition in Chapter 4.

Algebraically, the Kronecker product of two matrices $W_{N\times N}$ and $A_{n\times n}$ is defined as,

$$W \otimes A = \begin{pmatrix} w_{11}A & w_{12}A & \cdots & w_{1N}A \\ w_{21}A & w_{22}A & \cdots & w_{2N}A \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1}A & w_{N2}A & \cdots & w_{NN}A \end{pmatrix} \tag{2.3}$$

where each block matrix $w_{ij}A$ is,

$$w_{ij}A = \begin{pmatrix} w_{ij}a_{11} & w_{ij}a_{12} & \cdots & w_{ij}a_{1n} \\ w_{ij}a_{21} & w_{ij}a_{22} & \cdots & w_{ij}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{ij}a_{n1} & w_{ij}a_{n2} & \cdots & w_{ij}a_{nn} \end{pmatrix} \tag{2.4}$$

Kronecker product can be applied for representation of consensus protocols over vector space. Consider consensus of agents over a vector state ξ,

$$\xi_k^i = \sum_{j=1}^{N} w_{ij}\xi_{k-1}^j \tag{2.5}$$

Concatenating the state vectors at all agents, the global state vector is in the following form,

$$\underline{\xi}_k = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \vdots \\ \xi^N \end{pmatrix}_k = \mathbf{1}_N \otimes \xi_k, \tag{2.6}$$

Then the consensus equation (2.9) can be written as the following matrix form,

$$\underline{\xi}_k = (I \otimes W_k)\underline{\xi}_{k-1}, \tag{2.7}$$

where I is the $N \times N$ identity matrix. We specifically employ such Kronecker product to derive the distributed formulation results in chapter 4.

2.3 Consensus Algorithms

Consensus algorithms have potential application in distributed data processing and synchronization. The objective of a consensus algorithm is to lead a group of agents to reach a common state or certain quantity of interest (see Fig. 2.2 as an example). This concept is introduced in preliminary works of [2,4] for average consensus. Later, complementary works on dynamically changing sparse networks [54], asynchronous updates [55, 56], nonlinear protocols [57], and finite-time convergence [58, 59] are discussed in the literature. We refer interested readers to [1,60,61] for a survey of related works and applications. Here, we state a brief review of the general case of discrete-time consensus. Consider N agents distributed over a network, $\mathcal{G}_W$, with adjacency matrix W defining the interaction weights of the agents on each other. Assume each

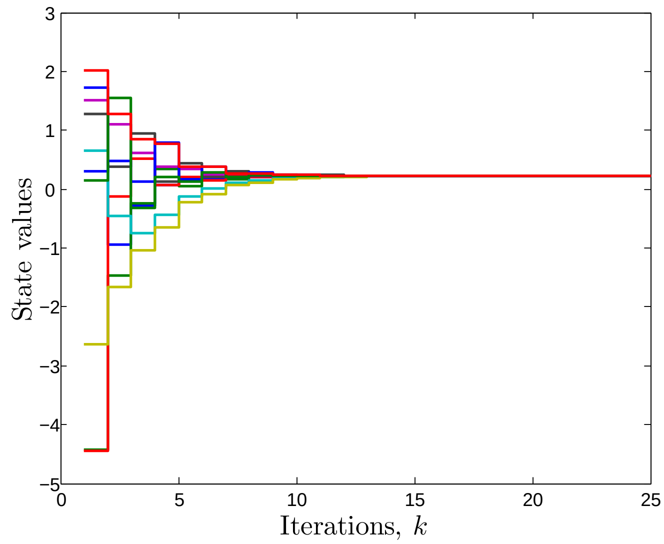

Figure 2.2: Consensus of 10 states in discrete-time.

agent is associated with a dynamic state ξ evolving as,

$$\xi_k^i = \sum_{j=1}^{N} w_{ij} \xi_{k-1}^j, \tag{2.8}$$

where k is the discrete-time index. Notice that in general the interaction topology $\mathcal{G}_W$ may change over time. In the matrix form the above equation can be written as,

$$\xi_k = W_k \xi_{k-1}, \tag{2.9}$$

where the subscript k in W_k is to consider possibly time-varying weight matrices.

Definition 2.3.1. *Matrix W is called row stochastic if $\sum_{j=1}^{n} w_{ij} = 1$. Matrix W is called column stochastic if $\sum_{i=1}^{n} w_{ij} = 1$. Matrix W is doubly stochastic if it is both row and column stochastic.*

Lemma 2.3.1. *[60] The dynamics in* (2.9) *reaches asymptotic consensus among* *agents if and only if:*

(i) Matrix W is row-stochastic.

(ii) The algebraic multiplicity of largest eigen-value of W is 1, i.e., $\lambda_W = 1$ is a simple eigen-value.

This lemma is a result of Gershgorin's disc theorem. If $\lambda_W = 1$ is simple and v is a column left eigen-vector of W (i.e., $vW = v\lambda$) then,

$$\lim_{k \to \infty} W^k \to \mathbf{1}v^T. \tag{2.10}$$

This implies that,

$$\xi_k = W^k \xi_0 \xrightarrow{k\to\infty} \mathbf{1}v^T \xi_0, \tag{2.11}$$

$$\xi_k^i \xrightarrow{k\to\infty} v^T \xi_0, \tag{2.12}$$

$$|\xi_k^i - \xi_k^j| \xrightarrow{k\to\infty} 0. \tag{2.13}$$

implying that consensus is reached. The second condition in Lemma 2.3.1 has an interpretation in graph theoretic sense,

Corollary 2.3.1. *[54] $\lambda_W = 1$ is a simple eigen-value of row-stochastic matrix W if and only if the interaction graph, $\mathcal{G}_W$, includes a directed spanning tree.*

A typical assumption in consensus literature is to assume $\mathcal{G}_W$ to be SC, or the matrix W to be *irreducible*. This is a sufficient condition for consensus follows from the Perron-Frobenius theorem.

Lemma 2.3.2. *[1] If the matrix W is doubly stochastic and the digraph $\mathcal{G}_W$ is SC, then the dynamics* (2.9) *reaches average-consensus. Mathmatically, $v = \frac{1}{n}\mathbf{1}$, and*

$$\xi_k^i \xrightarrow{k\to\infty} v^T \xi_0 = \frac{1}{n}\sum_{j=1}^{n} \xi_0^j. \tag{2.14}$$

In [54,59] authors investigate necessary and sufficient conditions on consensus over dynamically changing interaction graphs. It is shown that consensus can be reached if the union of interaction graph contains a spanning tree over sufficiently enough sequence of time-intervals. An immediate result is given in the following,

Lemma 2.3.3. *Interaction dynamics* (2.9) *reaches asymptotic consensus over fixed graphs with time-varying weights if and only if:*

(i) Matrix W remains row-stochastic at every time step.

(ii) Fixed topology $\mathcal{G}_W$ includes a spanning tree.

The consensus protocols found recent application in filtering, pioneering by works of Olfati-saber *et al.* [5, 62]. In these works the consensus state is tracking a time-varying input, as opposed to stable dynamics considered in [2]. Recently, such consensus based estimation and filtering are proposed in distributed estimation literature.

2.4 Brief Review of Centralized Estimation

Estimation and observability of linear dynamical systems has been discussed since the pioneering works by *Kalman* [12, 33][1]. Kalman filter is a recursive algorithm to optimally infer the variable x representing the dynamical state of the system from

[1]For a detailed introductory discussion on Kalman filter we refer the interested reader to [63].

the sample measurement y. The term recursive implies that, unlike the other filtering algorithms, this filter does not rely on the past system information. In more detail, to estimate x_k (state x at sample-time k) only system information at current step (y_k) and last step (x_{k-1}) are required. This along with its performance in terms of error optimality distinguishes the Kalman estimator for practical applications.

Let $\widehat{\mathbf{x}}^c_{k|k}$ and $\widehat{\mathbf{x}}^c_{k|k-1}$ be the centralized Kalman estimate at time k given all the observations, $\mathbf{y}^i_k$, and $\mathbf{y}^i_{k-1}$, respectively. The Centralized Kalman Filter (CKF) is defined as two recursive steps:

(i) Prediction step:

$$\widehat{\mathbf{x}}^c_{k|k-1} = A\widehat{\mathbf{x}}^c_{k-1|k-1}, \tag{2.15}$$

(ii) Correction step:

$$\widehat{\mathbf{x}}^c_{k|k} = \widehat{\mathbf{x}}^c_{k|k-1} + K^c_k\left(\mathbf{y}_k - H\widehat{\mathbf{x}}^c_{k|k-1}\right). \tag{2.16}$$

It can be shown that the Kalman estimator error, $\widehat{\mathbf{e}}^c_{k|k} = \mathbf{x}_k - \widehat{\mathbf{x}}^c_{k|k}$, is given by,

$$\widehat{\mathbf{e}}^c_{k|k} = (A - K^c_k HA)\widehat{\mathbf{e}}^c_{k-1|k-1} + \eta_k, \tag{2.17}$$

where K^c_k is the centralized Kalman gain and the vector η_k collects the remaining terms that are independent of $\widehat{\mathbf{e}}^c_{k-1|k-1}$. It is well known that the centralized Kalman error, $\widehat{\mathbf{e}}^c_{k|k}$ is stable if and only if Schur condition for stability is satisfied, i.e., $\rho(A - K^c_k HA) < 1$ [13]. However, to satisfy the Schur condition, the existence of the *Kalman gain matrix* K^c_k is tied with the concept of *observability* which is discussed next.

2.5 Observability

The concept of observability is a determining factor in state estimation. Generally speaking observability implies inferring the internal states of the system from its measurements in finite time. In linear but static cases, observability defines the solvability of the set of measurement equations to recover an n-dimensional state parameter, subsequently requiring at least as many measurements as the number of unknown states, $p \geq n$. Observability in LTI dynamics is more challenging since the number, p, of measurements is typically less than the number, n, of states. The reason is that many state variables of interest cannot be measured/observed directly, hence needed to be inferred from other measured quantities. Simply, an observable dynamic system has enough state dependencies that can be exploited towards estimation of unmeasured states. There are different approaches to check for observability of LTI systems: (i) algebraic method of finding the rank of the observability Gramian [13, 14]; (ii) the Popov-Belevitch-Hautus (PBH) test [64]; and, (iii) graph-theoretic analysis of the system graph [16–18, 21, 34, 65, 66]. We briefly go through these methods in the following.

2.5.1 Algebraic method: Gramian matrix

Algebraically, the linear system (2.1) is observable by the given set of measurements (2.2) if there exists a finite number of time steps $m \leq n$ such that sequence of measurements $\mathbf{y}_0, \mathbf{y}_1, ..., \mathbf{y}_m$ gives sufficient information to determine initial state of the system, $\mathbf{x}_0$. Consider the set of measurements as,

$$\begin{cases} \mathbf{y}_0 = H\mathbf{x}_0 \\ \mathbf{y}_1 = H\mathbf{x}_1 = HA\mathbf{x}_0 \\ \vdots \\ \mathbf{y}_{m-1} = H\mathbf{x}_k = HA^{m-1}\mathbf{x}_0 \end{cases} \tag{2.18}$$

These equations can be compactly expressed as,

$$\underbrace{\begin{pmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{m-1} \end{pmatrix}}_{Y} = \underbrace{\begin{pmatrix} H \\ HA \\ \vdots \\ HA^{m-1} \end{pmatrix}}_{\mathcal{O}_m} \mathbf{x}_0 \tag{2.19}$$

For the above to have a solution for $\mathbf{x}_0$ the matrix $\mathcal{O}_m$, known as observability Gramian, should be full-rank.

$$\mathrm{rank}(\mathcal{O}_m) = n. \tag{2.20}$$

It is known that for any $m > n$ we have $\mathrm{rank}(\mathcal{O}_m) = \mathrm{rank}(\mathcal{O}_n)$. This is a result of Cayley-Hamilton theorem and implies that taking more than n-step of measurements does not improve the rank of $\mathcal{O}_n$. Algebraic tests for observability, therefore, check the Gramian, $\mathcal{O}_n$, to be full-rank or the matrix $\mathcal{O}_n^T\mathcal{O}_n$ to be invertible.

2.5.2 Symbolic test: PBH

An alternative method is the PBH (Popov-Belevitch-Hautus) observability test [64], which checks the rank of the following matrix,

$$\widetilde{A} = \begin{pmatrix} A - sI \\ H \end{pmatrix}, \tag{2.21}$$

Observability implies $\widetilde{A}$ be full-rank for all values of $s \in \mathbb{C}$ where I is the $n \times n$ identity matrix. The matrix, $A - sI$, is full rank for all (probably complex) values of s, except for the eigenvalues of A. This simply implies that the PBH test has to be checked *only* for these values. In other words, an LTI system is *not* observable if and only if there exist a right eigenvector of A in the null space of measurement matrix, H, i.e.

$$\{\exists \mathbf{w} \in \mathbb{R}^n | \ A\mathbf{w} = \lambda \mathbf{w}, \ H\mathbf{w} = 0\}. \tag{2.22}$$

2.5.3 Graph theoretic test: Structural observability

Note that, both these algebraic and symbolic methods rely on the knowledge of exact values of each element in the matrices A and H. However, in many dynamical systems, only the sparsity (zero and non-zero pattern) of these matrices may remain fixed while the non-zero elements are subject to change, for example, when the entries of these matrices depend on certain parameters or operating points. Hence, the conventional methodologies fail to check for observability in such cases and graph-theoretic techniques are to be employed. To analyze such graph-based methods, we first introduce preliminaries on structured system theory and related graph notions.

2.6 Structured System Theory

Structural analysis deals with system properties that do not depend on the numerical values of the parameters but only on the underlying structure (zeros and non-zeros) of the system [16–21, 34, 65–68]. It turns out that if a structural property is true for one *admissible* choice of non-zero elements as free parameters it is true for *almost all* choices of non-zero elements and, therefore, is called a *generic* property of the system [69]. Furthermore, it can be shown that those particular (non-admissible) choices for which the generic property does not hold lie on some algebraic variety with zero Lebesgue measure [70]. In statistical sense, this implies that having random numbers as system parameter the probability that a random point lies on this subspace is almost zero. A simple geometrical example would be a line in $\mathbb{R}^2$ space. The area, as Lebesgue measure, of this line is almost zero. In probabilistic sense this means that the chance of having a randompoint in $\mathbb{R}^2$ to be on such line is approximately zero.

2.7 Structural Observability

Instead of using the algebraic/symbolic tests for observability, an alternate is a graph-theoretic approach. Structural (also known as generic) observability is based on structured systems theory. This method only relies on the structure of system matrices A and H (i.e. the topology of the system digraph $\mathcal{G}_{sys}$). The main theorem on generic observability–dual of the generic controllability result in [17]– is stated in the following,

Theorem 2.7.1. *A system is generically* (A, H)*-observable if and only if in its digraph* $\mathcal{G}_{sys}$*:*

(i) Every state x_i *is the begin node of a* $\mathcal{Y}$*-connected path, i.e.* $x_i \xrightarrow{path} \mathcal{Y}, \forall i \in \{1, \ldots, n\}$*;*

(ii) There exist a family of disjoint $\mathcal{Y}$*-connected paths and cycles covering all state nodes.*

The first condition is known as *accessibility* and the second as the *S-rank* or *matching* condition. The above conditions, however, are known to have algebraic meanings, which is discussed in Chapter 3 of this thesis. In LTI state-space observability, a significant question is to find a set of critical measurements to satisfy Theorem 2.7.1. Recent literature [15, 19, 66] discusses different aspects and approaches towards this problem. In these works, the LTI systems are modeled as digraphs and graph-theoretic algorithms are adapted to find the corresponding critical measurements. Since these results are structural, they ensure *generic observability*, i.e. the underlying LTI systems are observable for almost all choices of non-zeros in the corresponding matrices A and H [18]. In this regard the dynamical equations (2.1) and (2.2) may be represented in linear structured-invariant (LSI) form as,

$$\mathbf{x}_{k+1} = \mathcal{A}\mathbf{x}_k + \mathbf{v}_k, \tag{2.23}$$

$$\mathbf{y}_k = \mathcal{H}\mathbf{x}_k + \mathbf{r}_k, \tag{2.24}$$

where $\mathcal{A}$ and $\mathcal{H}$ represent the structured form (zero-nonzero pattern) of system matrix A and measurement matrix H, respectively.

2.8 Structural Rank

Another generic property of the system is structural rank or generic rank. The definition of structural rank (S-rank in short) and its properties can be described as follows.

Definition 2.8.1. *S-rank of a matrix, $A = [a_{ij}]$, is the maximal rank of structured matrix $\mathcal{A}$ over all numerical values of the non-zero entries $a_{ij} \neq 0$.*

In the algebraic sense, the S-rank implies maximum number of non-zero elements in *distinct* rows and columns of a matrix [71]. It is clear that for any given matrix A,

$$\text{rank}(A) \leq \text{S-rank}(A) \tag{2.25}$$

From the definition 2.8.1, it immediately follows that, a full rank system is also *structurally* full rank and a structurally rank deficient system is always rank deficient.

Lemma 2.8.1. *A system matrix, A, is full S-rank if and only if there exists a disjoint family of cycles spanning all the state vertices in its digraph $\mathcal{G}_A$; otherwise, the system is S-rank deficient. Further, condition (ii) in Theorem 2.7.1 on generic observability of $(A_{n\times n}, H_{N\times n})$ is equivalent to,*

$$S\text{-}rank \begin{pmatrix} A \\ H \end{pmatrix} = n. \tag{2.26}$$

Lemma 2.8.2. *A matrix, W, with all non-zero diagonals is full S-rank. This is because $\mathcal{G}_W$ includes a disjoint family of self-cycles at all nodes.*

Examples of S-rank deficient systems are shown in Fig. 2.3. For both graphs,

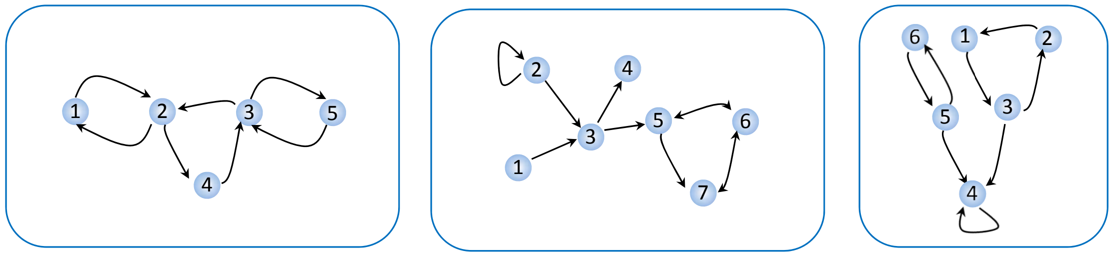

Figure 2.3: Examples of system digraph: (Left) an S-rank deficient SC digraph, (Middle) an S-rank deficient non-SC digraph, and (Right) a full S-rank digraph.

there are no family of *disjoint* cycles spanning all state nodes. More discussion on the structural observability and its algebraic implications are given in Chapter 3.

2.9 Distributed Estimation

Advances in distributed algorithms and data processing motivated estimation over a (distributed) network of observers. Instead of a central unit receiving all the observations and performing the estimation, the information processing is localized and shared among a group of agents. As mentioned in the last section, consensus algorithms are proved to be an efficient tool for information fusion among the agents. Particularly, in the context of sparsely-connected networks, where the main focus is to reduce the uncertainty of the individual estimates by averaging on collaborative data. The literature on this subject exists from earlier work in [23, 24] and references therein, where parallel Kalman filter architectures are considered, generally, for all-to-all connected networks, to more recent works in [5–8, 61, 62, 72], where average-consensus based Kalamn filtering has been studied as an effective method for distributed computing and estimation over sparsely connected networks. This approach requires two time-scales where communication is implemented over a much faster rate than the sampling of the dynamics (Fig. 1.1). Assuming a faster communication rate becomes practically in-feasible when the underlying system is operating under power constraints and has restricted communication and computation budgets.

To avoid these, single time-scale estimation is proposed recently in [10, 31, 40, 73] and has been used widely in the literature [9, 25, 26, 29, 32, 35–39]. In single time-scale estimation, communication is implemented at the same sampling rate as of the dynamics. Since the communication and sampling have the same time-scales, the communication network among the agents plays a key role in the observability of the agents. On the contrary, in average-consensus based approaches (two time-scales), the communication network becomes irrelevant due to more information exchanges among the agents (a sparsely connected graph looks like a fully connected graph when a large number of information exchanges are carried out). In single time-scale estimation a key consideration is the observability of the networked estimator. In general, it can be shown that an arbitrary communication network among the agents may not suffice to make the networked system observable. Hence, an important infrastructure design question is to design communication networks that can recover the observability of the distributed system. Distributed observability, however, is shown to be related to the *Kronecker product* of system matrix and adjacency matrix of the multi-agent network.

More recent solutions exists for distributed estimation, including diffusion-based schemes in Kalman filtering and smoothing [35] with application in distributed binary detection [36]. Meanwhile, incremental adaptive distributed strategies are proposed in [25, 74] along with distributed moving horizon estimation [37] to minimize estimation error variance for constrained problems. State estimators based on low-cost single-bit data transmission is proposed in [38] with binary *sign of innovations* (sign of difference of measurement and estimated value). [32] considers distributed estimation of time-varying *scalar* signals. The paper propose a distributed estimator with both measurement and estimate consensus update for tracking quasi-stationary systems. The filtering process on estimation and error covariance is *local* at each agent where the estimation bound is tied with the number of neighboring agents (network density). Further, the optimal estimation parameters (consensus weights) are defined locally to minimize the MSE. In other work, information theoretic approach based on consensus over the Kullback-Leibler average of Gaussian PDFs is exploited in [39]. It should be noted that not all of these literature are dynamic estimation but some consider the target state to be (quasi) stationary over time.[2] Dynamic estimation [5–9, 23, 37–39], however, tracks the time-evolution of the system.

As mentioned in Chapter 1, in single-time estimation, and specifically dynamic case, the structure of the communication network plays a key role on the stability of the distributed estimator [10, 11]. In this sense, [16, 18–20, 34] study observability with a structural point of view. The prevalent assumption requires the communication network to be (strongly) connected or cyclic as in [6,11,25,26,29,32,35,36,39]. On the other hand, [15, 16, 19] structurally determine the communication network such that $(A, \sum_{j \in \mathcal{N}(i)} H_j^T H_j)$ is observable, i.e. each agent i is observable in its neighborhood $\mathcal{N}(i)$. Reference [40] introduces a single-time consensus and innovation estimator over connected un-directed agent communication network. However, in their approach they

[2] As stated in [26], diffusion algorithms can be extended for non-stationary (dynamic) tracking when the target is not moving too fast, i.e. its state is relatively stationary over a period such that the algorithm can converge.

assume local observability, i.e. observability of $(A, \sum_{j\in\mathcal{N}(i)} H_j^T H_j)$. With the same assumption, [75] investigates distributed detectability of a single-time consensus-based estimator over the given multi-agent network, referred to as filter graph. Applying an algebraic approach on graph Laplacian, the necessary condition for observability is defined as each cluster having a spanning tree.

As compared to literature, our goal is to design the network with minimal communication. Specifically, our methodology is independent of exact system parameter values, relying on the system *structure*. Such *generic* approach is helpful when the parameters may vary depending on the system operating point (e.g. linearization of *smooth* non-linear dynamics [20]) and is, further, independent of the exact value of the weights chosen for data fusion. This leads to a robust estimator design where the analysis is not algebraic, as in the conventional Grammian or PBH observability tests, but graph-theoretic [18]. It should be emphasized that, as stated in Chapter 1, we make no assumption on the communication network, but we aim to design the network structure. Further, unlike [76,77], we do not impose any agent hierarchy, i.e., all agent duties are the same, increasing reliability to node/link failure.

2.10 Conclusions

In this chapter, we set the preliminaries on structural observability and corresponding graph-theoretic methods. In general, there exist combinatorial algorithms to check for observability conditions on system digraph, as given in the Appendix I of this thesis. The vector-space consensus algorithms are employed in Chapter 4 to derive the algebraic formulation for distributed observability.

Chapter 3

Measurement Partitioning and Observational Equivalence

In this chapter, we first consider the problem of finding a set of state measurements that are required for observability. Next, we show that each such set of critical measurements can be further partitioned into two types: α and β; these different types of measurements have different algebraic and graph-theoretic interpretations that we characterize. Further, for a *given set* of sufficient measurements (measurements satisfying observability condition) we classify the unnecessary (non-critical) ones as Type-γ. Contrary to [15, 19, 21] we further classify critical measurements for their role in distributed analysis in subsequent chapters.

The second problem we consider here is observational equivalence. This is to define the states that are *equivalent* in terms of observability–the equivalence relation is denoted by '$\sim$'. Indeed, the set of necessary measurements for observability is not unique; this motivates to search for all possible sets that ensure observability. In particular, if two states, x_i and x_j are observationally equivalent, i.e. $x_i \sim x_j$, then measuring any one of them suffices for observability. Hence, the corresponding measurements are also equivalent, i.e. $y_i \sim y_j$. We characterize this notion of observational equivalence towards state estimation in both algebraic and graph-theoretic sense.

To derive the results, recall from Chapter 2 that algebraic observability is related to the rank of Gramian matrix, and graph-theoretic observability calls for the two conditions in Theorem 2.7.1 stated below:

(i) *Accessibility condition*: every state x_i is the starting node of a $\mathcal{Y}$-connected path.

(ii) *Rank condition:* a family of *disjoint* $\mathcal{Y}$-connected paths and cycles ($\mathcal{L}$) going through *all* state nodes.

3.1 Advanced Graph-Theoretic Notions

In order to develop our results in the structural context, we need some advanced graph-theoretic concepts borrowed from [78]. These concepts provide the foundations and related preliminaries for measurement classification and deriving equivalence sets, and further, necessary conditions on designing the agent communication network in Chapter 5.

3.1.1 Contractions in digraphs

The graph-theoretic concepts stated in this section are built on the graph notations in Chapter 2.

- *Bipartite graph:* denoted by $\Gamma = (\mathcal{V}^+, \mathcal{V}^-, \mathcal{E}_\Gamma)$, includes two disjoint set of nodes: $\mathcal{V}^+$ and $\mathcal{V}^-$, with set of edges $\in \mathcal{E}_\Gamma$ originating in $\mathcal{V}^+$ and terminating in $\mathcal{V}^-$. To construct bipartite graph, Γ_A from system digraph $\mathcal{G}_A$ define $\mathcal{V}^+ = \mathcal{X}$ and $\mathcal{V}^- = \mathcal{X} \cup \mathcal{Y}$, and $\mathcal{E}_{\Gamma_A}$, as collection of edges (v_j^-, v_i^+) with $(v_j, v_i) \in \mathcal{E}_A$.

- *Matching:* denoted by $\underline{\mathcal{M}}$, is defined as subset of edge set $\mathcal{E}_A$ with no common end-states. In Γ_A, a matching is a subset of edges non-incident on the same node, i.e., edges are mutually disjoint. The number of edges in $\underline{\mathcal{M}}$ defines the size of the matching, $|\underline{\mathcal{M}}|$. A maximum size matching is called maximal matching, denoted by $\mathcal{M}$. It is known that $\mathcal{M}$ is not unique. in general. If $|\mathcal{M}| = n$ it is called perfect matching.[1]

- *Unmatched nodes:* Let $\partial\mathcal{M}^+$ and $\partial\mathcal{M}^-$ be the state nodes incident to edges in maximal matching $\mathcal{M}$, respectively, in $\mathcal{V}^+$ and $\mathcal{V}^-$. Denote by $\delta\mathcal{M}^+$ the unmatched nodes in $\mathcal{V}^+$, i.e., $\delta\mathcal{M}^+ = \mathcal{V}^+ \backslash \partial\mathcal{M}^+$.

- *Auxiliary graph:* denoted by $\Gamma_A^{\mathcal{M}}$, is a bipartite graph associated to $\mathcal{M}$ by reversing all edges in $\mathcal{M}$ while preserving direction of other edges, i.e., $\mathcal{E}_{\Gamma_A} \backslash \mathcal{M}$, in the bipartite graph, Γ_A.

- *Alternating path:* is defined as sequence of edges originating from an unmatched node in $\delta\mathcal{M}^+$ and every second edge in $\mathcal{M}$ in the auxiliary graph. The name comes from the fact that the edges alternate between $\mathcal{E} \backslash \mathcal{M}$ and $\mathcal{M}$.

- *Contraction:* denoted by $\mathcal{C}_i$, is assigned to an unmatched node, $v \in \delta\mathcal{M}^+$ such that it contains all states reachable by alternating paths from v_j in $\Gamma_A^{\mathcal{M}}$. In $\mathcal{G}_A$ a contraction represents state nodes connected (contracted) to less number of nodes. Further, define $\mathcal{C}$ as the set of all $\mathcal{C}_i$'s.

[1] A perfect matching indicates that all n columns are adjacent to all n rows of the associated matrix of the system digraph. This is known as the Hall property [79].

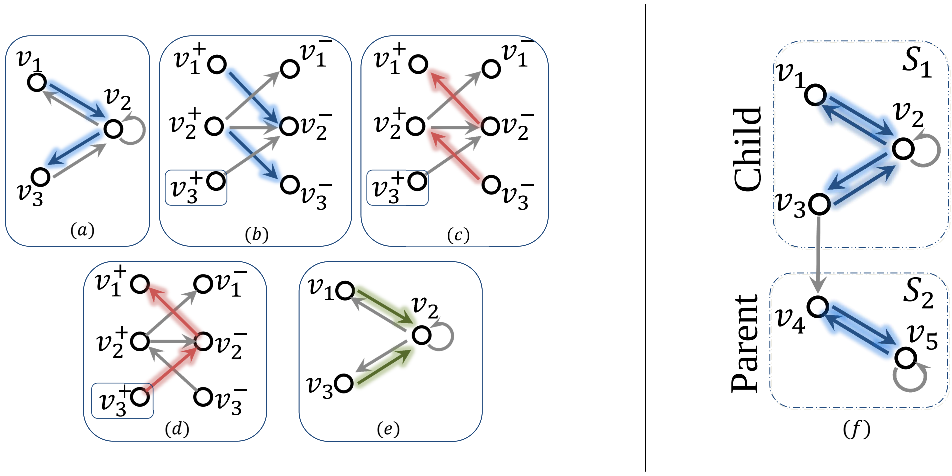

Figure 3.1: (a) System digraph $\mathcal{G}_A$. (b) Bipartite graph, Γ_A with maximal matching colored in blue. (c) Auxiliary graph $\Gamma_A^{\mathcal{M}}$. (d) Alternating path. (e) Contraction. (f) Parent/Child SCC classification of another simple digraph.

Example: In Fig. 3.1, a simple 3-node graph is given to illustrate the above definitions.

3.1.2 Parent/Child Strongly Connected Components

- *Matched SCC:* denoted by $\mathcal{S}_i^{\circlearrowleft}$, is an SCC including a union of disjoint cycles, also referred to as *Strong Components* [78]. $\mathcal{S}^{\circlearrowleft}$ denotes the set of all matched SCCs.

- *Parent SCC:* denoted by $\mathcal{S}_i^p$, is an SCC with no outgoing edge to any state out of it. A non-parent SCC is a *child*, denoted by $\mathcal{S}_i^c$. Let $\mathcal{S}^p$ be the set of all parent SCCs. Following the same convention, $\mathcal{S}_i^{\circlearrowleft,p}$ denotes a matched parent SCC, and $\mathcal{S}^{\circlearrowleft,p}$ denotes the set of all matched parent SCCs, and so on. The parent/child classification is also known as root/non-root SCC classification in [66]. It is known that the SCCs of an acyclic digraph can be uniquely and efficiently classified as a union of parent and child SCCs using the *depth first search* (DFS) or *Tarjan* algorithm [80, 81].

- *Partial order:* $\leq$, defines the existence of edge(s) from one component to another. For example, $\mathcal{S}_i \leq \mathcal{S}_j$ implies that $\mathcal{S}_i$ is a child component and some of it nodes have a path to some nodes in $\mathcal{S}_j$.

Remark 3.1.1. *In an SC graph, for every child SCC, $\mathcal{S}_i^c$, there exists a parent SCC, $\mathcal{S}_j^p$, i.e., $\mathcal{S}_i^c \leq \mathcal{S}_j^p$.*

Example: In the example of Fig. 3.1–(a), the graph is SC and the entire graph may be considered as one *unmatched* parent SCC since it has no outgoing link. On the other hand, Fig. 3.1–(f) gives another example of Parent/Child classification.

3.2 Measurement Partitioning

In this section, we describe the process of measurement partitioning. Given a set of observable measurements (matrix H)–such that (A, H) is observable–we partition the measurements into three types: α, β, and γ based on their role in generic observability. Type-α and Type-β are critical for observability (assuming fixed H) while Type-γ measurements are unnecessary.

Definition 3.2.1. *Given system matrices, A and H, a measurement is called critical if and only if removing it renders the system unobservable.*

For a given H, represent sub-matrices of partition, α, β, and γ, respectively, by H_α, H_β, and H_γ. Further, $H_{\alpha,\beta}$ represent the sub-matrix of both α and β measurements. Using this notation, the above definition can be summarized as:

$$\operatorname{rank}\left(\mathcal{O}\left(A, \begin{pmatrix} H_{\alpha,\beta} \\ H_{\gamma} \end{pmatrix}\right)\right) = \operatorname{rank}\left(\mathcal{O}\left(A, H_{\alpha,\beta}\right)\right) = n. \tag{3.1}$$

The rest of this chapter characterize these partitions in lieu of Theorem 2.7.1 and ensuring that each of the two (graph-theoretic) conditions (i) and (ii) are satisfied. It is straightforward to note that graph-theoretic interpretation is based on Theorem 2.7.1, while algebraic interpretation is described in Propositions 3.2.1 and 3.2.2.

3.2.1 Graph-theoretic

The first type of necessary measurements is characterized via maximum matching and graph contractions defined in Section 3.1. Having the maximum matching, $\mathcal{M}$, and set of unmatched states $\delta\mathcal{M}^+$, we define Type-α measurement as follows:

Definition 3.2.2. *Measurement of an unmatched node, $v_j \in \delta\mathcal{M}^+$ is Type-α.*

On the other hand, *Type-β* measurements are related to SCCs in the system digraph. In a non-SC digraph, define Parent/Child SCCs and partial order as given in Section 3.1.

Definition 3.2.3. *A Type-β measurement is the measurement of a state in a matched Parent SCC, $\mathcal{S}_i^{\circlearrowleft,p}$.*[2]

Example: in Fig. 3.1–(e) measurement of state v_1 or v_3 is Type-α, and in Fig. 3.1–(f) measurement of state v_4 or v_5 is Type-β.

[2]Throughout this thesis we name $\{\alpha, \beta, \gamma\}$ classification for the states, their measurements, and the assigned agents measuring those states.

3.2.2 Algebraic

Recall the algebraic implications of the structural system observability conditions in Chapter 2.

Proposition 3.2.1. *Accessibility is tied with the irreducibility of the matrix* $\begin{pmatrix} A^\top & H^\top \end{pmatrix}^\top$. *Having an inaccessible node in the system digraph implies the existence of a permutation matrix P such that,*

$$PAP^{-1} = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right), \; PH = [0 \;\; | \; H_1]. \tag{3.2}$$

Proposition 3.2.2. *S-rank condition is related to the structural rank of the system, i.e.*

$$S\text{-}rank \begin{pmatrix} A \\ H \end{pmatrix} = n. \tag{3.3}$$

Clearly, a measurement described in Proposition 3.2.2 may not satisfy a measurement given by Proposition 3.2.1, i.e. a measurement recovering accessibility may not improve the S-rank of $[A^\top \; H^\top]^\top$. In this sense, Type-α measurement improves the S-rank of $[A^\top \; H^\top]^\top$ by 1. Consequently, for full rank systems[3], there is no state measurement of Type-α, and Type-β measurements recover accessibility.

Definition 3.2.4. *In the algebraic sense, measurement α_i, the Type-α measurement*

[3]Notice that if system is full-rank, it is also structurally full-rank.

from state x_i, satisfies the following:

$$S\text{-}rank\begin{pmatrix} A \\ H_{\alpha_i} \end{pmatrix} = S\text{-}rank(A) + 1, \tag{3.4}$$

where H_{α_i} is a row vector of size n with only non-zero at its i location.

Each Type-α measurement thus improves the S-rank condition by exactly 1.

Definition 3.2.5. *The Type-β measurement of state x_i, denoted by β_i, does not improve the S-rank, i.e.*

$$S\text{-}rank\begin{pmatrix} A \\ H_{\beta_i} \end{pmatrix} = S\text{-}rank(A), \tag{3.5}$$

However, from Def. 3.2.1, a Type-β measurement satisfies Eq. (3.5) *and*

$$rank\left(\mathcal{O}\left(A, \begin{pmatrix} H_\alpha \\ H_{\beta_i} \end{pmatrix}\right)\right) = rank\left(\mathcal{O}\left(A, H_\alpha\right)\right) + 1. \tag{3.6}$$

3.3 Observational Equivalence

Set theory and abstract algebra literature defines the equivalence relation, '~', as having three properties: reflexivity, symmetry, and transitivity [82]. Towards observational equivalence in state estimation, reflexivity implies that every state is equivalent to itself, i.e. $x_i \sim x_i$; symmetry implies that if $x_i \sim x_j$ then $x_j \sim x_i$; and transitivity implies that if $x_i \sim x_j$ and $x_j \sim x_m$, then $x_i \sim x_m$. With this notation define observational equivalence as:

Definition 3.3.1. *Let H_i denote a row vector of size n with only non-zero at ith entry denoting measurement of state x_i. Observational equivalence among two states, $x_i \sim x_j$, is defined as*

$$rank\ \mathcal{O}(A, H_i) = rank\ \mathcal{O}(A, H_j) = rank\ \mathcal{O}\left(A, \begin{pmatrix} H_i \\ H_j \end{pmatrix}\right).$$

It can be easily verified that the above definition follows three properties of transitivity, reflexivity, and symmetry.

3.3.1 Graph-theoretic

Applying the concept of maximal matching and contractions in Section 3.1, we have the following results:

Lemma 3.3.1. *Any choice of maximal matching renders the same contraction set $\mathcal{C}$. Any state v_j reachable from $v_i \in \delta\mathcal{M}_1^+$ through an alternating path is unmatched in another maximal matching $\mathcal{M}_2^+$, i.e. $v_j \in \delta\mathcal{M}_2^+$ [78]. Also, any maximal matching, $\mathcal{M}$, includes only one unmatched state in every contraction, $\mathcal{C}_i$. [83]*

Lemma 3.3.2. *Given a contraction, $\mathcal{C}_i$, the unmatched node, $v_j(\mathcal{C}_i)$, within this contraction is not unique. In other words, for two vertices $v_j(\mathcal{C}_i)$ and $v_g(\mathcal{C}_i)$, $v_j(\mathcal{C}_i) \in \delta\mathcal{M}_1^+$ and $v_g(\mathcal{C}_i) \in \delta\mathcal{M}_2^+$, where $\mathcal{M}_1$ and $\mathcal{M}_2$ are two choices of maximum matching.*

Proof. The proof follows from the fact that maximum matching, in general, is not unique. This is from DM decomposition as given in [78]. To find a contraction, pick an unmatched state node, e.g. in $\delta\mathcal{M}_1^+$ and then in bipartite representation the contraction states can be reached. However, within a contraction $\mathcal{C}_i$ there are options for unmatched node, each of them related to an unmatched set $\delta\mathcal{M}_2^+$. In other words, every contraction includes exactly one unmatched node for any choice of $\mathcal{M}$. □

Lemma 3.3.3. *Two state measurements x_i and x_j from the same contraction, $\mathcal{C}_i$, are equivalent Type-α measurements.*

Proof. The equivalency of Type-α measurements is related to the choice of unmatched nodes in the corresponding contraction. Since measurement of each unmatched state increases the S-rank by 1 and considering that each contraction contributes to one rank-deficiency, it is straightforward to deduce three properties of the equivalence relation. □

Example: We illustrate these Lemmas in Fig. 3.2 with a contraction of 3 state nodes, $\mathcal{C}_1 = \{x_1, x_3, x_5\}$, into 2 nodes, $\{x_2, x_4\}$. The number of possible maximal matching is $\binom{3}{2} = 3$. From Fig. 3.2, a maximal matching, $\mathcal{M}$, gives one unmatched node in $\mathcal{C}_1$: e.g. in Fig. 3.2 (b), $\mathcal{M} = \{(x_5, x_4), (x_3, x_2)\}$ (highlighted edges) and x_1 is the unmatched node. $\delta\mathcal{M}^+ = \{x_1\}$ and after reversing the edges in that maximal matching, nodes x_3 and x_5 are reachable from x_1. Similarly, Figs. 3.2 (c) and (d) show other possible choices of maximal matching.

Next, we define Type-β equivalence relation.

Lemma 3.3.4. *Two Type-β measurements, β_i and β_j, of states x_i and x_j, are equivalent, $\beta_i \sim \beta_j$, if they belong to the same parent SCC, $\mathcal{S}_i^{\circlearrowleft p}$. Immediately follows that all states belonging to the same parent SCC are equivalent.*

Proof. The proof follows the strong connectivity of $\mathcal{S}_i^{\circlearrowleft p}$. This implies existence of a path through all state nodes to an observable node in the same SCC, and consequently, satisfying the second condition for structural observability. □

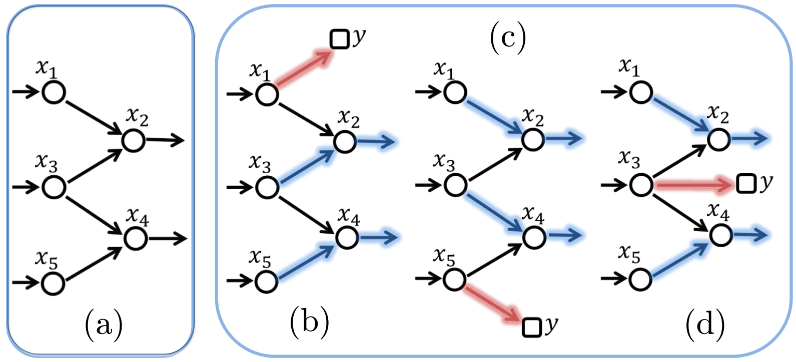

Figure 3.2: Possible maximal matching (shaded arrows) in a contraction. The red shaded arrow represents the measurement of the unmatched node.

Since all SCCs are disjoint in the system digraph, Type-β equivalent sets are disjoint. Notice that, an *unmatched* parent SCC has at least one Type-α measurement. This recovers both conditions for observability in Theorem 2.7.1 with no need of any other (Type-β) measurement.

3.3.2 Algebraic

Here, we provide the algebraic interpretation of equivalence among the Type-α and Type-β measurements.

Lemma 3.3.5. *Two Type-α measurements are equivalent, $\alpha_i \sim \alpha_j$, if and only if,*

$$\textit{S-rank}\begin{pmatrix} A \\ H_{\alpha_i} \end{pmatrix} = \textit{S-rank}\begin{pmatrix} A \\ H_{\alpha_j} \end{pmatrix} = \textit{S-rank}\begin{pmatrix} A \\ H_{\alpha_i} \\ H_{\alpha_j} \end{pmatrix} \tag{3.7}$$

implying that all equivalent Type-α measurements improve the S-rank only by 1.

Proof. Reflexivity and symmetry are directly induced by Eq. (3.7). For transitivity, consider three Type-α measurements, $\alpha_i, \alpha_j, \alpha_m$, with $\alpha_i \sim \alpha_j$ and $\alpha_j \sim \alpha_m$. Eqs. (3.4) and (3.7) give,

$$\text{span}\begin{pmatrix} A \\ H_{\alpha_i} \end{pmatrix} = \text{span}\begin{pmatrix} A \\ H_{\alpha_j} \end{pmatrix}, \qquad \text{span}\begin{pmatrix} A \\ H_{\alpha_j} \end{pmatrix} = \text{span}\begin{pmatrix} A \\ H_{\alpha_k} \end{pmatrix} \tag{3.8}$$

and transitivity follows. Similar arguments may be stated for sufficiency. □

The notion of (row) span in Lemma 3.3.5 is the maximal span over all possible choices of non-zeros in the corresponding matrix and is driven by structural rank.

Lemma 3.3.6. *Let H_α denote the Type-α measurement matrix. Two Type-β measurements, β_i and β_j, are equivalent, when*

$$\begin{aligned} rank\left(\mathcal{O}\left(A, \begin{pmatrix} H_\alpha \\ H_{\beta_i} \end{pmatrix}\right)\right) &= rank\left(\mathcal{O}\left(A, \begin{pmatrix} H_\alpha \\ H_{\beta_j} \end{pmatrix}\right)\right) = \\ rank\left(\mathcal{O}\left(A, \begin{pmatrix} H_\alpha \\ H_{\beta_i} \\ H_{\beta_j} \end{pmatrix}\right)\right) &= rank\left(\mathcal{O}\left(A, H_\alpha\right)\right) + 1. \end{aligned} \tag{3.9}$$

Proof. Reflexivity and symmetry are trivial. Transitivity follows from the fact that equivalent Type-β measurements belong to the *same irreducible block* of A (see [78]). Indeed, proper row-column permutation of the block structure proves the equivalency. □

3.4 Necessary Measurements for Centralized Observability

Given system (matrix), A, this section states the graph-theoretic conditions on the necessary measurements for centralized observability. This is the prerequisite of generic distributed analysis in the next chapters.

Theorem 3.4.1. *Both following conditions are critical for observability:*

(i) distinct measurement of (at least) one state in each contraction, $\mathcal{C}_i$.

(i) distinct measurement of (at least) one state in each parent SCC, $\mathcal{S}_j^{\circlearrowleft p}$.

Such set of measurements are critical for observability, and according to Assumption–(iii) in Chapter 1, we assume that any given set of measurements meet these two conditions for (minimum) observability.

3.5 Illustration

In this section we provide different examples which will be used in the next chapters to illustrate the results.

Example 3.5.1. *Consider a simple system digraph shown in Fig. 3.3 with the given set of measurements. The structured matrices A and H, are as follows:*

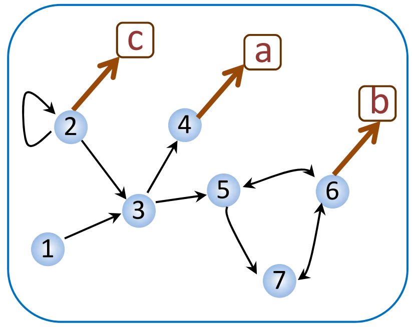

Figure 3.3: A digraph of state and measurement state nodes. This system is (A, H)-observable.

$$\mathcal{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 & 0 \\ \times & \times & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & \times & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & \times & 0 & \times \\ 0 & 0 & 0 & 0 & \times & \times & 0 \end{pmatrix}, \tag{3.10}$$

This system is S-rank deficient. The output matrix has the following structure:

$$\mathcal{H} = \begin{pmatrix} H_a \\ H_b \\ H_c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \times & 0 \\ 0 & \times & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.11}$$

Recall that system is globally observable by collection of the three measurements. By definition, agent a measuring the unmatched state $\{4\}$ is Type-α, agent b measuring a state in parent SCC $\{5, 6, 7\}$ is Type-β (crucial agents) and agent c is non-crucial (Type-γ).

Example 3.5.2. *Consider a 12-node graph given in Fig. 3.4. The structure of the adjacency matrix of this graph is as follows with each* $\times$ *sign representing a non-zero,*

$$\mathcal{A} = \begin{pmatrix} 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \times & 0 & \times & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 & \times & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \times & 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 \end{pmatrix}. \tag{3.12}$$

A possible maximum matching is

$$\mathcal{M} = \{(1,2),(2,4),(6,5),(5,6),(8,7),(7,8),(10,9),(9,10),(11,12),(12,11)\}$$

with set of unmatched states $\delta\mathcal{M}^+ = \{3,4\}$*. The assigned bipartite graph representation,* Γ_A *and the auxiliary graph,* $\Gamma_A^{\mathcal{M}}$ *are shown, along with the alternate path associated with each unmatched node. Node* x_1 *is reachable via alternate path from* x_3*; therefore,* $\mathcal{C}_1 = \{1,3\}$ *makes a contraction. Similarly,* $\mathcal{C}_2 = \{4,6,8,10,12,1\}$ *and the contraction set is* $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2\}$*. Further, this graph contains three SCCs; where* $\mathcal{S}_1 = \{5,6,7,8\}$ *and* $\mathcal{S}_2 = \{9,10,11,12\}$ *are matched and* $\mathcal{S}_3 = \{1,2,3,4\}$ *is unmatched. Further,* $\{\mathcal{S}_1, \mathcal{S}_2\}$ *have no outgoing edges and therefore make the Type-*β *equivalent set* $\mathcal{S}^{\circlearrowleft p}$*.* $\mathcal{S}_3$*, however, has outgoing edges to states* $\{x_5, x_9\}$ *and therefore is a child SCC:* $\mathcal{S}_3 \preceq \mathcal{S}_2$*,* $\mathcal{S}_3 \preceq \mathcal{S}_1$*. This example is an improvement to [21] where the parent cycle condition is revised to matched parent SCC condition. A possible set of system observations is* $\{x_3, x_6, x_{10}\}$*:*

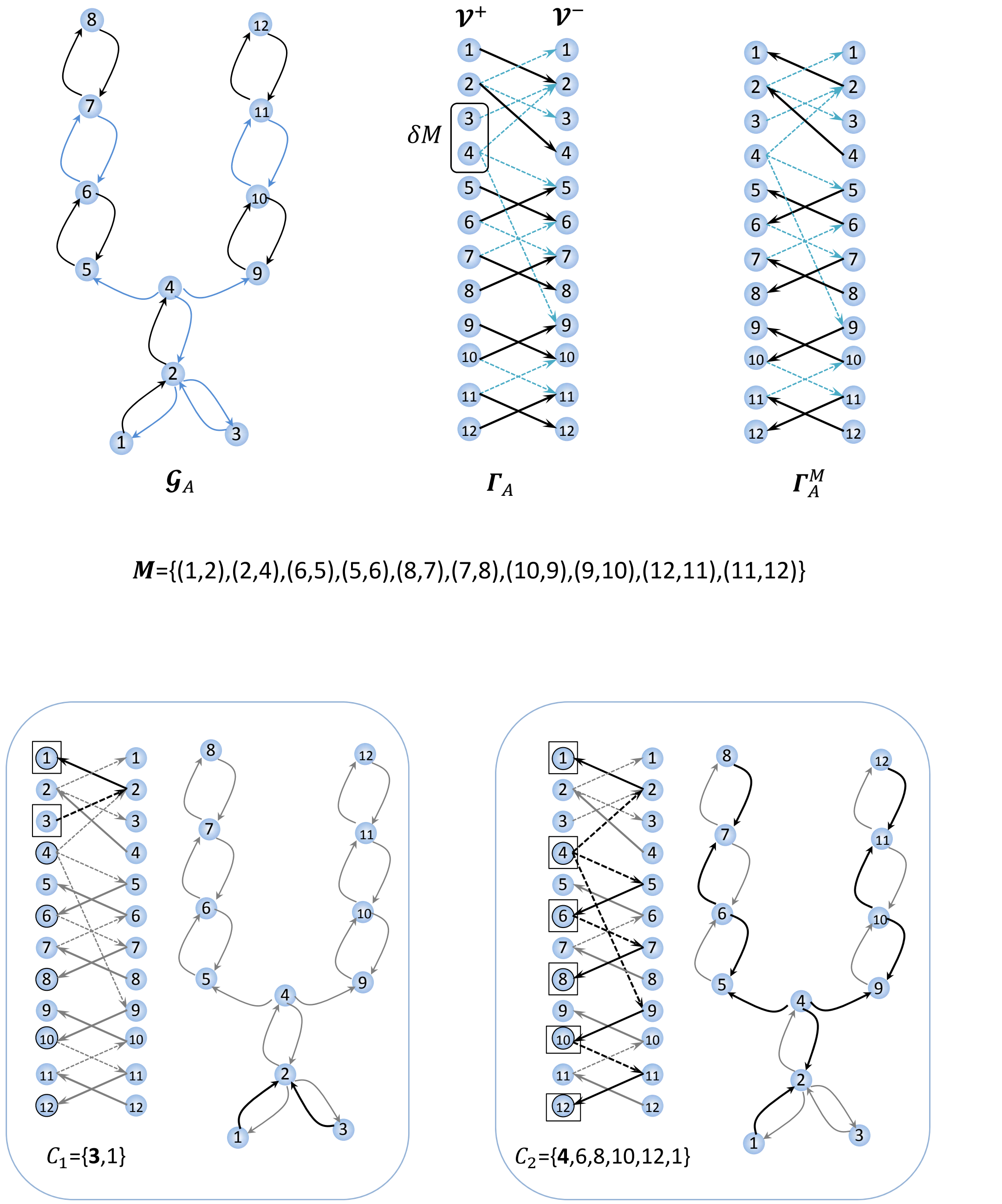

Figure 3.4: This figure shows the digraph of Example 3.5.2, a possible maximum matching, its bipartite graph, auxiliary graph, alternating paths, and contractions.

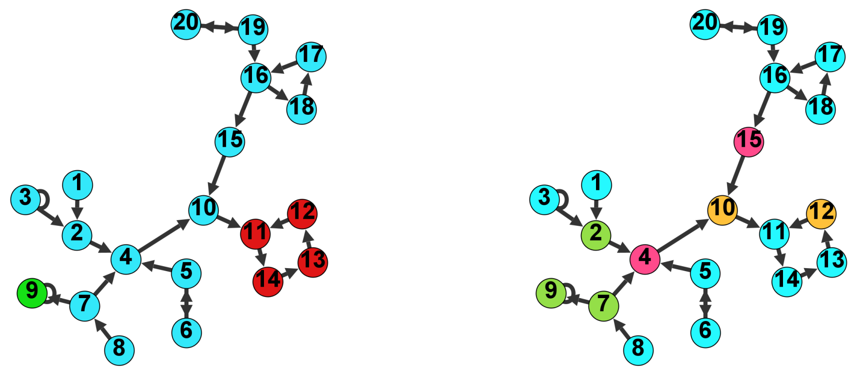

Figure 3.5: Type α and β equivalent sets for graph of Example 3.5.3. (Left) Type β equivalence sets (parent SCCs) are shown in red and green. (Right) Type- α equivalent sets (contractions) are shown in orange, purple, and green.

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 \end{pmatrix}. \tag{3.13}$$

Example 3.5.3. *Consider the system digraph given in Fig 3.5. This digraph has three contractions,* $\{\{2,7,9\},\{4,15\},\{10,12\}\}$, *constituting the equivalent Type-*α *sets; and two matched parent SCCs,* $\{\{11,12,13,14\},\{9\}\}$, *constituting the equivalent Type-*β *sets (the SCC,* $\{16,17,18\}$, *e.g., has an outgoing edge and hence is not parent). Three unmatched nodes each from a contraction make the Type-*α *sets:* $\alpha_1 \in \{2,7,9\}, \alpha_2 \in \{10,12\}, \alpha_3 \in \{4,15\}$. *Notice that both Type-*β *sets share nodes with the Type-*α *sets. Therefore, at least three measurements, e.g.* $\{4,9,12\}$, *are necessary.*

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.14}$$

In the case of not observing a shared α/β state, e.g. $\{12\}$, more than three observations are required; for example, $\{4, 9, 10, 13\}$ is another set of necessary measurements.

$$\mathcal{H} = \begin{pmatrix} 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \times & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.15}$$

3.6 Conclusions

In this chapter, we first derive both graph-theoretic and algebraic representations of two classes of critical measurements, Type-α and Type-β. This two-fold construction of partitions leads to establishing the notion of equivalence among both Type-α and Type-β measurements with different graph-theoretic and algebraic interpretations. We employ the graph-theoretic approach in the next chapters to formulate and solve the distributed observability problem in *generic* sense. As mentioned before, there exist combinatorial algorithms of *polynomial order* to define SCCs and their partial order, and matching/contractions; see Appendix I for more details.

The results of this chapter on observational equivalence specifically have application in fault detection. This is after the identification of faults, where observational equivalence provides a list of new states to be measured [84, 85] for system observability recovering. This may further improve estimation accuracy or provide low-cost benefits, to extend the work in [86, 87]. Another application is to define minimum order of input-output equations for Linear Time-Varying (LTV) filters as stated in [88].

Chapter 4

Distributed Observability from First Principles

The problem of estimating the state, $\mathbf{x}_k$, from the set of state measurements, $\mathbf{y}$, can be fundamentally considered in two different contexts;

- **Centralized observability:** all observations are transmitted to a central unit where the state estimate is processed. Such central coordinator is able to estimate the global state of the system with bounded error in steady state (e.g. Kalman Filtering) if and only if the pair (A, H) is *observable* [33].

- **Distributed observability:** the agents interact with each other over a network (graph). Each agent then estimates the state, $\mathbf{x}_k$, given the measurements *and/or* predictions of its neighbors. In this case, $(W \otimes A, D_H)$ observability implies system is observable in distributed sense, i.e., estimation of *every* agent based on its *local* information is bounded steady state error $[9, 27, 47]$.

In the previous chapter, we cast the necessary measurement, such that it includes (at least) one state observation from each set of equivalent sets. In this chapter, *given* a system with matrices (A, H) satisfying such observability condition, we establish the problem of distributed observability, i.e., the system to be observable locally at each agent. We derive the generalized mathematical term for observability of *any* distributed observer/estimator *from first principles*. Further, we support the results on distributed observability, via the example of our Networked Kalman-type Estimator (NKE). We derive the NKE error evolution and we show its stability under the proposed setup.

4.1 Distributed Observability

Recall $\mathcal{G}_{net} = (\mathcal{V}_{net}, \mathcal{E}_{net})$ as the agent interaction graph and $\mathcal{N}(i)$ as the neighborhood of agent i. Each agent i is to estimate the state, $\mathbf{x}_k$, with its observations, $\mathbf{y}_k^i$, and with its neighboring observations, $\{\mathbf{y}_k^j\}_{j\in\mathcal{N}(i)}$. Recall that, each agent, i, thus, estimates the state-vector, described by

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{v}_k, \tag{4.1}$$

from the following observations:

$$\mathbf{y}_k^j = H_j\mathbf{x}_k + \mathbf{r}_k^j, \qquad j \in \mathcal{N}_\alpha(i). \tag{4.2}$$

Let us assume that the neighbor set has $\mathcal{N}_\alpha(i)$ neighbors, and is indexed by $i_1, i_2, \ldots, i_{\mathcal{N}_\alpha(i)}$. Define matrix U as the adjacency matrix determining this neighborhood. Then, agent i is to estimate $\mathbf{x}_k$ from the neighboring observations, $\mathbf{y}_k^i, \mathbf{y}_k^{i_1}, \ldots, \mathbf{y}_k^{i_{\mathcal{N}_\alpha(i)}}$. Or, equivalently, with the following:

$$\widetilde{\mathbf{y}}_k^i \triangleq \begin{pmatrix} \mathbf{y}_k^i \\ \vdots \\ \mathbf{y}_k^{i_{N_i}} \end{pmatrix} = \begin{pmatrix} H_i \\ \vdots \\ H_{N_i} \end{pmatrix} \mathbf{x}_k + \begin{pmatrix} \mathbf{r}_k^i \\ \vdots \\ \mathbf{r}_k^{N_i} \end{pmatrix}. \tag{4.3}$$

The above observation model is equivalent to [9]:

$$\mathbf{z}_k^i = \begin{pmatrix} H_i^\top & \ldots & H_{N_i}^\top \end{pmatrix} \widetilde{\mathbf{y}}_k^i \triangleq \widetilde{H}_i \mathbf{x}_k + \widetilde{\mathbf{r}}_k^j, \tag{4.4}$$

$$\text{with} \quad \widetilde{H}_i \triangleq \sum_{j \in \mathcal{N}_\alpha(i)} H_j^\top H_j, \qquad \widetilde{\mathbf{r}}_k^i \triangleq \sum_{j \in \mathcal{N}_\alpha(i)} H_j^\top \mathbf{r}_k^j. \tag{4.5}$$

In fact, Eq. (4.4) is just a compact way of writing Eq. (4.3). The distributed estimation problem over the communication graph, $\mathcal{G}_{net}$, is now to estimate $\mathbf{x}_k$ at each agent, i, with the observations, $\mathbf{z}_k^i$. From the standard estimation theory arguments [13], we know that such an estimation is possible at any agent i, if and only if, the pair, $(A, \widetilde{H}_i)$, is observable. For observability at all of the agents, we must consider all such pairs, $(A, \widetilde{H}_1), (A, \widetilde{H}_2), \ldots, (A, \widetilde{H}_N)$, i.e. the observability of

$$\left(\underbrace{\begin{pmatrix} A & & \\ & \ddots & \\ & & A \end{pmatrix}}, \underbrace{\begin{pmatrix} \widetilde{H}_1 & & \\ & \ddots & \\ & & \widetilde{H}_N \end{pmatrix}}_{\triangleq D_H} \right), \tag{4.6}$$

compactly written as $(I \otimes A, D_H)$. It is straightforward to show that a centrally observable system does not necessarily imply that the distributed system is also observable, i.e.

$$(A, H)\text{-observability} \not\Rightarrow (I \otimes A, D_H)\text{-observability}. \tag{4.7}$$

We note that the above straightforward description of distributed observability is actually misleading. The primary reason is that although observation exchanges are considered, the agents may also exchange their local predictions. This latter exchange does not appear in the above characterization of distributed observability. In the following, we provide a novel construction to derive distributed observability that accommodates for both observation and prediction exchanges, and show that distributed observability *does not* require each agent to be observable in its neighborhood.

4.1.1 Derivation

Consider again the distributed estimation problem where we wish to estimate the dynamics in Eq. (2.1) via the observations in Eq. (1.4). Recall that $\widehat{\mathbf{x}}^i_{k|k}$ denotes the estimate of the state, $\mathbf{x}_k$, using all of the observations available at agent i, and its neighboring agents up to time k. Concatenating the estimates at all agents, the *global* state estimate in the network is

$$\underline{\widehat{\mathbf{x}}}_{k|k} \triangleq \begin{pmatrix} \widehat{\mathbf{x}}^1_{k|k} \\ \vdots \\ \widehat{\mathbf{x}}^N_{k|k} \end{pmatrix}. \tag{4.8}$$

Considering $\widehat{\underline{\mathbf{x}}}_{k|k}$ to be an estimate of some state, we seek the corresponding dynamical system to this state-estimate. Clearly, the corresponding dynamical system has the following *global* state vector:

$$\underline{\mathbf{x}}_k \triangleq \begin{pmatrix} \mathbf{x}_k \\ \vdots \\ \mathbf{x}_k \end{pmatrix} = \mathbf{1}_N \otimes \mathbf{x}_k. \tag{4.9}$$

where $\mathbf{1}_N$ is a column vector of N ones. To this end, let us assume that the dynamics associated to the above *global* state-vector, $\underline{\mathbf{x}}_k$, are given by some linear system:

$$\underline{\mathbf{x}}_{k+1} = Z\underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k, \tag{4.10}$$

where we have $Z \in \mathcal{Z}$, and $\mathcal{Z}$ is defined as a *class* of system matrices such that if we choose any matrix $Z \in \mathcal{Z}$, Eq. (4.10) remains a valid representation of the global state vector as given by concatenating the system dynamics of Eq. (2.1). We now characterize this class of system matrices, $\mathcal{Z}$. We have,

$$\begin{aligned} \underline{\mathbf{x}}_{k+1} &= \mathbf{1}_N \otimes \mathbf{x}_{k+1}, \\ &= \mathbf{1}_N \otimes (A\mathbf{x}_k + \mathbf{v}_k), \\ &= \mathbf{1}_N \otimes A\mathbf{x}_k + \mathbf{1}_N \otimes \mathbf{v}_k, \\ &= \underbrace{(W \otimes A)}_{\mathcal{Z}} \mathbf{x}_k + \underbrace{\mathbf{1}_N \otimes \mathbf{v}_k}_{\underline{\mathbf{v}}_k}, \end{aligned} \tag{4.11}$$

where the last equality follows if and only if W is *row-stochastic*, as described next,

$$(W \otimes A)\underline{\mathbf{x}_k} = \left(\begin{pmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{pmatrix} \otimes A \right) \underline{\mathbf{x}_k} \tag{4.12}$$

$$= \left(\begin{array}{c|c|c} w_{11}A & \cdots & w_{1N}A \\ \hline \vdots & \ddots & \vdots \\ \hline w_{N1}A & \cdots & w_{NN}A \end{array} \right) \left(\begin{array}{c} \mathbf{x}_k \\ \hline \vdots \\ \hline \mathbf{x}_k \end{array} \right)$$

$$= \begin{pmatrix} w_{11}A\mathbf{x}_k + \cdots + w_{1N}A\mathbf{x}_k \\ \vdots \\ w_{N1}A\mathbf{x}_k + \cdots + w_{NN}A\mathbf{x}_k \end{pmatrix} \tag{4.13}$$

$$= \begin{pmatrix} A\mathbf{x}_k \\ \vdots \\ A\mathbf{x}_k \end{pmatrix} = \mathbf{1}_N \otimes A\mathbf{x}_k \tag{4.14}$$

This leads to the conclusion that any matrix that cannot be decomposed as $W \otimes A$ is not a system matrix for the dynamics described by $\underline{\mathbf{x}}_{k+1}$, i.e.

$$\mathcal{Z} = \{Z \mid Z = (W \otimes A) \text{ and } W \text{ is stochastic}\}. \tag{4.15}$$

The propositions below follows the above arguments.

Proposition 4.1.1. *The distributed estimation of the dynamics in Eq.* (2.1) *monitored by measurements according to Eq.* (1.4)*, interacting over a communication* *graph, G_{net}, is equivalent to the centralized estimation of the following system:*

$$\underline{\mathbf{x}}_{k+1} = (W \otimes A)\underline{\mathbf{x}}_k + \underline{\mathbf{v}}_{k+1}, \tag{4.16}$$

$$\mathbf{z}_k \triangleq D_H \underline{\mathbf{x}}_k + \widetilde{\mathbf{r}}_k, \tag{4.17}$$

where W is stochastic.

We are now in a position to write the filtering equations for the centralized system (equivalent to the distributed estimation problem) in Eqs. (4.16)-(4.17):

$$\widehat{\underline{\mathbf{x}}}_{k|k-1} = (W \otimes A)\widehat{\underline{\mathbf{x}}}_{k-1|k-1}, \tag{4.18}$$

$$\widehat{\underline{\mathbf{x}}}_{k|k} = \widehat{\underline{\mathbf{x}}}_{k|k-1} + \underline{K}_k \left(\mathbf{z}_k - D_H \widehat{\underline{\mathbf{x}}}_{k|k-1}\right), \tag{4.19}$$

where $\underline{K}_k$ is the gain matrix (similar to Kalman filtering). The following theorem formally defines the distributed observability.

Theorem 4.1.1. *A dynamical system monitored by a network of interacting agents is distributively observable if and only if $(W \otimes A, D_H)$ is observable, where W is a row-stochastic matrix.*

Proof. The proof relies on the fact that the distributed estimation problem is equivalent to the centralized estimation problem with the pair of system matrices, $W \otimes A$ and D_H. □

As discussed in Chapter 2, the observability of the pair $(W \otimes A, D_H)$ can be checked using different observability tests, and among those we adapt structural observability. The structure of the matrix W that makes $(W \otimes A, D_H)$ observable, thus defines the sub-topology of the underlying agent communication, see Proposition 4.1.1. In the next chapter, we derive the necessary conditions on the communication topology, $\mathcal{G}_{net}$, to recover the distributed observability. But, we first give an example of a distributed observer in the next section.

4.2 Distributed Estimator

Although the centralized system, Eqs. (4.16)-(4.17), is equivalent to the distributed estimation problem, we still have to verify if the filtering equations, Eqs. (4.18)-(4.19), can be implemented in a distributed fashion. To this end we propose a distributed example in this section. Note that Eqs. (4.18)-(4.19) consists of two information fusion steps:

(i) Information fusion in the *prediction space*:

$$\underline{\widehat{\mathbf{x}}}_{k|k-1} = (W \otimes A)\underline{\widehat{\mathbf{x}}}_{k-1|k-1},$$

(ii) Information fusion in the *observation space*:

$$\underline{\widehat{\mathbf{x}}}_{k|k} = \underline{\widehat{\mathbf{x}}}_{k|k-1} + \underline{K}_k \left(\mathbf{z}_k - D_H \underline{\widehat{\mathbf{x}}}_{k|k-1}\right).$$

Consider information fusion in the prediction space (a priori step) to be implemented over the graph, $\mathcal{G}_\beta$, and information fusion in the observation space (a posteriori step) to be implemented over the graph $\mathcal{G}_\alpha$. Considering separate graphs is important since different connectivity conditions may be required for each fusion step. We may call $\mathcal{G}_\alpha$ and $\mathcal{G}_\beta$ respectively α-network and β-network throughout this thesis. With this two-layered approach, it is immediate to see that $\mathcal{G}_{net} = \mathcal{G}_\alpha \cup \mathcal{G}_\beta$. Finally, we denote the neighborhood at agent i (including agent i itself) as $\mathcal{N}_\alpha(i)$ and $\mathcal{N}_\beta(i)$, in $\mathcal{G}_\alpha$ and $\mathcal{G}_\beta$, respectively, with structured adjacency matrices U and W.

First consider the prediction-update in Eq. (4.18). It can be immediately observed that Eq. (4.18) is distributed:

$$\widehat{\mathbf{x}}^i_{k|k-1} = \sum_{j\in\mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}^j_{k-1|k-1}, \tag{4.20}$$

with $W = \{w_{ij}\}$. Next consider the observation fusion case, i.e. Eq. (4.19). Note that since the Kalman gain, $\underline{K}_k$, is a full matrix in general, Eq. (4.19) cannot be immediately distributed. In order to keep the implementation of Eq. (4.19) distributed and local, an alternate is to assume that the gain matrix, $\underline{K}_k$, is block-diagonal, leading to

$$\widehat{\mathbf{x}}^i_{k|k} = \widehat{\mathbf{x}}^i_{k|k-1} + K^i_k \sum_{j\in\mathcal{N}_\alpha(i)} H_j^\top \left(\mathbf{y}^j_k - H_j \widehat{\mathbf{x}}^i_{k|k-1}\right). \tag{4.21}$$

Now, recall the estimation error at time k,

$$\begin{aligned} \mathbf{e}^i_k &= \mathbf{x}_{k|k} - \widehat{\mathbf{x}}^i_{k|k}, \\ \mathbf{e}_k &= \begin{pmatrix} \mathbf{e}^1_k \\ \vdots \\ \mathbf{e}^N_k \end{pmatrix} \end{aligned} \tag{4.22}$$

Rewrite the estimation error at agent i as,

$$\mathbf{e}^i_k = \mathbf{x}_k - (\widehat{\mathbf{x}}^i_{k|k-1} + K^i_k \sum_{j\in\mathcal{N}_\alpha(i)} H_j^T (y^j_k - H_j \widehat{\mathbf{x}}^i_{k|k-1})) \tag{4.23}$$

Then from (4.20)-(4.21) we have the following,

$$\mathbf{e}_k^i = \mathbf{x}_k - \left(\sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j + K_k^i \sum_{j \in \mathcal{N}_\alpha(i)} H_j^T \left(y_k^j - H_j \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j \right) \right) \quad (4.24)$$

Replacing the system equations (2.1)-(1.4) we get,

$$\begin{aligned} \mathbf{e}_k^i = & (A\mathbf{x}_{k-1} + \mathbf{v}_{k-1}) - (\textstyle\sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j + \\ & K_k^i \textstyle\sum_{j \in \mathcal{N}_\alpha(i)} H_j^T (H_j \mathbf{x}_k + \mathbf{r}_k^i - H_j \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j)) \quad (4.25) \\ = & (A\mathbf{x}_{k-1} + \mathbf{v}_{k-1}) - \textstyle\sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j - \\ & K_k^i \textstyle\sum_{j \in \mathcal{N}_\alpha(i)} H_j^T (H_j (A x_{k-1} + \mathbf{v}_{k-1}) + \mathbf{r}_k^i - H_j \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j) \quad (4.26) \\ = & A\mathbf{x}_{k-1} - \textstyle\sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j - K_k^i \sum_{j \in \mathcal{N}_\alpha(i)} H_j^T (H_j A x_{k-1} - H_j \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j) \\ & + \mathbf{v}_{k-1} - \textstyle\sum_{j \in \mathcal{N}_\alpha(i)} H_j^T H_j \mathbf{v}_{k-1} - \sum_{j \in \mathcal{N}_\alpha(i)} H_j^T \mathbf{r}_k^i \quad (4.27) \end{aligned}$$

Note that the second term is a weighted linear function of the system and output noise and we can collect these noise terms in a new parameter $\mathbf{q}_k$.

$$\begin{aligned} \mathbf{e}_k^i &= A\mathbf{x}_{k-1} - \sum_{j \in \mathcal{D}_i} \widehat{\mathbf{x}}_{k-1|k-1}^j \\ &- K_k^i \sum_{j \in \mathcal{N}_\alpha(i)} H_j^T H_j \left(A x_{k-1} - \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j \right) + \mathbf{q}_k \quad (4.28) \end{aligned}$$

Using the fact that matrix W is stochastic, we get

$$A\mathbf{x}_{k-1} = \sum_{j \in \mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j \quad (4.29)$$

and thus,

$$\begin{aligned}\mathbf{e}_k^i &= \sum_{j\in\mathcal{N}_\beta(i)} w_{ij} A(\mathbf{x}_{k-1} - \widehat{\mathbf{x}}_{k-1|k-1}^j) \\ &- K_k^i \sum_{j\in\mathcal{N}_\alpha(i)} H_j^T H_j \sum_{j\in\mathcal{N}_\beta(i)} w_{ij} A(\mathbf{x}_{k-1} - \widehat{\mathbf{x}}_{k-1|k-1}^j) + \mathbf{q}_k \end{aligned} \tag{4.30}$$

Now define,

$$\begin{aligned}\underline{K}_k &= \begin{pmatrix} K_k^1 & & 0 \\ & \ddots & \\ 0 & & K_k^N \end{pmatrix}, \\ D_H &= \begin{pmatrix} \sum_{j\in\mathcal{N}_\alpha(1)} H_j^T H_j & & 0 \\ & \ddots & \\ 0 & & \sum_{j\in\mathcal{N}_\alpha(N)} H_j^T H_j \end{pmatrix},\end{aligned}$$

Every block diagonal, $\sum_{j\in\mathcal{N}_\alpha(i)} H_j^T H_j$, in matrix D_H, can be thought of as all the measurements in the (extended) neighborhood of agent i in sub-network $\mathcal{G}_\alpha$. This is because $H_j^T H_j$ is the square matrix of the measurement vector H_j, and have the same output information as in H_j. Finally, it can be shown that the networked error in the distributed estimator, Eqs. (4.20)-(4.21), evolves as,

$$\mathbf{e}_k = (W \otimes A - \underline{K}_k D_H (W \otimes A))\mathbf{e}_{k-1} + \mathbf{q}_k, \tag{4.31}$$

According to *Kalman Observability* [13], it is known that this error dynamics can be stabilized if the following pair is observable,

$$(W \otimes A, D_H). \tag{4.32}$$

Following Schur condition, a Kalman-type gain matrix, K, exists such that

$$\rho(W \otimes A - K D_H (W \otimes A)) < 1, \tag{4.33}$$

if $(W \otimes A, D_H)$ is observable[1]. Notice that, *theoretically*, to satisfy this $(W \otimes A, D_H)$ observability, we meet the observability conditions of Theorem 2.7.1 over distributed system digraph $\mathcal{G}_{Dist}$. However, in this thesis we constrain the gain matrix to be block-diagonal, denoted by $\underline{K}_k$. This implies that the resulting distributed estimator, Eqs. (4.20)-(4.21), is equal to the centralized counterpart, Eqs. (4.18)-(4.19), but with the extra condition of block-diagonal $\underline{K}_k$. In general, $\underline{K}_k$ cannot be computed locally from the standard procedures. However, computing such a constrained gain is possible via an iterative cone-complementary optimization algorithm stated in the Appendix II of this thesis. Nevertheless, if the centralized equivalent Eq. 4.31 has no (unconstrained) solution, then the distributed problem cannot have any solution to ensure observability of $(W \otimes A, D_H)$.

[1] In fact, this is straightforward to see by comparing to the Centralized Kalman-Filtering (CKF) (2.17) with error equation (4.31)

4.3 Conclusions

In this chapter, we formulate the distributed observability as the observability of the pair $(W \otimes A, D_H)$, where W and D_H are defined over different graphs (neighborhoods for each agent). We provide a novel construction to show that distributed observability *does not* require each agent to be observable in its neighborhood. This is by considering a step of *priori estimate* or *prediction* fusion as discussed. This is in contrast with the current trivial solutions for distributed observability, where each agent, i, is densely connected to be observable in its neighborhood [16, 40]. Intuitively, we expect distributed observability to be more relaxed, and this formulation holds for general distributed-type observers. In the next chapter, to satisfy this observability condition, we determine the *structure* of the network that agents communicate through.

Chapter 5

Necessary and Sufficient Network Connectivity

In this chapter, we provide the main results on the structure of the agent communication network to reach distributed observability. In simple words, we aim to define the *structure* of W and U (or D_H) such that the matrix $(W \otimes A, D_H)$ is observable in generic sense. Intuitively, these matrices define the communication links in the network, through which agents can share critical information on their observations and/or predictions. This information potentially may recover partial observability of each agent.

We define two graph topology $\mathcal{G}_\beta$ and $\mathcal{G}_\alpha$, the union of which builds the multi-agent network structure, $\mathcal{G}_{net}$. Assuming no information loss or package drop over the communication links, we first state *sufficient* condition(s) and the main theorem on the structure of the communication network. Then, we state a step-by-step proof of the theorem, first for full-rank and then for (structurally) rank deficient systems. Next, we use the same line of reasoning to discuss the *necessary* condition(s) on network topology. Finally, a system classification in terms of observability and, further, the network density of the monitoring distributed observer is given.

5.1 Graph Notations

In this work, we deal with three different graphs as described below:

(i) *System digraph*, $\mathcal{G}_{sys}$, representing the states of dynamic system (2.1) and (2.2), as defined in detail in Chapter 2.

(ii) *Communication network*, $\mathcal{G}_{net}$, defines the interaction of the agents. Let $\mathcal{G}_{net} = (\mathcal{V}_{net}, \mathcal{E}_{net})$, where $\mathcal{V}_{net} = \{1, \ldots, N\}$ is the vertex set consisting of N agents, $\mathcal{E}_{net} = \{(i,j) \mid i \leftarrow j\}$ is the set of edges (communication links), and $\mathcal{N}(i) = \{i\} \cup \{j \mid (i,j) \in \mathcal{E}_{net}\}$ denotes the extended neighborhood of agent i. Notice that, unlike many works in the literature we do not constrain $\mathcal{G}_{net}$ to be undirected. In fact, no assumption on the topology is considered here, as designing $\mathcal{G}_{net}$ is the main contribution of this chapter.

(iii) The distributed system digraph, $\mathcal{G}_{Dist}$, is the digraph associated with the distributed system $(W \otimes A, D_H)$. Indeed, the $\mathcal{G}_{Dist}$ is built on the Kronecker product of the matrices W and A [52], along with the (neighboring) measurements.

5.2 Agent Topology Design

In this section we define connectivity of $\mathcal{G}_{net}$ to satisfy the distributed observability condition. To this aim, we define the following graphs over $\mathcal{V}_{net}$:

Definition 5.2.1.

(i) *Define* $\mathcal{E}_0 = \{(i,i) \mid i \in \mathcal{V}_{net}\}$, *and* $\mathcal{G}_0 = (\mathcal{V}_{net}, \mathcal{E}_0)$. *Such graph consists of only self-edges at each node (agent).*

(ii) *Let* $\mathcal{E}_\alpha = \{(j,i) \mid i \in \mathcal{A}, j \in \mathcal{V}_{net}, j \neq i\}$, *i.e. there is a direct edge from every Type-*α *agent to all other agents; where* $\mathcal{G}_\alpha = (\mathcal{V}_{net}, \mathcal{E}_\alpha)$ *is the graph with such edges. Let* $\mathcal{N}_\alpha(i)$ *be the (extended) neighborhood of agent* i *in* $\mathcal{G}_\alpha \cup \mathcal{G}_0$. *Indeed, adjacency matrix* U *represents the graph* $\mathcal{G}_\alpha \cup \mathcal{G}_0$.

(iii) *Define* $\mathcal{G}_\beta$ *to be a SC graph over* $\mathcal{V}_{net}$. *Subsequently, let* $\mathcal{E}_\beta$ *to be the set of edges induced by* $\mathcal{G}_\beta$ *and let* $\mathcal{N}_\beta(i)$ *be the (extended) neighborhood of agent* i *in* $\mathcal{G}_\beta \cup \mathcal{G}_0$. *In this case, matrix* W *represents the adjacency of* $\mathcal{G}_\beta \cup \mathcal{G}_0$.

Remark 5.2.1.

(i) Each Type-α *agent is a hub of* $\mathcal{G}_\alpha$, *and the sub-graph of Type-*α *agents is a complete graph.*

(ii) An example of $\mathcal{G}_\beta$ *is a cycle graph, however,* $\mathcal{G}_\beta$ *is not necessarily cyclic.*

The network structure defining the communication among the agents is the union of these three sub-graphs, i.e.

$$\mathcal{G}_{net} = \mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\beta \tag{5.1}$$

Notice that, different type of information is shared over these networks. Unlike the existing works in the literature, e.g. [26, 89] among others, we do not constrain $\mathcal{G}_{net}$ to be undirected.

5.3 Distributed System Characterization

Refer $(W \otimes A, D_H)$ as the *distributed system* associated with digraph $\mathcal{G}_{Dist}$. Recall from last chapter that $W = \{w_{ij}\}$ is the weight matrix for *prediction-fusion* and is stochastic as defined in Chapter 2. On the other hand, *measurement-fusion* is defined by D_H over graph $\mathcal{G}_\alpha$.

$$D_H = \begin{pmatrix} \sum_{j \in \mathcal{N}_\alpha(1)} H_j^T H_j & & 0 \\ & \ddots & \\ 0 & & \sum_{j \in \mathcal{N}_\alpha(N)} H_j^T H_j \end{pmatrix}, \qquad (5.2)$$

To clarify the role of measurement and prediction fusion, we first consider $W = I$ and $D_H = \overline{D}_H$ defined as follows:

$$\overline{D}_H = \begin{pmatrix} H_1^T H_1 & & 0 \\ & \ddots & \\ 0 & & H_N^T H_N \end{pmatrix}, \qquad (5.3)$$

This implies no information exchange/fusion among the agents. This distributed system, $(I \otimes A, \overline{D}_H)$, can be thought of as N decoupled subsystems (as shown in Fig. 3.3). In the matrix representation, each of these subsystems is associated to an $n \times n$ block diagonal as in Fig. 5.2–(Right). Now consider W to have some non-zero off-diagonal entries. These entries define the *intra-connections* among the subsystems. As an illustrating example, consider Fig. 3.3–(Left), where we show a $n = 7$-state dynamical system with $N = 3$ agents/measurements, $\{a, b, c\}$. Agent a measures x_3, agent b measures x_5, and agent c measures x_7. Each agent is *required to* estimate the

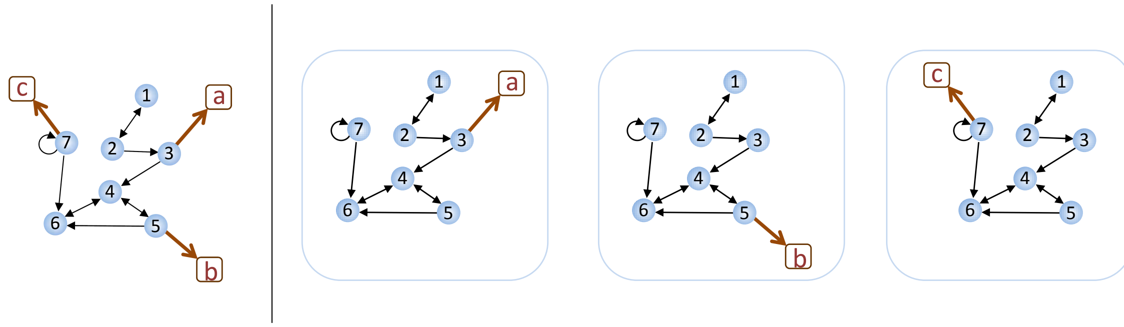

Figure 5.1: (Left) System digraph $\mathcal{G}_A$, and (Right) distributed digraph associated with $(I \otimes A, \overline{D}_H)$. This graph includes N decoupled subsystems each associated to a measurement/agent, where no subsystem is locally observable.

entire $n = 7$ dimensional state-vector. Without any information fusion each agent only has a local observation of the system as it is shown in Fig. 3.3–(Right). Therefore, each agent has to acquire the missing information (measurements and/or predictions) via communicating with agents in its immediate neighborhood. However, in this illustration, no agent finds any measurement in its neighborhood in addition to what it already has. Information sharing among the agents by applying prediction- and measurement-fusion provides more linking in $\mathcal{G}_{Dist}$. These extra links, respectively captured by the non-zeros in W and U (or the summation in D_H), implies more information sharing in the distributed system and *potentially* may improve generic distributed observability. For example, consider the case of measurement-fusion by adding the edge $b \leftarrow a$ in $\mathcal{G}_\alpha$. This implies $a \in \mathcal{N}_\alpha(b)$ and enables agent b to access the measurement of state x_3 observed by agent a.

The case for prediction-fusion is more challenging. A path, for example, from agent c to agent b (Type-β) may imply either $b \leftarrow c$, or $b \leftarrow a \leftarrow c$. Consider the latter case: the edge $a \leftarrow c$ implies $w_{ac} \neq 0$, which in the distributed system matrix, $W \otimes A$, represent as edges *from* parent SCC, $\{4, 5, 6\}$, in agent c's subsystem *to*

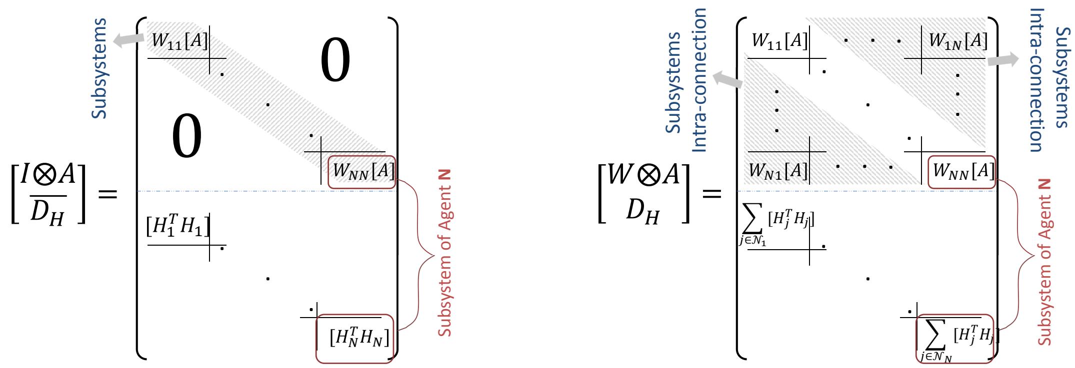

Figure 5.2: (Left) Matrix structure of the distributed system with no data fusion. Every block diagonal $W_{ii} \otimes A$ is a subsystem with a measurement monitored by an agent i. (Right) Adding information fusion, the intra-connections among the subsystems depends on the non-diagonals $W_{ij} \otimes A$, $i \neq j$.

the corresponding parent SCC in agent a's subsystem. Similarly, $b \leftarrow a$ adds edges *from* the parent SCC, $\{4, 5, 6\}$, in agent a's subsystem *to* the corresponding SCC in the subsystem of agent b. Since b is Type-β, it has a measurement of this SCC, say x_5 in this example; this entire setup allows the parent SCC at agent c (with no output) to be output-connected through a path (via agent b). This implies $\mathcal{Y}$-connectivity of states in the associated parent SCC and therefore enhances the output accessibility and observability.

A detailed discussion on the role of prediction-fusion–role of matrix W and graph $\mathcal{G}_\beta$– is stated in Section 5.5, and then the role of measurement-fusion–role of matrix U and graph $\mathcal{G}_\alpha$–in Section 5.6. The description of prediction and measurement-fusion is summarized in Table 5.1 [1]. Separating solutions for prediction and measurement-fusion is for intuition, obviously, in real applications if two agents are linked together

[1]We define the graph $\mathcal{G}^*_\alpha$ in section 5.6.

Table 5.1: Distributed system and network topology for information fusion.

Fusion level	Adj($\mathcal{G}_{Dist}$)	$\mathcal{G}_{net}$
No information fusion	$(I \otimes A, \overline{D}_H)$	$\mathcal{G}_0$
Only prediction-fusion	$(W \otimes A, \overline{D}_H)$	$\mathcal{G}_0 \cup \mathcal{G}_\beta$
Only measurement-fusion	$(I \otimes A, D_H)$	$\mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\alpha^\star$
Measurement & prediction-fusion	$(W \otimes A, D_H)$	$\mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\beta$

they *may* share all their information to maximally improve their estimation performance.

5.4 Sufficiency: Preliminary Analysis

In this section, we give some preliminaries to distributed observability analysis.

Remark 5.4.1. *All diagonal entries of W are nonzero ($w_{ii} \neq 0,\ \forall i$). This is because every agent is in its own (extended) neighborhood and uses its own information. From Lemma 2.8.2, this immediately gives S-rank$(W) = N$.*

In the following we state the main result on sufficient network connectivity; we defer the proof to Section 5.7 as it requires more development on prediction and measurement fusion from Sections 5.5-5.6.

Theorem 5.4.1. *Assume a given set of measurements H of system A, such that (A,H) is structurally observable; then $\mathcal{G}_{net}$ as in Definition 5.2.1 is sufficient for distributed observability in generic sense.*[2]

Note that, the notion of distributed observability in generic sense is similar to (centralized) generic observability but extended to a distributed estimator. As we showed

[2]This condition is not a tight sufficient connectivity on the network. In theory, there might be weakly connected networks that return distributed observability.

in the last chapters, having centralized observability is a necessary condition for distributed observability, but not sufficient. Therefore, as stated in the assumption–(iii) in Chapter 1, and particularly for the proof of sufficiency it is assumed that centralized (A, H)-observability is given (as in Theorem 2.7.1 and Theorem 3.4.1). With this assumption, Theorem 5.4.1 states that prediction and measurement fusion respectively over $\mathcal{N}_\beta(i)$ and $\mathcal{N}_\alpha(i)$ guarantee distributed observability. Let us assume for now that this statement is true, then the following corollaries are immediate:

Corollary 5.4.1. *For a full S-rank system, A, a strongly-connected network is sufficient for distributed observability.*

Proof. For a full S-rank system, there are no Type-α agents and thus $\mathcal{N}_\alpha(i) = \{i\}, \forall i$, implying no measurement sharing and therefore no $\mathcal{G}_\alpha$. However, according to Definition 5.2.1–(iii) a strongly-connected network ensures that the conditions on $\mathcal{G}_\beta$ are satisfied, see Remark 5.2.1–(i). □

A direct consequence of the above corollary is that for full S-rank systems no agent requires any measurement other than its own, i.e., measurement-fusion is not required.

Corollary 5.4.2. *If system, A, is S-rank deficient, distributed observability does not hold when $\mathcal{N}_\alpha(i) = \varnothing$, i.e. without measurement-fusion.*

Proof. Since S-rank deficiency implies existence of Type-α agents. This immediately verify from Theorem 5.4.1 that measurement-fusion is required (unlike the full S-rank case) as $\mathcal{N}_\alpha(i)$ includes more than self-measurements. This is true even for strongly-connected networks as they do include $\mathcal{G}_\beta$ as a sub-graph, but the connectivity requirements on $\mathcal{G}_\alpha$, see Definition 5.2.1–(ii), are not necessarily satisfied. □

The above corollary shows that when A is S-rank deficient, using prediction-fusion cannot guarantee distributed observability of the system and thus, the agents need access more measurements to recover their observability. Furthermore, we explicitly show that these additional measurements have to come only from Type-α agents and not from Type-β agents. This result is in contrast with existing work in the literature [15, 16, 19], because in these works: (i) only fusion in the measurement space is considered; and (ii) crucial agents are classified into a single category without recognizing their different roles towards distributed observability. Subsequently, they require *all Type-α and Type-β agents to be included in measurement-fusion.*

5.5 Prediction-fusion

In this section, according to Table 5.1, we analyze the structure of $\mathcal{G}_\beta$ for $(W \otimes A, \overline{D}_H)$ observability. First, we consider the system matrix, A, to be full S-rank. This is the case, for example, in linearization of nonlinear systems where the system matrix *almost always* has non-zero diagonal entries (e.g., [16] considers such a non-zero diagonal matrix).

Theorem 5.5.1. *With a full S-rank system, A, the pair $(W \otimes A, \overline{D}_H)$ is generically observable over $\mathcal{G}_\beta$.*

Proof. For $(W \otimes A)$ to be generically observable, the system diagraph, $\mathcal{G}_{Dist}$, should follow (i) and (ii) in Theorem 5.4.1. From Remark 5.4.1 and Lemma 2.8.2, $(W \otimes A)$ is full S-rank, which ensures condition (ii) in Theorem 2.7.1. To satisfy condition (i), according to Theorem 3.4.1–(ii), every parent SCC, $\mathcal{S}_i^{\circlearrowleft p}$, in every subsystem of $\mathcal{G}_{Dist}$ has to be $\mathcal{Y}$-connected, i.e. to reach a measurement (see Fig. 3.3 for illustration). Let

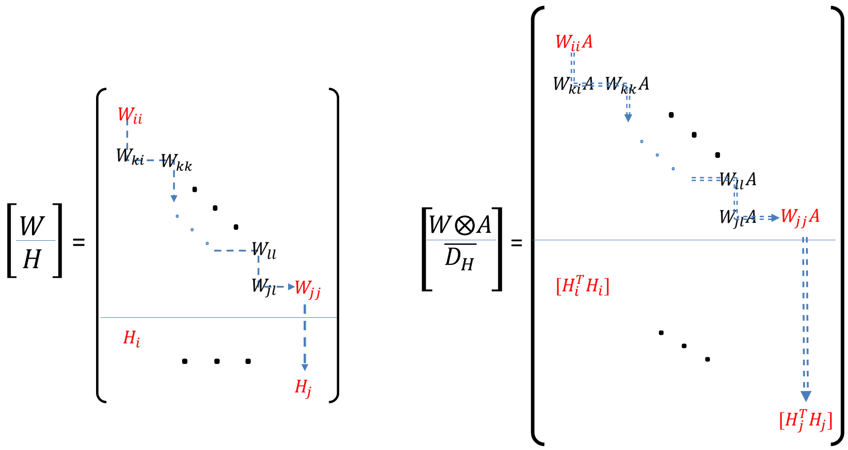

Figure 5.3: This figure illustrates the proof of Theorem 5.5.1, showing that A directed path from agent i to agent j in W implies a directed path from states in subsystem of i to subsystem of agent j in $(W \otimes A)$ and consequently y_j.

us assume that agent i has no measurement of parent SCC $\mathcal{S}_i^{\circlearrowleft p}$ in its subsystem. Then, according to $\mathcal{G}_\beta$, there is a path from i to another agent j measuring a state in SCC $\mathcal{S}_j^{\circlearrowleft p}$ (counterpart of SCC $\mathcal{S}_i^{\circlearrowleft p}$). This implies that every SCC in subsystem of agent i has a path to its counterpart SCC in subsystem of agent j, that implies $\mathcal{S}_i^{\circlearrowleft p} \rightarrow \mathcal{S}_j^{\circlearrowleft p}$ (see Fig. 5.3). Therefore, every state in SCC $\mathcal{S}_i^{\circlearrowleft p}$ is also connected to y_j, and thus, $\mathcal{S}_i^{\circlearrowleft p}$ is $\mathcal{Y}$-connected. Having this for every parent SCC in every subsystem, all SCCs in $(W \otimes A, \overline{D}_H)$ are $\mathcal{Y}$-connected and the theorem follows. □

For example, consider again the system in Fig. 3.3. Having nodes $\{4, 5, 6\}$ as parent SCC, agent b is Type-β. According to the above theorem any other agent without any measurement in $\{4, 5, 6\}$, like agent c, must have a path to agent b. This provides a connection from SCC $\{4, 5, 6\}$ in subsystem of c to the subsystem of b in $\mathcal{G}_{Dist}$ and, in turn, implies its $\mathcal{Y}$-connectivity.

Theorem 5.5.2. *If the system, A, is S-rank deficient, then $(W \otimes A, \overline{D}_H)$ is not generically observable.*

Proof. Let i be an agent for which condition (i) in Theorem 2.7.1 does not hold, i.e.

$$S\text{-rank}\begin{pmatrix} A \\ H_i^T H_i \end{pmatrix} < n. \tag{5.4}$$

Such an agent always exists because: (i) based on the Assumption–(iv) in Chapter 1, the entire system is not observable at any agent; and (ii) the matrix A is not full-rank. Now consider $(W \otimes A, \overline{D}_H)$ for the best-case scenario where $\mathcal{G}_\beta$ is a complete graph, and thus, W has all non-zero elements. Let W_i be the ith column of W. Obviously, $W_i \otimes A$ is the ith *block column* of $(W \otimes A)$, and contains block matrices $w_{ji}A$, $j = 1, \dots, N$. It follows that,

$$S\text{-rank}\begin{pmatrix} w_{ji}A \\ H_i^T H_i \end{pmatrix} < n, \tag{5.5}$$

for all $j = 1, ..., N$, as $w_{ji} \neq 0$ and scalar multiplication does not change the structure and the S-rank (maximum possible rank over all values). Since A is not full S-rank, $W_i \otimes A$ has rank less than n as stacking matrices with the same structure on top of each other (see Fig.5.4–(Left)) does not improve the S-rank. This immediately results in,

$$S\text{-rank}\begin{pmatrix} W_i \otimes A \\ H_i^T H_i \end{pmatrix} < n. \tag{5.6}$$

Consequently, according to Fig. 5.4, the structure of the matrix $W \otimes A$ is given as the side-by-side concatenation of the matrices $W_i \otimes A$. Thus we have,

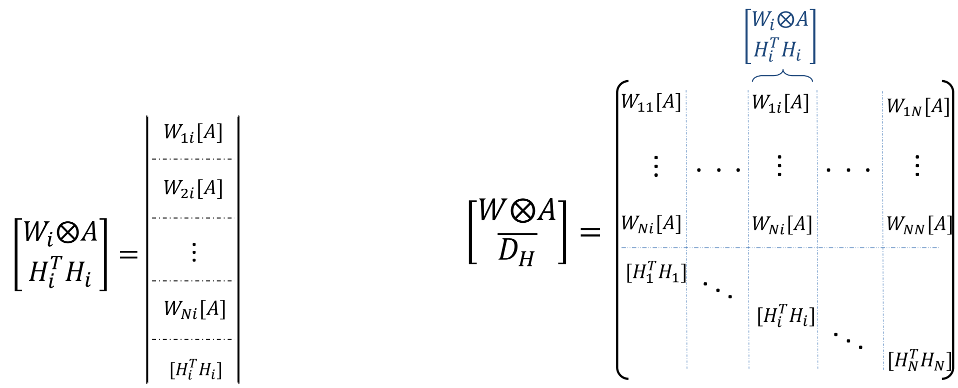

Figure 5.4: This figure illustrates the structure of $W_i \otimes A$ (Left) and matrix $W \otimes A$ (Right) in the proof of Theorem 5.5.2.

$$S\text{-rank}\begin{pmatrix} W \otimes A \\ \overline{D}_H \end{pmatrix} < Nn. \tag{5.7}$$

This holds for all choices of non-zero elements in full matrix W. Therefore, condition (ii) in Theorem 2.7.1 is violated and the theorem follows. □

The above theorem shows that when A is S-rank deficient, then using prediction-fusion *alone* cannot guarantee the distributed observability of the system, and thus, the agents need access to more measurement data to recover their observability, which is discussed next.

5.6 Measurement-Fusion

In this section, we discuss the other information fusion level, i.e. measurement-fusion. Each agent, i, shares its measurement with its direct neighbors and implements this as an innovation to update its prediction. According to the Table 5.1 measurement-fusion is tied with the observability of $(I \otimes A, D_H)$. Based on the definition of D_H, in the distributed system graph $\mathcal{G}_{Dist}$, this is equivalent to adding all neighboring measurements to its subsystem. However, with no prediction-fusion the only way to update is to directly include Type-β (necessary) measurements similar as for Type-α. Thus, here we need to define a new communication graph for Type-β agents as follows:

Definition 5.6.1. *Define $\mathcal{G}_\alpha^*$ to be a graph among $\mathcal{V}_{net}$ such that for every matched parent SCC in A, say $\mathcal{S}_l^{\circlearrowleft p}$, if agent i does not have a measurement of a state in $\mathcal{S}_l^{\circlearrowleft p}$, then it receives a direct link from any agent j with state measurement in $\mathcal{S}_l^{\circlearrowleft p}$. Subsequently, $\mathcal{E}_\alpha^*$ is the set of edges induced by $\mathcal{G}_\alpha^*$ and let $\mathcal{N}_\alpha^*(i)$ be the neighborhood of agent i in $\mathcal{G}_\alpha^* \cup \mathcal{G}_0$.*

Applying this definition the main result on measurement-fusion is stated below.

Theorem 5.6.1. *The system $(I \otimes A, D_H)$ is distributedly observable in generic sense over $\{\mathcal{G}_\alpha \cup \mathcal{G}_\alpha^*\} \cup \mathcal{G}_0$.*

Proof. Sufficiency: With the given conditions (i) and (ii), each agent has access to all necessary measurements. In this case, every agent is generically observable similar to a centralized case.

Necessity: If agent, i, is not connected to a crucial agent (α or β), then it is missing a necessary measurement and the statement follows. □

Notice that, $\mathcal{G}_\alpha^*$ contains all the Type-β agents but with a stringent connectivity requirement as compared to $\mathcal{G}_\beta$. In $\mathcal{G}_\alpha^*$, every Type-β agent is directly connected to all other agents; a restriction imposed by only considering measurement fusion, see [16] for related works. Clearly, this requires stronger connectivity as compared to strong connectivity in $\mathcal{G}_\beta$. In this work, we combine both measurement and prediction-fusion to obtain necessary and sufficient connectivity, where we need direct links *only* from the Type-α agents. This is particularly of interest in resource-constrained applications, where we cannot afford possibly long-distance links in the network.

5.7 Proof of Theorem 5.4.1

Finally, the developments of Section 5.5 and 5.6 lead to the proof of Theorem 5.4.1.

Proof. The proof of Theorem 5.4.1 is a direct consequence of the Theorems 5.5.1, 5.5.2, and 5.6.1 stated in previous sections. □

Recall that Theorem 5.5.1 sets the condition for prediction-fusion for full S-rank systems, i.e. conditions for $(W \otimes A, \overline{D}_H)$ generic observability. Theorem 5.5.2 states that for general S-rank deficient systems distributed observability cannot be achieved via the prediction-fusion alone. Measurement-fusion, i.e. generic observability of $(I \otimes A, D_H)$, is discussed in Theorem 5.6.1. Combining these results, the proof for generic observability of the distributed system $(W \otimes A, D_H)$ is immediate. Loosely speaking, in our approach prediction-fusion is applied to the full S-rank part of the system, while measurement-fusion covers the S-rank deficient part. In the following we provide some additional comments.

Remark 5.7.1.

(1) In the case of Type-β agents, the connectivity is either through a directed path (as in Theorem 5.4.1) or a direct link (as in Theorem 5.6.1); either one of these

is sufficient for observability. Notice that, the first strategy may exploit nearest-neighbor or similar communication topology, while the second may require long-distance communication.

(2) *An agent may have no system measurement and still be able to estimate the global system states via the proposed strategies. Such agents, for example, may play a role to provide and maintain connectivity of the communication network as in [90], or assist in providing directed paths to Type-β agents in $\mathcal{G}_\beta$.*

(3) *If system is not (A, H) observable then even using a fully-connected communication network does not recover observability irrespective of any estimation strategy. Clearly, the only way to recover observability is by increasing the number of state measurements to recover centralized observability [19].*

(4) *In general, adding more measurements of the system and/or more communication among the agents improve the estimation efficiency by decreasing the MSEE. In simple words, having more information/understanding of the system renders tighter bounds on the estimation error. This is the case for Type-γ agents as compared to Type-α and Type-β agents. The reason is that γ agents receive necessary information from the critical agents, but they benefit from their own non-critical measurement, which improves the bound on MSEE.*

5.8 Necessity: Recovering Observability at Each Agent

In this section, we define necessary connectivity for distributed observability based on the necessary measurement partitioning and equivalent sets defined in Chapter 3.

Theorem 5.8.1. *Consider system, A, with the necessary equivalent measurement sets given. The system is generically observable in distributed sense if every agent, i, in the network has the followings:*

(i) For every contraction set, $\mathcal{C}_l$, agent i receives a direct link from an α-agent, k, sharing a state measurement in $\mathcal{C}_l$;

(ii) Either one of the following for every matched parent SCC, $\mathcal{S}_l^{\circlearrowleft p}$:

a. Agent i receives a direct link from a β-agent, j, sharing a state measurement in $\mathcal{S}_l^{\circlearrowleft p}$;

b. Agent i is connected through a sequence of agents to a β-agent, j, observing a state in $\mathcal{S}_l^{\circlearrowleft p}$ (sharing predictions).[3]

Proof. Necessity follows a similar argument as in sufficiency in Theorem 5.4.1 and the results given in Chapter 3. The proof of parts (i) and (ii)–(a) is directly from Theorem 5.4.1; in part (i), receiving a state observation from every contraction set $\mathcal{C}_l$ recovers S-rank condition, while, in part (ii)–(a), receiving a state measurement of every parent set $\mathcal{S}_l^{\circlearrowleft p}$ directly recovers the accessibility at agent i. Part (ii)–(b)

[3]Theoretically, this condition is also *almost* sufficient for distributed observability in generic sense of Theorem 2.7.1, i.e. the observability of graph $\mathcal{G}_{Dist}$.

indirectly recovers the accessibility in $W \otimes A$. By a directed path from agent i to β-agent j, the inaccessible parent SCC $\mathcal{S}_l^{\circlearrowleft p}$ of agent i become accessible through agent j. Intuitively, states in parent SCC $\mathcal{S}_l^{\circlearrowleft p}$ of agent i affect the states of the same parent SCC of agent b and therefore any measurement of the latter one renders inference of the first one. □

Again we mention that, condition (i) defines an α-network, $\mathcal{G}_\alpha$, where agents share their measurement directly with each other. This simply implies that the necessary α measurements are required to be known for all agents at every sampling time k. On the other hand, condition (ii)–b defines an SC β-network, $\mathcal{G}_\beta$, over which the agents ***only*** share their predictions; further, condition (ii)–a defines $\mathcal{G}_\alpha^*$ over which β agents share their measurements, as in [15, 16, 19]. Notice that, this connectivity requirement is more relaxed than the necessary condition in [27] where each agent requires to transmit/share ***both*** its observations and predictions to every other agent over the same network.

5.9 Design of W matrix

This section states the design of W matrix defining consensus weights on state predictions. In previous sections we discussed two conditions needed to be satisfied by W matrix:

(i) The structure of W is associated to the topology of the network $\mathcal{G}_\beta$. Implying that W has to be irreducible to satisfy the SC condition on $\mathcal{G}_\beta$. [4]

(ii) The elements in W are such that it is row-stochastic.

[4] It should be noted that the irreducible condition satisfies the condition (ii) in Lemma 2.3.1.

For implementation the consensus update matrix W need to satisfy the above conditions. The generic result implies that each agent i may (locally) assign different weights w_{ij} on its *fixed* neighborhood while $\sum_{j=1}^{n} w_{ij} = 1$. These weights may represent how each agent *trust* the prediction of its neighbors. There might be many factors to define this trust among agents. For example, reliability/noise of the communication channel, the SNR of the transmitted signal, or even the spatial location of the neighboring agent. This weight could be a measure of importance/reliability of the incoming data. The typical choice to design such matrix is to reach *average consensus* [2, 4, 54] implying W matrix to be doubly-stochastic. As discussed in Chapter 2, other than arithmetic mean, consensus weights could be designed such that agent i takes the geometric mean or mean of order p of the neighboring predictions. Further, in [91] authors design the optimal weights according to Metropolis-Hastings method to reach fast convergence. The convergence rate is of particular interest to improve the estimation performance in single time-scale filtering.

5.10 System Classification

In general systems can be classified based on their structural rank. Recall that structural rank or S-rank of a matrix, A, is the maximal rank over all numerical values of its non-zero parameters. We summarize the results given in this chapter and Chapter 3 in the following remark:

Remark 5.10.1.

- *For full-rank systems there is no unmatched state node and, therefore, no Type-α measurement.*

- *For structurally rank-deficient systems number of Type-α agents equals the system rank-deficiency (number of unmatched nodes).*

This remark immediately results in the following lemma:

Lemma 5.10.1.

- *For full-rank systems*
  1. *centralized observability only requires β measurements.*
  2. *for distributed observability, (i) there is no hub (α agent) in $\mathcal{G}_{net}$ (multi-agent network), and (ii) strong connectivity is sufficient.*
- *For rank-deficient systems*
  1. *centralized observability further requires measurements of αs.*
  2. *for distributed observability, (i) there are hubs (α agents) in $\mathcal{G}_{net}$, and (ii) in general, more than strong connectivity is required.*

5.11 Illustrative Examples and Simulation

Example 1: Reconsider Example 3.5.1 and the system digraph in Fig. 3.3 with measurements of $\{x_2, x_4, x_6\}$. Recall that system is (A, H) observable by collecting all measurements at a central unit. By definition, a is Type-α, b is Type-β, and c is Type-γ. We propose the communication graphs $\mathcal{G}_\alpha^*$ and $\mathcal{G}_\beta$ in Fig. 5.5, respectively, associated to matrices U, W_1 and W_2 in the following,

$$U = \begin{pmatrix} \times & 0 & 0 \\ \times & \times & 0 \\ \times & 0 & \times \end{pmatrix}, \qquad W_1 = \begin{pmatrix} \times & \times & 0 \\ 0 & \times & 0 \\ 0 & \times & \times \end{pmatrix}, \qquad W_2 = \begin{pmatrix} \times & 0 & \times \\ \times & \times & 0 \\ 0 & 0 & \times \end{pmatrix}. \tag{5.8}$$

Graph $\mathcal{G}_\alpha^*$ is based on Theorem 5.6.1, where each crucial agent directly shares its measurement to every other agent. On the other hand, $\mathcal{G}_\beta$ is based on Theorem 5.4.1; agent a (Type-α) directly communicates with all other agents, while SC network connect agent b (Type-β) to all other agents. For both case $(W \otimes A, D_H)$ is generically observable.

For simulation, apply NKE in Chapter 4. We choose all the system non-zeros randomly in $[0.2, 1.2]$. A typical unstable system with $\rho(A) = 1.2 > 1$ is chosen. Random link weights for W is chosen while satisfying stochastic condition. The noise chosen to be standard Gaussian $\mathbf{v}_k \sim \mathbb{N}(0, I_{n\times n})$ and $\mathbf{r}_k^i \sim \mathbb{N}(0,1)$ and gain matrix K with blocks defined as following,

$$K^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0.0012 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1072 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0704 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.7211 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$K^{(2)} = \begin{pmatrix} 0 & 0 & 0 & -0.0026 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.0490 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0992 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.7211 & 0 & 0.5588 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.5215 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.6925 & 0 \end{pmatrix},$$

$$K^{(3)} = \begin{pmatrix} 0 & 0.1144 & 0 & 0.0043 & 0 & 0 & 0 \\ 0 & 1.0720 & 0 & 0.0037 & 0 & 0 & 0 \\ 0 & 0.7114 & 0 & 0.0507 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.7211 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

which gives,

$$\rho(W \otimes A - \underline{K} D_H (W \otimes A)) = 0.8498 \tag{5.9}$$

The MSEE is averaged over the Monte-Carlo trials and is normalized by max(MSEE) as shown in Fig. 5.6. As it can be seen, despite system instability MSEE is bounded steady state stable. This is a good example of how more measurement of the system may improve the MSEE. As we can see in this example, agent C measures a non-critical state, which is not shared by the other two agents. Having more information improves the estimation performance of this agent. As we see in the simulation, the MSEE at agent C is smaller as compared to other two agents.

Example 2: For this example, we consider the necessary set of measurements and connectivity for distributed observability. Consider the Example 3.5.2. Choose one measurement from each equivalent set $\mathcal{C}_i$ and $\mathcal{S}_j^{\circlearrowleft p}$. According to Theorem 5.8.1, we need state measurements of each equivalent set. Consider 3 agents taking necessary measurements from states $\{3, 6, 11\}$. Matrices A and H and type of agents are given in Chapter 3. The necessary communications among the agents are defined as in Fig. 5.7. Every α agent communicates with all other agents directly, and β agent shares prediction over a directed path. The associated structured adjacency matrices are given as follows:

$$U = \begin{pmatrix} \times & \times & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{pmatrix}, \qquad W = \begin{pmatrix} \times & 0 & \times \\ \times & \times & 0 \\ 0 & \times & \times \end{pmatrix} \tag{5.10}$$

To verify our results we simulate the estimation using the proposed NKE. We choose initial state values $\underline{\mathbf{x}}_0$ randomly in $(-1.5, 1.5)$. Nonzero system parameters in Eq.(3.13) are randomly chosen such that $\rho(A) = 1.1$ and the measurement gains in H are equal to 1. We choose random values for the non-zeros in associated matrix of $\mathcal{G}_\beta$ and $\mathcal{G}_\alpha$ and divide it by row-sum to make it *stochastic* as discussed in Chapter 2. System and measurement noise are Gaussian $\mathbf{v}_k^i \sim \mathbb{N}(0, 0.05^2)$ and $\mathbf{r}_k^i \sim \mathbb{N}(0, 0.2^2)$, and the block-diagonal gain matrix K is defined as in the Appendix II. The block diagonals are as follows:

$$K^{(1)} = \begin{pmatrix}
0 & 0 & 0.2864 & 0 & 0 & 0.0001 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1.0017 & 0 & 0 & 0.0004 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.8391 & 0 & 0 & -0.0004 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.0006 & 0 & 0 & 1.0001 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -0.0168 & 0 & 0 & 0.1732 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.0706 & 0 & 0 & 0.0145 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.0078 & 0 & 0 & 0.0008 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

$$K^{(2)} = \begin{pmatrix} 0 & 0 & 0.2856 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.9994 & 0 & 0 & -0.0001 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8471 & 0 & 0 & 0.0010 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0005 & 0 & 0 & 0.9999 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0179 & 0 & 0 & 0.1730 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1345 & 0 & 0 & 0.0273 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0439 & 0 & 0 & 0.0078 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$K^{(3)} = \begin{pmatrix} 0 & 0 & 0.2856 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0380 & 0 \\ 0 & 0 & 0.9994 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8390 & 0 & 0 & -0.0002 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0073 & 0 \\ 0 & 0 & -0.0007 & 0 & 0 & 0.9999 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.0058 & 0 \\ 0 & 0 & -0.0115 & 0 & 0 & 0.1745 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3058 & 0 \\ 0 & 0 & 0.2417 & 0 & 0 & 0.0249 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1.1976 & 0 \\ 0 & 0 & 0.1099 & 0 & 0 & 0.0090 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For this K matrix, we have

$$\rho(W \otimes A - \underline{K} D_H (W \otimes A)) = 0.7669 \tag{5.11}$$

Next, for each agent, we record the sum of Mean Squared Estimation Errors (MSEE) at all states at every iteration k, average it over 1000 Monte-Carlo simulation, and finally normalize it by the max(MSEE). The simulation results over $k = 50$ time-iterations is given in Fig. 5.8. As shown in the figure, error at every agent is bounded steady state *even though the system dynamics is unstable*. This simulation verifies that the linear structured system in Eq.(3.13) over the multi-agent network in Fig 5.7 is *almost always* observable. In other words, $(W \otimes A, D_H)$ is generically observable.

5.12 Conclusions

In this chapter we state the main result on necessary and sufficient connectivity of the multi-agent network for distributed observability. It is noteworthy that as opposed to the current approaches in the literature we didn't constrained our approach to only measurement sharing, but we added prediction sharing. In particular, when the system is full S-rank this makes a big difference on necessary network connectivity. In such systems, any strongly connected network ensures distributed (generic) observability. This is a prevalent network assumption to guarantee stability of distributed estimation schemes, e.g. in [11, 26, 29, 30], however, as we have shown, it is only applicable to full S-rank systems. For S-rank deficient systems more connectivity of α agents as hubs of the network is required. We used the NKE protocol and simple academic examples to verify our results. However, the combinatorial algorithms are scalable and practically feasible for any large-scale system and any distributed estimator. Application in larger scale is provided in the next chapter.

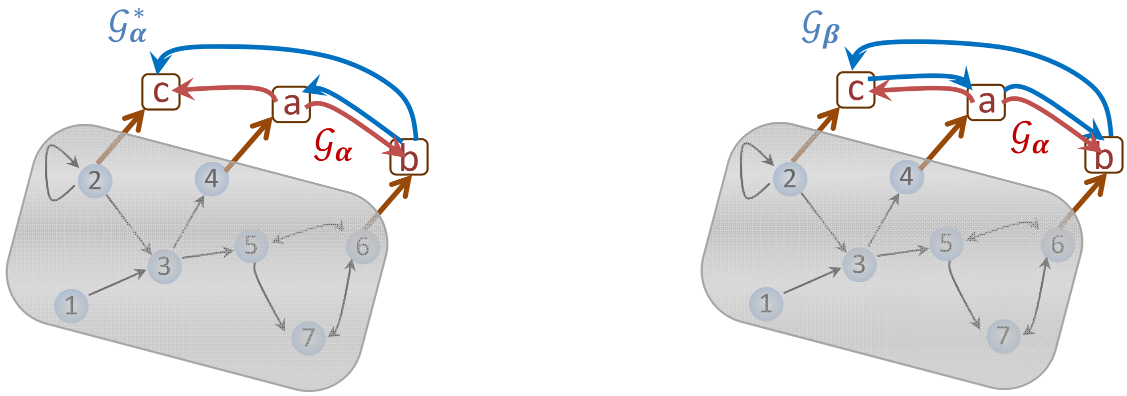

Figure 5.5: (Left) system digraph, and sufficient communication networks for Example 3.5.1: (Right) the graph $\mathcal{G}_{net_1} = \mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\alpha^*$, and (Right) the graph $\mathcal{G}_{net_2} = \mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\beta$.

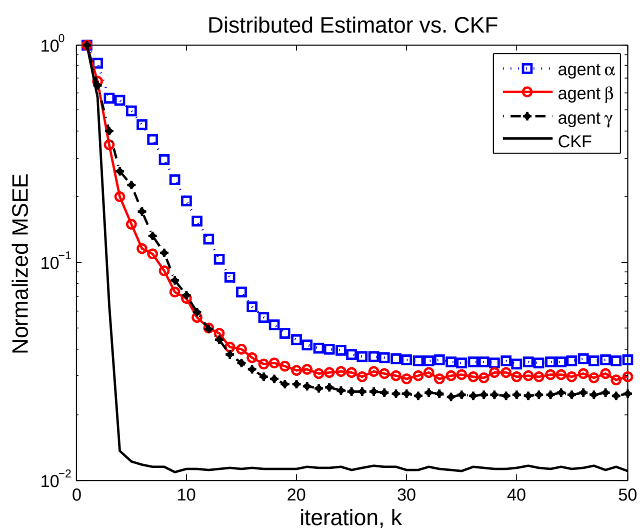

Figure 5.6: Performance of agents applying NKE over $\mathcal{G}_{net_2}$ in Fig. 5.5 as compared to CKF.

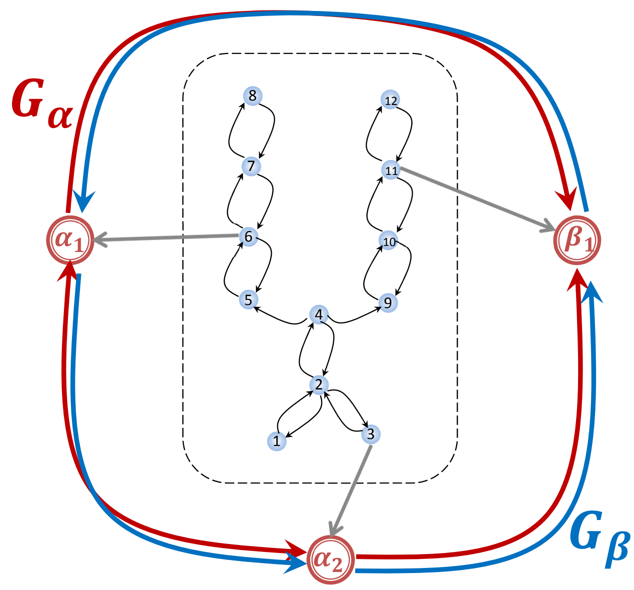

Figure 5.7: Necessary network connectivity for Example 3.5.2: $\mathcal{G}_{net} = \mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\beta$. Self edges are not shown for simplicity.

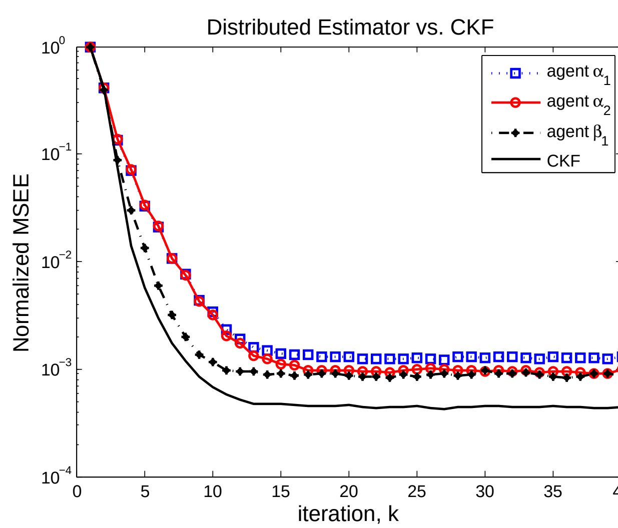

Figure 5.8: Evolution of MSEE for agents estimating the system states in Example 3.5.2 over the network $\mathcal{G}_0 \cup \mathcal{G}_\alpha \cup \mathcal{G}_\beta$ in Fig. 5.7.

Chapter 6

Applications

As mentioned previously, complexity and larger-scale in practical applications makes traditional estimation solutions obsolete. Generic single time-scale distributed solution provides estimators robust to system disturbance, with no central processor, and less communication/processing load at agents. However, the challenge is to ensure distributed observability with measurements distributed over the multi-agent network. In this context, the challenges are two-fold:

(i) the system parameters may change over time, e.g. due to dynamic loading conditions and the intermittent nature of the renewable sources in power systems;

(ii) the properties of the underlying estimator depends on the underlying agent communication graph and data fusion principles.

For example, in electric power systems, the parameters depend on the physical quantities that may change over time, whereas, in the case of linearization of a non-linear model, system parameters depend on the operating point [20]. This is the motivation behind this research, where the design methodologies are generic, independent to exact system values and only rely on the system structure.

6.1 Applications in Social Networks

In the context of the above discussion, distributed estimation is particularly relevant to social and economic networks. In such networks, the state of each node may represent an opinion [92], or a belief [93], of an individual, while the interconnections between such states represent social interactions and opinion sharing, see [94, 95], for additional details on social phenomena. The estimation problem is to estimate the state (opinions, beliefs, origin of a rumor [96, 97]) with the help of a *multi-agent network*, where each agent has observations from certain *observer nodes* and is able to exchange information with the neighboring agents, see Fig. 6.1.

Clearly, the estimation problem now is distributed and no agent may be able to infer the network opinion (global state-vector) from its observations (or neighboring observations) alone. A natural question in this regard is the following: what communication network among the agents result in *distributed observability* given the agent observations. The idea is to provide interagent communication topology to ensure distributed observability. If a social network is distributedly observable the social opinion can be discovered by adversaries. Towards security of social networks, our results can also be utilized to ensure that the social network is *not* distributedly observable by adversaries [98].

Social networks and complex networks, in general, have been modeled using both linear and nonlinear dynamics, see [94, 95, 99], and references within. Examples of linear models are in consensus/agreement problems [100–102] and Markov-based opinion formation [92, 103]. Two well-known linear models are social influence networks by Freidkin and Johnson [102] and French model [104]. The French model formulates the formation of opinions (states) under the interpersonal influence of peers. Similarly,

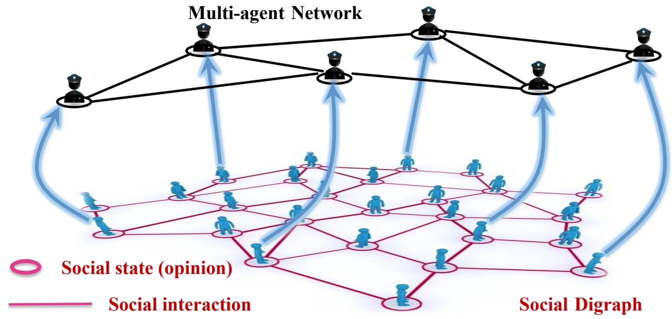

Figure 6.1: A multi agent network monitoring a social digraph.

Freidkin and Johnson model the process of social influence on opinion evolution. Another socio-economic example is [105], where product prices as states linearly evolve on a daily basis according to a competitiveness matrix (auction game).

Of significant relevance to this work is the characterization in [106] and [107] where the structure of the linear model is assumed to be fixed but with time-varying interaction weights. In particular, Reference [106] describes examples of a linear state-space on the social networks resulting from email communication, and social interaction of Monks (members of a particular religious order). On the other hand, Reference [107] discusses a linear state-space for influence networks, where attitudes, sentiments, or expectations (states) evolve over time-varying influences of other actors. For nonlinear social dynamics, simplified modeling methods have been considered, e.g. [20, 66]. Particularly, observability of nonlinear dynamics is characterized by the structural observability of the corresponding linearized system [66]. Hence, it is natural to model the social phenomena as LSI systems, where any (non-zero) element of the system matrix may change (modeling distinct or time-varying phenomena) as long as the structure (social digraph) is not violated, e.g. see [106] and [107].

Towards modeling the social systems we specifically address the following questions:

(i) Which individuals (observer nodes in Fig. 6.1) are critical for the global inference of the social system? and,

(ii) Given such critical opinions observed by a network of agents, what are the connectivity requirements on the agent network to ensure distributed observability?

The first question aims the contribution of each observation towards the understanding of the social phenomena, based on the results of Chapter 3; and the latter question, defines the communication of the observing agents, based on the results of Chapter 5. In this case, the critical observer nodes can be divided into α and β types, for which α types directly share their state observation, while β types share their state prediction over a SC network. The analysis and design procedures addressed in this thesis have significance in very large-scale social networks and related physical phenomena. This is primarily because of the structural analysis where (distributed) observability is considered as a *generic* property of the system.

6.1.1 Large-scale examples

We provide some insights of our results towards inference in social networks. Consider a social group of actors with states, e.g. opinions, sentiments, emotions, etc., that evolve over social interactions. The influence network, e.g. friendship, co-authorship, swarming, etc., is time-invariant but the influence weight of actors may vary over time, and different weight assignment model the evolution of different states resulting into different social phenomena. Our aim is to infer such phenomena by observing some critical states without considering any particular dynamics but only the social interactions (digraph). For distributed inference, first, we classify these states (and

Table 6.1: Social network examples: summary of inference features.

Networks	$n = \|\mathcal{V}_A\|$	$E = \|\mathcal{E}_A\|$	n_α ●	n_β ●
Monks	18	88	0	1
Blogs	1224	19025	436	0
Books	105	882	0	1
Dolphins	62	159	2	0
Coauthorship	1461	5484	37	248

the agents observing them) according to Definitions 3.2.3 and 3.2.2 in Chapter 3. The necessary network of agents is defined according to Theorem 5.4.1 in Chapter 5.

Following the discussion in Chapter 5, the structure of any social digraph is highly relevant to the dynamics that may take place over the social network. In this context, we use some of the well-known social network models [108, 109] and explore the graphical observability results developed in this work. These networks have been used for the estimation of corresponding social phenomena modeled as LSI systems. Each node (circles in Figs. 6.3–6.5) represents a state, e.g., heading, opinion, buying habits, etc., in the social digraph and evolves over social interactions. Theorem 2.7.1 characterizes the necessary observerations. These observations (and their associated agents) are classified as Type-α (red circles) and Type-β (green circles). Finally, Theorem 5.4.1 characterizes the network of these agents accordingly (a typical illustration of such network is presented in Fig. 6.2). The results are summarized in Table 6.1.

(a) *Political Blogs*: A social digraph of hyperlinks between weblogs on US politics [110], shown in Fig. 6.3–(Left). Each node represents a blog linked to other political blogs; the state at each node could be the popularity of the blog evolving via political commentary [111]. The blogs can be seen to have two dominant clusters constituting blogs that are more followed and hyperlinked. The digraph has

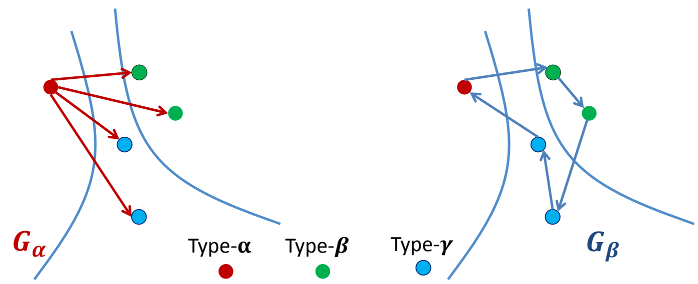

Figure 6.2: A typical symbolic sufficient linking of the agents based on their type following Theorem 5.4.1.

$n_\alpha = 436$ unmatched nodes. We may observe that most of the Type-α agents appear on the boundary of the network where the blogs are less cited (hyperlinked), and thus, may not be inferred from the interior nodes. This specific example of inference of the popularity of such blogging network shows that: (i) hubs (nodes with high degrees) are not critical for observability; and, (ii) to extract the popularity of all blogs in a distributed way, a fully-connected network is necessary (and sufficient [47]).

(b) *Books on US Politics*: Amazon.com data–undirected edges represent co-purchasing of books by the same buyers [112], digraph is shown in Fig. 6.3–(Right). The network has full Srank, thus $n_\alpha = 0$, and is further connected so $n_\beta = 1$, and can be an observation from any node.

(c) *Sampson's Monastery Network* is a directed network of interactions among the Monks in a monastery. The digraph from [109] is shown in Fig. 6.3–(Left). The network is full Srank, implying $n_\alpha = 0$, and is strongly-connected so $n_\beta = 1$. To illustrate agent connectivity, assume a collection of such monasteries, each observed by a β-agent. From our results, it is necessary for the agents to communicate over

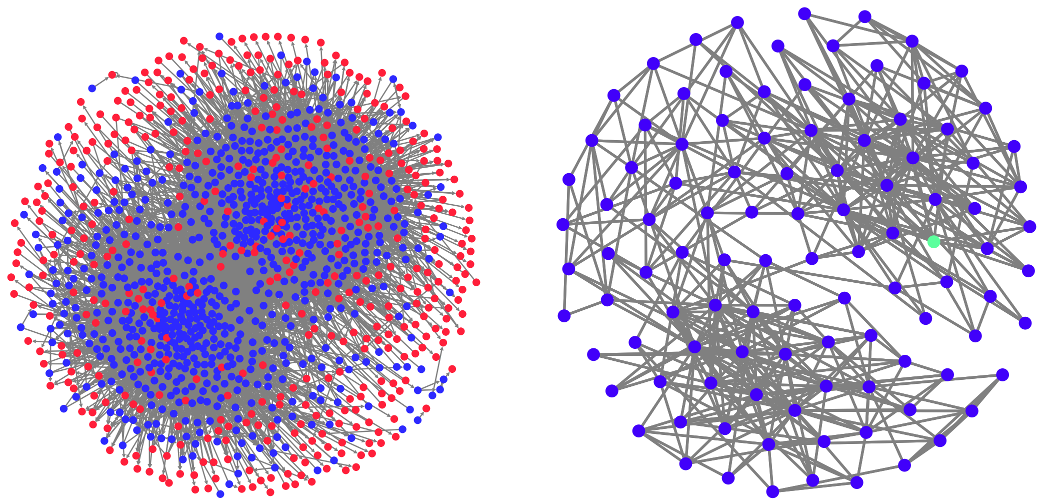

Figure 6.3: Social digraphs: (Left) Political blogs during the 2004 US Elections with 1224 nodes; (Right) Network of political books with 105 nodes.

a strongly-connected network in order to estimate any social phenomena on the union of the corresponding social digraphs.

(d) *Dolphin Social Network*: An undirected social network of frequent associations between dolphins [113], see Fig. 6.4–(Right). It can be verified that this network contains one connected component with $n_\alpha = 2$ unmatched nodes.

(e) *Co-authorship in Network Science*: A graph of researchers in network theory [114],

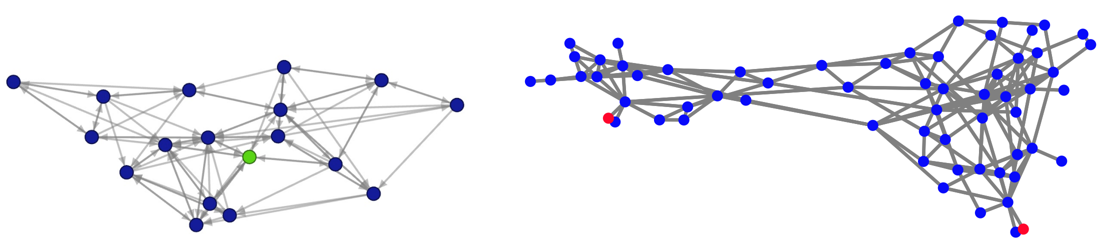

Figure 6.4: (Left) Directed Sampson's network with 18 actors; (Right) Undirected dolphin social network with 62 nodes.

shown in Fig. 6.5. The states may model a novel concept or a result and the links represent the influence among the authors. The digraph contains 268 components out of which 248 are matched. All of matched components are parent resulting into $n_\beta = 248$; and, $n_\alpha = 37$. Wiring according to Theorem 5.4.1, each agent may infer any phenomena that evolves over this social digraph.

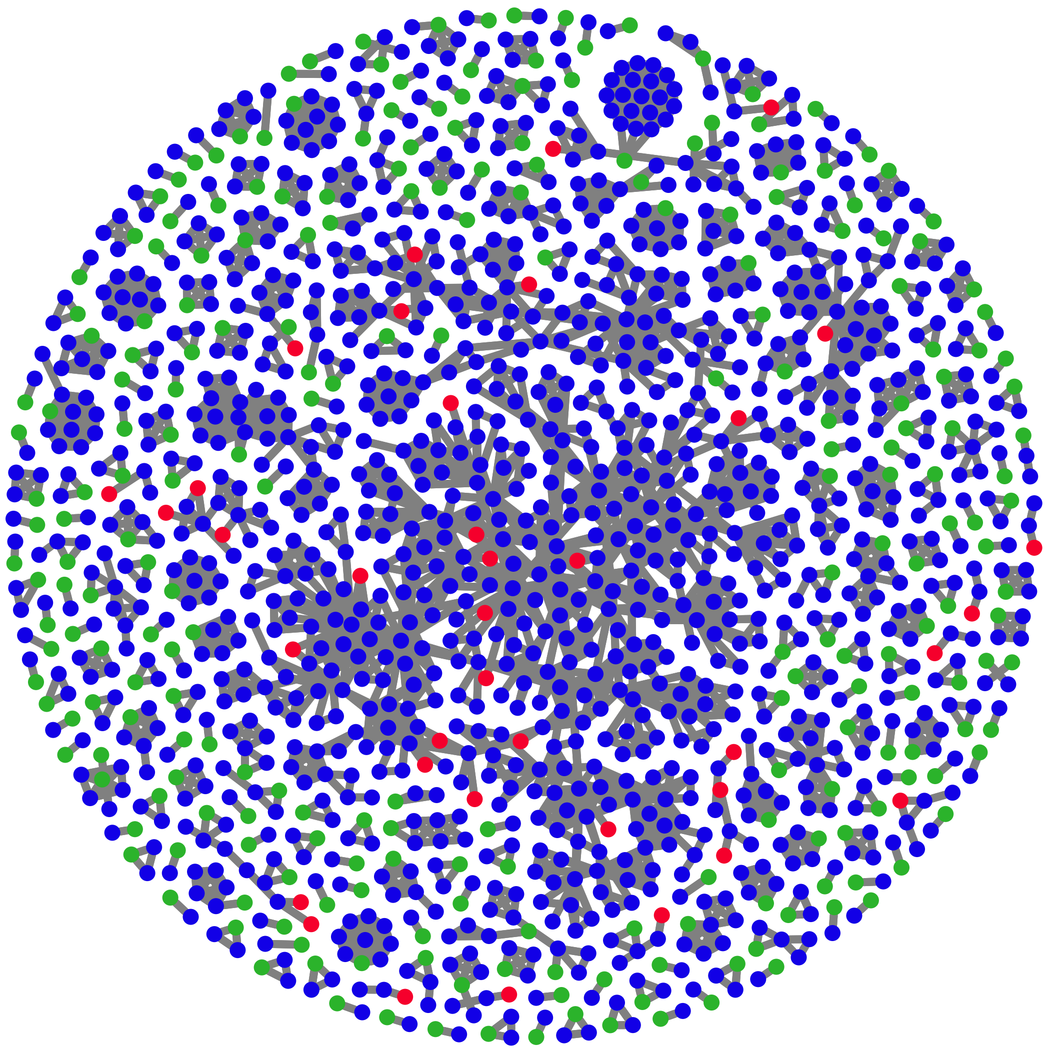

Figure 6.5: Co-authorship network with 1461 nodes

6.2 Applications in Power Systems

The future smart grid is envisioned to be a diverse mixture of conventional power plants (natural gas, coal) and renewable energy sources (solar, wind) providing new opportunities for electricity generation and distribution with a focus on an eco-friendly *green* planet. The overall operation is required to efficiently integrate this green initiative with modernized sensing/communication infrastructure while considering the possibility of an intelligent consumer and the physical limitations, e.g., the intermittent nature of renewable energy sources. The novel sensing methodologies hold the promise of allowing faster and distributed response to perturbations of the grid behavior. As mentioned before, such time-varying systems requires generic methodologies irrespective of system parameters.

In this regard, graphical model of a power system consists of nodes representing voltages and phases at different buses and edges defined by the topology and impedance parameters, see e.g. [115, 116] on related estimation scenarios where our results on measurements partitioning in Chapter 3 and scenarios on sensor placement [117, 118] are applicable. Moreover, the advent of fast sensing devices, like phasor measurement units (PMUs), also presents the possibility of instrumenting in large scale (say, regional, or national) power grid with a wide sensing network. Such networked systems typically consists of power-constrained and relatively cheap sensors/robots that can locally communicate (limited) system measurements/information.

6.2.1 Example

In order to motivate the concepts on structural equivalent partitioning, we illustrate the results on the Western States power grid borrowed from [108, 119] where the original data and description of the state nodes can be found. Since the sparsity of the actual system matrix has some resemblance to this structure and the corresponding dynamics may depend on what state-vector is being modeled (power flow, voltage, and/or angles), we choose to illustrate the results on the power network instead of a particular system matrix. The inference diagram of the network is presented in Fig. 6.6.

This network includes 6594 interaction edges connecting 4941 state nodes including 575 unmatched nodes. Without loss of generality, let assume *each* unmatched state is assigned to an α-agent, and since the network is one connected component, there is no β-agent. In the case of failed sensing of any of red states (in Fig. 6.6), the observability might be recovered by choosing any state in the corresponding contraction. Two examples are given in Fig. 6.7. The left figure is the largest contraction (blue colored) in the power system with 52 state nodes implying that a failed observation in this set may be recovered– in terms of observability– by observing any of 51 other states. On the other hand, the recovery of any failed observation of the green colored contraction in the right figure is restricted to only 2 options.

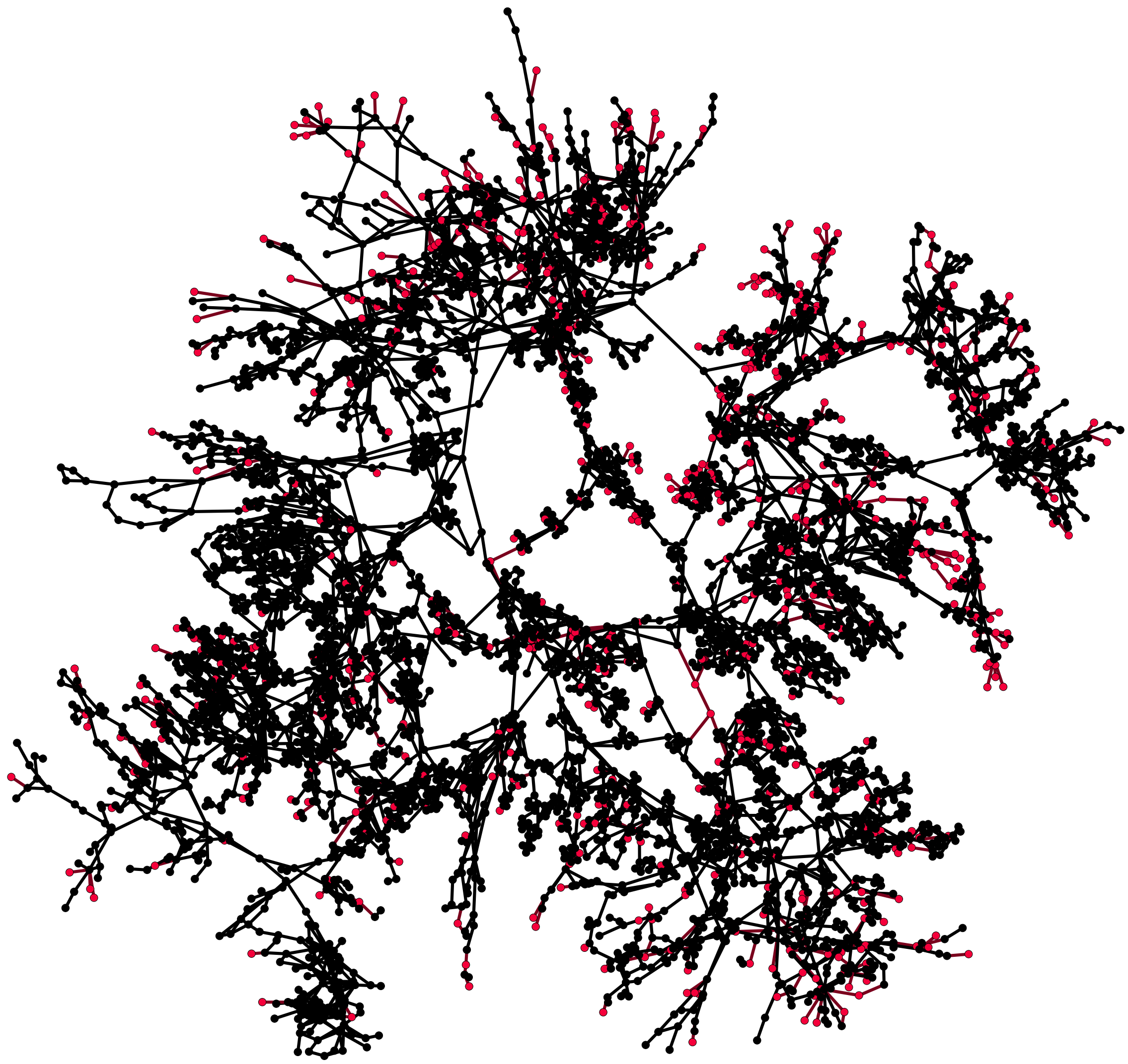

Figure 6.6: A power-grid system with 4941 state nodes and 6594 edges: red states in the network represent unmatched states.

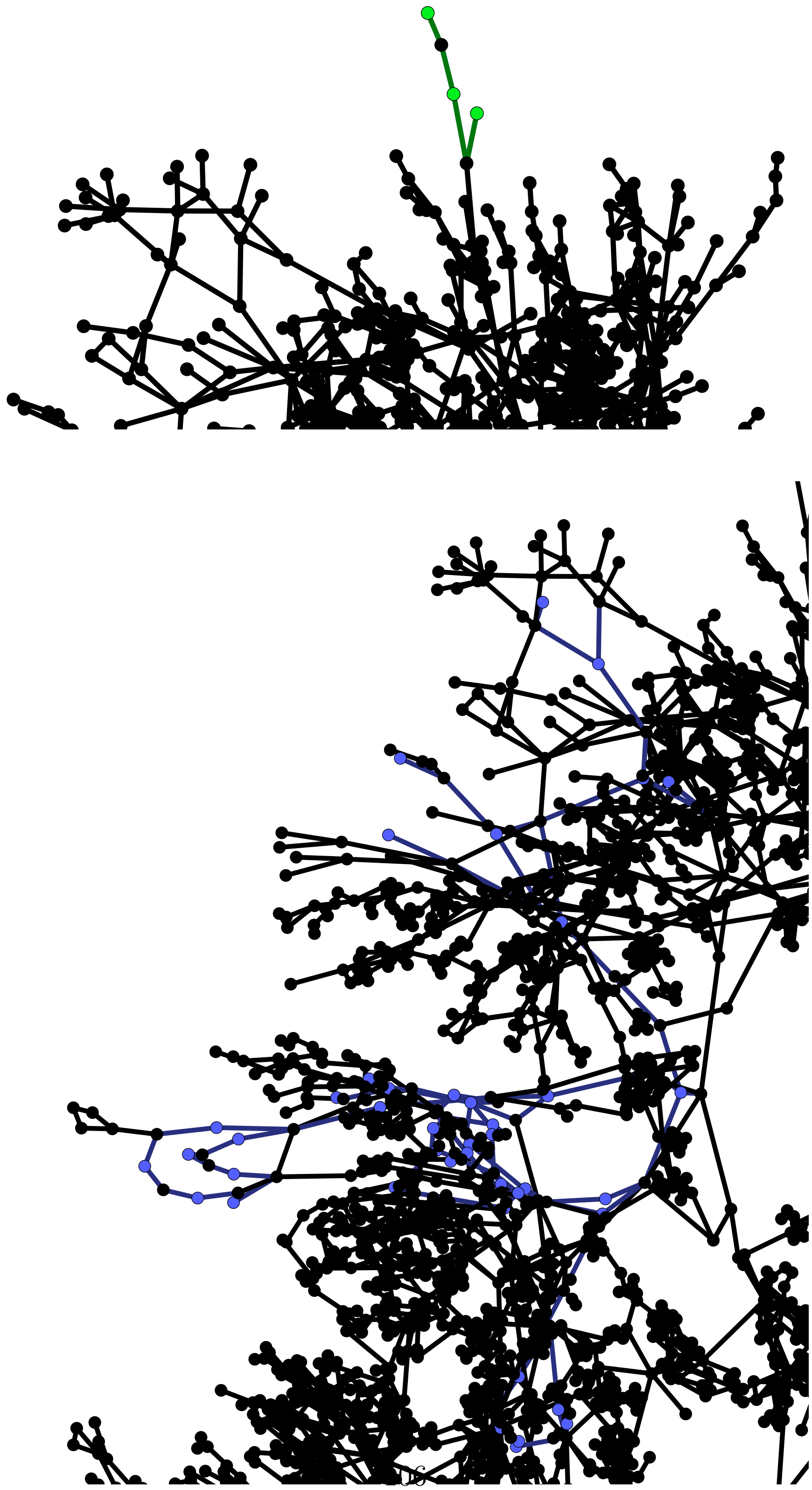

Figure 6.7: Two examples of equivalent states in contractions, represented as colored state nodes.

Chapter 7

Conclusions and Future Work

In this thesis, we formulate the distributed observability in the context of *single time-scale* distributed estimation, where communication and dynamics operate at the same time-steps. This is different from the multi time-scale consensus-based estimator that is prone to very large communication in between every two dynamical time-steps. Therefore, that does not require the availability of all crucial measurements but only a strongly-connected network. In contrast, the centralized and semi-centralized estimation approach requires all the crucial measurements. In the specific context of single-time distributed estimation, first, this thesis challenges the following question: *Do we require the availability of all crucial measurements at each agent?* And, the answer in the context of [15, 19, 40], and other related references, is 'Yes'. The **prime contribution** of this thesis is to show that the answer to the aforementioned question is 'Not necessarily.' In particular, we subdivide the agents taking crucial measurements into two classes, namely: Type-α and Type-β. We show that:

(i) Only the measurements of Type-α are required at each agent;

(ii) The measurements of Type-β are not necessarily required as long as strong connectivity holds, i.e. they are required to be shared through a path.

Note that (i) and (ii) are significant in terms of both the analysis (arriving at the results) and their applicability (e.g., in topology design).

We generalize formulation of distributed observability; for example, similar works [15, 19] in the literature are a special case of our formulation when distributed observability is given by $(I \otimes A, D_H)$, where $I \otimes A$ represents no fusion on the state predictions. In Chapter 5, Theorem 5.6.1, we show that with no information fusion on prediction space, all Type-β agents must have the same connectivity requirements as Type-α agents– represented as $\mathcal{G}_\alpha^*$. A very simple illustration where such estimator is practically infeasible is distributed estimation of a full rank system when the agents are randomly deployed, e.g., random field estimation. According to Chapter 5–Corollary 5.4.1, we show that no direct connection is required in a full Srank system and strong-connectivity is sufficient; it is clear that strong-connectivity can be assumed given a reasonable density vs. communication radius relation. On the contrary, estimation over $\mathcal{G}_\alpha^*$, requires direct connections to each agent from every crucial agent; for a random (e.g., geometric) deployment, this requires a communication radius as large as the maximum distance between any two nodes in a random graph. Notice that we do not necessarily assume (A, H_i) to be observable at any agent i, see assumptions–(iv) in Chapter 1. In our proposed formulation, even when all Type-α agents are included in the neighborhood of agent i, agent is *not* assumed to be locally observable in its neighborhood (unless there are no Type-β agents). In other words, we do not make any assumption on the observability of $(A, \sum_{j\in\mathcal{N}_\alpha(i)} H_j^T H_j)$. This is contrary to another typical approach in the existing literature [16, 40], where local observability of every agent in its neighborhood is a pre-assumption.

In the context of topology design, we view the multi-agent network as an SC graph of β-agents *plus* a hub network of α agents.[1] It can be readily seen that although strong-connectivity plus additional direct Type-α links are sufficient, weakly-connected networks may satisfy (i) and (ii). To the best of our knowledge, in literature, strong-connectivity is almost always assumed unless strict assumptions of *either* local observability, *or* availability of all crucial measurements at each agent are employed. As mentioned before, we introduce two new graph constructs, namely $\mathcal{G}_\alpha$ and $\mathcal{G}_\beta$:

(a) Over SC graph $\mathcal{G}_\beta$ every agent shares its prediction $A\widehat{\mathbf{x}}_{k-1|k-1}$ over a *directed-path*;

$$\text{Prediction fusion:} \qquad \widehat{\mathbf{x}}^i_{k|k-1} = \sum_{j\in\mathcal{N}_\beta(i)} w_{ij} A \widehat{\mathbf{x}}^j_{k-1|k-1}, \tag{7.1}$$

(b) The graph $\mathcal{G}_\alpha$ is such that every Type-α agent *directly* shares its observation with every other agent;

$$\text{Measurement fusion:} \qquad \widehat{\mathbf{x}}^i_{k|k} = \widehat{\mathbf{x}}^i_{k|k-1} + K^i_k \sum_{j\in\mathcal{N}_\alpha(i)} H_j^T (y^j_k - H_j \widehat{\mathbf{x}}^i_{k|k-1}). \tag{7.2}$$

The other contribution of this thesis is to realize that the set of all dynamical systems can be partitioned into systems with full Srank matrices and systems with Srank deficient matrices. This partitioning leads to a novel estimator formulation and the subsequent analysis. This is because each of the prediction- and measurement-fusion steps are now particularly designed in order to address the structural aspects of the system matrix. For example, we show that Type-α agents only exist in Srank

[1] Note that for topology design, a strongly-connected network always satisfy this property, however, strong-connectivity is not necessary.

deficient systems, and thus, measurement-fusion is not required in full Srank systems, i.e., prediction-fusion alone leads to bounded MSEE. From the topology perspective, this formally proves that strongly-connected network is sufficient for all full Srank systems–as no α agent exist in the system. This justifies the reason why invertability of system matrix is a typical assumption in distributed estimation literature.

We further address problem of finding sets of all equivalent states for observability. We provide graph theoretic algorithms to search for these equivalent sets in system digraph and also the algebraic implications of these sets. Graphically, contractions build equivalent α-sets while parent SCCs contain set of equivalent β-agents. This equivalence formulation is of interest specifically to find the minimal number of measurements for observability, and also, to recover failed observability. For example, for a failed α agent its observation might be recovered by measuring its equivalent state in a contraction, and a failed β communication link might be recovered by any path connecting the two agents in communication network. An interesting result is that having enough redundancy in $\mathcal{G}_\beta$ improves resiliency of the distributed estimator to failed/disrupted communications. As we mentioned in the introduction, in this research we guarantee the inference of the global state of the system. As future works, we may consider partial state estimation at agents. For example, consider the case where each agent only intend to track its spatially neighboring states. In such scenario only partial system observability is required at each agent.

The results of this thesis are independent of what fusion rule (e.g., Metropolis-Hastings [91]) is chosen in (7.1). The reason is that we analyze observability problem *generically*, such that it is true for almost all possible choices of the fusion rule (weight matrices). Furthermore, generic properties are, in general, easily verified. For example, there are efficient graph theoretic, [34], flow theoretic, [120], and linear programming, [121], methods that can be employed to check for generic properties. In terms of computational complexity and application, the structural-based algorithms in this work are of polynomial order and scalable for large scale, motivating application in power systems and social networks. As discussed in Chapter 6, in power systems, continually evolving nature of these systems along with disturbances make the precise representation of exact system parameters unavailable, while on the other hand, social networks typically have *time-invariant structures* but *time-evolving parameters.*

Appendix I

Combinatorial Algorithms

There are known combinatorial algorithms to define SCCs and their partial order in the system digraph. Depth-First-Search (DFS) algorithm [81] and Tarjan algorithm [80] are two common and well-known examples of polynomial order $\mathcal{O}(n^2)$ [81]. On the other hand, maximal matching can be defined via, for example, *Hopcraft-Karp* algorithm with running time of $\mathcal{O}(n^{2.5})$ [122]. The size of the maximum matching of the bipartite graph, further, defines the S-rank(A).

A general algorithm to determine SCCs, their partial order, maximum matching, and contractions in the system digraph is *Dulmage-Mendelsohn decomposition* (DM decomposition in short) [78, 123]. Thus, this algorithm yields both Type-α and β agents. The computational complexity of the algorithm is $\mathcal{O}(\sqrt{n}|\mathcal{E}_A|)$, i.e. maximum order of $\mathcal{O}(n^{2.5})$. A modified version of DM decomposition is given in Algorithm 1. The algorithm gives unique decomposition of the bipartite graph into *irreducible* subgraphs–called Dulmage-Mendelsohn components–and their partial order. In Algorithm 1, the unmatched nodes are included in $\mathcal{V}_0$ while matched parent SCCs are in $\{\mathcal{V}_1, ..., \mathcal{V}_k\}$. Further, the partial order identifies the parents in $\{\mathcal{V}_1, ..., \mathcal{V}_k\}$.

Algorithm 1 Dulmage-Mendelsohn Decomposition algorithm.
Given: Maximum matching $\mathcal{M}$, Auxiliary graph $\Gamma_A^{\mathcal{M}}$.

1. Let $\mathcal{V}_0 = \{\mathcal{V}_0^+ \cup \mathcal{V}_0^-\}$ such that $\{v \in \mathcal{V}^+ \cup \mathcal{V}^- |\ u \xrightarrow{path} v$ for some $u \in \delta\mathcal{M}$ on $\Gamma_A^{\mathcal{M}}$. This set defines the contractions.
2. Let $\mathcal{V}_\infty = \{\mathcal{V}_\infty^+ \cup \mathcal{V}_\infty^-\}$ such that $\{v \in \mathcal{V}^+ \cup \mathcal{V}^- |\ v \xrightarrow{path} u$ for some $u \in \mathcal{V}^- \backslash \partial\mathcal{M}^-$ on $\Gamma_A^{\mathcal{M}}$.
3. Define $\Gamma_{-\{\mathcal{V}_0 \cup \mathcal{V}_\infty\}}^{\mathcal{M}}$, i.e. the bipartite graph obtained from $\Gamma_A^{\mathcal{M}}$ by deleting the nodes (and associated edges to) $\{\mathcal{V}_0 \cup \mathcal{V}_\infty\}$.
4. Define a partial order $\leq$ on $\{\mathcal{V}_0, \mathcal{V}_1, ..., \mathcal{V}_k, \mathcal{V}_\infty\}$ as $\{\mathcal{V}_i \leq \mathcal{V}_j\} \iff \{v_i \xrightarrow{path} v_j$ for some $v_i \in \mathcal{V}_i, v_j \in \mathcal{V}_j\}$
5. Define $\mathcal{V}_i = \{\mathcal{V}_i^+ \cup \mathcal{V}_i^-\}$, $i \in \{1, ..., k\}$ in the graph $\Gamma_{-\{\mathcal{V}_0 \cup \mathcal{V}_\infty\}}^{\mathcal{M}}$ as matched SCCs of the system digraph.

Appendix II

Local Estimator Gain Design

We consider the design of the estimator gain matrix, K_k, in Eq.(4.19). Notice that having $(W \otimes A, D_C)$ observable guarantees a *full* gain matrix, K_k, to stabilize the MSEE. However, according to protocol (4.19), we need a local gain matrix, $\underline{K}_k$, which is *block-diagonal* with N blocks of $n \times n$ matrices. Here, we design a constant estimator gain matrix, $\underline{K}$, independent of time, k.

A partial list of references devoted to find constrained estimator gain for control and estimation are [11, 124–127]. Here, we use the Linear Matrix Inequality (LMI) approach in [125, 126]. However, in general, the corresponding LMIs do not have a solution, because of the structural constraints (block-diagonal) on the gain matrix, K. This is the main difficulty in distributed estimation and control as convex/semi-definite approaches are not directly applicable. To this end, an iterative procedure is implemented to solve LMIs under structural constraints. In this regard, the estimator gain matrix, K, is the solution of the following optimization problem. Interested readers may find more details in [11, 125].

$$
\begin{aligned}
&\min \quad \mathbf{trace}(XY) \\
&\text{subject to} \quad X, Y > 0, \\
&\qquad\qquad \begin{bmatrix} X & \widehat{A}^T \\ \widehat{A} & Y \end{bmatrix} > 0, \qquad \begin{bmatrix} X & I \\ I & Y \end{bmatrix} > 0, \\
&\qquad\qquad K \text{ is block-diagonal.}
\end{aligned} \tag{7.3}
$$

where,

$$\widehat{A} = W \otimes A - \underline{K} D_H (W \otimes A) \tag{7.4}$$

In fact, we need a $\underline{K}$ such that $\widehat{A}$ is Schur (i.e. $\rho(\widehat{A}) < 1$). Notice that, the solution to the second LMI is equivalent to $X = Y^{-1}$, which gives the minimum trace and the optimal value as nN. The nonlinear product of X and Y can be replaced with a linear approximation [125, 126, 128], $\phi_{lin}(X, Y) = \mathbf{trace}(Y_0 X + X_0 Y)$ and an iterative algorithm [126] can be used to minimize $\mathbf{trace}(XY)$ under the given constraints.

It is shown in [126] that $\mathbf{trace}(Y_k X + X_k Y)$ is a non-increasing sequence that converges to $2nN$. In this regard, a stopping criterion in step 3 of the above algorithm can also be established in terms of reaching within $2nN + \epsilon$ of the trace objective. The iterative procedure given above is centralized, however, the center has to implement this process only once, off-line; then it transmits the estimator gains to each agent and plays no further role in the implementation of local estimators; each agent, subsequently, observes and performs in-network operations to implement the estimator. A single time-scale algorithm can also be employed, where the above iterative procedure is implemented at the same time-scale k as of the dynamical system. With this approach, the estimator gain iterations, $\underline{K}_k$, at each k is applied to the estimator at time-step, k, and may be transmitted to each agent at each step k. This is helpful when the implementation is assumed in real-time.

Algorithm 2 Iterative calculation of local estimator gain matrix, K.

1. Find feasible points $X0, Y0, K$. If no such points exist, Terminate.
2. At iteration $k > 0$ minimize $\textbf{trace}(Y_k X + X_k Y)$ under the constraints given in (7.3) and find X, Y, K.
3. If $\rho(\widehat{A}) < 1$ terminate, otherwise set $Y_{k+1} = Y,\ X_{k+1} = X$ and run the step 2 for next iteration $k = k + 1$.

Bibliography

[1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[2] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.

[3] R. Olfati-Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *42nd IEEE Conference on Decision and Control*, 2003.

[4] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," vol. AC-48, no. 6, pp. 988–1001, June 2003.

[5] R. Olfati-Saber, "Distributed Kalman filters with embedded consensus filters," in *44th IEEE Conference on Decision and Control*, Seville, Spain, Dec. 2005, pp. 8179–8184.

[6] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering using consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, May 2008.

[7] U. A. Khan and J. M. F. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, Oct. 2008.

[8] E. J. Msechu, S. D. Roumeliotis, A. Ribeiro, and G. B. Giannakis, "Decentralized quantized Kalman filtering with scalable communication cost," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3727–3741.

[9] U. A. Khan and A. Jadbabaie, "On the stability and optimality of distributed kalman filters with finite-time data fusion," in *American Control Conference*, 2011, pp. 3405–3410.

[10] U. Khan, S. Kar, A. Jadbabaie, and J. Moura, "On connectivity, observability, and stability in distributed estimation," in *49th Conference on Decision and Control*, Atlanta, GA, 2010, pp. 6639–6644.

[11] U. A. Khan and A. Jadbabaie, "Coordinated networked estimation strategies using structured systems theory," in *49th IEEE Conference on Decision and Control*, Orlando, FL, Dec. 2011, pp. 2112–2117.

[12] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *ASME Journal of Basic Engineering*, vol. 83, pp. 95–108, 1961.

[13] J. Bay, *Fundamentals of linear state space systems*, McGraw-Hill, 1999.

[14] D. Luenberger, “Introduction to dynamic systems: theory, models, and applications,” 1979.

[15] C. Commault, J. M. Dion, and D. H. Trinh, “Observability preservation under sensor failure,” *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1554–1559, Jul. 2008.

[16] T. Boukhobza, F. Hamelin, S. Martinez-Martinez, and D. Sauter, “Structural analysis of the partial state and input observability for structured linear systems: Application to distributed systems,” *European Journal of Control*, vol. 15, no. 5, pp. 503–516, Oct. 2009.

[17] C. Lin, “Structural controllability,” *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201–208, Jun. 1974.

[18] J. M. Dion, C. Commault, and J. van der Woude, “Generic properties and control of linear structured systems: A survey,” *Automatica*, vol. 39, pp. 1125–1144, Mar. 2003.

[19] T. Boukhobza and F. Hamelin, “State and input observability recovering by additional sensor implementation: A graph-theoretic approach,” *Automatica*, vol. 45, no. 7, pp. 1737–1742, 2009.

[20] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, “Controllability of complex networks,” *Nature*, vol. 473, no. 7346, pp. 167–173, May 2011.

[21] S. Pequito, S. Kar, and A. P. Aguiar, “A structured systems approach for optimal actuator-sensor placement in linear time-invariant systems,” in *American Control Conference*, Washington, DC, 2013, pp. 6123–6128.

[22] J.J. Slotine and W. Li, *Applied nonlinear control*, Prentice-Hall, 1991.

[23] H. Hashemipour, S. Roy, and A. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 88–94, Jan. 1988.

[24] A. Mutambara, *Decentralized estimation and control for multisensor systems*, CRC Press, Boca Raton, FL, 1998.

[25] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4064–4077, Aug. 2007.

[26] S. Tu and A. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6217–6234, 2012.

[27] S. Park and N. Martins, "Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers," in *51st IEEE Conference on Decision and Control*, 2012, pp. 7431–7436.

[28] D. Acemoglu, A. Nedić, and A. Ozdaglar, "Convergence of rule-of-thumb learning rules in social networks," in *47th IEEE Conference on Decision and Control*, Dec. 2008, pp. 1714–1720.

[29] J. Chen and A. H. Sayed, "Diffusion adaptation strategies for distributed optimization and learning over networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4289–4305, 2012.

[30] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-based algorithms for distributed filtering," in *51st IEEE Conference on Decision and Control*, 2012, pp. 794–799.

[31] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," submitted for publication, available at http://arxiv.org/abs/0809.0009, Aug. 2008.

[32] A. Speranzon, C. Fischione, K. H. Johansson, and A. Sangiovanni-Vincentelli, "A distributed minimum variance estimator for sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 609–621, May 2008.

[33] R. Kalman, "A new approach to linear filtering and prediction problems," *Trans. of the ASME - Journal of Basic Engineering*, vol. 82, no. 2, pp. 35–45, 1960.

[34] K. J. Reinschke, *Multivariable control, a graph theoretic approach*, Berlin: Springer, 1988.

[35] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, Sep. 2010.

[36] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, May 2011.

[37] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2462–2475, Nov. 2010.

[38] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4782–4795, Dec. 2006.

[39] G. Battistelli, L. Chisci, S. Morrocchi, and F. Papi, "An information-theoretic approach to distributed state estimation," in *18th IFAC World Congress*, 2011, pp. 12477–12482.

[40] S. Das and J. M. F. Moura, "Distributed kalman filtering with dynamic observations consensus," *IEEE Transactions on Signal Processing*, 2015, to appear.

[41] M. Doostmohammadian and U. A. Khan, "Communication strategies to ensure generic networked observability in multi-agent systems," in *45th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2011, pp. 1865–1868.

[42] U. A. Khan and M. Doostmohammadian, "A sensor placement and network design paradigm for future smart grids," in *4th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, San Juan, Puerto Rico, Dec. 2011, pp. 137–140.

[43] M. Doostmohammadian and U. A. Khan, "On the distributed estimation of rank-deficient dynamical systems: A generic approach," in *38th International*

Conference on Acoustics, Speech, and Signal Processing, Vancouver, CA, May 2013, pp. 4618–4622.

[44] M. Doostmohammadian and U. A. Khan, “Topology design in network estimation: a generic approach,” in *American Control Conference*, Washington, DC, Jun. 2013, pp. 4140–4145.

[45] M. Doostmohammadian and U. A. Khan, “On the characterization of distributed observability from first principles,” in *2nd IEEE Global Conference on Signal and Information Processing*, Atlanta, GA, Dec. 2014, pp. 914–917.

[46] M. Doostmohammadian and U. A. Khan, “Vulnerability of CPS inference to DoS attacks,” in *48th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2014, pp. 2015–2018.

[47] M. Doostmohammadian and U. Khan, “On the genericity properties in distributed estimation: Topology design and sensor placement,” *IEEE Journal of Selected Topics in Signal Processing: Adaptation and Learning over Complex Networks*, vol. 7, no. 2, pp. 195–204, 2013.

[48] M. Doostmohammadian and U. Khan, “Graph-theoretic distributed inference in social networks,” *IEEE Journal of Selected Topics in Signal Processing: Signal Processing for Social Networks*, vol. 8, no. 4, pp. 613–623, Aug. 2014.

[49] M. Doostmohammadian and U. A. Khan, “Measurement partitioning and observational equivalence in state estimation,” submitted.

[50] C. Godsil and G. Royle, *Algebraic graph theory*, New York: Springer, 2001.

[51] R. Distel, *Graph theory*, Springer-Verlag, NY, 3rd edition, 2005.

[52] J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani, "Kronecker graphs: An approach to modeling networks," *The Journal of Machine Learning Research*, vol. 11, pp. 985–1042, 2010.

[53] M. Xue and S. Roy, "Kronecker products of defective matrices: Some spectral properties and their implications on observability," in *American Control Conference, 2012*, 2012, pp. 5202–5207.

[54] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[55] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 169–182, 2005.

[56] F. Xiao and L. Wang, "Asynchronous consensus in continuous-time multi-agent systems with switching topology and time-varying delays," *IEEE Transactions on Automatic Control*, vol. 53, no. 8, pp. 1804–1816, 2008.

[57] D. Bauso, L. Giarré, and R. Pesenti, "Non-linear protocols for optimal distributed consensus in networks of dynamic agents," *Systems & Control Letters*, vol. 55, no. 11, pp. 918–928, 2006.

[58] J. Cortés, "Finite-time convergent gradient flows with applications to network consensus," *Automatica*, vol. 42, no. 11, pp. 1993–2000, 2006.

[59] H. Sayyaadi and M. Doostmohammadian, "Finite-time consensus in directed switching network topologies and time-delayed communications," *Scientia Iranica*, vol. 18, no. 1, pp. 75–85, 2011.

[60] W. Ren, R. Beard, and E. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control systems magazine*, vol. 27, no. 2, pp. 71–82, 2007.

[61] U. A. Khan, *High-dimensional consensus in large-scale networks: Theory and applications*, Ph.D. thesis, Carneigie Mellon University, PA, Aug. 2009.

[62] R. Olfati-Saber, "Kalman-consensus filter : Optimality, stability, and performance," in *48th IEEE Conference on Decision and Control*, Shanghai, China, Dec. 2009, pp. 7036–7042.

[63] P. S. Maybeck, *Stochastic models, estimation, and control*, vol. 141 of *Mathematics in Science and Engineering*, 1979.

[64] M. Hautus, "Controllability and observability conditions of linear autonomous systems," *Nederlandse Akademie van Wetenschappen*, vol. Serie A, no. 72, pp. 443–448, 1969.

[65] S. Jafari, *Structural Controllability of Multi-Agent Systems Subject to Partial Failure*, M.Sc. thesis, Concordia University, July 2010.

[66] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, "Observability of complex systems," *Proceedings of the National Academy of Sciences*, vol. 110, no. 7, pp. 2460–2465, 2013.

[67] M. Ji and M. Egerstedt, "Observability and estimation in distributed sensor networks," in *46th IEEE Conference on Decision and Control*, Dec. 2007.

[68] M. Egerstedt, "Complex networks: Degrees of control," *Science*, vol. 473, pp. 158–159, May 2011.

[69] E. J. Davison and S. H. Wang, "Properties of linear time-invariant multivariable systems subject to arbitrary output and state feedback," *IEEE Transactions on Automatic Control*, vol. 18, no. 1, pp. 24–32, Feb. 1973.

[70] E. J. Davison and S. H. Wang, "Properties and calculation of transmission zeros of linear multivariable systems," *Automatica*, vol. 10, pp. 643–658, 1974.

[71] F. Harary, "The determinant of the adjacency matrix of a graph," *SIAM Review*, vol. 4, no. 3, pp. 202 –210, Jul. 1962.

[72] S. Li and Y. Guo, "Distributed consensus filter on directed graphs with switching topologies," in *American Control Conference*, Washington, DC, 2013, pp. 6151–6156.

[73] D. Acemoglu, A. Nedić, and A. Ozdaglar, "Convergence of rule-of-thumb learning rules in social networks," in *47th IEEE Conference on Decision and Control*, Dec. 2008, pp. 1714–1720.

[74] D. P Bertsekas, "Incremental gradient, subgradient, and proximal methods for convex optimization: A survey," Tech. Rep., MIT, 2010.

[75] V. Ugrinovskii, "Conditions for detectability in distributed consensus-based observer networks," *arXiv:1303.6397*, 2013.

[76] I. D. Schizas, G. Mateos, and G. B. Giannakis, "Distributed LMS for consensus-based in-network adaptive processing," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2365–2382, Jun. 2009.

[77] W. Abbas and M. Egerstedt, "Hierarchical assembly of leader-asymmetric, single-leader networks," in *American Control Conference*, San Francisco, CA, Jun. 2011.

[78] K. Murota, *Matrices and matroids for systems analysis*, Springer, 2000.

[79] P. Hall, "On representatives of subsets," *Journal of London Mathematic Society*, vol. 10, no. 1, pp. 26–30, 1935.

[80] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM Journal on Computing*, vol. 1, no. 2, pp. 146–160, 1972.

[81] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, 2009.

[82] D. S. Dummit and R. M. Foote, *Abstract Algebra*, Wiley, 2004.

[83] C. Berge, *Graphs and Hypergraphs*, Elsevier Science Ltd, 1985.

[84] C. Commault, J. Dion, D. Trinh, and T. Do, "Sensor classification for the fault detection and isolation, a structural approach," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 1, pp. 1–17, 2011.

[85] C. Commault, J. Dion, and S. Yacoub Agha, "Location of additional sensors for FDI," in *4th Workshop on Advanced Control and Diagnosis*, 2006.

[86] S. Pequito, S. Kruzick, S. Kar, J. M. F. Moura, and A. Pedro Aguiar, "Optimal design of distributed sensor networks for field reconstruction," in *21st IEEE European Signal Processing Conference*, 2013, pp. 1–5.

[87] S. Pequito, S. Kar, and A.P. Aguiar, "Minimum number of information gatherers to ensure full observability of a dynamic social network: a structural systems approach," in *IEEE 2nd Global Conference on Signal and Information Processing*, Atlanta, GA, Dec. 2014.

[88] C. Zhang, S. Wang, and Y. F. Zheng, "Minimum order input-output equation for linear time-varying digital filters," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 171–173, 1998.

[89] R. Olfati-Saber and P. Jalalkamali, "Collaborative target tracking using distributed kalman filtering on mobile sensor networks," in *American Control Conference*, San Francisco, CA, Jun. 2011.

[90] A. N. Kopeikin, S. S. Ponda, L. B. Johnson, O. Toupet, and J. P. How, "Real-time dynamic planning to maintain network connectivity in a team of heterogeneous unmanned systems," in *2nd International Workshop on Wireless Networking for Unmanned Autonomous Vehicles at the IEEE GlobeComm Conference*, Dec. 2011.

[91] L. Xiao, S. Boyd, and S. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, pp. 33–46, 2005.

[92] D. Acemoglu and A. Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, vol. 1, no. 1, pp. 3–49, 2011.

[93] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210–225, 2012.

[94] N. Friedkin, *A structural theory of social influence*, vol. 13, Cambridge University Press, 2006.

[95] M. Newman, A. Barabási, and D. Watts, *The structure and dynamics of networks*, Princeton University Press, 2006.

[96] D. Shah and T. Zaman, "Rumors in a network: Who's the culprit?," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5163–5181, 2011.

[97] A.G. Dimakis, S. Kar, J.M.F. Moura, M.G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1847–1864, 2010.

[98] Y. Wan M. Xue, S. Roy and S. K. Das, "Security and vulnerability of cyber-physical infrastructure networks: A control-theoretic approach," *Handbook on Securing Critical Cyber-Physical Infrastructures: Foundations and Challenges*, to appear in 2012.

[99] V. Blondel, J. Hendrickx, and J. Tsitsiklis, "On krause's multi-agent consensus model with state-dependent connectivity," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2586–2597, 2009.

[100] K. Birla and S. Kamalapur, "State Space Model for Link Mining," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 2, no. 2, pp. 296–298, 2013.

[101] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks: a local interaction game with stubborn agents," in *American Control Conference*, Washington, DC, Jun. 2013, pp. 1985–1990.

[102] N. Friedkin and E. Johnsen, "Social influence networks and opinion change," *Advances in Group Processes*, vol. 16, no. 1, pp. 1–29, 1999.

[103] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002.

[104] J. R. P. French, "A formal theory of social power," *Psychological Review*, vol. 63, no. 3, pp. 181–194, 1956.

[105] J. van Dalen, W. Ketter, and G. Lucchese, "A Kalman filter approach to analyze multivariate hedonics pricing model in dynamic supply chain markets," in *12th International Conference on Electronic Commerce*.

[106] E. Xing, W. Fu, and L. Song, "A state-space mixed membership blockmodel for dynamic network tomography," *The Annals of Applied Statistics*, vol. 4, no. 2, pp. 535–566, 2010.

[107] N. Friedkin and E. Johnsen, "Attitude change, affect control, and expectation states in the formation of influence networks," *Advances in Group Processes*, vol. 20, pp. 1–29, 2003.

[108] L. D. Christopher, "UCI network data repository," 2008, `http://networkdata.ics.uci.edu`.

[109] "Social Network Image Animator (SoNIA)," *http://www.stanford.edu/group/sonia*.

[110] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 US Election," in *Workshop on the Weblogging Ecosystem*, 2005.

[111] K. Lerman and T. Hogg, "Using a model of social dynamics to predict popularity of news," in *19th international conference on World wide web*. ACM, 2010, pp. 621–630.

[112] V. Kerbs (compiled by), "Books about us politics (unpublished)," `http://networkdata.ics.uci.edu/data.php?d=polbooks`.

[113] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations," *Behavioral Ecology and Sociobiology*, vol. 54, pp. 396–405, 2003.

[114] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Physcial Review Letters*, vol. 74, May 2006.

[115] Y. 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," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 33–43, 2012.

[116] G. Giannakis, V. Kekatos, N. Gatsis, S. Kim, H. Zhu, and B. Wollenberg, "Monitoring and optimization for power grids: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 107–128, 2013.

[117] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, Oct 2011.

[118] E. J. Msechu and G. B. Giannakis, "Sensor-centric data reduction for estimation with WSNs via censoring and quantization," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 400–414, Jan 2012.

[119] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[120] V. Hovelaque, C. Commault, and J.-M. Dion, "Analysis of linear structured systems using a primal-dual algorithm," *System and Control Letters*, vol. 27, pp. 73–85, February 1996.

[121] S. Poljak, "Maximum rank of powers of a matrix of a given pattern," *Proceedings of the American Mathematical Society*, vol. 106, no. 4, pp. pp. 1137–1144, 1989.

[122] J. Hopcroft and R. Karp, "An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs," *SIAM Jour. of Computation*, pp. 225–231, 1973.

[123] A. Dulmage and N. Mendelsohn, "Coverings of bipartite graphs," *Canadian Journal of Mathematics*, vol. 10, pp. 517–534, 1958.

[124] U. A. Khan and A. Jadbabaie, "Networked estimation under information constraints," available at http://arxiv.org/abs/1111.4580.

[125] M. Pajic, S. Sundaram, J. Le Ny, G. Pappas, and R. Mangharam, "The wireless control network: Synthesis and robustness," in *49th Conference on Decision and Control*, Orlando, FL, 2010, pp. 7576–7581.

[126] L. El Ghaoui, F. Oustry, and M. Ait Rami, "A cone complementarity linearization algorithm for static output-feedback and related problems," *IEEE Trans. on Automatic Control*, vol. 42, no. 8, pp. 1171–1176, 1997.

[127] A. I. Zečević and D. D. Šiljak, "Control design with arbitrary information structure constraints," *Automatica*, vol. 44, no. 10, pp. 2642–2647, 2008.

[128] O. L. Mangasarian and J. S. Pang, "The extended linear complementarity problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 2, pp. 359–368, Jan. 1995.